Purification of Human Smooth Muscle Filamin and Characterization of Structural Domains and Functional Sites†

Rick S. Hock,‡ Gary Davis,§ and David W. Speicher*,‡

Protein Chemistry, The Wistar Institute, 3601 Spruce Street, Philadelphia, Pennsylvania 19104, and Molecular Therapeutics, Inc., 400 Morgan Lane, West Haven, Connecticut 06516

Received March 8, 1990; Revised Manuscript Received May 18, 1990

ABSTRACT: A method was developed to purify human smooth muscle filamin in high yield and structural domains were defined by using mild proteolysis to dissect the molecule into intermediate-sized peptides. Unique domains were defined and aligned by using high-resolution peptide mapping of iodinated peptides on cellulose plates. The amino- and carboxyl-terminal orientation of these domains within the molecule was determined by amino acid sequence analysis of several aligned peptides. In addition to the three unique domains which were identified, a number of smaller and larger fragments were also characterized and aligned within the intact molecule. These structural domains and related peptides provide a useful set of defined fragments for further elucidation of structure–function relationships. The two known functionally important binding sites of filamin, the self-association site and the actin-binding site, have been localized. Self-association of two monomers in a tail-to-tail orientation involves a small protease-sensitive region near the carboxyl terminal of the intact polypeptide chain. Sedimentation assays indicate that an actin-binding site is located near the blocked amino terminal of the filamin molecule. Sequences derived from large peptides mapping near the amino terminal show homology to the amino-terminal actin-binding site of α-actinin (chicken fibroblast and *Dictyostelium*), *Dictyostelium* 120-kDa actin gelation factor, β-spectrin (human red cell and *Drosophila*), and human dystrophin. This homology is particularly interesting for two reasons. The functional form of filamin is single stranded, in contrast to α-actinin and spectrin which are antiparallel double-stranded actin cross-linkers. Also, no homology to the spectrin-like segments which comprise most of the mass of spectrin, α-actinin, and dystrophin was found. Instead, the sequence of a domain located near the center of the filamin molecule (tryptic 100-kDa peptide, T100) shows homology to the published internal repeats of the *Dictyostelium* 120-kDa actin gelation factor. On the basis of these results, a model of human smooth muscle filamin substructure is presented. Also, comparisons of human smooth muscle filamin, avian smooth muscle filamin, and human platelet filamin are reported.

Microfilaments are complex structural arrays that are involved in diverse cellular processes. They are modulated by many actin-associated proteins that influence linear actin assembly or affect the three-dimensional architecture of actin filaments. Included in the former category are proteins that control actin monomer polymerizability or nucleation, block filament ends, or sever actin filaments. In the latter category are proteins that cross-link actin oligomers or organize actin filaments into loose bundles, tight bundles, or orthogonal networks [reviewed in Stossel (1984, 1989), Frieden (1985), Hartwig et al. (1985), and Pollard and Cooper (1986)].

Filamin, a major component of the microfilament network, plays a critical role in determining the three-dimensional arrangement of actin filaments in vivo and in vitro [reviewed in Weihing (1985)]. Filamin is a long flexible end-to-end homodimer (M_r 500 000) that contains mainly β-strand conformation (Tyler et al., 1980; Koteliansky et al., 1982). This actin cross-linking protein was first isolated from avian smooth muscle (Wang et al., 1975; Shizuta et al., 1976; Wang, 1977) and has been purified from platelets (Rosenberg et al., 1981), macrophages (Hartwig & Stossel, 1975), guinea pig vas deferens (Wallach et al., 1978), baby hamster kidney (BHK-21) cells (Schloss & Goldman, 1979), thyroid gland (Roustan et al., 1982), HeLa cells (Weihing, 1983; Weihing & Franklin, 1983), and motile *Dictyostelium* amoebae (Hock & Condeelis, 1987; Condeelis et al., 1988). Filamin is now known to exist in most cell types as a family of isoforms. The number of isoforms has not been defined, although multiple isoforms are apparently present within a single human cell type, HeLa cells (Mangeat & Burridge, 1983). Also, isoform switching occurs during myogenesis (Gomer & Lazarides, 1981, 1983).

Filamin gels actin filaments at low molar ratios of filamin dimer to actin monomer (Brotschi et al., 1978) and can organize actin filaments in vitro into either parallel bundles at high filamin:actin ratios or three-dimensional networks at lower ratios (Wang & Singer, 1977; Sobue et al., 1982). Self-association of filamin monomers to form actin cross-link-competent dimers is a critical requirement for filamin function. The reversible self-association of avian smooth muscle filamin has been demonstrated (Wang et al.,1975; Wang, 1977; Davies et al., 1980), and the two HeLa cell filamin isoforms have different association properties since one form is predominantly monomer and the other is mostly dimer as isolated (Mangeat & Burridge, 1983). Mammalian smooth muscle (Wallach et al., 1978a,b) and platelet filamins (Carroll & Gerrard, 1982; Cox et al., 1984) are phosphorylated in vivo by cAMP-dependent protein kinase. Phosphorylation by protein kinase C has also been reported (Fuchiwaki, 1987). Filamin may be

† This work was initiated by G.D. and D.W.S. in the Pathology Department, Yale University School of Medicine, with partial support from BRSG Grant RR 05358 (Yale University). This work has also been supported by NIH Grants AR 39158, HL 38794, CA 10815, and RR 05540 (The Wistar Institute).

* Author to whom correspondence should be addressed.

‡ The Wistar Institute.

§ Molecular Therapeutics, Inc.

0006-2960/90/0429-9441$02.50/0

associated with the cell membrane under some conditions in platelets (Okita et al., 1985), macrophages (Hartwig & Shevlin, 1986), and *Dictyostelium* amoebae (Hock & Condeelis, 1987; Condeelis et al., 1988).

The avian smooth muscle filamin isoform has been partially characterized (actin-binding, self-association, hydrodynamic, and morphologic properties), while mammalian filamins have been less extensively characterized (Weihing, 1985). To date, a linear alignment of structural domains within the intact molecule has not been presented, and a definitive assignment of functional sites (actin-binding and self-association) has not been made.

In order to generate a linear domain alignment of a human filamin isoform, human smooth muscle filamin was purified. Unique intermediate-sized structural domains were produced and aligned into a linear sequence within the intact human smooth muscle filamin molecule. The actin-binding site has been assigned to the amino-terminal end of filamin, while the self-association site has been assigned to the carboxyl-terminal end. Evidence is presented that indicates that human smooth muscle filamin is comprised of repetitive structure homologous to that described for the *Dictyostelium* 120-kilodalton (kDa)[1] gelation factor (Noegel et al., 1989) and for human endothelial cell actin-binding protein (Gorlin et al., 1989a,b). A model of human smooth muscle filamin substructure is presented which provides a basis for the investigation of the regulation of the functional sites. The human smooth muscle filamin isoform has been compared to avian smooth muscle and human platelet filamins in order to help define species- and tissue-specific isoform diversity.

Experimental Procedures

Filamin Purification. Human smooth muscle filamin was purified essentially by the procedure of Wang (1977) with several critical modifications. Fresh or frozen uterine tissue was obtained from the Yale Medical School Surgical Pathology Department. The endothelial lining was scraped off in isotonic buffer, and any large leiomyomas (nonmalignant tumors common in perimenopausal uterine smooth muscle) were removed. The tissue was then minced into 5-mm cubes with a scissors, washed once in high-salt extraction buffer [50 mM potassium phosphate buffer, 0.6 M KCl, 10 mM ethylenediaminetetraacetic acid (EDTA), 1 mM dithiothreitol (DTT), 0.1 mM diisopropyl fluorophosphate (DFP), and 2 μg/mL leupeptin, pH 7.5], and homogenized in extraction buffer using a Polytron (Brinkmann Instruments, Co.) at full speed for 2–3 s. The homogenate was squeezed through two layers of cheesecloth. The solid material remaining on the cheesecloth was sequentially rehomogenized and filtered 2 times as above to improve yield. The high-salt extracts were combined, and the pool was clarified at 95000g for 20 min. The supernatant was dialyzed against 20 volumes of 10 mM EDTA, 0.1 mM DFP, and 2 μg/mL leupeptin, pH 7.5, and precipitated myosin was removed at 70000g for 20 min. Crude filamin was precipitated from the supernatant with 40% ammonium sulfate and collected at 15000g for 15 min. The pellet was gently homogenized in a small volume of buffer G [0.3 M KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 5 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), 10 mM sodium phosphate buffer, 0.005% azide, 5 mM β-mercaptoethanol, 0.1 mM DFP, and 2 μg/mL leupeptin, pH 7.8] and centrifuged at 215000g for 30 min. The supernatant was applied to a Sepharose CL-4B (Pharmacia) column (5 × 90 cm) and eluted with buffer G. Pooled fractions were dialyzed against 20 mM Tris–acetate, 1 mM EDTA, 0.1 mM DFP, and 2 μg/mL leupeptin, pH 7.2, and clarified at 215000g for 30 min.

