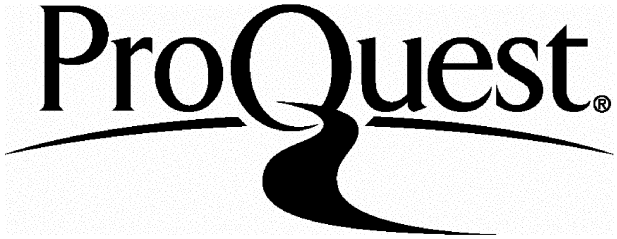
PURIFICATION AND SEQUENCING OF C4l: A NEW ACTIN-ASSOCIATED PROTEIN.

By: **MARTIN A. SMITH.**
Of: **UNIVERSITY COLLEGE LONDON.**

Thesis presented for the degree of **Doctor Of Philosophy**
at **University College London**, 1994.



ABSTRACT

The actin network is crucial for many cellular events such as cell movement, phagocytosis, cell division and movement of cell surface receptors. Control of the actin network lies with a large number of actin-associated proteins that regulate the polymerisation status, interaction and geometry of actin. Protein C4 is an actin-associated polypeptide doublet described by Shapland et al (1988; 1993). **C4^{l}** is the only protein C4 isoform present in motile cells such as lymphocytes and transformed mesenchymal cells, and the objectives of this work were to see if **C4^{l}** is associated with actin filament bundles in transformed cells and to characterise the molecule at the level of the gene.

My initial studies have shown that **C4^{l}** is associated with vestigial actin filament bundles found in transformed mesenchymal cells. To further investigate C4^{l} I have purified the molecule to more than 90% homogeneity from human T cell lymphoma (HTCL), and then used degenerate oligonucleotides to obtain internal **C4^{l}** sequence from reverse transcribed HTCL mRNA (RT-PCR). **C4^{l}** gene-specific oligonucleotides were used to obtain both the 5'- end (anchor-PCR) and 3'- end (3'- rapid amplification of cDNA ends) from reversed transcribed HTCL mRNA. Cloning and sequencing of these overlapping PCR products gives the full length coding sequence for **C4^{l}** from HTCL. The translated product of **C4^{l}** open reading frame is 199 amino acids in length, with a calculated molecular weight of 22,381Da and an estimated pI of 8.09. Northern blotting reveals that **C4^{l}** is expressed as a single message of 1.44 Kb which is apparently up-regulated in oncogenically transformed lymphocytes.

Database searches indicate that **C4^{l}** is a previously undescribed molecule. Regions of homology, including a potential actin-binding domain and phosphorylation sites, which are shared between **C4^{l}** and other proteins such as rat **transgelin** (Prinjha et al, 1994), chicken **SM22α** (Pearlstone et al, 1987), rat **NP25** (unpublished), chicken **calponins alpha** and **beta** (Takahashi & Nadal-Ginard, 1991), and *Drosophila* **mp20** (Ayme-Southgate et al, 1989) suggest that all of these proteins may be classified as members of a new transgelin-like multigene family.

ACKNOWLEDGEMENTS

I would like to acknowledge the help given to me by the following people:

Durward Lawson for allowing me to carry out my project within his group. The whole atmosphere of his laboratory was nothing less than friendly and inspiring. From him I have learned a great deal, and was urged to continue with my own ideas. His advice and supervision were excellent and his encouragement enthusiastic. I would like to thank him for a fulfilling an enjoyable time in his laboratory that I will never forget.

R. Prinjha for his practical help, lively discussion and friendship - thanks Rab!

C. Shapland for her constructive criticism and valuable comments throughout my project.

S. de Petris for allowing me to tap his vast knowledge of the cytoskeleton.

J. Hsuan and N. Totty for their microsequencing expertise that proved invaluable to this thesis.

I would like to thank the Medical Research Council for financial support during this study.

As ever in my life, my family has been a constant source of encouragement and support. I know they would not expect recognition, but I thank them for always being there for me.

Last, but not least, I would like to thank Sandra Wyatt for her endless patience and strong support under ever-changing conditions, and it is to her that I gratefully dedicate this thesis.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

A. CYTOSKELETON

The cytoskeleton of eukaryotic cells is involved in a number of fundamental cellular processes; these include cell motility, cell division, adherence, and the maintenance of cell shape and rigidity (Way & Weeds, 1990; Bray, 1992; Alberts et al, 1994).

The cytoskeleton is composed of three distinct groups of filamentous polymers: **microtubules** (Dustin, 1984; MacRae, 1992), **intermediate filaments** (see Shoeman & Traub, 1993; Stewart, 1993) and **microfilaments** (see Sheterline & Sparrow, 1994). The number of proteins associated with all these polymeric networks are increasing rapidly, and they are involved in a variety of processes such as the regulation of network structure, function and intermolecular interactions (Foisner & Wiche, 1991; Bloom, 1992; Kreis & Vale, 1993; Hartwig & Kwiatkowski, 1991).

A brief resume of some aspects of microtubules, intermediate filaments and their associated proteins are reviewed in sections B and C respectively. Microfilaments are reviewed in section D and functions of the cytoskeleton described in section E. As the molecule under investigation in this study ($C4^{l}$) is a proposed actin-associated protein, an extensive review of other such molecules is presented in section F.

B. MICROTUBULES

Bundled **microtubules** (**MTs**), arrayed in parallel and interconnected, are a structural feature of many eukaryotic cells (Dustin, 1984). The MT arrangements may be transient, as best exemplified by mitotic spindles and segregating chromosomes, or the MT assemblies may exhibit varying degrees of longevity, such as neuronal processes, the marginal band of erythrocytes, ciliary axoneme or flagellum, and the structural supports for cells (see Cleveland & Sullivan, 1984; Luduena et al, 1992; Caplow, 1992 for reviews). MTs are required for cell polarisation of migrating cells, small and highly motile cells, however, do not require MTs for this function (Schliwa & Honer, 1993). MTs can therefore be seen to be involved in the organisation of the cytoplasm, maintaining cell shape in concert with actin filaments (intro, section D), intermediate

filaments (intro, section C) and microtubule associated proteins (MAPs), which join with MTs and affect their spatial organisation (MacRae, 1992). MTs are also involved in the distribution of organelles (Vale et al, 1985), in concert with microtubule-based motor proteins such as kinesin and cytoplasmic dynein (see Bloom, 1992; Skoufias & Scholey, 1993 for reviews).

1. TUBULIN

Tubulin, the structural subunit of MTs, is a 100KDa dimer of two, ≈450 residue, 50KDa polypeptides designated α and β, which are 36-42% identical to each other (Luduena et al, 1992; Burns & Surridge, 1990). In mammals there appears to be five isotypes of α-tubulin and six of β-tubulin, the different isotypes therefore may account for the different functions of microtubules (see Luduena et al, 1992). Sequence alignments between α- and β-tubulins from different species identify the presence of conserved GTP-binding sites (Linse & Mandelkow, 1988), with the positively charged C-terminal regions of both α- and β-tubulin varying greatly. The C-terminal regions are the proposed sites for MAP binding (Littauer et al, 1986). The tubulin molecule is subjected to four known types of post-translational modification: (i) acetylation of α-tubulin, increasing MT stability (Piperno et al, 1987); (ii) tyrosylation and detyrosylation of some α-tubulin isotypes, perhaps playing a role in the association with the golgi apparatus (Skoufias et al, 1990); (iii) phosphorylation of β_{III}-tubulin isotype (Alexander et al, 1991); (iv) glutamylation of α- and β_{III}-tubulin (Alexander et al, 1991). γ-tubulin has been described in a variety of organisms (Stearns et al, 1991) and is about 35% identical to both α- and β-tubulin with conserved GTP-binding sites. It is associated with the microtubule organising centres (MTOCs) and is apparently required for the assembly of MTOCs in vivo (Stearns et al, 1991). The MTOCs are required for tubulin assembly and MT formation (Brinkley, 1985). The assembly of tubulin is promoted by MAPs, GTP, Mg^{2+}, elevated temperature (37°c) (Dustin, 1984), and the drug taxol (Bray, 1992). MTs are built up from protofilaments, strings of α- and β-tubulin; with the most common arrangement being 13 protofilaments forming a ≈25nm diameter, hollow, MT (Dustin, 1984).

MTs show the properties of dynamic stability and spatial pattern formation; even after assembly to an overall steady-state, individual MTs continue to switch between

phases of disassembly and reassembly (Caplow, 1992). When the phases are synchronised, the solution oscillates and is capable of forming various scattered structures (Caplow, 1992). The energy of the dynamic behaviour is derived from the GTP hydrolysis during assembly (Caplow, 1992). Assembly takes place mainly by endwise addition of tubulin heterodimers (at the GTP-cap), disassembly by oligomers which subsequently disintegrates into dimers (Caplow, 1992).

A large number of MAPs have been described with the potential to cross-link or bundle microtubules (see Bloom, 1992; Skoufias & Scholey, 1993; MacRae, 1992; Kreis & Vale, 1993 for reviews). By definition, MAPs co-purify with tubulin through cycles of assembly and disassembly (kinesin and dynein are associated with MTs in vivo but do not recycle with them in vitro) (Luduena et al, 1992). One common structural theme among MAPs, such as tau (intro, section F. 19), is that many of them contain 3-4 imperfectly repeated sequences, which have been shown to bind to and stimulate tubulin polymerisation into MTs (Goedert & Jakes, 1990). The fact they exist as repeats suggests that, if each on binds to a tubulin molecule, then a single MAP molecule can interact with several tubulins and thereby stimulate MT assembly (Luduena et al, 1992). It is clear that MAPs exert a profound influence on cells (MacRae, 1992). They control the spatial organisation of MTs and this may be of more significance as a determinant of MT function than the isotubulin composition. Movement of cells, intracellular transport of membrane-bound organelles and chromosomes, generation of cell-shape and polarity all depend on MT organisation (MacRae, 1992).

Microtubule-based motor proteins are enzymes that couple nucleotide hydrolysis to movement along MTs (Skoufias & Scholey, 1993). **Kinesin** and cytoplasmic **dynein** families are microtubule-associated MgATPases with motor activity (Skoufias & Scholey, 1993), that have nucleotide- and MT-binding sites, and display MT-activated nucleotidase activity. The energy released from nucleotide hydrolysis is used to generate force and motion relative to MTs (Skoufias & Scholey, 1993). MT motor polypeptides are generally organised into globular motor domains, which split ATP and generate movement along MTs, with tails which are adapted to attach to particular structures such as organelles, chromosomes and vesicles to be moved along the MTs (reviewed in

Bloom, 1992; Skoufias & Scholey, 1993; Vallee, 1991; Goldstein, 1991; Kreis & Vale, 1993).

C. INTERMEDIATE FILAMENTS

Although **intermediate filament (IF)** protein molecules differ in sequence (Klymkowsky et al, 1989), they share three common characteristics: (i) the ability to polymerise into 10nm filaments, (ii) a multidomain secondary structure consisting of a central rod domain of about 310 residues (350 for nuclear lamins) with high propensities to form an α-helix as well as non-α-helical head and tail regions, and (iii) their gene organisation, which provides a convenient system for classification (Shoeman & Traub, 1993). The non-α-helical head and tail domains of the individual proteins, which are variable in size and sequence, are responsible for determining the unique characteristics of the individual subunit proteins (Shoeman & Traub, 1993). The central rod domain is highly invariant in its amino acid sequence: within members of a single class, the sequence homology is about 70%, while even between classes the homology remains at 30-50% (Shoeman & Traub, 1993). This feature is the reason why the members of this group of proteins all polymerise into a common higher order structure: the 10nm filament (see Stewart, 1993). IFs have been classified into six types but may be conveniently divided into three broad groups (Stewart, 1993; and below). Within any one tissue or cell type, the number of IF proteins actually expressed is normally restricted to one or a few of the cytoplasmic subunits (Stewart, 1993).

IFs are able to form 10nm filaments spontaneously, and so, must encode all the structural information necessary for such assembly (Stewart, 1993). These filaments are rope-like structures made up of ≈4.5nm protofibrils, which in turn are made up of smaller fibres called protofilaments. Protofilaments contain linear arrays of IF dimer subunits arranged in head-to-tail fashion (reviewed by Fuchs, 1994). The IF dimer is a coiled-coil of two in-register molecules aligned in a parallel fashion, the formation of which, is guided by heptad repeats of hydrophobic amino acids within the central rod domain (Parry & Steinert, 1992). The N-terminal domains contain phosphorylation sites, and it is likely that phosphorylation is a major factor in the regulation of IF polymerisation, subcellular organisation and dynamics (Eriksson et al, 1992; Quinlan & Stewart, 1991).

Little is known about the exact function of IFs, they are less dynamic than microfilaments and MTs, suggesting a role in structural support to the cell (Schliwa & Honer, 1993). Indeed, there is little evidence to suggest that cellular locomotion depends on IFs; the pseudopod in general is devoid of IFs (Schliwa & Honer, 1993). In addition to the mechanical stability of cells, IFs may also serve as attachment sites for other cellular structures (Eriksson et al, 1992 and references therein).

Members of the different groups and types of IFs are briefly described below.

1. GROUP A

1.1. Types I and II

The **cytokeratins** are the first IFs expressed during development (Kreis & Vale, 1993 for review). The human genome contains at least 20 different genes encoding epithelial cytokeratins, plus another ten cytokeratins characteristic of hair- and nail-forming cells, which are expressed in different patterns in the various cell types (see Kreis & Vale, 1993 for review). The individual cytokeratins differ in Mr (from 40-68KDa) and pI (from 4.8-8). They can be divided into two subfamilies: more acidic **type I** and the more basic **type II**, sharing only < 30% identity (Steinert & Roop, 1988). A characteristic of cytokeratins not shared by other IFs is their ability to form bundles, which is increased by certain cytokeratin-associated proteins such as **filaggrin** (Foisner & Wiche, 1991).

2. GROUP B

2.1. Type III

This group of proteins form IFs in various types of vertebrate non-epithelial cells.

Vimentin, a 54KDa IF protein, is characteristic, but not restricted to fibroblasts and mesenchymally-derived cell types (Kreis & Vale, 1993). It is a substrate for certain protein kinases, providing a means by which the cell can alter the otherwise insoluble vimentin network (Kreis & Vale, 1993). Vimentin expression is increased by rapid growth, indicating growth regulation (Ferrari et al, 1986).

Desmin is a 53KDa IF protein characteristic of myogenic cells (Raats et al, 1990). It is most abundant in smooth muscle where it forms an interconnected meshwork linking cytoplasmic dense bodies with dense plaques on the membrane, thus forming a structural lattice that provides support for the contractile machinery (Raats et al, 1990).

Glial fibrillary acidic protein (GFAP), a 49.8KDa IF protein (Reeves et al, 1989), is expressed almost exclusively in astrocytes and cells of astro-glial origin, forming 10nm filaments throughout the cytoplasm, from the nucleus to the cell periphery (reviewed by Kreis & Vale, 1993). GFAP filaments appear to perform an active structural role in process formation of cultured astrocytes (Kreis & Vale, 1993).

Peripherin, a 57-58KDa IF protein, is detected in well defined sets of neurons where it is coexpressed with neurofilament proteins (Djabali et al, 1991). The cDNA of mouse peripherin indicates three isoforms that arise by alternative splicing (Djabali et al, 1991). The expression of peripherin in neurons with peripheral axons and its attachment to nuclear membrane suggest a role in signal transduction (Djabali et al, 1991).

2.2. Type IV

The neurofilament triplet proteins are recognised as the major components of the **neurofilaments (NFs)**, the IFs typical for most mature neurons (Liem, 1990). NFs consist of three proteins (Liem, 1990) with molecular weights of 68KDa (NF-L), 160KDa (NF-M) and 200KDa (NF-H). NF-L, the major triplet protein, forms IF in vitro either on its own or co-assembles together with NF-M and NF-H (Liem, 1990).

α-internexin, a novel 68KDa neuronal IF is found predominantly in the CNS (Chiu et al, 1989). cDNA cloning and sequencing has shown that it belongs to the type IV IFs (Fliegner et al, 1990). α-internexin can bind to a number of other IF proteins (Fliegner et al, 1990) in vitro via the α-helical rod region.

2.3. Type VI

Nestin, a 200KDa IF protein, detected in the CNS, is not closely related to other IFs. Although nestin possesses the coiled-coil core region, it is only 16-29% identical to other IF core sequences, and has virtually no N-terminal head domain (reviewed in Kreis & Vale, 1993).

3. GROUP C

3.1. TypeV

Lamins comprise a distinct class of IF proteins that are assembled into a filamentous network lining the nuclear lamina, and are characterised in a wide variety of higher eukaryotes and range in Mr between 60-75KDa (see Kreis & Vale, 1993) that have been classified into 'A' and 'B' subgroups (Nigg, 1992). Lamins are similar to other

IFs in structure, but do not occur in vivo as 10nm filaments (Nigg, 1992). The nuclear lamina is thought to provide a framework for organising the structure of the nuclear envelope, thus lamins may play a role in defining nuclear architecture (Nigg, 1992).

Besides the major structural subunits that make up 10nm filaments, there are also a diverse array of associated proteins, the IFAPs, that maybe responsible for regulating IF supramolecular organisation and function (reviewed by Foisner & Wiche, 1991). These proteins are often specific to the cell and tissue type, and are classified as IFAPs because of their co-purification with IF proteins and colocalisation by immunomicroscopy (see **epinemin** (Lawson, 1983) for example).

D. MICROFILAMENTS

The third major cytoskeletal network is formed by microfilaments. These are 9-10nm in diameter (Milligan et al, 1990) and are composed of linear polymers of **actin** molecules. A large number of proteins are associated with actin filaments (Hartwig & Kwiatkowski, 1991), and their roles are discussed in section E.

1. ACTIN INTRODUCTION

Actin is a globular protein (G-actin) of Mr ≈ 43KDa found in all eukaryotes at a concentration of ≈ 4mg/ml (Bray & Thomas, 1975) in the cytoplasmic compartment and also in the nucleus (Sheterline & Sparrow, 1994). It polymerises into filaments (F-actin) for most of its biological functions (Korn et al, 1987; Carlier, 1991). The location, extent, polarity and timing of assembly is regulated by a large number of actin-binding proteins (ABPs) (intro, section F) which are ultimately regulated by signal transduction pathways initiated at the cell surface (reviewed by Luna & Hitt, 1992; Ridley, 1994).

The actin cytoskeleton is highly dynamic in most cells with measured half lives for filament populations in the order of minutes (Carlier, 1991). Each actin monomer catalyses the hydrolysis of one molecule of ATP during a single cycle of assembly and disassembly (Carlier, 1991). Thus, whilst the characteristic pattern of actin filaments in particular cell types exist continuously, the individual filaments that comprise these structures are continuously changing (Carlier, 1991).

The control of F-actin assembly and turnover is important for many cellular processes such as motility, morphological changes, cell division, intracellular movements

and exertion of force by cells on the extracellular matrix (Bray, 1992; Pollard, 1990; Davies & Tripathi, 1993). The assembly of filaments into the correct polarity and spatial organisation is important to form tracks along which myosin (intro, section F. 5) can generate force in the appropriate direction, for example the contraction of striated muscle (Cooke, 1986), and cytokinesis (Satterwhite & Pollard, 1992). Shape change and motility in the absence of myosin can also be generated by the microfilament network, for example ruffling of leading lamellipodia and phagocytosis (Sheterline & Sparrow, 1994).

2. ACTIN STRUCTURE

The actin molecule, with dimensions of 6.7x4.0x3.7nm (Kabsch et al, 1990), is effectively divided into two domains of roughly equivalent size by a cleft containing a single, exchangeable high-affinity divalent-cation site, with three or more intermediate and lower affinity cation-binding sites elsewhere in the molecule (Estes et al, 1992). The cleft is also the site for the bound nucleotide (Kabsch et al, 1990). The two domains created by the cleft are covalently linked by two strands of the peptide chain which are close together at the base of the molecule, allowing the potential for significant relative movements of the domains. The two domains can themselves be divided into smaller subdomains: IA, IB, IIA and IIB. Subdomains IA and IIA represent the barbed end of the molecule when in the form of a polymer (Sheterline & Sparrow, 1994). The F-actin 3D model has been determined (Holmes et al, 1990) and reveals that the monomers in the filament are arranged helically with a pitch of about 70-80nm (Holmes et al, 1990), with each monomer orientated the same way round with respect to the long axis of the filament. The filament is thus polarised in the sense that both ends of the filament expose opposite ends of the actin molecule and the same edge of each actin molecule is exposed on the outside of the filament. The overall polarity of the filament is defined by the binding pattern of myosin as the barbed or pointed ends. The barbed end appears to dominate assembly and disassembly events in cells (see Sheterline & Sparrow, 1994).

3. ACTIN POLYMERISATION

At low ionic strength, actin exists in the G- form, which has a single ATP occupying the cleft nucleotide-binding site (Janmey, 1991). It is well established that filaments are formed by the helical polymerisation of the G-actin monomers (Pollard, 1990). This mechanism involves the energetically unfavourable **nucleation** step followed

by **elongation** off the nuclei by endwise addition of subunits. Kinetic analyses of polymerisation curves indicate that the nucleus is a trimer (Pollard & Cooper, 1986), in agreement with the helical structure of the filament (Holmes et al, 1990). The standard model of nucleation involves monomer activation by binding Mg to intermediate-binding sites followed by the formation of dimers and trimers (Pollard, 1990). The formation of the nuclei is the rate-limiting step of polymerisation (Pollard, 1990).

Elongation of the filament consists of subunit addition and loss at the two ends of the polymer (Pollard, 1990). Association of subunits is very rapid at both ends and is limited only by diffusion at the rapidly growing barbed end (Drenckhahn & Pollard, 1986). There is a difference between the critical concentration (the concentration above which extra actin assembles into filaments) at opposite ends of the same filament due to the different relative rates of monomer exchange at the pointed and barbed ends; rates are some 20-fold faster at the barbed end (Korn et al, 1987). Subunit dissociation is relatively slow at both ends (Pollard, 1990), with the nucleotide bound to the associating or dissociating subunit affecting both the on and off reactions. ATP-bound subunits associate more rapidly than subunits bound to ADP at both ends. At the barbed end ADP-subunits dissociate more rapidly than ATP-subunits, with the opposite true at the pointed end (Pollard, 1990). Calculations indicate that subunits flux through actin filaments at steady-state in the presence of ATP (known as treadmilling) (Selve & Wegner, 1986); this flux will be damped by the slow rate of dissociation of ADP-actin at the pointed end. The actin filaments are able to join end-to-end in a final stage of **annealing** to form the microfilament (Pollard & Cooper, 1986).

As briefly mentioned above, actin assembly is accompanied by the stoichiometric hydrolysis of ATP (Carlier, 1991; Pollard, 1990). ATP bound to assembling monomers is hydrolysed not during the assembly step, but subsequent to incorporation within the filament to yield an ADP-actin filament (Korn et al, 1987). Once the monomer is incorporated into the polymer, there is a finite probability that the ATP will be hydrolysed. The immediate product of hydrolysis is a monomer carrying ADP.Pi in the filament. The release of phosphate has a lower probability than hydrolysis, thus after a period of time the filament which assembled as ATP-F-actin becomes ADP-F-actin via an ADP.Pi-F-actin intermediate (Carlier, 1987). These observations predict that under

conditions where the rate of assembly exceeds the rate of hydrolysis and phosphate release, an individual filament will consist of co-linear segments containing respectively ATP, ADP.Pi or ADP monomers in the same filament, in proportions according to the relative times monomers have been present since assembly and the rate constants of hydrolysis and Pi release (Sheterline & Sparrow, 1994). The slower rate of addition at the pointed end allows a high probability of ATP hydrolysis on the terminal monomer before the next subunit adds on, whilst the converse is generally true at the barbed end (Coue & Korn, 1986). Thus, the pointed end will have a terminal ADP monomer, the barbed end an ATP (or possibly ADP.Pi) monomer. The probable function of ATP hydrolysis after actin filament assembly is that it facilitates their rapid depolymerisation at a later time (Pollard, 1990).

4. ACTIN GENES AND ISOFORMS

4.1. Class I and Class II Actins

Actins form a family (≈ 103 unique actins are known) of eukaryotic proteins whose sequences (375 (-8 to +5) amino acids) are highly conserved (see Sheterline & Sparrow, 1994 for alignments). The variable sequence at the N-terminus has provided the simplest classification criteria, but there are key amino acid positions that are characteristic of actins from different kingdoms and phyla (Sheterline & Sparrow, 1994). Using N-terminal differences, actins can be divided into two classes (Herman, 1993): **class I** actins, the non-muscle isoforms, which present N-acetyl aspartic or glutamic acid residues. **Class II** actin genes encode muscle isoactins that contain an N-terminal methionine residue followed by a cysteine and then an aspartic acid or glutamic acid (Herman, 1993). Vertebrate actins have also been categorised into different isoforms (Vandekerckhove & Weber, 1979), isoelectric focusing separates these isoforms into three isoelectric variants, referred to as α, β and γ actin (see Kreis & Vale, 1993). In both class I and class II isoactins, the original methionines (and cysteines of class II) are acetylated and cleaved (Herman, 1993), and the new N-terminus is then re-acetylated (see below). In yeast, the N-terminal processing is not carried out in the same way as in higher organisms (Cook et al, 1992). The N-terminus, which is located in major subdomain I (Holmes et al, 1990), is likely to modulate isoform-specific functions,

including actin-actin, acto-myosin and actin-ABP interactions (Herman, 1993; Milligan et al, 1990).

