The human embryonic myosin alkali light chain gene: use of alternative promoters and 3′ non-coding regions

Max Rotter, Katrin Zimmerman, Annemarie Poustka[1] and Anna Starzinski-Powitz*
Institut für Genetik, Zülpicher Straße 47, D-5000 Köln 1 and [1]Deutsches Krebsforschungszentrum, Im Neuenheimer Feld, D-6900 Heidelberg, FRG

Received January 7, 1991; Revised and Accepted March 5, 1991 EMBL accession no. X55000

ABSTRACT

Recently we have found evidence that the human embryonic myosin alkali light chain (MLC1emb) gene has two functional promoters and that its mRNAs exhibit heterogeneity in their 3′untranslated regions (UTR). To study this more in detail we have isolated and characterized the human MLC1emb gene. We focussed in particular on 2 kilobases of 5′flanking region and the alternative 3′UTRs. RNA primer extension and S1 mapping analyses revealed that the MLC1emb gene can indeed be driven either by a proximal or a distal promoter, both in fetal and adult cardiac tissue. These MLC1emb RNAs can contain either the proximal or distal 3′UTR. In contrast to this, in fetal as well as adult masseter muscle MLC1emb mRNA is predominantly transcribed from the proximal promoter and contains mainly the distal 3′UTR. These results explain the known heterogeneity of MLC1emb mRNAs. Finally, we present evidence that the murine MLC1emb gene also contains a functional distal promoter element which has hitherto been undetected.

INTRODUCTION

In mammals, myosin alkali light chain (MLC) genes and cDNAs have been characterized for $MLC1_F$ and $MLC3_F$ from mouse (1), rat (2) and human (3), for $MLC1_{Sb}$ from mouse (4) and man (5), for human $MLC1_{Sa}$ (6; 7), and for embryonic MLC1 (MLC1emb) from mouse (8) and man (9; 10; 11; 12; this paper). It should be noted that MLC1emb and MLC1A are the same (15; 8). We will use the nomenclature MLC1emb throughout.

From these reports it is apparent that each of the MLC isoforms is encoded by a separate gene except for the fast skeletal muscle isoforms of MLC1 and MLC3 which have alternative starts of transcription producing two primary transcripts which are spliced independently to give rise to $MLC1_F$ and $MLC3_F$ mRNA (13; 1). Furthermore, the $MLC1/3_F$ gene is the only MLC gene so far which has been found to be driven by two promoters.

Recently, we reported that human muscle cells in culture contain different MLC1emb mRNAs with identical protein coding regions, but alternative 5′ and 3′ untranslated regions (UTR) (11). These data suggested that transcription of the human MLC1emb gene may be controlled by two alternative promoters and that splicing produces two alternative 3′UTRs.

In fetal life, MLC1emb is expressed in skeletal muscle (14) as well as in heart (15). Postnatally, MLC1emb has only been found in atria where this isoform is described as the atrial light chain MLC1A (16, 9), in masseter muscle (17) and in regenerating muscle fibres (18). This raises the possibility that the MLC1emb gene may be regulated by different promoters in the different cell context.

In order to investigate this question we characterized the human MLC1emb gene including about 2 kb of 5′flanking region. We show that the MLC1emb gene can be transcribed from two alternative promoter elements and that the alternative 3′UTRs can be both used in muscle tissues expressing MLC1emb. In addition, we show that a functional distal promoter element most likely also exists in the murine MLC1emb gene.

MATERIAL AND METHODS

RNA and MCL1emb cDNA

Cytoplasmic RNA was prepared (19) from secondary human myogenic cultures (11) at day 3 after induction of myotube formation. RNAs from human fetal ventricle and adult atria were a gift from P.Barton. RNA samples from adult and fetal human masseter and quadriceps muscle were provided by N.Soussi-Yanicostas and G.Butler-Browne. Isolation and characterization of MLC1emb cDNA Hemp.4 has been described recently (11). This cDNA clone contains the entire MLC1emb protein-coding region, 82 nucleotides of the distal 5′UTR as well as 3′UTR/7.

Isolation and characterization of genomic MLC1emb cosmid clones

A cosmid library made from human blood cell DNA and cloned into vector pcos2EMBL (20) was screened with MLC1emb cDNA Hemp. 4. Eight cosmid clones were isolated, DNA was prepared and tested by restriction analysis (21). One of these

* To whom correspondence should be addressed

cosmid clones, pcemb8, which showed the expected restriction pattern, was chosen for further analysis. pcemb8 DNA was prepared by the alkaline lysis method (21) and purified by centrifugation in CsCl. DNA prepared by this method was directly used for sequencing reactions with the dideoxy chain termination method (22) performed on double-stranded DNA. Sequencing was performed by walking along the gene with synthetic oligonucleotide primers derived from the sequence reactions. All nucleotide sequences were determined from both DNA strands.

