Structure and expression of rat muscle-specific enolase gene

Kenji Sakimura[1], Etsuko Kushiya[1], Yuka Ohshima-Ichimura[2], Hiromi Mitsui[2,*] and Yasuo Takahashi[1]

[1] *Department of Neuropharmacology, Brain Research Institute* and [2] *Department of Biology, Faculty of Science, Niigata University, Niigata 951, Japan*

Received 20 October 1990

The structure of rat muscle-specific enolase ($\beta\beta$ enolase) gene was determined. It comprises 12 exons of various lengths (59–223 bp) spanning about 6 kbp and its exon–intron organization is similar to that of neuron-specific enolase ($\gamma\gamma$ enolase) gene. A transcriptional start site was identified by a combination of S1 nuclease mapping and primer extension analyses. In the 5′-flanking region we found a TATA-box-like sequence and several MyoD-binding motifs. The in vitro cell free transcription of the truncated genomic DNA fragment using HeLa cell extract showed that the transcription start site has been correctly identified and the promoter sequences work well.

Gene; Exon; Intron; Enolase; Transcription; Muscle

1. INTRODUCTION

The $\beta\beta$ isozyme of enolase (2-phospho-D-glycerate hydrolase, EC 3.2.1.11), a glycolytic enzyme, is present in mature muscle and known as muscle-specific enolase (MSE). A switch-over from non-neuronal enolase (NNE; $\alpha\alpha$ enolase) to MSE occurs during the final stage of cell differentiation and the amounts of MSE increase with the functional maturation of myotubes.

We previously cloned cDNA for rat NNE, neuron-specific enolase (NSE; $\gamma\gamma$ enolase) and MSE and determined their nucleotide sequences [1–3]. We have also isolated a genomic clone coding for rat NSE, elucidated its structural organization and studied its expression [4]. After we published a paper on MSE cDNA [1], Lamandé et al. reported a similar result [5]. In this paper, we describe the isolation of a genomic clone coding for rat MSE, the exon–intron organization and the expression of this gene in a cell-free transcription system.

2. MATERIALS AND METHODS

A rat genomic library was kindly provided by Dr J. Bonner. Two DNA fragments from cDNA for rat MSE (M 35) labeled with [α-^{32}P]αCTP by nick translation or random primed labelling, were used as the probes. The genomic library was screened by the plaque-hybridization technique of Benton-Davis using the probes [6]. The DNA inserts of positive clones were subcloned into pUC8 or pUC118 vectors. Restriction maps generated by digestion with appropriate restriction enzymes and Southern blot hybridization analyses revealed the DNA fragments containing exons. Nucleotide sequences were determined by the method of Maxam-Gilbert [7] and dideoxy-sequencing method [8]. In order to determine the transcription initiation site, primer extension and S1 nuclease mapping analyses were carried out as described previously [4].

Preparation of the HeLa cell extract and in vitro transcription of MSE gene in this cell extract were performed by the procedure of Manley et al. [4,9]. Plasmid pMX1-4 as a template contained 1781 bp, 642 bp, 331 bp or 180 bp of the 5′-flanking region, exon 1, intron 1 and a portion of exon 2, of MSE gene. Each DNA template added was truncated at a unique restriction site (*Nco*I). Another template, plasmid pMX5 contained 141 bp of the 5′ flanking region, a portion of the exon 1 (22 bp) and a spacer (187 bp). The reaction conditions were described in detail previously [4] and in the legend of Fig. 4.

Correspondence address: Y. Takahashi, Department of Neuropharmacology, Brain Research Institute, Niigata University, Asahimachi-dori 1, Niigata 951, Japan

* *Present address:* Kyoto Pharmaceutical College, Yamashina, Kyoto 607, Japan

3. RESULTS AND DISCUSSION

3.1. *Isolation and characterization of the rat MSE gene*

From the rat genomic library, 4 clones were obtained using cDNA probe containing 3′-noncoding region which contained the nucleotide sequence unique to MSE mRNA. Three overlapping clones out of 4 were isolated and designated λME1, λME2 and λME3 and used for further analyses (Fig. 1). DNA of each clone was digested with restriction endonucleases and analyzed by Southern blot hybridization. Furthermore, two DNA fragments from λME2 (E3-4 and E7-2) and three DNA fragments from λME3 (HB6-6, H4-6 and H2-9) were subcloned into the multicloning sites of pUC8 vector. The subcloned DNA fragments were subjected to further restriction endonuclease analysis and to sequencing by the procedures of Maxam-Gilbert and Sanger et al. Fig. 1 also shows the restriction endonuclease map and the structure of the rat MSE gene. The *Bam*HI-*Pst*I region (8.4 kbp) was sequenced except some parts of introns. The rat MSE gene is about 6 kbp long and

