

Molecular cloning of a brain-specific calcium/calmodulin-dependent protein kinase

CHIJEN R. LIN*[†], MICHAEL S. KAPILOFF*, SALLY DURGERIAN*, KAZUHIKO TATEMOTO[‡],
ANDREW F. RUSSO*, PHYLLIS HANSON[§], HOWARD SCHULMAN[§], AND MICHAEL G. ROSENFELD*[¶]

[†]Howard Hughes Medical Institute, *Eukaryotic Regulatory Biology Program, School of Medicine, M-013, University of California, San Diego, La Jolla, CA 92093; and Departments of [‡]Psychiatry and Behavioral Sciences and [§]Pharmacology, Stanford University School of Medicine, Stanford, CA 94305

Communicated by Daniel Steinberg, April 20, 1987

ABSTRACT A calcium/calmodulin-dependent protein kinase type II (CaM-K) α -subunit cDNA has been cloned from rat brain. This enzyme is encoded by a 5.1-kilobase mRNA expressed exclusively in the brain. Hybridization histochemistry reveals that the CaM-K mRNA expression corresponds to the distribution of the immunoreactive α -subunit protein, suggesting that the high enzyme levels in specific brain areas reflect regional differences in gene expression. The sequence of CaM-K α -subunit cDNA indicates a 478-amino acid (54-kDa) protein with three functional domains. The domain organization suggests a structural model for calcium/calmodulin-dependent and independent states that might subservise short- and long-term responses to transient stimuli.

Neuronal phenotype and function are controlled by extracellular stimuli, which activate diverse intracellular effectors, including protein kinases (1, 2). Because of the important role of calcium in regulating neural function, it is likely that calcium/phospholipid-dependent and calcium/calmodulin-dependent protein kinases have assumed critical regulatory functions (1). One isozyme of the calcium/calmodulin-dependent protein kinase type II (CaM-K) is a serine and threonine kinase that is expressed specifically in the brain (3). The ability of activated CaM-K to sustain calcium/calmodulin-independent autophosphorylation raises questions regarding its potential role in short- and long-term events characteristic of neurons (4). The role of covalent modification of proteins in complex events such as memory has been the subject of speculation (5, 6).

CaM-K is a holoenzyme of over 500 kDa composed of "51"-kDa (CaM-K- α) and "60"-kDa (CaM-K- β) structurally distinct subunits in various ratios (7-11). CaM-K- α appears to be exclusively expressed in neural tissues at high intracellular concentrations, constituting, for example, about 2% of all protein in the hippocampus (3). The enzyme appears to be distributed in several cellular compartments, including the cytosolic surface of synaptic vesicles, but is most concentrated in the postsynaptic density (12). CaM-K substrates include the neuronal cytoskeletal proteins synapsin I and microtubule-associated protein 2 (MAP-2), as well as tyrosine hydroxylase, τ , tryptophan hydroxylase, and myelin basic protein (4). CaM-K enzymatic activities in nonneuronal tissues appear to be isozymes with different subunit size and composition (3, 13).

In this manuscript we report the molecular cloning of DNA complementary to mRNA encoding CaM-K- α from rat brain.[¶] The predicted primary structure of CaM-K- α suggests a model for regulation of CaM-K function.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

METHODS

Protein Sequencing. CaM-K was purified to apparent homogeneity from rat forebrain as previously described (14). Tryptic fragments were separated by using reverse-phase high-pressure liquid chromatography, consisting of a μ Bondapak C₁₈ column (3.8 \times 300 mm) with a gradient of acetonitrile in 0.1% aqueous trifluoroacetic acid. The amino acid sequences of the peptides forming the major 215-nm absorbance peaks were determined by a gas-phase sequencer (Applied Biosystems, Foster City, CA) by automated serial Edman degradation.

Cloning of CaM-K cDNA. cDNA complementary to rat brain poly(A)-containing RNA [poly(A)-selected RNA] was generated by using reverse transcriptase from avian myeloblastosis virus or Moloney murine leukemia virus and cloned in λ gt10 or λ gt11 arms as previously described (15) or by using a variant procedure (16). On the basis of codon usage the following mixed- and unique-sequence oligonucleotides were constructed: 5'-TGG CA(A or G) AT(A, T, or C) GTG CA(C or T) TT(C or T) CA(C or T) AA-3', corresponding to P₁₄; and 5'-GA(A or G) TGG GA(C or T) ACG GT(A, G, C, or T) AC(A, G, C, or T) CCG GA(A or G) GC-3' and (antisense) 5'-CTT GGC CTC AGG GGT CAC TGT GTC CCA CTC AGG GGA GGG GAA GTC ATA GGC GCC AGC-3', corresponding to P₆. Subsequently, 5'-CTT GGC TTC CGG GGT GAC GGT GAC CCA TTC TGG TGA TGG GAA ATC-3' was constructed. Screening with kinase-treated oligonucleotides was performed with washing at 40°C to 50°C in 2 \times SSC (1 \times SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7.4). Inserts were subcloned in phage M13-based vectors for nick-translation (15) or sequencing by the enzymatic dideoxynucleotide method of Sanger *et al.* (17) with deazaguanosine.

RNA and DNA Analysis. After size fractionation, hybridization analysis was performed on poly(A)-selected RNA and restriction digests of rat brain genomic DNA, as previously described (15, 18). ³²P-labeled probes used were nick-translated C3 cDNA (2 \times 10⁸ cpm/ μ g) which corresponds to nucleotides -41 to 943, and 3'-C31 complementary RNA (cRNA) "run-off" transcripts (1.4 \times 10⁹ cpm/ μ g), which correspond to nucleotides 1077 to 1471 (19). Washing was performed in 0.2 \times SSC/0.2% NaDodSO₄ at 65°C. For *in situ* hybridization histochemistry a male Sprague-Dawley rat (475 g) was perfused with a formaldehyde buffer, and the brain was postfixed overnight, without glutaraldehyde, as previously

Abbreviations: CaM-K, the multifunctional calcium/calmodulin-dependent protein kinase type II; CaM-K- α , the α subunit of CaM-K; CaM-K- β , the β subunit of CaM-K; MLCK, myosin light chain kinase.

