Volume 242, number 2, 425–430 FEB 06661 January 1989

cDNA cloning and nucleotide sequence of rat muscle-specific enolase ($\beta\beta$ enolase)

Yuka Ohshima, Hiromi Mitsui, Yoichi Takayama*, Etsuko Kushiya+, Kenji Sakimura+ and Yasuo Takahashi+

*Department of Biology, Faculty of Science, *Department of Hygiene, Faculty of Medicine* and *+Department of Neuropharmacology, Brain Research Institute, Niigata University, Niigata 951, Japan*

Received 15 September 1988; revised version received 11 November 1988

The nucleotide sequence of rat muscle-specific enolase cDNA was determined by sequencing three cDNA clones encoding this enolase isozyme. The nearly full-length cDNA consists of 13-bp 5′- and 84-bp 3′-noncoding regions and a poly(A) tail in addition to a 1302-bp coding region encoding a polypeptide composed of 434 amino acid residues. The deduced primary structure of this enolase isozyme is about 80% similar to those determined previously for rat neuron-specific and non-neuronal enolase isozymes. Southern blot analysis suggested strongly the existence of a single copy of the muscle-specific enolase gene per haploid genome. The mRNA for this enolase isozyme was detected in rat skeletal muscle on day 1 after birth and its level increased rapidly during 10–30 days after birth without any change in its size (1500 bases).

mRNA; Amino acid sequence; Enolase; cDNA cloning; (Muscle)

1. INTRODUCTION

Three isozymes of enolase (2-phospho-D-glycerate hydrolase, EC 3.2.1.11), a glycolytic enzyme, occur in mature mammalian tissues. They are all homodimers and designated as $\alpha\alpha$, $\beta\beta$ and $\gamma\gamma$ [1]. The $\alpha\alpha$ isozyme is found in many tissues as well as in glial cells of nervous tissue and called non-neuronal enolase (NNE). The $\gamma\gamma$ isozyme occurs in neurons and neuroendocrine cells and is termed neuron-specific enolase (NSE) [1]. The $\beta\beta$ isozyme is present in mature muscle and known as muscle-specific enolase (MSE) [2]. A switch-over from NNE to NSE or to MSE occurs during the final stage of cell differentiation and the amounts of NSE and MSE increase with the functional maturation of neurons and myotubes, respectively [1,2].

We have previously cloned cDNAs for rat NNE [3] and NSE [4] and determined their nucleotide sequences. We have also isolated a genomic clone coding for rat NSE, elucidated its structural organization, and studied its expression [5]. In this paper, we report the molecular cloning of cDNA of rat MSE ($\beta\beta$), the nucleotide sequence of this cDNA and developmental changes of the expression of the MSE gene.

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

Abbreviations: MSE, muscle-specific enolase; NSE, neuron-specific enolase; NNE, non-neuronal enolase; bp, base pairs; b, bases

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00979

2. MATERIALS AND METHODS

RNA was prepared from rat skeletal muscle by phenol-chloroform extraction, followed by the lithium precipitation procedure. Poly(A)$^+$ RNA was isolated by oligo(dT)-cellulose chromatography [6]. A cDNA library of rat skeletal muscle was constructed from this poly(A)$^+$ RNA by the procedure of Gubler and Hoffman [7]. The cDNA library was screened under conditions of low stringency with a *Bgl*II-*Pst*I fragment of pERN11 which contained a part of the coding region of NNE

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50

cDNA (6 × SSC, 0.5% SDS, 1% FBP, 500 μg/ml salmon sperm DNA and 55°C). The filters were washed with 2 × SSC at 55°C and subjected to autoradiography to select the desired clones [8].

For RNA blot analysis, RNA was isolated from various rat tissues using phenol-chloroform extraction, followed by lithium precipitation as described above. RNA was also isolated from rat skeletal muscle of various ages by using the guanidine thiocyanate-CsCl centrifugation procedure [9]. Poly(A)$^+$ RNA or total RNA was denatured in 2.2 M formaldehyde for 10 min at 60°C, subjected to electrophoresis on a 1.2% agarose gel containing 2.2 M formaldehyde and transferred by capillary blotting to a nitrocellulose filter. The filter was hybridized with the cDNA probe under the conditions described. The cDNA probe was labeled with [α-^{32}P]dCTP by a random-primed DNA labeling kit (Boehringer Mannheim, FRG) [9,10]. High-molecular mass genomic DNA was prepared from rat brain by the procedure of Bline and Stafford [11]. DNA blot analysis was carried out with a 3′-terminus/*Hin*fI fragment of pM35 cDNA as an MSE specific probe which was nick-translated with [α-^{32}P]dCTP [12]. The DNA sequence was determined by the method of Maxam-Gilbert [13].