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

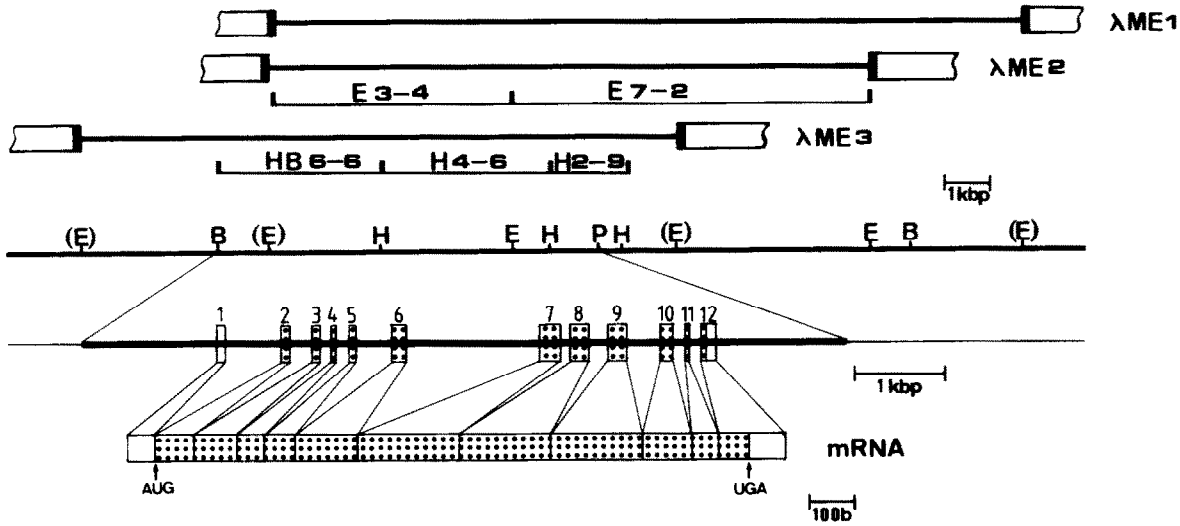

Fig. 1. Restriction map of the rat MSE gene cloned in λ phages and structure of the rat MSE gene. Three λ phage clones (λME1, λME2 and λME3) hybridized with MSE cDNA are shown with the subcloned DNA fragments from λME2 and λME3. A restriction endonuclease map and structure of MSE gene are indicated in the middle. Exons are shown by the numbered stippled and open boxes. Stippled boxes represent the coding regions and open boxes represent the noncoding regions. Thick bars show the introns and the flanking regions and these regions were sequenced except some parts of introns. The structure of MSE mRNA corresponding to that of MSE gene is shown below.

is separated into 12 exons by 11 introns, Fig. 2A shows the nucleotide sequence of 5′-flanking region, exon 1, intron 1 and exon 2 of MSE gene. The nucleotide sequence of the exon (1–12) regions completely corresponded to the cDNA sequence. The boundary sequences of the exon–intron junctional sequences in the gene were all referable to the typical 5′-GT-AG3′ rule as shown in Fig. 2B [10]. The length of each exon varied from 59 bp (exon 11) to 223 bp (exon 7). The 3′-noncoding region is 151 bp long and is located in exon 12. The 5′-noncoding region is 63 bp long and is split into exons 1 and 2 as described later. The translation start codon ATG is located in exon 2 and the stop codon TAG in exon 12. The nucleotides for the putative substrate-binding site, arginin residue, are split into exon 11 and 12. The nucleotides for the Mg^{2+}-binding site, histidine residue, are located in exon 7.