[†]To whom reprint requests should be addressed.

[¶]This sequence is being deposited in the EMBL/GenBank data base (Bolt, Beranek, and Newman Laboratories, Cambridge, MA, and Eur. Mol. Biol. Lab., Heidelberg) (accession no. J02942).

described (20). A 1 in 10 series of frozen sections (30 μm) was hybridized to the 3'-C31 [^{32}P]cRNA (10^8 cpm/ml) as previously described (19), washing with $0.1 \times \text{SSC}$ at 55°C .

RESULTS

Cloning of CaM-K. CaM-K was purified to apparent homogeneity from rat forebrain (14). Tryptic fragments were resolved by using high-performance liquid chromatography and subjected to microsequencing analysis (Fig. 1A). cDNA libraries from rat cortex or rat brain, excluding cerebellum, were screened with synthetic oligonucleotides whose sequences were based on the sequences of two peptides. A total of 58 reactive recombinants were identified from about 2×10^6 plaques with the oligonucleotides and subsequent nick-translated probes. Six clones were subjected to DNA sequencing; the sequences of four clones permitted the deduction of the entire coding sequence for CaM-K- α (Fig. 1B). An open reading frame composed of 1434 nucleotides indicated a 478-amino acid protein (Fig. 2A). The initiator methionine was identified by the presence of a preceding stop codon. Distributed throughout the predicted protein of actual M_r 54,111 were sequences corresponding to 15 tryptic peptides (Fig. 1A and 2A). A cell-free translation product of an *in vitro* RNA transcript encompassing the entire coding region is a 54-kDa protein that is specifically immunoprecipitated by anti-CaM-K- α -specific sera (data not shown).

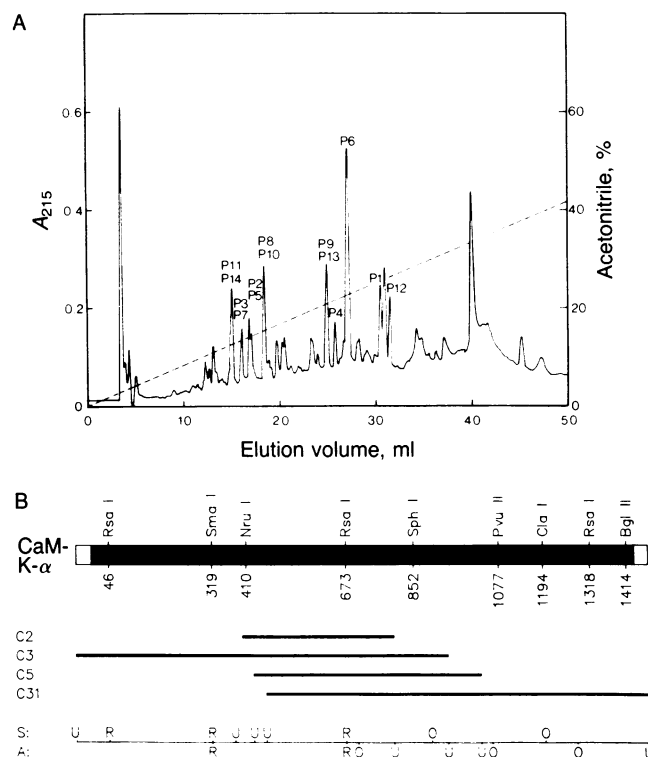


FIG. 1. Cloning of CaM-K- α . (A) Fourteen tryptic fragments of purified CaM-K were fractionated by HPLC. Sequences of these peptides, in the standard one-letter code, are as follows: P₁, FTEYQLFE; P₂, VLAGQEYAAK; P₃, HPNIVR; P₄, DLKPENLLA; P₅, LYQIQK; P₆, AGAYNFPSPEWDTVTPEAK; P₇, DLINK; P₈, ITAAEALK; P₉, ITAAEALKHPWI; P₁₀, HPWISHR; P₁₁, KQIQK; P₁₂, FYFENLWS; P₁₃, VWHR; P₁₄, WQIVHFHR. An additional peptide was found on a separate column fractionation: SGAPSV. (B) Representation of four of the six clones analyzed to confirm the sequence. DNA sequencing was performed by utilizing the "universal primer" oligonucleotide (U) to sequence the borders of the insert or ends created by restriction enzyme digestion (R) or by utilizing synthetic oligonucleotides (17-mer) corresponding to information from the clones (O). Clones were sequenced on both sense (S) and antisense (A) strands.

Analysis of CaM-K- α mRNA Expression. A probe corresponding to nucleotides -41 to 943 of CaM-K- α cDNA (amino acids 1-314) hybridized to a predominant 5.1-kilobase (kb) brain mRNA, which appears to represent the CaM-K- α mRNA. A 3.9- to 4.2-kb RNA hybridized at <3% of the signal strength of the 5.1-kb RNA (Fig. 3A), and other RNAs, 7.1, 6.1, and 3.3 kb in size, were detected in low variable levels. In human fetal brain, the migration of the dominant species is coincident with the rat brain 7.1-kb species. The ratio of signals from the 5.1- and 3.9-kb species in cerebellum was about 1.5:1. When a riboprobe corresponding to nucleotides 1077-1471 (amino acids 360-478) was used, the hybridization to the 3.9-kb RNA relative to the 5.1-kb RNA was greatly enhanced (Fig. 3B); it is suggested that the 3.9-kb RNA encodes the CaM-K- β (60-kDa) protein. In most rat tissues no RNA species hybridizing to the CaM-K- α probe was detected, including liver, kidney, spleen, and heart (Fig. 3A) and lung and adrenal gland (not shown); even if expressed at levels <1000-fold the level in the brain, CaM-K- α mRNA would have been detectable under these conditions. In pituitary (not shown) and skeletal muscle, a faintly hybridizing 3.9-kb RNA species comigrated with the putative brain CaM-K- β mRNA. In addition, 4.1-, 4.5-, and 4.9-kb homologous RNAs were detected in pituitary, skeletal muscle, and testes, respectively (Fig. 3A). A number of cell lines of different origin (Fig. 3A), including rat adrenal medulla (PC12), rat thyroid C cells (CA 1037), and mouse lymphocytes (A20) express reactive 3.9- to 4.5-kb RNA species, but no 5.1-kb CaM-K- α mRNA. Restriction analysis of rat genomic DNA fragments by using the C3 probe, which includes the kinase and regulatory domains of CaM-K- α (see below and Fig. 5A), suggests that these domains alone span >15 kb of genomic information (Fig. 3C).

