Cloning and Expression of a Smooth Muscle Caldesmon*

(Received for publication, April 11, 1989)

Joseph Bryan‡, Michiylo Imai, Robyn Lee, Paul Moore, Richard G. Cook§, and Wen-Gin Lin

From the Department of Cell Biology and the §Howard Hughes Medical Institute and the Department of Microbiology and Immunology, Baylor College of Medicine, Houston, Texas 77030

Caldesmon is a smooth muscle and nonmuscle regulatory protein that interacts with actin, myosin, tropomyosin, and calmodulin. Two overlapping clones, isolated from a chicken oviduct cDNA plasmid library and a chicken gizzard cDNA λ NMI1149 library, were used to generate a 4108-base pair sequence coding for one caldesmon. Expression of the coding sequence confirms this is one of the large smooth muscle caldesmons. The deduced protein molecular weight is 86,974, significantly less than the molecular weights estimated by sodium dodecyl sulfate gel electrophoresis. The protein has a high content of Gly, Lys, Arg, and Ala; there are two cysteine residues, one at either end of the molecule. Comparison with the Protein Identification Resource database demonstrates a similarity with a tropomyosin binding domain of troponin T, but none with any calmodulin or actin binding proteins. The center of the protein has an 8-fold repeat of a 13 amino acid sequence whose general motif is -Glu-(Lys/Arg)-Ala-Ala-Glu-(Lys/Arg)-X-(Lys/Arg)-Ala-to, where X is Glu, Gln, or Ala. Comparison with peptide sequences from a chymotryptic fragment that binds actin and calmodulin places this domain on the C terminus of caldesmon adjacent to the troponin T similarity. A tentative map of the major binding domains is proposed on the basis of available data.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) J00460.

To whom correspondence should be addressed: Dept. of Cell Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, email: jbryan@bem.tmc.edu.

1 The abbreviations used are: EGTA, [ethylenebis(oxyethylene-nitriilo)]tetraacetic acid; SDS, sodium dodecyl sulfate; TnT, troponin T.

The function of caldesmon is unclear. In reconstitution experiments, caldesmon inhibits the actin-activated ATPase of myosin (19–21). This inhibition is attenuated by calcium-calmodulin and is potentiated by tropomyosin. Several groups have reported that caldesmon will interact directly with tropomyosin (22, 23). The mechanism of the ATPase inhibition is complex (24, 25), and there are several reports that smooth muscle myosin binds directly to caldesmon (25–27). Ikebe and Reardon (26) have reported that smooth muscle heavy meromyosin and myosin subfragment 2, but not myosin subfragment 1, are specifically retained on a caldesmon-Sepharose column.

Other reports have aimed to define the organization of the actin, calmodulin, tropomyosin, and myosin binding sites on the caldesmon molecule. Proteolytic cleavage experiments (28-30) have localized the actin and calmodulin binding sites to a fragment with an estimated molecular weight of 38,000–40,000. On the basis of expression experiments using a cDNA that encoded part of caldesmon, we argued that two binding domains were on the C-terminal end of the protein (29). Recently Riseman et al. (31) have used partial cleavage at and labeling of two cysteines to establish this point more firmly. Sutherland and Walsh (27) have employed a similar strategy to argue the myosin binding domain is on the amino-terminal half, although these authors arrive at a quite different peptide order than Riseman et al. (31). Finally, Wang (32) has labeled the cysteine residues with a photolabile reagent and shown that caldesmon can bind two calmodulin molecules, one at either end.

We have cloned, sequenced, and expressed one of the high molecular weight caldesmons from chicken smooth muscle. The deduced protein sequence coupled with peptide sequencing confirms the localization of the actin-calmodulin binding fragment to the C terminus and confirms that the two cysteine residues are at opposite ends of the protein in agreement with the results of Mornet et al. (33), Wang (32), and Riseman et al. (31). The deduced molecular weight indicates that the smooth muscle forms are markedly smaller than estimated from SDS gels and significantly less than determined by sedimentation equilibrium, 93,000 ± 4,000 vs 86,974. In addition, the deduced protein sequence reveals a similarity to a tropomyosin binding site on troponin T, suggesting the location of the tropomyosin binding domain. Finally, the central region of the molecule contains a unique repetitive sequence of unknown function. The organization of sites on the extended molecule suggests that caldesmon would be well suited to act as a bridge between myosin and actin filaments.

EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes were from New England BioLabs; Sequenase was from United States Biochemical Corp., and common chemicals were from Sigma.
DNA Sequencing—The M13 dideoxy procedure was used to sequence DNA fragments cloned into M13mp18, M13mp19, or Phage-script (Stratagene, La Jolla, CA). All cloning procedures were performed as described by Maniatis et al. (34). Sequencase was used following the manufacturer's directions.

Peptide Sequencing—Chicken gizzard caldesmon was isolated as described by Bretscher (2); the molecular weight 40,000 calmodulin actin-binding peptide was isolated as described by Fujii et al. (29). The NH2-terminal 20 amino acids of this chymotryptic peptide were determined using the Applied Biosystems Inc. (Foster City, CA) model 470-A gas phase protein sequencer equipped with an in-line model 120-A phenylhydroxynitrocellulose analyzer.

Vector Construction—pLC-globin was obtained from K. Nagai (Medical Research Council, Cambridge, United Kingdom). The details of this vector are given by Nagai and Thogersen (35). This plasmid contains a β-globin cDNA insert between BamHI and HindIII sites. The construction of a full-length caldesmon expression vector, designated pCDM314, was done as follows. Starting with a Phage-script clone containing the caldesmon coding sequence, CDM57 in Fig. 1, we isolated a 1.8-kilobase BamHI-HindIII fragment containing the 5' end of caldesmon. This was subcloned into pUC119 obtained by restricting pLC-globin with BamHI and HindIII. The resulting intermediate was restricted with HindIII, filled with deoxynucleotide triphosphates using the Klenow fragment of DNA polymerase and then restricted with XhoI. A 2450-base pair XhoI-Smal fragment was isolated from a pUC19 vector carrying CDM57 and subcloned into this intermediate. This produced a plasmid carrying the full coding sequence. This plasmid was restricted with BamHI and dephosphorylated using calf intestinal alkaline phosphatase. The FX-oligomucleotide (see Ref. 35), GATCCATCGAGTGGCGCTACCCCGTATG, carrying a Stul site was cloned into the dephosphorylated BamHI site of this vector. The final expression vector, pLC-CDM57, was obtained by releasing the plasmid carrying the FX-oligo at its unique NcoI site, at the ATG start codon, blunting with mung bean nuclease, restricting with Stul, and religating. All of these manipulations were done in the 2940 cell strain described by Shatzman and Rosenberg (36). This carries a defective λ lysogen with a wild type λ repressor. For expression, the final plasmid was transformed into either NSI61 cells (36) or QY13 cells (35) carrying a defective λ lysogen with a temperature-sensitive λ repressor. Cells were grown to an OD600 of 1 at 30 °C and 1 ml aliquots were heat shocked at 43 °C for 20 min and then grown for an additional 3 h at 37 °C. Cells were pelleted in a Microfuge for 3 min and resuspended in 500 μl of SDS sample buffer for gel electrophoresis. Electrophoresis and immunoblotting were done as described by Fujii et al. (29). The antibodies used have been described by Dingus et al. (13) and Fujii et al. (29).

RESULTS

We previously reported the cloning of a cDNA coding for a part of chicken gizzard caldesmon. This was obtained by screening a λgt11 library with anti-caldesmon antibodies (29). This fragment was used to screen a chicken oviduct library constructed using the Okayama-Berg vector (37) and a chicken gizzard library constructed in the λ phase NM1149 (38). The relationship of these two clones, designated CDM2 and CDM57, respectively, is shown in the restriction map in Fig. 1.

