Volume 289, number 1, 105–109 FEBS 10101 September 1991

ADONIS 0014579391008237

cDNA sequence and differential expression of the mouse Ca^{2+}/calmodulin-dependent protein kinase IV gene

David A. Jones[1], John Glod[1], Donna Wilson-Shaw[1], William E. Hahn[2] and James M. Sikela[1]

[1]*Department of Pharmacology* and [2]*Department of Cellular and Structural Biology, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA*

Received 13 June 1991

We have isolated and sequenced a mouse brain cDNA encoding Ca^{2+}/calmodulin-dependent protein kinase IV. The sequence predicts an acidic protein (pI = 4.56) of 469 amino acids (M_r = 52 627) that contains kinase catalytic and calmodulin-binding domains. The carboxy region has several primary structural features that suggest it may be a readily cleaved attachment domain. This region is highly charged and hydrophilic and contains several PEST sequences, motifs associated with high turnover proteins. Of the tissues examined, expression of the CaM kinase IV gene is restricted to brain and testis, where transcripts are differentially expressed to produce a kinase in both tissues and a calmodulin-binding protein, calspermin, in testis that lacks a kinase catalytic domain.

Ca^{2+}; Calmodulin; Protein kinase; cDNA; Brain; Testis

1. INTRODUCTION

Ca^{2+}/calmodulin-dependent protein kinases (CaM kinases) are thought to be involved in a number of important processes in the nervous system including neurotransmitter release, long term potentiation, and gene expression (reviewed in [1–3]). Recently we discovered via cDNA cloning a new CaM kinase, called CaM kinase IV. A cDNA encoding part of the CaM kinase IV protein was originally isolated by using ^{125}I-calmodulin to screen a mouse brain λgt11 expression library [4]. Subsequently, additional partial-length cDNAs thought to code for CaM kinase IV from both mouse [5] and rat (called CaM kinase-Gr) [6] were independently isolated.

The CaM kinase IV gene is located on human chromosome 5 [5] and chromosome 18 in the mouse in a region containing the ataxic (*ax*) mutation [7]. *ax* Mice are sterile and are characterized by a progressive motor disability thought to involve cerebellar functions [8]. The chromosomal location and expression at high levels in the cerebellum [6] (and in the testis as described here) suggest CaM kinase IV as a candidate gene for the *ax* mutation. To facilitate investigation of this possibility and to allow a more precise analysis of the function of this kinase to be undertaken we report the full deduced amino acid sequence of CaM kinase IV from mouse. In addition, we provide a proposed model of the organization and differential expression of the CaM kinase IV gene.

Correspondence address: J. Sikela, Department of Pharmacology, C236 University of Colorado Health Sciences Center, 4200 E. 9th Avenue Denver, Colorado 80262, USA. Fax: (1) (303) 270 7097.

2. MATERIALS AND METHODS

2.1. *cDNA isolation and DNA sequencing*

A λgt11 Balb/C mouse brain cDNA library (generously provided by Y. Citri) was screened by plaque hybridization using a partial length CaM kinase IV cDNA (ICM-2) as probe. Plaque purification and subcloning into Bluescript plasmid (Stratagene) were carried out using standard procedures [9]. DNA sequencing was performed using the dideoxy strategy [10] with modified T7 DNA polymerase (Sequenase, United States Biochemicals (USB)) and either vector- or gene-specific primers. Computer analysis of DNA and protein sequences was carried out using the PC Gene software package (Intelligenetics).

2.2. *Northern blot analysis*

Total poly(A)$^+$RNA was isolated from mouse tissues using the Fastrack mRNA Isolation kit as directed by the supplier (Invitrogen). Equal amounts (5 μg) of poly(A)$^+$RNA were electrophoresed and transferred to nylon filters as described in [11]. Hybridization and washing were carried out using standard procedures [9]. The ICM-1 plasmid probe [4] was labeled by random oligopriming [12]. Oligonucleotide probes CK4A (5′-ATCTCCCCCTGGGATGAAGTGTCTTAA-3′), CK4B (5′-GTTTTGGGATCCCCAAGAAACGGCTGACTA-3′) and CK4T (5′-GGGCGCACAGCACCCTGGCCTCAATCTCAC-3′) were specific for the A, B and T exons respectively.

