Summary of the PhD Thesis

In Silico, *In Vivo* and *In Vitro* Analysis of the Calmodulin Gene Expression in the Central Nervous System of the Rat

Előd Körtvély

Supervisor: Károly Gulya
Professor and Chairman

Department of Zoology and
Cell Biology
Faculty of Science
University of Szeged

Szeged
2004

Summary of Ph.D. Thesis

In silico, in vivo and in vitro analysis of the calmodulin gene expression in the central nervous system of the rat

Calcium (Ca^{2+}) is an element that is necessary for various biological functions. The information encoded into and by spatially and temporally distinct Ca^{2+} spikes is translated into cellular activity by Ca^{2+}-binding proteins. Some proteins are directly regulated by Ca^{2+}, while the majority are intermediaries between the Ca^{2+} signal and the effectors. The prototypical Ca^{2+} sensor, calmodulin (CaM), is a multifunctional, highly conserved Ca^{2+}-binding protein. This EF-hand protein is found in all eukaryotic cells. It regulates a number of basic cellular functions, including cell division, differentiation, signal transduction, etc., by interacting with a large and heterogeneous array of target proteins. It is particularly abundant in the nervous tissue. Although certain glial cells contain considerable levels of CaM, there is general agreement that the bulk of the protein is synthesized by neurons within the CNS, where it resides in different intracellular pools.

In higher vertebrates, multiple CaM genes encode a single, highly conserved CaM protein. Three bona fide CaM genes (CaM I, II and III) expressed at high levels in the brain have been described in the rat. These three CaM genes transcribe eight major mRNA species of different sizes by means of alternative polyadenylation, i.e. 4.2 kb, 4.1 kb, 1.7 kb and 1.0 kb for CaM I, 1.4 kb for CaM II, and 2.3 kb, 1.9 kb and 0.9 kb for CaM III. Although the coding regions of the transcripts diverge only within the limits of degeneracy of the genetic code, their 5'- and 3'-untranslated regions are markedly different, suggesting a differential regulation for the expression of the three CaM genes.

In recent decades, much effort has gone into describing the exact pattern of expression of the CaM genes. For example, their physiological and experimentally modified expressions have been analyzed in adult and embryonic tissues and in different in vitro systems. Additionally, stage-specific changes in gene expression intensities have been detected during development and differentiation. For example, marked differences in CaM mRNA content were found in different brain areas on postnatal day 1 (PD1), and the distribution became

more uniform with age. Further, a marked increase in cerebellar CaM gene expression was observed, where the hybridization signal was confined to the external germinal layer on PD1, but the expression was much stronger in the newly-formed granular and Purkinje cell layers by PD16. A pronounced increase in the 4.0 kb transcript during development was also detected, whereas the 1.8 kb message exhibited little change. The regulation of the three CaM genes by hormonal factors was also investigated. Interestingly, in several neoplastic tissues CaM II proved to be the primarily expressed gene. The levels of the mRNAs corresponding to the three CaM genes are described by a unique developmental pattern during spermatogenesis. A majority of these studies, however, either provided mainly qualitative analyses of the experimental data, or focussed only on the expression of a particular CaM gene. Although the members of the apparently redundant CaM gene family are believed to be regulated in different ways in order to establish the precise spatial and temporal control of the CaM protein level, the mechanisms underlying the alterations in the expression profiles of the CaM genes in response to various stimuli remain unclear.

In the rodent brain, the CaM mRNAs are largely distributed in the neuronal cell soma. Additionally, it is known that the CaM I mRNAs are transiently targeted into the dendritic compartment in cortical and hippocampal pyramidal and cerebellar Purkinje cells in the newborn brain. In another study, by means of quantitative, radioactive in situ hybridization our laboratory demonstrated that mRNAs transcribed from all three CaMs are significantly more abundant in dendrite-rich than in axon-rich regions of the rat brain. The mRNA concentrations were consistently different; in all the dendrite-rich areas, the highest level was that of the CaM I gene, and the lowest that of the CaM II gene. These results suggest that a significant amount of CaM mRNAs is translocated into dendrites in a gene-specific manner. Studies of neuronal cell cultures also indicated the existence of a dendritic CaM mRNA pool. For example, in differentiated PC12 cells, the CaM I and CaM II mRNAs displayed a marked translocation into the processes, while the CaM III mRNAs were restricted to the perinuclear region. The number of known dendritically targeted mRNA species in mammalian neurons is growing, and the evidence indicates that the mRNAs found in the dendrites are translated there; the cells can thereby alter the dendritic protein concentrations in response to various cues.