3.2. *Characterization of the transcription initiation site and 5′-flanking region of the MSE gene*

For the determination of the 5′-terminus of the gene, we used S1 nuclease mapping analyses and the primer extension. The protected 5′ end by S1 nuclease mapping analysis was nucleotide C as shown in Fig. 3. Primer extension gave a product corresponding to C at 63 bp upstream to the translation initiation site. The most extended broad major band of 136 bases was found by the primer extension analysis (data not shown). Maxam-Gilbert sequencing of DNA eluted from this major band gave the same nucleotide sequence to MSE cDNA extended to base C at −63 on the base of the translation initiation site. Sequence analysis of DNA from the upper minor band did not give any cDNA sequence. Nucleotide C was therefore assigned as the transcription initiation site (+1). In the case of MSE gene, the transcription start site was single, different from those of NSE gene.

Fig. 2A also shows the nucleotide sequence of 5′-end of the MSE gene. The nucleotide sequence upstream from the MSE mRNA start point contains a TATA-box-like sequence at −26 to −22. CAAT and GC boxes are absent. About 1.8 kbp 5′ flanking sequence was examined by the homology search. Ten MyoD-binding motifs (CANNTG) were found in this region. Further, CA-rich G box sequence (−227 to −238) was also found. The significance of these findings will be described in the section of general discussion.

3.3. *The comparison of MSE gene with NSE gene*

The structure of MSE gene in Fig. 1 was compared with that of NSE gene [4]. Both genes were similar in the exon–intron organization, indicating the presence of 12 exons. The total size of the MSE gene from the transcription start point to the poly(A) addition site was about 6 kbp and considerably smaller than that of NSE gene. The reason for this difference was due to smaller size of all introns except introns 6 and 9. Particularly the size of the introns 1 and 7 of MSE gene was much smaller than that of NSE gene introns. However, the intron 6 of MSE was longer than that of NSE. As we previously reported in the cDNA, 3′ noncoding region in the exon 12 was much smaller than that of NSE gene. It is important that the transcriptional start site of MSE gene is single, indicating the difference from the multiple start sites in NSE gene. Furthermore, in the 5′ flanking region MSE gene contains a TATA-box-like sequence, but NSE gene did not contain any typical TATA-box-like sequence.