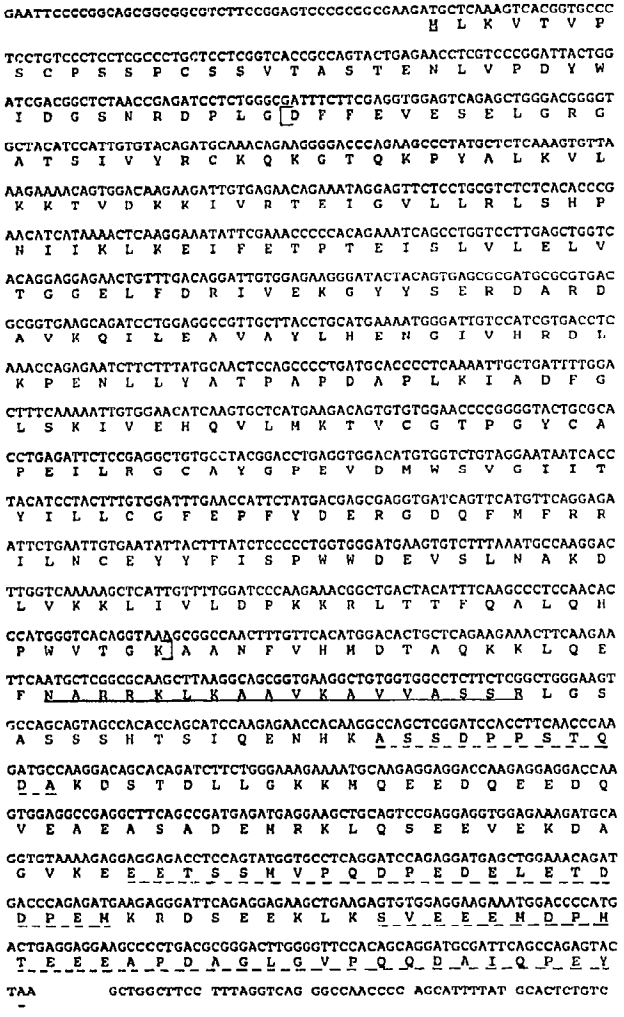
2.3. *cDNA synthesis, PCR amplification and DNA sequencing*

Single stranded cDNA was synthesized from testis mRNA and treated with *E. coli* ribonuclease H prior to being used for PCR [13]. PCR was carried out using a Perkin Elmer/Cetus Thermocycler and reagents from a Perkin Elmer/Cetus Gene Amp Kit. Primers for calspermin cDNA amplification were designed from CaM kinase IV cDNA and genomic sequences and contained restriction sites to facilitate subcloning if necessary. Primers were designated CK4T-Sal (5′-GCAGTCGACTGTCAGTAGATGAGCCAGGA-3′) and CK4G-Not (5′-GCAGCGGCCGCGGTTACCAACTGGTATATATGGG-3′). 5 μl of the testis cDNA was mixed with 15 pmol of each primer and amplified according to the following protocol: 5 cycles consisting of 94°C for 1 min, 37°C for 2 min, and 72°C for 3 min followed by 25 cycles consisting of 94°C for 1 min, 48°C for 2 min, and 72°C for 3 min. After gel purification PCR-amplified calspermin cDNA was sequenced according to the method of Carothers et al. [14].

3. RESULTS AND DISCUSSION

A partial length cDNA (ICM-2), containing part of the catalytic domain of CaM kinase IV [5] was used to screen a mouse brain cDNA library. A clone (Y11) was isolated that contained the complete coding region of the CaM kinase IV protein (Fig. 1). Sequencing of the insert indicates that it is 3165 bp in length and encodes a protein of 469 amino acids. The region around the proposed start ATG at position 46 exhibits similarity to consensus sequences associated with the start of translation [15] and therefore this ATG is probably the start site. This conclusion is consistent with Northern analysis which indicates that the largest CaM kinase IV mRNA is approximately 3.4 kb, the same approximate size of the Y11 cDNA if a poly(A) tail is included. A consensus polyadenylation signal occurs at position 2157 in Y11 and, if utilized, would generate an mRNA of approximately the size of the smaller CaM kinase IV brain mRNA. Thus, the two brain mRNAs could result from the use of alternative polyadenylation signals.

