Genetics and Molecular Biology, 25, 1, 43-47 (2002)

www.sbg.org.br

Three synonymous genes encode calmodulin in a reptile, the Japanese tortoise, *Clemmys japonica*

Kouji Shimoda, Toshihiro Miyake, Jun Kimura, Kazuyoshi Maejima

Laboratory Animal Center, Keio University School of Medicine, Tokyo, Japan

Abstract

Three distinct calmodulin (CaM)-encoding cDNAs were isolated from a reptile, the Japanese tortoise (*Clemmys japonica*), based on degenerative primer PCR. Because of synonymous codon usages, the deduced amino acid (aa) sequences were exactly the same in all three genes and identical to the aa sequence of vertebrate CaM. The three cDNAs, referred to as *CaM-A*, *-B*, and *-C*, seemed to belong to the same type as *CaMI*, *CaMII*, and *CaMIII*, respectively, based on their sequence identity with those of the mammalian cDNAs and the glutamate codon biases. Northern blot analysis detected *CaM-A* and *-B* as bands corresponding to 1.8 kb, with the most abundant levels in the brain and testis, while *CaM-C* was detected most abundantly in the brain as bands of 1.4 and 2.0 kb. Our results indicate that, in the tortoise, CaM protein is encoded by at least three non-allelic genes, and that the 'multigene-one protein' principle of CaM synthesis is applicable to all classes of vertebrates, from fishes to mammals.

Key words: reptile, cDNA, synonymousgenes.

Received: March 14, 2002; accepted: March 25, 2002.

Introduction

Calmodulin (CaM), a prototype of EF-hand Ca^{2+}-binding protein, is ubiquitously distributed in all eukaryotic cells, and is particularly abundant in the brain and testes of mammals (Kakiuchi *et al.*, 1982). CaM function is essential for the viability of cells and tissues, as well as for neuronal activity in the central nervous system (CNS) (Kennedy, 1989). When CaM binds to the Ca^{2+} ion, the protein becomes active, and it in turn activates many other enzymes. It thus plays a pivotal role as a cofactor regulating a wide variety of calcium-dependent proteins.

The amino acid (aa) sequence of CaM protein is well conserved in organisms whose evolution diverged millions of years ago. Three non-allelic *CaM* genes encoding exactly the same 148 aa protein (17,000 daltons) were identified in the rat (Nojima, 1989), mouse (Bender *et al.*, 1988; Danchin *et al.*, 1989), and humans (Fischer *et al.*, 1988); two were identified in *Xenopus* (Chien and Daxid, 1984), and four in the teleost fish medaka (*Oryzias latipes*) (Matsuo *et al.*, 1992). Although the reason for the presence of multiple synonymous genes in animals is unclear, differences in the 5' and 3' non-coding region of the rat *CaM* genes suggest that each gene may be differentially regulated. The overall distribution of the three CaM transcripts was coordinated in rat brain (Ikeshima *et al.*, 1993; Palfi *et al.*, 1999), however, temporal and spatial differences in the expression of *CaM* genes have been observed in some areas of the CNS, both under physiological and experimental conditions (Michelhaugh *et al.*, 1998; Palfi and Gulya, 1999; Palfi *et al.*, 2000, Palfi *et al.*, 2001, Shimizu *et al.*, 1997; Sola *et al.*, 1997; Weinman *et al.*, 1991). Knowledge of the evolutionary relationships between vertebrate *CaM* genes should provide useful information (in) regarding (to) this multiple-gene system. To this study, we attempted to isolate CaM cDNAs from a reptile, the Japanese tortoise (*Clemmys japonica*), which occupies a critical position in the phylogenetic tree of vertebrate evolution. The results indicate that reptiles also possess three active *CaM* genes, corresponding to *CaMI*, *CaMII*, and *CaMIII*, respectively, and thus the 'multigene-one protein' principle of CaM synthesis is as applicable to reptiles as it is to other vertebrates.

Send correspondence to Kouji Shimoda, Laboratory Animal Center, Keio University School of Medicine, 35-Shinanomachi, Shinjuku-ku, 160-8582, Tokyo, Japan. E-mail: shimodak@sc.itc.keio.ac.jp.

Materials and Methods

Isolation of total RNA

The Japanese tortoise was purchased from a commercial source. The animal was anesthetized by ethyl-ether and euthanized by decapitation on ice. Dissected organs (brain, liver, kidneys, intestines, heart, and testes) were immediately frozen in liquid nitrogen. Total organ RNA

preparations were obtained by the guanidine-thiocyanate-CsCl ultra-centrifugation method (Chirgwin *et al.*, 1979).

