Volume 239, number 1, 121–128 FEB 06431 October 1988

Characterization of an intronless human calmodulin-like pseudogene

Markus Koller and Emanuel E. Strehler

Laboratory for Biochemistry, Swiss Federal Institute of Technology (ETH), CH-8092 Zürich, Switzerland

Received 10 August 1988; revised version received 9 September 1988

We report the isolation and characterization of a human genomic clone encoding a calmodulin-like pseudogene. It contains an open reading frame of 444 nucleotides, not interrupted by introns. The nucleotide sequence of the open reading frame shows 80%, 71% and 69% identity to the previously reported human calmodulin cDNAs λht6 [17], hCWP [22], and λhCE1 [23], respectively. The derived amino acid sequence has only 85% identity to vertebrate calmodulin, but shows four potentially functional Ca^{2+}-binding loops. In the human tissues tested, this pseudogene is not expressed, though gene structure including promoter elements and a putative polyadenylation site seems to be intact.

Calmodulin; Calmodulin-like pseudogene

1. INTRODUCTION

Calmodulin (CaM), a highly conserved ubiquitous Ca^{2+}-binding protein, plays an important role in mediating intracellular calcium signals through a calcium-CaM complex [1]. Among its many functions, CaM has been shown to be involved in cell proliferation [2], DNA repair [3] and gene expression [4]. In addition, CaM binds to several nuclear proteins [5]. In proliferatively activated cells nuclear CaM increases and becomes predominantly associated with the matrix [6], which is known to be involved in DNA replication [7].

The regulatory mechanisms of CaM gene expression are not yet understood. In order to investigate the regulation of the CaM gene(s), it is necessary to know its (their) structure and organization. The structure of CaM genes in a variety of vertebrate [8–12] and invertebrate [13–16] species has been determined. In man, only a partial gene structure is known at present [17]. Beside these bona fide CaM genes there are additional CaM-related genes without introns in chicken [11] and rat [8,10]. These processed genes are thought to have originated from existing genes by reversed transcription.

We report here the cloning and sequencing of an intronless processed CaM-like gene from man. This processed gene, hGH6, encodes a CaM-like protein of 148 amino acids that differs significantly from all known CaM, yet may have retained four functional Ca^{2+}-binding sites and, if expressed, could represent a true 'isoform' of CaM. However, in the human tissues tested, no mRNA derived from hGH6 was detected.

2. MATERIALS AND METHODS

2.1. *Screening of a human genomic library*

Subfragments of the human CaM cDNA clone λht6 [17] were labelled with [α-^{32}P]dCTP by the oligolabelling method [18]. 10^6 phage plaques of a human leukocyte genomic library (Clontech Laboratories, Palo Alto, CA, USA) were screened following described procedures [19]. Positive phage clones were

Correspondence address: E.E. Strehler, Laboratory for Biochemistry, Swiss Federal Insstitute of Technology (ETH), CH-8092 Zürich, Switzerland