Localization of CaM-K mRNA in the Brain. Immunohistochemistry suggests a striking regional distribution of CaM-K, with highest levels in hippocampal pyramidal cells and cortex. Because of the high concentration in postsynaptic densities in the neuropil associated with these regions, the quantitative distribution could reflect either differential stability of the sequestered protein or regional variation in gene expression. To distinguish between these possibilities, hybridization histochemical analysis of a series of coronal rat brain sections was performed; several sections are shown in Fig. 4. In the forebrain, the hybridization was most predominant within the hippocampus and the superficial layers of the cortex. There was also considerable hybridization to the caudate-putamen, amygdaloid complex, thalamus, and hypothalamus. Relatively light hybridization was detected in the colliculi and several brainstem nuclei. Within the cerebellum there was hybridization in the Purkinje cell layer and, to a lesser extent, in the molecular and granular layers. In the spinal cord, light hybridization was detectable only in the superficial layers of the dorsal horns. The quantitative pattern of CaM-K RNA is in agreement with the immunohistochemical analysis of CaM-K protein (3, 31), implying that regulation of CaM-K gene expression in the brain determines the extremely high level of the enzyme in the hippocampus and cortex.

DISCUSSION

The cloning of the CaM-K- α cDNA has allowed us to predict a 54-kDa protein that encodes a protein kinase and contains three apparent domains. The NH₂-terminal domain is highly homologous to other known protein kinases, but it is most closely related to the γ subunit of phosphorylase *b* kinase (PbK- γ) and myosin light chain kinase (MLCK). Conserved regions are diagrammed in Fig. 2B. Two interesting sequence variations occur in this protein. First, the third Gly in the ATP-binding site sequence (2) is replaced by Ser in CaM-K- α and PbK- γ (Fig. 2B, block I). Likewise, the so-called "APE box" (Fig. 2B, block XII) is uniquely Ser-Pro-Glu in CaM-

