Genomic cloning of the Hsc71 gene in the hermaphroditic teleost *Rivulus marmoratus* and analysis of its expression in skeletal muscle: identification of a novel muscle-preferred regulatory element

Jong Hyuk Park, Jong Joo Lee, Sarah Yoon, Jae-Seong Lee, Soo Young Choe[1], Joonho Choe[2], Eun-Ho Park and Chul Geun Kim*

Department of Life Science and Research Institute for Natural Sciences, College of Natural Sciences, Hanyang University, Seoul 133-791, Korea, [1]School of Life Sciences, Chungbuk National University, Cheongju 361-763, Korea and [2]Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea

ABSTRACT

To further our understanding of the role of stress proteins in development as well as in adaptation of fish to adverse environmental conditions, we undertook molecular analyses of stress protein encoding genes from the hermaphroditic teleost *Rivulus marmoratus*. We isolated a genomic clone containing the Hsc71 gene (*rm-hsc71m*) and its upstream sequences. rm-Hsc71m is not induced by external stress, but is enriched in a tissue-specific manner during early development. In adult, the strongest expression appeared in skeletal muscle, whereas lower expression was seen in the gill, eye and brain. To understand the regulatory basis of high muscle expression of *rm-hsc71m*, transfection of *R.marmoratus* muscle tissue was performed using 5′ deletion fragments containing the *rm-hsc71m* promoter driving EGFP expression. An upstream region from −2.7 to −1.9 kb was identified as a muscle-specific regulatory region. Within this region, we identified at least three sites with the novel sequence TGTnACA interacting with a fish muscle factor having an M_r of 32 000. Our data indicate that *rm-hsc71m* expression in skeletal muscle is controlled by a muscle-specific regulatory element containing this novel motif.

INTRODUCTION

Heat shock proteins (HSPs) exhibit not only temporally induced expression in response to environmental stresses but complex patterns of spatial and temporal regulation during embryonic development in a wide range of organisms (1–4). Although most of what we know about HSPs comes from studies of a variety of organisms (5–20), fish are a valuable model system to investigate the functional role of HSPs as well as how their expression is regulated (21,22). Fish are highly diverse vertebrates living in a variety of aquatic habitats; therefore, to endure diverse environmental conditions, fish have acquired not only specialized organs and regulatory systems but likely have highly developed responses to stress. Furthermore, in recent years fish have emerged as a popular model for the study of vertebrate development and are amenable to a combination of embryological, genetic and molecular approaches (23,24).

Mangrove rivulus, *Rivulus marmoratus*, in particular is one of the best model systems available to examine the regulation of stress protein expression at the tissue level or as a whole organism. This mangrove-dwelling fish is the only vertebrate that is a synchronous, internally self-fertilizing hermaphrodite (25). This unique reproductive mode yields offspring with little genetic variation (26), and therefore, Mangrove rivulus are believed to rely on post-genomic means such as stress proteins to adapt to various environments. Besides its genetic homogeneity, this fish has several desirable attributes for studies on stress proteins. Most importantly, Mangrove rivulus shows strong physiological tolerance in both its natural habitat and in laboratory culture. Although the natural habitat of this species is shallow estuarine mangrove marshes, it has wide salinity tolerance ranging from freshwater to salinity twice that of seawater (70‰ salinity) and can survive within a water temperature range of 4–40°C (27,28). This species also exhibits a strong tolerance of pH toxicity, such that the 96-h LD_{50} values for acid and base are under pH 4.0 and over pH 10, respectively (29). Tolerance of environmental extremes

*To whom correspondence should be addressed. Tel: +82 2 2290 0957; Fax: +82 2 2296 5996; Email: cgkim@hanyang.ac.kr

Table 1. Oligonucleotides used in this study

Name	Description	Sequences of Oligonucleotide
Primer P1	rm-hsc71m -214 /-195 (forward)	5'-CTGCAGTAACCCAAGTCCCAGAATCC-3'
Primer P2	rm-hsc71m -79 /-60 (forward)	5'-CTGCAGGGGGGTCACCTGACACCTGA-3'
Primer P3	rm-hsc71m -2643/-2625 (forward)	5'-GCAGGGATGAGACGTCAAGA-3'
Primer P4	rm-hsc71m -2556/-2541 (reverse)	5'-GTATGTAGTTGTGTGTGGAG-3'
Primer P5	rm-hsc71m -2351/-2334 (forward)	5'-CCTCCAACTGCATATGTT-3'
Primer P6	rm-hsc71m -2235/-2216 (reverse)	5'-TGGCCACCCTAACAGAGGCA-3'
Primer P7	rm-hsc71m -2105/-2082 (forward)	5'-CTGATGTAAATTCATGTCTGCAAA-3'
Primer P8	rm-hsc71m -1996/-1973 (reverse)	5'-GACATTCACACAAATACTTCTTCT-3'
Primer a	rm-hsc71m +17/+36 (Exon I, forward)	5'-GGAGTGCTGGAGTGCTGGAG-3'
Primer b	rm-hsc71m +1715/+1734 (Exon II, reverse)	5'-CCTACTGCTGGTCCTTTAGA-3'
Primer c	rm-hsc71m +1837/+1856 (Exon II, forward)	5'-GGCTTTCACAGACACAGAGAGA-3'
Primer d	rm-hsc71m +2466/+2485 (Exon IV, reverse)	5'-GCTGCAGCTGTGGGTTCATT-3'
EGFP 1R	Jellyfish EGFP gene (forward)	5'-CACCCTGGTGAACCGCATCG-3'
EGFP 2R	Jellyfish EGFP gene (reverse)	5'-GCTGCACGCTGCCGTCCTCGA-3'
Lac Z 1R	*E. coli* Lac Z (forward)	5'-CGCCTTGCAGCACATCCCCC-3'
Lac Z 2R	*E. coli* Lac Z (reverse)	5'-CTGCCGTCACTCCAACGC-3'
rm-hsc71mE 1R	rm-hsc71m -2351/-2267 (forward)	5'-GACACAAACACAGATGTTTTTGTGACTTGTGACACCTGTTT-3'
rm-hsc71mE 2R	rm-hsc71m -2351/-2267 (reverse)	5'-ATGTCAAACAGGTGTCACAAGTCACAAAAACATCTGTGTTT-3'
MCKE 1R	MCKE (Apone and Hauschka, 1995)	5'-TTAACCCAGACATGTGGCTGCCCCCCCCCCCCCCCAACACCTGCTG-3'
MCKE 2R	MCKE (Apone and Hauschka, 1995)	5'-TCAGGCAGCAGGTGTTGGGGGGGGGGGCGGCAGCCACATGTCTGG-3'
a box F	rm-hsc71m -2627/-2606 (forward)	5'-GGGGAGATGTTGTAGCAGCTGGTTAC-3'
a box R	rm-hsc71m -2627/-2606 (reverse)	5'-GGGGGTAACCAGCTGCTACAACATCT-3'
b box F	rm-hsc71m -2296/-2275 (forward)	5'-GGGGGTGACTTGTGACACCTGTTTGA-3'
b box R	rm-hsc71m -2296/-2275 (reverse)	5'-GGGGACAAACAGGTGTCACAAGACAC-3'
bm box F	b box with 3 base substitutions (F)	5'-GGGGGTGACTTGATTCACCTGTTTGA-3'
bm box R	b box with 3 base substitutions (R)	5'-GGGGACAAACAGGTGAATCAAGACAC-3'
c box F	rm-hsc71m –2009/-1988 (forward)	5'-GGGGGGGAACCACTGTTACAGAAGAA-3'
c box R	rm-hsc71m -2009/-1988 (reverse)	5'-GGGGTTCTTCTGTAACAGTGGTTCCC-3'
m box F	c box with 6 base substitution (forward)	5'-GGGGTTCTTCACAATGTGTGGTTCCC-3'
m box R	c box with 6 base substitution (reverse)	5'-GGGGGGGGAACCACACATTGTGAAGAA-3'
MCKE-RH F	MCKE-right half (forward)	5'-GGGGCCCCCCCCCCAACACCTGCTG-3'
MCKE-RH R	MCKE-right half (reverse)	5'-GGGGCAGCAGGTGTTGGGGGGGGGG-3'

suggests that the fish have well-developed stress proteins, probably to compensate for the genetic rigidity.