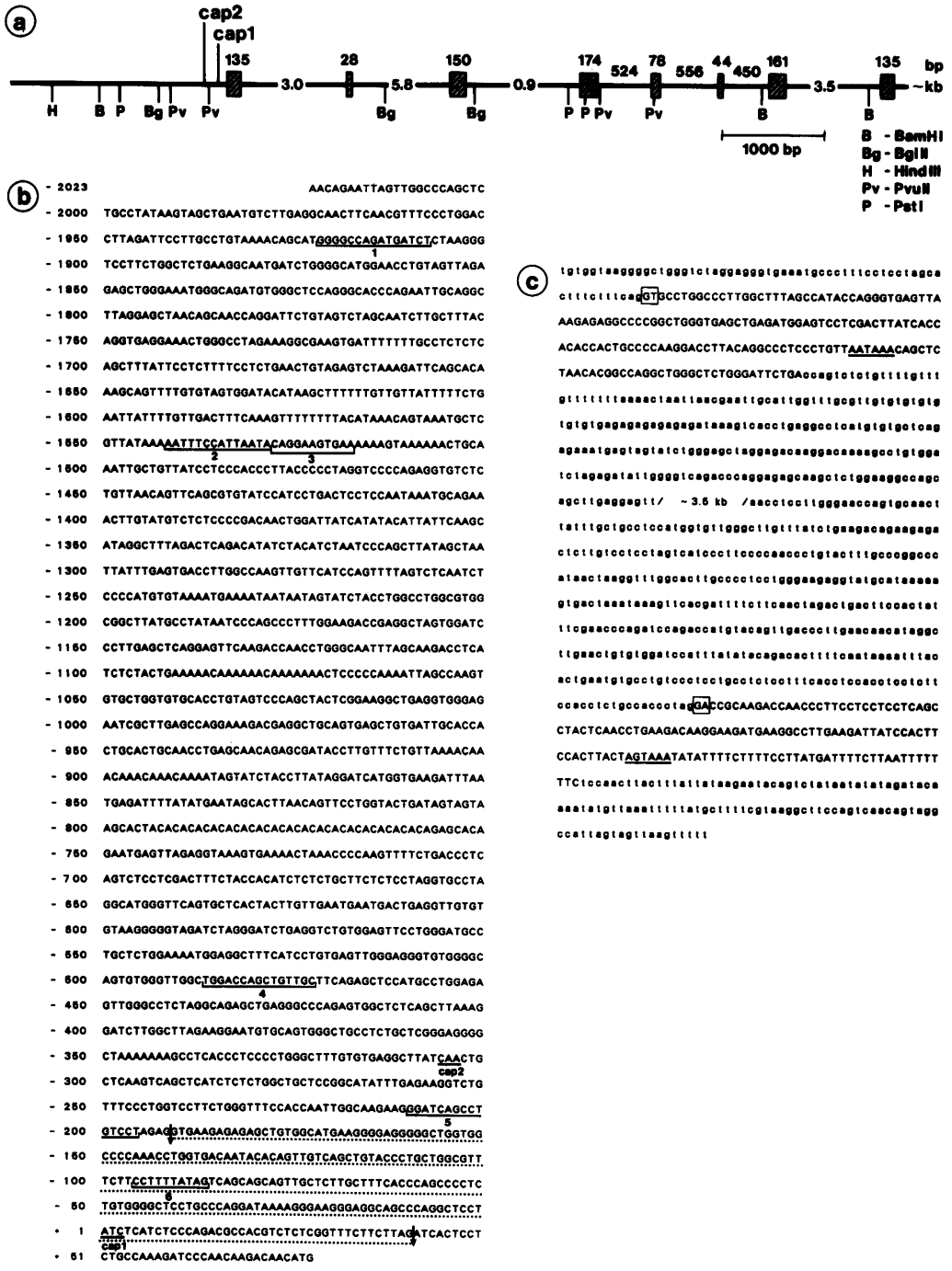

Figure 1. a) Scheme of exon/intron distribution in the MLC1emb gene as deduced from sequence analysis of cosmid pcemb8. Hatched boxes correspond to exons. The first six exons are protein-coding exons whereas Exon 7 and Exon 8 correspond to the 3'UTRs (3'UTR/7 and 3'UTR/8). Letters indicate restriction endonucleases as described. **b)** Nucleotide sequence of the 5'flanking region of the MLC1emb gene ending at the putative start of translation. Putative regulatory sequence motifs are underlined and numbered as listed in table 1. Cap sites of the proximal (cap1) and distal (cap2) promoters are marked. The 5' and 3' splice sites of the facultative 5'UTR intron (in case the distal promoter is active) are indicated by arrows and the intron itself is underlined with a dotted line. **c)** Nucleotide sequence of the 3'non-coding region of the MLC1emb gene. Capital letters represent sequences of exons 3'UTR/7 and 3'UTR/8, respectively whereas intron sequences are shown in small letters. The size of the intron (~4,1 kb in total) was estimated by restriction analysis (data not shown). Splice signals of the facultative 3'intron are boxed and putative polyadenlyation signals are underlined.

Table 1. Putative regulatory sequence motifs identified in the 5'flanking region of the human MLC1emb gene.

Number in the sequence (Fig.1b)	Sequence in MLC1emb gene	consensus sequence	Ref. for consensus sequence
1	Glucocorticoid responsive element		Jantzen et al., 1987 (28)
	GGGGCCAGATGATCT	GGTACANNNTGTTCT	
2	mouse $MLC1_F/3_F$ promoter motifs		Daubas et al., 1985 (27)
	AATTTCCATTAATA	AG(A/−)TTCCATT(A/T)ATA	
3	Enhancer consensus (E1a core)		Daubas et al., 1985 (27)
	CAGGAAGTGAA	C(A/G)GGAAGTGA(A/C)	
4	Myo D1-binding element	C GG CC CC C N CA TG N	Buskin et al., 1989 (24)
	GGACCAGCTGTTGC	G AA GG TT G	
5	Glucocorticoid responsive element		Jantzen et al., 1987 (28)
	GGATCAGCCTGTCCT	GGTACANNNTGTTCT	
6	MLC-specific sequence		Cohen et al., 1988 (26)
	CCTTTTATAG	CCTTTTATAG	

