Caldesmon mRNA splicing and isoform expression in mammalian smooth-muscle and non-muscle tissues

Annette M. PAYNE,* Ping YUE,† Kevin PRITCHARD and Steven B. MARSTON‡
Department of Cardiac Medicine, National Heart and Lung Institute, Dovehouse Street, London SW3 6LY, U.K.

The recent determination of the genomic sequence of human caldesmon indicates that eight caldesmon mRNA species could be generated by selection of exon 1 or 1′, exon 3a or 3ab and/or exon 4. We used reverse transcriptase PCR to determine which transcripts were produced in human, rabbit and sheep artery, vein, lung, intestine, kidney and liver. In all tissues the same three transcripts were present: exons 1′-2-3a-5-6...13, exons 1′-2-3a3b-5-6-...13 and exons 1′-2-3a3b-4-5-6...13. Exon 1 was not present and exon 4 was only present when exon 3b was also present. Three protein isoforms of caldesmon can be distinguished by electrophoresis on high-porosity 6% polyacrylamide gel: 130 kDa, 120 kDa and 70 kDa. The 70 kDa isoform lacks the sequence encoded by exon 3b. We investigated whether the two high-molecular-mass isoforms correspond to the presence and absence of exon 4 using an antiserum specific to the sequence encoded by exon 4. Western-blotting and immunoprecipitation experiments showed that both the 130 kDa and the 120 kDa isoforms were expressed with and without the exon 4 sequence. We therefore propose that the molecular-mass heterogeneity arises from additional first exons, possibly with separate promoter regions, which have not yet been characterized in the genomic sequence.

INTRODUCTION

Caldesmon is a calmodulin- and actin-binding protein which forms an integral part of thin myofilaments in smooth-muscle and many non-muscle contractile cells (for reviews see refs. [1]–[3]). It is believed to play a significant role in regulating thin-filament interaction with myosin in smooth-muscle and non-muscle tissue similar to the role of troponin in striated muscles. Two isoform classes have been identified, a high-molecular-mass form, caldesmon*h*, and a lower-molecular-mass form, caldesmon*l* (793 and 538 amino acids respectively from the human cDNA sequence [4]). On SDS/PAGE caldesmon*h* often appears as a closely spaced pair of bands [5–7]. Usually the caldesmon*h* isoform is predominant in smooth-muscle cells, and caldesmon*l* is expressed in non-muscle cells. The relative levels of caldesmon*h* and caldesmon*l* expression change when the smooth-muscle-cell phenotype changes, and this isoform change has been considered to be a marker for smooth-muscle cell differentiation [8–15].

The cDNAs coding for human and chicken caldesmon*h* and caldesmon*l* have been sequenced [4,16,17] revealing that the two isoforms have identical sequences except for the insertion of 229 amino acids in the centre of the molecule. Caldesmon is an extended molecule and the insertion results in an increase in length from 53 nm in caldesmon*l* to 74 nm in caldesmon*h* [18,19]. The insert forms an α-helix which is stable in aqueous solution and has been shown to interact with tropomyosin [20,21]. The remainder of the functional sites, including those involved in tropomyosin-dependent control of the thin filament, are in the invariant regions of the sequence [1,20].

It has been suggested that caldesmon*h* and caldesmon*l* may be produced from a single gene by alternative splicing of the mRNA [4,22,23]. This was recently confirmed by Hayashi et al. [17] who published a sequence from the human caldesmon gene. It was shown that all the known human caldesmon protein isoforms are transcribed from a single gene composed of at least 14 exons mapping to a locus of 7q33-q34. Caldesmon isoform diversity can be generated by the selection of exon 1 or 1′, exon 3a or 3ab and/or exon 4. Exon 1 and 1′ codes for the short N-terminal sequences (17 or 23 amino acids). Exon 3b encodes the 229-amino acid insert present in caldesmon*h*. mRNA transcripts with exon 3a generate caldesmon*l*, whereas exon 3ab encodes caldesmon*h*. Exon 4 codes for 26 amino acids including an extension of the stable α-helical motif. Expression of caldesmon with and without exon 4 might be responsible for the two caldesmon*h* species observed on SDS/PAGE, although all the sequences of caldesmon*h* published to date contain the exon 4 sequence [4,17,24,25]. More recently Haruna et al. [26] have published evidence for at least four possible first exons on analysis of the chicken caldesmon genomic clone.

Very little work has been carried out to determine which of the possible caldesmon transcripts are actually present in human and other mammalian tissues and which protein isoforms are expressed. We have therefore examined a number of smooth-muscle and non-muscle tissues from human, sheep and rabbit by reverse transcriptase PCR (RT-PCR) and have found that only three of the possible eight combinations of exons were actually present. We then raised antiserum specific to the sequence encoded by exon 4. We have determined that the exon combinations observed do not account for the protein isoform diversity observed; in particular the two caldesmon*h* bands seen on SDS/PAGE are not due to the presence or absence of exon 4. Thus still more exons coding for human caldesmon variants remain to be discovered.