As an initial effort to analyze the role of stress proteins in Mangrove rivulus, we report here the genomic cloning of a heat shock cognate protein 71 gene (*rm-hsc71m*) and its upstream regulatory sequences, and present expression profiles of *rm-hsc71m* during development as well as in response to external stresses. We have also identified a novel muscle-preferred regulatory element with the sequence TGTnACA interacting with a fish muscle factor having an M_r of 32 000.

MATERIALS AND METHODS

Accession number

The sequence for *rm-hsc71m* was deposited in DDBJ/EMBL/GenBank under accession number AF227986.

Genomic library screening, DNA subcloning and sequencing

A λGEM-II genomic library (30) was screened using a [α-^{32}P]dCTP-labeled human *hsp70* cDNA probe. After hybridization and washing, the filter was air-dried and exposed to X-ray film. For restriction enzyme mapping, the DNA from an individual clone was digested with various restriction enzymes, individually or in combination. Nucleotide sequencing of a series of deletion mutants generated by the erase-A-base method (Promega) was performed manually with a sequencing kit (Pharmacia). Sequence data was compiled and analyzed by the DNASIS program. Homology searches, at both the nucleotide and amino acid levels, were performed using the NIH BLAST program.

Primer extension analysis

Primer extension was carried out according to the procedure of Williams and Mason (31). A chemically synthesized oligonucleotide primer (from +1715 to +1734) from the 5'-end of the coding exon II (primer b in Table 1) was 5'-end labeled with [α-^{32}P]ATP. This primer was annealed to 30–50 μg of total RNA purified from Rivulus trunk muscle at 50°C in 30 μl of hybridization buffer. Reaction products were analyzed by electrophoresis through a 6% polyacrylamide gel including 8 M urea with reference to a sequencing ladder of M13 DNA and visualized by autoradiography.

Construction of expression plasmids

DNA containing 2792 bp of upstream sequences and 44 bp of exon I of *rm-hsc71m* was cloned upstream of an enhanced green fluorescent protein (EGFP) gene in the pEGFP-N1 vector (Clontech), and the resulting vector was designated pRM2792. The vectors pRM2654, pRM1943, pRM1572, pRM1278, pRM775 and pRM389 were generated by erase-A-base mutagenesis of the pRM2792 vector. The terminal junction sequences of each construct were confirmed by direct DNA sequencing. The pRM214 and pRM79 vectors were generated by subcloning the PCR fragments. The pRM1943-3'Sen and pRM1943-3'Asen vectors were constructed by inserting an ~1.2 kb *Pvu*II fragment from –2617 to –1418 bp of the *rm-hsc71m* upstream sequences downstream of the EGFP gene in pRM1943 in both the sense and antisense orientations, respectively. All plasmids were purified using a Qiagen plasmid Midi kit (Qiagen).

Laboratory culture of *R.marmoratus*

Fish were bred and reared in a 40 l glass container containing aerated brackish water of $10 \pm 1‰$ salinity at $25 \pm 1°C$ (32,33). Specimens used in this study were from the 27th to 29th generations of a single progenitor raised in Hanyang University until analysis. This progenitor originated from the Zoologisches Institut und Zoologisches Museum, University of Hamburg, Germany, in 1981 and the Hamburg stock was derived from Floridian wild ancestors in the mid 1970s. Developmental stages were defined as days post-fertilization or were determined by morphological criteria (34). For the test of stress inducibility, a minimum of three 1-year-old fish were placed as a group in containers with water of different temperatures (25, 38 and 40°C) or with water at pH 4.5 for 0.5, 1.0 or 2 h (29).

Tissue culture and transient transfection

For tissue culture of rivulus skeletal muscle and liver, fully matured 1-year-old individuals were anesthetized in ice-cold water and the skeletal muscle and liver tissues were dissected in individual 100 mm Petri dishes containing PBS with 2× penicillin/streptomycin. After washing several times with PBS, tissues were cultured in RPMI medium (Gibco-BRL) supplemented with 10% bovine calf serum at 30°C in a CO_2 incubator for 3–7 days. To determine transfection efficiency we transfected pCMV-EGFP or pCMV-lacZ using various transfection protocols and then determined the percentage of positive cells either by fluorescence microscopy in the case of EGFP or after X-gal staining in the case of lacZ. Although the efficiency was low (~1%), transfection using DOTAP reagents (Boehringer Mannheim) gave the best result. The cultured tissues were transfected using DOTAP liposomal transfection reagent with 1 µg of the reporter plasmid and 0.1 µg of a plasmid (pCMV-lacZ) bearing the cytomegalovirus enhancer/promoter upstream of the lacZ gene. The tissue mass was fragmented to ~1 mm^3 with a 25 gauge needle immediately before transfection. At 48 h after transfection, cells were harvested and RT–PCR analysis was performed to detect both lacZ and EGFP transcripts. As a transfection control, EGFP expression was normalized with that of lacZ. All transfections were repeated at least three times.

RNA preparation and RT–PCR analysis

Total RNAs were isolated using RNAzol B (Tel-Test). To purify total RNA from embryos and tissues, fresh samples were frozen in liquid nitrogen, and ground into powder with a mortar before solubilization in RNAzol B solution. RT–PCR assay was performed essentially as described (35), except that total RNA was treated with 20 µg/ml DNase I (DN-EP, Sigma) for 15 min before the RT reaction to remove contaminating DNA. Primers used for RT–PCR analysis were as follows: EGFP 1R/2R, lacZ 1R/2R and rm-hsc71m primer a/d or c/d (Table 1).

Extract preparation and gel mobility shift assay

Cellular S150 extracts from fish tissues were prepared according to the short protocol of Roy *et al.* (36). For preparing template DNAs, PCR products (A, B and C) amplified by primer pairs of P3/P4, P5/P6 and P7/P8 (Table 1), respectively,

were subcloned into pT7-blue vectors. The probe DNAs for rm-hsc71mE-boxes within the sub-elements in the 5′-distal sequence of *rm-Hsc71m* and the muscle creatine kinase gene enhancer (MCKE) were prepared by annealing chemically synthesized complementary oligonucleotides (Table 1). All other probes or competitors used in this study were chemically synthesized (Table 1). Binding reactions were according to the procedure previously described (37) with minor modifications. Briefly, 10 µg of S150 extracts were incubated with a [α-^{32}P]dCTP-labeled probe for 15 min at room temperature in the presence of 1 or 2 µg poly(dI–dC) (Pharmacia) as carrier. The reaction mixtures were electrophoresed on a 5% native polyacrylamide gel, and then dried gels were autoradiographed.

Southwestern hybridization

Southwestern hybridization was performed by the procedure of Silva *et al.* (38) with minor modifications. S150 protein extracts prepared from Rivulus liver and muscle tissues were fractionated on 15% SDS–PAGE with a 4% stacking gel, along with a protein molecular weight standard (Bio-Rad). After electrophoresis, the gels were washed three times in 1 h in renaturation buffer with gentle agitation, and the proteins were transferred onto nitrocellulose filters by electroblotting. To prevent non-specific binding, the filters were blocked by shaking in renaturation buffer for 2 h. Hybridization was carried out in a sealed plastic bag with gentle agitation at room temperature for 3 h. Filters were briefly rinsed three times with binding buffer, and DNA–protein complexes were visualized by autoradiography.