A

```

-41      A GTCCGAGCC TAAAGCTCC CCGCTGCC CAGTCCAGG ATG GCT ACC ATC ACC TGC ACC CGA TTC ACG
1        Met Ala Thr Ile Thr Cys Thr Arg Phe Thr
31      GAA GAG TAC CAG CTC TTC GAG GAA CTG CGA AAG GGA GGC TTC TCC GTG GTG CGC AGG TGT GTC AAG GTG CTG GCT
11      Glu Glu Tyr Gln Leu Phe Glu Glu Leu Gly Lys Gly Ala Phe Ser Val Val Arg Arg Cys Val Lys Val Leu Ala
106     GGC CAG GAG TAT GCT GCA AAG ATT ATC AAC ACC AAG AAG CTC TCA GCC AGA GAT CAC CAG AAG TTG GAA CGC GAG
36      Gly Gln Glu Tyr Ala Ala Lys Ile Asn Thr Lys Lys Leu Ser Ala Arg Asp His Gln Lys Leu Glu Arg Glu
181     GCC CGC ATC TGC CGC TTG TTG AAG CAC CCC AAT ATC GTC CGA CTC GAT GAG AGC ATC TCC GAG GAG GGC CAC CAC
61      Ala Arg Ile Cys Arg Leu Leu Lys His Asn Ile Val Arg Leu His Asp Ser Ile Ser Glu Glu Gly His His
256     TAC CTT ATC TTC GAT CTG GTC ACT GGT GGG GAG CTG TTC GAA GAC ATT GTG GCC GGG GAG TAT TAC AGT GAG GCT
86      Tyr Leu Ile Phe Asp Leu Val Thr Gly Gly Glu Leu Phe Glu Asp Ile Val Ala Arg Glu Tyr Tyr Ser Glu Ala
331     GAT GCC AGC CAC TGT ATC CAG CAG ATC CTG GAG GGT GTG CTA CAC TGT CAG CAG ATG GGG GTG CTG CAT CGC GAC
111     Asp Ala Ser His Cys Ile Gln Gln Ile Leu Glu Ala Val Leu His Cys His Gln Met Gly Val Val His Arg Asp
406     CTC AAC CCT GAG AAT CTG TTG CTG GCT TCC AAG CTC AAG GGT GCT GCG GTG AAG CTG CCA GAC TTT GGC CTG GCC
136     Leu Lys Pro Glu Asn Leu Leu Ala Ser Lys Leu Lys Gly Ala Ala Val Lys Leu Ala Asp Phe Gly Leu Ala
481     ATA GAG GTT GAG GGA GAG CAG GCA TGG TTT GGG TTC GCA GGG ACA CCT GGA TAC CTC TCC CCA GAA GTG CTG
161     Ile Glu Val Glu Glu Gln Gln Ala Trp Phe Gly Phe Ala Gly Thr Pro Gly Tyr Leu Ser Pro Glu Val Leu
556     CGG AAG GAC CCA TAC GGG AAG CCT GTG GAC CTG TGG GCC TGT GGC GTC ATC CTG TAT ATC TTG CTG GTT GGG TAT
186     Arg Lys Asp Pro Tyr Gly Lys Pro Val Asp Leu Trp Ala Cys Gly Val Ile Leu Leu Val Gly Tyr
631     CCC CGA TTC TGG GAT GAG CAG CAG CAC GCG CTG TAC CAG CAG ATC AAA GCT GCT GCC TAC GAT TTC CCA TCA CCA
211     Pro Pro Phe Trp Asp Glu Asp Gln His Arg Leu Tyr Gln Gln Ile Lys Ala Gly Ala Tyr Asp Phe Pro Ser Pro
706     GAA TGG GAC ACC GTC ACC CCG GAA GCC AAG GAT CTG ATC AAT AAG ATG CTG ACC ATC AAC CCG TCC AAA CGC ATC
236     Glu Trp Asp Thr Val Thr Pro Glu Ala Lys Asp Leu Ile Asn Lys Met Leu Thr Ile Asn Pro Ser Lys Arg Ile
781     ACG GCC GGT GAG GCT CTC AAG CAC CCC TGG ATC TCG CAC GCG TCC ACT GTG GCC TCC TGC ATG CAC AGA CAG GAG
261     Thr Ala Ala Glu Ala Leu Lys His Pro Trp Ile Ser His Arg Ser Thr Lys Ala Ala Ser Cys Met His Arg Gln Glu
856     ACC CTG GAC TCC CTG AAG AAG TTC AAT GCC AGG AGG AAA CTG AAG GGA GGC ATC CTC ACC ACT ATG CTG GCC ACC
286     Thr Val Asp Cys Leu Lys Lys Phe Asn Ala Arg Arg Lys Leu Lys Gly Ala Ile Leu Thr Met Leu Ala Thr
931     AGC AAC TTC TCC GGA GGG AAG ACT GGA GGA AAC AAG AAG AAT GAT GGC GTG AAG GAA TCC TCT GAG AGC ACC AAC
311     Arg Asn Phe Ser Gly Gly Lys Ser Gly Gly Asn Lys Lys Asn Asp Gly Val Lys Glu Ser Ser Thr Thr Asn
1006    ACC ACC ATC GAG GAT GAA CAG ACC AAA GTG CGC AAA CAG GAA ATT ATC AAA GTG ACA GAG CAG CTG ATC GAA GCC
336    Thr Thr Ile Glu Asp Glu Asp Thr Lys Val Arg Lys Gln Glu Ile Ile Lys Val Thr Glu Gln Leu Ile Glu Ala
1081    ATA AGC AAT GGA GAC TTT GAA TCC TAC ACG AAG ATG TCG CAC CCT GCA ATG ACA GCC TTT GAA CCG GAG GCC CTG
361    Ile Ser Asn Gly Asp Phe Thr Thr Lys Thr Met Cys Asp Pro Gly Met Thr Ala Phe Glu Pro Glu Ala Leu
1156    GGG AAC CTG GTC GAG GGC CTG GAC TTT CAT CGA TTC TAT TTT GAA AAC CTG TGG TCC CGG AAC ACC AAG CCC GTG
386    Gly Asn Leu Val Glu Gly Leu Asp Phe His Arg Phe Tyr Phe Glu Asn Leu Trp Ser Arg Asn Ser Lys Pro Val
1231    CAC ACC ACC ATC CTG AAC CCT CAC ATC CAC CTG ATG GGT GAC GAG TCA GCC TCC ATC GCC TAC ATC CGC ATC ACT
411    His Thr Thr Ile Leu Asn Pro His Ile His Leu Met Gly Asp Glu Ser Ala Cys Ile Ala Tyr Ile Arg Ile Thr
1306    CAG TAC CTG GAT GCG GGT GGC ATC CCC CGC ACG GGC CAG TCA GAG GAG ACC CGT GTC TGG CAC GCG AGG GAT GGA
436    Gln Tyr Leu Asp Ala Gly Gly Ile Pro Arg Thr Ala Gln Ser Glu Glu Thr Arg Val Trp His Arg Arg Asp Gly
1381    AAA TGG CAG ATC GTC CAC TTC CAG ACA TCT GGG GGG CCC TCC GTC CTG CCC CAT TGA AGC ACCAGGCCAG
461    Lys Trp Gln Ile Val His Phe His Arg Ser Gly Ala Pro Ser Val Leu Pro His 478
1451    GGTCCCTGGC CTCTGCTTC G 1471
    
```

