1838–1845 Nucleic Acids Research, 1994, Vol. 22, No. 10

Promoter elements and transcriptional control of the chicken β tropomycin gene

Madeleine Toutant, Cécile Gauthier-Rouviere[1], Marc Y.Fiszman* and Marguerite Lemonnier
Unité de Biochimie, Département de Biologie Moléculaire, C.N.R.S., Institut Pasteur, 25–28 rue du Dr Roux, 75724 Paris Cedex 15 and [1]CRBM, CNRS-INSERM, BP 5051 Montpellier Cedex, France

Received February 4, 1994; Revised and Accepted April 22, 1994

ABSTRACT

The chicken β tropomyosin (βTM) gene has two alternative transcription start sites (sk and nmCAP sites) which are used in muscle or non muscle tissues respectively. In order to understand the mechanisms involved in the tissue-specific and developmentally-regulated expression of the βTM gene, we have analyzed the 5′ regions associated with each CAP site. Truncated regions 5′ to the nmCAP site were inserted upstream to the bacterial chloramphenicol acetyl-transferase (CAT) reporter gene and these constructs were transfected into avian myogenic and non myogenic cells. The maximum transcription is driven by the CAT construct (–168/+216 nt) in all cell types. Previous deletion analysis of the region 5′ to the βTMskCAP site has indicated that 805 nt confer myotube-specific transcription. In this work, we characterize an enhancer element (–201/–68 nt) which contains an E box (–177), a variant CArG box (–104) and a stretch of 7Cs (–147). Mutation of any of these motifs results in a decrease of the myotube-specific transcriptional activity. Electrophoretic mobility shift assays indicate that these *cis*-acting sequences specifically bind nuclear proteins. This enhancer functions in an orientation-dependent manner.

INTRODUCTION

Tropomyosins (TM) are a family of highly conserved proteins which bind to actin. They are present in non muscle and in muscle tissue and different isoforms are characteristic of specific cell types. The diversity of TM isoforms is generated from a limited number of genes, each of which encodes several isoforms via multiple transcription start sites, alternative splicing and multiple 3′ end processing.

In previous studies, we have shown that the chicken βTM gene encodes three TM isoforms by the use of two promoters, alternative terminal exons and two mutually exclusive internal exons (1, 2). Two isoforms are specific to muscle tissue, one is present in smooth muscles (βTMsm), the other (βTMsk) accumulates in fast and slow skeletal muscles in variable amounts throughout development (3) with the exception of the Pectoralis muscle. In this last muscle, accumulation of βTMsk mRNA stops at hatching as a consequence of an arrest of transcription mediated in part by neuronal input (4). The two transcripts encoding βTMsk and βTMsm are synthesized from a unique CAP site (1). These transcripts differ in the use of alternative internal exons and in their 3′ coding and non coding regions. The third isoform (βTMnm) encoded by the βTM gene is expressed in non muscle cells and undifferentiated myoblasts and transcription of the cognate mRNA is initiated from a distinct region of the gene (2). This is in contrast with the mouse βTM gene where synthesis of all transcripts starts from a unique CAP site (5). Differing from both these genes, transcription of chicken αfastTM gene is initiated from multiple CAP sites in non myogenic and in myogenic cells (6).

In order to understand how the expression of the βTM gene is controlled, we have analyzed the 5′-flanking sequences associated with each CAP site. We show that transcription of the βTMnm in transfected cells is driven by an ubiquitous promoter whose activity correlates with the accumulation of the endogenous transcript in all cell types. Previous deletion studies of the region 5′ to the βTMskCAP site fused to the CAT gene (pβTMskCAT constructs) have indicated that 805 nt are sufficient to confer maximum tissue specific transcription. We show here that a subfragment of this sequence is an enhancer required for maximum muscle-specific activity. This enhancer, which functions in an orientation dependent manner, contains several *cis*-acting motifs which bind nuclear factors.

MATERIAL AND METHODS

In vitro transcription in isolated nuclei and hybridization to DNA

Nuclei of quail fibroblasts were prepared as reported previously (4), and RNAs elongated *in vitro* (4) in the presence of α-^{32}P UTP (Amersham, 3000 Ci/mmole) were hybridized to

*To whom correspondence should be addressed

recombinant plasmid DNAs transfered to nitrocellulose filters (7). After washings, filters were exposed to Kodak XAR films with intensifying screens.

Cell culture and transfections

Proliferating quail myoblasts were isolated from the hindlimb of 10-day embryonic quails and transformed with a temperature-sensitive mutant of the Rous sarcoma virus (8, 9). The non muscle cell line used was the QT6 quail fibroblast cell line (10). Myogenic cells were seeded at approximately 10^5 cells/6 cm plate while non muscle cells were seeded at 2×10^5 cells/6 cm plate.

DNA was transfected into all cells by Calcium Phosphate co-precipitation (11) as described before (12). Transient transfections were carried out with 4μg of CAT reporter construct and 0.5 μg of pRSVLacZ plasmid. The pRSVLacZ plasmid which contains the coding region of the β-galactosidase gene under the control of Rous sarcoma virus long terminal repeat sequences was used as an internal control for normalization of transfection efficiency. Cells were harvested after 48 h (myoblasts) or 72 h (myotubes). Non chromatographic CAT assays were performed according to Seed and Sheen (13). For each cell type, the volume of the extract and the incubation time were chosen in a range where there was a linear response between conversion of chloramphenicol and time.

Stable transfectants were obtained as reported previously (14) by co-transfection with 10μg of a βTM-CAT construct and 1 μg of pBS/LTR-neo and selection in the presence of Geneticin (BRL).

Plasmid constructions

All CAT constructs were made from the promoterless plasmid pBLCAT3 or from pBLCAT2 (15) supplied by Dr G.Schütz (Heidelberg, Germany). pBLCAT3 contains the coding region of bacterial chloramphenicol acetyltransferase (CAT) gene, and the small t intron and polyadenylation signal from SV40. In the pBLCAT2 plasmid, the CAT gene is under the control of the Thymidine kinase (TK) promoter of the herpes simplex virus. All constructs are numbered relative to the CAP site set at +1. Contructs made from the region located 5′ and including the muscle CAP site are referred to as βTMsk constructs while constructs assigned βTMnm refer to the non muscle CAP site. Inserted fragments and mutations were confirmed by dideoxy chain termination sequencing of the double-strand plasmids using Sequenase (USB). Oligonucleotides were synthesized on a DNA assembler (Pharmacia).