MATERIALS AND METHODS

Tissue samples

Samples of human, rabbit and sheep artery, vein, stomach, lung, intestine, kidney and liver tissue were obtained up to 6 h *post mortem*, snap-frozen in liquid nitrogen and stored in liquid

Abbreviation used: RT-PCR, reverse transcriptase PCR.

* Present address: Department of Molecular Genetics, Institute of Ophthalmology, Bath Street, London EC1 9EL, U.K.
† Present address: Cardiology Department, Department of Veterans Affairs Medical Center, 4150 Clement Street, San Francisco, CA 94121, U.S.A.
‡ To whom correspondence should be addressed.

Table 1 Sequences of oligonucleotides used for PCR analysis of caldesmon exon expression

Primer	Sequence 5′ to 3′	Exon specificity
a (sense)	^{359}CGAGCCCGACAGGAACGGCTGCGGCA^{384}	2
b (antisense)	^{2541}TCTCCTGGTCTCAAGTCAGAAGGTTT^{2516}	11,12
c (sense)	^{209}CCCTGAACCTGAAATCACACC^{229}	5′ Untranslated
d (sense)	^{705}GATACGAGATAGAGGAAACAG^{725}	3a
e (antisense)	^{1586}CTTGGAGCTTTTCTCTTTTAGC^{1565}	4
f (antisense)	^{1616}CATCTTTGATCTG^{1625,1537}TTTCTTTAGGA^{1521}	Junction of 3b/5
g (antisense)	^{1616}CATCTTTGATCTG^{1625,850}CTGATTTTCTCCAAC^{832}	Junction of 3a/5
h (sense)	^{43}CGCAGCCTGGCCGCGCTCTCCC^{63}	1
i (sense)	^{273}GGGAGGAGATGCGACTCGAAGC^{294}	1′

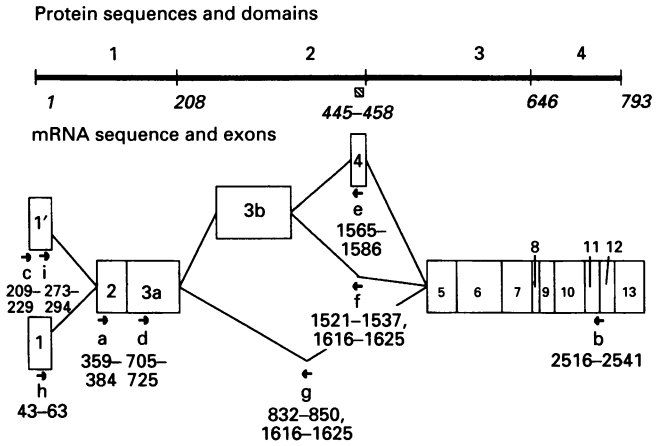

Figure 1 Domain structure of caldesmon*h* [1] related to the exon structure of the human caldesmon gene, based on the genomic sequence published by Hayashi et al. [17]

Amino acid numbers (italics) and nucleotide numbers refer to the human caldesmon cDNA sequence [4]. The short arrows indicate the position and orientation (→, sense; ←, antisense) of the primers used for PCR amplification (Table 1). The peptide sequence from exon 4 is indicated by a hatched box.

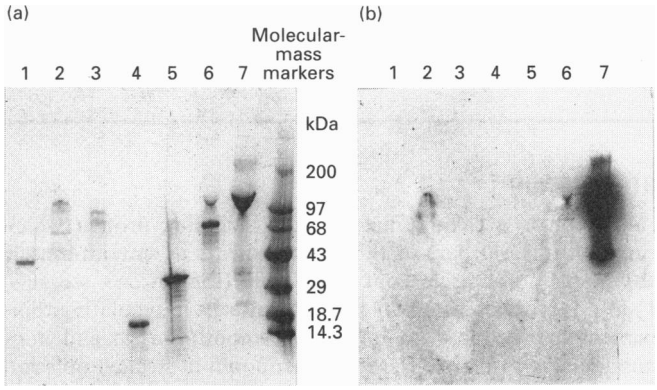

Figure 2 Demonstration of specificity of antiserum to exon 4 protein sequence

(**a**) Caldesmon fragments separated by 8–18% gradient SDS/PAGE and stained with PAGE Blue 83. The exons contained in each fragment are indicated. Lane 1, 0.75 μg of 203–419 exon 3b (central helix); lane 2, 0.25 μg of partially purified sheep aorta caldesmon; lane 3, 0.3 μg of N578 exons 1′, 2 3a, 3b, 4, 5 (chicken); lane 4, 1 μg of N128 exons 1′, 2 (chicken); lane 5, 1 μg of H1 exons 5, 6–13; lane 6, 1 μg of purified caldesmon/ exons 1′,2 3a,5–13; lane 7, 1 μg of purified caldesmon*h* exons 1′,2 3a, 3b, 4, 5–13. (**b**) Autoradiograph from a Western blot of the same gel slab probed with anti-(exon 4) serum and ^{125}I Protein G. The only cross-reactivity is with caldesmon*h* and the caldesmon*h* impurity in the caldesmon/ sample.

nitrogen for up to 2 months. For extraction of protein and mRNA the frozen tissue was ground to a powder under liquid nitrogen using a percussion mortar.