Abbreviations: CaM, calmodulin; cDNA, complementary DNA

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

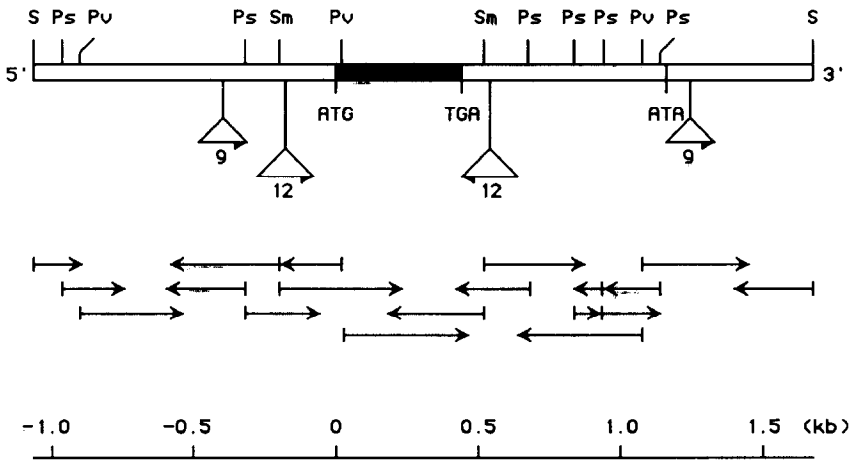

Fig.1. Partial restriction map and sequencing strategy for the human CaM-like pseudogene hGH6. The CaM-related open reading frame between the initiation codon (ATG) and the termination codon (TGA) is shown as a black box. A possible polyadenylation site is indicated by ATA. Arrows with numbers represent a direct repeat (9 nucleotides) and a palindromic sequence (12 nucleotides). Restriction enzyme sites are marked by: S, *Sst*I; Ps, *Pst*I; Pv, *Pvu*II; Sm, *Sma*I. Horizontal arrows indicate length and direction of sequence runs. Nucleotide numbering starts at the initiation codon (ATG) in the direction 5′ to 3′ of the sense strand. Negative numbers indicate nucleotides upstream of the initiation codon.

```
                                                5' - GAGCTCCTGGGGCAA   -1051
TGTGGACCAGAGAATGATAGTGGGACTTCCTGGCTGAGCCTTCAAGGTCTCCTTTGTCTCTCCTCCTTGT   -981
CCCCTCCCTCACAGGCCTGCAGCCCACTGATCATCTTCCAGCCCTTGGCCGTGCCATGGCCACTCCGGCC   -911
TAGACAGCTGCTTCTCAGCAGGAGGTCTGTCCTTGACCAGGCTGAAAGGTGACCTCGGGGTCCTGTCCAT   -841
CACCACCTTGTATGTCCATCAGCATTTATCAATAATCCCTAGTCACCTTGTTTGCATTCCTCGATCAGTT   -771
TACCTGCAAACTAGCTGCTTGCTAGTTAACCTCACGAGGACTGGGGGTGGCTTGCGGTCGCACTGTGTTC   -701
CCAGCTTCCAGAGTGACACTGCACCTAGTAGGCACTCAAATATCTGTGAAAGGGTGGATGAGTGAATGGG   -631
TTGGAGGAAACCCCTCCCTATTCATGTCCATTCTGAAGATAAGAAAACTCGCTGCTTCTACCTGAAGGGA   -561
AAACCTTCCTCCTCCAAAAACCCTCCTGTTCTATATGAGGCTTAGAAACCAAAGGCAAGGACGTTTCTTT   -491
CATCTGCCTTTCATCCTGCAAAAAGCCCTCCACTCAAGGCAAGAAAGGGCACTGAAGTTTATTGAGTGGG   -421
AGACAGCGGCAGGTGGGAGACCGGGGAGGGAGGAGAGAAAGGGAAATTCAGGAGGAAGGAGTCCAGCGTG   -351
GATTGCTCCAAAGCTCACCCACCACGCCCTGACTGCAGGTGTGATTCGGGGCCCCCGTGGCTCTGCTGGG   -281
TCCAGGTGCAAGCAGGCAAGAGGTGTGGCGTCAGCTCGATTCGCAGGCCCTGGACTACTGTCTAAACAGG   -211
ACAGGCCCGGGCAAGCAGGGCAGGGGCGTCTGCAATGATGGGGGAGGACTCTGCTGCTTCTTAAGCTCCA   -141
GCGTCTCAAGCCAGGGCGAGACAGCCCGCGGCCGCCCGGATCTCCACCTGCCACCCCAGAGCTGGGACAG    -71
AGCCGGGCTGCGGCACTGGGAGGGAGACCCCACAGTGGCCTCTTCTGCCACCCACGCCCCCACCCCTGGC     -1

--- --t --- --- --- --- --- --- --- a-t g-- --g --- --- --- --- --- ---
ATG GCC GAC CAG CTG ACT GAG GAG CAG GTC ACA GAA TTC AAG GAG GCC TTC TCC     54
Met Ala Asp Gln Leu Thr Glu Glu Gln Val Thr Glu Phe Lys Glu Ala Phe Ser

--c --- --- --- --- --a --t --- act --- --- --- aag --- t-- --g --a --g
CTG TTT GAC AAG GAT GGG GAC GGC TGC ATC ACC ACC CGC GAG CTG GGC ACG GTC    108
Leu Phe Asp Lys Asp Gly Asp Gly Cys Ile Thr Thr Arg Glu Leu Gly Thr Val
```

```
--- a-a --- --- --a --- --- --- --t --a --a --- --- -a- --t --- --c -a-