The structural topology of the kinase, based on its deduced amino acid sequence, exhibits several features typically found in CaM kinases. These include the size (261 amino acids) of the catalytic domain, the presence of several amino acid sequence motifs conserved in protein kinase catalytic domains, and the positioning of the catalytic domain approximately 20–30 amino acids upstream of the calmodulin binding site (Fig. 2). The sequence of amino acids in the carboxy region downstream of the calmodulin binding site suggests the presence of a number of functional features (Fig. 2). This region is almost entirely hydrophilic and contains an unusually high percentage of charged acidic amino acids, in particular glutamic acid and aspartic acid, a feature found in many nuclear proteins [16]. In this regard, antibodies against CaM kinase IV used as in situ probes show substantial localization of this kinase in



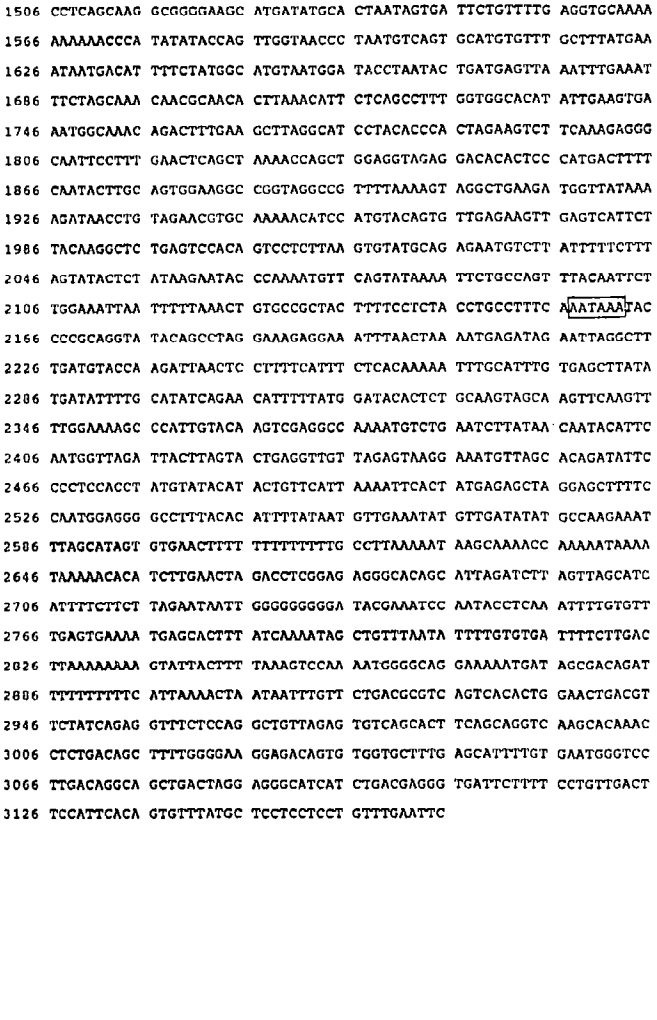


Fig. 1. cDNA and deduced amino acid sequence of mouse CaM kinase IV. Bracketed region represents the proposed catalytic domain. Regions underlined with solid and broken lines represent the predicted calmodulin-binding domain and PEST regions, respectively. A possible alternative polyadenylation signal is boxed.