The supernatant was fractionated by high-performance liquid chromatography (HPLC) using a preparative polymer-based ion-exchange column (Bio-Rad, Bio-Gel TSK DEAE-5-PW, 7.5 × 150 mm) and a nonlinear salt gradient developed to maximize separation of filamin from other components. Buffer A was 20 mM Tris–acetate, 1 mM EDTA, and 2 μg/mL leupeptin, pH 7.2, and buffer B was 20 mM Tris–acetate, 1 mM EDTA, 0.5 M KCl, and 2 μg/mL leupeptin, pH 7.2. Sample was injected onto the DEAE column equilibrated at 0% B. Proteins were eluted at a flow rate of 5 mL/min at room temperature as follows: 10-min hold at 0% B, step to 25% B, 45-min hold at 25% B, 60-min linear gradient from 25% B to 40% B. Filamin eluted as a single symmetrical peak at 33% B. The column was washed with 100% B prior to reequilibration. Purified fractions were concentrated by vacuum dialysis and were stored up to 2 months at 0 °C in isotonic buffers. All steps except the HPLC separation were performed at 0–4 °C, and the presence of 0.1 mM DFP (Aldrich Chemical Co., Inc.) and 2 μg/mL leupeptin (Sigma) in all solutions was essential to minimize proteolysis.

Avian smooth muscle filamin was purified from fresh chicken gizzards according to the procedure of Wang (1977).

Polyacrylamide Gel Electrophoresis. One-dimensional sodium dodecyl sulfate (SDS) gel electrophoresis was performed according to the procedure of Laemmli (1970) using 1.5 × 110 mm slab gels. Either uniform or gradient gels were used as indicated in the text.

Two-dimensional polyacrylamide gels were prepared by the method of O'Farrell (1975) with modifications as previously described (Speicher et al., 1982). Briefly, 110 mm long isofocusing gels were prepared in glass tubes (125 × 3.0 mm). After preelectrophoresis, samples were loaded and electrophoresed for 16 h at 340 V. Gels were removed and placed in 4 mL of equilibration buffer (3% SDS, 4 mM EDTA, 10% glycerol, 2.5% β-mercaptoethanol, and 0.01% bromophenol blue) for 10 min at 37 °C and then electrophoresed in the second dimension on slab gels (3 × 110 mm) prepared by the method of Laemmli (1970) that contained linear acrylamide gradients from 3 to 18%.

Mild Proteolytic Digestion of Filamin. Samples of human and avian smooth muscle filamin were dialyzed into 20 mM Tris, 0.02% azide, and 1 mM β-mercaptoethanol, pH 7.6, and digested at 0 °C with trypsin [L-1-(tosylamino)-2-phenylethyl chloromethyl ketone treated trypsin, Worthington]. Intermediate-sized peptides representative of the filamin domains were produced by using an enzyme-to-substrate ratio (w/w) ranging from 1:20 to 1:100 and incubation times ranging from 1 to 24 h. Specific conditions are indicated in the text. Enzymatic hydrolysis was terminated by the addition of 1 mM DFP or by boiling in SDS gel sample buffer (6% SDS, 4 M urea, 125 mM Tris-HCl, 4 mM EDTA, and 3% β-mercaptoethanol, pH 6.9).

Chemical Hydrolysis at Cysteine Residues. Samples of human and avian smooth muscle filamin were cleaved at cysteine residues with 2-nitro-5-thiocyanobenzoic acid (NTCB) in 200 mM Tris-HCl, 7 M guanidine hydrochloride, and 1 mM

[1] Abbreviations: DFP, diisopropyl fluorophosphate; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; FPLC, fast protein liquid chromatography; HPLC, high-performance liquid chromatography; kDa, kilodalton(s); NTCB, 2-nitro-5-thiocyanobenzoic acid; PIR, protein identification resource; PTH, phenylthiohydantoin; SDS, sodium dodecyl sulfate; Tris, tris(hydroxymethyl)aminomethane.

EDTA, pH 8.0. The reaction was initiated by the addition of NTCB at 10 times the cysteine content and incubated for 1 h at room temperature. The pH was then raised to 9.0 by the addition of 1 M Tris base, and the solution was incubated overnight at 37 °C. The reaction was terminated by the addition of 50 mM β-mercaptoethanol, and salts were removed by dialysis (1:1 dilution of isotonic buffer) prior to one- or two-dimensional polyacrylamide gel electrophoresis.

Two-Dimensional Cellulose Peptide Maps. Two-dimensional peptide maps were prepared by using fixed and Coomassie blue stained protein bands excised from either one- or two-dimensional polyacrylamide gels as originally described by Elder et al. (1977) with the following modifications. Excised gel bands were washed and radiolabeled with $Na^{125}I$ (carrier-free, 17 Ci/mg, New England Nuclear) using between 200 and 300 μCi of $Na^{125}I$ per gel slice. After iodination, free iodine was removed by extensive dialysis against 10% methanol in water (v/v), and iodinated gel slices were digested with 50 μg of α-chymotrypsin (Worthington, 3× crystallized) in 1 mL of 50 mM ammonium bicarbonate, pH 8.0, for 24 h at 37 °C. Duplicate 5-μL aliquots of supernatant solution containing the solubilized peptides were quantitated for radioactivity, and the remaining sample was lyophilized. Lyophilized fractions were dissolved in 20 μL of acetic acid/formic acid/water, 15:5:80 (v/v). Aliquots containing 20 cpm/dalton (original peptide molecular mass) were spotted on cellulose-coated chromatography sheets (Baker JT4480-4 without fluorescent indicator; 20 × 20 cm) and electrophoresed in the above buffer on an LKB flat-bed electrophoresis apparatus (2117-001 Multiphor II) at 35-W constant power (about 1400 V) using a recirculating water cooler (Savant Instruments RWC-50) to maintain the temperature at 7 °C. Electrophoresis was terminated when basic fuschin (1.0 μL of a 0.5% solution in the above buffer) migrated 11.5 cm (approximately 40 min). Chromatograms were dried and developed in the second dimension in 1-butanol/pyridine/acetic acid/water, 32.5:25:5:20 (v/v), containing 7% 2,5-diphenyloxazole (w/v, National Diagnostics). A series of exposures (1–24 h at −70 °C) was prepared for each cellulose map using Kodak XRP-1 or XAR-5 film with X-ray intensifying screens (Dupont Lightning Plus L847550). For comparison of maps, exposures with the closest grain density of major spots were selected. For each peptide, at least two maps were made from different gels that were iodinated separately. Polyacrylamide gel slices devoid of protein were iodinated and processed under the same conditions to determine background.

Fast Protein Liquid Chromatography (FPLC) Self-Association Assay. Filamin monomers, dimers, and larger complexes were separated by using a Superose HR 6 (Pharmacia) column (1 × 30 cm) at a flow rate of 0.5 mL/min. The column and injector were equilibrated at 0 °C in an ice bath, and a 1-m precolumn coil precooled the elution buffer (10 mM Tris, 20 mM NaCl, 130 mM KCl, 1 mM EDTA, 1 mM β-mercaptoethanol, 0.005% azide, 1 mM DFP, and 2 μg/mL leupeptin, pH 7.4). Protein amount was determined by the absorbance at 280 nm using a Nelson Analytical data acquisition system. Dimer and monomer concentrations were determined by integrating the peak area at 280 nm using a response factor determined by amino acid analysis. In most experiments, fractions were also collected and analyzed by SDS–polyacrylamide gel electrophoresis.

Amino Acid Compositional Analysis. Amino acid analysis was performed after hydrolysis at 110 °C for 20 h in vacuo in 200 μL of 6 N HCl containing 0.2% phenol. Tryptophan was determined after 20-h hydrolysis in 20 μL of 3 N mercaptoethanesulfonic acid (Penke et al., 1974). Cysteine was determined as cysteic acid after performic acid oxidation (Moore, 1963). Time course hydrolyses and analyses were performed ranging from 20 to 110 h. Maximum values were taken for valine and isoleucine, threonine and serine were extrapolated to zero time, and values for the remaining amino acids were averaged from four time points. Amino acid analysis was performed on a Dionex D-500 amino acid analyzer using a ninhydrin detection system.

Sequence Analysis. Initial peptide sequencing was performed on a modified Beckman 890C sequencer with 0.1 M Quadrol, Beckman reagents and solvents, and autoconversion with methanolic HCl (Speicher et al., 1983). A blank cycle was run before addition of the sample. Phenylthiohydantoin (PTH) derivatives were analyzed by on-line HPLC using a Rainin gradient HPLC system with Baker HPLC-grade solvents. Some sequences were also determined on an upgraded gas-phase sequencer (Model 475A, Applied Biosystems, Inc.) with an on-line PTH analyzer (Model 120A, Applied Biosystems, Inc.) using standard programs and reagents.

Computer Analysis. Sequences derived from filamin peptides were compared to the protein identification resource (PIR) protein database using the program FASTP (George et al., 1986). More detailed, statistical comparisons of filamin sequences with selected other proteins were performed by using the program ALIGN. This program compares the test alignment to randomized sequences as described (Barker et al., 1978). For single pairwise comparisons, the Mutation Data Matrix (250 PAMs) was used (Barker et al., 1978; Schwartz & Dayhoff, 1978) with 50 random runs and gap penalty of either 20 or 40. The PIR database and programs used were purchased from the National Biomedical Research Foundation (Washington, D.C.).

Rotary Shadow Electron Microscopy. Purified human smooth muscle filamin was rotary shadowed with platinum/carbon at low angle according to the procedure of Tyler et al. (1980).