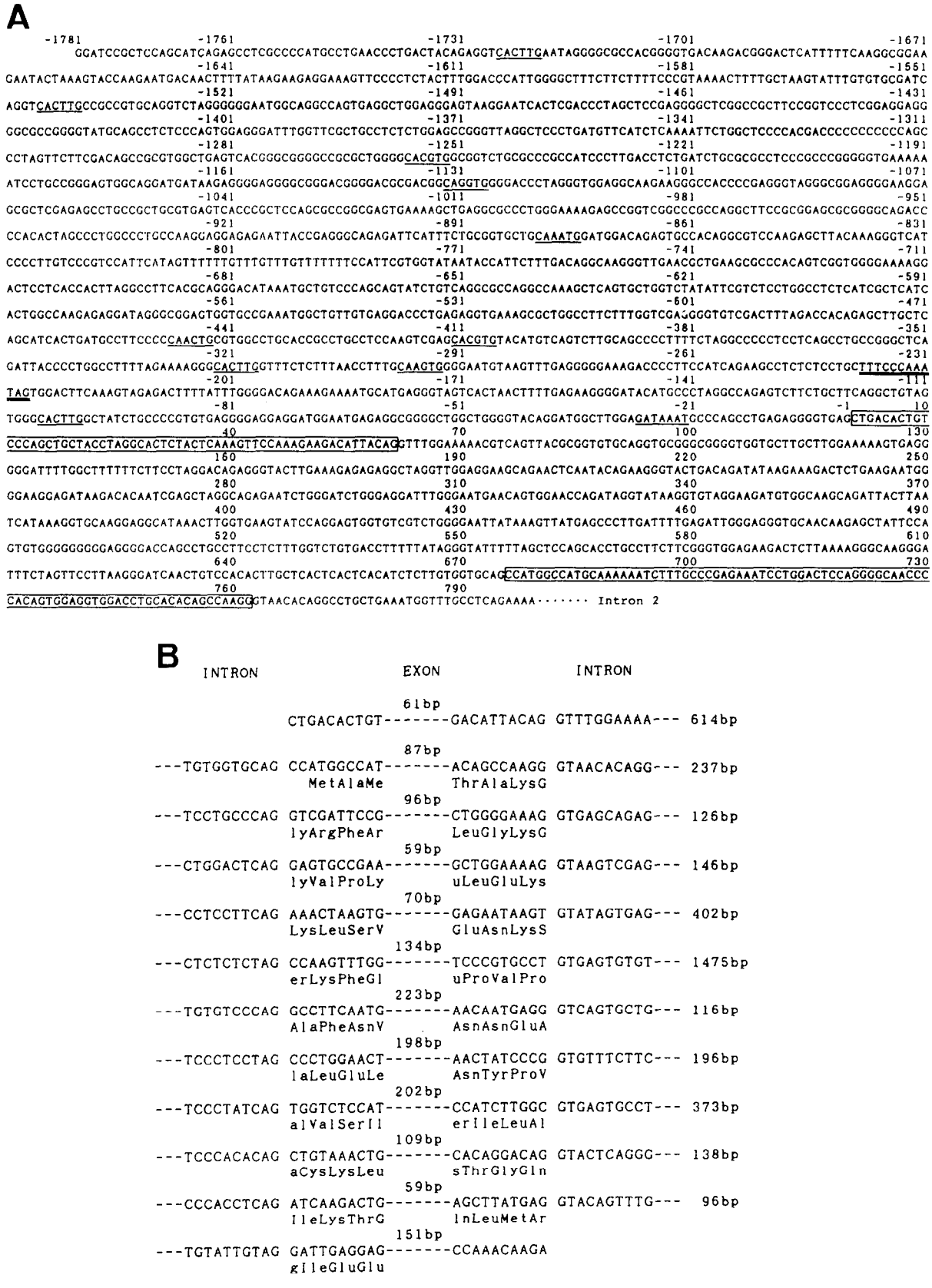

Fig. 2. The nucleotide sequence of the 5′-end and the exon–intron junctional regions of MSE gene. (A) Nucleotide position +1 is assigned to the C of the transcriptional start site; negative numbers refer to the nucleotides on the 5′ side of nucleotide 1. The TATA-box-like sequence and MyoD binding motifs are underlined and the CA-rich G-box sequence is double-underlined. The nucleotides in the exons are boxed. (B) Exon–intron junctional regions.

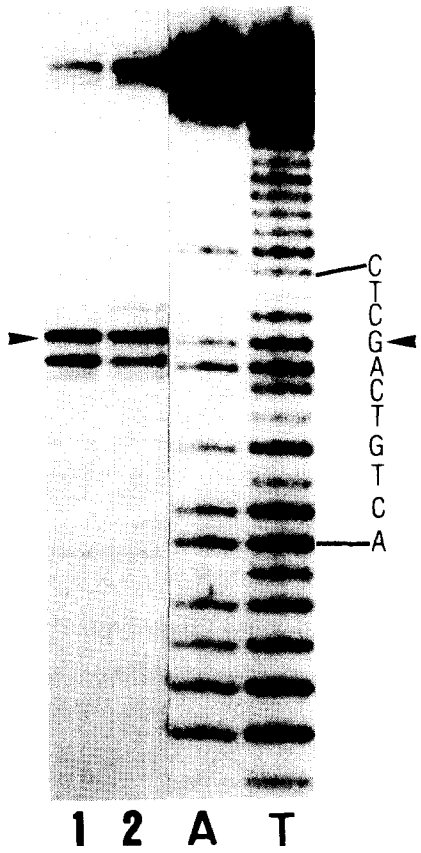

Fig. 3. S1 nuclease protection analysis. The arrow indicates the protected bands. The concentrations of S1 nuclease were 120 and 240 units/ml respectively in lanes 1 and 2. A and T show the sequence ladder of the same probe DNA.

The boundary amino acids in the exon-intron junctional regions were almost similar between MSE and NSE except two: Lys → Ser in the 3′ end of the exon 4 and Lys → Gly in the 5′ side of the exon 5. Considering this similarity of the exon-intron organization in the MSE and NSE genes, it was evolutionally presumed that both enolase isozyme genes should come from the same ancestral gene.

3.4. *In vitro cell-free transcription of the MSE gene*

The MSE gene is predominantly expressed in the mature muscle and the regulation of its expression may take place at the level of transcription. To examine the potential activity of the putative promoter region for the transcription initiation of MSE gene and to look into the mechanism of its muscle-specific expression, we first attempted to study in vitro cell-free transcription of the truncated genomic DNA fragment as a template using HeLa cell extract.