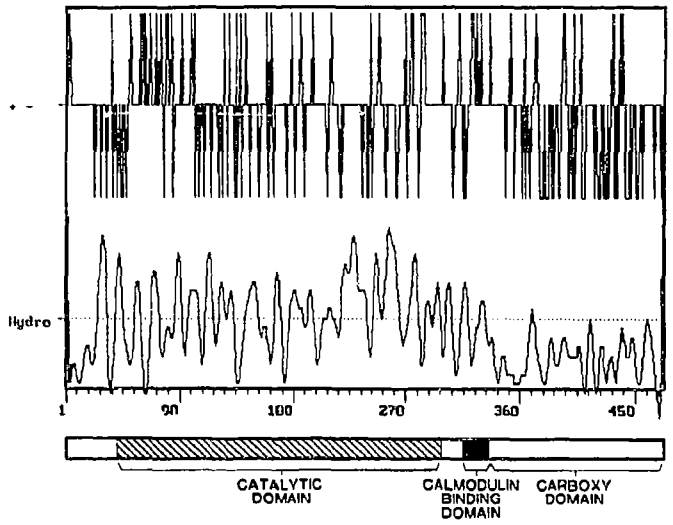


Fig. 2. Structural features of the CaM kinase IV protein. The domain organization of CaM kinase IV is depicted at the bottom. Upper profile represents the occurrence of positively (above line) and negatively (below line) charged amino acids in the CaM kinase IV protein. Middle profile represents a hydropathy analysis with hydrophobic and hydrophilic regions above and below the line, respectively.

certain neuronal nuclei (Hahn and Lasher, unpublished data; [17]). Additionally the three highest scoring PEST sequences, motifs thought to be involved in rapid protein cleavage and degradation [18], are found in this region. The clustering of PEST sequences, including a high number of Asp/Pro linkages, in this region of CaM kinase IV suggests the presence of a functional domain

```
  1 CTGCTTTGGA GCCCAAGGCT AGGGATTATT TCATTTGTTG ATCTTCCTGG
 51 CAATTCATTA TCTGTTCATC TTCAGGAACA AGAGCAAATC TTCCAGAATG
101 GGATGCTTTT AGATGAATTT AGTCATGCCC TGCAGGTCAC ACTCCTCAAG
151 CCAATGACTT AATTCTATGC TATAGTGGCT TTTATTCTGG CTCTGAACTT
201 ATTCCTTCAA GTACTTTCAA TGAGTATTTT AAAAAATGGA CTGTAGATCT
251 CAACAGTGTC ATAAAAGAGT GTTCAATAAG GGTTTGAACA CTAACCATTT
301 TGTCATTTAG TCCCCACTGA GGGAATGATG TCATTGGAGT TCTGCGTTCT
351 AAGCGAGCTC ACAGTGTGTG TGGCCTAGAA TTATGCAGAT GCTGGTTCAT
                    Exon T
401 CTCAGAAGTC ATTTGGTCAG TAGATGAGCC AGGACCTAAT CGTTTGAACT
451 CCGAAGACAG GAATCTGTGA GATTGAGGCC AGGGGTGCTG TGCGCCCTTG
501 TCTTCTAGAA TTTCCTTAGC TGCTGCTTTC CTGAGGTTGG TCCTACTCCT
551 GAGGGTGTGT GGACTTGGTG CTAGTGATAA CGGAGTAGCC AGACTTTCAC
601 AACGTGAAAC GTCCAATTTC TATAACTCTA TATTTCTTTA TGATTGCAG
          Exon B
650 GTC AAA AAG CTC ATT GTT TTG GAT CCC AAG AAA CGG CTG ACT
      V  K  K  L  I  V  L  D  P  K  K  R  L  T
692 ACA TTT CAA GCC CTC CAA CAC CCA TGG GTC ACA GGT AAA GCG
      T  F  Q  A  L  Q  H  P  W  V  T  G  K  A
734 GCC AAC TTT GTT CAC ATG GAC ACT GCT CAG AAG AAA CTT CAA
      A  N  F  V  H  M  D  T  A  Q  K  K  L  Q
776 GAA TTC
      E  F
```

Fig. 3. Sequence of mouse genomic clone p780. The proposed T and B exons are indicated and exon boundaries denoted by arrows. An in-frame stop codon found 9 bp upstream of the B exon is overlined, and consensus splice sequences found at the 3′ end of introns are indicated by dashed underlines. The deduced amino acid sequence corresponding to the 5′ part of the B exon is listed. The overlined dashes at the 3′ end of the sequence indicate that the 3′ end of the B exon is not found on the p780 sequence.

that serves to anchor the kinase to the cytoskeleton or nuclear matrix in a cleavage-sensitive manner, allowing the kinase to be released and perhaps relocated.