ATG CGG TCC CTG GGC CAG AAC CCC ACG GAG GCC GAG CTG CGG GAC ATG ATG AGT   162
Met Arg Ser Leu Gly Gln Asn Pro Thr Glu Ala Glu Leu Arg Asp Met Met Ser

--- g-g --t gca --t --g --- --g --- a-t --- --- --g --- --- --- ac- ---
GAG ATC GAC CGG GAC GGC AAC GGC ACC GTG GAC TTC CCC GAG TTC CTG GGC ATG   216
Glu Ile Asp Arg Asp Gly Asn Gly Thr Val Asp Phe Pro Glu Phe Leu Gly Met

--- --- --a --- --- --- --- --a --- -gt --- --- --- --- --a --- --g ---
ATG GCC AGG AAG ATG AAG GAC ACG GAC AAC GAG GAG GAG ATC CGC GAG GCC TTC   270
Met Ala Arg Lys Met Lys Asp Thr Asp Asn Glu Glu Glu Ile Arg Glu Ala Phe

--t --c --t --- --- --t --g --t --- -a- a-- --- --- --a --- --- --t ---
CGC GTG TTC GAC AAG GAC GGC AAC GGC TTC GTC AGC GCC GCC GAG CTG CGA CAC   324
Arg Val Phe Asp Lys Asp Gly Asn Gly Phe Val Ser Ala Ala Glu Leu Arg His

--a --- --g aac --- --- --- --- --- -cc --t --- --- --- --t --- --- ---
GTC ATG ACC CGG CTG GGG GAG AAG CTG AGT GAC GAG GAG GTG GAC GAG ATG ATC   378
Val Met Thr Arg Leu Gly Glu Lys Leu Ser Asp Glu Glu Val Asp Glu Met Ile

a-- -ag --t --- -tc --t --- --t --c --- --c --t --t --a --- --- --a -ag
CGG GCC GCG GAC ACG GAC GGA GAC GGA CAG GTG AAC TAC GAG GAG TTT GTC CGT   432
Arg Ala Ala Asp Thr Asp Gly Asp Gly Gln Val Asn Tac Glu Glu Phe Val Arg

a-- a-- act g-a --- ---
GTG CTG GTG TCC AAG TGA GGCCGGCGCCCACCATGCTCCTGGGCGCCCACGCGGCCCACAGGGC   496
Val Leu Val Ser Lys ***

AAGAACCCGGGGCCTCCCGCCTCCTCCCCCATCCCCCTGCCTCCCCTGGGCACTGTGGCTTCCTCCTGCG   566

CCTGGTTGATTCAGCCCACCTCTCTGCATCCCGCTTCCCGCGTCTCTTCTCTGCACTCCTGCCGACCTTC   636

CCACCTGCTCATCTGAATGACACGGAACGCTCCCACTGCAGGCAAACCGTGACGCCCTCCCCACTCGGGA   706

GAAGCAGAGCTGACCTTAGGACCGAGCACCAGGGCAGGTTGCGCTGACTCTGCGGCCCTCCAGGACGGAC   776

ACCGGGTGACCCTTAGGCACAGGCAAGATCCCTAACAGGCACCAATGCCAGGCAGGGGGCTGCAGCCCTC   846

AGCCCCCGCCAAGATTCCCGCAGGCTCCTGGACTGGAAGCTCCCTCCGCGGTCGGATTCTGGAGGGTGGG   916

AGGCATCTTGGCCTGCAGTAAGCGGTGCTGACGGGGACTCTGGCCACAGAGGTCAGGCCTCCTGAAAACA   986

GCACTGCCTTCCGCGCTGCCCCAGCTTGCCCCATTCCTTGTCCGCCAACCCACCGTGATTCATCTTCTGA  1056

AGCTGGGAGTGAAACTGGGTCAGCTGTAACCTGTTCCTATTCATCTGGAAGGAGGGAGGCTTGGATGAGC  1126

AGGGGATGAGAGCTGCAGGGAAATAAATGAGATATTCGTCCTTATTTCATTTCCTTTTTTTTTTGGTTT  1196

CCATGGAAATGATCCTTGTTAAATTCAGGGTTGAAACGAGGCAGGAATCTCCATTTTTGTGCTTTTTGAA  1266

AATGCAATGAATTCCTATACGGGGGAGCGGGAAAGGTGCCTCAGAGAGAGACAAGTCTGGATGAGGGAAA  1336

TATTGAATATTCTCAATCAAATGGATACGCTGGCAGCAAAGAGTGGTTAAAGTCCATCAGGACTTGAAAG  1406

ACCTGAGTCCATTACGTTGAGAAGGGACCTGCTGATTGCTTTGATTCCCCCTGGCAAGTGCTCCCTGGTT  1476

GTGAATGCCAGGCACTAGAGATGGTGAGGGGTTGGGGGGCAGTTGGGCACACACAGTGTAAGAGCAATTC  1546

AGAGCCGTTAGTCCTGCACTAGCCCTCAAGCTGCTGGCAACACCTAGAGAAGGTCGAAGGGCCCTGCCAG  1616

AGATCCCTTCAATTCTAAAGGGAGGTATGTTTTGCGGGATGGGCACTAGACGGAGCTC - 3'          1674
```