Plasmid pMX1-4 containing 1781 bp, 642 bp, 331 bp or 180 bp of the 5′-flanking region, exon 1, intron 1 and exon 2 was truncated with *Nco*I and used as a template. Some α-amanitin sensitive transcripts were observed. A faithful major transcript (about 676 bp), which corresponded to the transcription start point, was obtained. Another plasmid, pMX5 containing 141 bp of the 5′-flanking region, a portion of exon 1 (22 bp) and a spacer (187 bp), was used as a template. In this case, a faithful transcript was also obtained (209 bp). These transcripts were all α-amanitin-sensitive (Fig. 4).

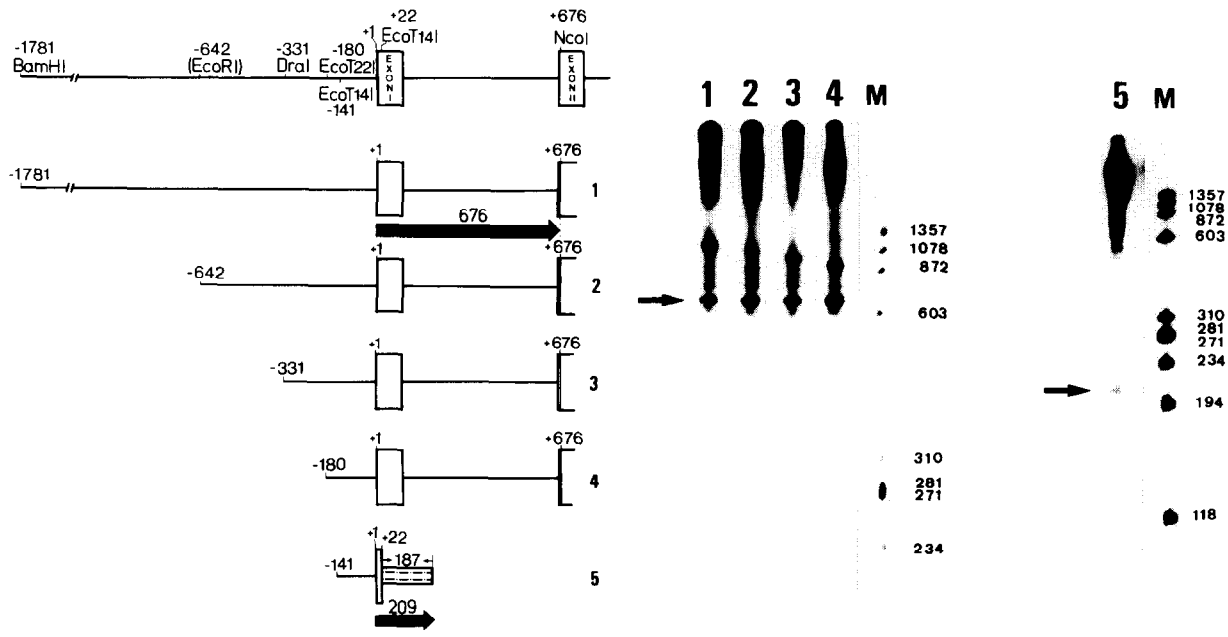

Fig. 4. Analysis of RNA transcripts synthesized in vitro. *Nco*I-truncated pMX1-4 DNAs contain 1781 bp, 642 bp, 331 bp or 180 bp of the 5′-flanking region, the first exon, intron 1 and exon 2 of MSE gene. Another plasmid, pMX5, contains 141 bp of the 5′-flanking region, a portion of exon 1 (22 bp) and a spacer (187 bp). These structures are shown on the left side. The putative cap site is marked with +1. The standard 12.5 μl reaction mixture contained 12 mM Hepes, pH 7.9, 7 mM $MgCl_2$, 60 mM KCl, 0.2 mM EDTA, 1.3 mM DTT, 10% (v/v) glycerol, 10 mM creatine phosphate, 600 μM GTP and CTP, 25 μM ATP, 50 μM UTP (10 μCi of [α-^{32}P]UTP), 5–7.5 μl of HeLa cell extract and 0.5 μg of cleaved plasmid DNA. The reaction mixtures were incubated at 30°C for 60 min. The transcripts were analyzed on the 5% polyacrylamide – 7 M urea gel. Lanes 1–5 indicate the transcripts from pMX1-5, and do not contain α-amanitin. The concentration of α-amanitin is 1 μg/ml. Lane M shows the molecular markers. The arrows indicate the bands of the faithful transcripts (lanes 1–4, 676 bp; lane 5, 209 bp).