B

	I	II	III	IV	V	VI	VII	
CaM-K	8--FTTEYQLPELCKGAFSVVRRG--7--EYAKII--2--KLSLA--5--LEREAR--6--HFW IVRLHD--7--HYLLFDL--2--CGELPEDI--							
PbK	14--FYENTPEKRIELGRGVSSVRRG--7--EYAVKII--3--GFSFA--4--ELRAT--12--HFW IQLKQ--7--FPIVFDL--2--KGLPFDVL--							
MLCK-G	>600--DVNIEKERLCSGKFGQVRL--7--VIAKFF KAYSA--7--DEISIM--4--HPK LWQVD--7--IVVLEM--2--CGELPEDI--							
MLCK-S	>300--KALGGCGKFGVCTG--7--KLAAKVI KQTF--7--LRIIVM--4--HBM LIQ Y--8--IVLMEY--2--CGELPEDI--							
cAMP-K	48--LWCGDFRW--10--HYAKIL--13--TLNKR--21--LDMVEY--2--CGENPESH--							
cGMP-K	364--LWCGDFRW--11--TFANKIL--13--IRSEKQ--5--HSDIIVRL Y--8--LTMLEA--2--CGELPEDI--							
PKC-I	16--LWCGDFRW--10--LVAIKIL--13--TLNKR--11--RHFLLTQMS--7--LVEYMEY--2--GGDLTYHI--							
v-raf	29--LWCGDFRW--10--DVAVKIL--12--FRNEVA--6--HYVLLFMYC--5--LAIVTQW--2--GSLYKHL--							
EGF-R	717--LWCGDFRW--14--PVAIKEL--11--ILDRAV--8--HYCRLLGICL--3--VQLITQL--2--FGCLLDVY--							
v-src	272--LWCGDFRW--9--RVAIKTL--9--FLQAAQ--6--HEK LKQL Y--7--IYVIEY--2--KGSLLDFL--							
	VIII	IX	X	XI	XII	XIII	XIV	
CaM-K	16--QLKAVLHCH--4--VERDLKPELLL--9--KLADFG--16--GTFYLSPEVL--3--PTGKFDVLA--1--GVILYLIL VGYFFF--							
PbK	16--ALLEVICALE--4--VERDLKPELLL--6--KLTDFG--15--GTFYLSAPEII--9--CFGKVDVMS--1--GVDTITLL AGSFFF--							
MLCK-G	17--QSGEVYIIE--4--LELDLAPENIC--8--KLIDFG--15--GTFEFLAPEVI--3--PIGTYETMS--1--GVICITLV SGLSFFM--							
MLCK-S	17--QICDGLFGR--4--LELDLAPENIC--8--KLIDFG--15--GTFEFLAPEVI--4--ISDK TMMS--1--GVITITML SGLSFFL--							
cAMP-K	16--QVILTFEYLE--4--IYRDLKPELLL--6--QVTDFG--13--GTFEFLAPEII--3--GVNKAVDWA--1--GVLIYEMA AGYFFF--							
cGMP-K	16--CVVEAFAYLE--4--IYRDLKPELLL--6--KLVDGF--15--GTFEFLAPEII--3--GHDISADVS--1--GILMYELL TGSFFF--							
PKC-I	16--EIAIGLFFLE--4--IYRDLKPELLL--6--KLTDFG--16--GTFDYIAPEII--3--PTGKVDVMS--1--GVLIYELM AQGFFF--							
v-raf	17--QTAQHDYLE--4--IBRDMESNIFL--6--KLIDFG--18--GSVLWMAPEVI--6--PFSQSDVVS--1--GVLVYELM AGELFYA--							
EGF-R	17--QIAGHDYLE--4--VERDLAARNVLL--6--KLTDFG--18--VPIKVALESI--3--IYTHQSDVVS--1--GVTVELMTGSKPYD--							
v-src	18--QIASGHAYVE--4--VERDLAARNVLL--6--KVADFG--2--RLIEDNEY--7--FPIKVAPEAA--3--RFTKSDVVS--1--GILITELTKGRVFPY--							
	XV	XVI	XVII			XVIII		
CaM-K	13--GAYDFPSPEVD--2--TPKADL--8--PSKRITAAEAALKHP--21--KFF NARR ELKGAITLTLATRNFSG--163							
PbK	13--GNYQFQSPVD--2--SDTVKDL--8--PKRYTAAEAALAP--20--RIYQYRNVKVFTRIVRDPALRFL--43							
MLCK-G	13--ATWDFDFAFD--2--SDDAKDF--8--MKSRLNCTQCLQRP--19--KKY NARR KWKQTKGAVTAIGRLSSM--156							
MLCK-S	13--GNYVDFEETPE--2--SDEAKDF--8--QGAHNSAAQCLARP--22--KKYL KR RWKNTIATVSAANRFKXI--7							
cAMP-K	23--SDI EDL--82							
cGMP-K	112							
PKC-I	22--SRLVAEI--112							
v-raf	17--SFDL--71							
EGF-R	30--SRP KFR--248							
v-src	13--GTRMPCPEEC FESLADL--32							

FIG. 2. Primary structure of CaM-K- α . (A) The nucleotide sequence encompassing the coding region of CaM-K- α and the predicted encoded protein. The 15 peptides listed in the legend to Fig. 1A are present in numerical order in the predicted sequence and are underlined. (B) Homology between CaM-K- α and other protein kinases, including the following: PbK, γ subunit of rabbit phosphorylase *b* kinase (21); MLCK-G, chicken gizzard myosin light chain kinase (22); MLCK-S, rabbit skeletal muscle myosin light chain kinase (23); cAMP-K, catalytic subunit of bovine cardiac muscle cAMP-dependent protein kinase (24); cGMP-K, catalytic subunit of bovine lung cGMP-dependent protein kinase (25-27); PKC-I, rat brain protein kinase C type I (28); v-raf, murine sarcoma virus v-raf (29); EGF-R, human epidermal growth factor receptor (15, 30); v-src, Rous sarcoma virus strain Prague C v-src [data of D. Schwartz, R. Tizard, and W. Gilbert in EMBL/GenBank Genetic Sequence Database (1982) GenBank (Bolt, Beranek and Newman Laboratories, Cambridge, MA), Tape Release 47]. Numbers indicate the number of intervening residues not listed due to lack of obvious homology, and blank spaces indicate an artificial gap created in alignment. Identity between CaM-K- α and PbK- γ or between CaM-K- α and any two other listed kinases is highlighted by bold print.

K- α and rabbit skeletal muscle MLCK. A series of homologies are found within the protein kinase domain (amino acid residues 1 to 220-250, Fig. 2B, blocks I to XIV-XVI). Amino acid residues 250 to 320-350 appear to reflect a second domain, unique to the calcium/calmodulin-dependent protein kinases. This domain includes a basic amino acid-rich sequence, *Lys-Lys-Phe-Asn-Ala-Arg-Arg-Lys-Leu-Lys* (Fig. 2B, block XVIII), which is postulated to represent the calmodulin-binding site. In support of this assignment, proteolytic fragments of PbK- γ and MLCK containing portions of block XVIII have been reported to bind calmodulin (32-34). Synthetic peptides corresponding to this domain of CaM-K- α bind calmodulin and inhibit kinase activity. Finally, the 163-residue COOH-terminal region, lacking homology to other protein kinases, is highly hydrophilic and contains two areas of high charge density (Fig. 5A). The function of

this region is unknown, but it could be important in the assembly of subunits of holoenzyme or in the association of the holoenzyme with other proteins or cellular organelles. The three-domain structure of the protein suggests a model for the mechanism of the known regulation of CaM-K enzyme (Fig. 5). In the inactive state the regulatory region of each subunit may be folded into a closed configuration, with the calmodulin-binding site sterically blocking the entry of substrates into the active site (Fig. 5B), analogous to similar events in MLCK (34). Binding of calcium/calmodulin displaces the inhibitory domain and activates the enzyme. Kinetic studies have shown that autophosphorylation at one or two threonine residues (L. L. Lou and H.S., unpublished data) converts the holoenzyme to an autonomous state (4, 36). Because limited proteolysis of CaM-K releases a monomeric 30-kDa autonomous enzyme devoid of autophos-