Fig.2. Nucleotide sequence of the human CaM-like pseudogene hGH6 (capital letters) and alignment with the coding sequence of the human CaM cDNA λht6 [17] (lower case letters). Hyphens indicate identical nucleotides between these two sequences. The translated amino acid sequence of hGH6 is given below the nucleotide sequence. Direct repeats and palindromic sequences are shown in italics. Putative regulatory sequences are underlined. Nucleotide numbering is the same as in fig.1.

recovered, replated and the plates were screened as before, until single lambda clones could be selected.

2.2. *Southern blotting, DNA subcloning and sequencing*

Lambda phage DNA was amplified and extracted as described [19]. Digestion with restriction enzymes, separation on 0.8% agarose gels and blotting onto nitrocellulose filters were as described [20]. The filters were hybridized and washed with the same probes and under the same conditions as used for the screening. A 2.7-kb *SstI* fragment containing CaM-related sequences was ligated into pUC18. The recombinant plasmid was amplified in *E. coli* JM101 and its DNA was purified in a CsCl density gradient [19]. DNA subfragments of the recombinant pUC18 plasmids were cloned in M13mp18 and M13mp19. The nucleotide sequence was determined by the dideoxy chain termination method [21]. Sequences were analyzed using the University of Wisconsin genetics computer group sequence analysis software package (Version 5.2, February 1988).

2.3. *Northern blotting*

RNA was separated on 1% agarose-formaldehyde gels, transferred to nitrocellulose and hybridized with ^{32}P-labelled [18] cDNA probes as described [19]. The final stringency of washing was 1 × SSC and 0.1% SDS at 50°C for 20 min. Filters were exposed to X-ray film for 24 h at −70°C.

3. RESULTS AND DISCUSSION

About 10^6 phage plaques of a human genomic DNA library were screened using as probes the *SmaI-EcoRI* fragment (from nucleotide 946 to the 3′-end) of the human CaM cDNA clone λht6 [17], and the *SmaI-SmaI* fragment (from nucleotide 88 to 560) of the same clone. One positive clone, which hybridized only to the coding region of the CaM cDNA, was selected for subcloning and sequence analysis. A hybridization-positive 2.7-kb *SstI* fragment of this clone, named hGH6, was subcloned into pUC18 and sequenced according to the strategy shown in fig.1. There is an open reading frame of 444 nucleotides, that specifies a 148-amino acid long CaM-like protein sequence. In the coding region of this intronless gene 354 out of 444 bases (80%) are identical to those of the CaM cDNA clone λht6 [17] (fig.2). Compared to the other two identified human CaM cDNAs, hCWP [22] and λhCE1 [23], the coding region of hGH6 shows identity values of 71% and 69%, respectively. Of the 90 differences between the coding regions of hGH6 and λht6 only 17 occur at first positions of codons but 59 at third positions. Obviously, there must have been (or still is) a selective pressure to keep the protein in a functional condition.