3. RESULTS AND DISCUSSION

Since it was expected that there is close sequence similarity between the coding regions of MSE and NNE cDNAs, we screened a rat muscle cDNA library under conditions of low stringency using a part of the coding region of NNE cDNA as a probe and isolated several positive clones. Sequence analysis of the longest clone (pM35) showed that its insert has a poly(A) tract at its 3′-terminus and a poly(A) addition signal (AATAAA). A stop codon was also found 84 bp upstream from the poly(A) addition site. The amino acid sequence deduced from the nucleotide sequence of its coding region was closely similar to those of the C-terminal halves of NSE and NNE. We therefore concluded that the insert of pM35 is actually a truncated MSE cDNA that lacks the 5′-terminal

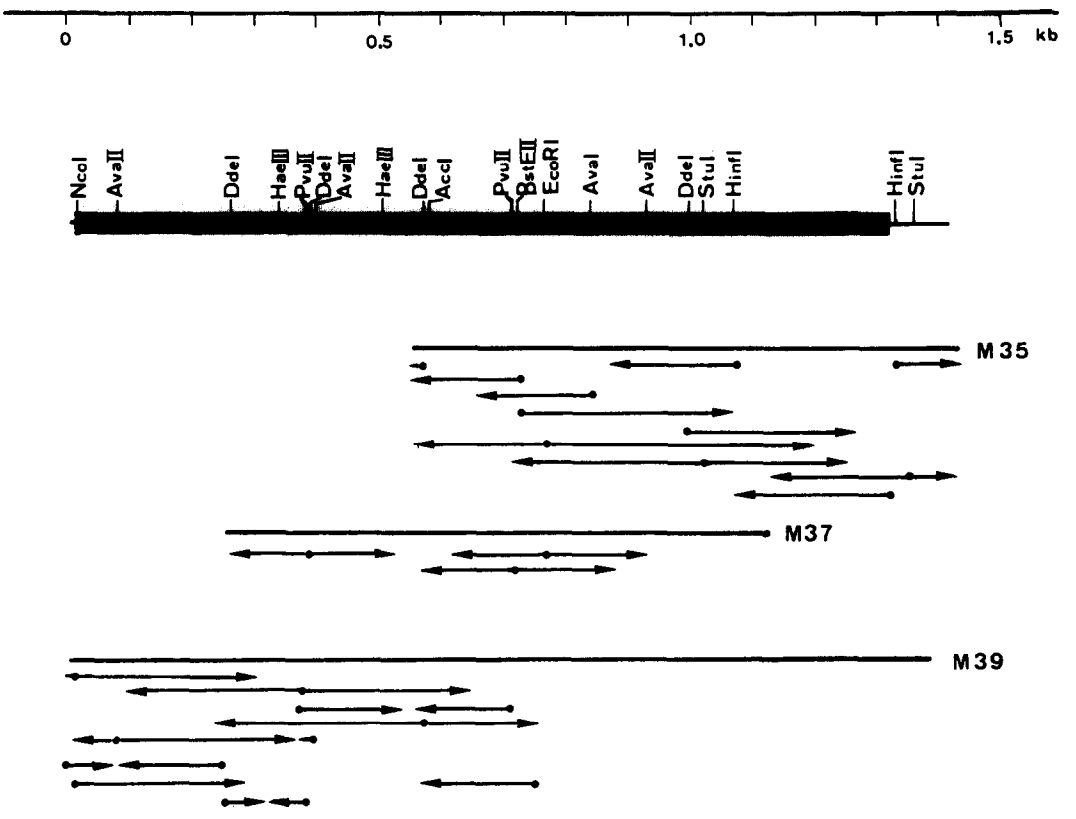

Fig.1. Strategy for nucleotide sequence determination. The arrows indicate the direction and extent of cDNA sequenced. All the nucleotides were determined in both directions. The filled box indicates the coding region. The relevant restriction endonuclease sites are shown.

```
        10         20         30         40         50         60         70         80         90
AGACATTACAGCCATGGCCATGCAAAAAATCTTTGCCCGAGAAATCCTGGACTCCAGGGGCAACCCCACAGTGGAGGTGGACCTGCACAC
              M  A  M  Q  K  I  F  A  R  E  I  L  D  S  R  G  N  P  T  V  E  V  D  L  H  T

       100        110        120        130        140        150        160        170        180
AGCCAAGGGTCGATTCCGAGCGGCTGTGCCAAGCGGAGCTTCCACGGGTATCTATGAAGCACTGGAGCTACGAGATGGAGACAAATCACG
 A  K  G  R  F  R  A  A  V  P  S  G  A  S  T  G  I  Y  E  A  L  E  L  R  D  G  D  K  S  R

       190        200        210        220        230        240        250        260        270
ATACCTGGGGAAAGGAGTGCCGAAGGCTGTGGAGCACATCAACAAGACTCTCGGCCCTGCTCTGCTGGAAAAGAAACTAAGTGTTGTGGA
 Y  L  G  K  G  V  P  K  A  V  E  H  I  N  K  T  L  G  P  A  L  L  E  K  K  L  S  V  V  D

       280        290        300        310        320        330        340        350        360
TCAAGAAAAAGTTGACAAATTTATGATTGAGCTGGACGGAACCGAGAATAAGTCCAAGTTTGGGGCCAATGCCATCCTGGGTGTGTCCCT
 Q  E  K  V  D  K  F  M  I  E  L  D  G  T  E  N  K  S  K  F  G  A  N  A  I  L  G  V  S  L

       370        380        390        400        410        420        430        440        450
GGCTGTTTGTAAGGCTGGAGCAGCTGAGAAAGGGGTCCCTCTCTACCGACACATTGCAGACCTCGCAGGGAATCCCGACCTTGTACTTCC
 A  V  C  K  A  G  A  A  E  K  G  V  P  L  Y  R  H  I  A  D  L  A  G  N  P  D  L  V  L  P

       460        470        480        490        500        510        520        530        540
CGTGCCTGCCTTCAATGTGATCAACGGGGGCTCTCATGCTGGAAACAAGTTGGCCATGCAGGAGTTCATGATTCTGCCAGTGGGAGCCAA