RESULTS

Purification and Characterization of Human Smooth Muscle Filamin. Initially, the purification method of Wang (1977) was used in an attempt to isolate human uterine filamin. After experiencing poor recovery and poor purity, the method of Shizuta et al. (1976) was used with little improvement. However, when avian smooth muscle filamin was purified by using either of these methods, a high yield of homogeneous filamin was obtained. This was the first indication that human and avian smooth muscle filamins were substantially different.

Therefore, a modified method was developed to purify human smooth muscle filamin in high yield and purity (Figure 1). Precipitation of myosin at low ionic strength was critical since it could not be separated from filamin by any of the other purification steps without drastically reducing yield. Ultracentrifuge spins were required to effectively clarify supernatants at most steps. The use of HPLC ion exchange increased the purity of the final product, in large part due to the high-efficiency HPLC column and because complex gradients could be conveniently generated to optimize separation. Concentration by vacuum dialysis was the only method tested which was suitable for concentration of human filamin. The use of protease inhibitors and low temperatures was absolutely essential. Leupeptin was especially important, presumably to inhibit a calcium-activated protease, since high EDTA concentrations could be substituted for leupeptin with only a slight increase in proteolysis. This method yielded about 30 mg of

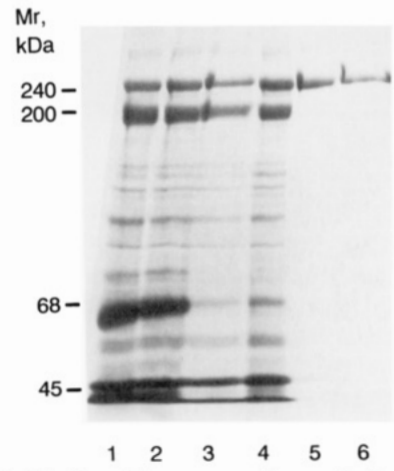

FIGURE 1: Purification of human smooth muscle filamin. Lane 1, whole uterine tissue homogenate (100 μg). Lane 2, supernatant after ultracentrifugation of homogenate at 95000*g* for 20 min (100 μg). Lane 3, 40% ammonium sulfate pellet (60 μg). Lane 4, Load to Sepharose CL-4B gel filtration column (resuspension of ammonium sulfate pellet in column buffer followed by ultracentrifugation at 215000*g* for 30 min, 70 μg). Lane 5, filamin pool after Sepharose CL-4B column chromatography (10 μg). Lane 6, filamin pool after HPLC ion-exchange chromatography (10 μg). Samples were run on a 7% polyacrylamide gel. Molecular weight standards included were α-spectrin (240K), myosin (200K), bovine serum albumin (68K), and ovalbumin (45K).

purified filamin per 100 g of tissue (wet weight). Purity was typically >95%, and purities of >98% were achieved at slightly reduced yields. Persistent contaminants were primarily trace levels of filamin proteolytic degradation products as indicated by immunoblot analysis using anti-filamin antibodies.

Since human smooth muscle filamin behaved very differently from avian smooth muscle filamin during purification, properties of the human isoform were examined based upon the known properties of the avian molecule. The amino acid composition of human uterine filamin was similar to avian smooth muscle filamin (Shizuta et al., 1976; Wang, 1977), nearly identical with guinea pig vas deferens filamin (Wallach et al., 1978a), and quite different from human erythrocyte spectrin (Table I). Human filamin was examined by rotary shadow electron microscopy using the method of Tyler et al. (1980). As expected, filamin was visualized as extended flexible rods of two molecular lengths which apparently corresponded to monomers and dimers (Figure 2, top and bottom). Spectrin contamination was not detectable in the purified human smooth muscle filamin fraction by immunoblot analysis using polyclonal antibodies prepared against human brain spectrin (gift of Dr. Jon S. Morrow, Yale University).

Table I: Amino Acid Compositions of Filamin Isoforms and Spectrin (Mole Percent)

amino acid	human smooth muscle filamin	guinea pig vas deferens filamin[a]	avian smooth muscle filamin[a]	human erythrocyte spectrin
Asp	8.5	9.2	7.6	10.8
Thr	6.6	6.4	5.7	3.9
Ser	8.0	6.8	6.4	5.2
Glu	11.1	10.8	9.4	20.1
Pro	8.0	8.7	8.4	2.1
Gly	11.9	12.0	12.9	4.7
Ala	7.3	7.7	9.2	9.3
Val	10.1	9.8	11.5	4.8
Met	1.0	1.1	1.0	1.1
Ile	3.9	4.1	3.1	3.5
Leu	5.4	6.0	5.7	14.2
Tyr	3.2	2.7	1.9	2.1
Phe	2.8	2.2	2.7	3.4
His	2.4	2.1	2.2	3.0
Lys	6.2	6.4	5.2	7.1
Arg	3.6	3.5	5.5	6.4

[a] Wallach et al. (1978a). [b] Shizuta et al. (1976).

Production of Human Smooth Muscle Filamin Domain Peptides. Numerous proteases, buffers, ionic strengths, and temperatures were tested to determine the optimal conditions for fragmenting filamin into a relatively small number of intermediate-sized (domain) peptides. Mild tryptic treatment (0 °C, 20 mM Tris, 0.02% azide, and 1 mM β-mercaptoethanol, pH 7.6; E:S = 1:100) provided the best fragmentaion of the molecule (Figure 3). Avian smooth muscle filamin, tested in parallel, was more resistant to cleavage by most proteases tested, and individual regions of the molecule were cleaved more heterogeneously, producing many more fragments (Figure 3). When these tryptic digests were run on two-dimensional gels, the pattern was only slightly more complex than suggested by the one-dimensional gels. Corresponding human and avian domain peptides usually had similar but nonidentical p*I*'s (Figure 4A).

Chemical cleavage of human and avian smooth muscle filamins with NTCB (cysteine cleavage) produced peptides which were very different in size and isoelectric point (Figure

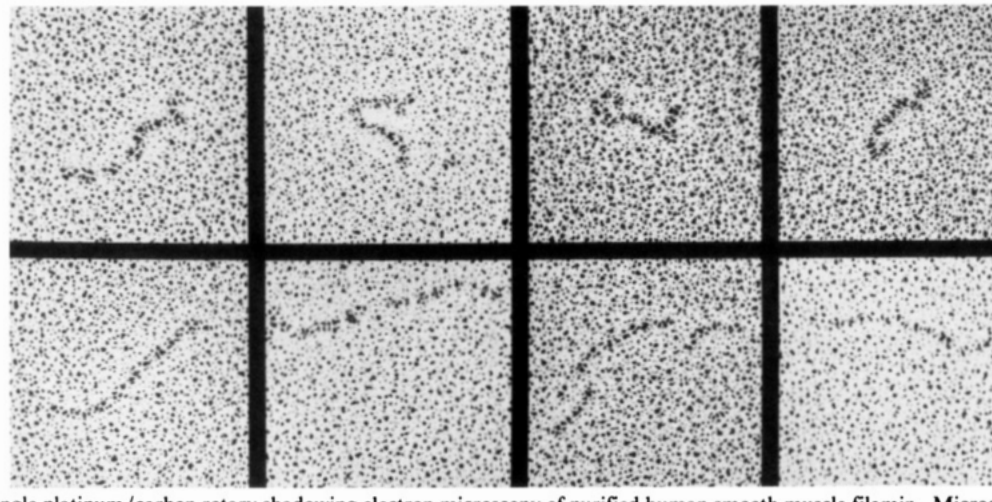

FIGURE 2: Low-angle platinum/carbon rotary shadowing electron microscopy of purified human smooth muscle filamin. Micrographs representative of the predominant species seen upon rotary shadowing of purified human smooth muscle filamin (200000×). The top rows shows filamin monomers (approximately 85 nm in length) while the bottom row shows filamin dimers (approximately 170 nm in length).

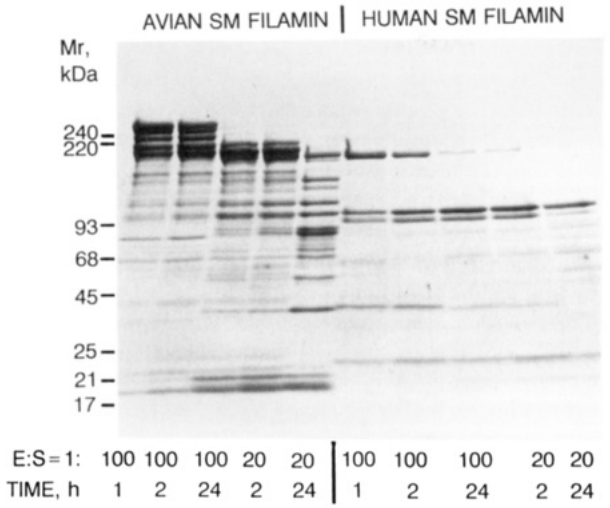

FIGURE 3: Comparison of mild tryptic digests of human and avian smooth muscle filamin. The enzyme-to-substrate ratio (E:S) was based on weight. Digestion times (in hours) and E:S ratios are indicated below each lane. Trypsin digestion was performed at 0 °C in 20 mM Tris, 0.02% azide, and 1 mM β-mercaptoethanol, pH 7.6. Samples were run on a 12% polyacrylamide gel. Molecular weight standards included were α-spectrin (240K), β-spectrin (220K), phosphorylase B (93K), bovine serum albumin (68K), ovalbumin (45K), chymotrypsinogen A (25K), soybean trypsin inhibitor (21K), and myoglobin (17K).