RNA primer extension and S1 mapping analysis

RNA primer extension analysis was carried out according to published procedures (21). Probes were prepared as described (see legend of Fig. 2) in the presence of ^{32}P-labelled αdATP and recut with the appropriate restriction endonuclease. Probes were separated on polyacrylamide gels and eluted. Hybridization of the probe to RNA was done in 80% formamide at 40°C overnight. Extension was done with Superscript reverse transcriptase (Gibco, BRL) at 42°C for 45'. Extension products were separated on 6% denaturing polyacrylamide gels. Size markers, including those for the S1 mapping analysis (see below) were sequence ladders performed on a genomic MLC1emb BglII/PstI DNA fragment (containing part of the first intron, the first exon and ~600 bp of 5' flanking DNA) which was cloned into plasmid vector pUC19. For S1 mapping analysis (11) ^{32}P-labelled probes were prepared as described in the legend to Fig.2, recut with restriction enzyme, eluted from the polyacrylamide gel and hybridized to RNA as indicated in 75% formamide at 50°C overnight. S1 nuclease-digested samples, as well as the sequence ladders, were separated on 6% denaturing polyacrylamide gels and exposed to X-ray film.

Polymerase chain reaction

For amplification of mRNA sequences, first strand cDNA was primed on 3 μg of total RNA using specific oligonucleotide primers and 200 U Superscript reverse transcriptase (Gibco BRL, Bethesda). Subsequently, polymerase chain reactions (PCR) were carried out according to published procedures (23; 11). Amplification cycles with 2.5 U Taq-Polymerase (Gibco, BRL) were: 1' annealing at 55°C; 3' extension at 70°C; 1' denaturing at 95°C. Aliquots of the PCR products were separated in 0,8% agarose gels, blotted onto nylon membranes and hybridized to ^{32}P-labelled cDNA Hemp.4 (21). Blots were washed at a final stringency of 0,1×SSC, 0.1% SDS at 65°C for 60'.

RESULTS AND DISCUSSION

Structure of the human MLC1emb gene including 2 kb of 5'flanking sequence

Recently, we characterized a cDNA (Hemp.4) which encodes human MLC1emb (11). Using this cDNA as probe, we isolated the MLC1emb gene from a human genomic cosmid library (20). Characterization of the organization of the MLC1emb gene was mainly done by direct sequencing of cosmid pcemb8 which contained the MLC1emb gene without subcloning the insert. From the comparison of human and mouse MLC1emb cDNA sequences it is apparent that the basic organization of this gene is conserved. Therefore, oligonucleotides, derived from positions in the human MLC1emb cDNA which correspond to exon/intron boundaries in the murine MLC1emb gene, were used to sequence those in the human gene. From the nucleotide sequences obtained, further oligonucleotides were derived in order to sequence complementary strands and to complete sequences of the shorter introns. In this way the organization of the MLC1emb gene shown in Fig. 1a was determined.

As in the murine MLC1emb gene (8), the protein-coding information of human MLC1emb is encoded in six exons (Fig. 1a; exons 1−6, (12)). The alternative 3'untranslated regions (3'UTR) identified in the MLC1emb cDNA clones Hemp.4 and Hemp.3 (11) were found in exons 7 (3'UTR/7 from cDNA Hemp.4) and 8 (3'UTR/8 from cDNA Hemp.3) separated from each other by about 4.1 kb of DNA as analysed by Southern blot analysis of the related cosmid DNA (data not shown). The sequences of exons 7 and 8 and their putative exon/intron boundaries are shown in Fig. 1c.