These data suggest that for the active transcription HeLa cell extract could recognize the putative promoter sequence (at least 141 bp) which contains a TATA-box-like sequence and the putative transcription start point is correct and single. This 141-bp region contains the promoter sequence enough for transcription of MSE gene.

3.5. *General discussion*

Recently, the *cis*-acting regulatory regions required for transcriptional regulation have been examined for several muscle-specific genes; quail troponin I [11,12], human cardiac actin [13], and mouse muscle creatine kinase genes [14,15]. Some studies have identified an enhancer element in the quail muscle troponin I [11,12] and mouse muscle creatine kinase genes [14,15]. Furthermore, some *trans*-acting factor proteins like MyoD and Myf-5 etc. were found in muscle creatine kinase [15] and myosin light chain genes [16]. The sequence homology search in 1.8 kbp of the 5′ flanking region of MSE gene could clarify the presence of at least ten muscle-specific elements homologous to the sequence CANNTG motif to which a muscle-specific trans-acting factor, MyoD binds. A CA-rich G-box sequence was also found as previously described in actin gene by Minty and Kedes [13]. The regulatory function of these nucleotide sequences and a *trans*-acting factor from muscle nuclear extract should be examined in detail in future. Further, to study the molecular mechanism that controls the expression of MSE gene in the specific early developmental stages of the muscle cells, the gene structure of NNE should be clarified. Very recently the structure of human NNE gene was described [17]. Already we found a switching appearance and increase of MSE mRNA from NNE mRNA in the course of differentiation from myoblast to myotubules (unpublished observation). The MSE gene will be used as a model system to examine the mechanism of muscle-specific gene expression.

REFERENCES

[1] Ohshima, Y., Mitsui, H., Takayama, Y., Kushiya, E., Sakimura, K. and Takahashi, Y. (1989) FEBS Lett. 242, 425–430.
[2] Sakimura, K., Kushiya, E., Obinata, M. and Takahashi, Y. (1985) Nucleic Acids Res. 13, 4365–4378.
[3] Sakimura, K., Kushiya, E., Obinata, M., Odani, S. and Takahashi, Y. (1985) Proc. Natl. Acad. Sci. USA 82, 7453–7457.
[4] Sakimura, K., Kushiya, E., Takahashi, Y. and Suzuki, Y. (1987) Gene 60, 103–113.
[5] Lamandé, N., Mazo, A.M., Lucas, M., Montarras, D., Pinset, C., Gros, F., Legault-Demare, L. and Lazar, M. (1989) Proc. Natl. Acad. Sci. USA 86, 4445–4449.
[6] Benton, W.D. and Davis, R.W. (1977) Science 196, 180–182.
[7] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[8] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[9] Manley, J.L., Fire, A., Cano, A., Sharp, P.A. and Gefter, M.L. (1980) Proc. Natl. Acad. Sci. USA 77, 3855–3859.
[10] Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) Proc. Natl. Acad. Sci. USA 75, 4853–4857.
[11] Konieczny, S.F. and Emerson Jr, C.P. (1985) Mol. Cell. Biol. 5, 2423–2432.
[12] Konieczny, S.F. and Emerson, Jr, C.P. (1987) Mol. Cell. Biol. 7, 3065–3075.
[13] Minty, A. and Kedes, L. (1986) Mol. Cell. Biol. 6, 2125–2136.
[14] Lassar, A.B., Buskin, J.N., Lockshon, D., Davis, R.L., Apone, S., Hauschka, S.D. and Weintraub, H. (1989) Cell 58, 823–831.
[15] Jayes, J.B., Johnson, J.E., Buskin, J.N., Gartside, C.L. and Hauschka, S.D. (1988) Mol. Cell. Biol. 8, 62–70.
[16] Donoghue, M., Ernst, H., Wentworth, B., Nadal-Ginard, B. and Rosenthal, N. (1988) Genes Dev. 2, 1779–1790.
[17] Giallongo, A., Oliva, D., Cali, L., Barba, G., Barbieri, G. and Feo, S. (1990) Eur. J. Biochem. 190, 567–573.