Histochemistry and RNA *in situ* hybridization

A 414 bp and a 690 bp cDNA fragment containing part of exons II–IV and part of initial non-coding exons I–IV of *rm-hsc71m* amplified by RT–PCR with the rm-hsc71m primer c/d and a/d, were subcloned into pGEM-*Xho*I, respectively. Sense and antisense RNA probes were generated by transcribing the linearized templates *in vitro* with digoxigenin-11-UTP using T7 and Sp6 RNA polymerases, respectively, according to the manufacturer's instructions (Genius system, Boehringer Mannheim). For whole-mount *in situ* hybridization, Rivulus embryos were prepared at the age of 3 or 13 days after fertilization as described by Harland (39). To prepare paraffin-embedded whole body sections of adult fish, we used the method described by Stern and Holland (40). Fish histology was guided by the description of Groman (41). Nucleic acids were unmasked by exposing tissue sections to 0.002 N HCl/0.01% Triton X-100 at room temperature for 90 s and fixed again with 4% paraformaldehyde for 20 min at 4°C. Whole body embryos and tissue sections were treated with proteinase K (50 µg/ml in PBS) for 7.5 min at 37°C and then washed with PBS containing 2 mg/ml glycine. Hybridization and washing of tissue sections were as described by Springer *et al.* (42). Hybridization was monitored with the alkaline phosphatase-conjugated anti-digoxigenin antibody (Genius system). To inhibit endogenous phosphatase activity, levamisole (2.4 mg/10 ml, Sigma) was added to the color reaction.

RESULTS

Molecular genomic cloning of the *R.marmoratus* Hsc71 (*rm-Hsc71m*)

To examine heat shock gene expression in Rivulus, we isolated *hsp70*-related genes from a Rivulus genomic library (30). Rivulus genomic clones were screened by low stringency hybridization using a human *hsp70* cDNA as a probe. Seventy-six clones were isolated in the primary screen and were placed into three groups depending on the intensity of the hybridization signal. Restriction enzyme analysis and Southern blotting of the strong signal group identified six clones with similar patterns (data not shown). One was sequenced, and contained sequences encoding the *R.marmoratus hsc71* gene, which showed high muscle-specific expression (see below); hence it was designated *rm-hsc71m*.

To characterize *rm-hsc71m*, we sequenced 11 041 nt of the 13.8 kb insert DNA. As indicated in the sequence deposited in GenBank (accession no. AF227986), the clone contained one putative gene composed of eight exons capable of encoding 655 amino acids and ~4.4 kb of upstream sequences. A homology search of this protein using the Blast program revealed its high homology to the cytosolic and constitutively expressed Hsc70 family (data not shown). Similar to other cytosolic and constitutively expressed Hsc70 family genes, *rm-hsc71m* has one non-coding exon at 1704 bp upstream of the first coding exon. A 690 bp RT–PCR product was obtained when total RNA was subjected to RT–PCR analysis with a primer in the putative non-coding exon and a primer in the third coding exon (Fig. 1A and data not shown). The splice junction of the first non-coding exon was confirmed by direct sequencing of the RT–PCR products (data not shown). As predicted, primer extension analysis showed that the transcription start site was 68 bp upstream of the splice donor of the non-coding exon I (data not shown). Thus, the *rm-hsc71m* gene is composed of nine exons including the first non-coding exon and eight introns (Fig. 1B). It is noteworthy that the first intron of fish is much longer than the first introns seen in mammals, and the fourth intron of Rivulus (1580 bp) is the longest among the known species.

External stress dependency of the *rm-hsc71m* gene expression

To determine whether *rm-hsc71m* is induced by external stress, we examined its temporal expression pattern in response to stress in 3-month-old fish by RT–PCR. Primer sequences for RT–PCR analysis were designed to detect selectively *rm-hsc71m* transcripts. One primer contains sequences from +17 to +36 in the initial non-coding exon I which is not conserved at all, whereas other primers in exon II or IV is one of the least conserved sequences among the known Hsc70 families in other species (see Materials and Methods, and Discussion). Indeed, these primers were unique to show a single cDNA product in RT–PCR of rivulus total RNA (data not shown). For induction by either pH-shock or heat-shock, rivulus were reared in water at pH 4.5 for 0.5, 1 or 2 h, or in water temperatures of 25, 38 or 40°C for 1 h, respectively. Under these conditions expression of *rm-hsc71m* was unchanged although the basal levels of *rm-hsc71m* expression
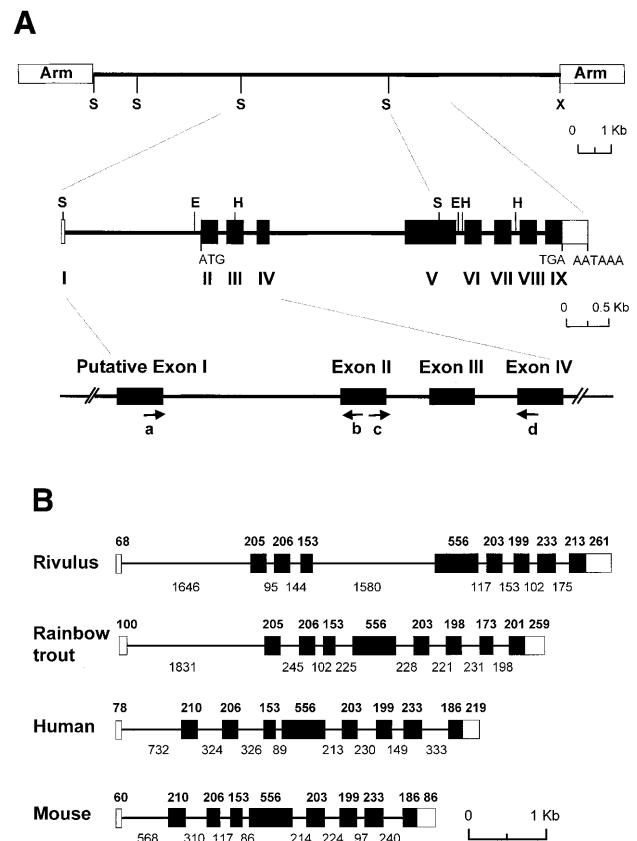

Figure 1. Genomic cloning of the *R.marmoratus hsc71* gene. (**A**) Schematic representation of the restriction map and gene structure of the *R.marmoratus hsc71* gene (*rm-hsc71m*). The open box represents the untranslated exon, whereas the filled boxes are the protein-coding exons. The translation start site (ATG), termination codon (TGA) and poly(A) signal sequences (AATAAA) are also indicated. A putative transcription start site was deduced using CGG genomic analysis web tools. Horizontal arrows with letters represent the location, name and direction of oligonucleotide primers. Vertical upward lines indicate restriction enzyme sites. E, *Eco*RI; H, *Hin*dIII; S, *Sac*I; X, *Xho*I. (**B**) Comparison of the *rm-hsc71m* gene structure to homologous genes from other species. Exons are shown as boxes, in which open boxes represent untranslated regions and introns are indicated as lines between the boxes. The numbers above and below the drawing represent the nucleotide numbers of each exon and intron, respectively. Accession numbers of the sequences are: Rivulus, AF227986; Rainbow trout, S85730; Human, Y00371; Mouse, U73744.

varied among tissues examined (data not shown). Similar results were obtained in experiments using fish of different ages (data not shown) indicating that expression of *rm-hsc71m* is not induced by acute external stress.