 V  P  A  F  N  V  I  N  G  G  S  H  A  G  N  K  L  A  M  Q  E  F  M  I  L  P  V  G  A  K

       550        560        570        580        590        600        610        620        630
GCTCTTTCAAGGAAGCCAGCGCATTGGCGCTGAGGTCTACCACCACCTCAAGGGGGTCATCAAAGCCAAGTATGGGAAGGATGCCACCAA
 L  F  Q  G  S  Q  R  I  G  A  E  V  Y  H  H  L  K  G  V  I  K  A  K  Y  G  K  D  A  T  N

       640        650        660        670        680        690        700        710        720
TGTGGGGGACGAGGGTGGCTTTGCACCCAACATCCTGGAGAACAATGAGGCCCTGGAACTACTAAAGACAGCCATTCAGGCAGCTGGTTA
 V  G  D  E  G  G  F  A  P  N  I  L  E  N  N  E  A  L  E  L  L  K  T  A  I  Q  A  A  G  Y

       730        740        750        760        770        780        790        800        810
CCCAGACAAGGTGGTGATTGGCATGGATGTGGCAGCATCTGAATTCTATCGCAATGGGAAGTATGATCTGGACTTCAAGTCACCTGATGA
 P  D  K  V  V  I  G  M  D  V  A  A  S  E  F  Y  R  N  G  K  Y  D  L  D  F  K  S  P  D  D

       820        830        840        850        860        870        880        890        900
TCCTGCGAGGCACATCAGTGGGGAGAAGCTCGGGGAGCCGTACAAGTCCTTCATCAAGAACTATCCCGTGGTCTCCATTGAAGACCCCTT
 P  A  R  H  I  S  G  E  K  L  G  E  P  Y  K  S  F  I  K  N  Y  P  V  V  S  I  E  D  P  F

       910        920        930        940        950        960        970        980        990
TGACCAGGATGACTGGGCAACATGGACCTCATTCCTCTCTGGGGTGGACATCCAGATTGTGGGAGATGACCTGACAGTCACCAACCCCAA
 D  Q  D  D  W  A  T  W  T  S  F  L  S  G  V  D  I  Q  I  V  G  D  D  L  T  V  T  N  P  K

      1000       1010       1020       1030       1040       1050       1060       1070       1080
GAGGATTGCTCAGGCTGTTGAGAAGAAGGCCTGCAATTGCCTGCTGCTGAAGGTCAACCAGATCGGCTCTGTGACAGAATCCATCTTGGC
 R  I  A  Q  A  V  E  K  K  A  C  N  C  L  L  L  K  V  N  Q  I  G  S  V  T  E  S  I  L  A

      1090       1100       1110       1120       1130       1140       1150       1160       1170
CTGTAAACTGGCACAGTCTAATGGCTGGGGAGTGATGGTGAGCCACCGCTCTGGGGAGACCGAAGACACTTTCGTTGCTGATCTTGTGGT
 C  K  L  A  Q  S  N  G  W  G  V  M  V  S  H  R  S  G  E  T  E  D  T  F  V  A  D  L  V  V

      1180       1190       1200       1210       1220       1230       1240       1250       1260
GGGACTCTGCACAGGACAGATCAAGACTGGTGCTCCCTGCCGTTCAGAGCGTCTGGCAAAATACAACCAGCTTATGAGGATTGAGGAGGC
 G  L  C  T  G  Q  I  K  T  G  A  P  C  R  S  E  R  L  A  K  Y  N  Q  L  M  R  I  E  E  A

      1270       1280       1290       1300       1310       1320       1330       1340       1350
TCTCGGAGACAAAGCTGTCTTTGCTGGACGAAAGTTCCGTAATCCAAAGGCCAAATGAGAAGCTGGAGATTCCACGATTCCACAGGACAG
 L  G  D  K  A  V  F  A  G  R  K  F  R  N  P  K  A  K  *

      1360       1370       1380       1390
ACACAGGCCTCCAAGCCCTTCTCCCAGAAATAAACACTGCCAAACAAGA...poly(A)...
```