Other amino acid substitutions are distributed in regions throughout the actin sequences. These regions include: the helix in subdomain IIB involved in tropomyosin-binding, the proposed hydrophobic plug involved in interstrand interactions on subdomain IIA, and proposed as a major self assembly site (reviewed by Sheterline & Sparrow, 1994). The regions of variability, however, are few, with only 15.2% of positions in the amino acid sequence being invariant in all class I and II actins (Sheterline & Sparrow, 1994). Skeletal muscle α-actins in human, mouse, rat, rabbit and chicken are identical; as are the cytoplasmic β-actins in human, mouse, rat, cow and chicken (Hennessey et al, 1993).

4.2. Actin-Like Sequences: Class III Actins

Although most actins are highly conserved in sequence and length, proteins with more divergent sequences, known collectively as actin-like, or **class III** proteins have been identified in several species (Hennessey et al, 1993; Herman, 1993). In addition to its single actin gene **ACT1**, *S. cerevisiae* contains the **ACT2** gene which encodes a protein of 391 residues that is 53% divergent from ACT1 actin (Schwob & Martin, 1992). *S. pombe* has an ACT2 protein (427 residues) which is 52.6% diverged from ACT1 actin (Lees-Miller et al, 1992b). Genetic experiments show that ACT1 cannot substitute for ACT2, or vice versa, arguing that both genes are essential for cell viability and have different functions (Schwob & Martin, 1992; Lees-Miller et al, 1992b). Two identical actin-like sequences, with an identity of about 55% compared to class I and II actins and 40% identity to the ACT2 genes of yeast, have been isolated and studied in man (Lees-Miller et al, 1992a) and dog (Clark & Meyer, 1992). These sequences, named respectively **actin-RPV** (related protein vertebrate) and **centractin**, are associated with microtubule based cellular motor systems (Lees-Miller et al, 1992a; Clark & Meyer, 1992). Class III actins have also been reported in *Aspergillus*, *Physarum* and *Caenorhabditis* (Schwob & Martin, 1992). Like all actins the class III proteins conserve a region corresponding to the ATP-binding site (Kabsch et al, 1990).

4.3 Isoform Function

Actin isoforms segregate into different structures within cells (Herman, 1993), suggesting they have specialised functions. Purified isoforms, and actins from different species, can differ in thermodynamic stability and polymerisation characteristics (see Hennessey et al, 1993). Actins from different species also differ in their affinity for profilin, DNaseI, α-actinin, myosin and tropomyosin (Hennessey et al, 1993).

The physiological significance of isoform differences can be tested by substituting one actin for another in a whole cell. A mutation in the promoter sequence, which regulates expression, of the mouse cardiac actin gene reduces its expression to 16.5% of its normal level (Garner et al, 1986). There is a corresponding increase in the expression of the α-skeletal muscle actin gene. Despite containing predominantly the skeletal muscle actin isoform the heart functions normally, suggesting the two isoforms are functionally equivalent (Garner et al, 1986). This functional equivalence can also be seen in *Drosophila* when only two out of nine single amino acid substitutions in the **Act88F** gene produce an altered muscle structure (Reedy et al, 1991).

The isoforms of actin in muscle and non-muscle cells are expressed in temporally and spatially regulated patterns (Rubenstein, 1990). During skeletal musclogenesis and development of heart, α-skeletal, α-cardiac and α-vascular smooth muscle aswell as non-muscle β- and γ-actins are expressed to varying degrees (Herman, 1991), with the isoactin expression dependent on tissue and cell type.

There is evidence to indicate that there is functional diversity between isoforms (see Herman, 1993; Hill & Gunning, 1993). By use of isoform specific antibodies, isoactin sorting can be observed: non-muscle isoforms are localised at the sub-plasma membrane region, with γ-actin associated with the mitochondria (Herman, 1993); skeletal and cardiac myocytes restrict α-vascular actin within the stress fibres (Herman, 1993); also, there is 3 to 4-fold more skeletal muscle actin incorporated into myofibrils than β-non-muscle actin (Peng & Fischman, 1991). Examination of endothelial cell locomotion after injury reveals that the β-actin isoform is exclusively distributed within the leading lamellae, membrane ruffles and psuedopods (Hoock et al, 1991). β-actin localisation persists only in actively motile membrane domains. This behaviour suggests that β-actin

nucleation and filament assembly from the plasma membrane influences the location of forward protrusions and membrane ruffles (Hoock et al, 1991).

The translation of specific actin isoforms within discrete cytoplasmic domains is likely to be influenced by mRNA position (Singer, 1992). Probes which identify all isoforms show actin mRNA at the motile periphery of chick fibroblasts, myoblasts and in association with the cytoskeleton (Sundell & Singer, 1991). mRNA association with the cytoskeleton for transportation and positioning is probably the mechanism used in both development and in mature cells (Herman, 1993). Mammalian endothelial cells localise β-actin mRNA to peripheral regions of motility (Hoock et al, 1991). Transport of mRNA is probably a process involving either the 3'-untranslated region (3'-UTR) sequence or its secondary structure binding to a carrier (Hill & Gunning, 1993). The β- and γ-actin isoforms share no sequence homology in their 3'-UTRs (Erba et al, 1986). If the 3'-UTRs of these two mRNAs have a specific function, then it is possibly different for each of these actin isoforms (Hill & Gunning, 1993).

The microfilaments of the cell cytoskeleton are made from the two non-muscle (class I) cytoplasmic β- and γ-actin isoforms. While both actin isoforms are very similar, differing by only four amino acids at the N-terminus (Herman, 1993), they have different roles in overall cell cytoarchitecture. The β:γ ratio within cells is strictly regulated and perturbation of this ratio in myoblasts by transfection with either isoform has significant and opposite effects (Schevzov et al, 1992). Increasing β-actin results in larger cells whereas increasing γ-actin decreases cell size (Schevzov et al, 1992). Myoblasts therefore differentiate between isoforms and use the proteins differently (Hill & Gunning, 1993). The above combined approaches all point to the functionally diverse roles that the α-, β- and γ-actin genes play in producing isoactins that coordinate contractile potential (α), cell motility (β), or cytoskeletal array (γ) (Herman, 1993).

5. POST-TRANSLATIONAL MODIFICATION

The actin molecule is modified after translation by a variety of methods.

N-terminal processing appears to proceed during or shortly after translation (Hennessey et al, 1993). The major consequence of processing is the addition of an acetyl group, the functional significance of which is unknown (Hennessey et al, 1993). The

precise mechanisms of acetylation differ due to the different sequences in class I and II actins (Rubenstein, 1990).

Most actins contain a methylated histidine at the equivalent of position 73 of vertebrate skeletal muscle actin. Methylation occurs shortly after translation, and is of unknown function (Sheterline & Sparrow, 1994).

There are very few reports supporting a role for phosphorylation as a regulatory post-translational modification in vivo (Sheterline & Sparrow, 1994). Tyrosine phosphorylation of actin in *Dictyostelium* is associated with cell shape changes, the depolymerisation of subcortical actin, and disruption of focal contacts; whereas dephosphorylation is linked to the assembly of filaments and formation of spikes (see Carlier, 1993).

Ribosylated actin will not polymerise, ribosylated monomers capping the barbed ends of filaments (Aktories, 1990). The functional significance of such ribosylation is not yet known (Aktories, 1990).

6. ACTIN MICROFILAMENT ORGANISATION

A role for the actin cytoskeleton has been implicated in many cellular functions, including motility, chemotaxis, cell division, endocytosis and secretion (Bretscher, 1991). Precise temporal and spatial control of actin filament organisation is essential for these activities. In fibroblasts, actin filaments exist principally in three types of structure: (i) the cortical actin isotropic arrays (Alberts et al, 1994); (ii) contractile bundles, as found in stress fibres, the adhesion belt of epithelium and the contractile ring (Byers & Fujiwara, 1986; Satterwhite & Pollard, 1992); and (iii) parallel bundles forming protrusive structures such as ruffles, microspikes, microvilli and filopodia (Sheterline & Sparrow, 1994).

6.1. Cortical Actin Network

A layer of actin on the cytoplasmic surface of the cell membrane, known as the cortical layer seems to be a universal feature of eukaryotic cells (Bray et al, 1986). The cortex does not have a distinct filamentous structure but is arranged as a 3D network (Lawson, 1986). The network is rich in actin and is cross-linked at 0.05-0.1μm, the cross-links mediated by ABPs (Bray, 1992). Organelles are excluded from the cortex, thus,

during exocytosis the layer is probably signalled to move aside by contracting or disassembly, so that secretory vesicles can fuse with the cell membrane (Bray, 1992).

6.2. Contractile bundles: Stress Fibres

Stress fibres (SFs) are temporary, loosely spaced (30-60nm apart), bundles of actin arranged with opposite polarities. They are typically 0.5μm wide and about 5μm long (Langanger et al, 1986). SFs are intimately associated with myosin and other ABPs which may regulate their activity (Langanger et al, 1986).

If fibroblasts in tissue culture are confluent, actin can be seen to be present in the pattern of SFs which are connected to the plasma membrane at focal contacts (intro, section E. 2. 3), where clusters of integrin receptors bind to the extracellular matrix (ECM) proteins such as fibronectin and collagen (see Luna & Hitt, 1992 for review). The end of the SF is inserted into a second adhesion plaque or into a meshwork of IFs that surround the cell nucleus (Alberts et al, 1994).

SFs are associated with the retardation of cell motility and concomitant increase in substrate adhesion (Janmey, 1991). If SFs are cut from their attachment sites by laser, they can be made to contract when supplied with Mg^{2+} and ATP (see Bray, 1992). This force that SFs are capable of producing is thought to be involved in tension maintenance of the cell (Davis & Tripathi, 1993). SFs are also involved in the maintenance of cell shape: infection of fibroblasts with transforming virus (Shapland et al, 1988) causes SF disruption and produces large scale changes in cell shape. There is also evidence to suggest that SFs are capable of creating cell surface domains which may segregate receptors and as a result inhibit their free diffusion (Byers et al, 1984).

Little is known of the mechanism of SF formation (Ridley & Hall, 1992). Some indication is provided by the fact that serum-starved Swiss 3T3 cells when stimulated by a variety of growth factors form new SFs and focal adhesions; this assembly may be regulated by the small GTP-binding protein rho in response to the growth factors (Ridley & Hall, 1992).

6.3. Protrusive Structures

These structures consist of parallel bundles of filaments that are orientated in the same polarity and are often closely spaced (10-20nm apart) (Bray, 1992). Microspikes are stubby protrusions from the actin cortex about 5-10μm long that contains about 20

filaments arranged in a loose bundle with their barbed ends pointing outwards (Small, 1989). The growing tip (growth cone) of a developing nerve cell axon extends even longer microspikes called filopodia which can be up to 50μm long (Small, 1989). Microvilli increase the surface area of cells (especially epithelial) and are finger-like structures forming what is known as the intestinal brush border (Mooseker, 1985). Each microvillus is ≈ 100nm in diameter and 1-2μm long; it is supported by a core of cross-linked actin filaments (Mooseker, 1985). The bundles are extensively crosslinked by fimbrin and villin (see refs in intro, section F. 1 and 2), and connected to the surrounding membrane by brush border myosin (see Titus 1993a). At the base of the microvillus the actin bundle is anchored into a specialised, stabilising, region of cortex known as the terminal web that contains a dense network of spectrin (Hitt & Luna, 1994) overlying a layer of IFs (Bray, 1992). The brush border cytoskeleton is not a static structure, in particular lengths of microvilli vary with time, indicating that actin filaments in the microvilli cores are able to grow (Mooseker, 1985).

E. CYTOSKELETAL FUNCTIONS

1. CELL MOTILITY

Cell movement plays a fundamental role in embryonic development and the maintenance of adult tissue (Bray, 1992). The crawling of cells is very complex and difficult to explain at the molecular level (Stossel, 1993). For example, different parts of the cell change at the same time. Actin is a ubiquitous element in such movements; it undergoes spatial and temporal transformations with each micron the cell advances, assembling into actin-containing processes, associating with focal contacts and forming contractile bundles with myosin (Stossel, 1993; Luna & Hitt, 1992; Zigmond, 1993).

All types of cell crawling involve a cyclic movement of actin molecules involved in three distinct stages (Stossel, 1993): (i) **extension** of the leading margin; (ii) **attachment** of the new extension with the substratum; (iii) **contraction** of the cortex that draws the remainder of the cell forward, using adhesion points as anchorage (Bray, 1992).

External signals must polarise the changes in cortical actin so that cells move in a specific direction (eg. chemotaxis) (Condeelis, 1993b). The initial event in the guidance

mechanism is probably a local stimulation of actin polymerisation (Cooper, 1991; Condeelis, 1993a). Newly formed actin filaments provide the skeletal framework for pseudopodial extensions, generating a sudden flurry of microspikes and ruffles on the cell surface. These become restricted to one end, due to a gradient of chemoattractant such as cAMP, as the cell becomes polarised and begins to migrate (Condeelis, 1993b).

The basic molecular mechanisms underlying the locomotion of cells along solid substratum are thought to be shared by different cell types, including those that appear morphologically distinct (Lee et al, 1993).

1.1. Retrograde Actin Flux.

Observations of rearward-moving particles, receptors and surface ruffles (Ridley, 1994) are thought to reflect a mechanism for locomotion (Lee et al, 1993). There has been a failure to detect a retrograde flow of lipid in locomoting cells to account for this phenomenon (Lee et al, 1993). It was found that plasma membrane lipid advance involves the dynamic meshwork of actin filaments (Bray & White, 1988). Retrograde actin flux could arise from the treadmilling of actin filaments, or an actomyosin contractile mechanism (Cooper, 1991; Lee et al, 1993).

1.2. Protrusion.

Extension, or protrusion, of the leading margin is the first step in cell locomotion, with the most active extending region of a moving cell being the lammellipodium (Bray, 1992). Three general mechanisms have been proposed for the generation of force for protrusive activity (Condeelis, 1993a): (i) Strong actomyosin contractions at the rear of a cell increases the hydrostatic pressure within the cytoplasm such that it is forced towards the tip of the lamella (Bray & White, 1988; Condeelis, 1993a). A localised weakening of the cortical cytoskeleton at the tip of the cell may then allow extension to occur. (ii) Severing of actin filaments leads to osmotic swelling of the actin gel within the lamella (Condeelis, 1993a) and would result in protrusion until a new equilibrium is reached between osmotic pressure and the strength of the actin meshwork (Lee et al, 1993). (iii) Actin polymerisation at the cell edge directly provides the necessary force for protrusion (Lee et al 1993). In many cells actin polymerisation has been shown to occur predominantly at the leading edge (Cooper, 1991). Protrusive forces may result from actin polymerisation with the action of myosin I motors (intro, section F. 5. 2) at the

leading edge (Lee et al, 1993). Binding of myosin I to the membrane-associated cortex may facilitate protrusion by 'walking' towards the barbed ends of actin filaments that are immobilised by linkages to the substratum, thus allowing actin filaments to slide relative to each other (Cramer & Mitchison, 1993). However, such sliding has not been detected in locomoting cells (Condeelis, 1993b). The above considerations indicate that the force derived individually from either actin polymerisation, the actomyosin interaction, or gel osmotic pressure is sufficient to account for the extension of pseudopods and filopods (Condeelis, 1993a).

1.3. Regulation of Locomotion.

Insights into the mechanism of protrusion implicates localised actin polymerisation, filament cross-linking and filament severing as major events during protrusive activity (reviewed by Stossel, 1993; Condeelis, 1993a).

Upon chemoattractant binding the localisation and timing of actin polymerisation in cells is probably regulated by the concerted release of actin from sequestered and/or inactive forms and by the regulation of nucleation sites for actin polymerisation. These reactions may be coordinated through ABPs, second messengers such as intracellular Ca^{2+}, and the metabolism of PPIs that is under the control of receptors that regulate phospholipase activity. Regulation of actin polymerisation may also involve small GTP-binding proteins (see Condeelis, 1993a; Ridley, 1994; Stossel, 1993; Carlier, 1993).

Cross-linking and severing are essential steps if protrusive force is to be generated. Only cross-linked filaments can be used to push against the membrane to form protrusion or to form a cortical gels that resists deformation of cell shape by surface tension (Condeelis, 1993b). In *Dictyostelium* and animal cells, it can be seen that actin filaments in the cell cortex, and extensions of it such as pseudopods and filopods, are cross-linked into networks and bundles (Condeelis, 1993b). Proteins such as α-actinin, filamin, ABP120, spectrin, fimbrin (intro, section F. 1) and myosin I (intro, section F. 5. 2) have all been identified as cortical cross-linking proteins (Stossel, 1993). Many of these cross-linking proteins probably localise in surface extensions after protrusion is complete so as to endow the actin cytoskeleton with properties such as membrane binding, substrate adhesion and rigidity (Stossel, 1993). Proteins, such as ABP280 and ABP120, may associate with polymerising filaments during network and bundle

formation, and contribute to the initial forces of protrusion (Condeelis, 1993a; Cunningham et al, 1992). Severing of cross-linked filaments accompanies the stimulation of extension (Hartwig, 1992). A gradient of intracellular Ca^{2+} is suggested in maintaining a polarised cell shape (Maxfield, 1993), and relatively low levels of Ca^{2+} at the extending edge are thought to promote lamellar severing proteins such as gelsolin (intro, section F. 2), which may then provide more sites for actin polymerisation (Lee et al, 1993). Localisation of the ABPs ponticulin (intro, section F. 19) and profilin (intro, section F. 13) at the extending edges of motile cells are thought to increase actin polymerisation in these regions (Luna & Hitt, 1992). At the rear of the cell, increases in Ca^{2+} are thought to promote retraction of the cell margin by increasing the contractility of the actin filament network (Lee et al, 1993). Retraction of the cell margin generates folds in the dorsal surface that provide surface area for subsequent lamellar extension (Lee et al, 1993).

1.4. Cell-Substrate Contacts.

Cell-substrate attachments allow traction forces generated by the cell to be exerted on the substratum (Lee et al, 1993). Focal contacts (intro, section E. 2. 3; Burridge et al, 1988) are found in slower-moving cells; in rapidly moving cells adequate cell contact is provided by diffuse, highly labile regions of close contact called podosomes (Marchisio et al, 1987). Lamellar formation can only succeed if the ventral surface is attached to the substratum; newly formed actin filaments must be anchored to the substratum (Lee et al, 1993). Cell-substrate contacts also transduce chemical and mechanical signals across the cell membrane that are important in regulating actin filament dynamics during motility (Luna & Hitt, 1992). Signalling events are triggered by the clustering of integrin molecules, which are mechanically coupled to the actin cytoskeleton, at the sites of cell-substratum contact (Hynes, 1992).

2. CELL ADHESIONS

There are four major families of receptors that mediate cell adhesion (Hynes & Lander, 1992): (i) cadherins; (ii) members of the Ig superfamily including N-CAM and I-CAM-1; (iii) proteins that bind to specific carbohydrate groups via their lectin domains called selectins; and (iv) integrins. Relatively little is known about the intercellular

interactions of members of the Ig superfamily, and the selectins are thought not to be attached to the cytoskeleton (Hynes & Lander, 1992). The junctional assemblies in which the cadherins and the integrins are mixed and matched with components of the actin and IF networks form four junctions that mediate cell-cell and cell-substrate attachments (Alberts et al, 1994; Luna & Hitt, 1992).

2.1. Cell-to-Cell Adherens Junctions (AJs).

The cadherin superfamily is a group of proteins that mediate Ca^{2+}-dependent homophilic adhesion between cells and cell sorting (Takeichi, 1993). Classical cadherins include E- (epithelial, also called uvomorulin), N- (neuronal, also called A-CAM), and P- (placental) cadherin, but at least 12 different members of the family are known (Takeichi, 1993). Classical cadherins each comprise of an extracellular domain, a transmembrane domain, and a highly conserved cytoplasmic domain (56-80% identity) (Takeichi, 1990). The cytoplasmic domains of the classical cadherins mediate interactions with the cytosolic catenins (see refs in intro, section F. 11).

Cells that express both cadherins and catenins form belt-like sites of cell-cell attachment, in the presence of millimolar extracellular Ca^{2+} called adhesion belts, zonula adherens or **AJs** (Takeichi, 1990). Cadherin mediated cell adhesion is not seen in cells that lack α-catenin or in cells expressing cadherins lacking the cytoplasmic domains (Hirano et al, 1992). Of the three catenins, β-catenin binds tightest to cadherin (Tsukita et al, 1992)

Actin is the principal component of the AJ undercoat (Alberts et al, 1994), and it is associated with ABPs such as α-actinin (intro, section F. 1), zyxin (intro, section F. 19), vinculin (intro, section F. 11), and radixin (intro, section F. 3) (Luna & Hitt, 1992; Hitt & Luna, 1994). Other proteins immunolocalised to the AJ undercoat include tenuin and the 220KDa protein ZO-1 (Hitt & Luna, 1994). Junctional integrity appears to be regulated by tyrosine phosphorylation (Tsukita et al, 1991). Elevated levels of phosphotyrosine rapidly, but reversibly, induce junctional disassembly (Luna & Hitt, 1992).

2.2. Desmosomes.

Desmosomes are dense, button-like points of intercellular contact that mediate Ca^{2+}-sensitive cell-cell adhesion (Schwarz et al, 1991) and serve as anchoring sites for

IFs (intro, section C). Intercellular attachment is mediated by transmembrane proteins: desmogelin I and the related splice-variants desmocollins I and II, which are members of the cadherin superfamily (Takeichi, 1993). The cytoplasmic domain of desmogelins and desmocollins possess unique characteristic features not found in classical cadherins, and probably reflects the fact that the desmosome is not associated with actin filaments, but with IFs (Luna & Hitt, 1992). Desmosomal plaque components include plakoglobin, which binds directly to the cytoplasmic domain of desmogelin I, and the desmoplakins (I and II) which are is involved in IF organisation (Stappenbeck & Green, 1992).

2.3. Focal Contacts.

Cell-matrix junctions enable cells to get a hold on the extracellular matrix (ECM) by connecting their actin filaments to the matrix (Alberts et al, 1994). Cultured fibroblasts migrating on an artificial substratum coated with ECM molecules, for example, grip the substratum at specialised regions of the plasma membrane called **focal contacts**, or **adhesion plaques** where actin stress fibres terminate (Geiger & Ginsberg, 1991). Although the number of focal contacts is inversely correlated to the rate of cell translocation, small areas at the leading edge of the cell may be required for traction during fibroblast locomotion (Luna & Hitt, 1992). The transmembrane adhesion proteins called integrins link ECM proteins on the substratum, such as fibronectin, to a dense cytoplasmic plaque associated with stress fibres (Hynes, 1992).

Integrins are a family of transmembrane $\alpha\beta$ heterodimers (approx 14 known α subunits and 8 β subunits) (Hynes, 1992). Many integrins have an affinity for the amino acid sequence RGD in their ECM ligand, and they also have a small cytoplasmic region that binds to elements of the actin cytoskeleton such as talin (intro, section F. 3) and α-actinin (Burridge et al, 1988; Hynes, 1992).

The interactions among the proteins at focal contacts appear to be very complex (Luna & Hitt, 1992; Hitt & Luna, 1994). The cytoplasmic domain of the integrin β-subunit binds talin and α-actinin. Talin also binds vinculin (intro, section F. 11) (Gilmore et al, 1993). Vinculin binds paxillin (intro, section F. 19), the C-terminus of α-actinin and may oligomerise with itself (Otto, 1990). Zyxin (intro, section F. 19) and actin both bind to the N-terminus of α-actinin (Crawford et al, 1992). Tensin (intro, section F. 9) which may bind vinculin, also binds and caps the barbed ends of actin filaments (Davis et al,

1991). Other proteins localised to focal contacts include tenuin and VASP (both intro, section F. 19) (Luna & Hitt, 1992).

Changes in protein phosphorylation apparently regulate the structure and function of focal contacts in vivo (Schaller & Parsons, 1993). Several components of the focal contacts are phosphorylated in response to ligand activation of integrin (see Hitt & Luna, 1994). These proteins include paxillin (Glenney & Zokas, 1989) and tensin (Davis et al, 1991) and the protein tyrosine kinase pp125^{FAK} (Schaller & Parsons, 1993).