Non muscle promoter. The NcoI sites flanking the 3.6 Kb genomic fragment encompassing the region between the two ATG codons that initiate translation of the muscle and non-muscle βTM RNAs were filled-in with the Klenow enzyme and cloned at the Sma site of Blue Scribe+ (Stratagene). Directed mutagenesis (16) was used to mutate the downstream ATG codon used in non muscle tissue into an artificil XhoI site. This fragment was cloned in front of the CAT gene (HindIII−XhoI sites) and gave construct pβTMnm-3163. The mutated plasmid was also used to generate construct pβTMnm-2222, by cutting with SplI, filling-in with the Klenow enzyme, then cutting with XhoI and cloning the insert in front of the CAT gene. A genomic 1.38 Kb fragment was produced by cutting with EcoRI, filling-in with Klenow, then cutting with XhoI and inserting the fragment between the XbaI blunt-ended and XhoI sites of the polylinker of pBLCAT3 (construct pβTMnm-1168). Constructs pβTMnm-26 and pβTMnm-168 were obtained from pβTMnm-1168 by cutting at SphI or PstI sites respectively, dilution and ligation. Construct pβTMnm-1783 was obtained by inserting the 2 Kb BamHI−BamHI fragment (+ orientation) in the BamHI site of the polylinker of pBLCAT3. CAT construct pβTMnm-919 was obtained by creating a HindIII site by site-directed mutagenesis (16) at position βTMnm-919 and cloning the HindIII−XhoII fragment in front of the CAT gene.

Genomic fragments were cloned in front of the TK promoter in pBLCAT2 (both orientations). Oligonucleotides were synthesized covering the sequence βTMnm-1168 to 1120nt and cloned at the HindIII−SphI sites of the polylinker. A HindIII fragment βTMnm-1120 nt to −919 nt delineated by artificial HindIII sites was cloned at the HindIII site.

Muscle promoter. Constructs pβTMsk-805, -201, and -68 were obtained as reported before (17). *Cis*-acting sequences of construct pβTMsk-805 were mutagenized (16) and transfered into pBLCAT3 at the PstI−XhoI sites. When required, oligonucleotides delineated by HindIII sites, were synthesized and cloned in both orientations in various constructs.

All plasmids were grown in the XL1Blue strain of *E.Coli*, harvested (7), and purified through Quiagen columns according to the indications of the manufacturer.

cDNA synthesis and polymerase chain reaction (RT-PCR)

Total RNA was isolated (18) from cell cultures and analyzed by the RT-PCR technique. cDNA synthesis was initiated with 10 picomoles of an antisense oligonucleotide complementary to exon 3 of the βTM gene. cDNAs were amplified between this same oligonucleotide and oligonucleotides located in exon 1a or 1b. A first denaturation cycle was followed by 15−20 cycles of amplification (denaturation: 30s at 94°C; annealing: 30s at 56°C; extension: 60s at 72°C). PCR products were analyzed by electrophoresis on 6% (w/v) polyacrylamide gels.

Nuclear extracts and electrophoretic mobility shift assays (EMSA)

Cells were harvested and washed in cold phosphate-buffered saline and nuclear extracts were prepared in the cold room on ice according to Dignam *et al.* (19). Following dialysis, extracts were spun at 4000 rpm, supernatant was aliquoted and stored immediately in liquid nitrogen. Proteins were quantified with Bio-Rad protein reagent. Lysis and extraction buffers contained 1mM PMSF, 5 μg/ml Aprotinin, Pepstatin and Leupeptin (Sigma).

Oligonucleotide probes were generated by labeling one DNA strand with T4 polynucleotidyl kinase. Unincorporated γ-^{32}P ATP was removed by passing the reaction through a chroma-spin10 column (Clontech). The single-stranded, labeled DNA was mixed with a 10-fold molar excess of the unlabeled complementary oligonucleotide, heated to 90°C for 5 min, and allowed to anneal at room temperature. Standard reactions were performed in 20 μml by mixing the labeled probe (20000 cpm; 0.5 ng) in 50mM Tris−HCl (pH 7.8), 10% Ficoll, 250mM NaCl, 5mM EDTA, 5mM DTT, with 12 μg of nuclear extract and 2μg poly(dIdC): poly(dIdC) (Pharmacia Inc) as a non-specific competitor for 15 minutes at room temperature. When appropriate, unlabeled double-stranded competitor oligomer (10 to 50 fold-excess) or antibodies α-SRF1, α-SRF2 and αSRF-DB were added 10 minutes before the labeled probe. DNA−protein complexes were resolved on 4% polyacrylamide

gels in 0.5× TBE (50mM Tris borate pH 8.3, 1mM EDTA). Gels were dried and autoradiographed at −70°.

The sequences (upper strand) of the oligonucleotides are as follows:

E box βTM: 5′ CCTTCGGTGCCAGGTGCCGGGGCTGC
MutE box βTM: CCTTCGGTGCTTGGAACCGGGGCTGC
E box MCK (EMCK): 5′ CCCCCCCAACACCTGCTGCCTGAG
CArG box βTM: 5′ AGTCTGTCCCTAAAAGGTGCTGGCG
CArG box α skeletal actin (αsk act) 5′ CCGACACCCAAATATGGCGCGGCC
C box βTM: 5′ CGGGTGCCGACCCCCCCGTCCCTCTCG

RESULTS

A/ Tissue-specific expression of the chicken βTM gene in cultures of avian cells

We have previously reported that transcription of the βTM gene *in vivo* is initiated at two different sites in a tissue-specific manner (1, 2). In non muscle tissues, transcription starts at a proximal CAP site located inside the gene (nm in Figure 1), whereas in striated or in smooth muscle tissue, transcription starts at a distal CAP site (sk in Figure 1). Exon 1b or exons 1a-2 are spliced to exon 3 in all mature transcripts, and the regulation of the tissue-specific alternative splicing of the primary transcript which occurs further downstream is independent of the transcription initiation site (14).