Fig.2. Nucleotide sequence of cDNA and deduced amino acid sequence of rat MSE. The nucleotides are numbered in the 5′ to 3′ direction. The poly(A) signal is underlined.

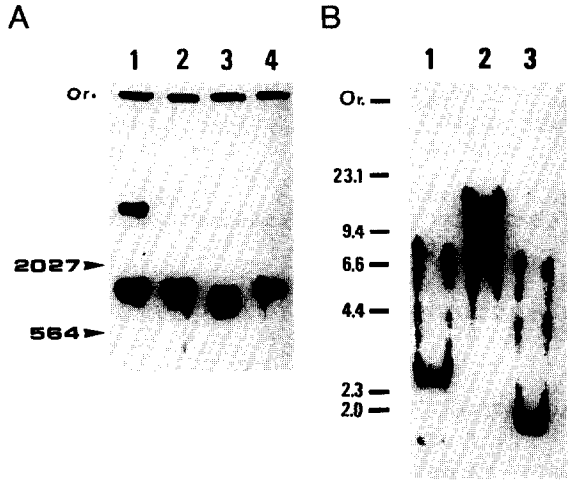

Fig.3. (A) Northern blot of poly(A)$^+$ RNA (5 μg) from various rat tissues. Lane 1, heart; lane 2, small intestine; lane 3, skeletal muscle; lane 4, kidney. The coding region of rat NNE cDNA was used as a probe. Washing was carried out as described in the text. The arrowheads indicate the size marker in bases. (B) Southern blot analysis. Rat genomic DNA (10 μg) digested with several endonucleases was electrophoresed. The size markers are indicated in kbp. A 3′-terminus/*Hin*fI fragment of pM35 containing the 3′-noncoding region was used as the probe. Lane 1, *Hin*dIII; lane 2, *Bam*HI; lane 3, *Pst*I.

portion. To obtain clones carrying longer inserts, the cDNA library was rescreened using the 5′-terminus/*Eco*RI fragment of pM35 cDNA as a probe, resulting in the isolation of two clones (pM37 and pM39). The restriction endonuclease maps of pM35, pM37 and pM39 as well as the strategy adopted for their sequencing are shown in fig.1. Although our sequencing of pM37 and pM39 did not cover the entire length, the sequences of the three cDNAs were in perfect match with one another at their overlapping portions, indicating that these three cDNAs had been derived from the same mRNA. The nucleotide sequence composed of the sequenced portions of the three cDNAs as well as the amino acid sequence deduced therefrom are shown in fig.2. The cDNA is 1399 bp long and contains a 13-bp 5′-noncoding region and an 84-bp 3′-noncoding region, which is followed by a poly(A) tail. The coding region encodes a polypeptide consisting of 434 amino acid residues.