2.4. Hemidesmosomes.

Hemidesmosomes (HDs) are relatively small domains (≈0.4μm in diameter) of the basal plasma membrane in epithelial and mesenchymal cells that mediate attachment to the underlying basal lamina (Schwarz et al, 1990). Ultrastructurally, HDs look like half desmosomes and contain a trilayered cytoplasmic plaque associated with IFs (Alberts et al, 1994). Attachment to the basal membrane, via the HD, is through fine anchoring filaments that contain three non-identical disulphide-linked proteins called epiligrin (see Luna & Hitt, 1992). A transmembrane protein known to localise at HDs is the $\alpha_6\beta_4$ integrin (Quaranta & Jones, 1991). The cytoplasmic domain of $\alpha_6\beta_4$ integrin is specific for interaction with IFs (Quaranta & Jones, 1991).

3. CYTOKINESIS

All cells have a finite life span, therefore, if an organism is to continue to live it must create new cells at a rate as fast as that at which its cells die. For this reason, cell division is central to the life of all organisms. The process of cell division consists of two sequential processes: nuclear division (mitosis) and the division of the mitotic cell by **cytokinesis** (see Salmon, 1989).

Most study of cytokinesis has concentrated on the 'sliding-filament contractile ring' model (see Satterwhite & Pollard, 1992 for review). The contractile ring is a transient bipolar array of actin filaments with their barbed ends attached to the plasma membrane at sites around the equator of the dividing cell (Satterwhite & Pollard, 1992). These filaments interact with bipolar myosin-II filaments (intro, section F. 1. 1), which are essential for cytokinesis as indicated by mutations to the RLC of *Drosophila* myosin-II (Karess et al, 1991) and *S. cerevisiae* MYOI (Rodriguez & Paterson, 1990). The contractile ring is confined to an equatorial band of cortex, it constricts the cell locally

and pinches it into two like a purse string (reviewed by Salmon, 1989; Satterwhite & Pollard, 1992)

Several proteins are found at the contractile ring, located at the cleavage furrow (Satterwhite & Pollard, 1992), that may contribute to the structure and dynamics of it. For example, as well as actin and myosin-II, inner centromere proteins of 135KDa and 150KDa (Earnshaw & Cooke, 1991), which are tightly associated with chromosomes during metaphase, move to the continuous microtubules of the anaphase spindle and then redistribute to the cortex just before the furrow forms (Earnshaw & Cooke, 1991). In the *Drosophila*, *serendipity-α*, *nullo*, *bottleneck* (intro, section F. 18), and *kelch* (intro, section F. 17) gene products are all located at ring canals, and stabilise these specialised cleavage furrows (Theurkauf, 1994; Xue & Cooley, 1993). Members of the ERM subfamily (intro, section F. 3), such as radixin, are highly concentrated in both the cleavage furrow and AJ (Tsukita, et al 1989). Actin filaments within the contractile ring are attached to the membrane via their barbed ends (Mabuchi et al, 1988), and radixin, a barbed end capping protein, could participate in the end-on association of actin filaments with plasma membranes (Sato et al, 1991). Barbed end association of actin filaments to the membrane is consistent with the contractile ring model because myosin-II is a barbed end directed motor (Satterwhite & Pollard, 1992).

The contractile ring consists of a very thin overlapping array of actin filaments with opposite polarities, this is important in the contractile ring model because the sliding-filament mechanism depends on tension generated by the interaction of bipolar myosin-II filaments with oppositely polarised actin filaments (Satterwhite & Pollard, 1992). Some filaments within the furrow appear to lie at random angles to the equator, thus forming a complex 3D network of filaments (Fishkind & Wang, 1993). This provides a structural framework for the action of myosin-II reflecting the direction of forces required for cell cleavage (Fishkind & Wang, 1993).

Formation of the contractile ring is part of a reorganisation process of the whole cytoskeleton during mitosis that begins long before cytokinesis (Cao & Wang, 1990). In cultured vertebrate cells, stress fibres breakdown during prophase and the constituent actin and myosin-II becomes concentrated at the equatorial cortex where the furrow will form (Mittal et al, 1987). Actin filaments of the contractile ring accumulate in the cortex

late in metaphase and then migrate by an unknown mechanism within the cortex to the cleavage furrow (Cao & Wang, 1990). The contractile ring undergoes contraction during early stages of furrowing then disassembles during completion of the cleavage; indeed, myosin-II filament disassembly is essential for contractile ring function (Egelhoff et al, 1991).

The formation of the cleavage furrow requires both a signal from the mitotic apparatus and a responsive cortex (Satterwhite & Pollard, 1992). The signal that emerges from the mitotic spindle may stimulate a furrow at anaphase or earlier, could also determine the position of the cleavage furrow perpendicular to the spindle and midway between its poles (Satterwhite & Pollard, 1992). Inhibition of cytokinesis by EGTA has identified calcium as a messenger, where a calcium wave could be a positive signal that activates MLCK (intro, section F. 19),which appears to play a pivotal role in activating the contraction of the cleavage furrow (Sellars, 1990) by phosphorylating the RLC of myosin-II which is linked with the initiation of contraction, and assembly or disassembly of the contractile ring (Sellars, 1990). Maturation-promoting factor (MPF), the enzyme consisting of cyclin and $p34^{cdc2}$ kinase, cyclically catalyses entry of eukaryotic cells into mitosis (Maller, 1991), and also coordinates the reorganisation of the cytoskeleton by phosphorylating the RLC of myosin-II (Satterwhite & Pollard, 1992; Lamb et al, 1990). Also, phosphorylation of caldesmon (intro, section F. 7) by MPF causes it to dissociate from actin filaments, contributing to the reorganisation of the cytoskeleton (Yamashiro et al, 1991).

4. NUCLEAR ACTIN

It has been demonstrated that translocation of G-actin into cell nuclei and formation of short nuclear actin rods is induced in non-muscle cells by various types of stimuli, eg. heat-shock (Ohta et al, 1989) and treatment with chemical agents such as DMSO (Sanger et al, 1980). Actin, in addition to its presence in the cytoplasm of non-muscle cells, has also been demonstrated to occur in nuclei without these stimuli (Bremer et al, 1981). Ultrastructural studies have shown actin is a structural component of the nuclear matrix, and is associated with the inner nuclear membrane (Amankwah & De Boni, 1994).

It has been proposed that the actin present in the nuclei may be cytoplasmic, migrating during experimental isolation of nuclei or physiological compromisation (Ohta et al, 1989; Sanger et al, 1980). However, biochemical evidence has demonstrated a unique form of actin native to nuclei in *Acanthamoeba* cells (Kumar et al, 1984), that indicates that nuclear actin does not represent contamination from cytoplasmic pools (Amankwah & De Boni, 1994). A distinct species of actin, with a more acidic pI, than the major non-muscle cytoplasmic isoform, has been observed to be selectively associated with isolated nuclei (Bremer et al, 1981), and has been identified as α-sarcomeric actin (Milankov & De Boni, 1993).

The function of actin in the nucleus remains unclear. It has been suggested that short actin filaments may function as a structural framework within the nucleus, providing a scaffold (Amankwah & De Boni, 1994). Nuclear actin may cooperate with small ribonucleoproteins in the processing and transport of RNA (Sahlas et al, 1993). In addition, it has been proposed that chromatin domains which undergo dynamic repositioning within the global confines of the interphase nucleus, into tissue-specific, non-random patterns, may be driven by an actin-based intranuclear motor (Milankov & De Boni, 1993). Also, actin associated with the nuclear envelope of *Amoeba proteus* may be responsible for final separation of daughter nuclei, by the formation of a nuclear constricting ring (Pomorski & Grebecka, 1993).

An increasing number of ABPs have been isolated from nuclei, suggesting that nuclear actin structures are regulated. These include myosin (Milankov & De Boni, 1993), cofilin (intro, section F. 4; Ohta et al, 1989) and ADF (intro, section F. 4; Ono et al, 1993), gCap39 (intro, section F. 2; Onoda et al, 1993), c-Abl (intro, section F. 19; Van Etten et al, 1994), NAB-34 (intro, section F. 19; Rimm & Pollard, 1989), and members of the capZ family, maintaining a population of actin in the monomeric form (intro, section F. 8; Ankenbauer et al, 1989).

Myosin present in the nucleus (Milankov & De Boni, 1993) is associated with the nuclear envelope. The presence of myosin suggests a contractile role for actin in modifying the structural state of chromatin (Milankov & De Boni, 1993).

ADF and cofilin are involved in the formation of intranuclear rods, which consist of bundles of F-actin (Ono et al, 1993). These molecules are transported to the nucleus

upon treatment of cells in culture with DMSO (Ono et al, 1993). A putative nuclear localisation signal, NLS (a short stretch of basic residues followed by a proline or glycine, allowing targeting to the nucleus), homologous to that of SV40 large T antigen (see fig 16 this study; Garcia-Bustos et al, 1991), has been detected in cofilin and ADF (Ono et al, 1993). Mutation of the putative NLS (KKRKK), shows that cofilin molecules are unable to translocate to the nuclei (Iida et al, 1992). Because an NLS sequence is not detectable in actin and both cofilin and ADF can make complexes with G-actin in the cytoplasm, actin may be transported into nuclei as an actin-cofilin and/or ADF complex (Ono et al, 1993). In SV40 large T antigen, 15 residues adjacent to the NLS sequence modulate the efficiency of nuclear transportation, possibly by phosphorylation (Garcia-Bustos et al, 1991). Therefore, it is a possibility that the nuclear transport of ADF and cofilin is regulated by phosphorylation of sites adjacent to the NLS (Ono et al, 1993). Indeed, dephosphorylation of cofilin has shown to cause its nuclear translocation upon heat shock and DMSO treatment of cells (Ohta et al, 1989.

The presence of gCap39 in the nucleus suggests filament end capping may be important in the regulation of nuclear actin under a variety of conditions (Onoda et al, 1993). In addition to interactions with actin, nuclear gCap39 may also bind DNA since it has a sequence which is related to the c-myc-basic/helix-loop-helix DNA binding motif (Prendergast & Ziff, 1991). gCap39, with an Mr of 39KDa, is at the upper limit for free diffusion into the nucleus across channels of the nuclear envelope (Onoda et al, 1993). It may therefore have to be actively transported after targeting with its own NLS or bind to another protein with a targeting sequence (Onoda et al, 1993).

5. TRANSFORMATION

In normal cells division, growth, movement and adhesion are controlled (Alberts et al, 1994). When such cells loose these controls they proliferate and become transformed. Viral infection can cause transformation either by the increased synthesis of an existing gene (protoncogene), or by the insertion of a transforming gene (oncogene) into a highly expressed region of the genome (Alberts et al, 1994). During tumor development, oncogenes are activated, and tumor suppressor genes are inactivated (Harris, 1993). P^{53} is a tumor suppressor in that it halts abnormal growth in normal cells and thus prevents tumor development (Harris, 1993). Mutation in P^{53} abolishes its

surveillance activity and allows a tumor to grow (Harris, 1993). Malignant tumor phenotypes include overproliferation, transformation, dedifferentiation, invasion and metastasis (Tsukita et al, 1993). Therefore, tumor suppressors are defined as proteins which suppress one or several malignant phenotype criteria of tumor formation (Tsukita et al, 1993).

Malignant oncogenic transformation of cells is characterised by major alterations to cell growth rates, cell shape, ruffling, contact inhibition, adhesion, and migratory behaviour (Grieg et al, 1985; Boschek et al, 1981). These changes, which involve alterations of the actin cytoskeleton (Lombardi et al, 1990; Leavitt et al, 1985), do not appear to be due to alterations in G:F actin ratios (Heacock et al, 1984) but are thought to be the result of the reorganisation of actin from stress fibres (SFs) to other, more complex, ABP-controlled, polymeric forms (Felice et al, 1990). The relationships between the changes in growth properties and the altered cell structure of transformed cells are unknown (Gluck et al, 1993). For example, it is not known whether the characteristic changes in cell morphology of transformed cells results from alterations in the rate of growth, or whether changes in cell-adhesion and cytostructure are responsible for the disruption of adhesion-dependent growth control (Gluck et al, 1993).

Cellular transformation by oncogenic tyrosine kinases is characterised by overt changes in cellular growth control and dramatic changes in cell morphology and cell surface topology (Jove & Hanafusa, 1987). CEF cells transformed by $pp60^{v\text{-}src}$, the oncogenic tyrosine kinase encoded by RSV, are morphologically round. This phenotype is accompanied by loss of actin SFs and focal contacts, and aggregation of F-actin into podosomes (Marchisio et al, 1987; Felice et al, 1990). $Pp60^{v\text{-}src}$ has been shown to be associated with detergent insoluble cytoskeletal elements (see Felice et al, 1990), and the study of the cellular substrates of $pp60^{v\text{-}src}$ has also pointed to the involvement of cytoskeletal proteins (Kellie et al, 1991) in the process of cellular transformation and cell signalling (Schaller & Parsons, 1993). Substrates include: FAK, a focal adhesion protein kinase (Schaller & Parsons, 1993); pp110, a phosphotyrosine-containing protein that localises to SFs in CEF cells (Flynn et al, 1993); tensin (intro, section F. 9); paxillin, a focal adhesion protein associated with vinculin (intro, section F. 11; Turner et al, 1990); cortactin (intro, section F. 19; Wu & Parsons, 1993), an ABP enriched in cell cortex;

MARCKS (intro, section F. 14; Joseph et al, 1992), an ABP down-regulated upon $pp60^{v\text{-}src}$ transformation; also, talin, ezrin, connexin 43, clatherin heavy chain, and integrin β-5 (see Kreis & Vale, 1993 for review). The Abl protoncogene (intro, section F. 19) is a non-receptor protein tyrosine kinase that is localised to F-actin (McWhirter & Wang, 1993). An activated form of Abl, typified by the $p120^{bcr/abl}$ complex activates tyrosine kinase actin-binding and transforming (leukaemias in mice and humans) functions of Abl (McWhirter & Wang, 1993).

5.1. Adhesion.

Transformed cells show weak substrate adhesion (Hedrick et al, 1993). They possess few, if any, focal contacts (Juliano & Varner, 1993) and this may account for the weak substrate adhesion. Upon transformation, distinctive adhesive structures called podosomes are formed, and are representative of dynamic cells with high motility (Marchisio et al, 1987).

Transformed cells express less fibronectin than their normal counterparts, and this contributes to the alterations in adhesiveness and motility observed in transformed cells (Juliano & Varner, 1993). Changes in the expression of fibronectin-binding integrins have also been observed in some types of transformed fibroblasts; indeed, transfection of $\alpha_5\beta_1$–integrin into various transformed cells results in increased adhesion and tumor suppression (Juliano & Varner, 1993).

Cell-cell association is often disorganised in tumors and is thought to be a cause of the unregulated behaviour of tumor cells including invasion and metastasis (Takeichi, 1993). Tumor metastasis involves two independent processes relevant to cell adhesion: detachment of the cells from primary tumors, and reattachment of cells to metastasis sites (Takeichi, 1993). For tumor cells to metastasise, the tight cell associations seen between epithelial cells (the origin of carcinomas), must be disrupted. It is known that cadherins are major cell-cell molecules (intro, section E. 2. 1) in tumors as well as in normal tissues (Takeichi, 1993). The perturbation of cadherin function causes disaggregation of tumor cells, and thus, may promote invasion and metastasis (Takeichi, 1993). Introduction of E-cadherin into invasive cell lines greatly reduces the invasive ability of tumor cells in vitro, suggesting that E-cadherin expression may inhibit tumor invasion (Hedrick et al, 1993).

5.2. Actin.

There is no large difference between the actin concentration in transformed and normal cells (Heacock et al, 1984). The reduction of actin SF number in transformed cells may therefore be due to the molecular composition of the actin pool (Leavitt et al, 1985). Within 3T3 cells, in addition to the nonmuscle forms of actin (β- and α-actin), the smooth muscle α-isoform is also expressed; upon transformation, this latter isoform is grossly diminished (Leavitt et al, 1985; Lombardi et al, 1990), and thus may damage the assembling of SFs. However, observations in 3T3 cells show a diffuse distribution of α-actin, lacking affinity for SFs (Lombardi et al, 1990), therefore, there may not be a direct connection between α-actin down-regulation and the derangement of SFs (Lombardi et al, 1990).

Transformation can also occur as a direct result of alteration in the primary sequence of an actin isoform: high level expression of transfected copies of a mutant β-actin gene induces tumorigenic conversion of human (KD) fibroblasts (Leavitt et al, 1987). The ability to induce a tumorigenic cell line from nonmutorigenic cells by transfection of a defined segment of DNA (namely mutant β-actin) seems to qualify as an oncogenic event (Leavitt et al, 1987).

5.3. Actin-Binding Proteins.

In the transformed state, fibroblasts and epithelial cells exhibit several morphological alterations that is accompanied by a dramatic reorganisation of elements of the cytoskeleton (see above). These underlying changes in microfilament organisation upon transformation may be caused by a coordinated down-regulation of several ABPs; with some of these proteins regulated at the transcriptional level (Kwiatkowski, 1988; Prinjha et al, 1994), but their decrease may also (in part) result from a loss of organisation, by which they become more prone to degradation (Vandekerckhove et al, 1990a). The ABPs that are down-regulated upon transformation include: α-actinin (intro, section F. 1. 3; Gluck et al, 1993), vinculin (intro, section F. 11; Rodriguez-Fernandez et al, 1992b), l-calponin (intro, section F. 20; Draeger et al, 1991), MARCKS (intro, section F. 14; Joseph et al, 1992), high Mr tropomyosins whose 3' UTR suppresses anchorage-independent-growth and tumor formation (Hendricks & Weintraub, 1981; Rastinejad et al, 1993), nonmuscle caldesmon (intro, section F. 7; Koji-Owada et al, 1984), gelsolin

(intro, section F. 2; Vandekerckhove et al, 1990a), smooth muscle myosin light chain 2 (intro, section F. 19; Kumar et al, 1989), ABP280 (intro, section F. 1. 2; Cunningham et al, 1992), LSP-1 (intro, section F. 19; Jongstra et al, 1988), α-catenin (intro, section F. 11; Tsukita et al, 1993), merlin (intro, section F. 3; Tsukita et al, 1993) and transgelin (intro, section F. 20; Shapland et al, 1988).

The down-regulation of these molecules is involved in the morphological changes seen in transformed cells (Vandekerckhove et al, 1990a). The lack of microfilament interaction with the membrane at AJs is seen with a down-regulation of α-actinin, vinculin, merlin and α-catenin, for example (Tsukita et al, 1993). Expression of these molecules in transformed cells re-establishes AJ formation and increased adhesive activity, indicating they act as tumor suppressors (Tsukita et al, 1993). Blebbing of the plasma membrane, which has been ascribed to peripheral instability such as in tumor cells, is a prominent feature of cells lacking ABP280 (Cunningham et al, 1992). The loss of caldesmon in transformed cells may correlate with the loss of Ca^{2+}-regulation in the transformed state (Koji-Owada et al, 1984). The down-regulation of transgelin, an actin cross-linking and gelling protein (Shapland et al, 1993) that is the only ABP to be completely down-regulated upon transformation (Prinjha et al, 1994), may explain the cytoskeletal activation and remodelling that occurs during metastasis (Van Roy & Mareel, 1992). It can be seen from the various functions of the down-regulated ABPs (see respective sections) that they may act synergistically to organise and control actin. Thus, upon transformation, and their down-regulation, this organisation is lost (Vandekerckhove et al, 1990a).

F. ACTIN-BINDING PROTEINS

The many actin-binding proteins (ABPs) present within eukaryotic cells (Pollard &Cooper, 1986; Hartwig & Kwiatkowski, 1991) are presented below, and categorised into their respective gene families that are based on sequence homology (Vandekerckhove & Vancompernolle, 1992; Kreis & Vale, 1993).

1. α–ACTININ SUPERFAMILY

A diverse family of proteins exists that can cross-link actin filaments, allowing them to assume many forms from tightly linked bundles to isotropic networks. To cross-

link actin filaments, a bundling or gelation protein must contain two separate actin-binding domains (ABDs) for each filament. Sequence analysis reveals that several actin-bundling, or gelation proteins share a common 250-270 residue sequence (de Arruda et al, 1990) that can be divided into two weakly homologous repeats (see fig 2 of Matsudaira, 1991), the N-terminal half being essential for binding (Bresnick et al, 1990). Two features representative of these proteins are: (i) their structural organisation (Hartwig & Kwiatkowski, 1991; Matsudaira, 1991) that determines actin-binding, oligomerisation, Ca^{2+}-regulation, or membrane association, and (ii) the presence of a spacer domain, determining the distance and orientation between ABDs and interaction with other proteins (Hartwig & Kwiatkowski, 1991). The spacer domains are made up of segments that consist of variable numbers of repeated α-helical or β-sheet motifs. Based on the arrangements of these domains and the secondary structure of the spacer segment, the actin cross-linking proteins can be grouped into three classes (see Hartwig & Kwiatkowski, 1991; Matsudaira, 1991).

1.1. Class I (Tandem actin-binding domains)

Proteins of this class are monomeric in solution, with no spacer domain, allowing the tandem arrangements of the family, **α-actinin**-like, ABD (de Arruda et al, 1990), thus forming closely packed actin bundles.

Fimbrin (Mr 70KDa; 630 residues) is a monomeric F-actin bundling protein in microvilli and filopodia (de Arruda et al, 1990). It is possibly the chick homologue of the human **plastins** (t-, l- and I-plastin) (Zu et al, 1990; Lin et al, 1994) with which it shares ≈70% identity (86% identical to I-plastin). L-plastin (627 residues) is found in only normal or transformed leukocytes (de Arruda et al, 1990), t-plastin (630 residues) found in cells of solid tissue (Zu et al, 1990), and I-plastin (629 residues) found only in intestinal brush borders (Lin et al, 1994). The plastins are encoded by three distinct genes which display tissue specific expression (Lin et al, 1994). Comparison of deduced amino acid composition of l-plastin and **accumentin** (Southwick & Stossel, 1981) suggest they are the same protein. A fimbrin homologue, encoded by the yeast ***SAC6*** gene, has been shown to bind actin both in vitro and in vivo; mutant alleles of *SAC6* result in a loss of actin cables and asymmetry in the distribution of actin (Adams et al, 1993).

The N-terminal region of both fimbrin and the plastins has two possible EF hand regions followed by a calmodulin-binding consensus sequence, with the two familial ABDs in the middle of the C-terminus (de Arruda et al, 1990; Zu et al, 1990). The homology between fimbrin and the plastins is weakest in the N-terminal domain, raising the possibility that the isoforms may be differentially regulated by calcium. Pacaud & Derancourt (1993) have observed that binding of fimbrin to F-actin is unaffected by Ca^{2+} (up to 5mM), whereas the cross-linking activity of l-plastin is inhibited by Ca^{2+} concentrations in the physiological range. The observation that l-plastin is phosphorylated after leukocyte stimulation by Il1 and Il2 (de Arruda et al, 1990; Pacaud & Derancourt, 1993) has led to the examination of potential regulation of fimbrin by phosphorylation by Messier et al (1993). They found that a fraction of fimbrin that is present in the cytoskeleton of macrophages is preferentially phosphorylated on its N-terminus relative to soluble fimbrin.

1.2. Class II (β-sheet proteins)

Proteins of this class contain the N-terminal familial consensus ABD separated by a spacer domain containing tandem repeats (of 96 residues) of a β-sheet motif (Gorlin et al, 1990; Noegal et al, 1989). Different numbers of these repeats account for the differences in sizes of the class members. Six repeats are found in *Dictyostelium* **ABP120** (Noegal et al, 1989) whilst human **ABP280** has 24 repeats (Gorlin et al, 1990).

ABP120 is a 857 residues rod-shaped homodimeric protein thought to be assembled by overlapping antiparallel subunit chains with high β-sheet potential (Noegal et al, 1989). ABP120 is present in the cell cortex and pseudopods suggesting that it may function in the formation of filament networks during psuedpod extension, eg. in chemotaxis of *Dictyostelium* (see Luna & Hitt, 1992 for review). Tryptic digestion followed by actin cosedimentation demonstrate that residues 89-115 are essential for actin-binding activity (Bresnick et al, 1990), and cDNA analysis (Noegal et al, 1989) has revealed that this region is ≈60% homologous within the α-actinin superfamily of proteins, and identifies their common ABD which is presumably able to bind to a common site on an actin filament.

ABP280 (nonmuscle **filamin**) is a ubiquitous dimeric cross-linking phosphoprotein (Chen & Stracher, 1989) that promotes orthogonal branching of actin

filaments, linking the actin filaments to the membrane glycoprotein Ib/Ix complex of platelets (Gorlin et al, 1990). ABP280 self-associates at the C-terminal ends of identical subunits, each having 2647 residues, corresponding to a Mr of 280.5KDa (Gorlin et al, 1990). 90% of the ABP280 primary sequence is made up of β-sheet runs, with the common ABD at the N-terminus. The V-shaped dimerisation of the two ABP280 subunits allow the ABDs to be well separated (Matsudaira, 1991). Sequence differences within the multiple repeats of endothelial ABP280 and smooth muscle filamin suggests a minimum of two isoforms are expressed in humans, however, Brown et al (1993) revealed bovine erythrocyte ABP280 to be identical to smooth muscle ABP280. ABP280 deficient cells such as human malignant melanomas, have impaired locomotion and peripheral cytoskeletal instability (blebbing), establishing ABP280 to function in stabilising cortical actin in vivo (Cunningham et al, 1992).

