Vol. 164, No. 1, 1989 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 16, 1989 Pages 503-511

PRIMARY STRUCTURE AND FUNCTIONAL EXPRESSION OF *h*-CALDESMON COMPLEMENTARY DNA

Ken'ichiro Hayashi[1,2], Keiko Kanda[1], Fusao Kimizuka[2], Ikunoshin Kato[2] and Kenji Sobue[1§]

[1]Department of Neurochemistry and Neuropharmacology
Biomedical Research Center
Osaka University Medical School, Nakanoshima, Kita-ku, Osaka 530, Japan

[2]Pharmaceutical Research Laboratories Takara Shuzo Co., Ltd.
3-4 Seta, Otsu-shi, Shiga 520-21, Japan

Received September 6, 1989

SUMMARY: Recently, the two *Mr* forms of caldesmon (*Mr's* in the range of 120-150kDa and 70-80kDa as judged by SDS-PAGE) have been identified. *h*-Caldesmon (high *Mr* 120-150kDa caldesmon) is predominantly expressed in smooth muscles, and *l*-caldesmon (low *Mr* 70-80kDa caldesmon) in non-muscle cells. In this paper, we report the nucleotide sequence of chick embryo gizzard *h*-caldesmon cDNA and its translation into amino acid sequence. This sequence predicts a protein of 771 amino acids with a *Mr* of 88,743. The central portion of this sequence is composed of a 10-fold repeat of conserved amino acid sequence containing 13-15 amino acids. Further, a recombinant protein produced in *Escherichia coli* containing the full-length *h*-caldesmon cDNA has been characterized. Although the *Mr* of *h*-caldesmon predicted from amino acid sequence is 88,743, native and recombinant proteins show the same mol. wt. with 150kDa as measured by SDS-PAGE. This discrepancy may be due to the acidic amino acid-rich sequences at the N-terminal and central portions. A recombinant protein produced in *E. coli* possesses calmodulin-, F-actin- and tropomyosin-binding abilities in common with the native *h*-caldesmon.

Caldesmon is a ubiquitous calmodulin- and F-actin-binding protein (reviewed in ref. 1). *h*-Caldesmon was initially purified from chicken gizzard smooth muscle (2). Nowadays, the two *Mr* forms of caldesmon (*h*- and *l*-caldesmons) have been identified in a variety of tissues and cells (3-7). From accumulating evidence, it is suggested that caldesmon plays a vital role in smooth muscle and non-muscle contraction (1).

Most recently, we have found the expressional change of caldesmon during phenotypic modulation of smooth muscle cells (7). In developing gizzards, the expression of caldesmon switches from *l*- to *h*-forms. When

§To whom all correspondence should be addressed.

0006-291X/89 $1.50

grown in cell culture of vascular smooth muscle cells, they convert from differentiated to dedifferentiated phenotypes. In accordance with this phenotypic modulation, the expression of caldesmon also changes from *h*- to *l*-forms. Additionally, the levels of mRNAs, which govern the synthesis of both caldesmons, are observed to be the same in proportion to protein levels. These findings suggest that the expressional change of caldesmon during the phenotypic modulation of smooth muscle cells may be regulated at the level of upstream of mRNA. As a first step to clarify the molecular events of this subject, we attempted to determine the molecular cloning of caldesmons. In our previous paper (8), we have demonstrated the nucleotide and deduced amino acid sequences of the C-terminal 35kDa domain in *h*-caldesmon. We report here the primary structure of full-length *h*-caldesmon cDNA and the characterization of recombinant protein encoded by this cDNA.

MATERIALS AND METHODS

Materials. The purification of *h*-caldesmon from chicken gizzard was performed according to the method as described previously (2) with a slight modification of heat treatment. *h*-Caldesmon-, C-terminal 35kDa fragment- and residual N-terminal fragment-specific antibodies were prepared as described elsewhere (3,8). Restriction enzymes and linker- and primer-DNAs were obtained from Takara Shuzo, Japan.