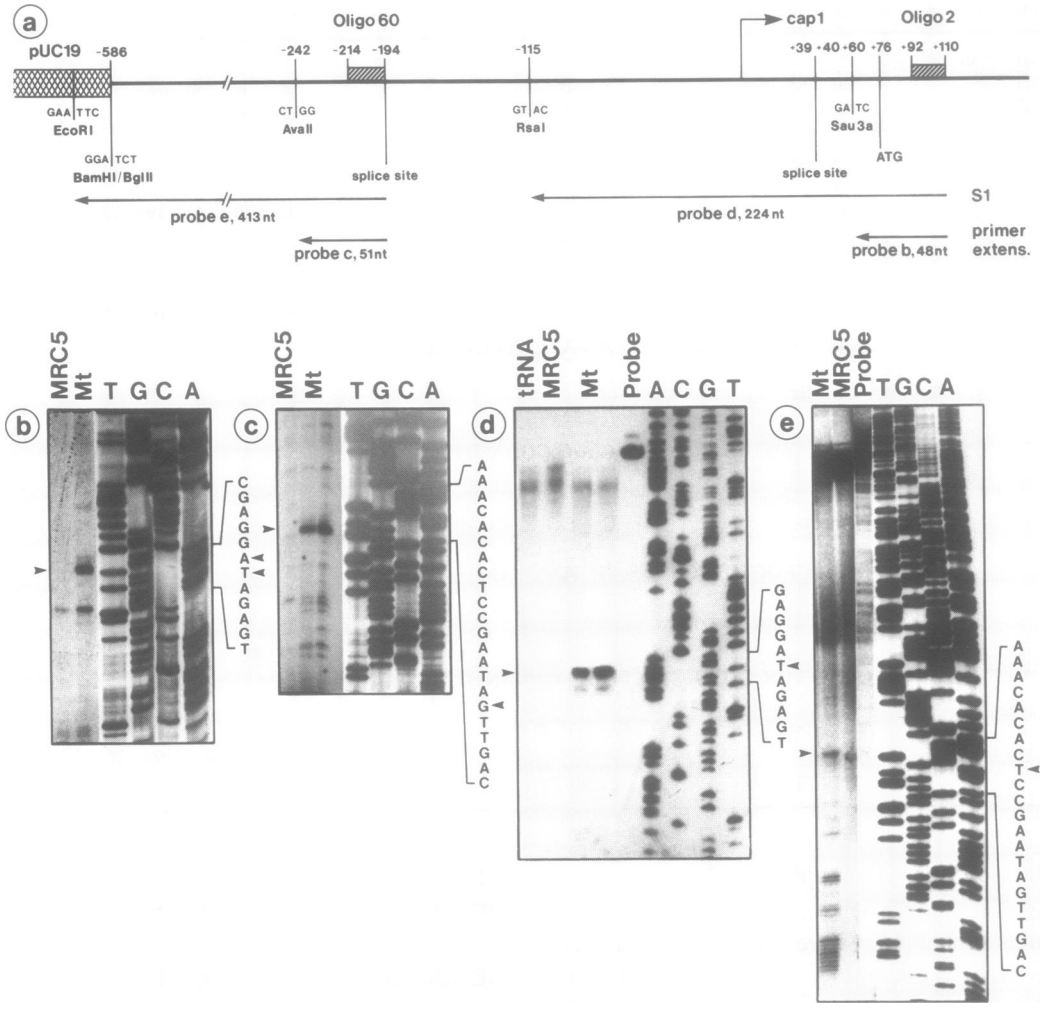

Figure 2. Primer extension and S1 mapping analyses of transcription start sites of the human MLC1emb gene. **a)** Schematic experimental design of the results shown in b–e. **b)** Primer extension analysis of the proximal cap site (cap 1; Fig. 1b) using probe b. Probe b was generated using oligo 2 as primer which was extended in the presence of MLC1emb 5′flanking DNA and ^{32}P-labelled nucleotides. The extension product was cut with restriction endonuclease Sau3A giving rise to probe b. 1×10^4 cpm of probe b were hybridized to 35 μg of total cytoplasmic RNA from either human myotubes (lane Mt) or MRC5 fibroblasts (lane MRC5) and extended. Extension products as well as the sequence ladder derived with oligo 2 were separated on a 6% polyacrylamide gel and exposed to Kodak XAR5 film at −70°C overnight. **c)** Primer extension analysis of the distal cap site (cap2; Fig. 1b). This was done essentially as described in b). Except that 100μg of RNA was used instead of 35 μg for hybridization to probe c. Priming of probe c was done with oligo 60 and the extension product was cut with Ava II. Oligo 60 was also used to derive the sequence ladder. Separated primer extension products were exposed to Kodak XAR 5 film at −70°C for 2 days whereas the sequence ladder is an exposure to KODAK X-omat S overnight. **d)** S1 mapping analysis of the proximal start site (cap 1) of transcription using P-labelled probe d which was primed with oligo 2 in the first protein-coding exon and recut with restriction endonuclease Rsa I. 5×10^4cpm were each hybridized with 30 μg of total cytoplasmic RNA from either human myotubes (lanes mt) or MRC 5 (lane MRC5) fibroblasts as well as to yeast tRNA (lane tRNA) and subsequently digested with S1 nuclease. The digestion products as well as the intact probe and the sequence ladder already used in b) were separated in 6% denaturing polyacrylamide gels which were exposed to Kodak X-OMAT S film at −70°C overnight. **e)** S1 mapping analysis of the distal start site (cap 2) of transcription. This was essentially performed as described in d) except that probe e was used. Probe e was primed with oligo 60 and recut with restriction endonuclease EcoRI in the linker region of vector pUC19. The sequence ladder is the same as that shown in c). Exposure of the separated digestion products was on Kodak XAR 5 at −70°C for 2 days. Arrowheads at the left side of b–e indicate MLC1emb specific reaction products whereas arrowheads at the right side of b–e indicate the first nucleotides identified in MLC1emb mRNA.

In order to determine the structure of the 5′region of the MLC1emb gene, over 2 kb of sequence upstream of the start codon was sequenced (Fig. 1b). The sequence contains elements which have been described in other genes as targets for transacting factors or as typical 'MLC' sequences (26) which have been conserved between different MLC genes during evolution. Some of these sequences are listed in table 1 and underlined in Fig. 1b.

Of the sequences identified one is homologous to the MEF1-binding enhancer element (No. 4 in table 1 and Fig. 1b) originally identified in the muscle creatine kinase gene (24) and found to be a binding site for the myogenic determination factor MyoD1 (25). It might well be that this site drives MyoD1-dependent transcription of MLC1emb in skeletal muscle in situ.

A so-called 'MLC sequence' motif CCTTTTATAG (No. 6 in table 1 and Fig. 1b) has, beside others, been found to be common for the fast (f), ventricular (v) and atrial (A; embryonic) MLC1 genes of mouse with similar locations in all three genes (around −100; 26). In the human MLC1emb gene this 'MLC sequence' is positioned at −105 which is comparable to the other MLC1 genes. Although the conservation of this sequence suggests a specific role in transcription of MLC1 genes, this role is still unresolved.

Also unknown is the functional role of another sequence motif

(AATTTCCATTAATA; No. 2 in table 1 and Fig. 1b) which has been identified in the promoter region of the murine MLC1/MLC3 gene (27). Directly adjacent to this sequence we found a putative enhancer element (No. 3 in table 1 and Fig. 1b) with 100% homology to the enhancer core sequence of the major enhancer of polyoma virus and of the adenovirus E1a gene which similarly, is also present in the murine MLC1/MLC3 promoter region (27). Again, the regulatory role of this sequence in the MLC1 genes has to be evaluated.

Finally, two sequence motifs with homology to glucocorticoid responsive elements (GRE) were identified (No. 1 and 5 in table 1 and Fig. 1b). These putative GREs show 73% and 80% homology to the GRE consensus sequence (28) and are even more similar to the GREs of the murine sarcoma virus MSV (87%) and of the human methallothionin gene (86%). For comparison of the GREs the three central nucleotides were always regarded as 'N'.