1.3. Class III (α-helical proteins)

The conserved features of this class include (i) an N-terminal ABD, (ii) a variable repeated α-helical spacer segment, (iii) a C-terminal calmodulin-like Ca^{2+}-binding domain, and (iv) oligomerisation into an antiparallel homodimer that allows a single ABD to be present in two copies in a single molecule at opposite ends (Matsudaira, 1991).

α-actinin is a 94-103KDa, rod-shaped F-actin binding and cross-linking protein forming bundles or isotropic networks in vitro depending on α-actinin concentration and affinity for actin of the α-actinin isoforms (see below) (Wachsstock et al, 1993). α-actinin was originally isolated from skeletal muscle (as a major Z disc component) and distinct isoforms can be found in smooth and nonmuscle cells (Blanchard et al, 1989). In smooth muscle, α-actinin is localised in membrane-associated dense plaques (see Otto, 1994 for review) and in the AJs of nonmuscle cells (Blanchard et al, 1989) where it is distributed along the actin filaments and reported to be bound to the cytoplasmic domain of the β-1 family of integrins (Otey et al, 1990). *Drosophila* α-actinin has 68% identity to chick isoforms and is associated with microfilaments in the vicinity of focal adhesion plaques (Fyrberg et al, 1990). An α-actinin-like protein, **ABP-46**, has been identified from *Physarum* (St Pierre et al, 1993), it is a 402 residue polypeptide, is 40% identical to the N-terminal, ABD region, of α-actinin.

It has been generally accepted that whereas muscle α-actinins (smooth and skeletal isoforms are 80% identical (Arimura et al, 1988)) are not regulated by Ca^{2+}, nonmuscle forms are regulated by virtue of their EF hand motif; such that in the presence of Ca^{2+} the level of actin cross-linking is substantially decreased (see Hartwig & Kwiatkowski, 1991 for review). However, Pacaud and Harricane (1993) concluded from purified macrophage α-actinin that it is not all regulated by Ca^{2+}. Sequencing of chick brain α-actinin (892 residues; 107KDa) indicates that it differs from the smooth muscle form in the region of the first EF hand, where 27 residues in brain α-actinin are reduced to 22 in the smooth muscle isoform (Waites et al, 1992). This probably accounts for the different Ca^{2+}-sensitivities of the muscle and nonmuscle isoforms. This region of sequence divergence is encoded by two separate exons, and therefore proportions of the two transcripts can be determined in various tissues by use of RT-PCR and isoform specific primers. Waites et al (1992) found that various tissues contain mRNA encoding both isoforms (ie. 11% of the smooth isoform is found in brain tissue), and therefore explains Pacaud and Harricane's (1993) observation. This technique showed cells lacking stress fibres only express the nonmuscle isoform, where a nonmuscle Ca^{2+}-sensitive isoform may serve to cross-link F-actin in a manner which can be regulated by local changes in $[Ca^{2+}]$. The Ca^{2+}-insensitive smooth muscle isoform may be required to maintain stress fibre integrity in the face of changes in intracellular $[Ca^{2+}]$ (Waites et al, 1992). Imamura and Masaki (1992) also found two α-actinin isoforms in chick lung (which has roughly equal amounts of the two isoforms according to Waites et al (1992)), and that one isoform (115KDa) is Ca^{2+}-insensitive, it probably being the smooth muscle isoform.

Dystrophin is a very large (3685 residues; 427KDa) actin cross-linking protein that can be separated into four domains (Koenig et al, 1988). The N-terminal domain I contains the family ABD (Hemmings et al, 1992). The large domain II contains a spacer domain of 25 repeats of a 100 residue sequence (Koenig et al, 1988). Domain III contains a degenerate EF hand sequence that is thought to be non-functional with a domain IV being unique to dystrophin (Koenig et al, 1988).

Dystrophin is a diverse family of proteins expressed in all tissues (reviewed by Ervasti & Campbell, 1993). One example, **urotrophin**, is similar in size (3433 residues) to dystrophin, and contains the unique domain IV, but it is expressed in a broader range of tissues, probably playing a similar role to dystrophin (Tinsley et al, 1992).

It has been demonstrated that domain IV of skeletal muscle dystrophin is tightly associated with an oligomeric complex of six sarcolemmal proteins (Ervasti & Campbell, 1993) of which one, a peripheral laminin-binding glycoprotein, provides a linkage between the complex and the ECM (see Luna & Hitt, 1992 for review). Patients with Duchenne muscular dystrophy (DMD) express a truncated dystrophin (caused by a mutation in the 14Kb gene) lacking the C-terminal domain IV (see Brown & Lucy, 1993 for review). The subsequent lack in linkage between F-actin/dystrophin and the sarcolemmal proteins of DMD patients could make muscle fibres susceptible to necrosis during contraction (Ervasti & Campbell, 1993).

Spectrin denotes a family of largely α-helical, high molecular weight actin-binding proteins found in association with the plasma membrane of mature cells (reviewed by Pumplin & Bloch, 1993). These proteins exist as dimers (α, β) and as heterodimers $(\alpha,\beta)_2$, with the N-terminal of the α-chain linking to the C-terminal of the β-chain (Bretscher, 1991). Spectrins bestow mechanical stability on the membrane and link transmembrane proteins to the cytoskeleton (Pumplin & Bloch, 1993).

There are three classes of spectrin: erythroid, intestinal brush border and nonerythroid spectrin. The erythroid spectrin consists of an α-subunit (2429 residues; 280KDa) (Sahr et al, 1990) and a β-subunit (2137 residues; 246KDa) (Winkelmann et al, 1990). The β–subunit appears to be the most conserved among all spectrins. This subunit contains the family ABD, and binding sites for band 4.1 (intro, section F. 3) and ankyrin (intro, section F. 19) which provide the chief linkages between actin/spectrin and the membrane (reviewed by Hitt & Luna, 1994). α-subunits have relatively small regions of homology among spectrins, except for an oligomerisation domain and an SH3 domain (Hitt & Luna, 1994).

The second class of spectrin is exemplified by a terminal web spectrin of avian intestinal brush borders called **TW260/240** (Coleman et al, 1989). It cross-links

microvillar core actin, but does not intimately associate with the plasma membrane due to a unique β-subunit (Coleman et al, 1989).

Fodrin represents the third class of spectrins (Moon & McMahon, 1990), with human α-fodrin (2472 residues; 284KDa) 54% identical to human erythrocyte α-spectrin. Nonerythrocyte α-spectrin binds calmodulin in a Ca^{2+}-dependent manner, an in vitro consequence of which is the down-regulation of the interaction between spectrin and band 4.1/F-actin (see Bretscher, 1991).

Adducin, a calcium/calmodulin sensitive, membrane-skeletal protein, is a candidate to promote assembly of spectrin-actin networks in erythrocytes and at sites of cell-cell contact in epithelial tissues (Joshi et al, 1991; Gardner & Bennett, 1987). Adducin is comprised of two subunits: alpha (103KDa) and beta (97KDa) and is known to be a substrate for protein kinase C (Waseem & Palfrey, 1988). The complete sequence of both subunits (Joshi et al, 1991), reveal they are 49% identical and 66% similar, suggesting evolution by gene duplication. Similarity has been noted between adducin and the highly conserved F-actin-binding ABD of the α-actinin superfamily. α-adducin head domain is 33% identical and 55% similar to it, while the β-adducin is 15% identical and 52% similar (Joshi et al, 1991). This similarity may reflect a distant relationship, however, the functional significance is not clear because head domains of adducin do not associate with F-actin (Joshi et al, 1991).

2. GELSOLIN SUPERFAMILY

The **gelsolin** superfamily is a class of actin-modulating proteins found in various eukaryotes. These proteins block the ends of F-actin and cause the fragmentation and depolymerisation of actin filaments in a Ca^{2+}-dependent manner (reviewed by Weeds & Maciver, 1993). Multiple alignment of these family of sequences indicate that similarities are greater in their N-terminal half, which contains ≈ 24% identity across all aligned sequences (Way & Weeds, 1988). The gelsolin family of proteins contain multiple structurally similar, highly conserved, actin-binding domains (ABDs) of ≈ 15KDa each (Kwiatkowski et al, 1986). Family members such as gelsolin and **villin** contain six of these domains, whereas some members such as **fragmin** contains only three of them (Vandekerckhove & Vancompernolle, 1992). Sequence analysis shows that gelsolin and

villin probably arose by duplication of an ancestral protein resembling fragmin. Fragmin may have arisen from an ancestral actin-monomer-binding protein that was tandemly repeated twice (3 repeats in fragmin; 6 repeats in gelsolin and villin) (see Hartwig & Kwiatkowski, 1991 for review).

Gelsolin is an 83KDa protein widely distributed in mammalian, bird and amphibian cells, as well as the ECM of mammals where it is known as **brevin** (Kwiatkowski et al, 1986). The cytoplasmic and secreted forms of gelsolin, which share 42% identity with human, has been identified in *Drosophila* (Stella et al, 1994). Gelsolin interacts with actin in several ways: (i) binds to actin monomers, stimulating the formation of nuclei, (ii) acts as a barbed end capping protein, (iii) severs actin filaments in a Ca^{2+}-dependent manner (Isenberg, 1991). It is well established that there are three distinct actin-binding sites unevenly distributed with the 6-fold (S1-S6) segmental repeats of gelsolin (Way et al, 1990; review in Weeds & Maciver, 1993). Briefly, the first segment, S1, contains a monomer-binding site. Segments S2-3 contain a calcium-independent F-actin binding site which competes for the same site on actin filaments with α-actinin (Way et al, 1992). S4-6 contains a calcium-dependent monomer site that is responsible for the calcium regulation of gelsolin, and is required for nucleation. Both S1 and S2-3 are required for severing and capping: without S2-3, S1 is not targeted to filaments; without S1, S2-3 decorates filaments but does not sever. Janmey et al, (1992) have identified PIP_2 binding sites at the N-terminus of S2 that inhibits gelsolin's severing activity.

The role of gelsolin in vivo is yet to be determined fully, but gelsolin-actin interactions are important in motile processes in mammalian cells. Indeed, transfection of gelsolin cDNA into fibroblasts results disruption of stress fibres (Finidori et al, 1992), and modestly increased levels of cytoplasmic gelsolin can enhance the stimulated locomotion of mouse fibroblasts, the increase in function being proportional to the amount of increased gelsolin expressed (Cunningham et al, 1991). Chemotactic stimulation of neutrophils results in the dissociation of pre-existing gelsolin-actin complexes concurrent with a rapid increase in F-actin content. These observations are consistent with a model in which gelsolin caps the barbed ends of filaments or nuclei in resting cells, being removed during stimulation through changes in PPI levels or

distribution (Isenberg, 1991), the uncapped nuclei/filaments providing sites for rapid elongation of filaments (Hartwig & Kwiatkowski, 1991). Gelsolin is down-regulated in transformed human fibroblasts and keratinocytes, with less gelsolin correlating with a more discorded cytoskeletal architecture and therefore a reduced 'need' for gelsolin's functional activity in regulating actin filaments (Vandekerckhove et al, 1990a). Severing and capping activity of gelsolin in vivo may be limited by interactions between tropomyosin (TM) and caldesmon (Ishikawa et al, 1989).

Full length cDNA for human (Kwiatkowski et al, 1986) and pig gelsolin (Way & Weeds, 1988) encode for a polypeptide of 802 amino acids. Gelsolin is unusual in that one form is produced for intrinsic intracellular use and another is made specifically for secretion. The secreted form differs by the presence of a 25 amino acid residue extension, but is encoded by the same single gene as cellular gelsolin, thus indicating different transcriptional initiation sites (Kwiatkowski et al, 1986).

Villin acts as a PPI inhibitable actin-bundling protein in the microvillus core filament of brush borders (Louvard, 1989). Villin has two unique features within the gelsolin superfamily: (i) it has a dual function in vitro regarding Ca^{2+}: at low [Ca^{2+}], villin acts as a bundling factor, while at high [Ca^{2+}] villin severs, nucleates and blocks the end of actin filaments; (ii) unlike many other ABPs villin presents a striking tissue-specific distribution (see Weeds & Maciver, 1993). Villin cDNAs from human (Arpin et al, 1988) and chick (Bazari et al, 1988) encode a polypeptide of 826 amino acids (92KDa) that share >80% identity. Villin sequence shares 50% identity with gelsolin, and also has the same 6-fold internal homology. A 75 amino acid C-terminal headpiece, present in villin, shows no homology with the 6-fold domain structure (Bazari et al, 1988). This headpiece contains an additional F-actin binding site; thus, by binding actin filaments to its headpiece and to a second site within the S2 domain of its core (de Arruda et al, 1992), bundles can be formed. The S2 domain has also been mapped for severing, with capping function being localised to the S1 domain (de Arruda et al, 1992). The roles of the similar ABDs of villin and gelsolin have been explored in vivo by transfection of fibroblasts with mutant forms of these proteins (Finidori et al, 1992; Friederich et al, 1992). The domains of villin required to form microvilli are present in S1-3 and the headpiece of this molecule. The presence of the basic KKEK cluster of

residues at the extreme C-terminus of its headpiece is essential for microvillus formation (Friederich et al, 1992) in which villin may be essential for maintaining the absorptive surface area of the epithelium.

A cDNA has been isolated from *Dictyostelium* encoding a protein termed **protovillin** (formally **cap100KDa**) (Hoffman et al, 1993). Protovillin is a PIP_2 inhibitable F-actin capping protein that displays neither severing, cross-linking nor nucleating activities (Hoffman et al, 1993). The cDNA encodes a protein of 959 residues (109 KDa) and shares 32% identity with villin, representing a *Dictyostelium* homologue. Protovillin has the villin/gelsolin-like core domain structure (plus headpiece); this indicates that duplication of an ancestral protein resembling fragmin gave rise to gelsolin or villin very early in the evolution of this family. The sequence differences between protovillin and villin, ie. a different N-terminal S2 domain sequence, and the lack of a KKEK motif in the headpiece of protovillin, is likely to reflect the differences in function between the two proteins (Hoffman et al, 1993).

A cDNA from human reticulocyte cDNA library has been isolated and encodes for the actin-bundling protein **dematin** (**band 4.9**) (Rana et al, 1993). The sequence includes a homologue to the headpiece of villin. As this region contains an actin-binding domain (Friederich et al, 1992), dematin is included in this family. Dematin has been identified in the human erythroid membrane skeleton with its actin-binding activity abolished upon phosphorylation and restored again after dephosphorylation (Rana et al, 1993). Dematin exists as a trimer and it is not known whether it contains another actin-binding site in its core, therefore it probably cross-links actin by forming higher order structures. With homology to villin in its headpiece, it is possible dematin may substitute for villin in villin-negative tissues (Rana et al, 1993).

gCap39, a Ca^{2+}-sensitive, PPI reversible barbed end capping protein of 42KDa shows weak filament-nucleating activity (Yu et al, 1990). An independent study by Dabiri et al (1992) led to the identification of a cDNA clone encoding a protein of length 351 residues and termed **macrophage capping protein** (**MCP**) that is identical to gCap39. gCap39/MCP shows high identity to the N-terminal half of gelsolin (48%), but is different from other gelsolin family members in that it does not sever actin filaments (Bearer, 1991). The functional differences between gCap39/MCP and gelsolin may mean

they play complementary roles in a physiologically significant manner. Gelsolin and villin are uncapped by PPIs only when cytosolic Ca^{2+} concentration falls to resting levels. gCap39/MCP uncapping is less restrictive, occurring when $[Ca^{2+}]$ or PIP_2 concentration rises. gCap39 may therefore promote actin polymerisation during the initial phase of agonist stimulation when intracellular Ca^{2+} is elevated (Yu et al, 1990). gCap39 is phosphorylated in vivo, and is present in the nucleus in a more highly phosphorylated form (Onoda et al, 1993). Nuclear gCap39 redistributes throughout the cytoplasm during mitosis, but its function in the nucleus is unknown (Onoda et al, 1993).

Adseverin is a 74KDa calcium-dependent and acidic phospholipid-modulated actin-binding protein isolated from bovine adrenal medulla (Maekawa et al, 1989; Sakurai et al, 1991). The same protein, isolated and called **scinderin** by Rodriguerz del Castillo et al (1990), is mainly localised in the subplasmalemma region of chromaffin cells. Adseverin/scinderin promotes nucleation, severs actin filaments, and caps the barbed ends of actin (Sakurai et al, 1991). These properties are retained in the N-terminal half of adseverin/scinderin that has a functional domain structure similar to that of the gelsolin and villin core.

Fragmin, from *Physarum* , is a 42KDa protein that severs F-actin and binds G-actin in the presence of micromolar concentrations of calcium. The fragmin/G-actin complex becomes the nucleus for actin polymerisation. The barbed ends of the short filaments that are formed are capped with fragmin (Furuhashi & Hatano, 1989). Fragmin peptide sequence shows it is divided into three weakly homologous regions that is 100% identical to the capping protein **cap42** and 36% identical to the N-terminal half of human gelsolin (Ampe & Vandekerckhove, 1987). Fragmin might be a candidate for actin nucleating sites along the plasma membrane where F-actin elongates (Furuhashi & Hatano, 1989).

Severin is a Ca^{2+} activated, PIP_2-inhibited, actin-binding protein that nucleates actin assembly, severs and caps the barbed ends of actin filaments (Weeds & Maciver, 1993). Severin cDNA isolated from *Dictyostelium* (Andre et al, 1988) encodes a protein of 361 amino acids (39.9KDa) and consists of the familiar, highly conserved, 3-fold domain structure (as seen in fragmin) that is duplicated in gelsolin and villin. Indeed,

there is 38% identity between severin and gelsolin N-terminal sequence (residues 1-391), and 30% identity between severin and gelsolin C-terminal sequence (residues 392-782), indicating duplication of ancestral severin. Eichinger and Schleicher (1992) have concluded that severin contains one actin-binding site in each of its three domains; namely, a capping site in S1 and two F-actin side-binding regions in S2 and S3. Two distinct Ca^{2+}-binding activities are associated with S1 and S2, with PIP_2-binding regions in S1 and in S2-3.

3. BAND 4.1-RELATED SUPERFAMILY

Members of 4.1-related family (see below) have a common structural organisation (see Arpin et al, 1994; Algrain et al, 1993 for reviews), and have also been defined between similarities in their primary structure. Their putative role as actin-plasma membrane linkers is based on the following observations: (i) they have been found in membrane extensions (Arpin et al, 1994); (ii) they share sequence homology with **talin** (a component of focal adhesions) and **band 4.1** (a spectrin-actin-membrane linker), whose function in membrane-cytoskeletal interactions is known (see Isenberg, 1991 for review); (iii) three members of the superfamily, **ezrin**, **radixin** and **moesin** (the ERM subfamily) have been identified as molecular linkers between cell surface CD44 and the actin-based cytoskeleton (Tsukita et al, 1994). The sequence similarities between all the family members reside within their N-terminal domains, with the ERM subfamily displaying 75% identity in their overall amino acid sequences (Sato et al, 1992).

Ezrin was first identified as one component of intestinal microvilli (Gould et al, 1989). Gene sequence was first isolated, and named **cytovillin**, by Turunen et al (1989) and shortly afterwards by Gould et al, (1989) who identified ezrin from a Hela cell library to be a polypeptide of molecular weight 69KDa. The C-terminal half of ezrin has been implicated in actin binding (Algrain et al , 1993). *Drosophila* ezrin cDNA transfected into fission yeast results in the recruitment of F-actin to the subplasalemmal cortex, indicating a functional membrane-cytoskeleton connection (Edwards et al, 1994). This transfected cDNA lacks a full N-terminus, and it is therefore surprising that it can apparently target actin to the plasma membrane. Ezrin is a substrate of tyrosine kinases

(Kreig and Hunter, 1992), being phosphorylated on tyrosine residues after induction with EGF or PDGF (Arpin et al, 1994).

Radixin, a Ca^{2+}- independent barbed-end capping protein is highly concentrated in the undercoat of AJ and cleavage furrows (Tsukita et al, 1989). As actin filaments associate with plasma membranes at barbed-ends (Mabuchi et al, 1988), it strongly suggests that radixin may be a key protein in end-on association of actin filaments with plasma membranes (Tsukita et al, 1989). Radixin cDNA derived from a mouse teratocarcinoma (F9 cells) library encodes for a 68.5KDa, 583 amino acid polypeptide that shares no homology with other barbed-end capping proteins such as gelsolin (intro, section F. 2) and CapZ (intro, section F. 8) (Funayama et al, 1991). Similarity between radixin and band 4.1, within their N-terminal region, follows the idea that radixin plays a crucial role in the association of the barbed-ends of actin filaments with the plasma membrane in the AJ and the cleavage furrow (Arpin et al, 1994). The binding site for a plasma membrane protein (glycophorin) has been mapped to the N-terminal domain of 4.1 (Bennett, 1989). As band 4.1 does not interact directly with actin and radixin specifically binds to barbed-ends of actin, it seems likely that the C-terminal domain of radixin may have a bearing on the interaction with cytoskeletal proteins (Arpin et al, 1994).

Moesin (membrane-organising extension spike protein) has been localised diffusely on or near the cell membrane with preference for retraction fibres, blebs, microspikes, and lamellipodia (Lankes & Furthmayr, 1991). Moesin cDNA, sequenced from a mouse HL60 cDNA library, encodes a 577 amino acid, 78KDa, polypeptide (Lankes & Furthmayr, 1991). Moesin contains no apparent signal peptide or transmembrane domain, and sequence indicates it to be a member of the ERM subfamily (Arpin et al, 1994). Moesin contains similarity with 4.1 and talin at the N-terminus and a similar membrane/cytoskeletal role is implied by structure and domain predictions derived from the cDNA-deduced peptide sequence (Lankes & Furthmayr, 1991).

The above ERM subfamily members have a close relationship to the tumor suppressor protein of type 2 neurofibromatosis (NF2), a human autosomal dominant disorder characterised by the occurrence of tumors of the CNS (eg: schwannomas) (see Trofatter et al, 1993; Rouleau et al, 1993 and references therein). Using a general strategy

of positional cloning on the basis of association between a disease and a known genetic marker, two groups have identified a gene that is a candidate for the tumor suppressor responsible for NF2. Trofatter et al (1993) have named the gene product **Merlin**, whilst Rouleau et al (1993) have called it **schwannomin**. Both groups have isolated cDNAs that encode identical, 595 amino acid, ≈66KDA, polypeptides. Merlin/schwannomin has 63% identity with the ERM subfamily in its first 342 residues (with ≈ 46% overall identity). With this homology in mind, it is tempting to speculate that merlin/schwannomin and the ERM subfamily members act as tumor suppressors by affecting actin filament/plasma membrane association, although, interaction of merlin/schwannomin with actin remains to be demonstrated (see Tsukita et al, 1993 for review).

Gu et al (1991) obtained a 3.7Kb cDNA, designated **PTPase MEG**, from a human megakaryoblastic cell line (MEG-01) cDNA library. cDNA encodes for a 926 amino acid polypeptide of calculated Mr 105KDa. The C-terminal region (residues 659-909) has strong (35-40%) identity to protein-tyrosine-phosphatase (PTP) domains, whilst the N-terminal region (residues 31-367) is 45% identical to the N-terminus of human erythrocyte band 4.1. Identification of a PTP that is related to cytoskeletal proteins implies that cell signalling activities reside not only in transmembrane receptors but in cytoskeletal elements aswell (Gu et al, 1991).

PCR primers corresponding to conserved sequences within the catalytic domains of PTP were used to amplify PTP-related cDNAs from a Hela cell library (Yang & Tonks, 1991). An obtained clone encodes for a protein of 913 amino acids with a calculated Mr of 104KDa and termed **PTPH1**. Sequence can be described in terms of 3 segments: (i) N-terminal region displays homology to the N-terminal domains of the 4.1 superfamily, (ii) central region bearing putative phosphorylation sites for ser/thr kinases, and (iii) C-terminal region shows homology to members of the PTP family. PTPH1 may act at junctions between the membrane and cytoskeleton (Yang & Tonks, 1991).

55% identity with the N-terminal region of the band 4.1 superfamily has been found in ***Echinococcus multilocularis*** **10 (EM10)**. EM10 is characterised as a surface antigen of the tapeworm that causes echinococcus (Frosch et al, 1991).