With the aim of a better understanding of the physiological significance of the CaM multigene family and the regulation of the expression of individual CaM genes, we have carried out a series of investigations. The specific aims of our investigations were as follows:

- Cloning of the cDNA region corresponding to the 3'-end of the 4.2 kb CaM I transcript in order to determine its nucleotide sequence.

- A computer analysis of the mRNAs transcribed from the three CaM genes in order to find possible regulatory elements affecting their stability, intracellular localization and translational efficiency.

- A gene-specific in vivo analysis of the distribution of CaM transcripts corresponding to the three CaM genes in the rat brain during early postnatal development. Quantitative determination of the changes in expression of the CaM genes in the time frame examined.

- An in vitro analysis of the CaM genes in primary hippocampal cell cultures, especially as concerns the intracellular targeting of their transcripts.

1. In silico analysis. A knowledge of the nucleotide sequences of the expressed CaM genes is a prerequisite for an understanding of the phenomena observed at a gene expression level. Most of the experiments addressing the CaM gene expression, however, were carried out in rodents, where information on these sequences was to some extent incomplete. Thus, as a first step, we completed identification of the cDNA sequence of the rat CaM I gene by using 3'-RACE methodology. Thereafter, the full-length cDNA sequences were compared against different databases, including those of the cis-acting regulatory elements (TransTerm, UTRdb). Although not coding for the actual protein sequence, the UTRs control diverse functions of eukaryotic mRNAs, such as their stability, translation efficiency, cytoplasmic localization and coding capacity. To date, most control elements have been identified in the 3'-UTR, although the regulatory role of the 5'-UTR and even the coding region has also been implicated.

We characterized the alternative polyadenylation of the different CaM transcripts, too. Our results indicate that the CaM transcripts often use noncanonical polyadenylation signals. Tissue-specific polyadenylation might well explain why the relative amounts of the different CaM transcripts vary from tissue to tissue. Alternatively, the individual CaM transcripts possess tissue-specific half-lives. The 13 AREs found in the 3'-UTR of the 4.2 kb CaM I transcripts favor this latter possibility. Additionally, certain AREs appear to be critical parts of some cytoplasmic mRNA localization signals, and therefore might explain what we also observed in our in vivo and in vitro experiments (see below). Moreover, the 3'-UTRs of a number of rat CaM transcripts contain CPEs, as do their human counterparts. These elements have been found to influence the efficiency of dendritic translation and thereby even synaptic plasticity. Translational control may operate in a poly(A) tail-dependent manner, the mechanism used by the CPEs too, or in a poly(A) tail-independent manner. The DICEs found in CaM I and CaM III transcripts are examples of this latter case, where they are recognized by trans-acting factors and the resulting complex then inhibits the initiation pathway of the translation. CaM mRNAs belong in the subset of mRNAs that are targeted to specific intracellular domains. The localization signals (zip codes) are bound by trans-acting factors to form a transport complex, which is then moved along the cytoskeleton. To date, few studies have been carried out to identify dendritic targeting elements in neurons. The 4.2 kb and the 1.7 kb rat CaM I transcripts contain a 70 nucleotide-long stretch highly similar to the zip code occurring in the β-actin and angiotensin II receptor mRNAs. Additionally, a 91% homologous element is present in the corresponding human sequences. Although its role has not yet been proved, this zip code might well be responsible for the prominent dendritic trafficking of the CaM I mRNAs. On the other hand, as described earlier, the mRNAs of the other two CaM genes are also targeted to dendrites, especially during early development. Since the only region highly conserved among these transcripts is the coding sequence, its role in targeting might be envisaged. Accordingly, the 70 nucleotide-long zip signal may act only in the adult, while other elements are responsible for the CaM mRNA localization in the developing brain.

2. In vivo analysis. To investigate whether differential expression takes place in the developing rat brain, a quantitative in situ hybridization analysis involving 15 brain areas at 6 ages between embryonic day 19 (E19) and postnatal day 20 (PD20) was carried out with gene-specific [^{35}S]cRNA probes. In a similar comparative study involving 70 brain areas in the adult rat brain, our laboratory earlier demonstrated the differential expression of the three CaM genes. From practical considerations, we selected 15 of these 70 brain areas for our current study, the first comprehensive and comparative description of the ontogeny of CaM transcription in the developing rat brain.