Amplification, subcloning and sequencing of CaM cDNA

A set of degenerate primers was synthesized by a DNA synthesizer, as described in a previous paper (Matsuo *et al.*, 1992): Amino-terminal: 5'-ATG GCN GAY CAR CTN ATN GA and Carboxyl-terminal: 5'-GGGTCTAGA YA YTT NGC NGT CAT CAT YTG. N represents an A, G, C, or T base; Y, a C or T base; and R, an A or G base. The underlined bases indicate an *Xba*I site. First-strand cDNA was synthesized from 2 μg of total RNA using 50 pmol of the Carboxyl-terminal degenerate primer as described by Kawasaki (1990). RCR was performed according to the methods described in a previous paper (Matsuo *et al.*, 1992). PCR products were treated with Klenow (large) fragment of *E. coli* DNA polymerase I, digested with *XbaI*, and fractionated in 1.5% agarose gel. About 450-bp bands were cloned between the *Sma I* and *Xba I* sites of pUC19. Nucleotide sequence analysis was performed by a DNA sequencer (373A, Applied Biosystems, CA) with Cycle Sequencing Kit (Takara, Tokyo), according to the manufacturer's instructions. Each nucleotide sequence was confirmed from both directions. The sequences of these cDNAs appeared in the DDBJ/EMBL/GenBank under accession Nos. AB055384 (*CaM-A*), AB055385 (*CaM-B*) and AB055386 (*CaM-C*).

Southern blot hybridization

One hundred ng of cloned CaM genes, CaM-A, -B and -C, were digested with XbaI and EcoRI restriction enzymes, submitted to electrophoresis in 0.8% agarose gel, and then blotted onto a NYTRAN membrane (Schleicher & Schuell). The membrane was pre-hybridized at 65 °C in hybridization solution (6 x SSC/50 mM sodium phosphate buffer, pH7.0/2 x Denhardt's solution/0.5% SDS/0.2 mg salmon sperm DNA/mL). A random hexamer-primed ^{32}P-labeled DNA probe of each coding region was added and allowed to hybridize for 20 h at 65 °C. The membrane was washed twice in 2 x SSC/0.1% SDS at room temperature for 5 min, twice in 0.1 x SSC/0.1% SDS at 65 °C for 20 min, and exposed to X-ray film overnight at -80 °C.

Northern blot analysis

Five μg of total tortoise RNA were denatured with glyoxal and dimethyl sulfoxide, submitted to electrophoresis in 1.0% agarose gel, and then blotted onto a NYTRAN membrane. The membrane was pre-hybridized for 4 h at 68 °C in hybridization solution (50% formamide/5 x SSC/50 mM sodium phosphate buffer, pH7.0/2 x Denhardt's solution/0.5% SDS/0.2 mg salmon sperm DNA/mL). A random hexamer-primed ^{32}P-labeled DNA probe of each coding region was added and allowed to hybridize for 20 h at 68 °C. The membrane was washed three times in 0.2 x SSC/0.1% SDS at 68 °C for 20 min, and exposed to X-ray film overnight at -80 °C.