4B). This indicated poor conservation of cysteine residues between the smooth muscle isoforms. Additionally, the NTCB cleavage peptides proved useful for overlaps of tryptic domains.

Comparison of Intact Human Smooth Muscle, Avian Smooth Muscle, and Human Platelet Filamins. Since avian and human smooth muscle filamins exhibited differing trypsin susceptibility and NTCB cleavage patterns (Figures 3 and 4), two-dimensional cellulose peptide mapping was then used to further compare these isoforms. Identical spots on the two peptide maps usually indicate few if any amino acid substitutions. When two-dimensional cellulose peptide maps of intact human and avian smooth muscle filamin were compared (Figure 5), approximately half the spots were different, indicating extensive sequence diversity. Two-dimensional cellulose peptide mapping was also used to compare human smooth muscle filamin with human platelet filamin. In contrast to the results described above with the smooth muscle filamin isoforms, the chymotryptic peptide maps of the human tissue isoforms were nearly identical (Figure 5), strongly suggesting that the same filamin isoform or nearly identical isoforms predominate in both human tissues.

Linear Alignment of Human Smooth Muscle Filamin Domain Peptides. The intermediate-sized tryptic peptides of human smooth muscle filamin (Figures 3 and 4) were subsequently analyzed by two-dimensional peptide mapping on thin-layer cellulose plates to establish precursor–product relationships and to align unique peptides. Figure 6 shows the cellulose peptide maps of the three unique tryptic domains identified, represented by 110-, 100-, and 44-kDa peptides, and several other trypic peptides. The three unique peptides comprise most of the mass of the intact filamin molecule. Alignment of these peptides was established by mapping larger tryptic fragments from shorter digestion times (Figure 3). In addition to the peptides represented in Figure 6, all other peptides produced by mild trypsin digestion were mapped and placed within the alignment. Cyanogen bromide and NTCB (Figure 4B) cleaved peptides were also mapped and provided useful multiple overlaps. For example, NTCB 45-kDa fragment overlaps the tryptic 100/90-kDa peptides and also the

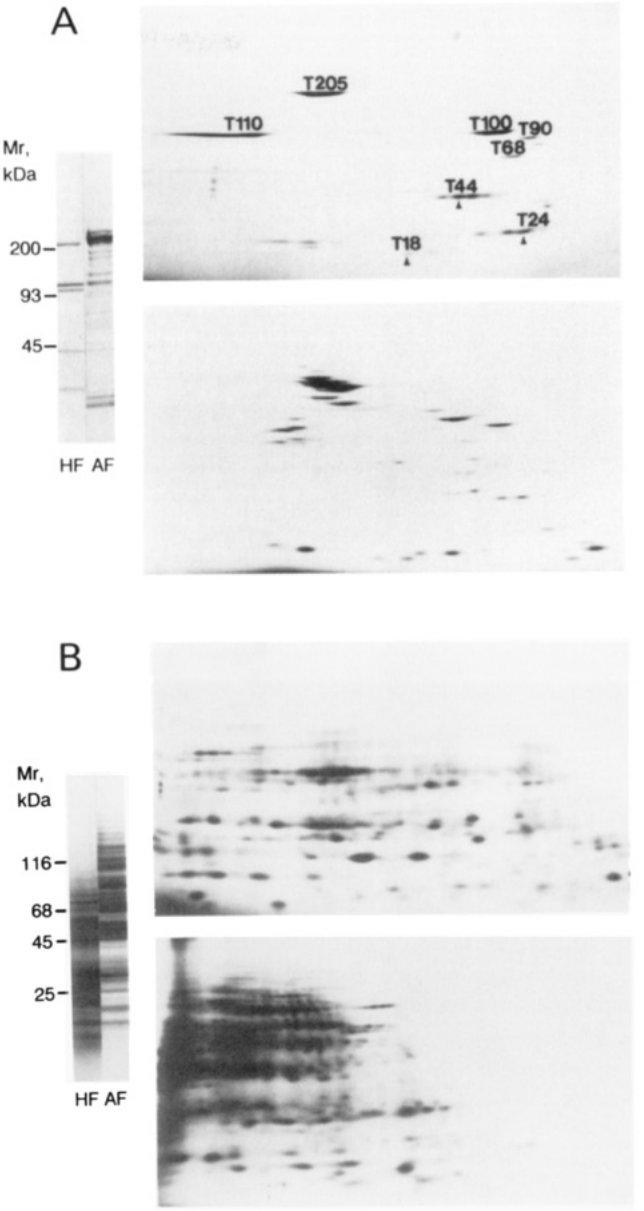

FIGURE 4: Enzymatic digestion and chemical cleavage of human and avian smooth muscle filamin. (A) Mild trypsin digest of human smooth muscle filamin (HF) or avian smooth muscle filamin (AF) electrophoresed on one-dimensional (left) and two-dimensional polyacrylamide gels (HF, top right, and AF, bottom right). Spots are labeled with the peptide's molecular weight $\times 10^{-3}$; e.g., T110 is a tryptic M_r 110000 fragment. (B) NTCB cleavage products of human smooth muscle filamin (HF) or avian smooth muscle filamin (AF) electrophoresed on one-dimensional (left) and two-dimensional polyacrylamide gels (HF, top right, and AF, bottom right). One-dimensional gels were 3–18% polyacrylamide gradients. Two-dimensional gels contained a nearly linear pH gradient from about pH 8.0 to pH 4.5 (basic left, acidic right). Molecular weight standards electrophoresed on the one-dimensional gels were myosin (200K), β-galactosidase (116K), phosphorylase B (93K), bovine serum albumin (68K), ovalbumin (45K), and chymotrypsinogen A (25K).

tryptic 44/24-kDa peptides. At least two clear overlaps for each domain were identified.

Characterization of the Self-Association Site. An FPLC gel filtration assay employing a Superose HR 6 column (see Experimental Procedures) was used to directly quantitate filamin monomers, dimers, and associated peptides. The FPLC elution profile of human smooth muscle filamin purified from the leading edge of the Sepharose CL-4B filamin peak is shown in Figure 7A. This fraction is mostly dimer since rotary shadow electron microscopic analysis of an analogous prepa-

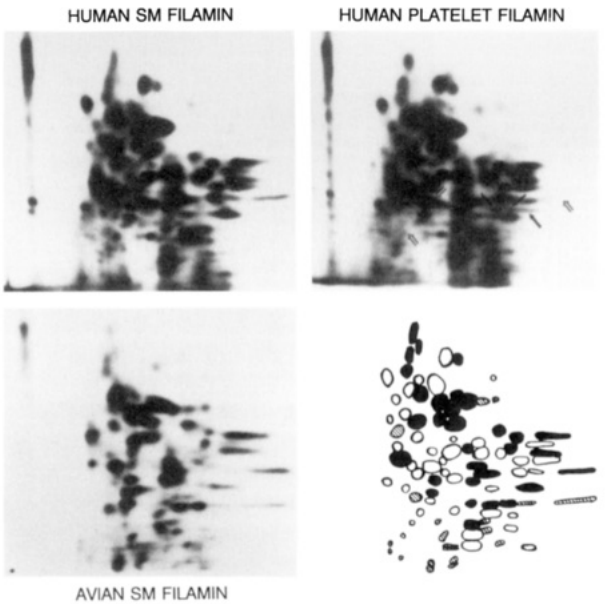

FIGURE 5: Comparison of the two-dimensional cellulose peptide map of human smooth muscle filamin with the peptide maps of avian smooth muscle and human platelet filamins. Intact filamin bands were cut from two-dimensional polyacrylamide gels after Coomassie blue staining and then mapped. Samples were spotted (lower left corner) on thin-layer cellulose plates and electrophoresed to the right followed by ascending chromatography. The lower right panel is a peptide map comparison of human and avian smooth muscle filamins. Closed spots denote common peptides, open spots represent peptides present in human smooth muscle filamin but not in avian smooth muscle filamin, and hatched spots represent peptides present in avian smooth muscle filamin but not in human smooth muscle filamin. A comparison between human smooth muscle and human platelet filamins is indicated in the upper right panel. The common peptides are not labeled, solid arrows indicate peptides present in platelet filamin but not in smooth muscle filamin, and open arrows indicate peptides present in smooth muscle filamin but not in platelet filamin.

ration visualized molecules that were primarily 170 nm long (see Figure 2, two 85-nm monomers associated end-to-end). Figure 7B shows the chromatograph of filamin purified from the tailing side of the Sepharose CL-4B filamin peak. As expected, this sample is primarily monomer since rotary shadowing of an analogous preparation visualized molecules that were primarily 85 nm long.

A specific cleavage of filamin into 220- and 35-kDa peptides occurred when both leupeptin and EDTA were deleted from the filamin storage buffer. This proteolysis was probably due to trace contamination by calcium-activated proteases rather than intrinsic autolytic activity since cleavage was more pronounced with less highly purified samples. When samples of purified dimer were intentionally incubated in the absence of protease inhibitors (leupeptin and EDTA) to produce the 220-kDa peptide and then separated on the FPLC column, most of the dimer was converted into a peak migrating near the monomer region (Figure 7C). The slightly earlier elution of the major peak relative to monomer (panel C) was reproducible and significant since reproducibility of retention times on this system was remarkably good. Analysis by SDS gels showed that the leading edge of this peak contained primarily 250- and 35-kDa peptides while the tailing side of this peak contained primarily a 220-kDa peptide.