Alignment of the catalytic domain of CaM kinase IV with the like domain of other CaM kinases indicates that it is most closely related (40–45% identity) to the isoforms of CaM kinase II, and to kinases PSK H1 and PSK C3, two kinases of unknown function isolated by a homology probing strategy [19].

In order to clarify the relationship of CaM kinase IV and calspermin, a non-catalytic calmodulin-binding protein present in testis [20,21], we isolated and partially

```
  1                    T GGTCAGTAGA TGAGCGAGGA CCTAATCGTT TGAACTCCGA
 42 AGACAGGAAT CTGTGAGATT GAGGCCAGGG GTGCTGTGCG CCCTTGTCTT CTAGAATTTC
102 CTTAGCTGCT GCTTTCCTGA GGTCAAAAAG CTCATTGTTT TGGATCCCAA GAAACGGCTG
162 ACTACATTTC AAGCCCTCCA ACACCCATGG GTCACAGGTA AAGCGGCCAA CTTTGTTCAC
222 ATG GAC ACT GCT CAG AAG AAA CTT CAA GAA TTC AAT GCT CGG CGC AAG CTT
  1  M   D   T   A   Q   K   K   L   Q   E   F   N   A   R   R   K   L
273 AAG GCA GCG GTG AAG GCT GTG GTG GCC TCT TCT CGG CTG GGA AGT GCC AGC
 18  K   A   A   V   K   A   V   V   A   S   S   R   L   G   S   A   S
324 AGT AGC CAC ACC AGC ATC CAA GAG AAC CAC AAG GCC AGC TCG GAT CCA CCT
 35  S   S   H   T   S   I   Q   E   N   H   K   A   S   S   D   P   P
375 TCA ACC CAA GAT GCC AAG GAC AGC ACA GAT CTT CTG GGA AAG AAA ATG CAA
 52  S   T   Q   D   A   K   D   S   T   D   L   L   G   K   K   M   Q
426 GAG GAG GAC CAA GAG GAG GAC CAA GTG GAG GCC GAG GCT TCA GCC GAT GAG
 69  E   E   D   Q   E   E   D   Q   V   E   A   E   A   S   A   D   E
477 ATG AGG AAG CTG CAG TCC GAG GAG GTG GAC AAA GAT GCA GGT GTA AAA GAG
 86  M   R   K   L   Q   S   E   E   V   E   K   D   A   G   V   K   E
528 GAG GAG ACC TCC AGT ATG GTG CCT CAG GAT CCA GAG GAT GAG CTG GAA ACA
103  E   E   T   S   S   M   V   P   Q   D   P   E   D   E   L   E   T
579 GAT GAC CCA GAG ATG AAG AGG GAT TCA GAG GAG AAG CTG AAG AGT GTG GAG
120  D   D   P   E   M   K   R   D   S   E   E   K   L   K   S   V   E
630 GAA GAA ATG GAC CCC ATG ACT GAG GAG GAA GCC CCT GAC GCG GGA CTT GGG
137  E   E   M   D   P   M   T   E   E   E   A   P   D   A   G   L   G
681 GTT CCA CAG CAG GAT GCG ATT CAG CCA GAG TAC TAAGCTGGCT TCCTTTAGGT
154  V   P   Q   Q   D   A   I   Q   P   E   Y
734 CAGGGCCAAC CCCAGCATTT TATGCACTCT GTCCCTCAGC AAGGCGGGGA AGCATGATAT
794 GCACTATAGT GATTCTGTTT TGAGGTGCAA AAAAAAAACC CATATATACC AGTTGGTAAC
854 C
```

Fig. 4. DNA and deduced amino acid sequence of mouse calspermin. The putative start codon is indicated and proposed calmodulin binding domain underlined. The T exon extends from bases 1–122.