RT-PCR analysis

Total RNA was extracted from 0.3–1 g of the powdered tissue samples with guanidinium isothiocyanate followed by ultra-centrifugation in caesium chloride solution [27]. First-strand cDNA was synthesized from 1 μg of RNA using (dT)12–18 (Pharmacia) primer and Superscript reverse transcriptase (Gibco–BRL) at 42 °C [27]. This cDNA mixture (1 μl) was then used as the starting material for PCR.

PCR primers were designed to detect mRNA coding for the exons expected to be differentially spliced. Their specificities and sequences are given in Table 1 and the positions of the primers relative to the intron/exon map of the genomic sequence are shown in Figure 1. By using different pairs of primers it was possible to amplify unique sequences for each of the theoretical isoform transcripts in a given tissue or library. The optimum Mg^{2+} concentration and annealing temperature were determined empirically for each pair of primers used. For reactions using primers binding within the coding region of caldesmon cDNA a nested PCR approach was used to ensure greater specificity and detection of transcripts present in low concentration. Primers a plus b, h plus b and i plus b (see Table 1 and Figure 1) were used as the first primer pairs for PCR to amplify caldesmon cDNA selectively. Then 1 μl of PCR products was added to 100 μl total volume of the second PCR mixture designed to amplify specific sequences within caldesmon cDNA.

PCR was also performed on a human aorta λgt10 cDNA library (Clontech). cDNA isolated from the clones grown up from the cDNA library (1 μl of a 1:100 dilution) was added directly to the PCR mixture.

The products of the various PCRs were separated by electrophoresis in 2% agarose gel; the DNA fragments were visualized by ethidium bromide and their size was determined by comparison with internal standards (see Figure 3).

Protein extraction

Extraction of caldesmon took advantage of its heat stability [6]. Heat-stable proteins were extracted from frozen tissue samples as described by Lehman et al. [28] with the following revisions; 10 μg/ml leupeptin, 10 μg/ml E64 [*trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane] and 10 μg/ml chymostatin were included during extraction in addition to the standard phenylmethane sulphonyl fluoride to inhibit proteolysis, and the temperature of the homogenates was maintained at 90 °C for 10 min. After the sample has been chilled on ice, the denatured proteins were removed by centrifugation at 47000 *g* for 10 min. Caldesmon was partially purified by precipitation with 65% $(NH_4)_2SO_4$ and dialysis of the water-solubilized pellets against 10 mM citric acid/NaOH buffer, pH 3.0 (caldesmon is acid soluble [29]). The dialysed mixture was centrifuged for 10 min at 47000 *g* leaving caldesmon of greater than 80% purity in the supernatant. The soluble proteins were neutralized to pH 7.0 with 1 M Tris. Protein concentration was determined using the Lowry method. Pure sheep caldesmon*h* and caldesmon*l* were separated from partially pure caldesmon by chromatography on Q Sepharose as described previously [30].

Gel electrophoresis

SDS/PAGE was performed using high-porosity discontinuous gels as described by Cuda et al. [31]. The separating gel contained 6% acrylamide and 0.8% bisacrylamide and the gels were run at 75 V for 3 h. Isoelectric focusing was performed using Pharmacia Immobiline DryPlate gels with an ampholine gradient, range pH 4–7. The gels were rehydrated in 6 M urea/2% CHAPS/2% carrier ampholines pH 4–6.5 and used according to the manufacturer's instructions. Western immunoblotting was performed by the method of Towbin et al. [33] on nitrocellulose membranes. ^{125}I-labelled Protein G (Amersham) was used to detect primary antiserum binding.

Antiserum production

Antibodies were raised against a unique amino acid sequence in exon 4 of human caldesmon (see Figure 1). The peptide ^{445}Gln-Ala-Lys-Arg-Glu-Lys-Leu-Gln-Glu-Asp-Lys-Pro-Thr-Phe^{458} linked to the ϵ-amino groups of a polylysine carrier peptide (molecular mass 17626 Da) was synthesized by Alta Biochemicals (Birmingham, U.K.). Rabbits were immunized by fortnightly repeated subcutaneous injection of 0.5 mg of the peptide emulsified in Freund's complete adjuvant for the primary injection and Freund's incomplete adjuvant for subsequent boosters. All the results reported here were obtained with a single batch of serum diluted 1:10. We also used a pan-caldesmon antiserum which was raised against pure sheep caldesmon*h* by Dr. C. Redwood [34].