The interpretation of the separation in panel C is that the leading edge of the major peak contained intact filamin monomers (250 kDa) associated to a 35-kDa peptide left after proteolysis of one monomer in the original dimer. The tailing edge of the peak contained the complementary 220-kDa part of the molecule. The filamin monomer/35-kDa complex could not be separated by simple dilution or manipulation of salt conditions, analogous to the behavior of authentic human smooth muscle dimers. In these experiments, a 35-kDa peptide dimer or 35-kDa peptide not associated to filamin monomer

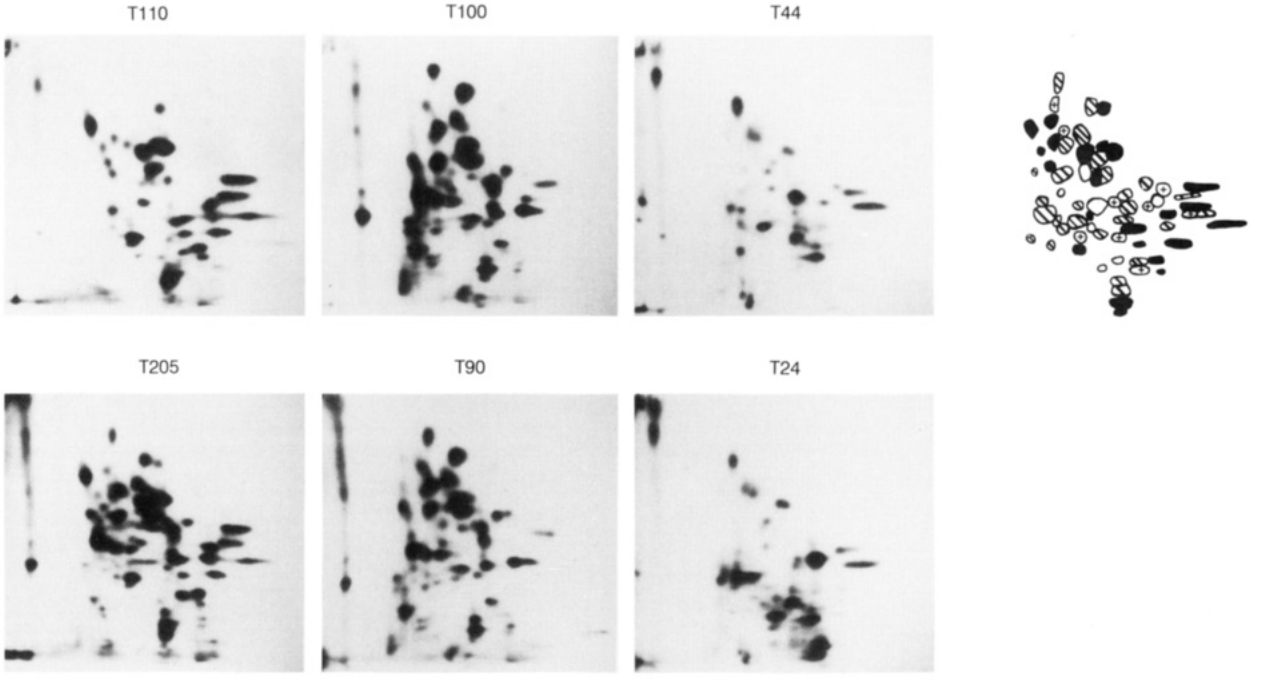

FIGURE 6: Two-dimensional cellulose peptide maps of human smooth muscle filamin tryptic fragments. Tryptic peptides were cut from two-dimensional polyacrylamide gels after Coomassie blue staining and then mapped. These peptide maps correspond to peptides which were sequenced (see Figure 8). The upper row shows the cellulose peptide maps of the tryptic 110-, 100-, and 44-kDa peptides and shows a comparison of these peptide maps to demonstrate the uniqueness of these peptides (top, right). With the cellulose peptide map of intact human smooth muscle filamin (Figure 5) as the template, closed spots represent peptides from the T110 kDa, hatched spots represent peptides from the T100 kDa, spots containing a plus (+) sign represent peptides from the T44 kDa, and open spots represent unassigned peptides. Relationships of peptides in the lower row are given in Figure 8. Samples were spotted (lower left corner) on thin-layer cellulose plates and electrophoresed to the right followed by ascending chromatography.

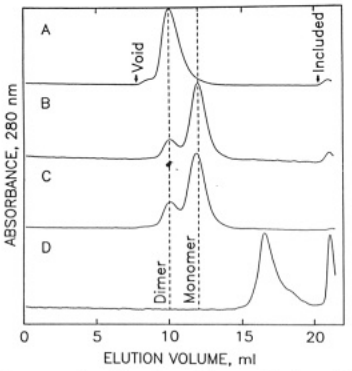

FIGURE 7: FPLC assay for filamin self-association. Filamin fractions were chromatographed on a Superose HR 6 column at 0 °C with a flow rate of 0.5 mL/min. Peak heights do not represent relative amounts (largest peak set to full scale). The void and included volumes of the column are indicated by arrows, while the vertical dashed lines indicate elution positions of filamin dimers and monomers. (A) Filamin dimer fraction (7.8 mg/mL in isotonic buffer in the presence of leupeptin and EDTA) and (B) filamin monomer fraction from Sepharose CL-4B separation in the presence of leupeptin and EDTA. (C) Filamin dimer fraction incubated in the absence of leupeptin and EDTA showing conversion into a peak migrating near the monomer region. (D) Rechromatography of the 35-kDa peptide (purified from the monomer region of run C by gel filtration chromatography in 8 M urea to dissociate the filamin monomer/35-kDa peptide complex followed by removal of urea).

were never observed. This indicated that the filamin dimer may be cleaved easier than the filamin monomer/35-kDa peptide complex, implying a subtle allosteric conformational change limits protease susceptibility of the second subunit.

When the filamin monomer/35-kDa complex was chromatographed on the FPLC column in the presence of urea, the two peptides were well separated. Also, after removal of urea, purified 35-kDa peptide eluted close to the included volume of the FPLC column (Figure 7D), supporting the idea that the peak in panel C near the monomer position was really a complex between a monomer and the 35-kDa peptide and not an incidental coelution.

The location of this cleavage site was determined by peptide map analysis of the 220- and 35-kDa peptides. Cellulose peptide maps indicated that the 35-kDa peptide overlapped the T44-kDa peptide and contained all the spots from the T18-kDa peptide but did not contain all the spots from the T44- or T24-kDa peptides. There were also several spots present on the map of the 35-kDa peptide which were not present on the T44-kDa peptide. The cellulose peptide map of the 220-kDa peptide was nearly identical with that of the T205-kDa peptide (see Figure 6, T205).

Sequence Analysis of Intact Human Smooth Muscle Filamin and Domain Peptides. Several attempts to sequence the intact protein using automated Edman degradation did not produce any sequence. This indicated that the amino terminal of human filamin was blocked. A sequence at a level <5% the amount loaded could have easily been detected. The 220-kDa peptide produced by intrinsic proteolysis was blocked. In contrast, the 205-kDa peptide produced by mild tryptic digestion was not blocked.

In addition, most of the other peptides produced by mild proteolysis were sequenced. Peptides were excised from two-dimensional gels (similar to Figure 4A, top right), electroeluted, dialyzed, quantitated by amino acid analysis, and sequenced. A total of 138 residues of sequence were determined as summarized in Figure 8A.

Characterization of the Actin-Binding Site. Intact human smooth muscle filamin and peptide digests were assayed for actin-binding activity in a cosedimentation assay (Correas et