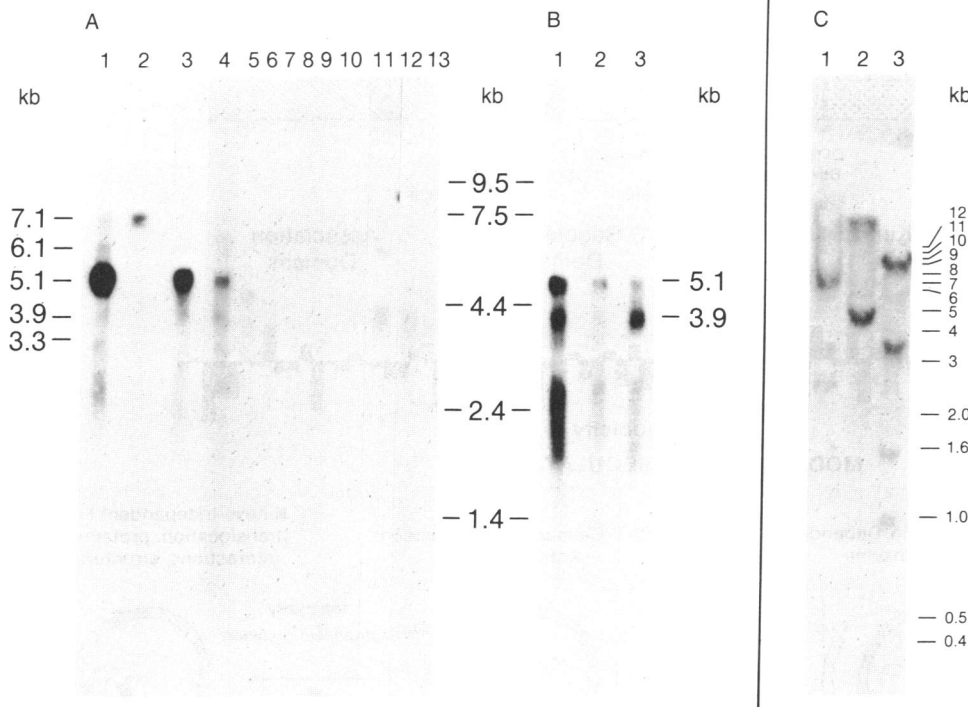


FIG. 3. RNA and genomic DNA hybridization to CaM-K- α cDNA. (A and B) Poly(A)-selected RNA prepared from various tissues (5 μ g from rat, unless otherwise specified, size-fractionated, transferred to nitrocellulose, and hybridized to 32 P-labeled probes. In A, the probe was C3 cDNA. Autoradiographs were exposed for 10 hr (lanes 1–3) or 24 hr (lanes 4–13). Lanes: 1, temporal lobe; 2, human fetal brain; 3, brain without cerebellum; 4, cerebellum; 5, skeletal muscle; 6, testes; 7, kidney; 8, liver; 9, spleen; 10, heart; 11, adrenal medullary cell line PC12; 12, thyroid C cell line CA 1037; 13, mouse lymphoma cell line A20. In B, the probe was a 3'-C31 cRNA. Lanes: 1, brain without cerebellum (5-hr exposure); 2, brain without cerebellum (1 μ g, 5 hr); 3, cerebellum (20 hr). Migration of RNA standards (Bethesda Research Laboratories RNA Ladder) is indicated. (C) Rat brain genomic DNA was digested to completion with *Bam*HI (lane 1), *Eco*RI (lane 2), or *Hind*III (lane 3), size fractionated and hybridized with C3 nick-translated probe. Migration of DNA standards (Bethesda Research Laboratories 1-kb DNA Ladder) is indicated.

phorylation sites (37), it is likely that these sites are in the regulatory or association domains. The four sequences most homologous to the consensus sequences for cAMP-dependent protein kinase and CaM-K substrates (4) are marked with asterisks in Fig. 5A and are found clustered throughout the "hinge" region and calmodulin-binding site. Phosphorylation of one or more of these residues may be the modification necessary to sustain the open configuration required for autonomy of CaM-K, potentially permitting short-term enhancement of responses to transient calcium signals (4, 6).

Autonomous autophosphorylation at a site(s) distinct from the aforementioned threonine residues results in the inhibition of enzyme activity for exogenous substrates (4, 5). This inhibitory autophosphorylation may occur within the calmodulin-binding site at the serine residue marked by an asterisk (Fig. 2B, block XVIII, and Fig. 5A), potentially inducing a conformational alteration of the COOH-terminal domain. Phosphorylation of gizzard MLCK by cAMP-dependent protein kinase at a comparable serine residue within the calmodulin-binding site leads to enzyme inhibition (32).

Autophosphorylation of CaM-K may also result in translocation of the holoenzyme to other areas of the neuron (38), altered protein-protein interactions, or activation of struc-

tural function(s) of the protein. The high concentration of such an autophosphorylated or translocated protein in neurons provides a potential facilitatory mechanism for short-term and long-term neural response to stimuli. The cloning of CaM-K- α cDNA permits a structure-function analysis of its potential biological functions.

Note Added in Proof. Comparison of the sequence of CaM-K- α with the recently published CaM-K- β sequence (39) suggests that they are the products of two highly homologous transcription units, with amino acid insertions carboxyl-terminal to the calmodulin-binding domain accounting for the larger size of the β subunit.

We thank G. Makk, Harry Mangalam, and Stuart Leff; Jeff Arriza and Ronald Evans for a rat brain cDNA library; and Tony Means for sharing sequence information in the course of these studies. This work was supported by grants from the National Institutes of Health. M.S.K. is supported by the Medical Science Training Program of the National Institutes of Health, and A.F.R. by the Jane Coffin Childs Memorial Fund for Medical Research.