rm-hsc71m expression during developmental stages

To examine how *rm-hsc71m* expression is regulated in early development, total RNA was prepared from unfertilized eggs through juvenile stage larvae, and the steady state levels of *rm-hsc71m* transcripts were analyzed by RT–PCR. As shown in Figure 2A, the expression of *rm-hsc71m* was apparent in 6-day-old embryos (lane 4), and the highest expression was seen in 10-day-old embryos (lane 8). It is possible that inability to observe *rm-hsc71m* expression in younger embryos is due to cell- or tissue-specific expression of low but undetectable levels of *rm-hsc71m*. To examine cell- or tissue-specific expression of *rm-hsc71m*, whole-mount *in situ* hybridization
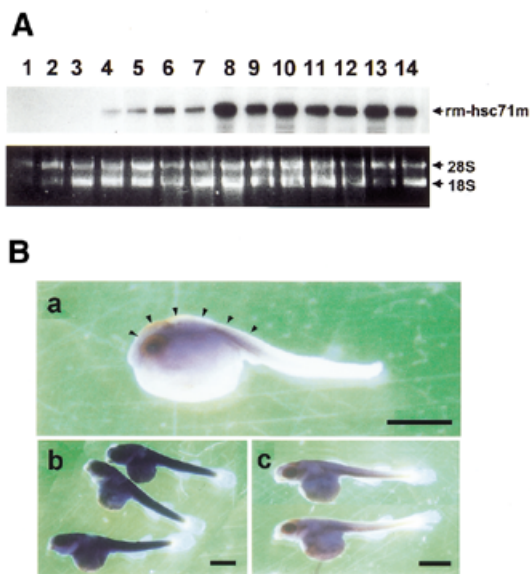
Figure 2. Expression of the *rm-hsc71m* gene is modulated during early development. (**A**) Expression of *rm-hsc71m* during early development. Total RNAs were prepared from developing embryos or larvae at each stage of development (*n* = 3–10) and used to monitor the level of *rm-hsc71m* expression by RT−PCR using a set of *rm-hsc71m*-specific primers in exon II and exon IV, respectively (rm-hsc71m primer c/d in Table 1). Staging of embryos was determined by the criteria of Harrington (34). As a loading control, photographs of ethidium bromide-stained 28S and 18S rRNA are also shown in the lower panel. Lane 1, unfertilized eggs; lane 2, blastula embryos; lane 3, 2-day-old embryos; lanes 4−12, 6- to 14-day-old embryos; lane 13, hatched larvae; lane 14, juvenile stage larvae. (**B**) Lateral views of a whole-body of a 3-day-old (a) and a 13-day-old embryo (b) showing spatial expression patterns of *rm-hsc71m* detected by whole-mount *in situ* hybridization with an antisense RNA probe containing a 414 base sequences corresponding to exons II−IV of *rm-hsc71m* (see Materials and Methods). Control 13-day-old embryos (c) were hybridized with a sense probe. High expression of *rm-hsc71m* in the head-fold and trunk region is marked by an arrowhead. Bar, 100 µm.

Figure 3. *rm-hsc71m* expression is differentially regulated in a tissue-specific manner. (**A**) Expression profile of *rm-hsc71m* in adult tissues. Total RNAs were isolated from several organs of a 1-year-old fish and subjected to RT–PCR analysis using a set of *rm-hsc71m*-specific primers as described in the legend to Figure 2A. The lower panel shows photographs of ethidium bromide-stained 28S and 18S rRNAs as a loading control. (**B**) Views of paraffin-sectioned 1-year-old fish showing spatial expression patterns of *rm-hsc71m* detected by *in situ* hybridization with antisense RNA probes as described in Figure 2B. Sections were stained with hematoxylin/eosin (a, d, f and h) or hybridized with an antisense (b, e, g and i) *rm-hsc71m* RNA probe. As a negative hybridization control, a sense *rm-hsc71m* RNA probe was also employed (c). Arrows indicate regions showing strong expression of *rm-hsc71m* (open arrows in panels e, g and i indicate skeletal muscle tissue showing the strongest expression). B, brain; CL, cerebellum (corpus cerebelli); E, esophagus; GF, gill filament; I, infundibulum; L, liver; MO, medulla oblongata; O, oral cavity; OL, olfactory lobe (telencephalon); ON, optic nerve; OT, optic tectum; S, spine; SK, skeletal muscle. Scale bars: a–c, 1.25 µm; d–i, 0.32 µm.

assay was performed. RNA probes were prepared as described in Materials and Methods. Whole-mount *in situ* hybridization showed *rm-hsc71m* expression in regions of the head and in somites of 3-day-old embryos (Fig. 2B, panel a). Similarly, strong expression of *rm-hsc71* was apparent in the entire body in 13-day-old embryos and comparable signals were not seen using a sense RNA probe (compare panels b and c in Fig. 2B). Thus, these data suggest that *rm-hsc71m* expression is regulated temporally and spatially during early development, such that early, low expression begins at day 3 in the head and somites and maximal expression occurs in the entire area of the trunk starting at day 10.

To determine whether *rm-hsc71m* has any tissue-tropism in its expression in adult fish, RT–PCR analysis was performed with RNAs isolated from adult tissues. As shown in Figure 3A, the strongest expression was in skeletal muscle, whereas lower expression levels were in the gill, eye and brain. Such tissue-specific expression was confirmed by RNA *in situ* hybridization of paraffin-sectioned 1-year-old fish. Similar to RT–PCR analysis, expression of *rm-hsc71m* was prominent in skeletal muscle (Fig. 3B, panels b, e and i), whereas lower expression levels were seen in gill filament and brain tissues (Fig. 3B, panels b and g). Thus, our data indicate that *rm-hsc71m* expression is regulated from early embryonic stages to adult stages
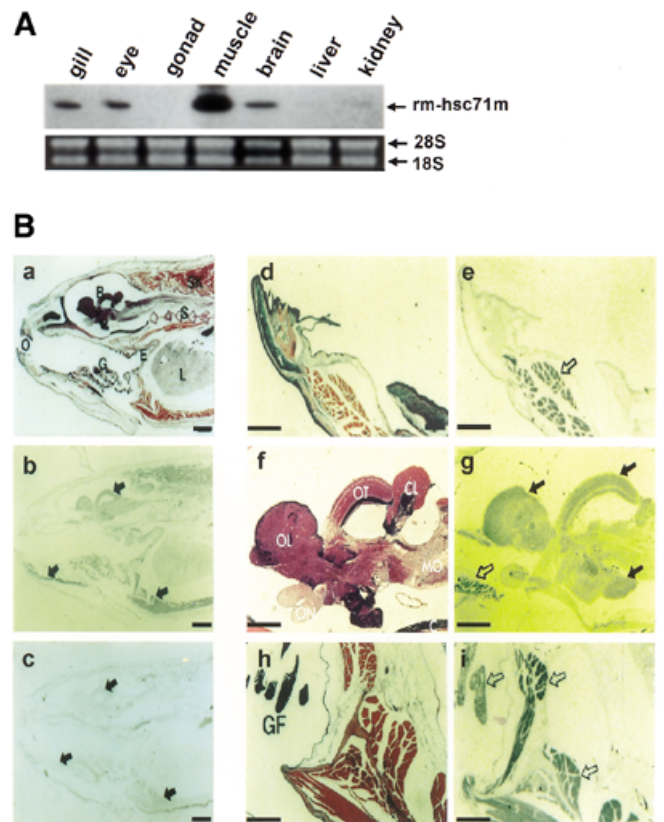
with strong tissue-tropism. The data also suggest that *rm-hsc71m* may play a role in development of skeletal muscle or in the maintenance of muscle tissue in adult Rivulus.

Identification of the enhancer region for the muscle-specific expression of *rm-hsc71m*

To identify specific regions of the genomic clone containing elements responsible for high muscle-specific expression of *rm-hsc71m*, we used transient transfection of Rivulus primary skeletal muscle tissue. Liver tissue, in which expression of *rm-hsc71m* is negligible, was included as a negative control. Tissue culture was conducted as described in Materials and Methods. Following co-transfection of varying amounts of pCMV-EGFP with a constant amount of pCMV-lacZ, EGFP expression was measured by RT–PCR and normalized to lacZ
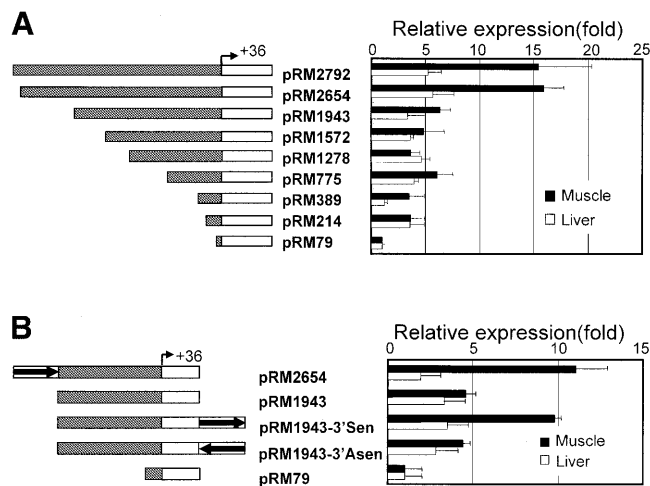
Figure 4. Identification of a muscle-specific regulatory region in *rm-hsc71m* upstream sequences. (**A**) Examination of the activity of 5′-upstream sequences in muscle based on reporter activity. Schematic drawings of a series of 5′-deletion constructs containing the EGFP reporter gene flanked by variable lengths of *rm-hsc71m* 5′-upstream sequences, which are shown on the left side. The relative levels of reporter gene expression in liver and muscle tissues are shown on the right side. The constructs were transiently co-transfected into cultured Rivulus muscle or liver tissue, along with a pCMV-lacZ control vector. The level of EGFP expression was monitored 48 h after transfection by RT–PCR. Transfection efficiency was normalized to the level of lacZ expression. In all cases, EGFP/lacZ values obtained from at least two different plasmid preparations and three experiments were normalized to that of the basal construct (pRM79 containing only a TATA box). (**B**) Position- or orientation-dependency of the rMME. Reporter constructs containing the rMME (1.2 kb *Pvu*II fragment from −2617 to −1418 bp of rm-hsc71m) downstream of the EGFP reporter in a sense or antisense orientation were transiently transfected into cultured muscle or liver tissue. Schematic drawings of each construct and their relative expression levels are shown in the left and right panels, respectively.