Results

Isolation of three CaM-coding regions

Our previous study showed that the degenerate primers encoding the Amino- and Carboxyl-terminal six aa of CaM are useful for reverse transcription and PCR of total teleost fish RNA (Matsuo *et al.*, 1992). In this study, we also employed the same pair of degenerate primers to isolate CaM cDNA from transcripts of tortoise brain (Figure 1). Eleven of the 60 cDNA clones obtained fell into three groups of distinctive CaM sequences, designated *A*, *B*, and *C*. The nucleotide sequences of these three groups of cDNA are shown in Figure 1. The deduced aa sequences were

```
  1  A  D  Q  L  T  E  E  Q  I  A  E  F  K  E  A  F  S  L  F  D
  1  ----------------AGAACAGATTGCTGAGTTCAAGGAAGCTTTCTCCCTATTTGAC
                     ...G........A..A.....A.....C.....A.........
                     G........C..A..A...........C........T..C...

 21  K  D  G  D  G  T  I  T  T  K  E  L  G  T  V  M  R  S  L  G
 61  AAAGATGGGGATGGCACCATCACAACAAAAGAACTTGGGACTGTCATGAGGTCATTGGGT
     ..G.....T.....T..T..A..C..C..G...T.G.....A..G........C.....A
     ..G.....T...........T..C..C..G..G..G..T..C...........C..A..G

 41  Q  N  P  T  E  A  E  L  Q  D  M  I  N  E  V  D  A  D  G  N
121  CAAAACCCAACAGAAGCAGAATTACAGGATATGATCAATGAGGTAGATGCTGATGGTAAT
     ........C..C.....T..G.....................A.............C...
     ..G.....C..C........G..G.....C.................C........G...

 61  G  T  I  D  F  P  E  F  L  T  M  M  A  R  K  M  K  D  T  D
181  GGTACTATCGATTTCCCTGAATTCTTGACCATGATGGCCAGAAAAATGAAGGACACTGAC
     ..C..C..T..C.....A.....TC....T........A...........A..T..A...
     ..C..C..T..C........G...C.A...........A..G...........T..A...

 81  S  E  E  E  I  R  E  A  F  R  V  F  D  K  D  G  N  G  Y  I
241  AGTGAAGAAGAAATCCGTGAAGCATTCCGAGTCTTTGACAAGGATGGGAATGGGTACATC
     ..............TA.A.....C.....T..G.....T.................T..T
     ..C..G..G..G........G..C.....G..G..............G..........T

101  S  A  A  E  L  R  H  V  M  T  N  L  G  E  K  L  T  D  E  E
301  AGTGCTGCAGAATTACGCCACGTGATGACAAACCTGGGGGAAAAGCTGACAGATGAAGAA
     ............C.T.....T...........T..T..A..G.....A...........
     .....G..C..GC.G.....T........C..TT.A..C..............C..G..G

121  V  D  E  M  I  R  E  A  D  I  D  G  D  G  Q  V  N  Y  E  E
361  GTAGACGAAATGATTCGGGAAGCAGACATAGATGGGGATGGCCAAGTAAACTATGAAGAG
     ..T..T..G.....C..............T.....T.....T...................
     ........G.....C..A..G..T........C..A.....G..G..C..T.........

141  F  V  Q  M  M  T  A  K
421  TTTGTA------------------
     ......
     .....G
```

Figure 1 - The coding sequences of tortoise CaM cDNA and their deduced aa sequence. The deduced common aa sequence is shown in the first line, and the coding sequences of the cDNA of the three tortoise CaMs, *CaM-A*, *-B*, and *-C*, are shown in the second, third, and fourth line, respectively. Dashes indicate nucleotide corresponding to the primers. Dots indicate the nucleotide identical to those of *CaM-A*. The sequence of these cDNAs will appear in the DDBJ/EMBL/GenBank under accession Nos. AB055384 (*CaM-A*), AB055385 (*CaM-B*) and AB055386 (*CaM-C*).

exactly the same, and they were identical to those of the vertebrate *CaM* genes. Since the nucleotide substitutions were scattered throughout the coding regions, the isolated coding regions were unlikely to have been generated by alternative splicing from a single gene. We therefore concluded that at least three distinct *CaM* genes, *CaM-A*, *-B*, and *-C*, were transcribed in the tortoise. The existence of three different *CaM* genes was further confirmed by Southern blot hybridization (Figure 2). Each probe of *CaM-A*, *-B*, and *-C* could detect its own clone, but not others, in stringent conditions (Tm-7 °C).

Comparative analysis of the tortoise CaM sequences and codon bias of glutamate

The sequence homologies between any two of the three tortoise cDNAs were 82% to 85%, which is very low considering that they encode CaM, *i.e.*, a gene product with identical aa sequences. Comparisons with the CaM sequences of other animals showed that *CaM-A*, *-B*, and *-C* exhibited the highest homology to *CaMI*, *CaMII*, and *CaMIII*, respectively, within each species (Table I). We therefore concluded that *CaM-A*, *-B*, and *-C* should be classified as *CaMI*, *CaMII*, and *CaMIII* genes, respectively. This conclusion was also supported by the fact that the codon usage of glutamate, which accounts for 21 residues of the CaM protein, was biased to the GAA codon in *CaMI* and *CaMII*, as opposed to the GAG codon in *CaMIII* (Table I). However, further confirmation of sequence homology in the non-coding and regulatory regions will be necessary.