1. Nairn, A. C., Hemmings, H. C., Jr., & Greengard, P. (1985) *Annu. Rev. Biochem.* **54**, 931–976.
2. Hunter, T. & Cooper, J. A. (1985) *Annu. Rev. Biochem.* **54**, 897–930.
3. Erondy, N. E. & Kennedy, M. B. (1985) *J. Neurosci.* **5**, 3270–3277.

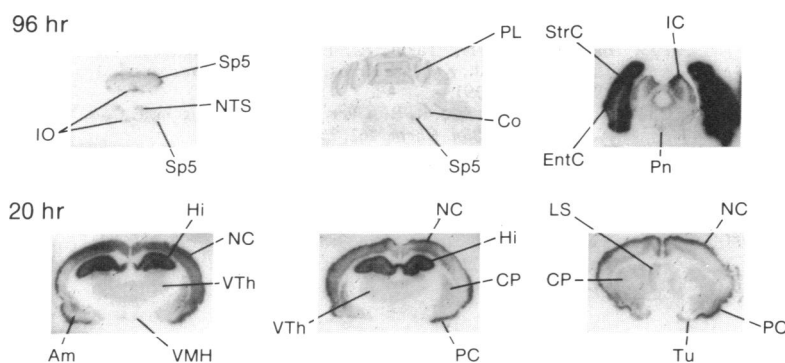


FIG. 4. *In situ* hybridization histochemistry of CaM-K RNA. Rat brain sections were hybridized to 32 P-labeled 3'-C31 riboprobe and exposed for 96 hr (Upper) or 20 hr (Lower) to Cronex 4 film without an intensifying screen. Brainstem sections (medulla, cerebellum, and midbrain, left to right) are in Upper, and forebrain sections (cortex, hippocampus, and neostriatum) are in Lower. Am, amygdaloid complex; Co, cochlear nucleus; CP, caudate-putamen; EntC, entorhinal cortex; Hi, hippocampus (including CA pyramidal cells and the dentate gyrus); IC, inferior colliculus; IO, inferior olive; LS, lateral septum; NC, neocortex (both frontoparietal-motor and frontoparietal-somato-sensory areas); NTS, nucleus tractus solitarius; PC, piriform cortex; PL, Purkinje layer of the cerebellum; Pn, pontine nuclei; Sp5, spinal tract of trigeminal nerve; StrC, striate cortex; Tu, olfactory tubercle; VMH, ventromedial hypothalamic nucleus; VTh, ventral thalamic nuclei (not exclusive).