Cloning and sequencing of *h*-caldesmon cDNA. Double-stranded cDNAs derived from chick embryo gizzard poly(A)$^+$ RNA were synthesized with oligo(dT) or a specific oligodeoxyribonucleotide as a primer, using the cDNA synthesis system (Amersham) (9). The cDNAs were blunted with T_4 DNA polymerase, methylated with *Eco*RI methylase, and ligated with *Eco*RI linker. *h*-Caldesmon cDNA clone was isolated from this cDNA library (constructed from size-fractionated *Eco*RI-digested DNA). λ_gt11 was used as an expression vector. This library was screened using a C-terminal 35kDa fragment-specific antibody (8) visualized with alkaline phosphatase-conjugated second IgG (Cooper Biochemical). The cDNA library constructed with an 18 nucleotides specific primer, complementary to nucleotides 1609-1626, was further screened by a similar immunological method using a residual N-terminal fragment-specific antibody (the flow-through fraction of the C-terminal 35kDa fragment-conjugated affinity chromatography). DNA restriction fragments of *h*-caldesmon cDNA were subcloned in M13 vectors: M13mp18*h*CaDN (carrying 1 to 1615) and M13mp18*h*CaDC (carrying 1487 to 2564 and poly(A) tract) (see Fig. 4). The nucleotide sequence was determined by the dideoxy chain termination method (10). The hydropathy plot was made according to the index of Kyte and Doolittle (11). The secondary structure prediction was analyzed using the method of Chou and Fasman (12).

Construction of expression plasmid (pTCG*h*CaD) in *E. coli*. The expression plasmid pTCG*h*CaD was constructed as follows. The outline of the construction method is shown in Fig. 4. M13mp18CaDN RF-DNA was digested with *Nco*I, filled in with Klenow fragments, and then partially digested with *Eco*T14I. A 1.3 kbp fragment encoding the N-terminal portion of *h*-caldesmon was isolated. M13mp18*h*CaDC RF-DNA was digested with *Bam*H I, filled

in with Klenow fragment, and then digested with *Eco*T14I. An 8 kbp fragment containing the C-terminal portion of *h*-caldesmon was also isolated. Both fragments were ligated to construct M13mp18*h*CaD carrying the entire *h*-caldesmon cDNA. The *Nco*I site was regenerated during this construction. The cDNA was isolated from M13mp18*h*CaD by *Ava*II digestion, blunting with Klenow fragment, and then *Nco*I digestion. This cDNA was ligated between *Nco*I and *Sma*I sites of pTV118N (13) to construct pTCG*h*CaD. Then, *E. coli* (JM109) was transformed by pTCG*h*CaD.

Binding assays of recombinant protein. *E. coli* transformed by pTCG*h*CaD was harvested and suspended in an extraction buffer (20 mM imidazole•HCl, pH 7.2, 1mM dithiothreitol, 300 mM KCl, 1 mM EGTA and 0.25 mM phenylmethylsulfonyl fluoride). This suspension was sonicated, boiled for 5 min, chilled on ice, and then centrifuged at 15,000 x g for 30 min. The resulting supernatant (heat-resistant extract) was dialyzed against a buffer (20 mM imidazole • HCl, pH 7.2, 0.5 mM dithiothreitol, 2 mM $MgCl_2$ and 20 mM KCl), and was used for the following binding assay. The calmodulin- and tropomyosin-binding activities were measured by the respective affinity column assays (2,8). The F-actin-binding assay was performed by the cosedimentation method (2,14). A recombinant protein crossreacted with a *h*-caldemon antibody was detected by immunoblotting.

Others. The actin-activated ATPase activity of gizzard myosin was measured under the same conditions as described in ref. 14. Blot hybridization analysis of chick embryo gizzard mRNAs was carried out by the method described elsewhere (15,16). Chymotryptic and CNBr-treated digests of *h*-caldesmon were electroblotted onto PVDF membrane. The partial sequences of the above digests were directly determined by the cut pieces of PVDF membrane with the aid of Applied Biosystems gas phase sequencer (8). SDS-polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the method of Laemmli (17). Immunoblotting was carried out by the method of Towbin *et. al.* (18).