The specificity of the antiserum for exon 4 was demonstrated by Western blotting of human and chicken caldesmon and recombinant caldesmon fragments separated by SDS/PAGE (Figure 2). The antiserum cross-reacted with human caldesmon*h*, which is the only caldesmon species known to contain the sequence encoded by exon 4. It did not cross-react with human fragments containing exons 1′,2,3a,3b,5,6,7–13, nor with N578 derived from a chicken cDNA clone [20] the exon 4 sequence of which is substantially different from human caldesmon [4]. In addition, antiserum binding to caldesmon*h* was competed out when the peptide used to raise the antiserum was added to the primary incubation solution during the immunoblotting procedure.

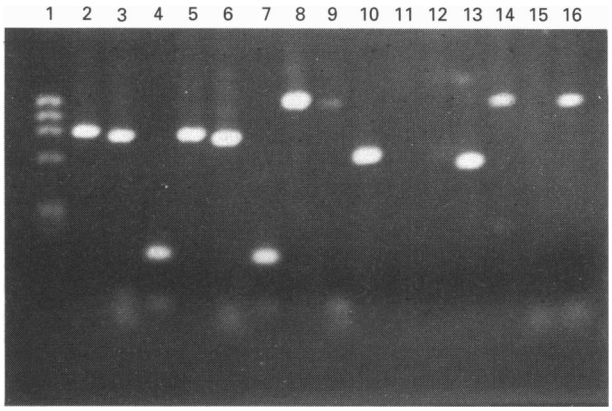

Figure 3 Electrophoresis of PCR products on 2% agarose gel

The primers used and the exons detected are given in Table 2.

RESULTS

Exons present in caldesmon mRNA

Exons 2,3a,5,6,7,8,9,10,11,12 and 13 are reported to be invariant in human and chicken mRNA, whereas exons 3b and 4 may be retained or excluded from the transcript. The first exon is believed to exist in at least two variants (exon 1 and 1′ in human mRNA) each possessing separate promoters [17,26] (Figure 1). RNA preparations from a number of human, rabbit and sheep smooth muscles and non-muscle tissue were analysed by RT-PCR for the presence or absence of exons 1,1′,3a,3b and 4. The primers used (Table 1, Figure 1) are based on the human genomic sequence published by Hayashi et al. [17]; however, they worked equally well with human, rabbit and sheep mRNA, indicating strong conservation of sequence between species [26].

The human aorta cDNA library was examined for the presence of exons 1 and 1′. PCR with primers i and g, i and e or i and f yielded products of the predicted size containing exon 1′ (Figure 3 lanes 13, 14 and 16). PCR with primers h and g, h and e or h and f yielded no product (Figure 3, Lanes 11, 12 and 15 and Table 2); thus exon 1 was not present in any transcript. As no products corresponding to the amplification of the cDNA clones containing exon 1 were detected in the human aorta λgt10 library, the majority of tissue samples were not tested for expression of this exon.

PCR using primer c, which is located in the 5′ untranslated region of exon 1′, and e, f or g (Figure 3, lanes 8, 9 and 10) yielded single bands on the agarose gel corresponding to the transcripts with 5′ exons 1′-2-3a etc. There was no evidence for any exon insertion or deletion in this region of the mRNA.

mRNA extracted from tissue was amplified by RT-PCR using the primer pair a and b to select out cDNA containing caldesmon exons 2–13. The cDNA was then subjected to further PCR using primer pairs d and e, which amplifies cDNA containing exon 4, d and f, which amplifies cDNA containing exon 3b but not exon 4, and primers d and g, which amplifies cDNA containing neither exon 4 nor exon 3b (lanes 2–7 in Figure 3). cDNA coding for exon combinations -3a-3b-4-5-, -3a-3b-5- and -3a-5- was detected in all the tissue samples (see Table 2).

Exon 4 was only present in cDNA that also contained exon 3b, coding for caldesmon*h* isoforms. Caldesmon*l* mRNA containing exon 4 (i.e. -3a-4-5-) should give a 252 bp RT-PCR product with primers d and e (lanes 2 and 5 in Figure 3), but none was observed. The positive result from amplification of -3a-3b-4-5-

Table 2 Detection of caldesmon mRNA transcripts on 2% agarose–gel electrophoresis after PCR.

The results shown in Figure 3 are analysed.