Band 4.1 has been identified as one of the major accessory proteins of human erythrocyte membranes and associated with spectrin (intro, section F. 1. 3) and actin in

membrane cytoskeletons (Bennett, 1989). Proteolytic digestion studies of band 4.1 revealed a 67 amino acid segment which is responsible for promoting spectrin-actin interactions, and also contains sites for cAMP induced phosphorylation (Bennett, 1989). The state of phosphorylation of 4.1 regulates its capacity to stimulate spectrin-actin assembly and the attachment of 4.1 to the erythrocyte membrane. Isolated 4.1 binds tubulin and myosin, suggesting 4.1 may play a role in integrating the membrane skeleton with the larger cytoskeleton (see Bretscher, 1991 for review). The band 4.1 family have an homologous N-terminal domain and may give a clue to understanding their physical functions. In band 4.1 this domain contains a binding site for glycophorins and band 3 protein (see Isenberg, 1991 for review). Full length cDNA for erythroid 4.1 has been isolated from human reticulocytes (Conboy et al, 1986), encoding a protein of 80KDa. Of particular interest is the finding that forms present in reticulocytes and lymphoid cells are spliced to eliminate the spectrin-actin binding domain of erythroid band 4.1 (Tang et al, 1990). This implies that 4.1-related molecules are involved in membrane-cytoskeletal associations other than with spectrin. Further studies have shown the generation of multiple isoforms by alternative mRNA splicing, with some of the species performing novel structural functions because they appear to be localised in intracellular junctions and in the nucleus (Tang et al, 1990).

Talin is a high molecular weight protein (270KDa) concentrated at regions where bundles of actin filaments attach to and transmit tension across the plasma membrane to the ECM (Rees et al, 1990). Talin binds in vitro to the cytoplasmic domains of the integrin family of ECM receptors, vinculin (intro, section F. 11) and actin (see Beckerle & Yeh, 1990 for review). Talin has been cloned and sequenced by Rees et al (1990) from a 3T3 cell cDNA library, and reveals a 2451 amino acid protein. Restriction enzyme analysis and partial sequence data on a large number of overlapping cDNA clones revealed no evidence for variant forms (eg: alternative splicing) (Rees et al, 1990). The C-terminal tail of talin shows weak relationships with various large structural proteins with extended coiled α–helices (eg. tropomyosin). A segment of the N-terminal domain of talin shows identity with band 4.1 (20%) and ezrin (23% over 220 amino acids of the N-terminal domain) (Rees et al, 1990).

4. COFILIN-RELATED

A group of small (16-20KDa) actin-binding proteins exist which share extensive sequence homology. They are G-actin binding and F-actin severing proteins (see Hartwig & Kwiatkowski, 1991 for review) that may regulate actin polymerisation in the cell by severing filaments and/or sequestering actin monomers. Physiological functions of these proteins are far from clear (Hartwig & Kwiatkowski, 1991).

Cofilin is a 20KDa, widely expressed vertebrate protein that binds to G and F-actin stoichiometrically. At pH > 7.5 cofilin depolymerises actin filaments, whilst at lower pH binds with high affinity to actin subunits in F-actin (Hartwig & Kwiatkowski, 1991). Cofilin also inhibits the ability of tropomyosin or myosin to interact with F-actin, and its interactions are specifically inhibited by various phosphoinositides (see Hartwig & Kwiatkowski, 1991). Cofilin cDNA from a mouse brain cDNA library encodes a protein of 166 amino acids (Moriyama et al, 1990a) which shares 99% identity with porcine brain cofilin (Matsuzaki et al, 1988). Abe and coworkers (1990) have identified chick cofilin that is over 80% identical to mammalian cofilin. Moon et al (1993) identified a *Saccaromyces* homologue of cofilin, it is 41% identical to mammalian cofilin. This yeast cofilin (**COF1**) is associated with the membrane cytoskeleton, and is essential for cell division. The sequence of cofilin reveals a motif that is homologous to a nuclear localisation sequence (NLS) of SV40 large T antigen (Kalderon et al, 1984; results, fig 16), and it has been shown to form intranuclear actin/cofilin rods in cultured fibroblasts exposed to heat shock or 10% DMSO (Ono et al, 1993). Cofilin dephosphorylation is closely related to its nuclear accumulation and it is likely that cofilin is dephosphorylated in the cytoplasm prior to its nuclear accumulation (Ohta et al, 1989). It has been proposed (Ohta et al, 1989) that under normal conditions cofilin exists uniformly in the cytoplasm and nuclei, where upon heat shock or 10% DMSO treatment, the signal sequence in cofilin might become exposed and work as an NLS. Cofilin's dephosphorylation and subsequent nuclear translocation occur spontaneously in autonomously proliferating T-lymphoma cells (Samstag et al, 1994). COF1 is not detected in the nucleus after treatment with DMSO or heat shock (Moon et al, 1993).

It has been shown that residues 104-115 of cofilin form a potential actin-binding site (Yonezawa et al, 1991). This region is very similar to sequences found in the other

proteins of the cofilin-related gene family, and therefore may be a consensus sequence essential for actin-binding and depolymerising activity (Yonezawa et al, 1991). A hexapeptide, DAIKKK, present in all the family, except depactin (see below), is identical to an N-terminal sequence of tropomyosin. The DAIKKK motif is located close to the cross-linked residues of cofilin-actin conjugate, it may therefore constitute a minimal actin-binding motif (see Vandekerckhove & Vancompernolle, 1992 for review).

Actin depolymerising factor (ADF) is a widely distributed, 165 residue, 18.5KDa protein that has Ca^{2+}-insensitive, pH-independent F-actin-binding and severing/depolymerisation activities that are inhibited by phosphinositides (reviewed by Hartwig & Kwiatkowski, 1991). A phosphorylated form of ADF is seen in chick myocytes that is unable to depolymerise actin filaments or bind monomers; representing a regulated form of ADF (Morgan et al, 1993). ADF is a constituent of a variety of tissues, cells and cultured cell lines, where it is localised in high concentrations along the leading edge of ruffled membranes in fibroblasts and in the growth cones of neurones (Morgan et al, 1993). Chicken ADF (Abe et al, 1990; Adams et al, 1990) has over 95% sequence identity with porcine **destrin**, the mammalian homologue of ADF (Moriyama et al, 1990b), with which it shares similar biochemical activities (Hawkins et al, 1993). Comparison between human ADF/destrin and chick ADF/destrin shows 8 amino acid substitutions, of which most are conservative (Moriyama et al, 1990b). ADF/destrin sequence shares 70% identity with either chicken or mammalian cofilin. Although very similar ADF/destrin and cofilin are shown by Southern analysis to be derived from different genes (Abe et al, 1990). The major difference between ADF/destrin and cofilin relates to their F-actin binding properties. Between pH 6.3 and 7.5 cofilin cosediments with F-actin (representing an increase in polymerisation and filament binding), but at pH values between 7.5 and 8.3 it severs actin filaments but has no stable binding (representing depolymerisation and monomer binding). ADF/destrin, on the other hand, shows no filament binding at lower pH values (Moriyama et al, 1990b). ADF/destrin contains the same NLS sequence as seen in cofilin, under conditions of heat shock or DMSO treatment ADF/destrin is also translocated to the nuclei of cultured myogenic cells (Ono et al, 1993).

Actophorin, from *Acanthamoeba* (Maciver et al, 1991), has two established biochemical activities: (i) by forming a 1:1 complex with monomeric actin, actophorin can regulate the pool of actin available for polymerisation; (ii) actophorin severs F-actin in a dose-dependent manner (Maciver et al, 1991). Immunofluorescence studies reveal actophorin to colocalise with actin filaments in the the leading edge of cells (Quirk et al, 1993). Actophorin cDNA sequence encodes a 138 amino acid polypeptide of calculated Mr 15.5KDa the peptide sequence of which is closer to that of cofilin (41% identity) than to ADF/destrin (37% identity) with which it shares functional properties (Quirk et al, 1993). Actophorin sequence contains high similarity with the putative actin-binding domain and the DAIKKK motif of cofilin, but lacks the NLS sequence and consequently no actophorin is detected in the nuclei of investigated cells (Quirk et al, 1993).

Depactin is an echinoderm F-actin depolymerising/monomer binding protein of 150 amino acids (Mr 17.6KDa) (Takagi et al, 1988) that acts in much the same way as ADF/destrin and actophorin (Hartwig & Kwiatkowski, 1991). Sequence of depactin reveals high similarity to the dodecapeptide in cofilin that has been shown to bind actin and deemed essential for actin cross-linking and severing (Yonezawa et al, 1991). This is the only major region of high sequence homology between depactin and the other members of the cofilin-related gene family. As this region contains a possible actin-binding site, depactin warrants inclusion within the gene family.

5. MYOSINS

Myosins are a superfamily of structurally diverse, actin-based motor proteins characterised by the presence of a conserved ≈80KDa motor domain attached to a variety of structurally distinct tail domains (Cheney et al, 1993). Myosins have usually been divided into two classes, based on whether they form monomers or dimers. **Myosins-II** are the two-headed and filament-forming myosins, and the **myosins-I** are the smaller monomeric myosins (reviewed in Pollard et al, 1991). Since it is the presence of a myosin-like head domain that defines the members of the myosin superfamily, the myosins have been classified based upon a comparison of their head-domain amino acid sequences (Cheney et al, 1993), and has led to the conclusion that myosins fall into at least nine distinct classes (see Cheney et al, 1993; Mooseker, 1993; Titus, 1993a for

reviews). The additional classes are numbered in order of their discovery, each class being evolutionary ancient.

A 'consensus' plan of the domain structure of the myosins can be described (Mooseker, 1993). The heavy chains of myosins are operationally defined by the presence of an N-terminal head-domain (except **ninaC** myosins, class-III (Porter et al, 1992)). The head domain is the myosin motor responsible for converting the energy of ATP hydrolysis into mechanical work through interaction with the actin filament. The head region of all myosins contains an actin binding site (Titus, 1993a), a 21KDa fragment of which has been shown to bind directly to F-actin (Muhlrad, 1991). The direction of movement of the head domain relative to the polarity of the actin filament is towards the barbed end (Mooseker, 1993).

Myosin heavy chain head and tail domains are connected by a neck domain, consisting of a variable number (1-6) of tandem repeats of an approximately 24 residues sequence known as the 'IQ' motif, each of which may be a light chain-binding site (Espreafico et al, 1992; Porter et al, 1992). Most myosins bind regulatory subunits called light chains that are calcium-binding proteins such as calmodulin (Rowe & Kendrick-Jones, 1992). Many myosins bind multiple calmodulin molecules, whereas myosins-II are associated with two types of specialised light chains: regulatory (RLC) and essential (ELC) light chains (Rowe & Kendrick-Jones, 1992).

The tail domain, the structure and size of which is highly variable amongst the various myosin classes (see Mooseker, 1993; Cheney et al, 1993 for diagrams), may define the functional specificity of a given myosin and/or define its subcellular localisation. The tail domain also defines the native oligomeric state of a given myosin (Mooseker, 1993).

5.1. Class-II Myosins

Class-II myosins are characterised by the presence of a rod-like α–tail that allows them to form two-headed dimers, which can further assemble into bipolar antiparallel filaments (Endow & Titus, 1992). The bipolar antiparallel filament arrangement of myosin-II is crucial for its function, which is to move groups of oppositely orientated actin filaments past each other, eg. as in muscle contraction. Myosins-II are associated with one RLC and one ELC at the neck region of each head

(Rowe & Kendrick-Jones, 1992). Vertebrate RLCs can be divided into two classes: (i) in smooth/nonmuscle myosins-II, phosphorylation of the light chains by a calcium-calmodulin dependent kinase (myosin light chain kinase) increases both interaction of the myosin head with actin and assembly of the myosin into filaments; (ii) the light chains of skeletal myosins-II are similarly phosphorylated but play no part in regulation (Rowe & Kendrick-Jones, 1992).

Myosins-II can be divided into two major groups - a striated muscle versus a nonmuscle/smooth group. The skeletal muscle myosin-II group may reflect the specialisations associated with striated muscle (Cheney et al, 1993). The nonmuscle/smooth myosin-II group can be subdivided into: nonmuscle myosins (eg. **myosin-IIA** and **B** found in chicken) and the smooth muscle myosins (Cheney et al, 1993). Like the 'nonmuscle' myosins-II of lower eukaryotes, the nonmuscle myosins in metazoans are believed to function in cytokinesis, receptor capping, and generation of cortical tension (reviewed by Aguado-Velasco & Kuczmarski, 1993).

5.2. Class-I

Myosins-I are monomeric and relatively small myosins (mammalian myosin-I is ≈ 110KDa). They were originally thought to be unique to either intestinal epithelial cells or protozoan ameoboid cells, but are now known to be ubiquitous (Barylko et al, 1992); class-I members being found in brain (see Titus, 1993a) and rat liver (Williams & Coluccio, 1994). In higher eukaryotes, myosin-I is best represented by **brush border myosin-I**. The 20KDa C-terminal region of brush border myosin-I is phosphorylated in vitro by protein kinase C (see Titus, 1993a). This phosphorylation, results in enhanced binding to phospholipid vesicles and actin activation. The actin activation of *Acanthamoeba* myosins-I require phosphorylation of the myosin heavy chain by myosin-I heavy chain kinase (Pollard et al, 1991).

Molecular cloning of myosins-I (reviewed by Hammer, 1991) show that the tails contain a highly basic region suggested to play a role in binding to membranes, and an SH3 domain that is present in many proteins associated with the membrane cytoskeleton (see Pollard et al, 1991). Since myosins-I are present in vertebrates, flies, amoebae,

Dictyostelium , and yeast (**MYO3** gene) (see Cheney et al, 1993 for review), the myosins-I represent an evolutionary ancient class of myosins.

5.3. Class-III

Class-III myosins are represented by a protein encoded by the ***ninaC*** gene in Drosophila (Montell & Rubin, 1988). *ninaC* is present in the rhabdomere, a part of the photoreceptor cell consisting of packed arrays of microvilli that contain cytoskeletal elements. *ninaC* mutants cause reduced rhabdomere size that appears to result in light and age-dependent retinal degeneration (Porter et al, 1992). The *ninaC* gene (Montell & Rubin, 1988) encodes two novel proteins: P174 localised to rabdomeres, and p132 found in the extra-rhabdomeric cytoplasm of the receptor cell body (Porter et al, 1992). Both proteins contain a ≈330 residue N-terminal protein kinase domain joined to myosin heavy chain-like head domain. The kinase domain has a role in phototransduction and in maintaining the rhabdomere (Porter & Montell, 1993). At present, *ninaC* is the most divergent member of the myosin superfamily so far discovered (Cheney et al, 1993).

5.4. Class-IV

An *Acanthamoeba* myosin heavy chain has been identified as being unique to other *Acanthamoeba* myosins (Horowitz & Hammer, 1990). The gene encodes a 177KDa polypeptide and is termed ***Acanthamoeba*** **high molecular weight (HMW) myosin**. The N-terminal 90KDa of this polypeptide is highly similar to the head sequences of myosins-I and II, the tail region does not share any similarity to any other myosins apart from a ≈50 residue SH3 domain at the tip of the tail (Horowitz & Hammer, 1990). Sequence analysis indicates that *Acanthamoeba* HMW is incapable of forming a myosin-II-like coiled-coil structure, implying that the protein is single headed and non-filamentous. *Acanthamoeba* HMW myosin is precipitated from crude extracts using F-actin and released from the pellet by ATP, supporting its inclusion as a member of the myosin superfamily (Horowitz & Hammer, 1990).

5.5. Class-V

A fifth class of myosins is formed by **mouse *dilute*** (Mercer et al, 1991), yeast **MYO2** (Johnston et al, 1991), yeast **MYO4** (Haarer et al, 1992) and the chicken **myosin-V (P190)** (Espreafico et al, 1992). Sequence analysis reveals that MYO2 shows 52%

identity to mouse *dilute* and chicken-V in its N-terminal myosin-like motor domain, and 22-36% identity along the remaining length (Endow & Titus, 1992).

Class-V myosins are characterised by the presence of six tandem IQ repeats responsible for calmodulin binding (Espreafico et al, 1992). Their tail regions consist of a proximal region predicted to form one or more segments of coiled-coil that does not form filaments under any conditions (see Titus, 1993b). Chick myosin-V has two mobile heads that are larger than those of conventional myosin, and the tail appears to contain two ≈45KDa globular projections (Titus, 1993b). Analysis of the *dilute* and MYO2 mutants indicate that both are necessary for survival and they appear to be involved in membrane targeting or vesicle trafficking (Mercer et al, 1991; Johnston et al, 1991). MYO2 is suppressed after overexpression of a kinesin microtubule motor and raises the possibility of interplay between microtubules and the actin-based cytoskeleton (Lillie & Brown, 1992).

5.6. Class-VI

A myosin, referred to as **95F**, that defines class-V myosins, has been identified in *Drosophila* (Cheney et al, 1993; Kellerman & Miller, 1992). An apparent homologue of this protein has been identified in a pig kidney proximal tubule cell line (Hasson & Mooseker, 1992). The tail domain of these myosins share no homology with other classes, and contains a segment predicted to form a coiled-coil α–helix. Although the functions of class-VI myosins are not fully known, *Drosophila* 95F myosin has a punctate and cytoplasmic location (Kellerman & Miller, 1992), and it has also been implicated in vesicle transport in the subcortical region of *Drosophila* embryos (Mermall et al, 1994).

5.7. Class-VII and Class-VIII

A seventh class of myosins comprises of the ***Drosophila* 35B, C** (Chen et al, 1991), and a similar protein identified in a pig kidney cell line (Hasson & Mooseker, 1992). These two myosins-VII are more closely related to each other than any other myosins according to phylogenic analysis by Cheney et al (1993). With only the head domains of the myosins-VII sequenced, nothing is known about their tail structure (Cheney et al, 1993).

Knight and Kendrick-Jones (1993) cloned a cDNA encoding a myosin-like protein from the plant *Arabidopsis thaliana* , and identified it as the first molecular motor of any kind from a higher plant, and termed it ***Arabidopsis* myosin-VIII**. The 131KDa polypeptide has a head domain similar to those of other myosins. The neck region contains four IQ motifs, with the tail region containing a segment of α–helical coiled-coil, indicating the molecule could be dimeric. The function of myosin-VIII in rigid plants remains unclear, but it may be involved in intracellular motile processes such as cytoplasmic streaming (Knight & Kendrick-Jones, 1993).

6. TROPOMYOSIN FAMILY

Tropomyosins (TM) are a family of actin filament binding proteins, identified in a variety of organisms, that exhibit extensive cell type specific isoform diversity (see Pittenger et al, 1994; Lees-Miller & Helfman, 1991 for reviews). TMs are elongated proteins that possess a simple dimeric α-helical coiled-coil structure along their entire length. The coiled structure is based on a repeated pattern of 7 amino acids (heptad repeat) that is highly conserved in all TM isoforms (Lees-Miller & Helfman, 1991). TMs are associated with the thin filaments in the sarcomeres of muscle cells and the microfilaments of non-muscle cells and bind to themselves in a head-to-head manner with each molecule interacting with 6 or 7 actin monomers (Lees-Miller & Helfman, 1991). Analysis of periodic distribution of the TMs reveals the presence of 7 or 14 quasi equivalent actin-binding sites. One actin-binding site corresponds with a half turn of the TM coiled-coil and spans 40 residues (reviewed in Vandekerckhove & Vancompernolle, 1992).

TM isoform diversity is generated by a combination of multiple genes, some of which contain alternative promoters and some of which exhibit alternative splicing of primary RNA transcripts (reviewed by Lees-Miller & Helfman, 1991). Four different TM genes have been characterised in vertebrates. *α* and *β* genes are named after striated muscle α and β TM respectively. The ***TM4*** and ***TMnm*** genes are named after the rat fibroblast TM-4 and human fibroblast TM30nm isoforms respectively (see Pittenger et al, 1994 for list). The *α-TM* is the most complex of vertebrate genes and encodes at least 9 different isoforms, with only five exons common to all mRNAs expressed from the gene in rat (Ruiz-Opazo & Nadal-Ginard, 1987). Mutations in the *α-TM* cause familial

hypertropic cardiopathy (FHC), which is a disease of the sarcomere (Thierfelder et al, 1994). Alternative promoters in this gene result in mRNAs that encode two distinct N-terminal amino acid sequences. Alternative promotion along with alternative splicing results in TMs belonging to the 284 residue (high Mr) class or the 248 residue (low Mr) class of TMs (Pittenger et al, 1994). The *β–TM* gene (Libri et al, 1989) of both chicken and rat encode skeletal muscle and smooth muscle isoforms. In rat, the smooth muscle β–TM is also identical to a major fibroblast isoform termed TM-1 (see Pittenger et al, 1994). *TMnm* gene has been identified in human encoding a slow-twitch skeletal muscle isoform (284 residues) and in rat where the gene product is referred to as TM-5 (248 residues) (Pittenger et al, 1994). Rat *TM-4* gene is the simplest of vertebrate TM genes (Lees-Miller et al, 1990b) and contains 8 functional exons and encodes a single nonmuscle isoform, fibroblast TM-4 (248 residues). The significance of this TM isoform diversity is not fully understood, nonetheless, the repertoire of different isoforms suggests that each isoform is required to carry out specific functions in actin-based filaments of various muscle and nonmuscle cell types (see Pittenger et al, 1994).

In striated muscle the α–isoform is more prominent in cardiac muscle of mammals and in fast-twitch skeletal muscle, while the β–isoform is abundant in slow-twitch muscles (Lees-Miller & Helfman, 1991). The function of TM in skeletal muscle is, in association with troponin complex, to regulate the calcium sensitive interaction of actin and myosin heads (Lees-Miller & Helfman, 1991).

In smooth muscle, the differences in sequence of the α– and β– TMs (which are encoded by the same genes as in striated muscle) reflects the mechanism of contraction in the respective cell types (Pittenger, et al, 1994). Smooth muscle does not contain the troponin complex, thus, its actomyosin is activated through phosphorylation of myosin light chain and caldesmon (Pittenger et al, 1994).

Nonmuscle cells express a larger number of TM isoforms than do muscle cells, and reflects the greater diversity of cellular functions rather than functional redundancy (Goodwin et al, 1990). Rat fibroblasts contain 7 isoforms of TM (see Pittenger et al, 1994; Goodwin et al, 1990), and brain expresses 3 unique isoforms (TMBr-1, 2 and 3) (Lees-Miller et al, 1990a). Bacterial expression of individual nonmuscle isoforms have revealed unique F-actin-binding features to each isoform (Pittenger & Helfman, 1992)

that may be due to differences in an internal exons (Pittenger et al, 1994). Microfilaments in a fibroblast represent dynamic structures that can exist in different supramolecular forms such as stress fibres, microfilament meshworks, polygonal networks and contractile rings (see Satterwhite and Pollard, 1992 for review). Thus, as TMs are associated with actin, the multiplicity of TM isoforms in fibroblasts raises the possibility that specific associations of given isoforms with themselves and other cellular proteins lead to these distinct actin structures (Lees-Miller & Helfman, 1991). TM may also play a pivotal role in stabilising actin filaments in nonmuscle cells. This role is inferred from the fact that TM protects actin filaments against severing action of gelsolin (intro, section F. 2) or ADF (intro, section F. 4) (experiments reviewed in Pittenger et al, 1994). The stabilising role for TM in non-muscle cells has also been implied from studies of transformed cells demonstrating that changes in microfilament structure are accompanied by alterations in TM expression (high Mr TMs are down-regulated) (Vandekerckhove et al 1990a).

Tropomodulin is a 43KDa TM-binding protein first isolated from human erythrocyte membrane skeleton that exhibits isoform-specific binding to one of the ends of TM and blocks TM head-to-tail association along the actin filament (Fowler et al, 1993). cDNA derived amino acid sequence of tropomodulin reveals no homology to any other TM-binding proteins (Fowler et al, 1993; Pittenger et al, 1994). Tropomodulin can act as an actin filament pointed end capping protein (Fowler et al, 1993) and may cap the TM polymers at the pointed end of the thin filament and play a role in regulating thin filament length at the I-A band junction of muscle fibres (Fowler et al, 1993).

A TM homologue identified in *S. cerevisiae* , designated **TPM1**, is a ≈23.5KDa, 199 residue polypeptide, that is localised with actin cables in wild type cells; disruption of which results in the loss of actin cables and slow growing cells (Liu & Bretscher, 1989). Many cells contain an abundance of secretory vesicles, and studies with the above null mutants suggest that TM interacts with yeast class-V myosin, MYO2 (intro, section F. 5. 5), which is thought to be involved in vesicle transport (Liu & Bretscher, 1992).

The *S. pombe* cdc8 gene product is a 161 amino acid protein, associated with the F-actin contractile ring, containing features that identify it as a TM: (i) sequence similarity (37% identity to TPM1); (ii) high predicted α–helical content (98.5%); (iii) a

repeating 7 residue pattern (Balasubramanian et al, 1992). In addition, Haploid cells with a disrupted cdc8 gene can be rescued by expression of rat fibroblast TM-2 (Balasubramanian et al, 1992).

A chick intestinal cDNA screened with brush border myosin-I antibody isolated clones encoding 239 amino acids (≈28KDa) containing the characteristic TM heptad repeats, and was designated **leucine zipper**. Leucine zipper has 34% identity (45% in the first 75 amino acids) with human skeletal α–TM. Localisation of this TM-like protein at the brush border region indicate a function in regulating the interaction of myosin-I with actin bundles (Bikle et al, 1993).