The hypothesis that pseudogenes are derived from intron-containing genes by a reverse transcriptional event is readily applicable to the rat pseudogenes λSC8 and λSC9 [8]. In the case of the human processed gene hGH6, however, the 'ancestor' gene is not easily recognized among the CaM genes corresponding to the three presently known CaM cDNAs. Because clone hGH6 shows no significant homology to any of these cDNAs in the flanking regions, a fourth bona fide CaM gene, not characterized so far, might be a more likely candidate as the ancestor gene of hGH6. Therefore, the human CaM multigene family may in fact consist of at least four separate members, excluding pseudogenes derived from particular CaM genes.

Codon usage and G+C content of the four known CaM(-like) sequences in man suggest a separation into two groups: the G+C-rich group (λht6 and hGH6) and the A+T-rich group (λhCE1 and hCWP). Members of a particular group prefer certain codons, resulting in a higher sequence homology in comparison to the homology between members of different groups (see table 1). A possible mechanism for the generation of these two groups of genes with divergent codon usages could be the duplication of an ancestral CaM gene followed by dispersion of the two resulting genes to different regions of the genome. It is thought that stretches of relatively uniform and distinct

Table 1

Codon usage and G+C content in the coding region of the CaM cDNAs λhCE1 [23], hCWP [22], λht6 [17] and the CaM-like pseudogene hGH6

		A+T rich group		G+C rich group	
		λhCE1 (%)	hCWP (%)	λht6 (%)	hGH6 (%)
Asp	G A C	41	18	41	94
Glu	G A G	19	5	90	95
Leu	C T G	22	22	78	100
Ala	G C C	0	9	27	89
Gly	G G T	36	27	0	0
Thr	A C A	75	50	17	10
Lys	A A A	50	63	0	0
G+C content at third codon position		26	33	68	92
Total G+C content		39	42	54	63

base composition, called isochores, influence the codon usage of genes in them [24]. The two ancient CaM genes, embedded in different isochores, may then have diverged in codon usage as well as in total G+C content. After a second duplication the four generated CaM genes again diverged mainly at third codon positions, giving rise to the proposed four human CaM genes.

```
Domain 1                          helix   Ca2+-loop   helix
                                10        20          30
                                . |       |.          . |       |
Mammalian [1]                 ADQLTEEQIAEFKEAFSLFDKDGDGTITTKELGTVMRSL
Scallop [31]                  ---------------------------------------
Trypanosome [14]              ----SN---S-----------------------------
Spinach [32]                  -ZZ--D-------------------C-------------
cCM1 [11]                     -ER-S---------------R----C---M---------
Lambda SC8 [8]                ----------------P---R---Q-V------M--G--
Lambda SC9 [8]                ---------------------------------------Q
hGH6                          --------VT---------------C---R---------

Domain 2                    40        50        60        70
                             .        | .     |     .       | .     |
Mammalian                     GQNPTEAELQDMINEVDADGNGTIDFPEFLTMMARK
Scallop                       ---------------------------------------
Trypanosome                   -----------------Q--S----------L----
Spinach                       -------------------------------NL----
cCM1                          ------------VG------S---------SL----
Lambda SC8                    -------------S----------N-----------
Lambda SC9                    -------------------------S-----------
hGH6                          ---------R--MS-I-R-----V------G-----

Domain 3                       80        90        100       110
                                .     |     . |       .     |     . |
Mammalian                     MKDTDSEEEIREAFRVFDKDGNGYISAAELRHVMTNL
Scallop                       -----------------------F-------------
Trypanosome                   -Q-S------K------------F---------I----
Spinach                       ---------LK---------Q--F-------------
cCM1                          -R-S-----------------------------------
Lambda SC8                    --G--V---------------**-T----F-----T-
Lambda SC9                    ------K-----------------F------------
hGH6                          -----N-----------------FV----------R-

Domain 4                           120       130       140
                                     .|      | .         .|
Mammalian                     GEKLTDEEVDEMIREADIDGDGQVNYEEFVQMMTAK
Scallop                       -------------------------------T---S-
Trypanosome                   -----------------V-----I------K--MS-
Spinach                       -----------------V-----I------KV-M--
cCM1                          -------------K---CNN----------R---E-
Lambda SC8                    ----------G------M------I-------V---
Lambda SC9                    -------DI-------------------------S-
hGH6                          ----S---------A--T------------RVLVS-
```