Lane	Source of DNA	Primers	Size of PCR product (bp)	Exon transcripts detected
1	ϕX174 RF DNA, *Hea*III digest	–	–	–
2	Human aorta RNA	a,b then d,e	882	3a,3b,4
3	Human aorta RNA	a,b then d,f	843	3a,3b,5
4	Human aorta RNA	a,b then d,g	156	3a,5
5	Human aorta RNA	a,b then d,e	882	3a,3b,4
6	Human aorta RNA	a,b then d,f	843	3a,3b,5
7	Human aorta RNA	a,b then d,g	156	3a,5
8	cDNA library (human aorta)	c,e	1378	1′ (5′ untranslated)...3b,4
9	cDNA library (human aorta)	c,f	1338	1′ (5′ untranslated)...3b,5
10	cDNA library (human aorta)	c,g	653	1′ (5′ untranslated)...3a,5
11	cDNA library (human aorta)	h,g		None
12	cDNA library (human aorta)	h,e		None
13	cDNA library (human aorta)	i,g	588	1′.....3a,5
14	cDNA library (human aorta)	i,e	1314	1′.....3a,3b,4
15	cDNA library (human aorta)	h,f		None
16	cDNA library (human aorta)	i,f	1275	1′.....3a,3b,5

Table 3 Caldesmon-encoding mRNA species in human, rabbit and sheep tissues

+, Transcript present; −, transcript absent.

	Caldesmon*h*		Caldesmon*l*	
Tissue (no. of specimens)	1′+3a3b-4	1′+3a3b	1′+3a-4	1′+3a
Rabbit				
Aorta (1)	+	+	−	+
Carotid (3)	+	+	−	+
Stomach (1)	+	+	−	+
Trachea (1)	+	+	−	+
Small intestine (1)	+	+	−	+
Lung (1)	+	+	−	+
Skeletal muscle (1)	+	+	−	+
Kidney (1)	+	+	−	+
Sheep				
Aorta (3)	+	+	−	+
Femoral artery (1)	+	+	−	+
Human				
Aorta (6)	+	+	−	+
Umbilical cord (4)	+	+	−	+
Liver (1)	+	+	−	+
Kidney (1)	+	+	−	+
Oesophagus (1)	+	+	−	+
Carotid (2)	+	+	−	+
Femoral artery (1)	+	+	−	+
Mesenteric artery (1)	+	+	−	+
Renal vein (1)	+	+	−	+
Pulmonary artery (1)	+	+	−	+

(882 bp) in the same lanes rules out any defect in the experimental protocol. Thus the PCR products derived from transcripts coding for caldesmon*l* (-3a-5-) lacked exon 4 as well as exon 3b.

PCR with primers d and e and d and f produced products indicating that transcripts containing exons 3b-4 and transcripts containing only exon 3b were present (Figure 3, lanes 2, 3, 5, 6). This is the first evidence of mRNA diversity within the caldesmon*h* class.

Table 3 summarizes the results of RT-PCR on mRNA from a variety of species and tissues. The pattern of caldesmon transcripts was the same in all three species and did not differ between gut and vascular smooth muscles. The presence of caldesmon*h* mRNA in the non-muscle tissue samples may represent transcripts derived from vascular tissue contained within them.

Analysis of protein isoform expression

Caldesmon separates into high- and low-molecular-mass isoform classes, with the high-molecular-mass isoforms frequently ap-

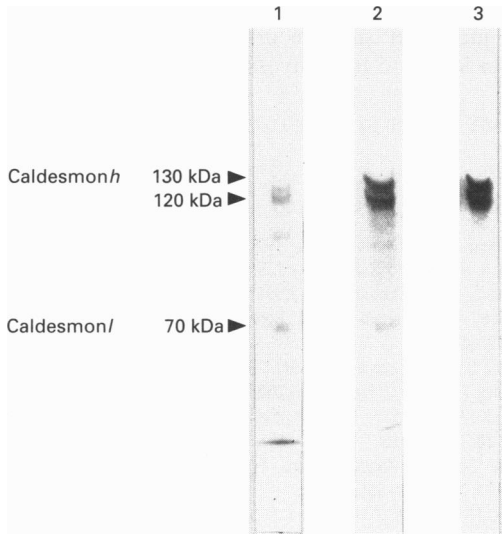

Figure 4 Electrophoresis on high-porosity SDS/6% polyacryalmide gel of sheep aorta caldesmon

Partially purified caldesmon (mixed isoform; 1 μg per lane) was analysed. Lane 1, stained with PAGE Blue 83; lane 2, immunoblot probed with anti-pan-caldesmon serum; lane 3, immunoblot probed with anti-(exon 4) serum. In each lane two caldesmon*h* bands are clearly visible. The caldesmon*l* band is recognized by the anti-pan-caldesmon serum in lane 2 but not by the anti-(exon 4) serum in lane 3.