The widespread and differential expression of all CaM genes was found. The developmental stage-specific gene expression of the cells in the different brain areas obviously plays a crucial role in the establishment of the observed expression profiles. However, the enormous changes and rearrangements in the number of cells within the developing brain itself could be reflected in changes in mRNA levels.

The characteristic changes in the signal intensities in the examined time frame were utilized to classify the brain regions in three categories. While the three CaM genes exhibited a similar expression profile in a given brain area, the different brain areas frequently displayed different expressional profiles. In spite of the similar in-region expressional profiles of the three CaM genes, the magnitudes of the signals often varied significantly from gene to gene. For a majority of the areas (e.g. the piriform cortex for CaM III), the signal intensities peaked at around PD10 and the expression profile was symmetric (type 1). Other regions (e.g. the cerebral cortex, layer 1 for CaM II) displayed their highest signal intensities at the earliest age measured, followed by a gradual decrease (type 2). The signal intensities in the regions in the third group (e.g. the hypothalamus for CaM III) fluctuated from age to age (type 3). Marked CaM mRNA levels were measured for each transcript corresponding to the three CaM genes in the molecular layers of the cerebral and cerebellar cortices and hippocampus, suggesting their dendritic translocation. The highest signal intensity was measured for CaM II mRNA, followed by those for CaM III and CaM I mRNAs on PD1. However, the CaM II and CaM III mRNAs subsequently decreased steeply, while the CaM I mRNAs were readily detected even on PD20.

Although the CaM expression has been investigated intensively, few studies have related to its ontogeny. Previous studies dealt with the localization of the CaM protein in the brains of rodent embryos and neonates. From a comparison of these results with our findings, it seems very likely that CaM gene transcription appreciably precedes protein translation in several regions. However, because of the identity of the CaM proteins translated from different mRNAs, the gene-specific control in the translational regulation is a complex field to investigate. Methodological dissimilarities (e.g. the specificity of the probes, the quantitative assessment of the results, etc.) mean that the data obtained from earlier experiments are difficult to interpret in comparison with our gene-specific and more precisely quantified results. Nevertheless, we also found CaM signals in the regions of strongest expression (the piriform cortex and the hippocampus), peaking at around PD10. We detected roughly similar changes in CaM gene expression in the thalamic region and the striatum, as described earlier, although our study indicated a rather different transcriptional activity for the three CaM genes in the latter case. In the cerebellum, an increase was observed between PD1 and PD32. Our results, however, pointed to a more complex regulation of each CaM gene in the cerebellum, characterized by several rapid changes in signal intensities. In the rat, this region undergoes significant postnatal development, which might be the cause of the observed fluctuation in CaM gene expression.

It would be especially important to establish whether the changes in CaM expression coincide with those coding its major target proteins. The developmental expression of CaM-binding proteins (CaM-BPs) was also studied. It was found that the hippocampus exhibited the earliest onset of CaM-BP expression (PD5), followed by the cerebral cortex and the striatum, while the onset of CaM-BP expression occured latest in the cerebellum (PD9). Moreover, the profile of CaM-BP expression in the maturing cerebellum was different from that in other regions, i.e. different CaM-BPs predominate in the cerebellum and in other brain regions. We measured the highest CaM mRNA content at around PD5 and PD10 in the hippocampus and the cerebral cortex, which means that CaM and its major target proteins are expressed simultaneously in these areas. The apparently irregular expression profiles observed for each CaM gene in the cerebellum as a whole might be a result of a unique utilization of the CaM-BPs in this region. Additionally, the genes for the subunits of the

multimeric Ca^{2+}/CaM-dependent protein kinase II (CaM-KII), a major CaM target protein, have also been reported to be differentially regulated during development.

The main findings of the present study were as follows: 1) The three CaM genes exhibited a similar expression profile in a given brain area. 2) Typically, the expression profiles of areas rich in cell somata peaked at around PD10, when the most intense synaptogenesis takes place. 3) In spite of the general similarities in expression of the three CaM genes, significant differences emerged from the quantitative analysis. Their physiological role, however, is uncertain. 4) The marked mRNA levels corresponding to each CaM gene in the developing molecular layers may hint at their dendritic targeting, a phenomenon that decreases with age.