Fig. 1. Partial restriction map of caldesmon. The relationship of the two clones sequenced in this study is shown along with a partial restriction map. CDM2 was isolated from a chicken oviduct cDNA library constructed using the Okayama-Berg plasmid vector. CDM57 was isolated from a λ NM1149 chicken gizzard cDNA library. The position of the caldesmon open reading frame is shown for reference. aa, amino acid.

The longest clone from the plasmid library contained a 2606-base pair insert which terminated at a HindIII site at its 5' end and at a poly(A) tract at its 3' end. Because the complete sequence shows two closely spaced HindIII sites at this location, we assume the mRNA/cDNA heteropolymers were cleaved by HindIII during construction of the library, resulting in premature termination. CDM2 includes 1004 base pairs of coding sequence and 1602 base pairs of 3'-untranslated sequence.

The longest clones isolated from the NM1149 library, all approximately 2500-2600 base pairs, were initiated by oligo(dt) primers hybridized to a stretch of 12 internal A residues starting at position 2541 which is located about 35 base pairs inside the 3'-untranslated region. The sequences of CDM2 and CDM57 overlap for 1038 base pairs (Fig. 1), positions 1502-2540 in Fig. 2. We have found only a single base pair difference at position 2115, a T in CDM2 and a C in CDM57. This change does not lead to an amino acid change in the deduced protein sequence. The longest cDNAs isolated from the λ NM1149 library contain the entire coding sequence but lack most of the 3'-untranslated region.

Polyadenylation Signals—The 3'-untranslated region has two AATAAA polyadenylation signals at positions 3514 and 4085, respectively. In poly(A)+ mRNA from chicken gizzard we detect a major 4100-nucleotide species and a minor, estimated <5%, 3500-nucleotide RNA on Northern blots (data not shown).

Deduced Protein Sequence—The longest open reading frame (Fig. 2) begins with a good Kozak (39) consensus sequence at position 235 and terminates at position 2503, specifying a protein of 756 amino acid residues. The calculated molecular weight is 8698; the amino acid composition, shown in Table I, is in excellent agreement with the recent report from Graceffa et al. (16). Caldesmon has an unusually high content of glutamic acid and lysine, 22 and 15%, respectively, and a low content of aromatic amino acids. The calculated pl is 5.07; the calculated partial specific volume is 0.748 g/ml. In agreement with the protein chemical evidence (16, 31-33), caldesmon has two cysteine residues which are located at positions 153 and 580, respectively, in the amino acid-terminal halves of caldesmon. The predicted molecular weight is markedly less than the values, M, 120,000-150,000, determined by SDS gel electrophoresis and is approximately 6000 less than the sedimentation equilibrium determination, 93,000 + 4,000 reported by Graceffa et al. (16).

The calculated extinction coefficient for caldesmon is εM = 33,270 M-1 cm-1, derived using the extinction coefficients for Trp (εM = 5850 M-1 cm-1) and Tyr (εM = 1340 M-1 cm-1) given by Magne et al. (40). The εM is 3.8 in agreement with the value reported by Graceffa et al. (16) which was obtained by determining protein concentration from the refractive index measured in the analytical ultracentrifuge.

The amino acid sequences in the two underlined regions in Fig. 2 have been determined using protein sequencing methods. We determined the first sequence, beginning at Asp451, described by Szpacenko and Dabrowska et al. (29, 30) as an actin and a calmodulin digests of chicken gizzard caldesmon (33). We have two other confirmatory peptide sequences obtained from Dr. Rob Cross at the Medical Research Council in Cambridge and from Dr. Albert Wang at the Boston Biomedical Research Institute.
performed as outlined under "Experimental Procedures." The se-