RESULTS AND DISCUSSION

A chick embryo gizzard cDNA library was screened with two antibodies against the C-terminal 35kDa and the residual N-terminal fragments of *h*-caldesmon. Both antibodies specifically crossreacted with the respective fragments, but not with the counterparts. Overlapping DNA restriction fragments of *h*-caldesmon cDNA clone were sequenced using the dideoxy method (10). Figure 1 shows the nucleotide sequence of the chick embryo gizzard cDNA clone encoding *h*-caldesmon and its deduced amino acid sequence. The ATG coding for the initiating methionine is assigned because the nucleotide sequence surrounding it agrees well with the consensus sequence established for the sites of translation initiation (19). The 3'-noncoding region of cDNA is 38 nucleotides long (excluding the poly(dA) tract). The ACTAAA sequence at 24 nucleotides upstream from poly(dA) tract seems to be the polyadenylation signal. The long open reading frame following the initial methionine predicts a protein of 771 amino acids with a *Mr* of

```
1    GCC TTT CTT GCT GCG CTC CTG CCC AAG CAG CCA GTT ATT CAT TTC CTT CCT ACT AGT TGA AAA CTT TTC ACC GCT

76   GAG TGA TCA CCT CTA GTT TCT CGA GTC AGC GTT CTG ATT TCG TCA GCA TCT GGT TCA TTA CCG AAA GGT TCA GAA

151  TAT TTC CTC TGA TTG CTT TCC TGT TTG GAT TGT TCC AGA GCT CTG AAC ACT TGA ACA CCC ACC ATG GAT GAC TTT
                                                                                  Met Asp Asp Phe    4

226  GAA CGC CGT AGA GAA CTC AGG AGG CAA AAG CGT GAG GAA ATG CGC CTT GAA GCA GAG AGA CTG TCC TAC CAG AGA
     Glu Arg Arg Arg Glu Leu Arg Arg Gln Lys Arg Glu Glu Met Arg Leu Glu Ala Glu Arg Leu Ser Tyr Gln Arg   29

301  AAT GAT GAT GAT GAG GAA GAA GCT GCC AGA GAA CGT CGT CGA CGA GCT CGA CAG GAA AGG CTG CGG CAA AAG GAA
     Asn Asp Asp Asp Glu Glu Glu Ala Ala Arg Glu Arg Arg Arg Arg Ala Arg Gln Glu Arg Leu Arg Gln Lys Glu   54

376  GAA GGA GAT GTA TCA GGG GAA GTA ACA GAG AAA TCA GAA GTT AAT GCC CAG AAC AGT GTG GCA GAA GAA GAA ACG
     Glu Gly Asp Val Ser Gly Glu Val Thr Glu Lys Ser Glu Val Asn Ala Gln Asn Ser Val Ala Glu Glu Glu Thr   79

451  AAA CGT AGT ACA GAT GAT GAA GCT GCA TTG TTG GAG AGA CTG GCA AGA CGG GAA GAG AGA CGC CAA AAA CGT CTA
     Lys Arg Ser Thr Asp Asp Glu Ala Ala Leu Leu Glu Arg Leu Ala Arg Arg Glu Glu Arg Arg Gln Lys Arg Leu  104

526  CAG GAA GCC CTG GAA CGT CAA AAG GAA TTT GAC CCA ACG ATC ACA GAT GGG AGC TTG TCA GTG CCC AGC AGG AGA
     Gln Glu Ala Leu Glu Arg Gln Lys Glu Phe Asp Pro Thr Ile Thr Asp Gly Ser Leu Ser Val Pro Ser Arg Arg  129

601  GAA GTA AAC AAT GTG GAA GAA AAT GAG ATC ACA GGG AAA GAG GAA AAG GTT GAA ACA CGC CAA GGA CGC TGT GAG
     Glu Val Asn Asn Val Glu Glu Asn Glu Ile Thr Gly Lys Glu Glu Lys Val Glu Thr Arg Gln Gly Arg Cys Glu  154