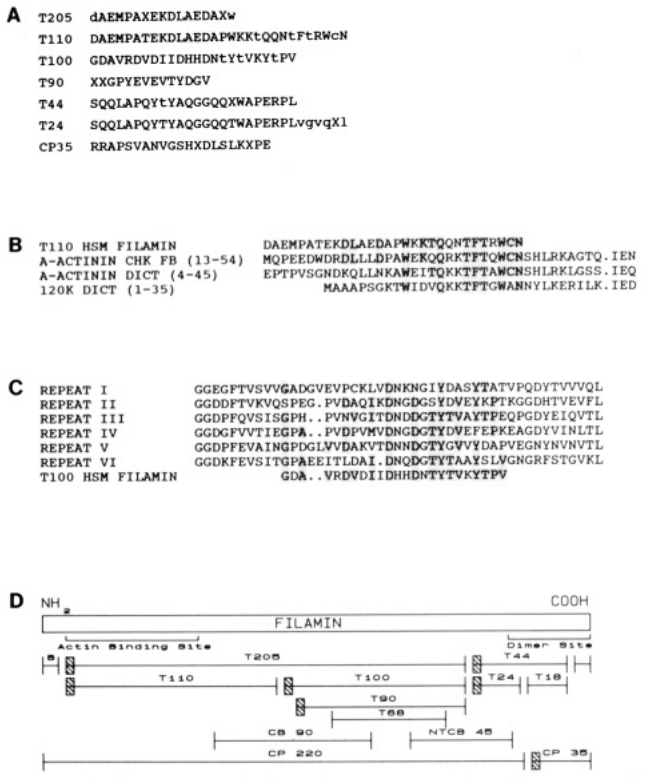

FIGURE 8: Sequence analysis and summary of the substructure of human smooth muscle filamin. (A) Amino-terminal sequences of filamin peptides. Unknown amino acid assignments are indicated by X and tentative by lower case letters. (B) Comparisons between the tryptic 110-kDa peptide (T110) sequence and amino-terminal sequences of proteins containing actin-binding sites: chicken fibroblast α-actinin, *Dictyostelium* α-actinin, and *Dictyostelium* 120-kDa gelation factor (Baron et al., 1987; Noegel et al., 1987, 1989). In single pairwise comparisons of the T110-kDa peptide sequence with these three sequences, the alignment scores were respectively 9.0, 10.0, and 5.5 standard deviations (parameters: Mutation Data Matrix-250 PAMs; 50 random runs; gap pentalty of 40). (C) Sequence comparison between the tryptic 100-kDa peptide (T100) peptide and the six internal repeats of the *Dicytostelium* 120-kDa gelation factor. In single pairwise comparisons of the T100-kDa peptide sequence with these six sequences (repeats I through VI), the alignment scores were respectively 2.2, 5.2, 7.4, 6.4, 5.8, and 6.2 standard deviations (parameters: Mutation Data Matrix-250 PAMs; 50 random runs; gap penalty of 40). Repeat sequences shown begin at about amino acid 30 from the amino-terminal end of published sequences and are shown as gapped by the authors. In panels B and C, an amino acid identity between the filamin peptide and the comparison sequence is indicated by shading. (D) Structural model of filamin and approximate location of binding sites. Hatched areas on peptides indicate sequenced regions. A short blocked peptide from the amino-terminal of filamin is indicated by B. Peptides are denoted by the reagent used to produce them: T, trypsin; CB, cyanogen bromide; NTCB, 2-nitro-5-thiobenzoic acid; CP, calcium protease.

al., 1986) with rabbit skeletal muscle F-actin. Intact filamin and the T205-kDa peptide cosedimented with F-actin, while control samples without actin remained in the supernatant. The T110-kDa peptide also apparently contained a competent actin-binding site as demonstrated by actin cosedimentation.

The computer program FASTP was used to compare the human smooth muscle filamin peptide sequences with the current PIR protein database and with an auxiliary database that was compiled in our laboratory from unpublished sequences and recently published sequences of interest to us. Selected sequences were further compared by using the statistical program ALIGN. The amino-terminal sequence of the T110-kDa peptide showed significant similarity to the amino-terminal actin-binding site of chicken fibroblast α-actinin (Baron et al., 1987), *Dictyostelium* α-actinin (Noegel et al.,

1987), and the *Dictyostelium* 120-kDa gelation factor (Noegel et al., 1989) (Figure 8B). In addition, this filamin sequence also showed reduced but significant similarity to amino-terminal regions of human dystrophin (Koenig et al., 1988), *Drosophila* β-spectrin (Byers et al., 1989), and human red cell β-spectrin.[2]

Repetitive Structure of Filamin. The amino acid compositions of human smooth muscle filamin domain peptides (T110, T100, and T44 kDa) were very similar to each other and to the intact molecule. While amino acid similarity does not prove internal homology exists, it is consistent with the occurrence of multiple internal repeats of a single type.

More importantly, the amino-terminal sequence of the T100-kDa peptide showed homology to the recently described six internal β-structure repeats present in the *Dictyostelium* 120-kDa gelation factor (Noegel et al., 1989). The similarity was highest for repeat III of the 120-kDa gelation factor, intermediate for repeats II and IV–VI, and lowest for repeat I (Figure 8C). The compositional and sequence homologies together provide evidence that human smooth muscle filamin contains repetitive structures similar to those in the 120-kDa gelation factor. These observations are consistent with recent reports of the cDNA sequence of human endothelial actin-binding protein that indicate that most of this filamin isoform is comprised of a single type of β-structure repeat related to the 120-kDa gelation factor repeats (Gorlin et al., 1989a,b).

Amino and Carboxyl Orientation of Structural Domains and Functional Sites. The amino and carboxyl orientations of mild tryptic cleavage domains within the molecule were deduced by comparison of sequence data with the domain alignments derived from peptide mapping (domain map, Figure 8D). These results indicate that the actin-binding site is located near the amino-terminal end of the molecule while the self-association site is near the carboxyl-terminal end. This orientation is supported by several observations. First, the T205- and T110-kDa peptides have identical amino-terminal sequences and the T44- and T24-kDa peptides have identical amino-terminal sequences (see Figure 8A). If the opposite orientation existed, the T100-kDa sequence should overlap the T205-kDa sequence, and the T44-kDa sequence might overlap the T18-kDa sequence. Second, this orientation and the assignment of the actin-binding site to the amino terminal are also consistent with the homology to the amino-terminal actin-binding site of spectrin, α-actinin, and dystrophin described above and in Figure 8B.

Several additional substructural features were derived. Since the amino terminal of the intact filamin molecule was blocked, the T205-kDa peptide was produced by cleavage of a small blocked peptide from the amino terminal in addition to cleavage of the T44-kDa peptide from its carboxyl-terminal end. The T90-kDa peptide must have been produced by cleavage of a small peptide from the amino-terminal end of the T100-kDa peptide, since the T100- and T90-kDa peptides had essentially identical cellulose peptide maps (Figure 6, middle, top, and bottom) and their amino-terminal sequences did not overlap (Figure 8A).

Discussion

The human smooth muscle filamin isoform was purified in high yield, and the peptide substructure of this multifunctional actin cross-linker was analyzed. A linear alignment of three unique tryptic peptides (T110, T100, and T44 kDa) within the the intact filamin molecule was formulated based upon peptide mapping analysis of these peptides, several smaller and larger tryptic fragments, and NTCB and cyanogen bromide cleavage products. This structural map of human smooth muscle filamin (peptide domain map, Figure 8D) has already facilitated the assignment of actin-binding and self-association functions to discrete regions of the molecule, and it should be invaluable in further characterization of this large molecule.

Multiple lines of evidence support the alignment of the three unique tryptic peptides presented here (T110–T100–T44). The combination of cellulose peptide mapping with amino-terminal sequencing allowed definitive alignment of peptides within the filamin molecule as well as assignment of amino- and carboxyl-terminal orientation. The assignment of T110 and T100 as neighbors comprising the T205 peptide is especially clear from both time course digestion (Figure 3) and peptide mapping (Figure 6). The common N-terminal sequences for the T205 and T110 confirm this relationship and support the orientation indicated in Figure 8. The placement of the T44 peptide is critical to completion of the map, and its potential location on either side of the T205 was considered. Placement of T44 on the carboxyl-terminal side of T205 as indicated in Figure 8 is clearly supported by the following observations: (a) the 220-kDa peptide produced by the calcium-activated protease, CP220, has a blocked amino terminal similar to intact filamin which indicates that this peptide extends to the amino terminal of the intact molecule as shown in Figure 8; (b) the CP220 cellulose peptide map is nearly identical with the T205 map and contains some, but not all, spots from T44—if T44 were on the amino-terminal side of T205 and CP220 extends to the amino terminal, the entire T44 map would be included in the CP220 peptide; (c) the CP35 which is complementary to the CP220 is not blocked, and its peptide map overlaps the map of T44, but does not overlap the T205 map; (d) the peptide map of an NTCB 45-kDa peptide contains spots from the T100 and T44 domains, indicating that they are adjacent to each other.

The observations that the CP35 peptide maps near the carboxyl terminal and that this peptide remains associated with the complementary subunit indicate the self-association site is located near the carboxyl-terminal end of the molecule. In contrast, the CP220 which contains the intact amino-terminal region does not remain associated with either another CP220 peptide or intact filamin monomer since it migrates slightly slower than monomers (Figure 7) rather than near the dimer position. These observations directly imply that the actin-binding site might be near the amino-terminal end of the molecule since dimers are long extended rods nearly twice the length of monomers (Figure 2), and dimers cross-link actin filaments near their distal ends. This location of the actin-binding domain is further supported by the following results: (a) peptides mapping to the amino-terminal end of the molecule (T205, T110) associate with actin filaments; (b) the amino terminal shows sequence homology to amino-terminal actin-binding domains of several other cytoskeletal proteins; and (c) the calcium-activated protease cleavage of human filamin into an actin-associating CP220 fragment containing the blocked amino terminal and a self-association competent CP35 peptide is closely analogous to avian smooth muscle heavy and light merofilamin as discussed below.

The peptide domain map presented here is more detailed than previous attempts to characterize filamin substructure in other tissues or species. The majority of published data is consistent with results reported herein although it is difficult to reconcile one report on calcium-activated proteolysis with

[2] John C. Winkelmann and Bernard G. Forget and co-workers, personal communication.

the domain map presented here. The avian smooth muscle filamin isoform is cleaved by the skeletal muscle calcium protease into a large fragment (heavy merofilamin, 240 kDa) and a smaller fragment (light merofilamin, 10 kDa). Heavy merofilamin binds to F-actin but cannot induce gelation of F-actin into networks (Davies et al., 1978) and is unable to self-associate (Davies et al., 1980). Therefore, light merofilamin is responsible for the association of avian smooth muscle filamin monomers into dimers and confers gelation activity on the intact molecule. These results are consistent with our observation that human smooth muscle filamin is cleaved into a 220-kDa fragment containing the actin-binding site and a 35-kDa fragment containing the carboxyl-terminal self-association site when calcium protease inhibitors are removed. The differences in molecular weight probably reflect a species difference and/or differing specificity of the proteases involved in each case.