Table I

| Amino acid | Occurrences | % of total | Amino acid | Occurrences | % of total |
|------------|-------------|------------|------------|-------------|------------|
| Ala        | 84          | 11.11      | Tyr        | 3           | 0.40       |
| Cys        | 2           | 0.26       | Asp        | 28          | 3.70       |
| Glu        | 168         | 22.22      | Phe        | 8           | 1.06       |
| Gly        | 29          | 3.84       | His        | 1           | 0.13       |
| Ile        | 11          | 1.46       | Lys        | 111         | 14.68      |
| Leu        | 27          | 3.57       | Met        | 8           | 1.06       |
| Lys        | 30          | 3.97       | Arg        | 72          | 9.52       |
| Ser        | 41          | 5.42       | Thr        | 31          | 4.10       |
| Val        | 35          | 4.63       | Trp        | 5           | 0.66       |

**Acidic**

| Polyamide | 196 | 25.53 | Basic | 184 | 24.34 |
|------------|-----|-------|-------|-----|-------|
| Nonpolar   | 213 | 28.17 |

**Fig. 3.** Self comparison matrix of the predicted caldesmon protein sequence. The similarity of the deduced caldesmon (Cam) protein sequence with itself was determined using an optimized search algorithm (53, 54). The window size was 10 residues; a unit cost matrix was employed that assigns a match a value of 1 and a mismatch a value of 0. 

These are located on the amino-terminal half of the deduced sequence.

Sequence Similarities and Structural Predictions—An analysis of the caldesmon protein sequence using a comparison matrix reveals the remarkable centrosymmetric pattern shown in Fig. 3. The protein has two main repetitive regions: a large central sequence which begins near residue 250 and ends near 400 and a second smaller and weaker pattern which begins near residue 510 and ends at about 570. The weaker repetitive pattern serves to define the head of caldesmon, which contains the actin, calmodulin, and perhaps the tropomyosin binding domains. The core of the longer central repetitive region consists of a 13 amino acid sequence repeated eight times with some minor variations. The general motif is

\[ \text{Glu} - \text{Lys (Arg)} - \text{Ala} - \text{Glu} - \text{Lys (Arg)} - \text{X} - \text{Lys (Arg)} - \text{Ala} - \text{X} \]

where X is Glu, Gln, or Ala. This region is shown in more detail in Fig. 4, where the nondegenerate positions have been boxed in. Four positions, 4, 5, 10, and 12, show a preservation of charge with the substitution of either arginine or lysine; position 11 is the most variable. This repeated motif sequencing methods are underlined. The first peptide was sequenced as described under "Experimental Procedures," the second sequence is from Morren et al. (33).
of caldesmon. The eight repeated sequences are aligned to show the
13-residue repeating unit. The regions of identity have been
boxed.

FIG. 4. Amino acid sequence in the central repeated region of caldesmon. The eight repeated sequences are aligned to show the
13-residue repeating unit. The regions of identity have been
boxed.

can be extended somewhat, particularly toward the C terminus, with the insertion of some padding but becomes increas-
ingly degenerate.

A search of the Protein Identification Resource database using the predicted caldesmon sequence indicates a similarity
with all of the published troponin T (TnT) sequences. The specific region identified in rabbit skeletal muscle troponin T was taken from the Protein Identification Resource database and compared with the weakly repetitive region in the head region of caldesmon. Identical amino acids are marked
by vertical lines; the alternating charge motif has been emphasized by outlining the acidic and basic amino acid blocks.

FIG. 5. Similarity of the weakly repetitive region in the
caldesmon head domain with the Ca"-insensitive a-skeletal
troponymosin binding site of troponin T. The sequence for rabbit skeletal muscle troponin T is from the Protein Identification Resource database and compared with the weakly repetitive region in the head region of caldesmon. Identical amino acids are marked
by vertical lines; the alternating charge motif has been emphasized by outlining the acidic and basic amino acid blocks.

The search of the Protein Identification Resource database using the predicted caldesmon sequence indicates a similarity
with all of the published troponin T (TnT) sequences. The specific region identified in rabbit skeletal muscle troponin T is
Glu86 to Lys147. This sequence is in a TnT peptide which has been identified as a calcium-insensitive skeletal a-tropomyosin binding site (41, 42). The region of similarity in caldesmon is the weakly repetitive region, Glu86 to Lys147, in Figs. 2 and 3. This similarity is shown in more detail in Fig. 5. The overall identity is 43%, 25 amino acids over a stretch of 58 residues. More striking is the similarity of the distribution of acidic and basic residues indicated by the + and - symbols. This repeat region in TnT has been analyzed in some detail by Parry (43). This sequence is in the head region and is a candidate for the site of tropomyosin binding to caldesmon identified by Graceffa (22) and localized to the head region by Fujii et al. (23).