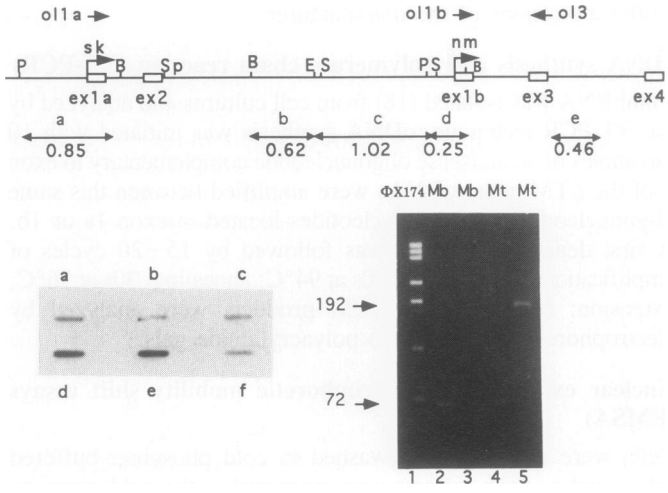

Figure 1. Transcription of the βTM gene in cultures of non myogenic and myogenic cells. The upper part of the figure shows a schematic representation of the genomic region of the βTM gene including the two CAP sites of the gene (broken arrows) used respectively in muscle (sk) and in non muscle (nm) tissues. The restriction sites are indicated as follows: P, PstI; B, BamHI; E, EcoRI; S, SphI; Sp, SplI. Exons are indicated by open boxes. The segments delineated by horizontal arrows refer to the sequences cloned in recombinant plasmids (a, b, c, d, e) which have been used in run-on assay. Ol 1a, ol 1b, and ol 3 refer to the oligonucleotides used in RT-PCR asays. The lower left side panel of the figure shows the result of a run-on assay performed with nuclei prepared from fibroblast cultures (prepared as described in Methods). Labeled nuclear RNAs were hybridized to recombinant plasmids a, b, c, d, and e previously blotted onto nitrocellulose filter (4). BlueScribe (f) was used as a control. Filters were washed and exposed to Kodak X-Ray films. The lower right side panel shows the use of the two CAP sites of the gene in myoblasts (Mb) and in myotubes (Mt), investigated by RT-PCR analysis between exons 1a or 1b and exon 3. RTs were initiated from the antisense oligonucleotide (located in exon 3, see the upper panel) and amplified by PCR between the same oligonucleotide and two sense oligonucleotides located either in exon 1b or exon 1a (respectively ol 1b and ol 1a, see the upper panel). The products of amplification were analyzed by electrophoresis on a 6% acrylamide gel and stained by ethidium bromide. A ΦX HaeIII digest was used as a marker (lane 1). lane 2 and 4: amplification between olignucleotides 1b and 3 lane 3 and 5: amplification between olignucleotides 1a and 3.

To confirm that in avian non muscle cells, such as fibroblasts, transcription is initiated only from the nmCAP site, we performed transcriptional run-on assays. Nuclear RNAs elongated *in vitro* in the presence of α^{32}P-UTP were hybridized to a series of recombinant plasmid DNAs encompassing the two CAP sites (see in Figure 1, upper panel: probes a−e). Comparison of the signals obtained by probes a−e with the signal of the control probe f (Figure 1, lower panel left) showed that labeled nuclear RNAs of fibroblasts hybridized only to probe d (encompassing the nmCAP site and 150 nt of the exon 1b) and to probe e located in the intron between exons 3 and 4 (see Figure 1, upper panel). This indicates that, in non muscle cells, transcription is initiated only at the nmCAP site.

To investigate whether the specificity of transcription initiation is identical *in vivo* and in myogenic cell cultures, total RNA was isolated from undifferentiated and differentiated cultures of myogenic cells and analyzed by RT-PCR (Figure 1, lower right panel). In agreement with results previously obtained by S1 nuclease mapping and primer extension analysis (2), the βTMsk transcript accumulates in differentiated cultures only (amplification of the 192 nt fragment). The βTMnm transcript is present both in undifferentiated and in differentiated cultures of myogenic cells (amplification of the 72 nt fragment), while it was not detected in muscle tissue.

B/ Analysis of the 5′-flanking sequence of the non muscle CAP site

To identify the 5′-flanking sequences of the nmCAP site which are essential for a functional promoter, several fragments were

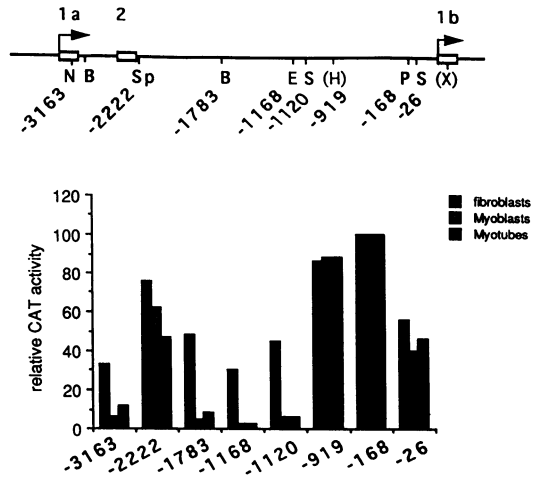

Figure 2. CAT activity of deletions of the βTMnm promoter. Constructs were made as described under Materials and Methods by placing upstream to the CAT gene, different fragments, located between the two ATG codons of exons 1a and 1b. This region of the gene is schematically representated on the upper part of the figure. The natural restriction sites are indicated as in Figure 1 (N indicates NcoI), while the artificial sites introduced by mutagenesis are distinguished by the use of parenthesis. Numbers indicate the 5′ end of the deletion, position +1 being assigned to the nmCAP site. These constructs were transfected into quail fibroblasts, myoblasts and myotubes as described in Materials and Methods. The results were standardized against β galactosidase activity of a co-transfected pRSV LacZ plasmid and are expressed as a percentage of the normalized CAT activity of the plasmid pβTMnm-168) set at 100%. Each value is the average of five experiments.

generated and fused to the CAT reporter gene in a promoterless vector. These fragments have a common 3′ end, at position nm +273 nt (artificial XhoI site), and contain decreasing amounts of 5′ sequence delineated by natural or artificial restriction sites (see the scheme on the top of Figure 2). Constructs were analyzed by transfection into myogenic cells and QT6 fibroblasts of quail. To take into account differences in transfection efficiency, all constructs were cotransfected with pRSVLacZ.