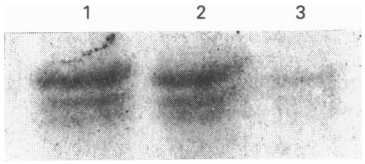

Figure 5 Immunoprecipitation experiments

A solution of 20 μg of caldesmon (2 mg/ml) was incubated with 100 μl of Protein G–Sepharose (Pharmacia) and sedimented to remove any proteins that bound non-specifically to the protein G. The caldesmon was then incubated and precipitated twice with 100 μl of Protein G–Sepharose + 100 μl of anti-(exon 4) serum. The protein precipitated from three steps is shown in Lane 1. The depleted caldesmon solution was then incubated and precipitated with 100 μl of Protein G–Sepharose + 100 μl of anti-pan-caldesmon serum (shown in lane 2). The precipitated proteins were separated on high-porosity 0.1% SDS/6% polyacrylamide gels and stained with PAGE Blue 83. Lane 1, caldesmon immunoprecipitated with anti-(exon 4) serum; lane 2, caldesmon immunoprecipitated with anti-pan-caldesmon serum; lane 3, protein immunoprecipitated with preimmune serum. All the caldesmon*h* bands are present in both the immunoprecipitates. Little caldesmon*h* binds to the Protein G–Sepharose non-specifically (lane 3).

pearing as a closely spaced pair of bands on high-resolution SDS/PAGE. The use of 6% polyacrylamide/0.08% bisacrylamide (high-porosity) gels enabled us to separate clearly the bands of the high-molecular-mass isoform with both purified caldesmon*h* and semi-pure tissue extracts (Figure 4). The bands ran at 130 kDa and 120 kDa. In some samples a band of intermediate mobility was seen (e.g. Figure 5). As this was variable, we suspect it to be a breakdown product of the higher band. Only one band corresponding to caldesmon*l* was seen (70 kDa).

When these gels were immunoblotted the pan-caldesmon antiserum detected two caldesmon*h* bands and one caldesmon*l* band (Figure 4, lane 2). The antiserum specific to the exon 4 sequence (see Figure 2) also detected two caldesmon*h* bands but did not detect any exon 4 sequence in caldesmon*l* (Figure 4, lane 3). The latter result is in accord with the RT-PCR results (Tables 2 and 3); however, the RT-PCR also predicted the presence of two caldesmon*h* isoforms, one with and one without exon 4 (Figure 3, Table 2). As both the caldesmon*h* bands detected by SDS/PAGE contain exon 4, the alternative splicing pattern suggested by Hayashi et al. [17] (Figure 1) cannot account for them.

To investigate this result further we performed an immunoprecipitation experiment (Figure 5). Protein was precipitated from a partially purified sheep aorta caldesmon solution in two steps: first using the antiserum to exon 4 and then a pan-caldesmon antiserum. The PAGE Blue 83-stained gel shows that the same pattern of caldesmon*h* bands was seen in the protein precipitated by anti-(exon 4) serum as in the remaining caldesmon when precipitated by anti-pan-caldesmon serum. In control immunoblots immunoreactive exon 4 was present in the samples precipitated by anti-(exon 4) serum and absent from the samples precipitated by anti-pan-caldesmon serum from the supernatant remaining after precipitation with anti-(exon 4) serum (results not shown).

These results demonstrate that each caldesmon*h* band seen on SDS/PAGE contains at least two isoforms: one with and one without exon 4. Thus there are at least four forms of caldesmon*h* present: 120 kDa and 130 kDa bands that contain exon 4 and 120 kDa and 130 kDa bands that do not contain exon 4.

DISCUSSION

The genomic sequence of human caldesmon described by Hayashi et al. [17] showed that its mRNA is made up of 13 exons. The same exon pattern was found for the chicken caldesmon gene [26]. The exon boundaries appear to correspond to the boundaries between protein domains previously identified (see Figure 1) [1]. The myosin- and tropomyosin-binding domain 1 corresponds to exons 1, 2 and 3a [20,35,36]. Domain 2, the tropomyosin-binding helical sequence [20,21], corresponds to exons 3b and 4. Domain 3 corresponds to exons 5, 6 and 7, and domain 4, which contains all the actin- and calmodulin-binding sites [1,20,37], is coded for by exons 8–13.

Hayashi and others [4,17,22] showed that alternative splicing could result in inclusion or exclusion of exons 3b and 4 and could account for the two major isoforms of caldesmon, caldesmon*h* and caldesmon*l*. This heterogeneity plus alternative promoters in exon 1 and 1′ gives eight possible isoforms of caldesmon (Figure 1) [17]. Hayashi's group did not determine which of these isoforms was expressed in human tissue. We have examined the expression of caldesmon in smooth-muscle and non-muscle tissues of human, sheep and rabbit at the mRNA and protein levels and found that the pattern of expression is in fact quite restricted.