676  ATT GAG GAA ACA GAA ACA GTT ACC AAA TCG TAC CAA AGG AAC AAT TGG AGG CAA GAT GGA GAA GAA GAG GGA AAA
     Ile Glu Glu Thr Glu Thr Val Thr Lys Ser Tyr Gln Arg Asn Asn Trp Arg Gln Asp Gly Glu Glu Glu Gly Lys  179

751  AAA GAA GAA AAA GAC TCA GAA GAG GAG AAA CCA AAG GAG GTA CCC ACA GAG GAA AAT CAG GTA GAT GTG GCA GTA
     Lys Glu Glu Lys Asp Ser Glu Glu Glu Glu Lys Pro Lys Glu Val Pro Thr Glu Glu Asn Gln Val Asp Val Ala Val  204

826  GAA AAG TCC ACA GAT AAA GAA GAG GTG GTA GAA ACA AAA ACT CTA GCT GTA AAT GCA GAG AAT GAT ACA AAT GCT
     Glu Lys Ser Thr Asp Lys Glu Glu Val Val Glu Thr Lys Thr Leu Ala Val Asn Ala Glu Asn Asp Thr Asn Ala  229

901  ATG CTG GAA GGG GAG CAG AGT ATA ACT GAT GCT GCA GAT AAA GAG AAG GAA GAG GCT GAG AAA GAA AGG GAG AAA
     Met Leu Glu Gly Glu Gln Ser Ile Thr Asp Ala Ala Asp Lys Glu Lys Glu Glu Ala Glu Lys Glu Arg Glu Lys  254

976  CTT GAG GCA GAA GAA AAG GAG AGG TTA AAA GCA GAA GAA GAA AAG AAG GCA GCT GAA GAA AAA CAG AAA GCA GAG
     Leu Glu Ala Glu Glu Lys Glu Arg Leu Lys Ala Glu Glu Glu Lys Lys Ala Ala Glu Glu Lys Gln Lys Ala Glu  279

1051 GAG GAA AAG AAG GCA GCT GAG GAA AGA GAG CGG GCT AAA GCA GAA GAG GAG AAG AGA GCA GCT GAG GAA AGA GAG
     Glu Glu Lys Lys Ala Ala Glu Glu Arg Glu Arg Ala Lys Ala Glu Glu Glu Lys Arg Ala Ala Glu Glu Arg Glu  304

1126 AGG GCT AAG GCA GAA GAG GAG AGG AAA GCA GCT GAG GAA AGA GAG CGG GCT AAG GCA GAA GAG GAG AGG AAA GCA
     Arg Ala Lys Ala Glu Glu Glu Arg Lys Ala Ala Glu Glu Arg Glu Arg Ala Lys Ala Glu Glu Glu Arg Lys Ala  329

1296 GCT GAG GAG AGG GCT AAG GCA GAA GAG GAA AGG AAA GCA GCT GAG GAG AGG GCT AAG GCA GAA GAG GAA AGG AAA
     Ala Glu Glu Arg Ala Lys Ala Glu Glu Glu Arg Lys Ala Ala Glu Glu Arg Ala Lys Ala Glu Glu Glu Arg Lys  354

1276 GCA GCT GAG GAG AGG GCT AAG GCA GAA AAG GAG AGG AAA GCA GCT GAG GAG AGA GAG AGG GCT AAG GCA GAA GAG
     Ala Ala Glu Glu Arg Ala Lys Ala Glu Lys Glu Arg Glu Lys Ala Ala Ala Glu Glu Arg Glu Arg Ala Lys Ala Glu Glu  379

1351 GAA AAG AGG GCA GCT GAA GAA AAG GCT AGG TTA GAG GCA GAA AAA TTA AAG GAA AAG AAA AAG ATG GAA GAG AAG
     Glu Lys Arg Ala Ala Glu Glu Lys Ala Arg Leu Glu Ala Glu Lys Leu Lys Glu Lys Lys Lys Met Glu Glu Lys  404

1426 AAA GCC CAA GAG GAA AAA GCT CAA GCA AAT TTG CTA AGA AAA CAG GAG GAA GAC AAA GAG GCT AAA GTG GAA GCT
     Lys Ala Gln Glu Glu Lys Ala Gln Ala Asn Leu Leu Arg Lys Gln Glu Glu Asp Lys Glu Ala Lys Val Glu Ala  429

1501 AAA AAG GAA AGC TTA CCA GAG AAG CTT CAA CCT ACC TCC AAA AAA GAT CAG GTA AAA GAC AAC AAG GAT AAA GAA
     Lys Lys Glu Ser Leu Pro Glu Lys Leu Gln Pro Thr Ser Lys Lys Asp Gln Val Lys Asp Asn Lys Asp Lys Glu  454

1576 AAA GCA CCC AAG GAG GAA ATG AAG AGT GTC TGG GAT CGT AAA AGG GGA GTT CCA GAA CAA AAG GCA CAG AAT GGA