Fig.3. Alignment of the translated amino acid sequence of hGH6 with CaM sequences from different sources. The derived amino acid sequences of pseudogenes cCM1, λSC8 and λSC9 (ignoring frameshift mutations by insertions in λSC9) are also included in the comparison. Residues identical to those of the mammalian sequence are indicated by hyphens. Deletions are shown by asterisks. Z stands for Glu or Gln. The four Ca^{2+}-binding domains are aligned such that the Ca^{2+}-binding loops are beneath each other.

700 base pairs downstream of the termination codon (TGA), there is a possible polyadenylation signal AATAAA (underlined in fig.2). In the upstream region of the hypothetical initiation codon no obvious TATA box was detected. A possible CAAT box was observed 175 nucleotides upstream of the hypothetical start of translation. This putative promoter region is characterized by repeat sequences, namely GCAGG-like and CACCC-like sequences. The GCAGG-box, repeated nine times within 200 nucleotides, is also observed in other genes [25], but its function is unknown. The CACCC-like repeats, also found in other genes [25,26], are thought to be a regulatory sequence of house-keeping genes [27]. 250 base pairs upstream of ATG, another possible regulatory sequence is observed: TGGCGTCA, a variant of the cAMP-responsive element, TGACGTCA [28]. This finding coincides with the observation that the cytosolic level of CaM is cAMP-dependent [29].

Transcript accumulation of the human CaM-like pseudogene hGH6 was tested on Northern blots of RNA from various tissues (fig.4). A 0.7-kb *Sma*I-*Sma*I fragment containing the coding region of clone hGH6, hybridized to a 2.3-kb mRNA in human erythroleukemic K562 cells (fig.4a). This 2.3-kb mRNA species corresponds to CaM cDNA λht6, which is very abundant in K562 cells [17]. The coding region of clone hGH6, showing an identity of 80% to clone λht6, cross-hybridized to this 2.3-kb mRNA. In the tissues tested no mRNA of higher homology to hGH6 was detected, indicating that hGH6 may be an unexpressed