The longest construct pβTMnm-3163 promoted in myogenic cells a CAT activity similar to that of the promoterless vector (referred to as the background level) (Figure 2). As the 5′ sequence flanking the non muscle CAP site was deleted, transcriptional activity increased slightly and construct pβTMnm-2222 promoted CAT activity consistently 4−5 fold that of the background. With further deletions, the CAT activity decreased again to background level, or even under background level (pβTMnm-1783, pβTMnm-1168, pβTMnm-1120). Higher CAT activities were promoted from shorter constructs (pβTMnm-919, pβTMnm-168, and pβTMnm-26) and reached a maximum (6% of RSVCAT activity) with construct pβTMnm-168. In quail fibroblasts, the highest CAT activity was also generated by construct pβTMnm-168 (Figure 2), and a biphasic pattern of CAT activities driven by 5′ deletions −3163 to −168 nt was also observed. Deletion of the CAP site abolished CAT activity (not shown). The levels of CAT activity in myogenic cells correlate with the accumulation of the endogenous non muscle βTM transcript in myoblasts and in myotubes evidenced by RT-PCR (Figure 1).

As the deletions studies delineated sequences having a silencer effect, strong in myogenic cells and more discrete in fibroblasts, we examined the ability of these sequences to modify the transcription from a heterologous promoter. Genomic fragments were cloned in both orientations in front of the TK promoter driving the CAT gene (TKCAT). Of all the fragments tested, only two:βTMnm-1168/-1120 and βTMnm-1120/-919 reduced consistently the CAT activity driven by the TK promoter to respectively 25% and 10% of that of the reference construct TKCAT, thus showing a silencer effect. However this silencer effect was not tissue-specific or differentiation-regulated as those constructs drove the same CAT activities in all transfected cell types (data not shown). This latter result was not in agreement with the results of deletion studies and we propose that combination of several motifs within −1168 and −168 nt is required to ensure full tissue-specific silencer activity.

C/ Multiple motifs participate in the regulation of transcription from the skeletal CAP site

In contrast to the low level of CAT activities driven by the chimeric constructs of the βTMnm promoter, constructs derived from the βTMsk promoter can drive CAT activities in avian myogenic cells which are of the same order of magnitude as those obtained with the MCK enhancer−promoter construct: (+enh 110)80MCK CAT made from the promoter of the muscle creatine kinase gene (20). In a preliminary report, we have shown that 805 nt located 5′ to the skCAP site of the chicken βTM gene (construct pβTMsk-805) drive maximum muscle-specific expression (17); 5′ deletion of this sequence to 201nt (construct pβTMsk-201) causes a loss of only 10% of maximum CAT activity whereas further deletions progressively reduce CAT activity (17). We have also shown that within these 201 nt a muscle-specific element (an E box) and two ubiquitous Sp1 motifs play a role in the control of the CAT activity since mutation of either of these sites causes a decrease of about 40% of the overall activity of construct pβTMsk-805 (17).

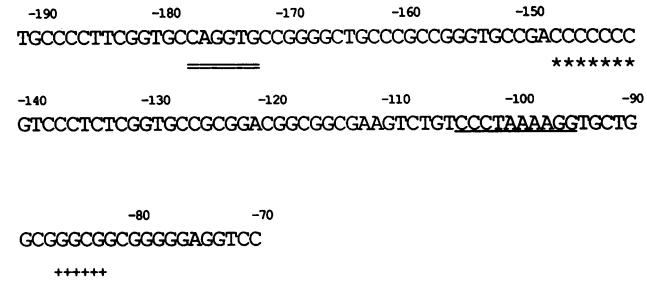

Figure 3. Partial nucleotide sequence upstream to the skeletal CAP site. The nucleotide sequence is numbered as negative from the transcription start site. Doubly underlined sequence marks consensus E box CANNTG motif (20) which matches to the consensus binding sequence of myogenin (30). The CArG-like box is underlined (22, 23), + marks the Sp1 motif, and * marks the C box.

Table 1. Mutational analysis of the *cis*-acting sequences which confer maximum activity to the construct pβTM-805

	Sequence	CAT activity %
E box	GC**CAGGTG**CC	100
Mutant A	GC**AAGCTT**CC	57
C box	GA**CCCCCCC**GT	100
Mutant B	GA**AGCTTC**GT	32
CArG-like motif	GT**CCCTAAAAGG**TG	100
Mutant C	GT**CCGGATCC**GT	46
Mutant D	GT**CGTTTAAAC**GT	44

Mutations were performed on the construct pβTMsk-805. The sequences susceptible to bind transcription factors are shown above the corresponding sequences of the oligonucleotides used for mutagenesis (mutated bases are in bold type). Data are means of four experiments and are expressed as a percentage of the normalized CAT activity of the plasmid pβTMsk-805 set at 100%.

As mutation of the E box did not completely abolish transcription of the CAT gene, we scanned the 201 nt upstream of the CAP site for other putative muscle-specific *cis*-acting DNA elements. Two other motifs located between −201 nt and −68 nt were noted: a variant CArG box and a stretch of 7Cs (Figure 3). The variant CArG box, $CCC(A/T)_5GG$, present at −104 nt, does not strictly match the CArG consensus sequence: $CC(A/T)_6GG$ (21), but an identical motif exists in the promoters of the chicken (22, 23) and of the human (21) α-cardiac actin genes. Contribution of this box to the control of the expression of the βTM gene was investigated using two distinct mutations with or without an A/T rich sequence (Table 1). The difference in the decrease of CAT activity caused by the two types of mutations of the variant CArG box is not significant: Mutant C 46%, and Mutant D 44%. We also investigated the possible role of a stretch of 7Cs located at position −147 by mutagenesis (Table 1 Mutant B). Surprisingly, mutation of the 7Cs motif caused a drop of 70% of the CAT activity while mutation of the variant CArG box (Mutants C and D) or mutation of the E box (Mutant A) caused a decrease of only 40−50% wild type activity (Table 1). No cumulative effect was observed in any constructs where two of the motifs: E box, variant CArG box, and the 7Cs were mutated at the same time (data not shown). These results suggest that all of these DNA motifs are required for maximum transcription from the βTMsk CAP site.