The Garnier-Osguthorpe-Robson (44) and Chou-Fasman (45) secondary structure prediction algorithms suggest that caldesmon could have a significant a-helical content; the predicted values are near 80-85%. The central repeated region is predicted by all of the algorithms to be helical, but how much would be stable in solution in an extended protein like caldesmon is not clear. The estimate of a-helix content by
deduced protein

In Vitro Expression—It could be argued, on the basis of apparent molecular weights, that we have cloned the smaller
nonmuscle form of caldesmon. To evaluate this possibility, we have subcloned the coding region into the pLcII expression plasmid described by Nagai and Thogersen (35). The fusion protein coded by this construct, pLc-CDM314, contains 231 amino acids from the large II gene, three from a BamHI linker, four from the factor X recognition site, and 755 from caldesmon. The expected molecular weight is about 91,000. Two cell
strains, QY13 and N5151 (35, 36), carrying defective lambda
lysogens and a temperature-sensitive lambda repressor were transfected with this construct and grown at 30 °C. Expression
was induced by heat shock at 45 °C for 20 min followed by
growth at 37 °C. After 3 h, cells were collected and prepared for SDS gel electrophoresis. Caldesmon expression was detected by blotting with anti-caldesmon antibodies as described by Fujii et al. (29). The results are given in Fig. 6. The largest polypeptide had an apparent molecular weight slightly larger than chicken gizzard caldesmon control (lane a) and is produced in larger amounts after
heat shock (compare lanes c and e with d and f). We have not established whether the M, ~40,000 species is derived from caldesmon. This species, and a small amount of the M, 125,000 material, is present in the 294cl" cell strain transformed with pCDM314 and
grown at 37 °C (lane b). Neither species is present in QY13 or N5151 cells alone or after transformation with pLC-8-globin (data not shown).

FIG. 6. Expression of a cII-caldesmon fusion protein using
pLc-CDM314. QY13 and N5151 cells were transformed with pLc-
CDM314 as described under "Experimental Procedures" and grown to an OD600 of 1.0 at 30 °C. 1-ml aliquots were maintained at 30 °C or heat-shocked at 43 °C for 20 min and then grown for 3 h at 37 °C. Cells were harvested in a microcentrifuge, dissolved in SDS sample buffer, and subjected to electrophoresis in 7.5% polyacrylamide gels containing 0.1% SDS. Immunoblots were done as described by Fujii et al. (29). The SDS extracts from the cells treated at 43 °C (lanes c and e) show two major bands (arrows) with relative molecular weights of about 125,000 and approximately 40,000 and a number of minor bands, presumably proteolytic fragments of the M, 125,000 species. The large band migrates slightly faster than the chicken gizzard caldesmon control (lane a) and is produced in larger amounts after
heat shock (compare lanes c and e with d and f). We have not established whether the M, ~40,000 species is derived from caldesmon. This species, and a small amount of the M, 125,000 material, is present in the 294cl" cell strain transformed with pCDM314 and
grown at 37 °C (lane b). Neither species is present in QY13 or N5151 cells alone or after transformation with pLC-8-globin (data not shown).