     Lys Ala Pro Lys Glu Glu Met Lys Ser Val Trp Asp Arg Lys Arg Gly Val Pro Glu Gln Lys Ala Gln Asn Gly  479

1651 GAA CGT GAA CTC ACT ACC CCC AAA CTT AAA TCT ACT GAA AAT GCT TTT GGA CGC TCC AAC TTG AAA GGA GCT GCA
     Glu Arg Glu Leu Thr Thr Pro Lys Leu Lys Ser Thr Glu Asn Ala Phe Gly Arg Ser Asn Leu Lys Gly Ala Ala  504

1726 AAT GCT GAG GCT GGC TCC GAG AAG CTG AAG GAG AAA CAG CAG GAG GCA GCT GTG GAG CTG GAT GAG CTG AAG AAA
     Asn Ala Glu Ala Gly Ser Glu Lys Leu Lys Glu Lys Gln Gln Glu Ala Ala Val Glu Leu Asp Glu Leu Lys Lys  529

1801 AGG CGG GAG GAG CGC CGG AAA ATC CTG GAG GAA GAG GAG CAG AAG AAG AAA CAG GAG GAG GCT GAG AGA AAA ATC
     Arg Arg Glu Glu Arg Arg Lys Ile Leu Glu Glu Glu Glu Gln Lys Lys Lys Gln Glu Glu Ala Glu Arg Lys Ile  554

1876 AGA GAG GAG GAG GAA AAG AAG AGG ATG AAG GAA GAA ATT GAA AGG AGA AGA GCT GAA GCT GCT GAG AAA CGT CAG
     Arg Glu Glu Glu Glu Lys Lys Arg Met Lys Glu Glu Ile Glu Arg Arg Arg Ala Glu Ala Ala Glu Lys Arg Gln  579

1951 AAA GTG CCA GAA GAT GGC GTA TCT GAA GAA AAG AAG CCA TTT AAA TGT TTC AGT CCT AAA GGT TCA TCT CTC AAG
     Lys Val Pro Glu Asp Gly Val Ser Glu Glu Lys Lys Pro Phe Lys Cys Phe Ser Pro Lys Gly Ser Ser Leu Lys  604

2026 ATA GAA GAA CGA GCA GAA TTT TTG AAC AAA TCC GCT CAG AAG AGT GGT ATG AAA CCT GCC CAC ACG ACA GCA GTT
     Ile Glu Glu Arg Ala Glu Phe Leu Asn Lys Ser Ala Gln Lys Ser Gly Met Lys Pro Ala His Thr Thr Ala Val  629

2101 GTC TCA AAG ATT GAC AGT AGA CTT GAG CAA TAT ACT AGC GCA GTT GTG GGC AAC AAG GCT GCA AAA CCC GCC AAA
     Val Ser Lys Ile Asp Ser Arg Leu Glu Gln Tyr Thr Ser Ala Val Val Gly Asn Lys Ala Ala Lys Pro Ala Lys  654

2176 CCA GCA GCC TCT GAC CTT CCT GTT CCA GCT GAG GGT GTC CGT AAT ATC AAG AGC ATG TGG GAG AAA GGG AAT GTT
     Pro Ala Ala Ser Asp Leu Pro Val Pro Ala Glu Gly Val Arg Asn Ile Lys Ser Met Trp Glu Lys Gly Asn Val  679

2251 TTT TCA TCA CCT GGG GGA ACA GGA ACA CCA AAT AAG GAA ACT GCT GGA CTG AAA GTT GGT GTC TCC AGT CGT ATC
     Phe Ser Ser Pro Gly Gly Thr Gly Thr Pro Asn Lys Glu Thr Ala Gly Leu Lys Val Gly Val Ser Ser Arg Ile  704

2326 AAC GAA TGG CTA ACC AAG ACC CCA GAG GGT AAC AAA TCG CCT GCT CCA AAA CCT TCT GAT TTA AGA CCA GGA GAT
     Asn Glu Trp Leu Thr Lys Thr Pro Glu Gly Asn Lys Ser Pro Ala Pro Lys Pro Ser Asp Leu Arg Pro Gly Asp  729

2401 GTA TCC GGC AAA CGT AAT CTC TGG GAG AAG CAG TCA GTT GAA AAG CCA GCT GCT TCT TCT TCA AAG GTA ACA GCT
     Val Ser Gly Lys Arg Asn Leu Trp Glu Lys Gln Ser Val Glu Lys Pro Ala Ala Ser Ser Ser Lys Val Thr Ala  754

2476 ACG GGG AAA AAA TCA GAG ACT AAT GGT TTG AGA CAA TTT GAG AAA GAA CCG TAG AAG GCT ACT AAA GAC GCT GGA
     Thr Gly Lys Lys Ser Glu Thr Asn Gly Leu Arg Gln Phe Glu Lys Glu Pro                                 771

2551 CCA ATC AGT TGG GGA AAA AAA AAA AAA AAA A
```