Glucocorticoids have two main functions the first of which is their involvement in energy metabolism as antagonists of insulin. Secondly, glucocorticoids can control differentiation processes such as that of chicken oviduct (29). However to date, none of the muscle-specific genes described has been found to be regulated by glucocorticoid. Therefore, it remains to be shown that the putative GREs in the human MLC1emb gene are indeed functional.

Initiation sites of transcription

Characterization of MLC1emb cDNAs from human muscle cells in culture by sequence and S1 mapping analysis previously revealed the existence of two 5'untranslated regions (5'UTR) (11); One 5'UTR (designated 5'UTR/d in this paper) has been found in MLC1emb cDNA Hemp.4 and, as verified by PCR analysis, in MLC1A mRNA of human atria (11). The second 5'UTR sequence (designated 5'UTR/p in this paper) has been published in MLC1emb cDNA of human skeletal muscle (10) and also in MLC1A cDNA of human atria (9). In the cDNAs, both 5'UTRs share 36 nucleotides directly adjacent to the ATG (11) and were also found in the 5'flanking sequence of the MLC1emb gene (Fig. 1b in this paper). From these results it seemed very likely that the human MLC1emb gene has two promoter elements, a proximal and a distal one which should be separated by a facultative intron.

In order to test this assumption and to determine the two putative alternative start sites of transcription, primer extension analyses and S1 mapping experiments were carried out. First, initiation of transcription was analyzed by primer extension. For analysis of the proximal cap site a primer covering nucleotides (nt) +92 to +110 of exon 1 was used for extension to generate probe b of 48 nt (Fig. 2a) as well as to obtain the sequence ladder. In the presence of human myotube RNA, the primer was extended for 62 nt (Fig. 2b, lane Mt) whereas in the presence of MRC5 RNA (negative control; Fig. 2b, lane MRC5) no MLC1emb-specific product was obtained. Thus, transcription of one of the MLC1emb mRNAs initiates at TATCTCATC. This finding is supported by the observation that a putative TATA box is located 24 nt upstream of this cap site (Fig. 1b). The cap site (cap1 in Fig. 1b and Fig. 4a) identified here corresponds to the one described in the murine MLC1emb gene (8).

For analysis of the distal promoter element probe c (Fig. 2a) ranging from nt −194 to −242 of the later on determined proximal start of transcription was prepared for extension. The extension product obtained in the presence of two different human

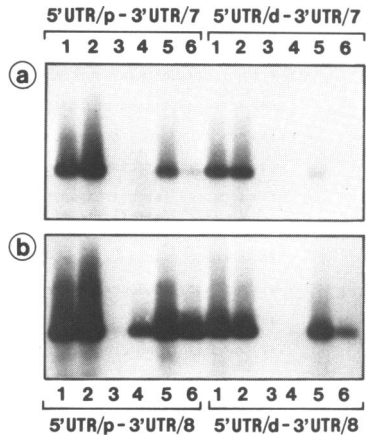

Figure 3. Analysis of MLC1emb mRNA species in human striated muscle by PCR: fetal ventricle (lane 1); adult atria (lane 2); adult quadriceps (lanes 3); adult masseter (lanes 4), fetal quadriceps (lanes 5) and fetal masseter (lanes 6). For PCR, first strand cDNA was primed either with oligonucleotide 426 (3a; specific for the proximal 3'UTR = 3'UTR/7) or with oligonucleotide 425 (3b; specific for the distal 3'UTR = 3'UTR/8). First strand cDNAs were each divided into two aliquots and amplified either with oligonucleotide 5'GGTTTCCACC-AATTGGCAAG (specific for the distal 5'UTR (5'UTR/d); right lanes 1−6 in 3a and 3b) or oligonucleotide 5'ACGTCTCTCGGTTTCTTCT (specific for the proximal 5'UTR (5'UTR/p); lanes 1−6 on the left side in 3a and 3b). The four possible combinations are indicated: 5'UTR/p-3'UTR/7; 5'UTR/d-3'UTR/7; 5'UTR/p-3'UTR/8; 5'UTR/d-3'UTR/8. Aliquots of specific amplification products (usually 1/20) were separated in agarose gels, blotted onto nylon membranes and hybridized to ^{32}P-labelled MLC1emb cDNA. Exposure was for 15−60 minutes on Kodak X-OMAT S film. All PCR reactions were internally controlled by amplification of actin mRNA sequences (data not shown).

myotube RNA samples was each a fragment of 113 nt (Fig. 2c, lanes Mt). At this point the sequence ladder shows the sequence TTATCAACT (around nt −310). Thus, a second MLC1emb RNA can be initiated at this sequence (cap 2 in Fig. 1b and Fig. 4a). Although no characteristic TATA box precedes this distal cap site there is an A-rich sequence stretch, which might serve as a TATA box, about thirty nucleotides upstream of this initiation site of transcription.

In order to confirm the primer extension analyses, S1 mapping experiments were performed. For these, single-stranded antisense probes (probe d; proximal region and probe e; distal region; Fig. 2a) were synthesized as described in legend to Fig. 2. Probe d was 224 nt in length including cap site 1 (Fig. 2a). Probe e contained 413 nt of 5'region und some pUC19 sequence starting with the first nucleotide of the distal 5'UTR (nt −194) which has been identified in MLC1emb cDNA Hemp.4 (11). Each of both probes was hybridized to total cytoplasmic RNA from developing human myotubes and from human MRC5 fibroblasts (negative control), and subsequently digested with S1 nuclease.