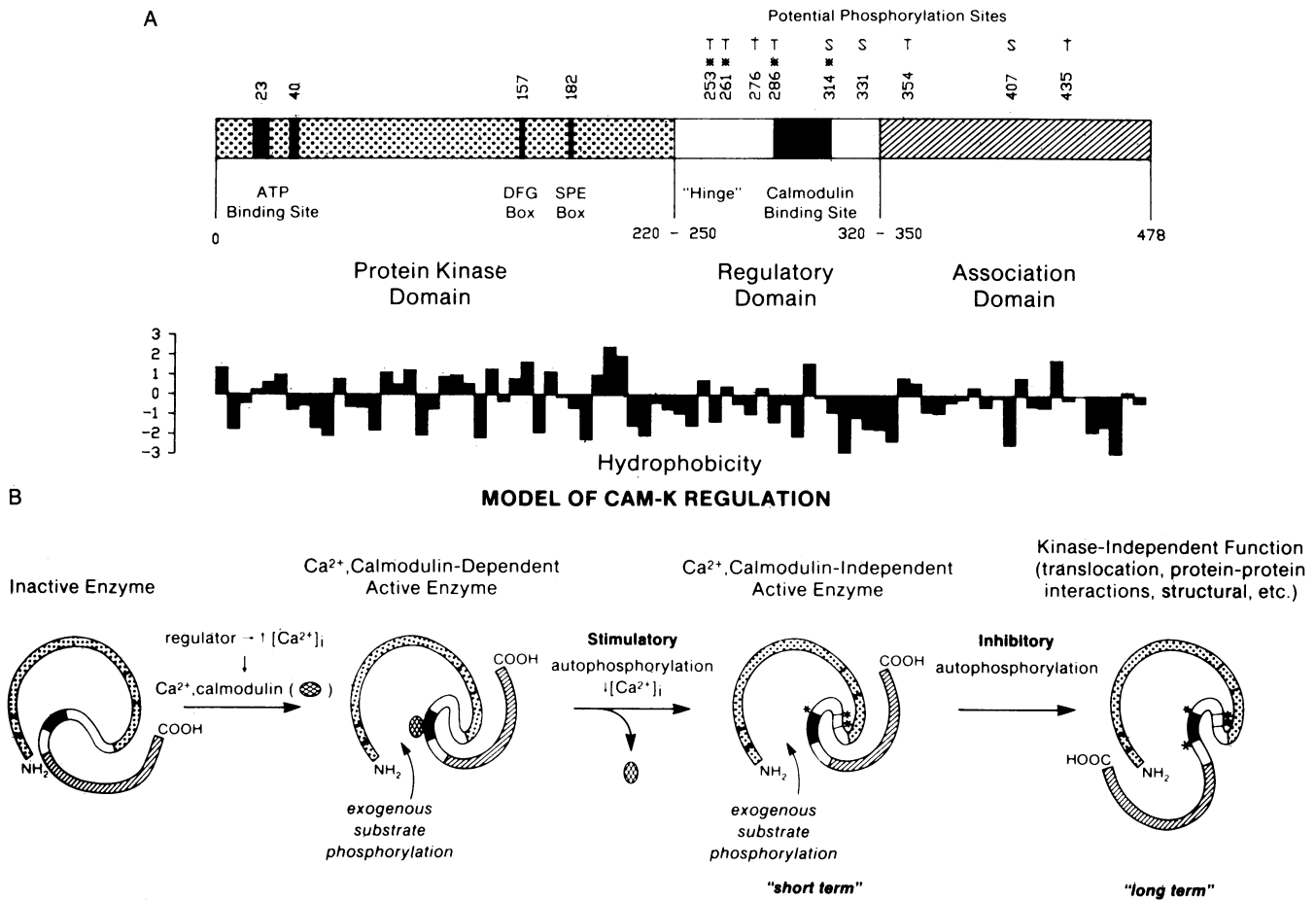


FIG. 5. Model of domain structure of Cam-K- α . (A) Potential serine (S) and threonine (T) sites for phosphorylation are noted; those with the greatest site homology are marked with an asterisk. The putative "hinge" region represents a region where the regulatory domain may bend so that further COOH-terminal regions may block the kinase active site (see Discussion). Below is a hydrophobicity plot drawn to scale and based on a window of six residues generated by the algorithm of Kyte and Doolittle (35). (B) A model for alternative conformations of Cam-K- α (see Discussion). Asterisks correspond to the threonine and serine sites shown in A.

- Schulman, H. (1987) *Adv. Cyclic Nucleotide Protein Phosphorylation Res.* 22, in press.
- Lou, L. L. & Schulman, H. (1986) *Adv. Cyclic Nucleotide Protein Phosphorylation Res.* 21, in press.
- Schwartz, J. H. & Greengard, S. M. (1987) *Annu. Rev. Neurosci.* 10, 459-476.
- Kuret, J. H. & Schulman, H. (1984) *Biochemistry* 23, 5495-5504.
- Bennett, M. K., Erondou, N. E. & Kennedy, M. B. (1983) *J. Biol. Chem.* 258, 12735-12744.
- Goldenring, J. R., Gonzalez, B., McGuire, J. S., Jr., & DeLorenzo, R. J. (1983) *J. Biol. Chem.* 258, 12632-12640.
- Miller, S. G. & Kennedy, M. B. (1985) *J. Biol. Chem.* 260, 9039-9046.
- McGuinness, T. L., Lai, Y. & Greengard, P. (1985) *J. Biol. Chem.* 260, 1696-1704.
- Kelly, P. T., McGuinness, T. L. & Greengard, P. (1984) *Proc. Natl. Acad. Sci. USA* 81, 945-949.
- Shenolikar, S., Lickteig, R., Hardie, D. G., Soderling, T. R., Hanley, R. M. & Kelly, P. T. (1986) *Eur. J. Biochem.* 161, 739-747.
- Schulman, H. (1984) *J. Cell Biol.* 99, 11-19.
- Lin, C. R., Chen, W. S., Kruiger, W., Stolarsky, L. S., Weber, W., Evans, R. M., Verma, I. M., Gill, G. N. & Rosenfeld, M. G. (1984) *Science* 224, 843-848.
- Gubler, U. & Hoffman, B. J. (1983) *Gene* 25, 263-269.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
- Southern, E. M. (1975) *J. Mol. Biol.* 98, 503-517.
- Crenshaw, E. B., Russo, A. F., Swanson, L. W. & Rosenfeld, M. G. (1987) *Cell*, in press.
- Swanson, L. W., Sawchenko, P. E., Rivier, J. & Vale, W. W. (1983) *Neuroendocrinology* 36, 165-186.
- Reimann, E. M., Titani, K., Ericsson, L. H., Wade, R. D., Fisher, E. H. & Walsh, K. A. (1984) *Biochemistry* 23, 4185-4192.
- Guerriero, V., Jr., Russo, M. A., Olson, N. J., Putkey, J. A. & Means, A. R. (1986) *Biochemistry* 25, 8372-8381.
- Takio, K., Blumenthal, D. K., Edelman, A. M., Walsh, K. A., Krebs, E. G. & Titani, K. (1985) *Biochemistry* 24, 6028-6037.
- Shoji, S., Ericsson, L. H., Walsh, K. A., Hischer, E. H. & Titani, K. (1983) *Biochemistry* 22, 3702-3709.
- Hashimoto, E., Takio, K. & Krebs, E. G. (1982) *J. Biol. Chem.* 257, 727-733.
- Takio, K., Smith, S. B., Walsh, K. A., Krebs, E. G. & Titani, K. (1983) *J. Biol. Chem.* 258, 5531-5536.
- Takio, K., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A. & Titani, K. (1984) *Biochemistry* 23, 4207-4218.
- Knopf, J. L., Lee, M. H., Sultzman, L. A., Kriz, R. W., Lloomis, C. R., Hewick, R. M. & Bell, R. M. (1986) *Cell* 46, 491-502.
- Kan, N. C., Flordellis, C. S., Mark, G. E., Duesberg, P. H. & Papas, T. S. (1984) *Science* 223, 813-816.
- Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Liberman, T. A., Schlessinger, J. & Seeburg, P. H. (1984) *Nature (London)* 309, 418-425.
- Ouimet, C. C., McGuinness, T. L. & Greengard, P. (1984) *Proc. Natl. Acad. Sci. USA* 81, 5604-5608.
- Lukas, T. J., Burgess, W. H., Prendergast, F. G., Lau, W. & Watterson, D. M. (1986) *Biochemistry* 25, 1458-1464.
- Blumenthal, D. K., Takio, K., Edelman, A. M., Charbonneau, H., Titani, K., Walsh, K. A. & Krebs, E. G. (1985) *Proc. Natl. Acad. Sci. USA* 82, 3187-3191.
- Kemp, B. E., Pearson, R. B., Guerriero, V., Jr., Bagchi, I. C. & Means, A. R. (1987) *J. Biol. Chem.* 262, 2542-2548.
- Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* 157, 105-132.
- Schworer, C. M., Colbran, R. J. & Soderling, T. R. (1986) *J. Biol. Chem.* 261, 8581-8584.
- LeVine, H. I. & Sahyoun, N. E. (1987) *Eur. J. Biochem.*, in press.
- Saitoh, T. & Schwartz, J. H. (1985) *J. Cell Biol.* 100, 835-842.
- Bennett, M. K. & Kennedy, M. B. (1987) *Proc. Natl. Acad. Sci. USA* 84, 1794-1798.