Figure 5. Identification of muscle-specific binding activity in the rMME of *rm-hsc71m*. (**A**) Schematic representation of the rMME (from −2654 to −1943) of *rm-hsc71m* is shown. Three putative muscle-specific factor-binding sub-elements identified by a DNASIS program were designated A, B and C, respectively. The rm-hsc71mE element represents sequences similar to those seen in the MCKE. (**B**) Electrophoretic mobility shift assay shows muscle-enriched rm-hsc71mE binding activity distinct from MCKE binding activity. A gel mobility shift assay was performed using a MCKE or an rm-hsc71mE probe in the presence of rivulus muscle or liver S150 extracts to assay for muscle-specific rm-hsc71mE binding activity. To determine the specificity of rm-hsc71mE binding activity, molar excess amounts of cold rm-hsc71mE competitor (10× and 50×) were included in the reaction. Filled and unfilled arrowheads show specific and non-specific binding activities, respectively. (**C**) A single predominant rm-hsc71mE binding activity exists in the B region. To look for additional binding activities in the B region, the B fragment was used as a probe in a gel shift assay in the presence of Rivulus muscle S150 extracts. To determine the specificity of binding, 10× or 50× molar excess amounts of cold competitor, B or rm-hsc71mE, were added during the assay. The arrowhead indicates the shifted binding activity. (**D**) Regions A and C also show muscle-specific binding activity similar to rm-hsc71mE binding activity in the B region. To determine whether muscle-specific DNA binding activities exist in A and C regions, probe DNAs prepared from A and C regions were used in gel shift assays containing Rivulus muscle or liver S150 extracts. The arrowhead indicates a shifted band.

expression. EGFP expression was always proportional to the concentration of pCMV-EGFP (data not shown), indicating that despite its low efficiency, transfection of cultured fish tissues is feasible.

To identify skeletal muscle regulatory elements, transfections were performed using 5′ deletion fragments containing the *rm-hsc71m* promoter linked to the EGFP reporter. The activity of each construct was expressed relative to that of the pRM79 construct (Fig. 4A). Transcriptional activity of the pRM79 construct containing an E-box and a TATA box was similar in both liver and muscle (data not shown). The relative transcriptional activity of pRM2792 and pRM2654 in muscle tissue was 3-fold higher than in liver. However, a 5′ deletion up to −1943 (pRM1943) reduced the transcriptional activity in muscle to a level seen in liver. Deletion of downstream sequences (−1572, −1278, −775, −389 and −214) did not significantly reduce promoter activity in either tissue. These data indicate that a 712 bp region between −2654 and −1943 functions as a muscle-specific element in the *rm-hsc71m* gene; therefore we designate this region as a Rivulus major muscle element (rMME).

To confirm that rMME is the major muscle-specific regulatory element, two reporters (pRM1943-3′Sen and pRM1943-3′Asen) containing rMME cloned 3′ of the reporter in both sense and antisense orientations, respectively, were constructed and analyzed in the transient assay (Fig. 4B).
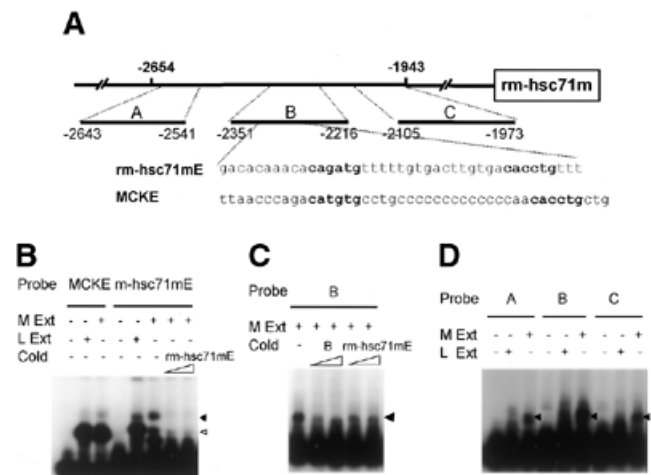
Activation by pRM1943-3′Sen was comparable to that of RM2654, whereas activation by pRM1943-3′Asen was comparable to that of pRM1943. Thus, this result suggests that the rMME is a muscle-specific upstream promoter element although we do not exclude the possibility that the rMME is part of an enhancer (see Discussion).

Identification of muscle-specific factor binding sites in the rMME of *rm-hsc71m*

To identify specific factors that might regulate *rm-hsc71m* expression in muscle, we searched mammalian databases for putative transcription factor binding sites within the genomic clone. We found several potential binding sites for muscle-specific factors, namely, four E-box (CANNTG) motifs, one element for transcriptional enhancer factor 1 and one myocyte enhancer factor 2/serum response element (43–46), within the rMME region. Of the four E-boxes, three are located in sub-region B (from −2417 to −2215 nt), whereas the fourth one is in sub-region A (Fig. 5A). Interestingly, sequences from −2351 to −2267 nt of sub-region B (rm-hsc71mE box) were similar to the sequence and spacing of two E-boxes found in the MCKE,

which is one of the best-characterized skeletal muscle cell-specific regulatory elements (45–47) (Fig. 5A). Thus, we examined whether binding to the rm-hsc71mE box was similar to that seen in the MCKE by an electrophoretic shift assay using probes derived from both the MCK enhancer and rm-hsc71mE. Strong MCKE-binding activity was observed in both muscle and liver extracts, although a weaker muscle-specific complex was also observed (Fig. 5B). The strong complex formed on the MCKE probe was competed with excess amounts of the cold MCKE but not with the rm-hsc71mE (data not shown). In contrast, a muscle-specific complex with different mobility from that seen in liver was formed on the rm-hsc71mE probe with muscle extracts and competed by excess cold rm-hsc71mE competitor (Fig. 5B, lanes 6–8) but not by the MCKE competitor (data not shown). This muscle-specific binding activity is the major one observed in region B (Fig. 5C), and it is also observed with two other DNA fragments A and C (Fig. 5D). These complexes were competed by rm-hsc71mE and DNA fragment B cold competitors (Fig. 5B and data not shown). It is noteworthy that an E-box motif is not found in region C. Thus, this investigation suggests that sequences responsible for muscle-specific rm-hsc71mE binding activity are not E-box motifs and that E-box binding activity may not be sufficient for muscle-specific high expression of *rm-hsc71m* in Rivulus.