Fig. 1. Nucleotide sequence of *h*-caldesmon cDNA clone and the deduced amino acid sequence. Nucleotide sequence numbering appears on the left, and amino acid sequence numbering on the right. Underlining of the predicted amino acid sequence indicates regions confirmed by amino acid sequencing of chymotryptic and CNBr-cleaved peptides derived from *h*-caldesmon.

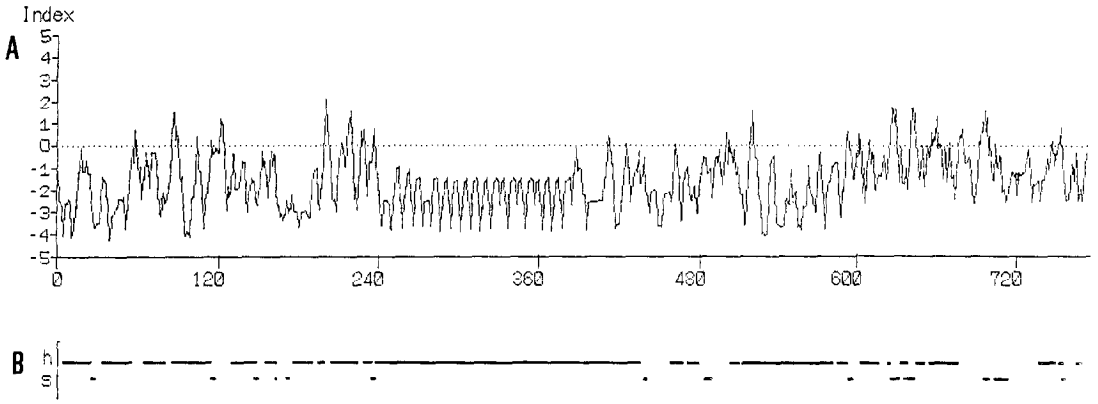

Fig. 2. Hydropathicity profile (A) and predicted secondary structures (B) of *h*-caldesmon. The hydropathicity was calculated by the method of Kyte and Doottitle (11) with a window of 5. The secondary structures were analyzed by the method of Chou and Fasman (12). The positions of α-helical and β-sheet structures are indicated by bars at h and S lines, respectively.

88,743. The amino acid composition predicted from the nucleotide sequence coincides with that from the direct determination of native *h*-caldesmon (20,21). All the partial amino acid sequences of from chymotryptic and CNBr-cleaved peptides derived from the native *h*-caldesmon are encoded by the cDNA in the same reading frame (underlined in Fig. 1).

In Figure 2A and B, the amino acid sequence presented was analyzed for local hydropathicity and secondary structure. Residues 239-434 show an uninterrupted α-helical and highly hydrophilic structure, and are composed of a 10-fold repeat of conserved amino acid sequence containing 13-15 amino acids; one repeating unit is based on a glutamic acid-rich sequence, EEER(or K)KAAEERERAKA (Fig. 3). However, these residues cannot form a coiled-coil structure. As demonstrated previously, residues 524-581 and 623-637 in the C-terminal 35kDa domain of *h*-caldesmon are highly homologous to those of troponin T (8), which are the conserved consensus sequences of the torpomyosin-binding domains (so-called T1 and T2 fragments) in the striated muscle troponin Ts (22,23). The predicted sequence of *h*-caldesmon does not share a significant homology with any other proteins. With computer searching, only a small homologous sequences (taking conservative substitutions into account, the homologies are less than 35%) are, however, found between the highly helical region of *h*-caldesmon and

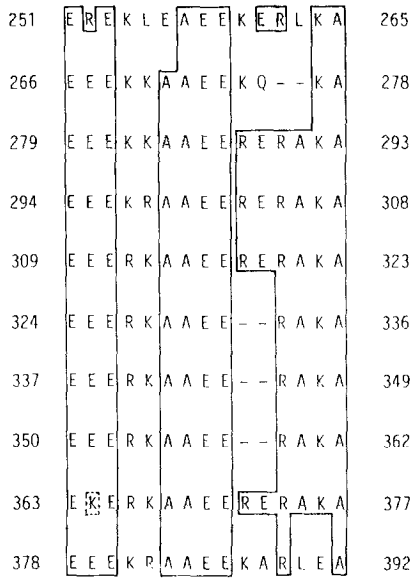

Fig. 3. Regions of repeating sequence of *h*-caldesmon. Numbers indicate residue positions within the *h*-caldesmon sequence. The conserved amino acid repeat sequences are enclosed with solid lines.

tropomyosin α and β subunits (24,25), rod portion of myosin heavy chain (26,27) or tail subdomain of neurofilament L protein (28,29). Ngai and Walsh have reported that *h*-caldesmon itself shows the Ca^{2+}- and calmodulin-dependent protein kinase activity (30). In contrast, significant homology in the primary structures between *h*-caldesmon and some protein kinases has not been observed. Despite the regulatory mode of this protein on the actomyosin system resembles that of troponin I, there is no structural homology between the two proteins.

It is interesting to note that a *Mr* of *h*-caldesmon calculated from the predicted amino acid sequence is much smaller than that measured by SDS-PAGE (120-150kDa). Using sedimentation equilibrium, Graceffa *et al.* have reported a *Mr* of this protein at 93,000 ± 4,000 (19). Our result, from the predicted sequence, is in good agreement with their value. A *Mr* of recombinant protein produced in *E. coli* containing the full-length *h*-caldesmon cDNA was compared with that of native *h*-caldesmon. The construction method for *h*-caldesmon cDNA is shown in Fig. 4 and described in MATERIALS AND METHODS. Using an *h*-caldesmon antibody, immunoblotting analysis of heat-resistant extract of *E. coli* revealed a discrete band with