DISCUSSION

We have sequenced two overlapping cDNAs coding for one
caldesmon and have provided evidence that it is one of the
"large" smooth muscle forms of this protein. We found a
single base pair difference in the overlapping sequence which
does not change the amino acid sequence. The sequence for
CDM57 is given in Fig. 2 with a C in position 2115. It is not
clear whether this single base change means the clone from the
Okayama-Berg chicken oviduct library is a second iso-
form, a polymorphism, or the result of an artifact introduced
during reverse transcription and cloning. The deduced protein

Molecular Cloning of Caldesmon
sequence confirms our earlier speculative positioning of the actin-calmodulin binding head region at the C terminus of caldesmon. The sequence allows us to resolve some of the conflicting reports about this protein and begin to order the binding domains. To facilitate the discussion and to begin to relate known caldesmon protein chemistry to the sequence, we have summarized some of the available data in Fig. 7. Chymotryptic cleavage defines the head region, whereas cleavage at the cysteine residues produces three peptides designated I–III. The assignments of the individual binding domains illustrated in Fig. 7 are discussed below.

**General Properties and Molecular Weight Problems—**Caldesmon has a high content of glutamic acid, lysine, and arginine and is low in aromatic amino acids. The acidic and basic residues are asymmetrically distributed with peptides I and II being acidic overall, whereas the head region and peptide III are basic in character, which agrees with the observation that the head region is not retained on DEAE resins. The protein has a calculated isoelectric point of 5.07, assuming the amino terminus is blocked.

All of the estimated molecular weights for caldesmon determined by SDS gel electrophoresis differ significantly from the expected values. A precise reason for this disagreement is not known, but Takano et al. (46) have pointed out a correlation of the misestimate with the content of acidic residues for several different proteins (see Ref. 47 for comparison of caldesmon with various other proteins). For caldesmon, the estimated $M_r$ in SDS ranges from 120,000–150,000 versus the expected 87,000. The in vitro expression of a cDNA coding for the overestimation of the molecular weight by sedimentation equilibrium by about 6–7% is not clear.

The difficulty with the molecular weight estimates extends to the various pieces. For example, the estimated versus expected values for the cysteine cleavage peptides are; peptide I, $M_r$ 28,000 versus 18,400; peptide II, $\approx M_r$ 80,000 versus 49,800, and peptide III, $M_r$ 25,000 versus 18,800. Similarly, the $M_r$ 38,000–40,000 head region, Asp$^{461}$ to Pro$^{509}$, has a true

cellular weight of 34,000 and is approximately 39% of the entire molecule. The fact that the head region and peptide III estimates are in error, even though they are not particularly rich in acidic groups, suggests that the slower mobility is due to something more than simple charge. These errors increase the difficulty of mapping proteolytic and CNBr fragments. It is difficult, for example, to correlate the recent CNBr map of BaJ and Kovala (48) with the expected patterns. These authors identify and order four (or five) unique fragments with estimated molecular weights in SDS greater than 16,000, whereas the sequence predicts a total of eight peptides. It is likely that the 17-residue amino-terminal peptide, and possibly both of the C-terminal CNBr peptides with 52 and 58 residues and acidic amino acid contents of less than 5%, would not have been resolved on the 6–15% SDS gel system used. On the positive side, the predicted sequence indicates that the head region is nicely broken up into four CNBr fragments that should prove very useful in mapping the actin, tropomyosin, and calmodulin binding domains more exactly using protein biochemical methods.

**Anti-CAL Desmodin Binding Domains—**Sequence comparison has not been particularly informative with respect to the location of the actin and calmodulin binding sites. A straight similarity search of the Protein Identification Resource database with the head region and with peptide III looking for actin and high affinity calmodulin binding sequences has not produced any significant matches that narrow the location of the binding domains to a region smaller than peptide III. This fragment presumably corresponds to the ~20-kDa binding peptide described by Szpacenko and Darbrowska (28) and by Fujii et al. (29). Some speculation is possible; Shirinsky et al. (8) have used changes in intrinsic tryptophan fluorescence to study calmodulin-caldesmon interactions. Peptide III contains three of the five tryptophan residues in caldesmon and both are near the C terminus, suggesting that the calmodulin binding site is C-terminal of the actin binding site. We are in the process of constructing expression vectors to produce caldesmons truncated at their C termini to test this notion.