Table 2. Activity of the enhancer fragment in different cell types

	Reference	u+	u−	d2+	d+	d−
pβMnm-168						
Fibroblasts	100	622	302	302	−	203
Myotubes	100	1581	1089	1032	−	403
TK CAT						
Fibroblasts	100	246	151	−	140	131
Myotubes	100	212	151	−	153	162
pβTMsk-68						
Fibroblasts	100	125	110	−	−	−
Myotubes	100	283	99	−	−	−

Constructs were transfected and CAT activities measured as described in Materials and Methods. Activities are normalized to that of an equal amount of the reference promoter and represent average of three experiments. Abbreviations: u+, indicates that the enhancer is inserted upstream to the reference promoter in the sense orientation and u− in the reverse orientation; d+ indicates that it is inserted downstream to the CAT gene of the promoter CAT constructs in the sense orientation (d2+ : double insertion) and d− in the reverse orientation.

To determine if these motifs could act together as an enhancer, an oligonucleotide covering the sequence between −201 and −68 nt was synthesized and cloned in both orientations (+ or − configuration) in construct pβTMnm-168 and in the enhancerless plasmid TKCAT, upstream (u+ or u−) of each promoter or downstream (d+ or d−) of the CAT gene (Table 2). When cloned in the pβTMnm-168 construct, enhanced CAT activity was observed in fibroblasts (2 to 6 fold increase) and myotubes (4 to 15 fold increase). It is worthwhile noting that higher responses were always observed in the u+ configuration suggesting that unlike classical enhancers, this region functions in an orientation-dependent manner. This enhancer was active in myogenic and non myogenic cells, although higher CAT activities were promoted in myotubes versus fibroblasts; this may be related to the presence of muscle specific and ubiquitous (Sp1) *cis*-acting DNA motifs in the enhancer (see Figure 3). When this fragment was cloned in the TKCAT plasmid, the increase of CAT activity was low in all cell types (1.3 to 2.4 fold increase). Higher responses were again observed in the u+ versus u− configuration. We then cloned the same oligonucleotide in front of the construct pβTMsk-68 (homologous promoter): as observed for the heterologous promoters, the u+ configuration drove CAT activities higher than the u− configuration in myotube cultures. Only very weak increase of CAT activity, if any, was observed in fibroblast cultures. Lastly, we cloned the enhancer fragment in a promoterless plasmid to test the ability of this fragment to drive transcription of the CAT gene. Background level CAT activities were driven by u+ and u− configuration (data not shown) indicating that this enhancer is unable by itself to promote CAT transcription.

D/ The E box, the variant CArG box and the stretch of 7Cs bind nuclear proteins

Having identified by mutagenesis three motifs important for enhancer activity, we investigated whether these motifs (E box, CArG box and C box) actually bind nuclear proteins.

Following incubation of the labeled E box βTM* oligomer with myotube nuclear extracts, multiple complexes were observed which are refered to as: Es, Ei1 and Ei2, Ef, according to their electrophoretic mobilities (Figure 4, lane 1). The complex Ei1 is E box specific. It was no longer formed in the presence of E box MCK (lane 3), and it was not disrupted in the presence of the unlabeled competitor mutE box βTM (lane 2) in which the sequence of the E box was mutated. Lastly, this Ei1 complex

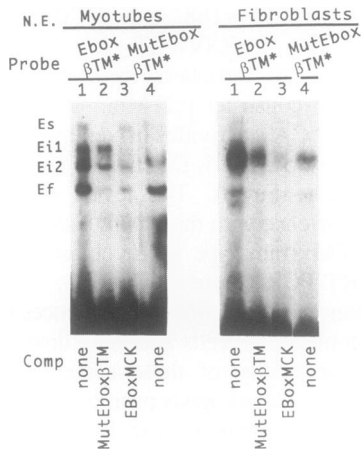

Figure 4. Gel mobility shift assay analysis of the interaction between nuclear extracts from myotubes or fibroblasts and the E box βTM. Following incubation of the 5'end-labeled double-stranded oligomer E box βTM* (E box −177 nt to the skCAP site of the βTM gene) with nuclear extracts (N.E.) from myotubes (left panel) and fibroblasts (right panel), the complexes formed were analyzed on nondenaturing polyacrylamide gels. Specificity of the binding (lane 1) was tested by adding 10-fold molar excess of competitors (comp): mut E box βTM (lane 2) or E box MCK (lane 3), over labeled E box βTM*. Specificity of the binding was also checked using the oligomer labeled mut E box βTM* as the probe (lane 4). Sequences of the oligomers are indicated in Materials and Methods.

was not formed after incubation of nuclear extracts with the labeled oligomer mut E box βTM* (lane 4). The E box specific complex Ei1 is probably formed by binding myogenic + ubiquitous βHLH heterodimers. However, we were not able to identify which myogenic factor is involved as no supershift was observed when the extracts were incubated with CMD1 or chicken myogenin specific antibodies. CMD1 and myogenin fusion protein were then tested for their ability to bind to E box βTM: no complex was formed with CMD1 whereas myogenin bound to E box βTM (data not shown). The pattern of complexes formed following incubation of the E box βTM oligomer with fibroblast nuclear extracts (Figure 4, right panel) was different to that obtained with myotube extracts; the presence of E box binding activities in non muscle cells have also been detected by other groups (24).

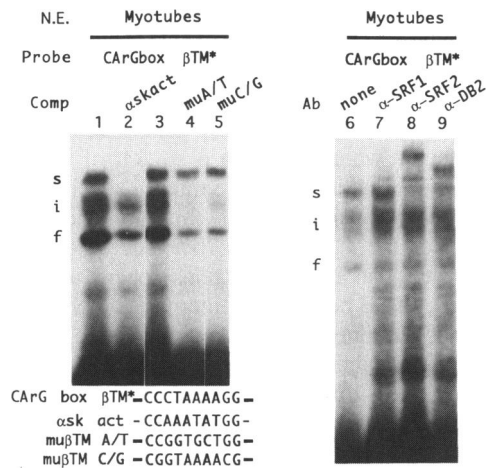

Figure 5. Gel mobility shift analysis of the interaction between nuclear extracts from myotubes and the variant CArG box βTM. Following incubation of the 5'end-labeled double-stranded oligomer CArG box βTM* (variant CArG box −104 nt to the skCAP site of the βTM gene) with nuclear extracts (N.E.) from myotubes, the complexes formed were analyzed on nondenaturing polyacrylamide gels (lanes 1 and 3). Specificity of the binding was tested by adding 10-fold molar excess of competitors (comp): αsk act (CArG box of the chicken skeletal α-actin promoter) (lane2) or muA/T (lane 4) and muC/G (lane 5). Mutated sequences of the CArG box βTM are indicated on the bottom of the panel. The complexes formed were also probed with anti $p67^{SRF}$ antibodies (Ab): α-SRF1 (lane 7), α-SRF2 (lane 8), and α-DB2 (lane9) (27, 42).