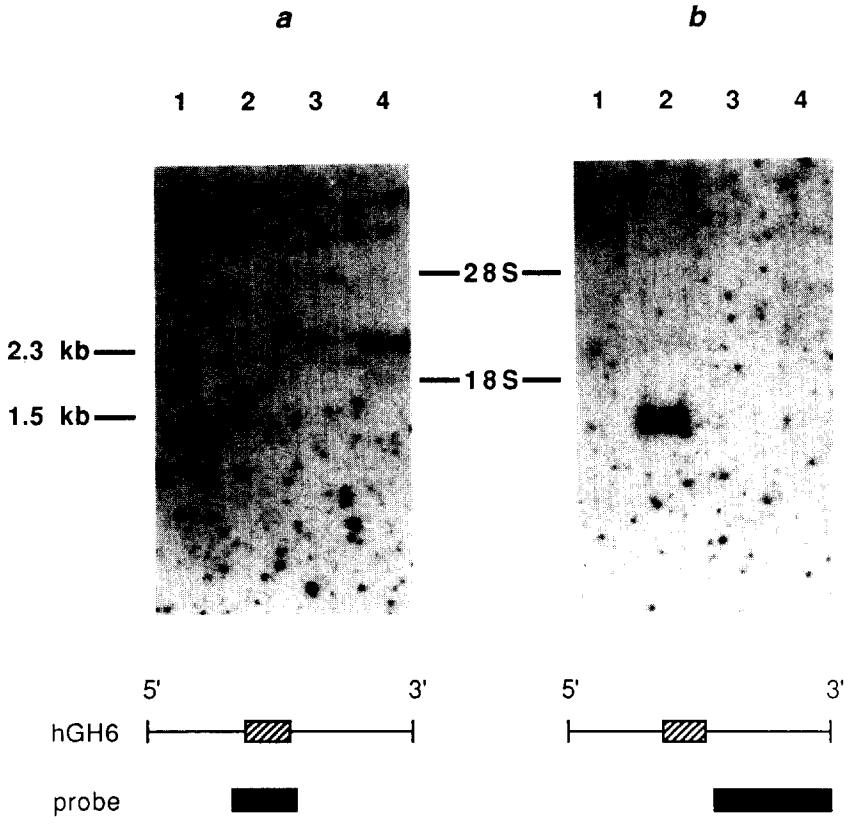

Fig.4. Identification of hGH6-related mRNA species in different human tissues by Northern blotting. Total RNA (10 μg) from human fibroblasts (lanes 1), SV-40-transformed human fibroblasts (lanes 2), human teratoma cells (lanes 3) and human erythroleukemic K562 cells (lanes 4) was used. Two equivalent blots were probed with (a) a 0.7-kb *Sma*I-*Sma*I fragment containing the coding region of clone hGH6, and (b) a 1.2-kb *Sma*I-*Sma*I fragment comprising 3'-flanking sequences of clone hGH6. The positions of 28 S and 18 S rRNA as well as of the two most prominent bands (2.3 kb and 1.5 kb) are indicated.

pseudogene. Expression in other tissues, however, cannot be ruled out at present. The second probe, a 1.2-kb *SmaI-SstI* fragment containing the 3′-flanking region of clone hGH6, hybridized specifically to a 1.5-kb mRNA species present in transformed fibroblasts (fig.4b). In contrast, the 0.7-kb coding probe of hGH6 hybridized only weakly to this 1.5-kb band. For an expressed CaM-like pseudogene hGH6 one would expect a hybridization signal of the same intensity for both, the coding and the 3′-flanking probe. Computer comparisons using the 'wordsearch' program showed that particularly the sequences adjacent to the open reading frame of hGH6 seem to have some relationship to viral sequences, indicating a virally mediated process of genome insertion as mechanism for the generation of the CaM-like pseudogene hGH6. These regions also contain direct repeats, for example AGGGCA, CCCACC, CCTGG or CCCACAG, occurring in the 5′- as well as the 3′-flanking region. Fibroblasts, transformed by SV-40, could express viral sequences which are recognized by probes derived from flanking regions of clone hGH6.

If expressed, hGH6 would encode a protein of 148 amino acids. 125 amino acids are identical with those in vertebrate CaM (85% identity). Many of the 23 changed amino acids are conservative substitutions (see fig.3). Other changes, not necessarily conservative, such as Cys for Thr in the first Ca^{2+}-binding loop and Ser for Ala at the C-terminal end, also occur in other species. All four Ca^{2+}-binding loops seem to have retained potential Ca^{2+}-binding activity. The derived amino acid sequence of hGH6 is diverging markedly from vertebrate CaM in the C-terminal region (see fig.3). It is known that this domain is the immunoreactive site of CaM [30]. Expression of the CaM-isoform hGH6 in different tissues could therefore be tested by immunological methods using specific antibodies against this protein. Furthermore, in vitro expression of this mutant CaM would provide a suitable system for studying the structural and functional behaviour of such a CaM isoform.

Acknowledgements: We are grateful to Roland Fischer for the gift of human CaM cDNA (λht6), to Marie-Antoinette Strehler-Page for initial library screening and to Beat Trüeb for his generous gift of human fibroblast RNA. We would also like to thank Professor Ernesto Carafoli for critical reading of the manuscript. This work was supported by Swiss National Science Foundation grant number 3.531.-0.86.