As shown in Fig. 2d, lanes Mt, a fragment of 111 bp of probe d was protected from S1 digestion by myotube RNA but not by MRC5 RNA (lane MRC5). When the size of this protected fragment was compared to the sequence performed with the same primer, cap site 1 already identified by primer extension (Fig. 2b) was confirmed.

With probe e, the longest fragment protected from S1 digestion by myotube RNA (Fig. 2e, lane Mt, arrow head) but not by MRC5 RNA (Fig. 2e, lane MRC5 was 113 bp in length. The shorter fragments seen in Fig. 2e occurred regularly in these S1 digestions. They are probably due to particular properties of the sequence tested and do not reflect heterogeneic initiation sites

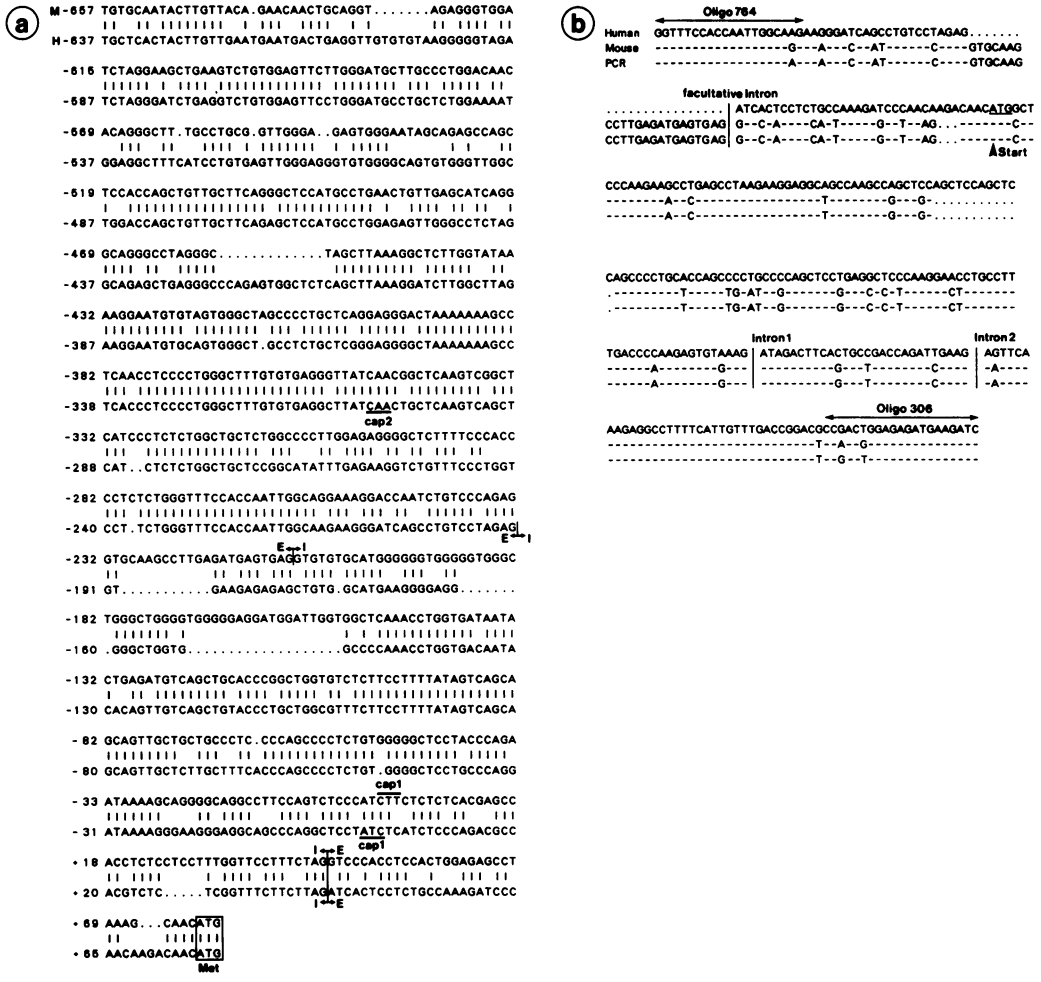

Figure 4. a) Comparison of 5'nucleotide sequences of the mouse (upper line) and human (lower line) MLC1emb gene. Known cap sites are underlined (human sequence) or overlined (mouse sequence) and exon (E)/intron (I) boundaries of the facultative introns are indicated by arrows both in the mouse and human sequence. The putative start codon is boxed in both sequences. **b)** Nucleotide sequence comparison of human MLC1emb cDNA (upper line) transcribed from the distal promoter with a murine MLC1emb cDNA fragment from heart (PCR, lower line) obtained by PCR using a pair of oligonucleotide primers (Oligo 306/Oligo 764) derived from the human MLC1emb sequence. Exon sequences only are shown. The positions of introns are indicated by the solid lines through all three sequences. First strand cDNA was primed with oligonucleotide 5'GATCTTCATCTCTCCAGTCGG (oligo 306) located in the protein coding region of MLC1emb cDNA (exon 4). For PCR, the second oligonucleotide was 5'GGTTTCCACCAATTGGCAAG (oligo 764) located in the distal 5'UTR. The middle line represents the MLC1emb cDNA sequence deduced from the murine genomic MLC1emb sequence using the standard exon/intron boundary rules.

since the primer extension analysis only showed one start site of transcription in two different RNA samples. Comparison of the longest fragment to the sequence ladder obtained with the primer which was used to synthesize probe e revealed a transcription start site at 5'AGGCTT instead of 5'CAACTG as it was found in the primer extension experiment (Fig. 2c). We interpret these results as indicating that the discrepency of 8 nt between primer extension and S1 mapping analyses is also due to the sequence properties of the distal 5'UTR.