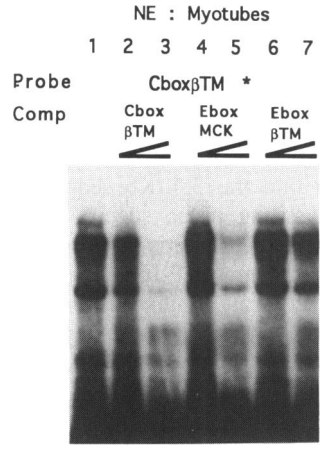

Figure 6. The C box βTM binds nuclear factors. Following incubation of the 5' end-labeled double-stranded oligomer C box βTM* (−147 nt to the skCAP site of the βTM gene) with nuclear extracts from myotubes, the complexes formed were analyzed on nondenaturing polyacrylamide gels. Specificity of the binding was tested by adding 10-fold and 50-fold molar excess of competitors: unlabeled C box βTM (lanes 2−3), EMCK which has a stretch of 7Cs (lanes 4−5), and E box βTM (lanes 6−7) which has no stretch of Cs (sequences of these oligomers are indicated in Materials and Methods).

With the CArG box βTM* oligomer, several complexes (**s**, **i**, and **f**) were formed after incubation with myotube nuclear extracts (Figure 5, lanes 1 and 3). A 10-fold molar excess of unlabeled αsk act (CArG box of the chicken α-skeletal actin gene) abolished the formation of the complex **s** (Figure 5, lane 2). This suggests that the oligomers αsk act and CArG box βTM bind to a common factor, i.e.: serum response factors (SRF) (25) or muscle nuclear factors (MAPFs) (26). Formation of the complex **s** was not prevented by unlabeled competitors in which the CArG box βTM was mutated either in the A/T sequence (mu A/T, lane 4) or at the CC/GG borders (mu C/G, lane 5), indicating that this complex is specific of the variant CArG box. The complex **s** was supershifted following incubation with anti-peptide antibodies directed against the $p67^{SRF}$ protein (α-SRF2) (27) (Figure 5, lane 8) or against a large portion (a.a. 133−264) of the human protein expressed in bacteria (α-DB2) (Figure 5, lane 9). No supershift was induced by a pre-immune serum (not shown) or by another antibody α-SRF1 raised against another region of SRF (lane 7). All these results indicate that the complex **s** is formed by binding of SRF to the variant CArG box of the βTM gene. Competition of the CArG box βTM* by 10-fold molar excess of any competitor muA/T or muC/G inhibited the formation of the intermediate complexes **i** suggesting that the proteins involved in this complex do not bind to the CArG box. Lastly, the protein(s) involved in the complex **f** (which is not disrupted in the presence of the competitors muA/T or muC/G), may also contribute to the control of the βTMsk promoter by binding to the variant CArG box of the βTM gene. These(this) protein(s) remain to be characterized. Band-shift assays with the oligomers muA/T or muC/G displayed similar patterns and this is consistent with our observation that both types of mutation caused the same decrease (55%) of the CAT activity as compared to the wild type activity of the reference construct pβTMsk-805.

Incubation of the oligomer C box βTM* with myotube nuclear extracts indicated that nuclear proteins bind to this sequence (Figure 6, lane 1). Competition assays with 10-fold or 50-fold molar excess of the unlabeled probe, or of EMCK (which includes an E box and also a stretch of 7Cs adjacent to the E box in the MCK promoter) (20), or of the E box βTM (no stretch of 7Cs) showed that the MCK oligomer abolished binding with nearly the same efficiency (lanes 4 and 5) as the unlabeled probe (lanes 2 and 3) whereas the oligomer E box βTM did not compete for the proteins that bind to the C box (lanes 6 and 7) suggesting a specificity in the binding of nuclear proteins.

DISCUSSION

Among the many genes whose transcription is induced during myogenic differentiation, the chicken βTM gene which encodes three isoforms, is of particular interest. While the transcription of the non muscle isoform (βTMnm) is progressively turned off on muscle differentiation, *in vivo* the transcription of the skeletal muscle βTM isoform (βTMsk) is initiated at second specific transcription start site. In the present study, we report an analysis of the *cis*-acting sequences which control the specificity of transcription initiation from these alternative transcription start sites.

We have first shown that in non muscle cells, transcription of the βTM gene is initiated only from the nmCAP site located inside the gene. We have confirmed that transcription is initiated exclusively from the nmCAP site in cultured myoblasts, whereas in differentiated cultures of myogenic cells, transcription is initiated from both the nm and skCAP sites. These results demonstrate that the expression and up regulation of βTMsk in differentiated muscle cells is controlled at the level of RNA transcription. The finding that transcription of the βTMnm isoform continues in cultures of differentiated muscle cells may

be due to the presence of undifferentiated cells in the culture, but also to *bona fide* synthesis of this isoform which participates in the cytoskeleton of cultured myotubes.

In order to delineate the DNA sequences which control the initiation of transcription from the nmCAP site, we have tested the ability of 5′ deletion constructs of pβTMnm-3163 to drive CAT activity. Within the 3.1 Kb sequence upstream of the nmCAP site, no strong tissue specific activating element is detected. The highest CAT activities are promoted by shorter construct such as pβTMnmCAT-168 and only reach about 8 times the background activity. Similar levels of CAT activity are measured in fibroblasts, myoblasts and myotubes, suggesting that the non muscle βTM promoter is an ubiquitous promoter. This correlates the expression of the non muscle endogenous mRNA in cell cultures. The presence of putative non-specific silencer which is suggested by deletion studies, is confirmed by analysis of chimeric βTM-TKCAT constructs. This region contains a 8 nt motif identical to a *cis*-acting DNA motif present in the silencer of the chicken vimentin gene (29).