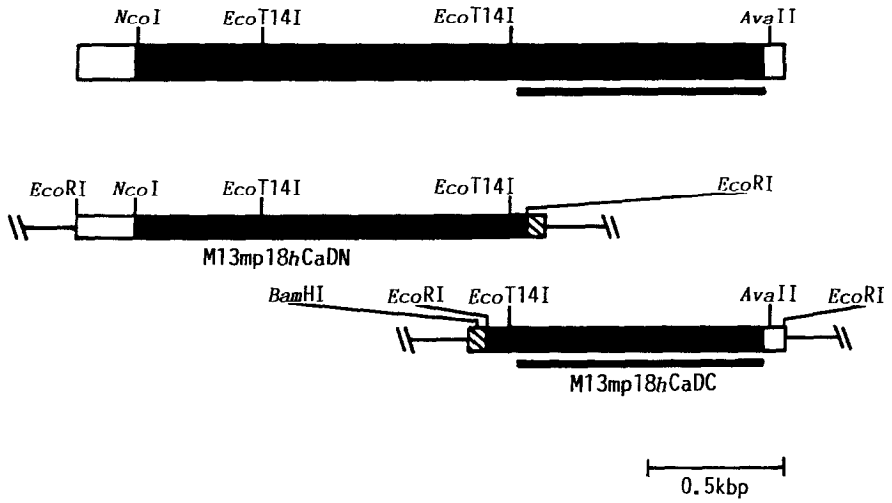

Fig. 4. Construction of expression plasmid pTCGhCaD carrying the chick embryo gizzard *h*-caldesmon cDNA. The detailes of the construction method are described in MATERIALS AND METHODS.

an apparent *Mr* of 150kDa and the following small *Mr* bands assuming the breakdown products of the parent 150kDa band. The largest *Mr* protein, which is heat-resistant and crossreacts with an *h*-caldesmon antibody, therefore corresponds to the native *h*-caldesmon as judged by SDS-PAGE (Fig. 5a and c). A *Mr* of the C-terminal 35kDa domain in *h*-caldesmon

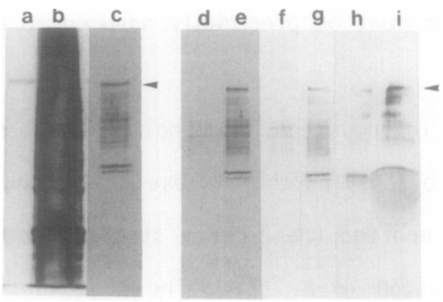

Fig. 5. Characterization of a recombinant protein in *E. coli* transformed by pTCGhCaD. Lanes a and b, Coomassie blue staining of SDS gels; lanes c-i, immunoreplica using a *h*-caldesmon antibody. *h*-Caldesmon purified from chicken gizzard (lane a); heat-resistant extract of *E. coli* transformed by pTCGhCaD (b); immunoreplica corresponding to lane b (c); flow through fraction of calmodulin affinity column (d); Ca^{2+}-dependent calmodulin-binding fraction (EGTA eluate fraction, e); flow through fraction of tropomyosin affinity column (f); tropomyosin-binding fraction (high-salt eluate fraction, g); lanes h and i indicate the supernatant and the precipitated fraction measured by cosedimentation assay. Arrowheads indicate a recombinant protein with a *Mr* of 150 kDa which crossreacts with a *h*-caldesmon antibody.

measured by SDS-PAGE is in good agreement with that calculated according to the predicted sequence. Thus, the discrepant result regarding a *Mr* of this protein may be explained by an unexpected structure of the N-terminal and central portions. Based on the present primary structure, These portions have a high content of acidic amino acid in addition to the unique repeating sequence mainly composed of glutamic acid, resulting in the anomalous mobility on SDS-PAGE. A recombinant protein with a *Mr* of 150kDa produced in *E. coli* also showed the same properties of native *h*-caldesmon with respect to the calmodulin-, F-actin- and tropomyosin-binding abilities (Fig. 5d-i). Judging from the actomyosin ATPase activity, the crude calmodulin-binding fraction of the heat-resistant extract from *E. coli* transformed by *h*-caldesmon cDNA inhibited the actin-myosin interaction in a dose-dependent manner (not shown). These results give support for the predicted sequence presented here to be the sequence of *h*-caldesmon.

In our preliminary experiment, using the deletioned *h*-caldesmon cDNA probe poly(A)$^+$ RNA preparation from 10-day-old embryo gizzards, in which *h*- and *l*-caldesmons are coexpressed, was subjected to blot hybridization. This probe encodes the C-terminal 35kDa domain of *h*-caldesmon. We have found that it hybridizes two major RNA species of ~3,000 and ~2,500 nucleotides (not shown). This result suggests that the hybridizable mRNAs of ~3,000 and ~2,500 nucleotides must be presumably assumed to be *h*- and *l*-caldesmon mRNAs.