Quantitatively, these results confirm the previously published indication (11) that the distal promoter of the MLC1emb gene is less active than the proximal promoter in human myotubes in culture.

Use of alternative promoters and 3'UTRs in different MLC1emb-positive muscle tissues

Data presented in this paper and previous results (11) indicated that transcription of MLC1emb mRNAs can be driven by two alternative promoters and, in addition, can use two alternative 3'UTRs (3'UTR/7 and 3'UTR/8; Fig. 1). Therefore, theoretically four different MLC1emb mRNAs could be expressed each being transcribed from either the distal or proximal promoter (thus containing different 5'UTRs) and having alternative 3'UTRs but identical protein coding regions.

In order to test which type of 5'/3'UTR combinations are used in MLC1emb-positive muscle tissues, polymerase chain reactions (PCR) were performed with RNA from human muscle tissues such as adult atria, fetal cardiac ventricle, adult and fetal masseter and quadriceps. First strand cDNA was specifically primed with an oligonucleotide derived from either 3'UTR/7 (all samples in Fig. 3a) or 3'UTR/8 (all samples in Fig. 3b). The second oligonucleotide for the PCR was derived either from the distal (5'UTR/d) or proximal (5'UTR/p) 5'non-coding region. After 20 PCR cycles aliquots of the amplification products were either reamplified with an internal 3'UTR-specific ^{32}P-labelled primer (data not shown) or subjected to Southern blot analysis (Fig. 3). PCR reactions were internally controlled by amplification of actin mRNA (data not shown). The results shown in Fig. 3 were

reproducable in several independent experiments. Nevertheless, PCR results are difficult to interpret in quantitative terms, especially if the individual PCR products are not derived from the same first strand cDNA synthesis.

One conclusion of these analyses is that cardiac muscle expresses all four types of MLC1emb RNAs (Fig. 3, lanes 1 and 2). Thus, it appears that the proximal promoter element is stronger than the distal promoter element both in fetal ventricle and adult atria. In contrast to this observation, fetal masseter muscle apparently contains only low levels of the combinations 5'UTR/p-3'UTR/7 (Fig. 3a, lane 6, left) and 5'UTR/d-3'UTR7 (fig. 3a, lane 6, right) but higher levels of combinations 5'UTR/p-3'UTR/8 (Fig. 3b, lane 6, left) and 5'UTR/d-3'UTR/8 (Fig. 3b, lane 6, right). Thus, fetal masseter seems to express preferentially MLC1emb RNA with the distal 3'UTR but from both promoters. Also in fetal quadriceps both promoters can be used as well as both 3'UTRs. Here it seems that 3'UTR/8 is predominatly expressed. Furthermore, in adult masseter muscle, which has been found to express MLC1emb in reasonable amounts (17), MLC1emb mRNA is mainly transcribed from the proximal promoter and contains preferentially the distal 3'UTR as it was also found for fetal masseter. As discussed previously (11) the differential use of 3'UTRs may reflect posttranscriptional regulation of MLC1emb mRNAs.

Adult quadriceps is almost negative for MLC1emb protein (17) a result which was confirmed at the RNA level (Fig. 3). The weak reactions seen with combinations 5'UTR/p-3'UTR/8 most likely came from regenerating fibres which have been found to express MLC1emb (18).

In summary, our observations from the PCR experiments indicate that there is no stringent control for either promoter or 3'UTR. But certain combinations of 5' and 3'UTRs seem to be used predominantly in particular muscle tissues.

The distal MLC1emb promoter appears to be conserved between man and mouse

Previous data (8, 11, 26) have suggested that the murine MLC1emb gene is different from the human MLC1emb gene insofar as it is only transcribed from a single promoter. Nucleotide sequence comparisons of the distal promoter region of the human gene with the corresponding region of the mouse gene, however, revealed a high degree of homology (Fig. 4a) which might reflect conserved regulatory functions. Therefore, we asked whether as an yet unidentified distal promoter might be active in mouse.

To investigate this, PCR experiments were performed with conserved primers from the coding region (Oligo 306; Fig. 4b) and the putative distal 5'UTR (Oligo 764; Fig. 4b) using RNA from mouse atria. To increase the specificity of the PCR, an internal ^{32}P-labelled primer was used in the last 5 amplification cycles. The resulting fragment which had the expected size (data not shown) was cloned into plasmid vector pUC19 and subjected to sequence analysis. The sequence of this PCR product (Fig. 4b) contained part of the previously published (8) 5'flanking sequence of the murine MLC1emb gene whereas the proximal promoter as well as part of the proximal 5'UTR, in analogy to the human situation, were spliced out (Fig. 4b). Thus, we conclude that in mouse as well as in human a second promoter is able to drive MLC1emb expression at least at low frequency. Splicing of this *facultative* 5'UTR intron appears to happen in very similar positions both in the mouse and human MLC1emb gene although the 5'-splice site in the mouse gene must be used at a slightly different position when compared to the human situation (indicated in Fig. 4b).

Taken together, our results provide evidence for the existence of two different promoter elements in the human as well as in the murine MLC1emb gene. Yet, it is not clear at all whether there are discriminating regulatory activities for these promoters in the different muscle tissues which express MLC1emb although the proximal and not the distal promoter appears to be preferentially used in adult masseter muscle.

Despite its different tissue specificity the myosin light chain gene L23 of embryonic chicken (30) is thought to be the equivalent of mammalian MLC1emb (8). Gene L23 is driven by one promoter and is mainly expressed in fetal smooth muscle and in brain throughout development. Speculatively, one might imagine that the L23 gene lost its proximal promoter during evolution and thus could rather easily acquire an expression pattern which is different from its mammalian homologue MLC1emb.