To identify specific sequences responsible for muscle-specific binding activity within the rm-hsc71mE, we searched for homologous regions on both strands within regions A, B and C, and identified the consensus sequence TGTnACA (Fig. 6D). To determine whether these sequences were responsible for muscle-specific binding activity, electrophoretic gel shift assays were performed using double-stranded probe DNAs containing conserved sequences from each of these regions. As a negative control, a right half E-box from the MCKE (MCKE-RH) was also included. Indeed, as shown in Figure 6A, a muscle-specific complex, similar to that found in the rm-hsc71mE probe, was formed in every reaction containing each probe, but not on the MCKE-RH probe. The binding affinity was of the order of b box > c box > a box, and was recapitulated in the competitive gel shift analysis (Fig. 6B). Furthermore, when a double-stranded mutant oligonucleotide (m box), in which all of the central conserved sequences of the b box were altered (TGTGACA→ACAGTGT) was used as a probe, muscle-specific rm-hsc71mE binding activity disappeared (Fig. 6C). On the other hand, the bm box with three base substitutions in the central conserved region of the b box was a poor probe for the muscle-specific factor, but was a fairly good competitor for the b box (Fig. 6A and B). These observations suggest that the central TGTnACA sequence is essential for muscle-specific factor recognition even though sequences flanking the central TGTnACA may also participate in binding activity.

Identification of a muscle-specific *trans*-acting factor recognizing a TGTnACA box

A southwestern assay was employed to identify and determine the size of the putative muscle-specific factor interacting selectively with the TGTnACA box DNA. Rivulus muscle and liver S150 extracts (100 μg) were separated on an SDS–polyacrylamide gel (Fig. 7A), and a southwestern assay was performed using the double-stranded probe containing either central
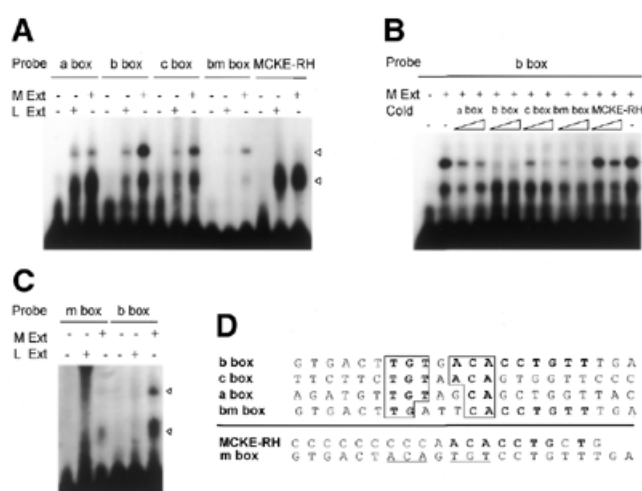

Figure 6. A novel muscle-specific factor is responsible for recognition of TGTnACA sequences in all three 5′-distal sub-elements of the rMME of *rm-hsc71m*. (**A**) Electrophoretic mobility shift analysis showing binding specificity of a novel muscle-specific factor to rm-hsc71mE-related sequences in each sub-region of the rMME of *rm-hsc71m*. The rm-hsc71mE-related sequences in each region (a box, b box and c box, respectively; see panel D for each sequence) were identified by a computer search. The bm box is derived from the b box with three base changes. The MCKE-RH is the right half E-box in the MCKE (see Table 1). Each box was used as a probe in a gel shift analysis containing Rivulus muscle or liver S150 extracts. Specific and non-specific bands are marked by filled and unfilled arrowheads, respectively. (**B**) Competitive electrophoretic gel mobility shift analysis showing the relative binding affinity of the muscle-specific factor to each box. Increasing amounts (0, 10× and 50×) of a molar excess of cold competitor were added to the reaction containing the b box probe and Rivulus muscle S150 extracts. Specific and non-specific bands are marked by filled and unfilled arrowheads, respectively. (**C**) Electrophoretic mobility shift analysis showing that the central TGTnACA sequences are critical for muscle-specific factor binding. A double-stranded mutant oligonucleotide (m box; see panel D) in which all of the central conserved sequences of b box were altered (TGTGACA→ACAGTGT) was used as probe, along with parent b box probe, containing Rivulus muscle and liver S150 extracts. (**D**) Comparison of sequences recognized by a muscle-specific factor (only a top strand of sequences is shown: a, b and c boxes corresponded to sequences from –2627 to –2606, from –2296 to –2275 and from –2009 to –1988, respectively). Consensus binding bases are indicated as a boldface letter in a box. Boldface letters indicate homologous sequences in the MCKE-RH and the b box, and the underlined sequences in the m box indicate mutagenized bases from the b box.

TGTnACA sequences (b box) or mutant ones (m box) (see Fig. 6D). A band with an apparent M_r of 32 000 was detected in muscle extracts but not in liver extracts to the b box probe (Fig. 7B), whereas no band was detected in both extracts to the m box probe (Fig. 7C). This protein band was efficiently competed in a reaction containing excess amounts of cold competitors, such as DNA from boxes a, b and c, but not by MCKE DNA or by the m box (data not shown). Thus, these data indicate that a muscle-specific 32 kDa protein is abundant in Rivulus muscle and responsible for the binding to the TGTnACA box.

DISCUSSION

By low stringency screening of Mangrove rivulus genomic clones, we identified at least six Hsp70 family genes (data not shown). Here, we have isolated a genomic clone containing *rm-hsc71m*, which shows high muscle-specific expression.
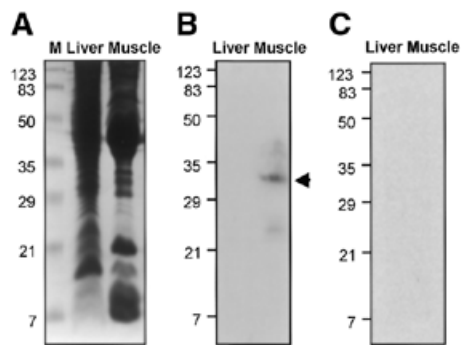
Figure 7. Identification of the TGTnACA box-binding protein in Mangrove rivulus muscle by southwestern blotting. S150 extracts from Mangrove rivulus liver and muscle tissues were subjected to SDS–PAGE and either stained with Coomassie brilliant blue (**A**) and filter-hybridized with an [α-^{32}P]dCTP-labeled b box probe (**B**) or m box probe (**C**) after electroblotting to the membrane. An arrow indicates a specific protein band hybridized with the probe DNA. Prestained protein molecular weight markers (Bio-Rad) were also electrophoresed in an adjacent lane.

The *rm-hsc71m* gene is composed of nine exons and eight introns including an initial non-coding exon, and its genomic structure and deduced amino acid sequences are similar to those of other cytosolic and constituvely expressed *hsc70* family genes (Fig. 1).

It is noteworthy that probes used for *in situ* hybridization are specific to *rm-hsc71m*, without knowing sequences of other Mangrove rivulus hsp70 homologs. Compared with all known Hsp70 family genes, the nucleotide sequences of *rm-hsc71m* coding exons have 48–88% homology whereas a non-coding first exon is not conserved at all (data not shown). Indeed, the tissue-specific expression profile of *rm-hsc71m* demonstrated by *in situ* hybridization using a probe corresponding to a 414 bp sequence in exons II–IV of *rm-hsc71m* was not different from that of RT–PCR (Fig. 3). Furthermore, the same result was obtained in the *in situ* hybridization using a probe corresponding to a 690 bp sequence in exons I–IV (data not shown and see Materials and Methods).

Like Hsc70 family genes in *Xenopus laevis* (16) and *Pleurodeles warltl* (17), the *rm-hsc71m* gene is not inducible at increased temperature or low pH (data not shown). On the other hand, *rm-hsc71m* mRNA is expressed at low levels in early embryogenesis, showing only localized expression in a fraction of somites and in the central nervous system of 3-day-old embryos. Maximal expression occurs in the whole trunk of the fish starting at day 10 (Fig. 2). The 10-day-old embryo exhibits morphological and behavioral characteristics, such as tail beating and the disappearance of the extra-embryonic membrane (34). Extra-embryonic membranes (chorion, allantois and yolk sac) protect the embryo from the environment and provide nutrients. Thus, high muscle-specific expression of *rm-hsc71m* correlates with muscle growth and differentiation and with adaptive responses to muscle-specific physiological stresses such as the demand for oxidative metabolism.