As mentioned in Introduction, sequencing analysis of caldesmons is an essential approach to studying the expressional change of this protein. In this paper, we have reported the primary structure of *h*-caldesmon cDNA and its functional expression in *E. coli*. The present results will, in near future, provide a large amount of information with respect to the relationship between structure and function of caldesmon, and new tools to study the biological significance of this protein.

ACKNOWLEDGMENTS-- This work was supported by grants from the Ministry of Education, Science and Culture of Japan to K. Sobue.

REFERENCES

1. Sobue, K., Kanda, K., Tanaka, T., and Ueki, N. (1988) J. Cell. Biochem., 37, 317-325.
2. Sobue, K., Muramoto, Y., Fujita, M., and Kakiuchi, S. (1981) Proc. Natl. Acad. Sci. USA, 78, 5652-5655.
3. Owada, M. K., Hakura, A., Iida, K., Yahara, I., Sobue, K., and Kakiuchi, S. (1984) Proc. Natl. Acad. Sci. USA, 81, 3133-3137.
4. Sobue, K., Tanaka, T., Kanda, K., Ashino, N., and Kakiuchi, S. (1985) Proc. Natl. Acad. Sci. USA, 82, 5025-5029.
5. Bretscher, A. and Lynch, W. J. (1985) J. Cell. Biol., 100, 1656-1663.
6. Dingus, J., How, S. and Bryan, J. (1986) J. Cell. Biol., 102, 1748-1757.
7. Ueki, N., Sobue, K., Kanda, K., Hada, T., Higashino, K. (1987) Proc. Natl. Acad. Sci. USA, 84, 9049-9053.
8. Hayashi, K., Yamada, S., Kanda, K., Kimizuka, F. Kato, I. and Sobue, K. (1989) Biochem. Biophys. Res. Commun., 161, 38-45.
9. Gubler, U. and Hoffman, B. J. (1983) Gene, 25, 263-269.
10. Sanger, F., Nicklen, S., and Coulson, A. (1977) Proc. Natl. Acad. Sci. USA, 74, 5463-5467.
11. Kyte, J. and Doolittle, R. F. (1982) J. Mol. Biol., 157, 105-132.
12. Chou, P. Y. and Fasman, G. D. (1978) Ann. Rev. Biochem., 47, 251-276.
13. Maki, M., Takano, E., Mori, H., Sato, A., Murachi, T. and Hatanaka, M. (1987) FEBS. Lett., 223, 174-180.
14. Sobue, K., Takahashi, K., and Wakabayashi, I. (1985) Biochem. Biophys. Res. Commun., 132, 645-651.
15. Goldberg, D. A. (1980) Proc. Natl. Acad. Sci. USA, 77, 5794-5798.
16. Thomas, P. (1980) Proc. Natl. Acad. Sci. USA, 77, 5201-5205.
17. Laemmli, U. K. (1970) Nature, 227, 680-685.
18. Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. USA, 76, 4530-4354.
19. Kozak, M. (1984) Nucleic Acids Res., 12, 857-872.
20. Graceffa, P.. Wang, C-L. A. and Stafford, W. F. (1988) J. Biol. Chem., 263, 14196-14202.
21. Lynch, W. P., Riseman, V. M. and Bretscher, A. (1987) J. Biol. Chem., 262, 7429-7437.
22. Pato, M. D., Mak, A. S., and Smillie, L. B. (1981) J. Biol. Chem., 256, 602-607.
23. Mak, A. S. and Smillie, L. B. (1981) J. Mol. Biol., 149, 541-550.
24. Ruiz-Opazo, N. and Nadal-Ginard, B. (1987) J. Biol. Chem., 262, 4755-4765.
25. Lewis, W. G., Cote, G. P. Mak, A. S. and Smillie, L. B. (1983) FEBS. Lett., 156, 269-273.
26. Karn, J. Brenner, S. and Barnett, L. (1983) Proc. Natl. Acad. Sci. USA, 80, 4253-4257.
27. Capony, J-P. and Elzinga, M. (1981) Biophys. J., 33, 148a.
28. Geisler, N. Kaufmann, E. Fischer, S. Plessmann, U. and Weber, K. (1983) EMBO. J., 2, 1295-1302.
29. Lewis, S. A. and Cowan, N. J. (1986) Mol. Cell. Biol., 6, 1529-1534.
30. Ngai, P. and Walsh, M. P. (1987) Biochem. J., 244, 417-425.