When rat skeletal muscle poly(A)$^+$ RNA was analyzed by Northern blot hybridization by using a part of the coding region of rat NNE cDNA as a probe under conditions of low stringency, a

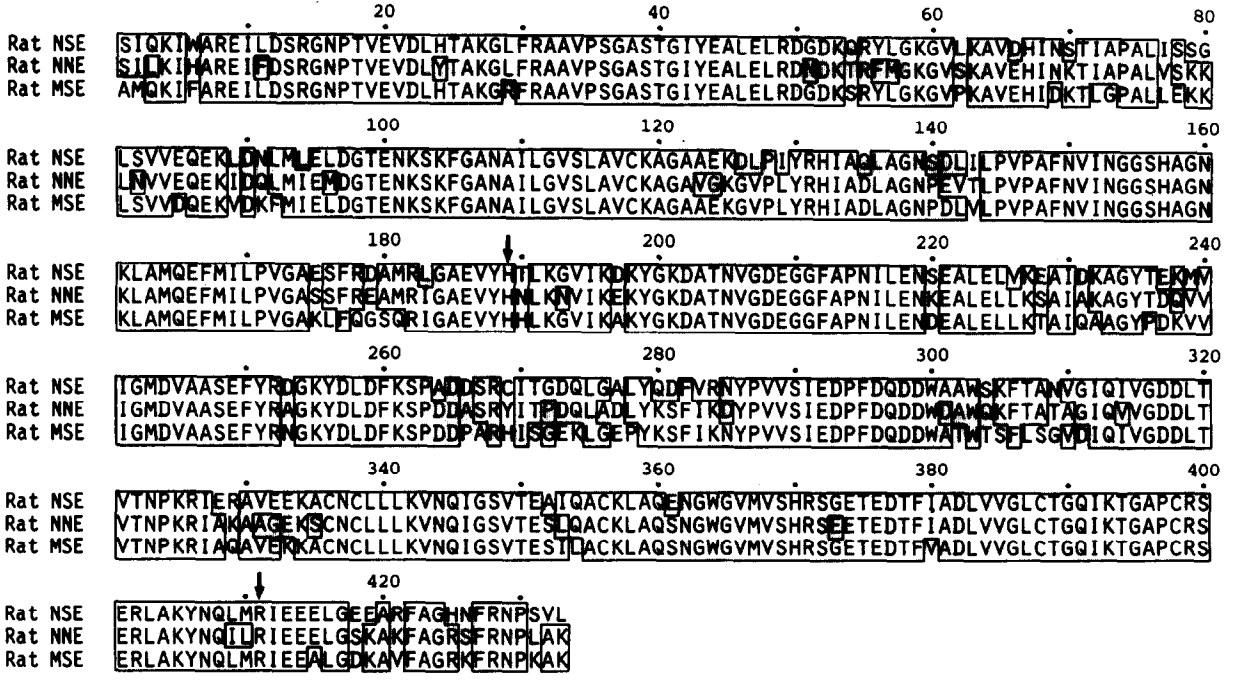

Fig.4. Comparison of NSE, NNE and MSE. The initiation methionine was omitted. One-letter symbols were used for amino acid residues. The homologous residues are boxed. The arrow at position 189 indicates the Mg ion-binding site. The arrow at position 411 indicates the substrate binding site.

single mRNA band of about 1500 b was detected (fig.3A, lane 3). If we assume that the mRNA contains a poly(A) tail of about 100 b, it can be concluded that our cDNA covers nearly the full length of MSE mRNA, even though the 5′-noncoding region is very short. The size of MSE is significantly smaller than that of NNE mRNA from heart, intestine and kidney (fig.3, lanes 1, 2, and 4) and that of NSE (not shown). This reflects the fact that the 3′-noncoding regions of NSE, NNE and MSE cDNAs are 848, 312 and 84 bp long, respectively. This finding is consistent with Sutcliffe's hypothesis that the 3′-noncoding regions in brain-specific genes are longer than those of the genes that are expressed in other tissues [14]. Southern blot analysis of *Hin*dIII, *Bam*HI and *Pst*I digests of rat genomic DNA using a 3′-noncoding fragment of pM35 as a probe showed one major and one or two minor bands (fig.3B). This suggests strongly that there exists only one copy of the MSE gene per haploid genome.