7. CALDESMON

Caldesmon (CaD) has been identified in a wide range of tissues and cells (see Marston & Redwood, 1991 for review).The primary sequences of caldesmon are divided into either the h, high molecular weight (120-150KDa; CaDh), or l, low molecular weight (70-80KDa; CaDl) isoforms, both of which can be split into α– and β– forms in chicken (Marston & Redwood, 1991). CaDh is predominantly expressed in smooth muscle, whereas CaDl is found in nonmuscle cells (Hayashi et al, 1989). The protein sequences of the chicken CaDh α and β isoforms identify a long predicted α-helical region containing a 15 amino acid motif repeated 8-12 times. CaDhα differs from CaDhβ by the addition of one of these repeats (Bryan et al, 1989; Hayashi et al, 1989). Sequencing of nonmuscle CaD indicates that it is smaller than the muscle form (Marston & Redwood, 1991). Both the N- and C- termini are conserved, with nonmuscle CaD lacking 232 residues in the central region. Human smooth muscle and nonmuscle CaD are very similar to those of chicken, but only has 3-4 central repeats (Humphrey et al, 1992). Rat CaDl has also been determined, its sequence differing in the N-terminal region to chicken CaDl (see Marston & Redwood, 1991 for alignments). Both smooth muscle and nonmuscle CaD are elongated (75nm), dumbbell shaped structures in solution (Matsumura & Yamashiro, 1993).

The N-terminal domain of CaD contains a binding site for myosin subfragment 2 and a weak calmodulin-binding site and the C-terminal head contains calmodulin and tropomyosin-binding domains (Marston & Redwood, 1991). CaD contains at least two actin-binding regions: one located to the C-terminus associated with a calmodulin-

binding site (Wang et al, 1991), and a second region also predicted to be adjacent to a calmodulin-binding site (see Vandekerckhove & Vancompernolle, 1992 for review).

Immunocytochemically, CaDl in normal fibroblasts, is found periodically distributed along stress fibres (Marston & Redwood, 1991), such localisation is in agreement with the concept that caldesmon is a regulatory protein for the actomyosin system (review by Marston & Redwood, 1991). Indeed, CaD is an inhibitor of the ATPase, binding near the actin N-terminus (the same region as myosin) suggesting it inhibits ATPase activity of actomyosin by sterically hindering actin-myosin interaction (Matsumura & Yamashiro, 1993). CaD may also inhibit by blocking the release ATP hydrolysis products from myosin (Marston, 1988). However, other studies suggest that CaD may inhibit by competing with myosin for a common binding site on actin (Velaz et al, 1990). Harricane et al (1992) have shown that the N-terminal region of actin is not available to a 50KDa, trypsin-split product, of myosin subfragment 1 when CaD is bound to actin, this indicates CaD is bound to actin and could therefore inhibit ATPase activity of actomyosin. The inhibitory effect can be rendered Ca^{2+}-dependent by the binding of calmodulin or caltropin in smooth muscle (Mani et al, 1992). Smooth muscle also contains calponin (intro, section F. 20) which has properties similar to CaD (Barany et al, 1992) even though there is little sequence homology between them (Takahashi & Nadal-Ginard, 1991)

CaD is phosphorylated in vitro by a number of kinases (see Matsumura & Yamashiro, 1993 for list), and dissociates from microfilaments when phosphorylated by a cdc2 kinase during mitosis; Yamashiro et al (1991) have demonstrated the kinase could be $p34^{cdc2}$. Phosphorylation of CaD may play an important role in the architecture of the cytoskeleton, but little is known regarding the function of CaD in vivo. CaD has an ability to cross-link actin and myosin, and in smooth muscle it is thought to play a role in the latch-state contraction, where much smooth muscle tissue maintains high tension even after Ca^{2+} levels has dropped and myosin dephosphorylated (see Marston, 1989). In nonmuscle cells CaD is thought to provide regulatory functions in cell motility. Nonmuscle CaD is down-regulated in RSV transformed fibroblasts (Koji-Owada et al, 1984) where such stringent regulation of microfilaments is not required.

8. CAPZ-RELATED: CAPPING PROTEINS

Capping proteins (see Pollard & Cooper, 1986 for review) are a highly conserved family of actin-binding proteins that cap the barbed ends of actin filaments, nucleate polymerisation of actin monomers, but neither sever filaments, nor require Ca^{2+} for their activity (Weeds & Maciver, 1993).

Capping proteins are heterodimers with non-similar subunits of Mr 32-36KDa (α) and 28-32KDa (β) and are found in all eukaryotes examined (Weeds & Maciver, 1993). Chicken CapZ α–subunit is encoded by two expressed genes, and the β–subunit by one (Casella et al, 1989; Caldwell et al, 1989; Cooper et al, 1991). The classification of capping proteins is confirmed by protein sequence derived from cDNA cloning for chicken (Casella et al, 1989; Caldwell et al, 1989; Cooper et al, 1991) and *Dictyostelium* (Hartmann et al, 1989). These sequences have been used to identify the gene for the β-subunit of capping protein in *S. cerevisiae* (Amatruda et al, 1990). The sequences share no homology with other known actin-binding proteins, but they are highly conserved, with 49% identity between the β-subunits of yeast and chicken (Amatruda et al, 1990).

The role of CapZ is probably the attachment of actin filaments to Z lines via their barbed ends, since CapZ binds to barbed ends of actin filaments in vitro and is located at the Z line (Casella et al, 1987). During myofibrillogenesis of cultured chicken muscle, capZ is found distributed at Z lines before actin filaments are organised in a pattern characteristic of I bands (Schafer et al, 1993). CapZ is a component of Z discs in cardiac myocytes and is colocalised with vinculin (intro, section F. 11) at cell-cell junctions (Schafer et al, 1993). Studies have shown that PIP_2 binds to the α-subunit of capZ and completely inhibits its ability to affect actin polymerisation (Heiss & Cooper, 1991). It has been speculated that, as capping protein binds to and caps barbed ends of actin filaments at the membrane, it is regulated by the phospholipid composition of the membrane, which varies spatially and temporally, enabling growth of actin filaments to be regulated by signal transduction from the membrane (Heiss & Cooper, 1991).

The role of capping protein in non-muscle cells is less clear. Its nucleating function may help control the number and length of actin filaments, key parameters in the properties of the actin network (Caldwell et al, 1989). The chicken genes are expressed in all muscle and non-muscle tissue (Caldwell et al, 1989; Cooper et al, 1991), so it is

probable that non-muscle cells contain the same capping protein as muscle cells and that it therefore serves a similar physiologic role. Mutations in the yeast capping protein β-subunit, **CAP2**, leads to abnormal actin distribution, including loss of actin cables (Amatruda et al, 1990). CAP2 has been colocalised with actin at the tips of growing buds, but not in actin cables or in the contractile ring during cytokinesis (Amatruda & Cooper, 1992). These data suggest capping protein in nonmuscle cells is associated with membrane-cytoskeletal contacts and regulating actin-filament distribution in vivo (Amatruda & Cooper, 1992). Nuclear staining is also a prominent feature of capping protein localisation in Xenopus non-muscle cells (Ankenbauer et al, 1989).

Hug et al (1992) have identified that capping requires both α– and β–subunits. In the chicken β-subunit a proposed ABD is identified, capping ability being lost when the C-terminal 12 residues of the β-subunit are removed (Hug et al, 1992).

The Z-line localisation and the size of CapZ are similar to that of **β–actinin** (Maruyama et al, 1990). Partial sequence data for β-actinin (Maruyama et al, 1990) confirms that it is identical to CapZ with minor sequence differences. Although β–actinin had been reported to bind to the pointed ends of actin-filaments, it has been shown that highly purified β–actinin binds to the barbed ends of filaments, in a manner similar to CapZ (Maruyama et al, 1990). The minor sequence differences noted, and multiple isoforms of β-actinin suggest that some isoforms of β-actinin/CapZ may exist which bind preferentially to the pointed rather than the barbed end (Hartwig & Kwiatkowski, 1991).

9. TENSIN AND INSERTIN FAMILY

Antibodies raised against the contaminants of crude vinculin (intro, section F. 11) preparations cross-react with bands of 150KDa and 200KDa (Wilkins et al, 1986). The 150KDa polypeptide was shown to interact with the barbed end of actin filaments in vitro (Wilkins et al, 1987). Resultant polyclonal antibodies raised against the 150KDa band recognised focal contacts and other submembranous cytoskeletal structures (Risinger & Lin, 1988) as well as both the 150KDa and 200KDa proteins. These two polypeptides were called **tensin** because of their putative function in maintaining tension in the microfilaments at their point of anchorage.

Tensin cDNA clones have been isolated from a chicken embryo fibroblast library (Davis et al, 1991) and an adult chicken heart library (Chuang et al, 1991). Overlapping

clones encode a polypeptide of about 200KDa. Deduced amino acid sequence reveals the presence of an SH2 domain. This domain is shared by a number of signal transduction proteins including non-receptor protein kinases (see $pp60^{c\text{-}src}$; intro, section F. 19),the transforming protein Crk, phospholipase C-γ and the ras GTPase activity protein (Davis et al, 1991). The SH2 domain of tensin binds specifically to a number of phosphotyrosine-containing proteins from v-src transformed cells (Davis et al, 1991). Tensin is phosphorylated on tyrosine residues suggesting that by possessing both actin-binding and phosphotyrosine-binding activities and being itself a target for tyrosine kinases, tensin may link signal tranduction pathways with the cytoskeleton (Davis et al, 1991).

Insertin, a 30KDa protein isolated from chicken gizzard smooth muscle, binds strongly to the barbed ends of actin filaments (Ruhnau et al, 1989). It has been suggested to be a functionally important part of the machinery that permits treadmilling of membrane-attached actin filaments in living cells (Ruhnau et al, 1989). Insertin is shown to retard, but not inhibit, actin polymerisation at the barbed ends of filaments, which can be explained by the 'walking insertion model' proposed by Gaertner & Wegner (1991). Briefly, it is known that two insertin molecules bind cooperatively to the barbed end of a filament (Ruhnau et al, 1989). The binding site at one strand of the double helical actin filament is free for incorporation of a new actin molecule. The terminal subunit of the other strand is occupied by one molecule of the insertin dimer. Following incorporation of an actin molecule the other subunit of the insertin dimer binds to a new actin subunit, while the previously bound insertin molecule dissociates from the other strand of the actin double helix to produce a new binding site for the next actin molecule. In this model insertin remains bound to the barbed end of filaments during polymerisation (see Gaertner & Wegner, 1991 for diagram of model).

The amino acid sequence of insertin has been determined (Weigt et al, 1992) and its primary sequence is almost identical (only 2 amino acids difference) to amino acid residues 862-1212 of tensin. Thus, tensin contains a domain that possesses actin-inserting activity. As tensin has been shown to occur at the membrane adhesion sites of actin filament bundles (see above), actin-inserting domains are localised at sites where actin filaments are linked to membranes. Therefore, membrane-bound actin filament bundles

in cells are likely to assemble by a mechanism in which actin monomers are inserted at the barbed ends while these ends remain attached to the protein molecules of the membrane adhesion sites (Weigt et al, 1992). It is possible insertin and tensin are coded by two different genes in which the actin-inserting domain has been highly conserved. It is also possible, however, that insertin is a proteolytic fragment of tensin (Weigt et al, 1992).

10. FASCIN GROUP

The fascin group of proteins constitute a family of actin-bundling proteins that share sequence similarity (Otto, 1994), with widespread representatives in nature. These group of proteins are not normally believed to be localised in classical microvilli and may represent an unusual filament organising principle. No extensive similarities between these proteins with other known actin-binding/bundling proteins has been detected (Matsudaira, 1991), indicating that this is a separate gene family.

Fascin, a 58KDa protein that bundles actin filaments into hexagonally packed linear arrays, was originally identified from extracts of unfertilised sea urchin eggs (reviewed by Bryan & Kane, 1982). It has been localised in the bundles of F-actin that form the cores of microvilli on the surface of fertilised sea urchin eggs and spikes on starfish oocytes (Otto, 1994). Sequencing of fascin cDNA from a larval stage sea urchin cDNA library reveals an encoded polypeptide of 496 residues (Bryan et al, 1993). Fascin peptide sequence identifies a gene family of proteins that includes the *Drosophila* ***singed*** (***sn***) gene product (Paterson & O'Hare, 1991) and a 55KDa (**p55**) actin-bundling protein in mammalian cells (Yamashio-Matsumura & Matsumura, 1985; 1986).

Studies indicate that the **sn** gene of *Drosophila* has a role in the female germline during oogenesis and in somatic cells during the formation of adult bristles and hairs (Paterson & O'Hare, 1991). *Sn* gene product is required for actin filament bundle formation in cytoplasm nurse cells during oogenesis (Cant et al, 1994); in *sn* mutants, the absence of cytoplasmic actin filament bundles allows nurse cell nuclei to lodge in ring canals and block nurse cytoplasmic transport (this is similar to the *kelch* gene product (intro, section F. 17)). Sn is also required for organising actin filament bundle formation in the cellular extension that forms a bristle (Cant et al, 1994); mutants displaying shortened, gnarled bristles. In addition to *sn*, mutations in the *Drosophila*

genes ***forked, chickadee*** (intro, section F. 13) and ***stubble-stubbloid*** also affect the actin filaments in bristles (Petersen et al, 1994; Verheyen & Cooley, 1994; Appel et al, 1993), and may therefore represent new actin-binding proteins in *Drosophila*. Cloning of the *sn* cDNA encodes a 57KDa protein (Paterson & O'Hare, 1991) that has 35% identity with fascin, and bundles actin filaments in vitro with the same stoichiometry as it (one fascin molecule per filament cross-link) (Bryan et al, 1993).

Purified from HeLa cells, **p55** is a monomeric globular, Ca^{2+}-insensitive actin-bundling protein (Yamashiro-Matsumura & Matsumura, 1985).The peptide sequence of p55 identifies it as part of the fascin family, indicating a mammalian homologue (see Bryan et al, 1993 for alignment). P55 causes F-actin to aggregate side-by-side into bundles, a process inhibited by tropomyosin, suggesting that their binding site on actin is in the same vicinity (Yamashiro-Matsumura & Matsumura, 1985). The function of a fascin group homologue in mammalian cells is poorly understood, but the functional studies in sea urchins suggest it may play a role in the organisation of a class of actin bundles and microvilli of cellular projections that differ from those defined by villin (intro, section F. 2) and fimbrin (intro, section 1. 1) (Yamashiro-Matsumura & Matsumura, 1985).

11. VINCULIN GROUP

Vinculin, a 117KDa protein, is associated with adhesion plaques and cell-to-cell contacts at the termini of microfilament bundles (Geiger & Ginsberg, 1991). It is also present in a very broad spectrum of cells including lymphoid and platelets which do not form stable adhesions. Among vinculin-binding proteins detected are talin (Belkin & Koteliansky, 1987), Paxillin (Turner et al, 1990), and α-actinin (Wachsstock et al, 1987). Vinculin expression may be modulated in cells due to different environmental stimuli; eg. fibroblasts regulate the expression of vinculin mRNA in response to changes to cell density and adhesiveness (Bendori et al, 1987), and its gene expression is rapidly, yet transiently, stimulated in quiescent 3T3 cells upon growth activation by serum factors (Bellas et al, 1991). Increased vinculin expression in 3T3 cells results in extensive spreading with abundance of stress fibres and adhesion plaques accompanied by a dramatic decrease in cell motility (Rodriguez-Fernandez et al, 1992a).

Study of the effect of the forced reduction of vinculin by use of vinculin antisense transfection in 3T3 cells showed that the cells had a more rounded phenotypes, fewer focal contacts, increased motility (manifested by faster closure of wounds), and larger colonies in soft agar than normal cells (Rodriguez-Fernandez et al, 1993). The results demonstrate that the regulation of vinculin expression in cells can effect cell shape and motility, inducing changes similar to those found in transformed cells. Elevation of vinculin expression in malignant fibroblastic and epithelial cells which express diminished levels of vinculin, resulted in the suppression of the tumorigenic ability and anchorage independence of these cells (Rodriguez Fernandez et al, 1992b).

Sequencing of chicken and mammalian vinculin cDNA (Price et al, 1987; Weller et al, 1990) shows they are highly homologous (>95%). The deduced polypeptide consists of 1066 amino acids and has a large N-terminal head domain which contains a talin binding site and three ≈110 amino acid repeats (Price et al, 1987; Weller et al, 1990). The N-terminal domain of human vinculin is 54% identical to nematode vinculin with the C-terminal domain being 61% identical (Weller et al, 1990), indicating the evolutionary importance of vinculin. A higher molecular weight variant of vinculin (124KDa) is detected only in muscle and denoted **metavinculin** (Gimona et al, 1988). Alignment of metavinculin sequence with that of chicken vinculin indicates the presence of an insert of 68 amino acids within metavinculin. Identity of the mapped vinculin and metavinculin sequences outside this difference insert is consistent with the two proteins arising via alternative splicing at the mRNA level (Gimona et al, 1988).

Vinculin is posttranslationally modified by phosphorylation on either serine or tyrosine residues in normal fibroblasts and platelets (see Vostal & Shulman, 1993). In transformed cells, vinculin can be phosphorylated by tyrosine-specific protein kinases encoded by the transforming genes of a number of retroviruses including $pp60^{src}$ (intro, section F. 19) (Kellie et al, 1991).

The peripheral cytoplasmic proteins **catenins** have been identified by their ability to interact with E-cadherin (uvomorulin), a protein responsible for Ca^{2+}-dependent cell-cell adhesion (see Tsukita et al, 1992 for review). α, β and γ-catenins are structurally distinct cytoplasmic proteins with apparent Mr of 102, 88 and 80 KDa respectively (Ozawa et al, 1989). Catenins link E-cadherin to actin bundles, with α–catenin playing a

key role in the association with actin filaments, β-catenin binding more directly to the cytoplasmic region of E-cadherin, and γ-catenin located in the periphery of the complex (Ozawa et al, 1989). Catenin-E-cadherin complex is shown to be essential for cell adhesive function (Ozawa et al, 1990).

Xenopus β-catenin has been sequenced and is shown to be 70% identical to the *Drosophila* **Armadillo** protein (McCrea et al, 1991), therefore identifying *Drosophila* β–catenin. Catenins have been grouped together with vinculin as comparison of deduced sequence of human α–catenin (906 amino acids) with that of human vinculin suggests they have a similar structure and function (Nagafuchi et al, 1991).

12. THE β- THYMOSIN FAMILY

The β- thymosins are a family of small proteins whose major function is to sequester actin monomers in different vertebrate cells (reviewed by Nachmias, 1993). Nine β- thymosins have been identified; six having been shown to bind to actin. The sequence of all β- thymosins (see Nachmias, 1993 for sequences and family list) are closely related, with the substitutions between the homologues minimal, many being conservative, retaining hydrophobicity or ionic character. One region is completely invariant, and is analogous to other actin-binding proteins (Safer et al, 1991). Vancompernolle et al (1992) showed this invariant hexapeptide motif (LKKTET) forms the major contact site through electrostatic interactions with actin. An N-terminal segment preceding this motif exerts an inhibitory effect on actin polymerisation, probably by steric hinderance (Vancompernolle et al, 1992). All vertebrates contain one or often two β- thymosins, and a sea urchin homologue (≈70% identity to vertebrate β-thymosins) has also been identified (Nachmias, 1993).

Two major β- thymosin isoforms are found in mammals and termed **thymosin-β4** (**Tβ4**) and **thymosin-β10** (**Tβ10**). Rat spleen cDNA for Tβ4 reveals a transcript of 0.8Kb encoding a 44 residue polypeptide; removal of the methionine initiator and acetylation of the N-terminal serine yields a mature Tβ4 of 43 residues with a Mr of 5KDa (Wodnar-Filipowicz, 1984). Immunofluorescence studies using Tβ4 and Tβ10 isoform specific antibodies in peritoneal macrophages showed that both β- thymosins are uniformly distributed within the cytoplasm (Yu et al, 1994). Overexpression of the β- thymosins in CV1 fibroblasts induces extensive stress fibre loss and a decrease in the number of focal

adhesions which indicates that both are effective regulators of a large subset of actin filaments in living cells (Yu et al, 1994).

Tβ4 is a major G-actin-binding protein, forming a 1:1 complex with α-actin (Weber et al, 1992), inhibiting salt induced polymerisation, but not capping or severing actin filaments (Safer et al, 1990). It has been identified as the main G-actin-sequestering agent in platelets (Safer et al, 1991) and neutrophils (Cassimeris et al, 1992), where it is present at concentrations in the $1x10^{-5}$ to $5x10^{-4}$M range.

In nonmuscle cells a high concentration of actin (0.1mM) is maintained unpolymerised by interaction with sequestering proteins, including Tβ4 (Cassimeris et al, 1992). Polymerisation of actin is induced upon appropriate stimulation, which implies that some method of desquestration controls the rapid formation of short filaments (Cassimeris et al, 1992). Tβ4 is found to have a 50-fold higher affinity for ATP-actin than for ADP-actin (Carlier et al, 1993). This suggests that in unstimulated (resting) platelets and neutrophils, actin is sequestered by Tβ4 as ATP-G-actin. This ATP/ADP dependence of Tβ4 affinity for G-actin can generate a mechanism of desequestration of G-actin by ADP, in the presence of physiological concentrations of Tβ4 (Carlier et al, 1993). The desequestration of G-actin by ADP is kinetically enhanced by profilin (Goldschmidt-Clermont et al, 1992). Low amounts of profilin (Which may be regulated by PPI metabolism) catalytically accelerates the dissociation of ATP from G-actin, promoting extensive actin assembly from the pool of actin/Tβ4 complex (Pantaloni & Carlier, 1993). Profilin acts like a pump that allows the shift of the pool of unpolymerised (sequestered) actin/Tβ4 into the F-actin pool (Pantaloni & Carlier, 1993), the presence of Tβ4 being essential to provide a reservoir of unpolymerised actin that can be shifted into the filament pool by profilin action.

13. PROFILINS

Profilin is a small (126-140 amino acids; 12-15KDa by SDS-PAGE), globular, cytoplasmic protein that binds actin monomers (Carlsson et al, 1977), and associates with plasma membranes (Hartwig et al, 1989). Profilins are highly conserved and present in all eukaryotic cells studied (Machesky & Pollard, 1993)

Profilin was originally identified as a component of cell extracts that inhibited actin filament growth (Carlsson et al, 1977), and was assumed to be the major sequestering factor in most cells. However, not enough profilin is found in a typical cell (Babcock & Rubenstein, 1993); Tβ4 (intro, section F 12) has now been identified as the major actin monomer sequester in most cells (Safer et al, 1990). The interaction between profilin and actin is now thought to be much more complex (reviewed by Theriot & Mitchison, 1993). Profilin strongly inhibits elongation of actin filaments at the pointed (slowly growing) end, but does not substantially inhibit elongation at the barbed end (Theriot & Mitchison, 1993). Infact, profilin-actin complexes bind directly to the barbed end of preexisting filaments, and then the profilin is released, leaving the actin monomer behind (Pring et al, 1992). It has been shown that elongation of actin filaments by the addition of the profilactin complex is thermodynamically more favourable than elongation through the addition of monomer alone. Because of this, profilin can lower the critical concentration for ATP-actin elongation at the barbed end. This effect is most pronounced in the presence of Tβ4 (Pantaloni & Carlier, 1993).

Profilin also acts as a nucleotide exchange factor for actin. When bound, profilin increases the rate constant for dissociation of adenine nucleotide from actin by 3-orders of magnitude (Goldschmidt-Clermont et al, 1991b). Relatively low concentrations of profilin can catalytically effect nucleotide exchange on actin monomers, even in the presence of Tβ4 (Goldschmidt-Clermont et al, 1992). Thus, spatial variations in profilin concentrations could control local ratios of ATP-actin to ADP-actin. Since ATP-actin polymerises more rapidly and with a lower critical concentration than does ADP-actin, the nucleotide exchange activity of profilin might substantially promote filament elongation in a local region of a cell (reviewed by Theriot et al, 1994).

The two effects that profilin has on actin monomer, lowering the critical concentration and stimulating nucleotide exchange, indicate that profilin activity should generally promote actin filament growth in cells. Yet in vivo, experiments yield mixed results. Microinjection of profilin into fibroblasts causes rapid loss of F-actin structures in the cell (Cao et al, 1992), but overexpression of profilin in CHO cells stabilises filaments (Finkel et al, 1994). Profilin may be necessary to promote formation or stability of F-actin structures: deletion of *Drosophila chickadee* is lethal, disrupting multiple

actin-dependent processes during development (Verheyen & Cooley, 1994); cytoplasmic actin networks critical for proper oocyte development fail to form the egg chambers of mutant flies (Cooley et al, 1992). When profilin is deleted in *S. cerevisiae,* actin cables fail to form and the cells unable to divide normally (Haarer et al, 1990). cDNA for *Drosophila* profilin transfected into *S. pombe* , strongly inhibits cell division and normal cell shape, and increases the overall concentration of F-actin (Edwards et al, 1994). Profilin has been localised in cells to regions of rapid filament growth (Buβ et al, 1992), and localised in cells at the surface of *Listeria monocytogenes* (Theriot et al, 1994). It is not known whether the effects of profilin on nucleotide exchange or on ATP-actin critical concentration are responsible for the above observations.