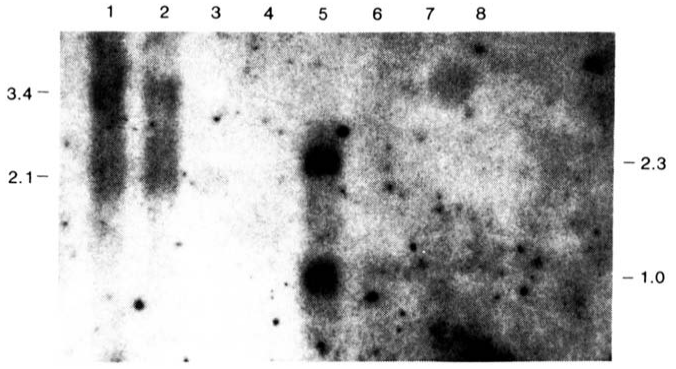


Fig. 5. Northern analysis of the tissue distribution of CaM kinase IV mRNAs. The blotted RNA was hybridized with the ICM-1 cDNA probe. Lane 1, brain; lane 2, brain (cytoplasmic mRNA); lane 3, liver; lane 4, kidney; lane 5, testis; lane 6, heart; lane 7, lung; lane 8, muscle. Size in kb is indicated for each hybridizing band.

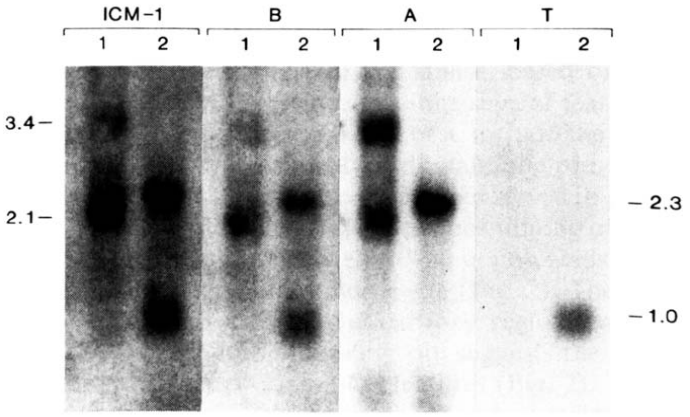


Fig. 6. Northern analysis using exon-specific probes. Blots were hybridized with ICM-1, a B exon-specific probe (CK4B), an A exon specific probe (CK4A) or a T exon-specific probe (CK4T). Lanes labeled '1' and '2' are brain and testis mRNAs, respectively.

sequenced a mouse genomic clone (λ8) (Fig. 3) containing part of the Cam kinase IV gene and also determined the sequence of mouse calspermin (Fig. 4). A 780 bp genomic restriction fragment (p780) from λ8 thought to contain sequences found in calspermin was sequenced in its entirety. Alignment of the CaM kinase IV, p780 and calspermin sequences indicates that, except for the untranslated 5′ end, calspermin mRNA is identical to the corresponding region of CaM kinase IV. The point at which the CaM kinase IV and calspermin mRNAs diverge corresponds to the region of p780 that contains consensus intron acceptor splice site sequences and an in-frame stop codon (Fig. 3). Approximately 114 bp upstream of this splice site, a sequence corresponding to the 5′ end of calspermin is found. As described below, this sequence is present in a testis-specific mRNA (calspermin) and therefore we refer to this region as the T exon. These data suggest that a testis-specific exon utilized only in calspermin is positioned slightly upstream from exons found in both CaM kinase IV and calspermin.