REFERENCES

[1] Klee, C.B. and Vanaman, T.C. (1982) Adv. Protein Chem. 35, 213–321.
[2] Rasmussen, C.D. and Means, A.R. (1987) EMBO J. 6, 3961–3968.
[3] Chafouleas, J.G., Bolton, W.E. and Means, A.R. (1984) Science 224, 1346–1348.
[4] White, B.A. (1985) J. Biol. Chem. 260, 1213–1217.
[5] Bachs, O. and Carafoli, E. (1987) J. Biol. Chem. 262, 10786–10790.
[6] Serratosa, J., Pujol, M.J., Bachs, O. and Carafoli, E. (1988) Biochem. Biophys. Res. Commun. 150, 1162–1169.
[7] Smith, H.C. and Berezney, R. (1983) Biochemistry 22, 3042–3046.
[8] Nojima, H. and Sokabe, H. (1986) J. Mol. Biol. 190, 391–400.
[9] Nojima, H. and Sokabe, H. (1987) J. Mol. Biol. 193, 439–445.
[10] Nojima, H., Kishi, K. and Sokabe, H. (1987) Mol. Cell. Biol. 7, 1873–1880.
[11] Stein, J.P., Munjaal, R.P., Lagacé, L., Lai, E.C., O'Malley, B.W. and Means, A.R. (1983) Proc. Natl. Acad. Sci. USA 80, 6485–6489.
[12] Chien, Y.H. and Dawid, I.B. (1984) Mol. Cell. Biol. 4, 507–513.
[13] Smith, V.L., Doyle, K.E., Maune, J.F., Munjaal, R.P. and Beckingham, K. (1987) J. Mol. Biol. 196, 471–485.
[14] Tschudi, C., Young, A.S., Ruben, L., Patton, C.L. and Richards, F.F. (1985) Proc. Natl. Acad. Sci. USA 82, 3998–4002.
[15] Goldhagen, H. and Clarke, M. (1986) Mol. Cell. Biol. 6, 1851–1854.
[16] Davis, T.N., Urdea, M.S., Masiarz, F.R. and Thorner, J. (1986) Cell 47, 423–431.
[17] Fischer, R., Koller, M., Flura, M., Mathews, S., Strehler-Page, M.-A., Krebs, J., Penniston, J.T., Carafoli, E. and Strehler, E.E. (1988) J. Biol. Chem., in press.
[18] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.
[19] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[20] Southern, E.M. (1975) J. Mol. Biol. 98, 503–517.
[21] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5468.
[22] Wawrzynczak, E.J. and Perham, R.N. (1984) Biochem. Int. 9, 177–185.
[23] SenGupta, B., Friedberg, F. and Detera-Wadleigh, S.D. (1987) J. Biol. Chem. 262, 16663–16670.
[24] Alonso, S., Minty, A., Bourlet, Y. and Buckingham, M. (1986) J. Mol. Evol. 23, 11–22.
[25] Nakanishi, T., Kohno, K., Ishiura, M., Ohashi, H. and Uchida, T. (1988) J. Biol. Chem. 263, 6384–6391.

[26] Myers, R.M., Tilly, K. and Maniatis, T. (1986) Science 232, 613–618.
[27] Lawn, R.M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980) Cell 21, 647–651.
[28] Montiminy, M.R., Sevarino, K.A., Wagner, J.A., Mandel, G. and Goodman, R.H. (1986) Proc. Natl. Acad. Sci. USA 83, 6682–6686.
[29] Soriano, M., Pujol, M.J. and Bachs, O. (1988) J. Cell. Physiol. 135, 345–350.
[30] Van Eldik, L.J. and Watterson, D.M. (1981) J. Biol. Chem. 256, 4205–4210.
[31] Toda, H., Yazawa, M., Kondo, K., Honma, T., Narita, K. and Yagi, K. (1981) J. Biochem. 90, 1493–1505.
[32] Lukas, T.J., Iverson, D.B., Schleicher, M. and Watterson, D.M. (1984) Plant Physiol. 75, 788–795.