In constrast to the nmβTM promoter, transcription from the skCAP site is strongly promoted in myotubes and we have previously shown that, of all the constructs tested, pβTMsk-805 (retaining 805nt upstream of the muscle CAP site) drives the highest CAT activity at levels similar to those of (+enh110)80 MCKCAT (20), whereas it promotes only background CAT activity in fibroblasts (17). As 90% of the CAT activity was retained in the pβTMsk-201 construct, we have focussed on the *cis*-acting elements present in this sequence: an E box, a variant CArG box and a stretch of 7Cs. The E box contributes to the maximum activity of construct pβTMsk-805 as shown previously by transfection assays (17). This E box most probably binds myogenic factors of the βHLH type as shown by EMSA, although no supershift was observed with the four tested antibodies directed against either chicken myogenin or CMD1. The binding of purified chicken myogenin fusion protein to the oligomer βTM E box, while purified CMD1 does not bind may be due to the similarity of the sequence surrounding this E box (GTGCCA-GGTGCCGG) to the consensus myogenin binding site reported by Wright *et al.* (30). Only one major E box specific complex is formed with the oligomer E box βTM and myotube nuclear extracts. This indicates that among the several myogenic factors and ubiquitous βHLH proteins (such as E12 and E47) present in cultures of chicken myogenic cells (31), the E box of the βTM gene has high affinity for limited number of heterodimers. Other muscle genes such as muscle creatine kinase (32), myosin light chain 1F/3F (33) (34), and the acetyl choline receptor α-subunit gene (31) require two E boxes for maximal activity. In contrast, a single E box is sufficient for both developmental and tissue-specific expression of the mouse acetyl choline receptor d subunit gene (35). Our results indicate that a single E box is an important element in the assembly of the transcription initiation complex at the skCAP site of the βTM gene. However, for the βTM gene, this single E box is not sufficient to ensure maximum transcription in differentiated myotubes.

An other motif, the variant βTM CArG box: $CCC(A/T)_5GG$, which is identical to a *cis*-acting sequence found in the promoters of the human (21) and chicken α-cardiac actin genes (22, 23) also contributes to the regulation of transcription from the muscle CAP site of the βTM gene. Disruption of this box impairs the transcriptional activity of the pβTMsk-805 reference construct (decrease of 55%). Nuclear factors bind to this motif and one of the complexes (complex s) is sensitive to competition by a molar excess of an oligomer αske act retaining the consensus CArG box of the chicken α-skeletal actin gene (36). This complex is supershifted when specific antibodies directed against $p67^{SRF}$ are preincubated with nuclear extracts and thus, we conclude that the variant CArG box of the βTM gene bind to SRF.

Another factor which binds to a stretch of 7Cs appears to participate in the regulation of the muscle-specific promoter of the βTM gene as mutation of this sequence impairs the transcription of the CAT gene (70% decrease). In addition, EMSA indicates that nuclear proteins specifically bind to this motif. Indeed, the E box βTM oligomer, which has no stretch of 7Cs, is unable to prevent complexes formation while the MCK E box oligomer (retaining a stretch of 7Cs adjacent to the E box) competes complexe formation. A stretch of 9 Cs is present in the promoter of the acetylcholine receptor α subunit gene (37): three of the Cs are part of an adjacent Sp1 site, and the authors propose that G homopolymer binding factor (GBF) which bind to the C string would interfere with Sp1 binding. Since the stretch of 7Cs of the βTM muscle promoter is not associated with an Sp1 site, it is likely that this *cis*-acting element acts by a different mechanism.

It is surprising to note that the double mutations: E box-variant CArG box, E box-7Cs and variant CArG-7Cs only impair transcription of the CAT gene to the same extent as individual mutations. However, similar finding has been reported for the mouse acetylcholinesterase promoter, where mutations of two Egr-1 motifs did not result in an additional decrease of promoter activity (38).

The *cis*-acting elements present in the βTM gene between −201 and −68 act as an enhancer element which is active in differentiated muscle cells, and which functions in an orientation-dependent manner. The position of *cis*-acting elements with respect to the transcription start site might play an important role in transcription as the enhancer is proximal to the CAP site, and thus susceptible to direct interactions with the basal transcription machinery. The absence of one of the *trans*-acting factors could destabilize the transcription initiation complex which assembles through higher order protein−protein interactions in what was defined by F.Jacob as 'aggregulate' (39).

As shown in this work, βTMskCAT constructs promote transcription of the CAT gene in differentiated myogenic cells and this correlates with the presence of the endogenous βTMsk mRNA. This suggests that the sequences cloned in βTMskCAT constructs contain the *cis*-acting elements controling the tissue-specific expression of the muscle promoter. In contrast, the same constructs transfected into undifferentiated myogenic cells drive CAT activity although there is no detectable transcription of the endogenous skβTM mRNA. Lack of differentiation-specific expression of CAT constructs from other muscle (40) or non muscle (41) genes has already been reported. Transcription of the endogenous βTM gene from the muscle CAP site may be repressed by the binding of a negative *trans*-acting factor to a *cis*-acting sequence which is not included in construct pβTMsk-805. Consistent with this, we have observed that βTMskCAT constructs retaining longer 5′ sequence pβTMsk-1800 and pβTMsk-4000) do not promote CAT activity when transfected in undifferentiated myogenic cells (17).

In conclusion, analysis of the elements involved in the control of expression of the βTM gene has shown several important points: dependence upon stereospecific configuration, absence of developmental control of the promoter-enhancer. Work is in progress to investigate these points.

ACKNOWLEDGEMENTS

D.Gros is acknowledged for her technical assistance. The MCK promoter was kindly provided by Dr J.Buskin (Washington, USA), plasmids pBLCAT3 and pBLCAT2 by Dr G.Schultz. The CMD1 or chicken myogenin fusion proteins and the antibodies raised against these proteins were gifts from Dr Hoang Oang Nghiem (Paris). The authors also wish to thank Prof. F.Gros for constant interest. Dr A.M.Prets, L.Balvay, and Dr R.Kelly are acknowledged for fruitful discussions and critical reading of the manuscript. This work was supported by grants from the I.N.S.E.R.M., (contrat INS920807), The C.E.A., the Ministère de l'Industrie et de la Recherche, the Fondation pour la Recherche Médicale, the ARC, the Ligue Française contre le Cancer, and the Association des Myopathes de France.