To further investigate the relationship of CaM kinase IV and calspermin the size and tissue distribution of mouse mRNAs that hybridize to CaM kinase IV and

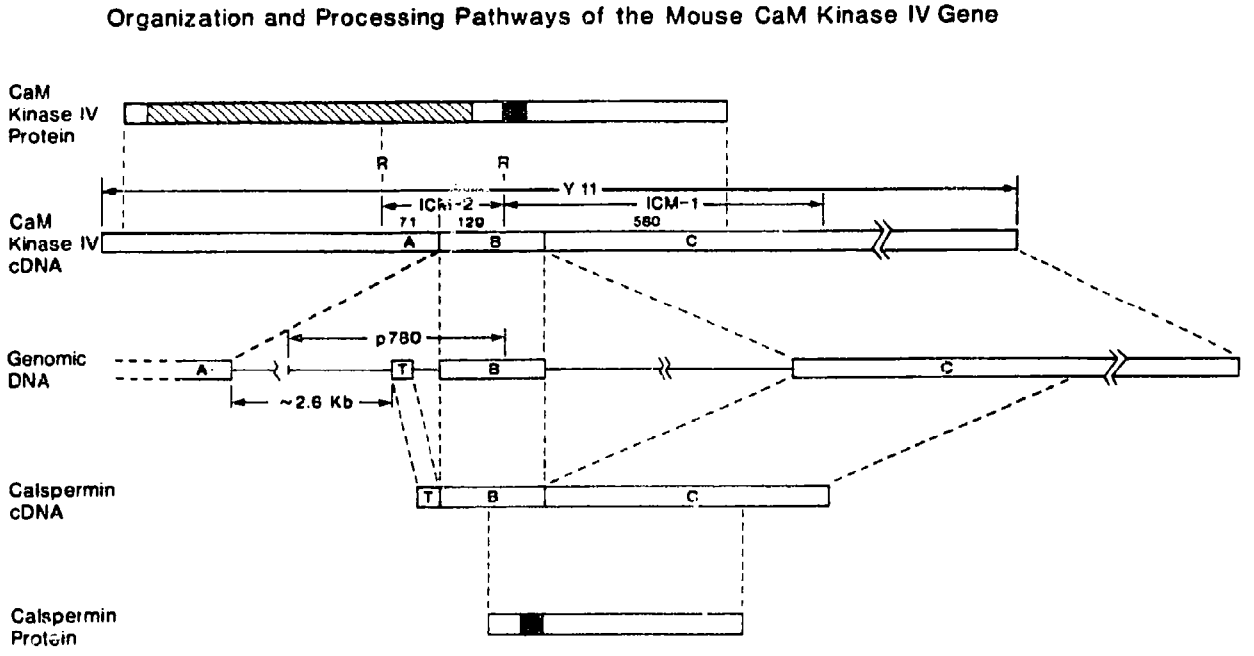


Fig. 7. Model of CaM kinase IV gene organization and processing. Letters A, B, C and T denote exons, while R refers to *Eco*RI restriction sites. The cross-hatched and solid black regions correspond to the kinase catalytic and calmodulin-binding domains, respectively. PCR amplification of mouse genomic DNA with A and T exon primers indicates these exons are approximately 2.6 kb apart (data not shown).

calspermin sequences were examined by RNA blot analysis. As demonstrated previously [4], the partial length CaM kinase IV cDNA, ICM-1, hybridized to brain mRNAs of 3.4 and 2.1 kb (Fig. 5). The ICM-1 probe also hybridized to two mRNA species in testis corresponding to 2.2 and 1.0 kb, while mRNA from other tissues showed no detectable hybridization.

Further characterization of the brain and testis messages was addressed using oligonucleotide probes corresponding to various exons in the mouse CaM kinase IV gene (Fig. 6). Blots were prepared as before using brain and testis mRNAs. The B oligo, generated from sequences upstream of ICM-1 and shared between CaM kinase IV and calspermin, hybridized to 3.4 and 2.1 kb messages in brain and 2.2 and 1.0 kb messages in testis giving the same pattern as the ICM-1 probe. The A oligo, representing the catalytic domain of CaM kinase IV did not hybridize to the testis 1.0 kb message. A fourth probe, derived from the T exon, hybridized only to the testis 1.0 kb message, consistent with T exon sequences being unique to calspermin.