The strongest expression of *rm-hsc71m* is in skeletal muscle, while moderate expression is seen in brain, eye and gill, and little or no expression is seen in liver and kidney of adult fish (Fig. 3). Similar but distinct tissue tropic expression of *hsc70* family gene has been reported in other species. Strong expression

of zebrafish *hsc70* mRNA is detected in regions of the central nervous system, the eye and in differentiating somites (22). A similar pattern of *hsc70* expression has been reported in *Drosophila* embryo (18). In mouse, expression of *hsc70* is ubiquitous (13), whereas the *hsc73* gene is expressed in a pattern correlated with neurogenesis (19). Chicken *hsc70* transcripts are largely restricted to neuroectoderm- and mesoderm-derived structures from gastrulation to early organogenesis (20). These observations suggest that *hsc70* has specific developmental functions in addition to its role in protein folding.

To understand regulation of muscle-specific expression of *rm-hsc71m*, a regulatory region 5′ of the *rm-hsc71m* coding sequences was defined using transient transfection analysis in cultured Mangrove rivulus tissues. The rMME containing 712 bp DNA from –2654 to –1943 was identified as a major element for muscle-specific expression (Fig. 4). Because the rMME functions only in a sense orientation when it is located 3′ of a reporter gene, the rMME might be a simple upstream promoter element. However, we do not exclude the possibility that the rMME is part of a muscle-specific enhancer. Thus, to understand the tissue tropic elements in *rm-hsc71m*, regulatory modules specific for the central nervous system, eye and gill expression should also be characterized.

Muscle gene expression is thought to occur via interactions between muscle-specific and ubiquitous transcription factors with a common subset of *cis*-regulatory elements (48,49). Because the most widely described muscle gene control elements are sequences containing the canonical E-box motif (45–47), we initially expected that E-box motifs in the rMME would be required for activity. However, a novel motif with the consensus of TGTnACA was identified as a major binding site for a muscle-specific factor having an M_r value of 32 000 (Figs 5–7). We do not know whether the interaction between the TGTnACA motif and the novel 32 kDa factor alone is sufficient for muscle-specific expression. Because there was only a single band shifted by crude cellular extracts of muscle tissue, the TGTnACA box-binding factor may be an abundant protein in Mangrove rivulus muscle tissue. However, further studies, such as fractionation of proteins and footprinting analyses, are required to fully understand the molecular mechanism of muscle-specific expression of *rm-hsc71m*.

Is the TGTnACA box a novel muscle-specific regulatory element conserved during evolution? Our preliminary data suggest that a muscle-specific binding activity is conserved in fish, but not in mammals. We observed a muscle-specific rm-hsc71mE-binding activity in the electrophoretic gel shift assay or in southwestern blot analysis using muscle extracts of sea bass (*Lateolabrzx japinicus*), bastard halibut (*Paralichthys olivaceus*), armorclad rockfish (*Sebastes hubbsi*), snake head (*Channa argus*) or mud loach (*Misgurnus mizolepis*), but not of mouse (data not shown).

It is noteworthy that our novel *cis*-element responsible for muscle-specific expression is similar to the *cis*-acting unfolded protein response element (UPRE) in yeast (50). UPRE is also a palindrome with a spacer of 1 nt but with a different palindromic sequence (CACCGTG). UPRE acts as a *cis*-acting element for the small basic-leucine zipper transcription factor, Hac1p/Ern4p (50–52), which is necessary for induction of Kar2p, an *Escherichia coli* DnaK homolog. Hac1 expression is induced by endoplasmic reticulum stress triggered by accumulation of unfolded proteins (53,54). A single UPRE sequence

in the promoter of at least six luminal proteins involved in the unfolded protein response (UPR) is necessary and sufficient for transcriptional induction (55). Thus, the high muscle-specific expression of *rm-hsc71* gene may be comparable to a UPR in yeast. We are currently investigating mechanistic interactions between our novel muscle-specific *cis*-element and its *trans*-acting factor.

ACKNOWLEDGEMENTS

REFERENCES

1. Morimoto,R.I., Tissieres,A. and Georgopoulos,C. (1994) *The Biology of Heat Shock Proteins and Molecular Chaperones*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Feder,M.E. and Hofmann,G.E. (1999) Heat-shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annu. Rev. Physiol.*, **61**, 243–282.
3. Sorger,P.K. (1991) Heat shock factor and the heat shock response. *Cell*, **65**, 363–366.
4. Boorstein,W.R., Ziegelhoffer,T. and Craig,E.A. (1994) Molecular evolution of the HSP70 multigene family. *J. Mol. Evol.*, **38**, 1–17.
5. Loones,M.T. and Morange,M. (1998) Hsp and chaperone distribution during endochondral bone development in mouse embryo. *Cell. Stress Chaperones*, **3**, 237–244.
6. Masuda,H., Hosokawa,N. and Nagata,K. (1998) Expression and localization of collagen-binding stress protein Hsp47 in mouse embryo development: comparison with types I and II collagen. *Cell. Stress Chaperones*, **3**, 256–264.
7. Giudice,G., Sconzo,G. and Roccheri,M.C. (1999) Studies on heat shock proteins in sea urchin development. *Dev. Growth Differ.*, **41**, 375–380.
8. Christians,E., Michel,E. and Renard,J.-P. (1997) Developmental control of heat shock and chaperone gene expression. Hsp 70 genes and heat shock factors during preimplantation phase of mouse development. *Cell. Mol. Life Sci.*, **53**, 168–178.
9. de la Rosa,E.J., Vega-Nunez,E., Morales,A.V., Serna,J., Rubio,E. and de Pablo,F. (1998) Modulation of the chaperone heat shock cognate 70 by embryonic (pro)insulin correlates with prevention of apoptosis. *Proc. Natl Acad. Sci. USA*, **95**, 9950–9955.
10. Freeman,B.C. and Morimoto,R.I. (1996) The human cytosolic molecular chaperones hsp90, hsp70 (hsc70) and hdj-1 have distinct roles in recognition of a non-native protein and protein refolding. *EMBO J.*, **15**, 2969–2979.
11. Ziemienowicz,A., Zylicz,M., Floth,C. and Hubscher,U. (1995) Calf thymus Hsc70 protein protects and reactivates prokaryotic and eukaryotic enzymes. *J. Biol. Chem.*, **270**, 15479–15484.
12. Dubrovsky,E.B., Dretzen,G. and Berger,E.M. (1996) The Broad-complex gene is a tissue-specific modulator of the ecdysone response of the *Drosophila hsp23* gene. *Mol. Cell. Biol.*, **16**, 6542–6552.
13. Hunt,C.R., Parsian,A.J., Goswami,P.C. and Kozak,C.A. (1999) Characterization and expression of the mouse Hsc70 gene. *Biochim. Biophys. Acta*, **1444**, 315–325.
14. Tavaria,M., Gabriele,T., Kola,I. and Anderson,R.L. (1996) A hitchhiker's guide to the human Hsp70 family. *Cell. Stress Chaperones*, **1**, 23–28.
15. Hahnel,A.C., Gifford,D.J., Heikkila,J.J. and Schultz,G.A. (1986) Expression of the major heat shock protein (hsp70) family during early mouse embryo development. *Teratog. Carcinog. Mutagen.*, **6**, 493–510.
16. Lang,L., Miskovic,D., Lo,M. and Heikkila,J.J. (2000) Stress-induced, tissue-specific enrichment of *hsp70* mRNA accumulation in *Xenopus laevis* embryos. *Cell. Stress Chaperones*, **5**, 36–44.
17. Delelis-Fanien,C., Penrad-Mobayed,M. and Angelier,N. (1997) Molecular cloning of a cDNA encoding the amphibian *Pleurodeles waltl*