Recently, Wang (32) has argued that caldesmon can bind two calmodulin molecules and that there is a second low affinity binding site near the amino-terminal cysteine. Wang has confirmed this by isolating and sequencing a CNBr fragment that binds to calmodulin affinity columns; comparison with the protein sequence shows this second calmodulin binding peptide begins at Met$^{16}$ (Footnote 2) and includes the first cysteine residue.

**Tropomyosin Binding Domain—**Graceffa (22) has shown that caldesmon will interact directly with tropomyosin, and Fujii et al. (23) have provided evidence that binding is to the head region. The similarity of the amino terminus of the head region, Glu$^{106}$ to Lys$^{505}$, with a tropomyosin binding region of tropomin T suggests that this is the most likely site for tropomyosin binding. However, this conclusion is not straightforward. Tropomin T has been reported to have two tropomyosin binding sites: one Ca$^{2+}$ sensitive site that binds near Cys$^{190}$ of rabbit $\alpha$-skeletal muscle tropomyosin about two-thirds of the way along the molecule and a Ca$^{2+}$-insensitive site that binds near the C terminus, possibly in the tropomyosin head-to-tail overlap. These two binding domains have been identified by cyanogen bromide and tryptic fragment studies (41, 42). The Ca$^{2+}$-sensitive site is on a trigger fragment, TnT-T2, residues 159–259; the Ca$^{2+}$-insensitive site is on a cyanogen bromide fragment, TnT-CB2, residues 71–151. TnT-CB2 encompasses the region of similarity with caldesmon. Pearlstone and Smillie (41) have re-

---

**Fig. 7. Summary of caldesmon fragments and binding domain data.** This is a diagramatic and, as indicated in the text, somewhat speculative map of the binding domains of smooth muscle caldesmon (CAM). The three fragments produced by cleavage at cysteine residues are designated peptides I–III; their molecular weights are given. The site of cleavage with chymotrypsin (CT) that produces the head region is marked. Note that this is not the only chymotryptic cleavage site in the molecule. The positioning of the various binding sites is discussed in the text. Tm, tropomyosin.

---

2. C. A. Wang, personal communication.
ported that TnT-CB2 does not bind to chicken gizzard or platelet tropomyosin at moderate salt concentrations, but TnT-T2 does. It is unclear therefore whether the Ca\(^{2+}\)-insensitive binding domain of tropomin T, the region with sequence similarity to caldesmon, actually has a binding site on smooth muscle tropomyosin. However, this region appears to be the best candidate for a tropomyosin binding domain in caldesmon. From the binding data and the sequence it is possible to make the simple prediction that peptide III will not bind to tropomyosin if the tropomin T-like sequence is important for binding because the second cysteine group is distal to the similarity.

**Myosin Binding Domain**—Our placement of the myosin binding domain is the most speculative. Several reports have shown that a whole molecule will bind to smooth muscle myosin affinity columns (25–27). Sutherland and Walsh (27) have shown that one of the two smaller fragments, peptides I and III in Fig. 7, obtained by cleavage with 2-nitro-5-thiocy-anobenzoic acid will bind to myosin-Sepharose. As discussed above, Riseman et al. (31) demonstrated that peptide III binds to actin and calmodulin and a comparison of the two papers indicates that the large of the two peptides, peptide I, binds to myosin. The caveat is that Sutherland and Walsh (27) have arrived at a completely different map for peptides I–III that cannot be reconciled with the deduced sequence.

**Phosphorylation Sites and Kinase Activity**—Scott-Woo and Walsh (49) have presented evidence that partially purified caldesmon preparations will autoprophosphorylate. These observations have led to the suggestion that caldesmon is a protein-serine/threonine kinase that can self-phosphorylate in the presence of Ca\(^{2+}\)-calmodulin. We have examined caldesmon with the elongated shape of caldesmon observed in the electron microscope and would suggest that the repetitive region beginning at Arg\(^{466}\)-Glu-Leu-Thr. This potential phosphor-

[References continued...]