Results obtained by calcium-activated proteolysis of mammalian filamins [guinea pig vas deferens (Wallach et al., 1978), human platelet (Truglia & Stracher, 1981; Fox et al., 1983, 1985; Ezzell et al., 1988), and HeLa cells (Weihing, 1988)] are also consistent with our results. These filamin isoforms are cleaved into a large fragment (190–200 kDa) and a smaller fragment of about 100 kDa (from which approximately 10-kDa peptide is cleaved to produce an approximately 90-kDa peptide). Two of these studies employed actin cosedimentation assays to determine which peptide(s) retained actin-binding activity [HeLa cell (Weihing, 1988) and platelet (Truglia & Stracher, 1981)]. Both studies demonstrated that the larger fragment contained the actin-binding domain and the platelet study showed that the larger fragment bound to F-actin but could not gel F-actin into networks. These results are again consistent with our observation that human smooth muscle filamin is cleaved into a large fragment containing the actin-binding site (220-kDa peptide) when calcium-activated proteolysis was allowed to occur.

In contrast, some results reported by Hartwig and co-workers for platelet actin-binding protein are more difficult to reconcile with our observations on human smooth muscle filamin. They reported that a self-association specific monoclonal antibody immunoprecipitated a 190-kDa calcium protease fragment of human platelet actin-binding protein (Ezzell et al., 1988). This result appears inconsistent with our results and previously published results since none of these studies indicated any evidence of a large peptide produced by calcium-activated proteolysis which contained the self-association site. Similarly, we could not produce such a fragment either by "autodigestion" or with added calcium-activated protease. In this same study, Hartwig and co-workers used monoclonal antibodies to the platelet filamin isoform and to actin in coprecipitation studies to demonstrate that the actin-binding site is contained within a 100/90-kDa peptide produced by calcium-activated protease (Ezzell et al., 1988). This appears inconsistent with both our present results and the data published by others. None of these studies using calcium-activated proteases showed evidence of peptides smaller than approximately 190–240 kDa that contained the actin-binding site. This inconsistency is especially surprising since two-dimensional cellulose peptide maps indicate that human platelet and smooth muscle isoforms are nearly identical (Figure 5). One possible explanation for the apparent inconsistency is that the endogenous calcium protease(s) that are stimulated by the ionophore used may be different from the calcium-activated proteases used in the other experiments discussed above. An alternative explanation is that differences in phosphorylation state or the potential presence of a cofactor in the whole cell platelet experiment may influence the cleavage pattern.

Recent abstracts by Hartwig and co-workers which report the cloning and cDNA sequence of human endothelial cell actin-binding protein (Gorlin et al., 1989a,b) do appear consistent with the results described here. Specifically, they report that the carboxyl-terminal four repeats of this filamin isoform are expressed as a secreted peptide (about 45 kDa) which spontaneously self-associates. Their localization of the self-association site within a carboxyl-terminal 45-kDa fragment is consistent with our peptide studies that implicate a carboxyl-terminal 35-kDa fragment in filamin self-association. They also report that the amino terminal has positively charged amino acid residues and propose that these charge clusters are related to an actin-binding site although they did not describe homology to other actin-binding domains. It appears likely that their positively charged amino-terminal region will be homologous to the actin-binding domain identified by us at the amino-terminal end of the human smooth muscle filamin tryptic 110-kDa fragment.

Hartwig and co-workers also reported that the endothelial cell filamin isoform DNA sequence indicated that 21 β-structure repeat units (each about 100 amino acids in length) follow "amino-terminal charge clusters" (Gorlin et al., 1989a,b). The *Dictyostelium* 120-kDa gelation factor contains an amino-terminal actin-binding site followed by six β-structure repeats. It is possible that the filamin class of proteins may have evolved from smaller gelation factors such as the 120-kDa protein or that these proteins evolved from a common ancestor. Since *Dictyostelium* contains an approximately 240-kDa filamin isoform (Hock & Condeelis, 1987; Condeelis et al., 1988) in addition to the 120-kDa gelation factor, it should be possible to study the evolutionary development of filamin and the related 120-kDa protein. This raises the possibility that gelation factors similar to the 120-kDa protein exist in mammalian cells although this has not yet been demonstrated. Both the 120-kDa protein (Condeelis et al., 1984) and filamin (Hartwig & Shevlin, 1986) organize actin into three-dimensional networks in the cell periphery, but may differ in their function and detailed subcellular localization. This observation has a interesting parallel to the spectrin-like family which also contains short (α-actinin) and long (spectrin, fodrin) actin cross-linkers (Dubreuil et al., 1989) which have subtle differences in function and subcellular localization.

The lack of homology between filamins and spectrins is also interesting. Both classes of proteins are long flexible rods with actin cross-linking activity. While their similar morphology and function might suggest a distant evolutionary relationship, this is clearly not the case with the exception of the amino-terminal actin-binding domain. Filamin and spectrin differ in their self-association mechanism (monomer/monomer vs heterodimer/heterodimer interaction), actin-binding region (single vs double stranded), and repetitive structure (β-structural vs α-helical repeats). Further examination of the regulation of the functional sites of filamin and spectrin should yield information on their putative differential roles as high molecular weight actin cross-linking proteins in the three-dimensional organization of cellular microfilament networks.

Additionally, further biochemical characterization of human filamin isoforms is especially important since the better characterized avian smooth muscle filamin molecule is quite different from mammalian filamin isoforms as demonstrated here. Human and avian smooth muscle isoforms behave differently upon purification, have differing protease suscep-

tibility and chemical cleavage patterns, and have dissimilar two-dimensional cellulose peptide maps. This indicates substantial species diversity between smooth muscle isoforms. At the same time, the human and avian smooth muscle isoforms do have similar amino acid compositions and substantial immunological cross-reactivity as might be expected for even substantially diversified species isoforms of the same protein. The substantial differences observed on the cellulose peptide map comparisons of the avian and human isoforms are consistent with related, but moderately conserved proteins. This mapping method is highly sensitive to sequence differences since almost any single mutation in a limit peptide would probably shift its position in at least one of the two dimensions of the cellulose plate.

In contrast, some tissue-specific isoform differences within a single species may be minimal since two human filamin isoforms (those predominating in smooth muscle and platelet) are nearly identical when compared by cellulose peptide mapping. Therefore, structural information learned about human smooth muscle filamin should be able to be readily applied to the platelet isoform.

ACKNOWLEDGMENTS

The expert technical assistance of Ray DeAngelis (Yale University) as well as Kevin Beam and Clem Purcell (The Wistar Institute) is gratefully acknowledged. We thank Kaye Speicher for assistance with the computer sequence analysis and Drs. Meenhard Herlyn and Nancy Joyce for critical comments on the manuscript. We especially thank Drs. John Winkelmann and Bernard Forget for providing us with sequences of human red cell β-spectrin prior to publication.

REFERENCES

Barker, W. C., Ketcham, L. K., & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, pp 359–362, National Biomedical Research Foundation, Silver Spring, MD.

Baron, M. D., Davison, M. D., Jones, P., & Critchley, D. R. (1987) *J. Biol. Chem. 262*, 17623–17629.

Brotschi, E., Hartwig, J. H., & Stossel, T. P. (1978) *J. Cell Biol. 253*, 8988–8993.

Byers, T. J., Husain-Chishti, A., Dubreuil, R. R., Branton, D., & Goldstein, L. S. B. (1989) *J. Cell Biol. 109*, 1633–1641.

Carroll, R. C., & Gerrard, J. M. (1982) *Blood 59*, 466–471.

Condeelis, J., Vahey, M., Carboni, J. M., DeMay, J., & Ogihara, S. (1984) *J. Cell Biol. 99*, 119–126s.

Condeelis, J., Hall, A., Bresnick, A., Warren, V., Hock, R., Bennett, H., & Ogihara, S. (1988) *Cell Motil. Cytoskel. 10*, 77–90.

Correas, I., Leto, T. L., Speicher, D. W., & Marchesi, V. T. (1986) *J. Biol. Chem. 261*, 3310–3315.

Cox, A. C., Carroll, R. C., White, J. G., & Rao, G. H. R. (1984) *J. Cell Biol. 98*, 8–15.

Davies, P. J. A. Wallach, D., Wallingham, M. C., & Pastan, I. (1978) *J. Biol. Chem. 253*, 4036–4042.

Davies, P. J. A., Wallach, D., Willingham, M., Pastan, I., & Lewis, M. S. (1980) *Biochemistry 19*, 1366–1372.

Dubreuil, R. R., Byers, T. J., Sillman, A. L., Bar-Zvi, D., Goldstein, L. S. B., & Branton, D. (1989) *J. Cell Biol. 109*, 2197–2205.

Edler, J. H., Pickett, R. A., II, Hampton, J., & Lerner, R. A. (1977) *J. Biol. Chem. 252*, 6510–6515.

Ezzell, R. M., Kenney, D. M., Egan, S., Stossel, T. P., & Hartwig, J. H. (1988) *J. Biol. Chem. 263*, 13303–13309.

Fox, J. E. B., Reynolds, C. C., & Phillips, D. R. (1983) *J. Biol. Chem. 258*, 9973–9981.