REFERENCES

1. Libri, D., Lemonnier, M., Meinnel, T. and Fiszman, M. (1989) *J. Biol. Chem.*, **264**, 2935–2944.
2. Libri, D., Mouly, V., Lemonnier, M. and Fiszman, M. (1990) *J. Biol. Chem.*, **265**, 3471–3473.
3. Toutant, M., Toutant, J.-P., Montarras, D. and Fiszman, M. Y. (1983) *Biochimie*, **65**, 637–642.
4. Meinnel, T., Libri, D., Mouly, V., Gros, D., Fiszman, M. Y. and Lemonnier, M. (1989) *Dev. Biol.*, **131**, 430–438.
5. Helfman, D. M., Cheley, S., Kuismanen, E., Finn, L. A. and Yamawaki-Kataoka, Y. (1987) *Mol. Cell Biol.*,, **6**, 3582–3595.
6. Lemonnier, M., Balvay, L., Mouly, V., Libri, D. and Fiszman, M. Y. (1991) *Gene*, **17**, 229–240
7. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: A laboratory manuel. Cold Spring Harbor University Press, **2nd edition**,,
8. Fiszman, M. Y. and Fuchs, P. (1975) *Nature (London)*, **254**, 429–431.
9. Montarras, D. and Fiszman, M. Y. (1983) *J. Biol. Chem.*, **258**, 3883–3888.
10. Moscovici, C., Moscovici, M. G., Jimenez, H., Lai, M. M. C., Hayman, M. J. and Vogt, P. K. (1977) *Cell*, **11**, 95–103.
11. Gorman, C. M., Moffat, L. F. and Howard, B. H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
12. Libri, D., Balvay, L. and Fiszman, M. Y. (1992) *Mol. Cell. Biol.*, **12**, 3204–3215.
13. Seed, B. and Sheen, J. Y. (1988) *Gene*, **67**, 271–277.
14. Libri, D., Marie, J., Brody, E. and Fiszman, M. Y. (1989) *J. Biol. Chem.*, **177**, 6449–6462.
15. Luckow, B. and Schutz, G. (1987) *Nucleic Acids Res.*, **15**, 5490.
16. Kunkel, T. A., Roberts, J. D. and Zakour, R. A. (1987) *Methods in Enzymol.*, **154**, 367–382.
17. Toutant, M., Fiszman, M. and Lemonnier, M. (1993) *C.R.Acad Sci. Paris*, **316**, 711–715.
18. Chomczynski, P. and Sacchi, N. (1982) *Anal. Biochem.*, **162**, 156–159.
19. Dignam, J. D., Lebovitz, R. M. and Roeder, R. G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
20. Buskin, J. N. and Hauschka, S. D. (1989) *Mol. Cell. Biol.*, **9**, 2627–2640.
21. Minty, A. and Kedes, L. (1986) *Mol. Cell. Biol.*, **6**, 2125–2136.
22. Elridge, J., Zehner, Z. and Paterson, B. M. (1985) *Gene*, **36**, 55–63.
23. Chang, K. S., Rothblum, K. N. and Schwartz, R. J. (1986) *Nucleic Acids Res.*, **13**, 1223–1237.
24. Simon, A. M. and Burden, S. J. (1993) *Mol. Cell. Biol.*, **13**, 5133–5140.
25. Treisman, R. (1987) *EMBO J.*, **6**, 2711–2717.
26. Santoro, I. M. and Walsh, K. (1991) *Mol.Cell.Biol.*, **11**, 6296–6305.
27. Gauthier-Rouvière, C., Cavadore, J.-C., Blanchard, J.-M., Lamb, N. J. C. and Fernandez, A. (1991) *Cell regulation*, **2**, 575–588.
28 Daubas, P., Klarsfeld, A., Garner, I., Pinset, C., Cox, R. and Buckingham, M. E. (1988) *Nucleic Acids Res.*, **16**, 1251–1271.
29. Farrell, F. X., Sax, C. M. and Zehner, Z. E. (1990) *Mol. Cell. Biol.*, **10**, 2349–2358.
30. Wright, W. E., Binder, M. and Funk, W. (1991) *Mol. Cell. Biol.*, **11**, 4104–4110.
31. Piette, J., Bessereau, J.-L., Huchet, M. and Changeux, J.-P. (1990) *Nature*, **345**, 353–355.
32. Weintraub, H., Davis, R., Lockshon, D. and Lassar, A. (1990) *Proc. Natl, Acad. Sci. U.S.A.*, **87**, 5623–5627
33. Wentworth, B. M., Doughue, M., Engert, J. C., Berglund, E. B. and Rosenthal, N. (1991) *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 1242–1246.
34. Fujisawa-Sehara, A., Nabeshima, Y., Komiya, T., Uetsuki, T., Asakura, A. and Nabeshima, Y. (1992) *J. Biol. Chem.* **267**, 10031–10038.
35. Wang, X.-M., Tsay, H.-J. and Schmidt, J. (1990) *EMBO J.*, **9**, 783–790.
36. Fornwald, J. A., Kuncio, G., Peng, I. and Ordahl, C. P. (1982) *Nucleic Acids Res.*, **10**, 3861–3876.
37. Bessereau, J.-L., Mendelzon, D., LePoupon, C., Fiszman, M., Changeux, J.-P. and Piette, J. (1993) *EMBO J.*, **12**, 443–449.
38. Li, Y., Camp, S., Rachinsky, T. L., Bongiorno, C. and Taylor, P. (1993) *J. Biol. Chem.*, **268**, 3563–3572.
39. Jacob, F. (1993) *C. R. Acad. Sci. Paris*, **316**, 331–333.
40. Minty, A., Blau, H. and Kedes, L. (1986) *Mol. Cell. Biol.*, **6**, 2137–2148.
41. Tamura, K., Tanimoto, K., Murakami, K. and Fukamisu, A. (1992) *Nucleic Acids Res.*, **20**, 3617–3623.
42. Gauthier-Rouvière, Caï Q.-Q., Lautredon N., Fernandez A., Blanchard J.M., and Lamb N. (1993) *Exp. Cell Res..*, **209**, 208–215.