In addition to regulating actin polymerisation, there is evidence that profilin has a role in regulating signal transduction. Profilin binds the membrane phospholipid, PIP_2, with high affinity (Lassing & Lindberg, 1985). In vitro, PIP_2 bound to profilin cannot be converted to the second messengers DAG and IP_3 by unphosphorylated phospholipase C-γ1 (PLC-γ1) (Goldschmidt-Clermont et al, 1990). When PLC-γ1 is phosphorylated by EGFR, it can hydrolyse PIP_2 bound to profilin, thus releasing profilin (Goldschmidt-Clermont et al, 1991a). This suggests profilin blocks PLC-γ1 function until that enzyme is phosphorylated, in effect protecting PIP_2 from unregulated hydrolysis. In addition, the liberated profilin would be available to interact with actin to modulate filament assembly. Furthermore, profilin may interact with the Ras pathway in *S. cerevisiae*; overexpression of profilin has been shown to restore normal morphology in yeast cells carrying a null mutation for CAP, an effector of Ras (Vojtek et al, 1991). Of the two profilin isoforms found in *Acanthamoeba,* profilin II is more efficient at suppressing the CAP mutant phenotype than profilin I Since profilin II has the higher affinity for PIP_2, it is possible that the CAP rescuing of profilin is related to the interaction of profilin with PIP_2 (Goldschmidt-Clermont & Janmey, 1991).

Polyproline-binding is a conserved feature in profilins of all species (Metzler et al, 1994; Machesky & Pollard, 1993). Mutations in the C-terminal region of yeast profilin eliminate the ability of profilin to bind polyproline in vitro, indicating a polyproline binding site (Haarer et al, 1993). More severe mutants in this region do not, however, effect actin-binding, suggesting that this region is not solely responsible for

actin-binding (Haarer et al, 1993). Profilin contains a src homology (SH) 3-like fold (Bjorkegren et al, 1993), and its tight interaction with polyproline is reminiscent of the binding activity exhibited by SH3 domains. However, replacement of hydrophobic amino acids in regions of profilin different from the SH3-like-domain, completely abolishes the profilin polyproline-binding capacity (Bjorkegren et al, 1993). Eight to ten sequential prolines are required to bind profilin. Stretches of prolines of this length occur in numerous regulatory proteins (eg. CAP), suggesting that the ability of profilin to bind polyproline may be an important component of its signalling capabilities (Metzler et al, 1994; Machesky & Pollard, 1993).

Many species have multiple isoforms of profilin. *Acanthamoeba* has at least three functional isoforms of profilin (**IA**, **IB** and **II**) (Ampe et al, 1985; Kaiser et al, 1986; Pollard & Rimm, 1991), profilin IA being the most abundant. *Acanthamoeba* profilin I and II are almost identical except for a region between residues 24 and 66 where 47% of the amino acids differ (Pollard & Rimm, 1991), but the functional significance of different isoforms of profilin is not clear. 3D analysis of *Acanthamoeba* profilin (Vinson et al, 1993) indicates that an outer region of profilin I contains the difference in amino acids between the isoforms, and therefore probably determines their difference in PPI-binding.

S. cerevisiae has one known profilin gene encoding a profilin I isoform of 126 amino acids (Magdolin et al, 1988). In *Drosophila* the 1.0 and 1.2Kb mRNAs from the ***Chickadee*** gene encode a protein of 126 amino acids that is 40% identical to profilins from yeast and *Acanthamoeba* (Cooley et al, 1992). Highest homology exists within the C-terminus, and it has been shown that the region between residues 95-125 of *Acanthamoeba* profilin I and II contain an actin-binding site (Vandekerckhove et al, 1989).

Three members of the profilin multigene family from the plant *Zea mays* have been isolated (Staiger et al, 1993). They encode for polypeptides of 131 residues (14KDa)-137 residues (14.8KDa). All three are highly homologous to each other (>90% identity) and 30-40% identical to eukaryotic profilins (Staiger et al, 1993).

Homo sapiens have two genes encoding profilin I (Kwiatkowski & Bruns, 1988) and profilin II (Honore et al, 1993). Both isoforms are polypeptides of 140 amino acids

and share 62% identity. Human profilin I shares 95% identity to bovine and mouse profilins and 25-28% identity with *Acanthamoeba* profilins. Human profilin II shares 63% identity to mouse and bovine profilins and 26-30% with profilins from *Acanthamoeba* (Honore et al, 1993).

14. THE MARCKS FAMILY

Diacylglyceral-activation of members of the protein kinase C (PKC) family leads to a wide variety of physiological effects. Study of the cellular substrates of PKC may help explain its down-stream effects. One class of PKC substrates is the **myristoylated, alanine-rich, C kinase substrate, MARCKS** protein (Blackshear, 1993).

The primary sequence of MARCKS from mouse (Seykora et al, 1991), bovine (Stumpo et al, 1989) and various other species have been identified (see Blackshear, 1993 for alignments) and range in size from 31.8KDa (bovine) to 27.7KDa (chicken). Overall homology between species is ≈70%; all are very alanine-rich (up to 31% in bovine) (Blackshear, 1993).

Three regions of conservation exist between all MARCKS proteins. The first consists of the N-terminus, of which the first 14 residues are identical in all species. The region contains a myristoylation site which comprises a membrane binding domain (Graff et al, 1989a). A second conservative region surrounds the site of intron splicing; differential use of this splice accounts for two forms of mRNA (2.6 and 4.5Kb) (Blackshear, 1993). A final conserved region is a basic α–helical domain in the middle of the polypeptide chain. This region contains C-kinase sites (Stumpo et al, 1989; Seykora et al, 1991), and also contains calcium-calmodulin-binding and actin-binding sites (reviewed in Aderem, 1992a).

MARCKS is a prominent substrate for PKC (see Blackshear, 1993 for review) and binds to calcium-calmodulin. Calmodulin-binding is prevented by PKC phosphorylation of MARCKS, and it has been suggested that MARCKS could serve as a reservoir for calmodulin in states of PKC inactivity; perhaps releasing calmodulin from the membrane upon PKC activation (Graff et al, 1989b). When N-terminal mutants of MARCKS were expressed in mammalian cells (Graff et al, 1989a), myristoylation and membrane association were both prevented, isolating the myristoylation site to the N-terminus. The myristate group of MARCKS binds to membranes as a result of

hydrophobic interactions (Blackshear, 1993). Rosen et al (1990) located MARCKS at the membrane in discrete structures, suggesting that the protein associates with the membrane through a receptor. In murine macrophages MARCKS is found in a punctate pattern at the interface of the substratum within pseudopodia and filopodia, where it colocalises with vinculin (intro, section F. 11) and talin (intro, section F. 3) (Rosen et al, 1990). MARCKS disappears from this punctate localisation upon phosphorylation by PKC in these cells (Rosen et al, 1990). Subsequent dephosphorylation of MARCKS is accompanied by its reassociation with the membrane (Aderem, 1992a).

It has been demonstrated (Hartwig et al, 1992) that MARCKS binds to and cross-links F-actin in vitro. The binding of calmodulin or the phosphorylation of MARCKS results in its translocation from membrane bound actin structures, where it remains associated with actin filaments. However, since phosphorylated MARCKS binds to actin but does not cross-link it (Hartwig et al, 1992), these filaments may be less rigid. Filament cross-linking requires that MARCKS must have either two binding sites or dimerise. Thus, phosphorylation or calmodulin either inactivates an actin-binding site or dissociates the dimers (Aderem, 1992a).

The observations of membrane-release and loss of actin cross-linking activity of MARCKS upon PKC phosphorylation, and the inhibition of cross-linking activity due to calmodulin-binding, clearly indicate methods of regulation of MARCKS interaction with both the actin cytoskeleton and the plasma membrane. This regulation could modulate the ability of MARCKS to serve as a cross-bridge between substrate-adherent plasma membrane and membrane cytoskeletal actin during chemotaxis (Hartwig et al, 1992; see also Aderem, 1992a; Blackshear, 1993 for reviews).

Immunoblot analysis has revealed that MARCKS is down-regulated in V-Src transformed 3T3 fibroblasts (Joseph et al, 1992). Since a change in cell shape is very often characteristic of transformation (Greig et al, 1985) MARCKS may be important in maintaining a non-transformed phenotype, with up-keep of a normal cell shape or in delivering anchorage dependent growth signals. It is possible, however, that MARCKS down-regulation during transformation is a consequence of shape change (Aderem, 1992a; Aderem, 1992b).

A MARCKS relative has been cloned from a mouse cerebella library and termed **F52** (also known as **MacMARCK** or **MRP**) (Umekage & Kato, 1991). F52 is a 20.1KDa polypeptide that is homologous to MARCKS in the three evolutionary conserved domains. Its properties are also very similar to that of MARCKS, it is: myristoylated, alanine-rich, calmodulin-binding, and a substrate for PKC. The presence of F52 and MARCKS identifies a gene family (see Blackshear, 1993).

15. SYNAPSINS

Synapsins are a family of closely related, neuronal specific, synaptic vesicle-associated phosphoproteins that have been implicated in the regulation of neurotransmitter release from nerve endings (Bahler et al, 1990). In mature neurons the synapsins are localised on the cytoplasmic side of synaptic vesicles and appear to be present in virtually all nerve terminals irrespective of the neurotransmitter released (reviewed by Valtorta et al, 1992a).

The synapsin family includes **synapsin I** and **synapsin II** which are encoded by two distinct genes (Sudhof et al, 1989).The **a** and **b** isoforms are derived by alternative splicing of the primary transcripts. In the rat, the mRNAs code for proteins of 704 (synapsin Ia), 668 (synapsin Ib), 586 (synapsin IIa), and 479 (synapsin IIb) amino acids which show differential distribution (Sudhof et al, 1989). The four proteins share a common, collagenase resistant, N-terminal region which is highly conserved and comprises more than half of each molecule, with the differences among the four proteins restricted to the C-terminal regions (Sudhof et al, 1989). The primary structure of the synapsins reveals that the common N-terminal region (78% identity between synapsins I and II) comprises of sites for cAMP-dependent protein kinase and CaM kinases (see Valtorta et al, 1992a; DeCamilli et al, 1990). An extra N-terminal region, present only in synapsin I, contains three phosphorylation sites. Analysis of fragments of synapsin I reveals the presence of an actin-binding site located in the N-terminus (Bahler et al, 1989), the presence of a second binding site has been postulated for its extra N-terminal region, which might explain its bundling effect (Bahler et al, 1989).

Synapsins are specifically associated with the membrane of small synaptic vesicles (Bahler et al, 1990). Synaptic vesicles exhibit a high affinity binding for the dephosphorylated form of both synapsin Ia and Ib. It has been reported that

microinjection of the dephosphorylated form of synapsin I inhibits the release of neurotransmitter (Llinas et al, 1985). Phosphorylation of synapsin I causes a decrease in the affinity of synaptic vesicle binding (see Valtorta et al, 1992a), and abolishes the inhibition of neurotransmitter.

Synapsin I interacts in vitro with various components of the cytoskeleton, namely F-actin, microtubules, neurofilaments and spectrin (see Valtorta et al, 1992a for review). Interaction of synapsin I with F-actin in vitro induces the formation of thick bundles of actin filaments, an effect which is strongly reduced after its phosphorylation by CaM kinase II (Valtorta et al, 1992b). Synapsin I has also been shown to increase the initial rate of actin polymerisation and to induce the rapid growth of a higher number of shorter actin filaments. These effects are markedly reduced upon phosphorylation of synapsin I (Valtorta et al, 1992b).

The involvement of synapsin I in the regulation of neurotransmitter release and its ability to bind to synaptic vesicles and to actin suggest a model in which synapsin I reversibly cross-links synaptic vesicles to the cytoplasmic network of the nerve terminal, thereby regulating the availability of synaptic vesicles for exocytosis (Bahler et al, 1990).

16. ANNEXINS

Annexins (previously called **lipocortins** or **calpactins**) are widely distributed phospholipid- and membrane-binding proteins. They were first identified and isolated by their Ca^{2+}-dependent binding to cellular membranes and their ability to promote membrane aggregation (see Crompton et al, 1988; Glenney et al, 1987 for review). The Ca^{2+}-dependent binding to membranes of some of these proteins is resistant to treatment with nonionic detergents, suggesting that they may interact with the cytoskeleton (Crompton et al, 1988).

The sequence organisation of the annexin family is highly conserved, and indicates there are 8 unique gene products (**I-VIII**) (see Smith et al, 1990 for review), each composed of a domain repeat of 60-70 amino acids. Annexins I-V, VII-VIII are composed of 4 of these domains, whereas annexin VI is comprised of 8 domains. The sequence conservation for each domain ranges between 40% and 60% when individual family members are compared (Smith et al, 1990). The N-terminus of each protein is unique suggesting that this region may confer functional differences (Crompton et al,

1988; Crumpton & Dedman, 1990; Smith et al, 1990). Although the sequences of these proteins are known, their specified cellular role remains undefined (Crumpton & Dedman, 1990).

The distinguishing features of the annexin family are their ability to bind phospholipids and the ability to bind Ca^{2+} without using the classical EF-hand motif (Smith et al, 1990). Both Ca^{2+} and phospholipids may modulate the function of these proteins and provide an independent mechanism for the mediation of intracellular calcium (Smith et al, 1990).

Annexin VI (**P68**) has been studied in detail (Diaz-Nunoz et al, 1990; Hosoya et al, 1992). Purified from bovine liver, annexin VI is a 68KDa, globular, phospholipid- and F-actin binding protein that showed a higher affinity for F-actin in the presence of Ca^{2+} than in its absence, with no F-actin severing or capping activity noted (Hosoya et al, 1992). Immunocytochemical studies show annexin VI localised along stress fibres as well as membrane ruffles, microspikes and focal contacts, raising the possibility that annexin VI may contribute to control membrane-microfilament interaction in the cell (Hosoya et al, 1992). The Ca^{2+}-dependent membrane binding of annexin I and II, and their in vitro association with actin suggests that they could be involved in the control of secretion and cell motility (Glenney et al, 1987).

17. SCRUIN FAMILY

Scruin, a 102KDa actin bundling protein, is found in a 1:1 complex in the acrosomal process of *Limulus* sperm (Schmid et al, 1991). The scruin molecule is organised into two domains which contact two actin subunits in different strands of the same actin filament. Protein sequence reveals that scruin is unrelated to any of the other known actin-binding protein families (Schmid et al, 1994), but reveals homology within each domain with *Drosophila* ***kelch*** (Xue & Cooley, 1993), and with the mouse gene ***MIPP*** (Chang-Yeh et al, 1991). Their similarity with scruin may suggest cytoskeletal function and formation of a family of actin-cross-linking proteins (binding site not identified). Schmid et al (1994) show that scruin shares a binding site on actin (subdomain 3) with tropomyosin and contacts with subdomain 1, a region that has not been described as a binding site for any other actin-binding protein.

Work by Xue and Cooley (1993) reveal the ***kelch*** gene encodes two large ORFs (1 and 2) separated by a single UGA stop codon. Monoclonal antibody against kelch ORF 1 peptide (688 amino acids) show that two kelch proteins are produced in the ovary. Peptide from ORF 1 is the most prominent kelch protein, while the full length protein (ORF 1 + ORF 2) is also detected suggesting that the UGA stop codon separating ORF 1 and ORF 2 is partially suppressed during mRNA translation. Together with actin, kelch is localised at the cellular structures called ring canals, suggesting a direct interaction of kelch with the cytoskeleton (Xue & Cooley, 1993).

MIPP gene product is of unknown function, isolated from a 8.5 day old mouse embryo cDNA library, encodes a 202 amino acid polypeptide with four, 48 amino acid, repeat units (Chang-Yeh et al, 1991). The 202 amino acids of MIPP has 31% identity with the ORF 1 peptide of *Drosophila kelch.*

18. DROSOPHILA ZYGOTIC CELLULARISATION GENES

A dynamic network of cortical microfilaments is associated with the cleavage furrow membranes during cellularisation of the *Drosophila* embryo (transition from syncytium to a cellular blastoderm). The gene products of the loci ***Bottleneck, nullo*** and ***serendipity-*** α have been described, mutations in which, disturb the sequence of rearrangements of this microfilament network (see Theurkauf, 1994 for review). All three gene products accumulate during late syncytial divisions and reach a peak just before cellularisation, it is therefore possible that cellularisation is triggered by these gene products (reviewed in Theurkauf, 1994).

Derived amino acid sequences encoded by *Bottleneck, nullo* and *serendipity-* α are not related to each other or any other identified proteins (Schejter & Wieschaus, 1993; Simpson & Wieschaus, 1990). However, by use of specific monoclonal antibodies, all zygotic cellularisation proteins colocalise with the actin cytoskeleton, suggesting they are novel components of the cytoskeleton (Schejter & Wieschaus, 1993) and could represent three new actin-associated groups.

19. UNCLASSIFIED ACTIN-ASSOCIATED PROTEINS

The following proteins have been identified as actin-associated proteins. Their

ABDs have either not been determined, or they appear to have no homology with any other ABPs. These unclassified proteins are presented below, alphabetically.

ABP30 is a monomeric actin filament-bundling protein, whose actin-binding activity is inhibited in the presence of micromolar calcium (see Otto, 1994 for review). It is present in mammalian cells in culture, and *Dictyostelium* ABP30 has been localised to pseudopodia and filopodia of motile cells and to cleavage furrow during cytokinesis, where it contributes to changes in the cortical architecture associated with cell movement and cytokinesis (Fechheimer et al, 1991; Otto, 1994). ABP30 cDNA encodes a protein of 295 residues (33.4KDa) and reveals two EF hand structures that provides a structural explanation for calcium regulation. No strong homology to previously identified ABPs is apparent in ABP30 (Fechheimer et al, 1991). Zigmond et al (1992) demonstrated that ABP30 inhibits depolymerisation at both ends of the actin filament in a dose dependent manner, suggesting that it may cause local accumulation of actin filaments. It has also been reported (Fechheimer & Furukawa, 1993) that a 27KDa proteolytic fragment of ABP30 can cross-link actin filaments into a meshwork, indicating that it contains two ABDs.

ABP50 is a 50KDa polypeptide isolated from *Dictyostelium* amoebae, which bundles actin filaments in vitro (Demma et al, 1990). Distribution of ABP50 is diffuse (99% bound to G-actin) in resting cells. Upon stimulation by cAMP, ABP50 becomes localised with F-actin to newly formed surface projections such as filopodia (Dharmawardhane et al, 1991). Sequencing of ABP50 cDNA from *Dictyostelium* reveals that ABP50 is **elongation factor 1α (EF1α)**, with 76% identity between them (Yang et al, 1990). This defines an element of the protein synthetic apparatus that binds to the actin cytoskeleton (see Edwards, 1993 for review). The binding of EF1α to actin may affect the activity or availability of it in protein synthesis, with binding to G- of F-actin having different effects (Yang et al, 1990). The reversible association of EF1α with actin may provide a cytoskeletal mechanism for temporal and spatial regulation of protein synthesis in eukaryotic cells (see Fulton, 1993; Singer, 1992 for reviews).

ABP52, a 52KDa, F-actin-bundling protein isolated from *Physarum* (Itano & Hatano, 1991) can induce actin filaments and microtubules to co-bundle in a Ca^{2+}-

insensitive manner (Itano & Hatano, 1991). ABP52 cross reacts with monoclonal antibodies raised against HeLa p55, of the F-actin bundling fascin group (intro, section F. 10) (Yamashiro-Matsumura & Matsumura, 1985) and may be classed within this group once it is sequenced. It is proposed that ABP52 may function as an F-actin-bundling protein in *Physarum* plasmodium during interphase, and may associate with microtubules in the mitotic spindle during mitosis (Itano & Hatano, 1991).

Actobindin is a 9.7KDa protein purified from *Acanthamoeba* (Bubb et al, 1991). The entire 88 amino acid sequence of actobindin is most notable for a nearly identical repeated sequence of approximately 33 residues (Vandekerckhove et al, 1990b). Depending on the molar ratios in which they are mixed, actobindin binds to either one or two G-actin molecules with a slight degree of negative cooperativity, the second actin monomer binding 5-8 times weaker than the first (Bubb et al, 1991). The repeat structure of actobindin could easily explain the bivalent affinity for actin. Both repeat regions contain a sequence of similarity to an 11 residue peptide in profilin (intro, section F. 13), and a hexapeptide common to several other actin-binding proteins (see results section, fig 14) (Vandekerckhove et al, 1990b). Lysine residues within this hexapeptide have been shown to bind with the actin N-terminus (Vancompernolle et al, 1991), and mutation of the residues within this sequence is known to alter the functional interaction of actobindin and actin (Vancompernolle et al, 1992). A protein such as actobindin could decrease the number of functional nuclei available for polymerisation of monomeric actin de novo, while having little effect on the rate of elongation of existing filaments (see Bubb et al, 1991).

Actolinkin isolated from sea urchin eggs and starfish oocytes is a monomeric protein of 20KDa on SDS-PAGE. The majority of actolinkin localises exclusively in the cortical fraction of the sea urchin egg and is considered to exist in a stable 1:1 complex with actin to which it binds an N-terminal segment. Both free actolinkin and actolinkin-actin complex bind to the barbed ends of the actin filament and block actin monomer addition at that end (see Kreis & Vale, 1993 for review).

Aginactin is a 70KDa, agonist-regulated, Ca^{2+}-insensitive, barbed end capping protein isolated from *Dictyostelium* (Sauterer et al, 1991). It may regulate the changes in actin nucleation activity that occur following chemotactic stimulation by inhibiting actin

polymerisation. Aginactin has been demonstrated to bind and cosediment F-actin and it also inhibits the depolymerisation and polymerisation of F-actin at the barbed end of *Limulus* acrosomal bundles, defining it as a capping protein (Sauterer et al, 1991). Aginactin co-purifies and shares 73% identity with the 70KDa heat shock protein, **Hsc70** (Eddy et al, 1993).

It is known that the glycolytic enzyme **aldolase** can interact with F-actin, it being localised within the I-band of skeletal muscle myofibrils (see Tellam et al, 1989 for review). It is capable of organising actin filaments into highly ordered three-dimensional structures indicating that the enzyme/actin interaction may have a functional duality, one catalytic and the other structural (Tellam et al, 1989; O'Reilly & Clarke, 1993). An ABD has been identified within the aldolase sequence (O'Reilly & Clarke, 1993) that is highly homologous to a region near the C-terminus of actin itself and is also found in the ABDs of a number of other actin binding proteins (Tellam et al, 1989).

Ankyrin, the major membrane-skeleton linker of erythrocytes, connects the cytoplasmic domain of band 3 with the C-terminus of β-spectrin (intro, section F. 1. 3) (Hitt & Luna, 1994; Bretscher, 1991). It is a globular 205KDa monomer that contains three major structural domains (see Lambert & Bennett, 1993 for review), a central domain binds spectrin and vimentin. A large number of ankyrin-like proteins are present in erythrocytes and other cell types. These isoforms arise from several distinct genes and a complexity of alternatively spliced exons (see Hitt & Luna, 1994 for review). Ankyrins that have been resolved at the level of primary structure include products of the **ANK1** and **2** genes, expressed in erythrocytes and brain (Davis et al, 1992; Hitt & Luna, 1994). Most of the variability in the predicted ankyrin-like proteins affects the allosteric control of ankyrin-binding to both band 3 and spectrin (Davis et al, 1992), the different isoforms modulating different interactions.

ASP56, a 56 KDa actin-binding protein, isolated from pig platelet lysate, sequesters actin in a non-polymerisable 1:1 ASP56: actin complex (Gieselmann & Mann, 1992). Tryptic digestion and microsequencing of ASP56 reveals 62% sequence identity with yeast **CAP70** (Gieselmann & Mann, 1992). It is therefore possible that pig ASP56 is acting on the adenylate cyclase pathway, whilst the yeast homologue, discovered as a

adenylate cyclase-binding protein may also act on the yeast cytoskeleton, by binding actin (Gieselmann & Mann, 1992).

C-Abl is a non-receptor PTK, whose activated forms, typified by the viral transforming proteins **p160$^{gag/v\text{-}abl}$** and **p210$^{bcr/abl}$**, are localised to the cytoplasm, plasma membrane, focal adhesions and microfilaments of transformed cells (Vann Etten et al, 1989; McWhirter & Wang, 1991; 1993). The normal c-Abl protein is present in two forms, the myristoylated form localised to the nucleus, plasma membrane, cytoplasm and stress fibres in 3T3 cells (Vann Etten et al, 1989). Analysis of deletion mutants of the myristoylated form of c-Abl maps the domain responsible for Ca^{2+}-independent F-actin interaction to its C-terminus (Vann Etten et al, 1994). In addition to the F-actin binding domain, the C-terminus of c-Abl contains a proline-rich region that mediates binding and sequestration of G-actin; the F- and G-actin-binding sites cooperate to bundle F-actin (Vann Etten et al, 1994). With c-Abl being a PTK, it is possible actin itself or another associated protein is an important target for phosphorylation by it (Vann Etten et al, 1994). Also, as c-Abl contains an NLS and a DNA-binding domain (Kipreos & Wang, 1992), it has the potential to transfer signals from the cell surface and the cytoskeleton directly to the nucleus.