Fox, J. E. B., Goll, D. E., Reynolds, C. C., & Phillips, D. R. (1985) *J. Biol. Chem. 260*, 1060–1066.

Frieden, C. (1985) *Annu. Rev. Biophys. 14*, 189–210.

Fuchiwaki, T. (1987) *Nippon Sanka Fukinka Gakkai Zasshi 39*, 1029–1036.

George, D. G., Barker, W. C., & Hunt, L. T. (1986) *Nucleic Acids Res. 14*, 11–15.

Gomer, R. H., & Lazarides, E. (1981) *Cell 23*, 524–532.

Gomer, R. H., & Lazarides, E. (1983) *J. Cell Biol. 96*, 321–329.

Gorlin, J. B., Kwiatowski, D. J., Bruns, G., Egan, S., & Hartwig, J. H. (1989a) *J. Cell Biol. 107*, 31a (Abstr.).

Gorlin, J., Yamin, R., & Kwiatkowski, D. J. (1989b) *J. Cell Biol. 109*, 272a (Abstr.).

Hartwig, J. H., & Stossel, T. P. (1975) *J. Biol. Chem. 250*, 5696–5705.

Hartwig, J. H., & Shevlin, P. (1986) *J. Cell. Biol. 103*, 1007–1020.

Hartwig, J. H., Niederman, R., & Lind, S. E. (1985) *Subcell. Biochem. 11*, 1–49.

Hock, R. S., & Condeelis, J. S. (1987) *J. Biol. Chem. 262*, 394–400.

Koenig, M., Monaco, A. P., & Kunkel, L. M. (1988) *Cell 53*, 219–228.

Koteliansky, V. E., Glukhove, M. A.,Shirinsky, V. P., Smirnov, V. N., Bushueva, T. L., Filimonov, V. V., & Venyaminov, S. Y. (1982) *Eur. J. Biochem. 121*, 553–559.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Mangeat, P. H., & Burridge, K. (1983) *Cell Motil. 3*, 657–669.

Moore, S. (1963) *J. Biol. Chem. 238*, 235–237.

Noegel, A., Witke, W., & Schleicher, M. (1987) *FEBS Lett 221*, 391–396.

Noegel, A. A., Rapp, S., Lottspeich, F., Schleicher, M., & Stewart, M. (1989) *J. Cell Biol. 109*, 607–618.

O'Farrell, P. H. (1975) *J. Biol. Chem. 250*, 4007–4021.

Okita, J., Pidard, D., Newman, P., Montgomery, R., & Kunicki, T. (1985) *J. Cell Biol. 100*, 317–321.

Penke, B., Ferenczi, R., & Kovacs, K. (1974) *Anal. Biochem. 60*, 45–50.

Pollard, T. D., & Cooper, J. A. (1986) *Annu. Rev. Biochem. 55*, 987–1035.

Rosenberg, S., Stracher, A., & Lucus, R. (1981) *J. Cell Biol. 91*, 201–211.

Roustan, C., Boyer, M., Fattoum, A., Jeanneau, R., Benyamin, Y., Roger, M., & Pradel, L. (1982) *Eur. J. Biochem. 129*, 149–155.

Schloss, J., & Goldman, R. D. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4484–4488.

Schwartz, R. M., & Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., Ed.) Vol. 5, Suppl. 3, pp 353–358, National Biomedical Research Foundation, Silver Spring, MD.

Shizuta, Y., Shizuta, H., Gallo, M., & Pastan, I. (1976) *J. Biol. Chem. 251*, 6562–6567.

Sobue, K., Morimoto, K., Kanda, K., Maruyama, K., & Kakiuchi, S. (1982) *FEBS Lett. 138*, 289–292.

Speicher, D. W., Morrow, J. S., Knowles, W. J., & Marchesi, V. T. (1982) *J. Biol. Chem. 257*, 9093–9101.

Speicher, D. W., Davies, G., Yurchenco, P. D., & Marchesi, V. T. (1983) *J. Biol. Chem. 258*, 14931–14937.

Stossel, T. P. (1984) *J. Cell Biol. 99*, 15s–21s.

Stossel, T. P. (1989) *J. Biol. Chem. 264*, 18261-18264.
Truglia, J. A., & Stracher, A. (1981) *Biochem. Biophys. Res. Commun. 100*, 814-822.
Tyler, J. M., Anderson, J. M., & Branton, D. (1980) *J. Cell Biol. 85*, 489-495.
Wallach, D., Davies, P. J. A., & Pastan, I. (1978a) *J. Biol. Chem. 253*, 3328-3335.
Wallach, D., Davies, P. J. A., & Pastan, I. (1978b) *J. Biol. Chem. 253*, 4739-4745.
Wang, K. (1977) *Biochemistry 16*, 1857-1865.
Wang, K., & Singer, S. J. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2021-2025.
Wang, K., Ash, J. F., & Singer, S. J. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 4483-4486.
Weihing, R. R. (1983) *Biochemistry 22*, 1839-1847.
Weihing, R. R. (1985) *Can. J. Biochem. Cell Biol. 63*, 397-413.
Weihing, R. R. (1988) *Biochemistry 27*, 1865-1869.
Weihing, R. R., & Franklin, J. S. (1983) *Cell Motil. 3*, 535-543.

Solution Structure of the Nogalamycin–DNA Complex†

Xiaolu Zhang and Dinshaw J. Patel*
Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Received May 21, 1990; Revised Manuscript Received July 10, 1990

ABSTRACT: The nogalamycin–d(A-G-C-A-T-G-C-T) complex (two drugs per duplex) has been generated in aqueous solution and its structure characterized by a combined application of two-dimensional NMR experiments and molecular dynamics calculations. Two equivalents of nogalamycin binds to the self-complementary octanucleotide duplex with retention of 2-fold symmetry in solution. We have assigned the proton resonances of nogalamycin and the d(A1-G2-C3-A4-T5-G6-C7-T8) duplex in the complex and identified the intermolecular proton–proton NOEs that define the alignment of the antitumor agent at its binding site on duplex DNA. The analysis was greatly aided by a large number of intermolecular NOEs involving exchangeable protons on both the nogalamycin and the DNA in the complex. The molecular dynamics calculations were guided by 274 intramolecular nucleic acid distance constraints, 90 intramolecular nogalamycin distance constraints, and 104 intermolecular distance constraints between nogalamycin and the nucleic acid protons in the complex. The aglycon chromophore intercalates at (C-A)·(T-G) steps with the long axis of the aglycon approximately perpendicular to the long axis of the flanking C3·G6 and A4·T5 base pairs. The aglycon selectively stacks over T5 and G6 on the T5-G6-containing strand with the aglycon edge containing OH-4 and OH-6 substituents directed toward the C3-A4-containing strand. The C3·G6 and A4·T5 base pairs are intact but buckled at the intercalation site with a wedge-shaped alignment of C3 and A4 on the C3-A4 strand compared to the parallel alignment of T5 and G6 on the T5-G6 strand in the complex. The nogalose sugar in a chair conformation, the aglycon ring A in a half-chair conformation, and the $COOCH_3$-10 side chain form a continuous domain that is sandwiched within the walls of the minor groove and spans the three base pair (G2-C3-A4)·(T5-G6-C7) segment. The nogalose ring is positioned in the minor groove such that its nonpolar face is directed toward the G6-C7 sugar–phosphate backbone while its polar face containing OCH_3 groups is directed toward the G2-C3 sugar–phosphate backbone in the complex. The intermolecular contacts include a nonpolar patch of aglycon (CH_3-9) and nogalose (CH_3-3′) methyl groups forming van der Waals contacts with the base–sugar residues in the minor groove and intermolecular hydrogen bonds involving the amino groups of G2 and G6 with the ether oxygens OCH_3-3′ and O7, respectively, on the nogalose sugar. The bicyclic amino sugar adopts a chair conformation and is positioned in the major groove. The intermolecular contacts include a hydrophobic patch involving the bridgehead H1″ and CH_3-5″ protons on the bicyclic amino sugar and the CH_3 groups of T5, establishing the contribution of the thymidine to intercalation specificity at (C-A)·(T-G) steps in DNA. A pair of intermolecular hydrogen bonds stabilizes the alignment of the bicyclic amino sugar in the major groove and is formed between the bicyclic amino sugar OH-2″ and OH-4″ protons which are directed toward the N7 and O6 atoms of G6 in the complex. The nogalose and bicyclic amino sugars of nogalamycin move toward each other on complex formation with DNA in order to maximize the intermolecular contacts in both grooves of the duplex. The above details of the solution structure of the nogalamycin–DNA complex provide insights into the intermolecular contacts involving the drug's intercalating, minor groove, and major groove components and explain the sequence specificity of this antitumor agent for intercalation between G·C and A·T pairs at the (G-C-A)·(T-G-C) trinucleotide segment on the d(A-G-C-A-T-G-C-T) duplex.

Nogalamycin (**1**), isolated from *Streptomyces nogalator*, is a member of the anthracycline family of antibiotics, which exhibit significant antitumor and antibacterial activity (Bhuyan & Reusser, 1970; Li et al., 1979). Nogalamycin binds to DNA and selectively inhibits DNA-directed RNA synthesis in vivo (Fok & Waring, 1972; Ennis, 1981).

† This research was funded by NIH CA-46778.

0006-2960/90/0429-9451$02.50/0