1. Sobe, K., Muramoto, Y., Fujita, M., and Kakiuchi, S. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 5652–5655
2. Bretscher, A. (1984) *J. Biol. Chem.* **259**, 12773–12880
3. Ngai, P. K., Carruthers, C. A., and Walsh, M. P. (1984) *Biochem. J.* **218**, 863–870
4. Marston, S. B., and Lehman, W. (1985) *Biochem. J.* **231**, 517–522
5. Furst, D. O., Cross, R. A., DeMey, J., and Small, J. V. (1986) *EMBO J.* **5**, 251–257
6. Clark, T., Ngai, P. K., Sutherland, C., Gr"{o}schel-Stewart, U., and Walsh, M. P. (1986) *J. Biol. Chem.* **261**, 8028–8035
7. Yamazaki, K., Itoh, K., Sobe, K., Mori, T., and Shibata, N. (1987) *J. Biochem. (Tokyo)* **101**, 1–9
8. Shirinsky, V. P., Bushueva, T. L., and Frolova, S. I. (1988) *Biochem. J.* **255**, 203–208
9. Oowada, M. K., Hakura, A., Iida, K., Yahara, I., Sobe K., and Kakiuchi, S. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 3133–3137
10. Sobe, K., Tanaka, T., Kanda, K., Ashino, N., and Kakiuchi, S. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5025–5029
11. Bretscher, A., and Lynch, W. (1985) *J. Cell Biol.* **100**, 1656–1663
12. Burgoyne, R. D., Cheek, T. R., and Norman, K. M. (1986) *Nature* **319**, 68–70
13. Dingus, J., Hwo, S., and Bryan, J. (1986) *J. Cell Biol.* **102**, 1748–1757
14. Podolska, R., and Galazkiewicz, B. (1986) *Biomed. Biochem. Acta* **45**, S153–S158
15. Furst, D. O., Cross, R. A., DeMey, J., and Small, J. V. (1986) *EMBO J.* **5**, 251–257
16. Graceffa, P., Wang, C.-L. A., and Staff, W. F. (1988) *J. Biol. Chem.* **263**, 14194–14202
17. Lynch, W. P., Riseman, V. M., and Bretscher, A. (1987) *J. Biol. Chem.* **262**, 7429–7437
18. Marston, S. B., and Lehman, W. (1985) *Biochem. J.* **231**, 517–522
19. Sobe, K., Morimoto, K., Inui, M., Kanda, K., and Kakiuchi, S. (1982) *Biochem. Res. * **3**, 186–196
20. Smith, C. W. J., Pritchard, K., and Marston, S. B. (1987) *J. Biol. Chem.* **262**, 116–122
21. Pritchard, K. P., and Marston, S. B. (1989) *Biochem. J.* **257**, 893–894...
Molecular Cloning of Caldesmon

44. Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) J. Mol. Biol. 120, 97–120
45. Chou, P. Y., and Fasman, G. D. (1978) Annu. Rev. Biochem. 47, 251–276
46. Takano, E., Maki, M., Mori, H., Hatanaka, M., Marti, T., Tiani, K., Kannagi, R., Ooi, T., and Murachi, T. (1988) Biochemistry 27, 1964–1972
47. Bryan, J. (1989) J. Muscle Res. Cell Motil. 10, 95–96
48. Ball, E. H., and Kovala, T. (1988) Biochemistry 27, 6093–6098
49. Scott-Woo, G. C., and Walsh, M. P. (1989) Biochem. J. 252, 463–472
50. Hanks, S. K. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 388–392
51. Hunter, T., and Cooper, J. A. (1986) in The Enzymes (Boyer, P. D., and Krebs, E. G., eds) pp. 191–246, Academic Press, Orlando, FL
52. Dillon, P. F., Aksoy, M. O., Driska, S. P., and Murphy, R. A. (1981) Science 211, 495–497
53. Lawrence, C. B., Goldman, D. A., and Hood, R. H. (1986) Bull. Math. Biol. 48, 569–583
54. Altschul, S. F., and Erickson, B. W. (1986) Bull. Math. Biol. 48, 617–632