18. Perkins,L.A., Doctor,J.S., Zhang,K., Stinson,L., Perrimon,N. and Craig,E.A. (1990) Molecular and developmental characterization of the heat shock cognate 4 gene of *Drosophila melanogaster*. *Mol. Cell. Biol.*, **10**, 3232–3238.
19. Walsh,D., Li,Z., Wu,Y. and Nagata,K. (1997) Heat shock and the role of the HSPs during neural plate induction in early mammalian CNS and brain development. *Cell. Mol. Life Sci.*, **53**, 198–211.
20. Vega-Nunez,E., Pena-Melian,A., de la Rosa,E.J. and de Pablo,F. (1999) Dynamic restricted expression of the chaperone Hsc70 in early chick development. *Mech. Dev.*, **82**, 199–203.
21. Zafarullah,M., Wisniewski,J., Shworak,N.W., Schieman,S., Misra,S. and Gedamu,L. (1992) Molecular cloning and characterization of a constitutively expressed heat-shock-cognate *hsc71* gene from rainbow trout. *Eur. J. Biochem.*, **204**, 893–900.
22. Santacruz,H., Vriz,S. and Angelier,N. (1997) Molecular characterization of a heat shock cognate cDNA of zebrafish, *hsc70*, and developmental expression of the corresponding transcripts. *Dev. Genet.*, **21**, 223–233.
23. Eisen,J.S. (1996) Zebrafish make a big splash. *Cell*, **87**, 969–977.
24. Nusslein-Volhard,C. (1996) Gradients that organize embryo development. *Sci. Am.*, **275**, 54–61.
25. Harrington,R.W.,Jr (1961) Oviparous hermaphroditic fish with self internal fertilization. *Science*, **134**, 1749–1750.
26. Laughlin,T.F., Lubinski,B.A., Park,E.-H., Taylor,D.S. and Turner,B.J. (1995) Clonal stability and mutation in the self-fertilizing hermaphroditic fish, *Rivulus marmoratus*. *J. Hered.*, **86**, 399–402.
27. Harrington,R.W.,Jr (1971) How ecological and genetic factors interact to determine when self-fertilizing hermaphrodites of *Rivulus marmoratus* change into functional secondary males, with a reappraisal of the modes of intersexuality among fishes. *Copeia*, **1971**, 389–432.
28. Lindsey,C.C. and Harrington,R.W.,Jr (1972) Extreme vertebral variation induced by temperature in a homozygous clone of the self-fertilizing fish *Rivulus marmoratus*. *Can. J. Zool.*, **50**, 733–744.
29. Kim,A.-R., Kweon,H.-S., Noh,J.-K., Choi,M.-G. and Park,E.-H. (1993) Tolerance to acidic and alkaline water of laboratory-reared hermaphroditic fish *Rivulus maromoratus*. The 4th Indo-Pacific Fish Conference Program and Abstracts of Papers, p. 114.
30. Lee,J.-S., Choe,J. and Park,E.-H. (1994) Absence of the intron-D-exon of c-Ha-ras oncogene in the hermaphroditic fish *Rivulus marmoratus* (Teleostei: Rivulidae). *Biochem. Mol. Biol. Int.*, **34**, 921–926.
31. Williams,J.B. and Mason,P.J. (1985) Primer extension methods. In Homes,B.D. and Higgins,S,J. (eds), *Nucleic Acid Hybridization: A Practical Approach*. IRL Press, Oxford, UK, pp. 139–160.
32. Park,E.-H. and Kim,D.S. (1984) Hepatocarcinogenicity of diethylnitrosamine to the self-fertilizing hermaphroditic fish *Rivulus marmoratus* (Teleostomi: Cyprinodontidae). *J. Natl Cancer Inst.*, **73**, 871–876.
33. Park,E.-H. and Yi,A.-K. (1989) Photoreactivation rescue and dark repair demonstrated in UV-irradiated embryos of the self-fertilizing fish *Rivulus marmoratus* (Teleostei; Aplocheilidae). *Mutat. Res.*, **217**, 19–24.
34. Harrington,R.W.,Jr (1963) Twenty-four hour rhythms of internal self-fertilization and of oviposition by hermaphrodites of *Rivulus marmoratus*. *Physiol. Zool.*, **36**, 325–341.
35. Kim,C.G. (1996) Knock-out effects of transcription factor GATA-1 on early erythropoiesis. *Mol. Cells*, **6**, 176–182.
36. Roy,R.J., Gosselin,P. and Guerin,S.L. (1991) A short protocol for micro-purification of nuclear proteins from whole animal tissue. *Biotechniques*, **11**, 770–777.
37. Kim,C.G., Swendeman,S.L., Barnhart,K.M. and Sheffery,M. (1990) Promoter elements and erythroid cell nuclear factors that regulate alpha-globin gene transcription *in vitro*. *Mol. Cell. Biol.*, **10**, 5958–5966.
38. Silva,C.M., Tully,D.B., Petch,L.A., Jewell,C.M. and Cildowski,J.A. (1987) Application of a protein-blotting procedure to the study of human glucocorticoid receptor interactions with DNA. *Proc. Natl Acad Sci. USA*, **84**, 1744–1748.
39. Harland,R.M. (1991) *In situ* hybridization: an improved whole-mount method for *Xenopus* embryos. *Methods Cell Biol.*, **36**, 685–695.
40. Stern,C.D. and Holland,P.W. (1993) *Essential Developmental Biology: A Practical Approach*. IRL Press, Oxford, UK.
41. Groman,D.B. (1982) *Histology of the Striped Bass, Monograph number 3*. American Fisheries Society, Bethesda, MD, USA.
42. Springer,J.E., Robbins,E., Gwag,B.J., Lewis,M.E. and Baldino,F.,Jr (1991) Non-radioactive detection of nerve growth factor receptor (NGFR)

70-kDa heat-shock cognate protein. *Biochem. Biophys. Res. Commun.*, **238**, 159–164.

mRNA in rat brain using *in situ* hybridization histochemistry. *J. Histochem. Cytochem.*, **39**, 231–234.

43. Perry,R.L. and Rudnicki,M.A. (2000) Molecular mechanisms regulating myogenic determination and differentiation. *Front. Biosci.*, **5**, D750–D767.

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

45. Buskin,J.N. and Hauschka,S.D. (1989) Identification of a myocyte nuclear factor that binds to the muscle-specific enhancer of the mouse muscle creatine kinase gene. *Mol. Cell. Biol.*, **9**, 2627–2640.

46. Lassar,A.B., Buskin,J.N., Lockshon,D., Davis,R.L., Apone,S., Hauschka,S.D. and Weintraub,H. (1989) MyoD is a sequence-specific DNA binding protein requiring a region of *myc* homology to bind to the muscle creatine kinase enhancer. *Cell*, **58**, 823–831.

47. Apone,S. and Hauschka,S.D. (1995) Muscle gene E-box control elements. Evidence for quantitatively different transcriptional activities and the binding of distinct regulatory factors. *J. Biol. Chem.*, **270**, 21420–21427.

48. Naya,F.S. and Olson,E. (1999) MEF2: a transcriptional target for signaling pathways controlling skeletal muscle growth and differentiation. *Curr. Opin. Cell Biol.*, **11**, 683–688.

49. Weintraub,H. (1993) The MyoD family and myogenesis: redundancy, networks, and thresholds. *Cell*, **75**, 1241–1244.

50. Mori,K., Kawahara,T., Yoshida,H., Yanagi,H. and Yura,T. (1996) Signaling from endoplasmic reticulum to nucleus: transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway. *Genes Cells*, **1**, 803–817.

51. Cox,J.S. and Walter,P. (1996) A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response. *Cell*, **87**, 391–404.

52. Nikawa,J., Akiyoshi,M., Hirata,S. and Fukuda,T. (1996) *Saccharomyces cerevisiae* IRE2/HAC1 is involved in IRE1-mediated KAR2 expression. *Nucleic Acids Res.*, **24**, 4222–4226.

53. Mori,K., Sant,A., Kohno,K., Normington,K., Gething,M.J. and Sambrook,J.F. (1992) A 22 bp *cis*-acting element is necessary and sufficient for the induction of the yeast KAR2 (BiP) gene by unfolded proteins. *EMBO J.*, **11**, 2583–2593.

54. Kohno,K., Normington,K., Sambrook,J., Gething,M.J. and Mori,K. (1993) The promoter region of the yeast KAR2 (BiP) gene contains a regulatory domain that responds to the presence of unfolded proteins in the endoplasmic reticulum. *Mol. Cell. Biol.*, **13**, 877–890.

55. Mori,K., Ogawa,N., Kawahara,T., Yanagi,H. and Yura,T. (1998) Palindrome with spacer of one nucleotide is characteristic of the cis-acting unfolded protein response element in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **273**, 9912–9920.