We detected mRNA containing exon 1′ in all the tissue samples and in the aorta cDNA library (Figure 3, Table 1); we did not detect any exon 1. This agrees with the published cDNA sequences of caldesmon from smooth-muscle and non-muscle sources, all of which contain the amino acid sequence encoded by exon 1′ [4,5,22–25].

The main diversity in caldesmon isoform expression is between the low-molecular-mass caldesmon*l* and the high-molecular-mass caldesmon*h*. The difference between caldesmon*h* and caldesmon*l* is the presence of exon 3b in the former, which codes for a stable α-helical 229-amino acid peptide [4,23] (see Figure 1). Regulation of expression is thought to depend on selection of the two 5′ splice sites within exon 3 [17] (Figure 1). We confirmed this and demonstrated that caldesmon-encoding mRNAs with and

without exon 3b were present in all tissues (Figure 3, Tables 2 and 3). We also found that exon 4 was only present in transcripts that also contained exon 3b. The caldesmon*h* isoform is predominant in smooth-muscle cells, whereas caldesmon*l* predominates in non-muscle cells. Changes in expression of the two major caldesmon isoforms are characteristic of smooth-muscle-cell phenotypic change and have been considered to be markers of cell differentiation [8–15]. On the other hand the pattern of expression of caldesmon*h* with and without exon 4 was invariant (Table 3).

On SDS/PAGE caldesmon*h* usually showed a pair of bands [5–7] which could be better separated in high-porosity 6% gels (Figure 4). The apparent mass of these bands was 120 and 130 kDa, corresponding to a difference of 80–90 amino acids. Caldesmon*l* was always a single 70 kDa band. One likely source of heterogeneity is the inclusion or exclusion of exon 4 which codes for 26 amino acids, by alternative splicing of mRNA (Figure 1) [17]. We did, indeed, find that caldesmon*h* mRNA was obtained with and without exon 4 but caldesmon*l* mRNA never contained exon 4 (Figure 2, Table 1). It appeared possible that the presence or absence of exon 4 could account for the two caldesmon*h* bands observed on SDS/PAGE with 6% gels, whereas the absence of exon 4 from caldesmon*l* transcripts could account for the single band. In fact this proved not to be the case. Direct Western blotting with an antiserum specific to the sequence encoded by exon 4 (Figure 4), confirmed by immunoprecipitation (Figure 5), showed that all the bands of caldesmon*h* on SDS/polyacrylamide gels contain isoforms with and without the sequence encoded by exon 4.

What is the origin of the molecular-mass heterogeneity of caldesmon*h*? It is not accounted for in the model derived from the genomic sequence of human caldesmon [17] (Figure 1). PCR amplification using primers c or h and b, e, f or g, failed to show any evidence of multiple mRNAs with differences in any of the known exons 1′,2–13 (Figure 3). The heterogeneity clearly arises from either additional exons inserted 5′ to exon 1′ or additional first exons with separate promoter regions. Unfortunately this part of the human caldesmon genomic sequence is not well defined [17]. However, a recent paper on the chicken genomic sequence proposed additional 5′ exons and two promoter regions [26], indicating potential for molecular-mass heterogeneity. Considerable further work is required to define this region of the caldesmon gene.

Two promoters and the inclusion or exclusion of exon 4 could account for up to four variants of caldesmon*h* and two variants of caldesmon*l*. Additional forms of heterogeneity are required to account for the five or more isoelectric variants of caldesmon*l* and the nine variants of caldesmon*h* observed to separate in the isoelectric focusing dimension of two-dimensional gels [10,28].

The function of the different isoforms is obscure. There seems to be no tissue-specific variation in isoform expression other than the well-documented expression of caldesmon*h* in smooth muscle and caldesmon*l* in embryonic and non-muscle cells [14,15] (Table 3). The major functional properties of caldesmon are located in invariant exons (actin and calmodulin binding in exons 8–13, myosin binding in exons 2,3 and 5–8 [1,20]). It has been noted that the protein sequence encoded by exon 3b is a tropomyosin-binding region [1] and it has been suggested that the length of the caldesmon molecule may be matched to the length of the tropomyosin molecule. Non-muscle caldesmon and tropomyosin molecules are shorter than smooth-muscle caldesmon and tropomyosin molecules [20]. Perhaps the two calesmon*h* isoforms of different molecular size are involved in precise matching of caldesmon length to tropomyosin length.

This work was supported by a grant from the Wellcome Trust.