3. In vitro analysis. CaM is intimately involved in a variety of physiological processes. In neurons, this protein regulates, among others, proliferation, differentiation, synapse formation, the release of neurotransmitters and the microtubule function. Intricate mechanisms regulate the available CaM level from intracellular pools and presumably from local de novo protein synthesis, which includes the transport of CaM mRNAs to these sites. However, little straightforward evidence points to the dendritic localization of the CaM mRNAs in the CNS neurons. Although the presence of CaM I transcripts has been demonstrated by color in situ hybridization in apical dendrites of pyramidal and Purkinje cells in the postnatal rat brain, all other studies were carried out in peripheral neuronal cultures and/or by radioactive in situ hybridization analysis. This latter technique is inherently unsuitable for direct intracellular localization.

We carried out CaM gene-specific, nonradioactive in situ hybridization cytochemistry on rat primary hippocampal cultures. This system was selected for the following reasons: 1) unlike many immortalized cell lines, it may represent the nearest approximation to in vivo CNS neurons, 2) it mimics several in vivo characteristics of these cells (e.g. dendrites and axons are present, and synapse formation takes place), 3) its monolayer structure allows resolution of the subcellular compartments, 4) although optimized for the culturing of neurons, it contains both neuronal and glial cells, and 5) notably, in vivo hippocampal pyramidal cells have been proved to target CaM I mRNAs into dendrites.

A majority of the cultured cells exhibited pronounced expressions of all the CaM genes. Both earlier and these data indicate that the cells that expressed CaM markedly were neurons, while the expression of glial cells was so weak as to be almost undetectable. We found a strong dendritic targeting of all CaM mRNAs, regardless of the corresponding CaM gene. This study provides the first evidence of the localization of the CaM II and III mRNAs in the dendrites of CNS neurons. The extent of the dendritic distribution of the mRNAs was independent of the CaM genes. Additionally, our present study revealed a punctate localization of the CaM mRNAs, which is consistent with previous observations that mRNAs translocate in granular structures.

These findings are not in full agreement with previous in vitro data obtained on PC12 cells, where only the CaM I and II mRNAs were present in both the soma and the processes, while the CaM III signals were restricted to the cell bodies. However, PC12 cells are derived from peripheral neurons, which are quite different from the neurons in the CNS, which may well explain this discrepancy.

Double-labeling experiments revealed that the cultured cells can be grouped into three discrete categories: 1) CaM mRNA-positive neurons, 2) S100-immunopositive glial cells, and 3) cells negative for both the CaM mRNAs and S100 protein.

For a full understanding of the functions of the CaM gene family, much work remains to be done. Our results providing data on this issue may be summarized in the following points:

a) We cloned and sequenced the 3'-end of the 4.2 kb CaM I transcript missing from the genbanks. We also used this information in our in silico experiments.

b) The identification of several putative cis-acting elements in the multiple CaM transcripts hints at their unique intracellular fates. The elements analyzed may be grouped into the following categories: 1) signals driving the nuclear polyadenylation, 2) elements affecting the cytoplasmic polyadenylation, 3) localization signals, 4) factors influencing the stability of the transcript, and 5) elements regulating the translation. Our computational

analysis revealed several putative regulatory elements along the CaM transcripts. Interestingly, AREs are present only on the CaM I transcripts, while DICEs are characteristic for the CaM III transcripts. On the other hand, CPEs are found in the sequences of each CaM gene. Thus, there is a striking distribution and clustering of these signals in the 3'-UTRs, while the 5'-UTRs appear to be silent.

c) The quantitative detection of the CaM mRNA levels during ontogeny revealed gene-specific differences. The most prevalent variations were observed in the molecular layers, which mostly consisted of dendritic processes. By comparing the quotients of the highest and lowest signals for each CaM gene within a given region, we observed, for example, that in the molecular layer (layer 1) of the cerebral cortex the signals corresponding to the CaM II mRNAs underwent an 8-fold decrease between PD1 and PD20, whereas the decrease for CaM III was 4-fold and that for CaM I in the same area during the same period was only 1.7-fold.

d) The in vitro experiments demonstrated that translocation into the dendrites of cultured hippocampal neurons is a general characteristic of the mRNAs derived from different CaM genes. Moreover, the extent of the dendritic distribution of the mRNAs was independent of the CaM genes. Additionally, our present studies revealed a punctate localization of the CaM mRNAs, which is consistent with previous observations that mRNAs translocate in granular structures. Double-labeling experiments revealed that the cultured cells can be grouped into three discrete categories: 1) CaM mRNA-positive neurons, 2) S100-immunopositive glial cells, and 3) cells negative for both the CaM mRNAs and S100 protein.