Northern blot analysis

Brain, liver, kidney, intestine, heart, and testis were analyzed for expression of tortoise *CaM* genes by Northern blot hybridization (Figure 3). The *CaM-A* and *-B* probes detected bands of about 1.8 kb, respectively. The strongest expression of both genes was observed in the brain and testis. The possibility remained that the 1.8kb band was at least partially the result of cross-hybridization between *CaM-A* and *-B*, however, the *CaM-A* and *-B* genes were specifically detected by Southern blot hybridization, and the expression pattern in various tissues was slightly different. The *CaM-C* probe detected two bands of about 1.4 and 2.0 kb, most prominently in the brain. The results of other studies indicated that utilization of multiple polyadenylation sites yielded CaM RNAs of different sizes in other organisms (Lagace *et al.*, 1983; Nojima, 1989; Nojima *et al.*, 1987).

Table I - Homology of tortoise CaM genes with mammalian genes and their glutamate codon usage.

	Tortoise[a]			Glutamates[c]	
	CaM-A	CaM-B	CaM-C	GAA	GAG
Human					
CaMI	88.0*[b]	84.6	80.4	20	1
CaMII	84.6	92.4*	79.2	17	4
CaMIII	82.4	83.4	85.1*	2	19
Rat					
CaMI	87.5*	81.7	81.2	14	7
CaMII	83.9	88.5*	80.2	16	5
CaMIII	83.4	83.6	86.1	6	15
Mouse					
CaMI	84.4*	83.4	82.2	15	6
CaMII	83.9	89.7*	81.2	12	9
CaMIII	82.2	84.3	86.8*	4	17
Tortoise					
CaM-A	-	84.8	81.7	17	4
CaM-B	-	-	83.1	16	5
CaM-C	-	-	-	4	17

a) The pairwise sequence alignment was performed in the 409 nucleotide indicated in Figure 1.
b) Sequence identities (%) with corresponding mammalian genes are shown. Asterisks (*) indicate the highest homology with CaM genes in the species.
c) CaM protein contains 21 glutamate residues.

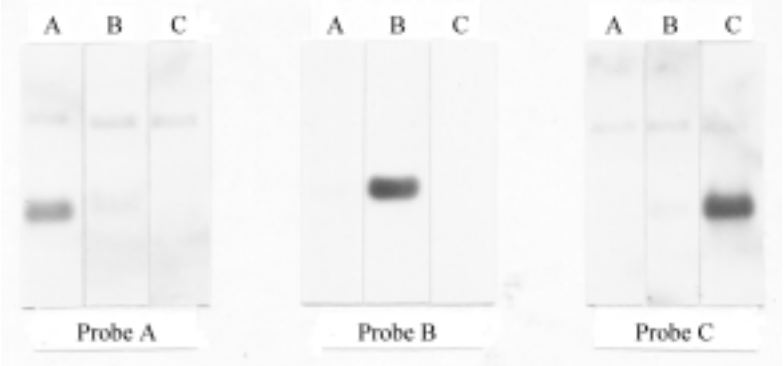

Figure 2 - Southern blot analysis of tortoise cDNAs. Tortoise cDNAs of *CaM-A*, *-B*, and *-C* were digested with *Xba*I and *EcoR*I restriction enzymes, and three sets of digested DNAs underwent electrophoresis and were then blotted onto membrane. Lanes A, B and C indicate cDNAs of *CaM-A*, *-B* and *-C*, respectively. Probes A, B, and C indicate a random hexamer-primed ^{32}P-labeled DNA probe of, respectively, the *CaM-A*, *-B*, and *-C* coding regions. Each *CaM-A*, *-B*, and *-C* probe could detect its own clone, but not others. Faint signals were visible in the position of vector plasmid pUC19 in probes A and C, and slight cross-hybridizations between *CaM-B* and probes A and C were also detectable.