ACKNOWLEDGEMENTS

We thank B.Jack and P.Barton for critical reading of the manuscript, C.Siewert for expert technical assistence and U.Ringeisen for photographs. Human RNA samples were kindly provided by P.Barton (London), N.Soussi-Yanicostas and G.Butler-Browne (Paris). This work was supported by the Bundesministerium für Forschung und Technologie and the Deutsche Forschungsgemeinschaft. Fellowships of the Fritz Thyssen-Stiftung to K.Z. and the Deutsche Forschungsgemeinschaft to A.SP. (Heisenberg fellowship) are gratefully acknowledged.

REFERENCES

1. Robert, B., P. Daubas, M.-A. Akimenko, A. Cohen, I. Garnier, J.-L. Guenet and M.E. Buckingham. 1984. Cell **39**:129–140.
2. Periasamy, M., E.E. Strehler, L. Garfinkel, R. Gubits, N. Ruiz-Opazo and B. Nadal-Ginard. 1984. J. Biol. Chem. **259**:13595–13604.
3. Seidel, U., E. Bober, B. Winter, S. Lenz, P. Lohse and H.-H. Arnold. 1987. Nucl. Acids Res. **15**:4989.
4. Barton, P.J.R., A. Cohen, B. Robert, M.Y. Fiszman, F. Bonhomme, J.L. Guenet, D.P. Leader and M.E. Buckingham. 1985. J. Biol. Chem. **260**:8578–8585.
5. Wade,K.R., D. Feldman, P. Gunning and L. Kedes. 1989. Mol. Cell. Biochem. **87**:119–136.
6. Hailstones, D.L. and P.W. Gunning. 1990. Mol. Cell. Biol. **12**:1095–1104.
7. Zimmermann,K. and A. Starzinski-Powitz. 1989. Nucl. Acids Res. **17**:10496.
8. Barton, P.J.R., B. Robert, A. Cohen, I. Gardner, D. Sassoon, A. Weydert and M.E. Buckingham. 1988. J. Biol. Chem. **263**:12669–12676.
9. Kurabayashi, M., I. Komuro, H. Tsuchimochi, F. Takaku and Y. Yazaki. 1988. J. Biol. Chem. **263**:13930–13936.
10. Arnold, H.H., Lohse, P., Seidel, U. and E. Bober. 1988. Eur. J. Biochem. **178**:53–60.
11. Zimmermann,K., Kautz,S., Hajdu,G., Winter,C., Whalen,R.G. and A. Starzinski-Powitz. 1990. J. Mol. Biol. **211**: 505–513.
12. Seharaseyon,J., Bober,E., Hsieh,C.-L., Fodor,W.L., Francke,U., Arnold,H.H. and E.F.Vanin (1990). Genomics **7**, 289–293.
13. Nabeshima, Y., Y. Fujii-Kuriyama, M. Muramatsu and K. Ogata. 1984. Nature (London) **308**:333–3381.
14. Whalen,R.G., Butler-Browne,G.S. and F. Gros. 1978. J. Mol. Biol. **126**, 415–431.
15. Whalen, R.G. and S.M. Sell. 1980. Nature **286**, 731–733.
16. Barton, P.J.R., Robert,B., Fiszman,M.Y., Leader,D.P. and M.E. Buckingham. 1985. J. of Muscle Research and Motility **6**, 461–475.
17. Soussi-Yanicostas,N., Barbet,J.P., Laurent-Winter,C., Barton,P. and G.S. Butler-Browne. 1990. Development **108**, 239–249.
18. Carraro,U., dalla Libera,L. and C. Catani. 1983. Exp. Neurol. **79**, 106–117.

19. Favaloro, J., R. Freisman, and R. Kamen. 1980. Methods Enzymol. **65**:718–749.
20. Ehrich, E., Craig, A., Poustka, A., Frischauf, A.-M., and H. Lehrach. 1987. Gene **57**:229–237.
21. Sambrook, J., Fritsch, E.F. and T. Maniatis. 1989. Molecular Cloning: A Laboratory Mamial. Cold Spring Harbor Laboratory Press. Cold Spring Harbor., N.Y., USA.
22. Sanger, F., S. Nicklen and A.R. Coulson. 1977. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
23. Saiki, R.K., Gelfaud, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, D.B. and H.A. Erlich. 1988. Science **239**:487–491.
24. Buskin, J.N. and S.D. Hauschka. 1989. Mol. Cell. Biol. **9**:2627–2640.
25. Lassar,A.B., Buskin,J.N., Lockshon,D., Davis,R.L., Apone,S., Hauschka,S.D. and H.Weintraub (1989). Cell **58**, 823–831.
26. Cohen, A., Barton, P.R., Robert, B., Garner, I., Alonso, S. and M.E. Buckingham. 1988. Nucl. Acids Res. **16**:10037–10052.
27. Daubas, P., Robert,B., Garner,I. and M.Buckingham (1985). Nucl. Acids Res. **13**, 4623–4643.
28. Jantzen, H.M., Strähle, U., Gross, B., Stewart,F., Schmid,W., Boshart,M., Miksicek,R. and G.Schütz (1987). Cell **49**, 29–38.
29. Palmiter, R.D., Mulvihill, E.R., McKnight, G.S. and A.W. Senear. 1980. Cold Spring Harbour Symp. Quant. Biol. **42**:639–647.
30. Nabeshima, Y., Nabeshima, Y., Kawashima, M., Nakamura, S., Nonomura, Y. and Y. Fujii-Kuziyama. 1988. J. Mol. Biol. **204**:, 497–505.