As is the case for NSE and NNE, rat MSE is composed of 434 amino acid residues. A comparison of the amino acid sequence of MSE with those of NSE and NNE is shown in fig.4. The sequence similarity between MSE and NSE is 81% and that between MSE and NNE is 80%. The amino acid sequence of rat MSE is 83% similar to that of chicken MSE, which has been determined by protein sequencing procedures [15]. This 83% sequence similarity between MSEs from a rodent and a bird is remarkable. Although the similarity of nucleotide sequences among rat NSE, NNE and MSE cDNAs is 75% in their coding regions, there is practically no sequence similarity in their 5′- and 3′-noncoding regions.

We also examined the developmental changes in the level of MSE mRNA in rat skeletal muscle by Northern blot hybridization using a 3′-noncoding fragment of MSE cDNA as a probe. As shown in fig.5 (lanes 1–5), MSE mRNA was detectable on day 1 after birth and its level increased rapidly until day 30, especially during days 10–30. No change in the size of mRNA (1500 b) was detected during the whole period examined. The level of MSE mRNA in skeletal muscle of adult rats was high (lane 6), but no hybridization signal was observed in rat liver and brain poly(A)$^+$ RNA (lanes 7 and 8).

We have recently isolated a genomic clone encoding rat MSE and its structural study is currently being performed. Such a study will provide an opportunity for clarifying the differences among the enolase isozyme genes and the mechanisms for tissue-specific expression of these genes.

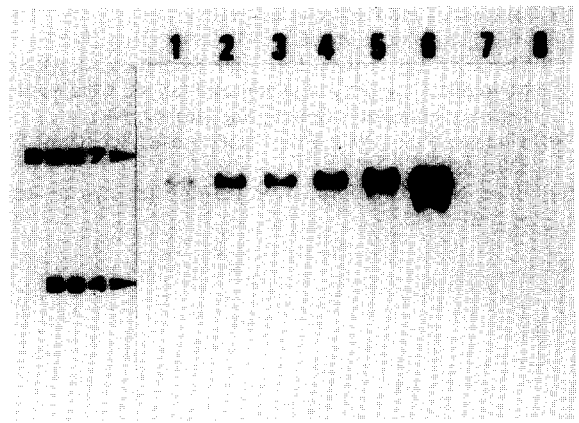

Fig.5. Developmental changes in the level of MSE mRNA. Northern blot of cytoplasmic RNA (10 μg) from rat skeletal muscle (lane 1, postnatal day 1; lane 2, day 5; lane 3, day 10; lane 4, day 20; lane 5, day 30) and poly(A)$^+$ RNA (3 μg) from adult skeletal muscle (lane 6), liver (lane 7) and brain (lane 8) was carried out as described in the text. The filter was hybridized with the ^{32}P-labeled 3′-noncoding region of MSE cDNA (50% formamide, 5 × SSC, 100 μg/ml salmon sperm DNA, 10% dextran sulfate and 42°C).

Acknowledgements: This study was supported in part by a grant-in-aid for scientific research from the Ministry of Education, Science and Culture of Japan to Y.T. and K.S. This study was carried out under the NIBB Cooperative Research Program (87-121).

REFERENCES

[1] Zomzely-Neurath, C.E. (1983) in: Handbook of Neurochemistry (Lajtha, A. ed.) 2nd edn, vol.4, pp.403–433, Plenum, New York.
[2] Rider, C.C. and Taylor, C.B. (1975) Biochim. Biophys. Acta 405, 175–187.
[3] Sakimura, K., Kushiya, E., Obinata, M. and Takahashi, Y. (1985) Nucleic Acids Res. 13, 4365–4378.
[4] Sakimura, K., Kushiya, E., Obinata, M., Odani, S. and Takahashi, Y. (1985) Proc. Natl. Acad. Sci. USA 82, 7453–7457.
[5] Sakimura, K., Kushiya, E., Takahashi, Y. and Suzuki, Y. (1987) Gene 60, 103–113.
[6] Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408–1412.

[7] Gubler, U. and Hoffman, B.J. (1983) Gene 25, 263–270.
[8] Hanahan, D. and Meselson, M. (1980) Gene 10, 63–67.
[9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[10] Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201–5205.
[11] Bline, N. and Stafford, D.W. (1976) Nucleic Acids Res. 3, 2303–2306.
[12] Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237–251.
[13] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[14] Sutcliffe, J.G., McKinnon, R.D. and Tsou, A.-P. (1986) in: Role of RNA and DNA in Brain Function (Giuditta, A. et al. eds) pp.23–31, Martinus Nijhoff, Boston.
[15] Russel, G.A., Dunbar, B. and Fothergill-Gilmore, L.A. (1986) Biochem. J. 236, 115–126.