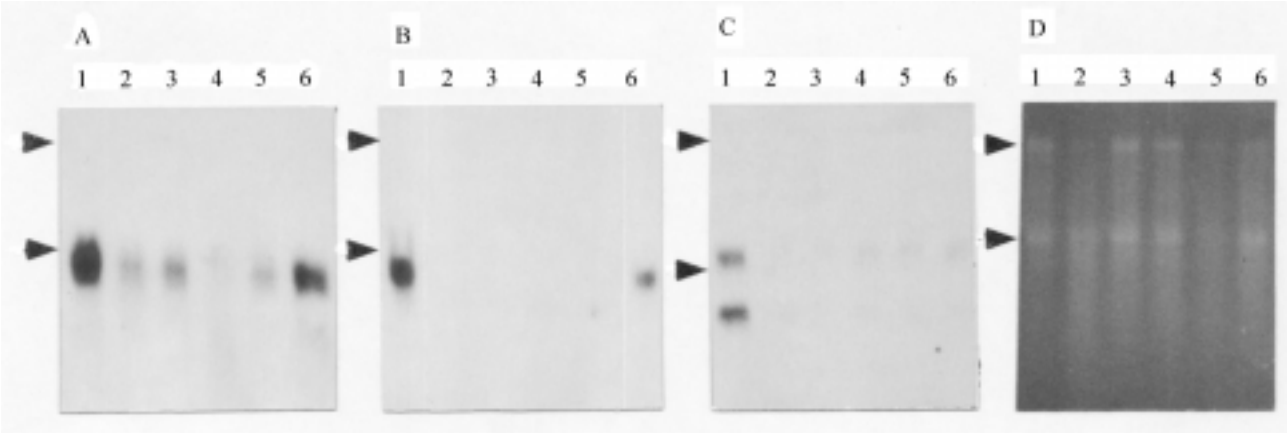

Figure 3 - Northern blot analysis of tortoise RNA in various tissues. Tortoise *CaM-A*, *-B*, and *-C* cDNAs were used as probes in panels A, B, and C, respectively, and an EtdBr-stained gel is shown in panel D. Lanes 1-6 represent total RNA isolated from brain, liver, kidney, intestine, heart, and testis, respectively. About 5 µg of total RNA underwent electrophoresis in 1.0% agarose gel. Arrowheads indicate the positions of the major ribosomal RNA bands of the Japanese tortoise.

Our results clearly demonstrate that the CaM protein of the tortoise is encoded by at least three independent CaM transcripts.

Discussion

The 'multigene-one protein' principle of CaM synthesis is conserved in the reptile, and thus seems to be common to all classes of vertebrates (Toutenhoofd and Strehler, 2000). The nt sequence homologies between any two of the three types of CaM cDNA are very low, considering that the aa sequence of CaM protein is identical in all vertebrates. Comparisons with the nt sequences of other animals may provide some information concerning the type of CaM cDNA (Table I). Our results also indicate that the codon usage of glutamate is an additional landmark for the classification of CaM genes.

Although we do not know why organisms that lost the multiple *CaM* gene system failed to survive natural selection, this multigene system may confer proper expression of CaM protein in the cells of vertebrate tissues. While all three *CaM* genes are coordinately expressed in adult rat brain (Ikeshima *et al.*, 1993), slight but definite differences in individual gene expression have been observed in CNS neurons under normal conditions (Palfi *et al.*, 1999). Differential responses of *CaM* genes in the rat brain after administration of psychotropic drugs (Michelhaugh *et al.*, 1998; Shimizu *et al.*, 1997) and a convulsant neurotoxic agent (Sola *et al.*, 1997) have also been reported. Our previous studies indicated that the disparate promoters of rat *CaMII* and *CaMIII* direct similar, but not identical, expression patterns in the CNS of adult and fetal transgenic mice (Matsuo *et al.*, 1993; Shimoda *et al.*, 1995). Further analysis of the CaM 5' and 3'-flanking regions in reptiles and other vertebrates is needed to confirm both the regulatory mechanisms and the evolutionary relationships of *CaM* genes.

Acknowledgments

We wish to express our sincere thanks to Dr. Hiroko Ikeshima for valuable suggestions. This study was partly supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan, and Grants from the Keio University.

References

Bender PK, Dedman JR and Emerson Jr CP (1988) The abundance of calmodulin mRNAs is regulated in phosphorylase kinase-deficient skeletal muscle. J. Biol. Chem. 263:9733-9737.

Chien YH and Dawid IB (1984) Isolation and characterization of calmodulin genes from *Xenopus laevis*. Mol. Cell. Biol. 4:507-513.

Chirgwin JM, Przybyla AE, MacDonald RJ and Rutter WJ (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18:5294-5299.

Danchin A, Sezer O, Glaser P, Chalon P and Caput D (1989) Cloning and expression of mouse-brain calmodulin as an activator of *Bordetella pertussis* adenylate cyclase in *Escherichia coli*. Gene 80:145-149.