Taken together, these data indicate that the single copy CaM kinase IV gene is differentially expressed to produce a protein kinase in brain and testis and also a Ca^{2+}-dependent calmodulin-binding protein (i.e. calspermin) in testis that lacks a protein kinase catalytic domain (Fig. 7). The production of calspermin transcripts from the CaM kinase IV gene could either be by differential transcriptional initiation sites or by alternative processing. If the former mechanism applies then the p780 sequences upstream of the T exon may contain signals for testis-specific transcriptional regulation and initiation.

A final point of interest pertains to the surprisingly high amino acid sequence divergence found between mouse and rat in the middle third of the ICM-1 sequence [6,21]. The data presented here indicate that this region is identical between CaM kinase IV and calspermin and therefore the divergence observed between rat and mouse in this region is the result of species differences. Such a dramatic difference in the deduced amino acid sequences of a single copy gene between two species as closely related as mouse and rat suggests that this region may be subject to an unusually high rate of mutation and positive selection.

Acknowledgements: We thank Misi Robinson, Brent Toland, Jan Hopkins and T.J. Stevens for excellent technical help. The mouse CaM kinase IV (Y11) and calspermin cDNA sequences have been deposited in the EMBL (accession number X58995) and Genbank (accession number M64266) databases, respectively. This work was supported by USPHS grant NS27322.

REFERENCES

[1] Schulman, H. (1988) Adv. Second Messenger Phosphoprotein Res. 22, 39–112.
[2] Kelly, P.T., Weinberger, R.P. and Waxham, M.N. (1988) Proc. Natl. Acad. Sci. USA 85, 4991–4995.
[3] Kapiloff, M.S., Mathis, J.M., Nelson, C.A., Lin, C.R. and Rosenfeld, M.G. (1991) Proc. Natl. Acad. Sci. USA 88, 3710–3714.
[4] Sikela, J.M. and Hahn, W.E. (1987) Proc. Natl. Acad. Sci. USA 84, 3038–3042.
[5] Sikela, J.M., Law, M.L., Kao, F.T., Hartz, J.A., Wei, Q. and Hahn, W.E. (1989) Genomics 4, 21–27.
[6] Ohmstede, C.A., Jensen, K.F. and Sahyoun, N.E. (1989) J. Biol. Chem. 264, 5866–5875.
[7] Sikela, J.M., Adamson, M.C., Wilson-Shaw, D. and Kozak, C.A. (1990) Genomics 8, 579–582.
[8] Green, M.C. (1989) in: Genetic Variants and Strains of the Laboratory Mouse (M.F. Lyon and A.G. Searle, eds) Catalog of Mutant Genes and Polymorphic Loci, Oxford Univ. Press, Oxford.
[9] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual, Cold Spring Harbor, N.Y.
[10] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[11] Fourney, R.M., Miyakoshi, J., Day III, R.S. and Paterson, M.C. (1988) Focus 10, 5–7.
[12] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[13] Saiki, R.K., Gelfand, D.H., Stoffel, S., Sharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) Science 339, 487–491.
[14] Carothers, A.M., Urlaub, G., Mucha, J., Grunberger, D. and Chasin, L.A. (1989) BioTechniques 7, 494–494.
[15] Kozak, M. (1986) Cell 44, 283–292.
[16] Earnshaw, W.C. (1987) J. Cell Biol. 105, 1479–1482.
[17] Jensen, K.J., Ohmstede, C.A., Fisher, R.S. and Sahyoun, N. (1991) Proc. Natl. Acad. Sci. USA 88, 2850–2853.
[18] Rogers, S., Wells, R. and Rechstein, M. (1986) Science 234, 364–368.
[19] Hanks, S.K., Quinn, A.M. and Hunter, T. (1988) Science 241, 42–52.
[20] Ono, T., Koide, Y., Arai, Y. and Yamashita, K. (1984) J. Biol. Chem. 259, 9011–9016.
[21] Ono, T., Slaughter, G.R., Cook, R.G. and Means, A.R. (1989) J. Biol. Chem. 264, 2081–2087.