REFERENCES

1 Marston, S. B. and Redwood, C. S. (1991) Biochem. J. **279**, 1–16
2 Sobue, K. and Sellers, J. R. (1991) J. Biol. Chem. **266**, 12115–12118
3 Yamashiro, S. and Matsumura, F. (1991) Bioessays **13**, 563–568
4 Humphrey, M. B., Herrera-Sosa, H., Gonzalez, G., Lee, R. and Bryan, J. (1992) Gene **112**, 197–205
5 Marston, S. B. and Lehman, W. (1985) Biochem. J. **231**, 517–522
6 Bretscher, A. (1984) J. Biol. Chem. **259**, 12873–12880
7 Sobue, K., Muramoto, Y., Fujita, M. and Kakiuchi, S. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 5652–5655
8 Ueki, N., Sobue, K., Kanda, K., Hada, T. and Higashino, K. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 9049–9053
9 Tanaka, J., Watanabe, T., Nakamura, N. and Sobue, K. (1993) J. Cell Sci. **104**, 595–606
10 Glukhova, M. A., Kabakov, A. E., Frid, M. G., Ornatsky, O. I., Belkin, A. M., Mukhin, D. N., Orekov, A. N., Koteliansky, V. E. and Smirnov, V. N. (1988) Proc. Natl. Acad. Sci. U.S.A. **85**, 9542–9546
11 Glukhova, M. A., Kabakov, A. E., Ornatsky, O. I., Vasilevskaya, T. D., Koteliansky, V. E. and Smirnov, V. N. (1987) FEBS Lett. **218**, 292–294
12 Glukhova, M. A., Frid, M. G. and Koteliansky, V. E. (1990) J. Biol. Chem. **265**, 13042–13046
13 Reckless, J., Pritchard, K., Marston, S. B., Fleetwood, G. and Tilling, L. (1994) Atheroscler. Thromb., in the press
14 Reckless, J., Pritchard, K. and Marston, S. B. (1994) Biophys. J. **66**, A197 (abstr.)
15 Frid, M. G., Shekhonin, B. V., Koteliansky, V. E. and Glukhova, M. A. (1992) Dev. Biol. **153**, 185–193
16 Novy, R. E., Lin, J. L.-C. and Lin, J. J.-C. (1991) J. Biol. Chem. **266**, 16917–16924
17 Hayashi, K., Yano, H., Hashida, T., Takeuchi, T., Takeda, O., Asada, K., Takahashi, E., Kato, I. and Sobue, K. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 12122–12126
18 Mabuchi, K. and Wang, C.-L. A. (1991) J. Muscle Res. Cell Motil. **13**, 146–151
19 Stafford, W. F., Jancso, A. and Graceffa, P. (1990) Arch. Biochem. Biophys. **281**, 66–69
20 Redwood, C. S. and Marston, S. B. (1993) J. Biol. Chem. **268**, 10969–10976
21 Wang, C.-L. A., Chalovich, J. M., Graceffa, P., Lu, R. C., Mabuchi, K. and Stafford, W. F. (1991) J. Biol. Chem. **266**, 13958–13963
22 Bryan, J. and Lee, R. (1991) J. Muscle Res. Cell Motil. **12**, 372–375
23 Hayashi, K., Fujio, Y., Kato, I. and Sobue, K. (1991) J. Biol. Chem. **266**, 355–361
24 Hayashi, K., Kanda, K., Kimizuka, F., Kato, I. and Sobue, K. (1989) Biochem. Biophys. Res. Commun. **164**, 503–511
25 Bryan, J., Imai, M., Lee, R., Moore, P., Cook, R. G. and Lin, W. G. (1989) J. Biol. Chem. **264**, 13873–13879
26 Haruna, M., Hayashi, K., Yano, H., Takeuchi, O. and Sobue, K. (1993) Biochem. Biophys. Res. Commun. **197**, 146–153
27 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1986) Molecular Biology: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
28 Lehman, W., Denault, D. and Marston, S. B. (1993) Biochim. Biophys. Acta **1203**, 53–59
29 Smith, C. W., Pritchard, K. and Marston, S. B. (1987) J. Biol. Chem. **262**, 116–122
30 Marston, S. B., Pinter, K. and Bennett, P. M. (1992) J. Muscle Res. Cell Motil. **13**, 206–218
31 Cuda, G., Fananazapir, L., Zhu, W. S., Sellers, J. R. and Epstein, N. D. (1993) J. Clin. Invest. **91**, 2861–2865
32 Reference deleted
33 Towbin, H., Staehelin, T. and Gordon, J. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4350–4354
34 Redwood, C. S. (1992) Identification of the Functional Domains of Caldesmon, Ph.D. Thesis, University of London
35 Hemric, M. E. and Chalovich, J. M. (1990) J. Biol. Chem. **265**, 19672–19678
36 Bogatcheva, N. V., Vorotnikov, A. V., Birukov, K. G., Shirinsky, V. P. and Gusev, N. B. (1993) Biochem. J. **290**, 437–442
37 Szpacenko, A. and Dabrowska, R. (1986) FEBS Lett. **202**, 182–186

Received 28 April 1994/22 August 1994; accepted 26 August 1994