Fischer R, Koller M, Flura M, Mathews M, Strehler-Page MA, Krebs J, Penniston JT, Carafoli, E and Strehler, EE (1988) Multiple divergent mRNAs code for a single human calmodulin. J. Biol. Chem. 263:17055-17062.

Ikeshima H, Yuasa S, Matsuo K, Kawamura K, Hata J and Takano T (1993) Expression of three non-allelic genes coding calmodulin exhibits similar localization on the central nervous system of adult rats. J. Neurosci. Res. 36:111-119.

Kakiuchi S, Yasuda S, Yamazaki R, Teshima Y, Kanda K, Kakiuchi R and Sobue K (1982) Quantitative determinations of calmodulin in the supernatant and particulate fractions of mammalian tissues. J. Biochem. 92:1041-1048.

Kawasaki ES (1990) Amplification of RNA. In: Innis MA, Gelfand DH, Sninsky JJ, and White, TJ (eds) PCR Protocols. Academic Press, San Diego, pp. 21-27.

Kennedy MB (1989) Regulation of neuronal function by calcium. Trends Neurosci. 12:417-420.

Lagace L, Chandra T, Woo SLC and Means AR (1983) Identification of multiple species of calmodulin messenger RNA using a full length complementary DNA. J. Biol. Chem. 258:1684-1688.

Matsuo K, Ikeshima H, Shimoda K, Umezawa A, Hata J, Maejima K, Nojima H and Takano T (1993) Expression of the rat calmodulin gene II in the central nervous system: a 294-base promoter and 68-base leader segment mediates neuron-specific gene expression in transgenic mice. Mol. Brain Res. 20:9-20.

Matsuo K, Sato K, Ikeshima H, Shimoda K and Takano T (1992) Four synonymous genes encode calmodulin in the teleost fish, medaka (*Oryzias latipes*): conservation of the multigene one-protein principle. Gene 119:279-281.

Michelhaugh SK, Pimputkar G and Gnegy ME (1998) Alteration in calmodulin mRNA expression and calmodulin content in rat brain after repeated, intermittent amphetamine. Mol. Brain Res. 62:35-42.

Nojima H (1989) Structural organization of multiple rat calmodulin genes. J. Mol. Biol. 208:269-282.

Nojima H, Kishi K and Sokabe H (1987) Multiple calmodulin mRNA species are derived from two distinct genes. Mol. Cell. Biol. 7:1873-1880.

Palfi A and Gulya K (1999) Water deprivation upregulates the three calmodulin genes in exclusively the supraoptic nucleus of the rat brain. Mol. Brain Res. 74:111-116.

Palfi A, Simonka JA, Pataricza M, Tekulics P, Lepran I, Papp G and Gulya K (2001) Postischemic calmodulin gene expression in the rat hippocampus. Life Sciences. 68:2373-81.

Palfi A, Tarcsa M, Varszegi S and Gulya K (2000) Calmodulin gene expression in an immortalized striatal gabaergic cell line. Acta Biologica Hungarica. 51:65-71.

Palfi A, Vizi S and Gulya K (1999) Differential distribution and intracellular targetting of mRNAs corresponding to the three calmodulin genes in rat brain: a quantitative in situ hybridization study. J. Histochem. Cytochem. 47:583-600.

Shimizu Y, Akiyama K, Kodama M, Ishihara T, Hamamura T and Kuroda S (1997) Alteration of calmodulin and its mRNA in rat brain after acute and chronic administration of methamphetamine. Brain Res. 765:247-258.

Shimoda K, Ikeshima H, Matsuo K, Hata J, Maejima K and Takano T (1995) Spatial and temporal regulation of the rat calmodulin gene III directed by a 877-base promoter and 103-base leader segment in the mature and embryonal central nervous system of transgenic mice. Mol. Brain Res. 31:61-70.

Sola C, Tusell JM and Serratosa J (1997) Differential response of calmodulin genes in the mouse brain after systemic kainate administraion. Neurosci. 78:155-164.

Toutenhoofd SL, and Strehler EE (2000) The calmodulin multigene family as a unique case of genetic redundancy: multiple levels of regulation to provide spatial and temporal control of calmodulin pools? Cell Calcium. 28:83-96.

Weinman J, Gaspera BD, Dautigny A, Dinh DP, Wang J, Nojima H and Weinman S (1991) Developmental regulation of calmodulin gene expression in rat brain and skeletal muscle. Cell Regul. 2:819-826.