Coactosin is a 17KDa protein purified from actin-myosin complexes of *Dictyostelium* . The protein binds to F-actin without significantly altering its viscosity (de Hostos et al, 1993). Coactosin cDNA encodes a polypeptide of 146 amino acids (de Hostos et al, 1993) that shares 26% identity to a 143 residue N-terminal region of yeast **ABP65** (Drubin et al, 1990), a protein of unknown function. Coactosin has 24% identity over 128 amino acids with **drebin A** (Kojima et al, 1988), a protein implicated in the extension of neuronal processes and in maintaining the plasticity of dendrites. There is also 24% overall identity between coactosin and vertebrate cofilin (intro, section F. 4), but does not share homology with cofilin's ABD (Yonezawa et al, 1991). Coactosin may provide a link between the cofilin family and neuronal drebrins (Yonezawa et al, 1991).

Comitin, identified in *Dictyostelium* (Noegal et al, 1990), is a membrane-associated protein which binds to G- and F-actin (Weiner et al, 1993). Comitin is present in *Dictyostelium* on (i) perinuclear structures with tubular or fibrillary extensions, and (ii) on vesicles distributed throughout the cell. In 3T3 cells, comitin has a similar staining

pattern, with the perinuclear structure identified as the golgi apparatus (Weiner et al, 1993). Comitin *Dictyostelium* cDNA encodes a 185 residue, 20.7KDa polypeptide (Noegal et al, 1990), that shows no homology to any other ABP, nor the presence of a transmembrane hydrophobic domain. As comitin is present on the cytoplasmic surface of the vesicle or golgi membranes, it could form a stabilising scaffold linking vesicles and golgi to the microfilament system; comitin may therefore be responsible for the correct location of these structures (Weiner et al, 1993), or part of an actin-based microfilament system involved in intracellular vesicle transport, as described by Kuznetsov et al (1992).

Coronin, a 55KDa protein that cosediments with actin filaments, is so named because of its association with crown-shaped, F-actin-rich, cell surface projections, during *Dictyostelium* growth phase (de Hostos et al, 1991). Coronin null mutant cells (cor^-) have normal filament localisation, but grow and migrate more slowly than wild-type cells (de Hostos et al, 1993). Cor^- cells become multinucleate in liquid medium, indicating a role in cytokinesis, a role reinforced by the observation that coronin accumulates at the distal portions and leading edges during separation of daughter cells (de Hostos et al, 1993). In aggregating *Dictyostelium* cells (which are more sensitive to the chemoattractant cAMP), coronin is accumulated at the front where surface projections are directed towards a cAMP source and is rapidly reshuffled within cells during motion and chemotactic orientation (de Hostos et al, 1991). *Dictyostelium* cDNA encodes a protein of 445 amino acids that shares no homology with any ABPs. The coronin N-terminal domain has similarity to the β-subunits of G-proteins, suggesting a role in the transmission of chemotactic signals from cAMP receptors, through G-proteins, to the cortical cytoskeleton (de Hostos et al, 1991).

Cortactin, an 80/85 KDa protein, is present in multiple cell types and enriched in cortical structures (Wu & Parsons, 1993). It binds specifically to polymerised actin and is phosphorylated in transformed and growth factor stimulated cells (Wu & Parsons, 1993). Cortactin cDNA from chick embryos identifies its N-terminal half to be comprised of multiple copies of a repeat sequence (five copies in the p80 isoform, with six in the p85 isoform), that is necessary to mediate actin-binding activity (Wu & Parsons, 1993). The C-terminal SH3 domain of cortactin may serve as a membrane localisation signal (Wu & Parsons, 1993).

DNase1 from bovine pancreas, an endonuclease of Mr 30.4KDa that degrades double stranded DNA (see Oefner & Suck, 1986), was one of the first nonmuscle ABPs identified (see Stossel et al, 1985 for review). It forms a tight complex with monomeric actin, and can therefore depolymerise a stoichiometric amount of actin, with actin-binding inhibiting its enzymatic effect. DNase1 can also bind with high affinity to the barbed end of F-actin in a Ca^{2+}-insensitive manner (Stossel et al, 1985). The binding of DNase1 to actin retards the exchange of nucleotide bound to the actin subunit, the extent of which depends on whether it is added to G- or F-actin (Stossel et al, 1985). DNase1 has been identified as a key enzyme involved in apoptosis (Peitsh et al, 1993). Dephosphorylation of cofilin, a protein that can depolymerise actin under certain conditions (intro, section F. 4), results in the nuclear accumulation of actin/cofilin complexes (Samstag et al, 1994). This complex may exert an anti-apoptotic activity, with the resultant G-actin formed in the nucleus, inhibiting DNAase1 degradation activity (Samstag et al, 1994).

The epidermal growth factor (EGF) receptor (**EGFR**) is a transmembrane glycoprotein of 170KDa with an extracellular EGF binding domain, a single hydrophobic membrane spanning stretch, and an intracellular tyrosine kinase domain (Ullrich et al, 1985) that has been shown to bind the actin filament system, thus representing, an example of receptor modulation of the cytoskeleton (den Hartigh et al, 1992).

Growth-associated protein (GAP-43) (reviewed by Moss et al, 1990), is a membrane-associated, nervous tissue-specific protein kinase C substrate, involved in the regulation of the dynamics of the submembranous cytoskeleton (Moss et al, 1990). Various assay systems have shown that both the phosphorylated and dephosphorylated forms of GAP-43 cosediment with F-actin, in a Ca^{2+}-independent manner, with no severing or capping activity (Hens et al, 1993). It remains to be established whether in growth cone and mature nerve terminals, membrane-bound GAP-43 serves as an anchor protein for actin filaments.

Hisactophilin, a 17KDa unique protein from *Dictyostelium* amoeboid cells (Scheel et al, 1989), is present in the submembranous region throughout growth and all stages of development. Hisactophilin binds to F-actin in a pH-dependent and saturable manner (Scheel et al, 1989). Hisactophilin cDNA encodes a polypeptide of 118 amino

acids, 31 of which are histidine residues (Scheel et al, 1989). 3D analysis has shown that 90% of all the histidine residues are located in the loops and turns at the surface of the molecule (Habazetti et al, 1992). The histidine content and location seem to be the basis of the pH-dependent actin-binding activity of hisactophilin that acts as an intracellular pH-sensor that links chemotactic signals to responses in the microfilament system (Scheel et al, 1989).

The **25KDa inhibitor of actin polymerisation (25KDa IAP)** is mainly detected in muscular tissue, and is hardly detectable in fibroblasts (Miron et al, 1988). The expression of 25KDa IAP increases dramatically (≈15-fold) in chick embryo fibroblasts after heat shock treatment at 45°c (Miron et al, 1991). 25KDa IAP inhibits actin polymerisation, possibly acting as a non-nucleating, barbed end capping protein (Miron et al, 1991). It also induces disassembling of F-actin cross-linked to α-actinin (Miron et al, 1988). 25KDa IAP cDNA encodes a 193 residue polypeptide that shares a high degree of homology (67% identity; 80% similarity) to human 27KDa heat shock protein, **HSP27** (Miron et al, 1991). HSP27 is rapidly phosphorylated after growth factor stimulation; its overexpression stimulates membrane ruffling and pinocytosis. A non-phosphorylatable mutant inhibits these processes in response to growth factors (Lavoie et al, 1993).

LSP1 is a lymphocyte specific intracellular protein, expressed in normal B and T cells and in transformed B cells but not (or in much smaller amounts) in T lymphoma lines (Jongstra et al, 1988). LSP1 derived amino acid sequence reveals a 330 amino acid polypeptide with two putative Ca^{2+}-binding domains (Jongstra et al, 1988). LSP1 binds along the sides of F-actin through its C-terminal domain, but does not bind G-actin (Jongstra-Bilen et al, 1992). There is 34% identity between the 20KDa C-terminal actin binding region of chick smooth muscle caldesmon (Bryan et al, 1989; Wang et al, 1991) and the C-terminal region of LSP1. However, unlike caldesmon, LSP1 binding to F-actin is Ca^{2+}-insensitive (Jongstra-Bilen et al, 1992).

Microtubule associated protein 2 (MAP2), is found specifically in neuronal cells, binding microtubules and promoting their assembly from purified tubulin (Riederer & Matus, 1985). Pollard et al (1984) have described the presence of two ABDs on MAP2, one located on, or close to, the tubulin-binding domain. Sattilaro (1986) has reported that actin binds to a 32KDa fragment of MAP2 which also contains the tubulin-

binding site. The tubulin-binding domain of MAP2 is composed of three repeated sequences (Lewis et al, 1988) that have been identified in tau protein (see below). As this sequence is involved in the in the binding of tubulin and actin, it could suggest that actin and tubulin share a similar sequence responsible for the binding of MAP2 and tau proteins (ie. residues 433-440 of tubulin and 362-368 of actin show a certain identity) (Correas et al, 1990).

The kinase responsible for phosphorylation of the two 20KDa light chains of myosin is **myosin light chain kinase (MLCK)**. Gizzard full length MLCK cDNA (Olson et al, 1990) encodes a 972 residue (108KDa) protein that has been demonstrated to bind to stress fibres in nonmuscle cells and the I band of skeletal and cardiac muscle (Guerriero et al, 1981). The ABD for smooth muscle MLCK has been identified in the MLCK of fibroblasts, but not in rabbit and rat skeletal muscle MLCK (Kanoh et al, 1993). MLCK may stay attached to thin filaments during muscle contraction, but it is possible that it attaches and detaches from the thin filament during different phases of the contractile cycle, detachment coinciding with an increase of intracellular [Ca^{2+}] and calmodulin binding to MLCK (Kanoh et al, 1993).

Nebulin is a family of giant actin-binding proteins with mass ranging from 600-900KDa found in most skeletal sarcomeres (Kruger et al, 1991). Nebulin is widely believed to constitute the set of long inextensible longitudinal filaments which span the space between the Z-disc and the distal end of the thin filament (reviewed by Tatsumi et al, 1993). Human nebulin cDNA encodes a polypeptide comprised, almost entirely, of a recurring motif of 35 residues (Labeit et al, 1991). Nebulin filaments are split into five subfragments on the treatment of myofibrils with 0.1mM $CaCl_2$ (Tatsumi et al, 1993), with all but one fragment released from the myofibril. Prior to treatment, all subfragments possess the ability to bind actin in vitro with high affinity (Jin & Wang, 1991). Nebulin appears to serve as a length determining protein scaffold in skeletal muscle, since the characteristic length of thin filaments is proportional to the size of expressed nebulin isoform (Kruger et al, 1991).

Nuclear actin binding protein (NAB-34) is a dimeric protein of 34KDa purified from *Acanthamoeba* (Rimm & Pollard, 1989). Polyclonal antisera against the purified protein show almost exclusive nuclear location in *Acanthamoeba* , skeletal muscle, glial

and retinal epithelial cells (Rimm & Pollard, 1989). It binds F-actin in an ATP-insensitive manner without cross-linking, severing or capping (Rimm & Pollard, 1989). Southwestern analysis reveal NAB-34 binds to DNA, indicating it may play a role in attaching DNA to an actin containing matrix (Rimm & Pollard, 1989).

P-39, a 39KDa membrane-bound protein from bovine adrenal medulla, is found to bind to F-actin with a stoichiometry of nearly one mole of P-39 to two moles of actin on the filament (Maekawa, 1992). P-39 is localised to the cortical regions of bovine adrenal medullary chromaffin cells and rat pheochromacytoma cells (PC12). It is suggested by Maekawa (1992) that P-39 being localised to the cell membrane, could be an anchoring protein of the actin filament.

P-58 is a 58KDa elongated, calcium-insensitive, monomeric protein found in microvilli and actin-containing transmembrane complexes of ascites tumor cells (Liu et al, 1989). It binds to F-actin and phospholipids in pelleting assays and inhibits polymerisation of pyrenyl-actin (indicating capping activity). The microfilament and phospholipid-binding properties of P-58 suggests role in stabilising membrane-microfilament interactions in the ascites cell microvilli (Liu et al, 1989).

Paxillin, a 68KDa cytoskeletal protein is localised to focal adhesions of cultured cells (Turner et al, 1990), where it may be involved in actin membrane attachment via interactions with vinculin (Turner et al, 1990). Immunoblot analysis reveals paxillin to be most abundant in smooth muscle (Turner et al, 1990), where it is localised to dense plaques (Turner et al, 1991). Paxillin is heavily phosphorylated by $pp60^{src}$ in chick embryo fibroblasts transformed by RSV (Glenney & Zokas, 1989) which may reflect the major reorganisation of the actin cytoskeleton and disassembly of cell-ECM linkages observed upon transformation (Burridge, 1986). It is also the major tyrosine kinase substrate during chick embryonic organogenesis, its phosphorylation becoming reduced upon formation of stable actin-membrane interactions (Turner, 1991). Phosphorylation of $p125^{FAK}$, a focal adhesion kinase, and paxillin is coordinately regulated in intact cells (Zachary et al, 1992), with purified preparations of paxillin tyrosine-phosphorylated by $p125^{FAK}$ (Turner et al, 1993). As $p125^{FAK}$ is phosphorylated by PDGF receptor (Rankin & Rozengurt, 1994), it seems likely that paxillin is a direct substrate of $p120^{FAK}$ upon PDGF stimulation.

Ponticulin is an actin-binding integral glycoprotein in the plasma membrane of *Dictyostelium* (Wuestenhube & Luna, 1987). Purified *Dictyostelium* ponticulin consists of six major 17KDa isoforms, five isoforms of 19KDa, and two isoforms of 15KDa which are all isolated on F-actin affinity columns (Wuestenhube & Luna, 1987). Because all of these polypeptides bind F-actin directly (Chia et al, 1991; Wuestenhube & Luna, 1987) they may form a related family. Ponticulin appears to be responsible for most of the basal actin-binding activity of *Dictyostelium* plasma membranes (Wuestenhube et al, 1989; 1991). Ponticulin, in addition to its identification as an actin-membrane anchor, has been identified, in the correct lipid matrix, to nucleate actin filaments at the plasma membrane during motile processes (Chia et al, 1993).

pp60$^{c\text{-}src}$, the 60KDa product of the c-src gene, is a protein tyrosine kinase (PTK), which is highly conserved throughout the vertebrate kingdom. All src family genes encode PTKs of about 525 residues that contain a membrane-association domain (it is myristoylated), regulatory SH3 and SH2 domains, and a C-terminal protein kinase catalytic domain (Koch et al, 1991). Much of the transforming counterpart of pp60$^{c\text{-}src}$ (**pp60$^{v\text{-}src}$**) is known to be associated with cytoskeletal elements (Hamaguchi & Hanafusa, 1987; Kellie et al, 1991). In non-transformed cells, phosphotyrosine has been principally localised in the cell-to-cell contact regions (Tanaka & Singer, 1988). Tsukita et al (1991) have found pp60$^{c\text{-}src}$ at the cell-to-cell contact areas in adult rat hepatocytes. During platelet aggregation (a form of cell-to-cell contact through GP IIb-IIIa) pp60$^{c\text{-}src}$ is associated with the membrane cytoskeleton (Oda et al, 1992). Association of pp60$^{c\text{-}src}$ with the cytoskeleton (possibly with actin, talin, vinculin and clathrin) suggests a regulatory role in tyrosine phosphorylation (Oda et al, 1992).

Rap2B is a Ras-related, low molecular weight GTP-binding protein from platelets (Ohmstede et al, 1990). Agonist-induced actin polymerisation in platelets is necessary for the translocation of Rap2B to the cytoskeleton, suggesting that Rap2B interacts with newly formed actin filaments (Torti et al, 1994). Rap2B associates with activated platelet cytoskeleton only when a functional GPIIb-IIIa (platelet fibrinogen receptor) is present on the plasma membrane (Torti et al, 1994), and that translocation of Rap2B to the cytoskeleton is paralleled by the translocation of comparable amounts of

GPIIb-IIIa. When neither Rap2B nor GPIIb-IIIa associate with the actin network, platelets cannot enter the irreversible phase of aggregation (Torti et al, 1994).

Tau, a family of MAPs, consists of four to six discrete isoforms ranging in apparent Mr 36-65KDa. All the isoforms contain a cluster of three or four imperfect amino acid repeat domains found within the C-terminal half of tau (Lee et al, 1988). Study has shown the existence of interaction of a specific tubulin-binding domain within tau with actin in vitro. The interaction is attributed to the first of these repeats, which induces the increase in bundling of actin filaments (Moraga et al, 1993). Data suggests that the tau repetitive sequence (also found in MAP2; see above) may constitute a functional regulatory domain of the microtubule-actin filament cytoskeleton, as it appears operationally connected to their functions of polymerisation and aggregation (Moraga et al, 1993; Correas et al, 1990).

Tenuin, a 400KDa protein, has been identified in various cells and tissue such as fibroblasts and skeletal muscle (Tsukita et al, 1989). The slender rod-shaped tenuin molecule does not bind actin filaments in vitro, however, immunofluorescence and EM studies have revealed that it is distributed in the undercoat of AJs, on stress fibres, and on the circumferential bundles in epithelial cells (Tsukita et al, 1989). The distribution of tenuin indicates a crossliking or filament strengthening role at focal adhesions (see Luna & Hitt, 1992 for review).

Titin (connectin) is a family of 2000-3500KDa polypeptides (Fulton & Isaacs, 1991) that spans Z-line to the M-line in adult sarcomeres (Furst et al, 1988). The expression of titin size variants or isoforms appears to be developmental-, tissue-, and species-dependent (Fulton & Isaacs, 1991; Wang et al, 1991). On SDS gels, titin appears as either a single band (T1: 2800-3200KDa) or a doublet (T1 and T2: 2100-2400 KDa) (Wang et al, 1991). T2 may be a proteolytic product of the longer T1 polypeptide. Titin binds to, and is rendered inextensible by thick filaments along most of its length (Furst et al, 1988), forming an elastic filamentous matrix in the sarcomere that provides structural continuity and an elastic restoring force (Fulton & Isaacs, 1991). In developing skeletal and cardiac muscle, titin participates in myofibrilogenesis by serving as a template or scaffold in the assembly of individual A-bands and I-Z-I bundles (see Fulton & Isaacs, 1991). *Drosophila* **projectin** (Ayme-Southgate et al, 1991) and the *C. elegans* gene,

unc22, product called **twitchin** (Benian et al, 1989) have been found to share sequence motifs, secondary structure, filamentous morphology and physical properties of partially sequenced vertebrate titin (Labeit et al, 1990). A novel, 4.0Kb sequence for titin has recently been obtained and is localised at the I-band of sarcomeres (Maruyama et al, 1993).

VASP (vasodilator-stimulated phosphoprotein), a 46-50KDa F-actin binding-protein, is phosphorylated by cGMP-dependent and cAMP-dependent protein kinases in intact human platelets (Reinhard et al, 1992). In human platelets spread on glass, VASP is associated predominantly with the distal parts of radial microfilament bundles and with microfilaments outlining the periphery (Reinhard et al, 1992). VASP is also detectable in cells such as fibroblasts, where it is concentrated at focal contact areas, and is phosphorylated in the same manner as in platelets (Reinhard et al, 1992; see Luna & Hitt, 1992 for review).

Vitamin D binding/Gc protein (DBP/Gc) is mainly a mammalian serum protein, but is also associated with the membrane of certain cells (Petrini et al, 1983). The rat gene spans 35Kb, containing 13 exons, and encodes a 58KDa protein (Ray et al, 1991). Monomeric actin binds to DBP/Gc with a higher affinity than for the elongating actin filament (McLeod et al, 1989). Therefore DBP/Gc is capable of sequestering monomeric actin away from polymerisation. This has been proposed to constitute an effective mechanism whereby actin, released to the circulation during cell destruction, could be prevented from polymerisation (Janmey et al, 1986). DBP/Gc binds to proteolytic fragments containing the C-terminal end of actin and competes with profilin for actin-binding, suggesting that this region corresponds to the binding-site for DBP/Gc (Goldschmidt-Clermont et al, 1986).

Zyxin, an 82KDa AJ component binds along actin stress fibres near where they terminate at adhesion plaques, suggesting that it may play some role in the regulation or organisation of membrane-cytoskeletal attachments (Crawford & Beckerle, 1991). Protein interactions at focal adhesions appear very complex (see Luna & Hitt, 1992 for review), and zyxin's association with α-actinin probably localises it to particular subcellular domains such as the ends of actin filaments (Crawford et al, 1992). Zyxin cDNA encodes a 542 residue polypeptide whose N-terminal region is proline-rich

(>35%), and C-terminal domain contains three tandemly arranged copies of a zinc-binding, LIM domain (Sadler et al, 1992). Zyxin and the 23KDa protein, **cCRP**, interact via their LIM domains (Crawford et al, 1994; Sadler et al, 1992). The demonstration that zyxin and cCRP share a sequence motif with proteins important for transcription (see Crawford et al, 1994 for list) raises the possibility that they are components of a signal transduction pathway that mediates adhesion-stimulated changes in gene expression.

20. TRANSGELIN-RELATED

A molecule of 22KDa termed **protein C4** has been found to be present as a detergent soluble doublet ($C4^{h/l}$) associated with actin filaments in all cells and tissues apart from neurones, erythrocytes and skeletal muscle (Shapland et al, 1988). The higher relative Mr polypeptide, $C4^{h}$, now known as **transgelin** (Shapland et al, 1993), is absent in both transformed mesenchymal cells where actin stress fibres are reduced in number, and in non-adherent cells such as lymphocytes (Shapland et al, 1988). The lower relative Mr polypeptide, $\mathbf{C4^{l}}$, is always present (Shapland et al, 1988; 1993; this study).

Transgelin, purified from sheep aorta (Shapland et al, 1993): (i) binds directly to actin filaments in vitro at a saturable ratio of 1:6 actin monomers, with a binding constant of $\approx 7.5 \times 10^{5} M^{-1}$, addition of transgelin at a concentration of 2.4μM results in the complete gelation of 9.3μM of actin; (ii) causes actin filament gelation within two minutes in a pH and Ca^{2+} insensitive manner that is controlled by ionic strength; and (iii) binds actin filaments in a permeabilised cell system (Shapland et al, 1993). Viscometric and light scattering assays suggest either a bundling or cross-linking role for transgelin in the formation of actin gels, and EM studies show that transgelin rapidly induces the formation of an actin filament meshwork (Shapland et al, 1993).

Sequencing and cloning of rat small intestine transgelin cDNA encodes a polypeptide of 201 residues and a calculated Mr of 22.5KDa (Prinjha et al, 1994). Transgelin is: (i) present as one copy in the mammalian genome; (ii) conserved between yeast, *Drosophila*, molluscs and humans; and (iii) expressed as a single message with apparent size of 1.45kb (Prinjha et al, 1994). Transgelin message is not present in lymphocytes and transformed cells such as SV40 3T3 cells (Prinjha et al, 1994). Since transgelin is known to cross-link and gel actin, the loss of this molecule following oncogenic transformation may, at least in part, explain the cytoskeletal activation,

remodelling and cell migration that occur during metastasis (Shapland et al, 1993; Van Roy & Mareel, 1992).

The ability of transgelin to gel actin suggests that the molecule either: (i) forms dimers; or (ii) contains two actin-binding sites. The former has been excluded by use of sucrose gradient centrifugation (Shapland et al, 1993). Sequence analysis has identified the peptide LKAAED within transgelin that is similar in sequence to a putative actin-binding domain in a number of ABPs (Vancompernolle et al, 1991). A two-plus-three cluster of positively charged residues which may also be involved in binding of transgelin to actin has also been identified (Prinjha et al, 1994; Shapland et al, 1993).

A 21 residue motif within the C-terminus of transgelin (and all the members of the transgelin multigene; including 3x in α-calponin, see below) shares homology with a 21 residue region that is repeated seven times within an alternatively spliced product from the *C. elegans* **unc87** gene (unpublished; EMBL accession Nº U04711). unc87 is a cytoskeletal protein, as mutants in it cause an altered body wall, animals that are slow-moving, and thin filaments of actin that are bunched at the ends of cells and in cell processes (Waterston et al, 1980). The 21 residue motif found in the transgelin-related multigene family contains an RGASQAG hexapeptide (or similar) which contains a putative phosphorylation site that is also found in gelsolin (Kwiatkowski et al, 1986).

SM22 is a protein widely distributed in the smooth muscles of chicken and mammals, and appears to be absent in skeletal muscle (Lees-Miller et al, 1987). It is reported that there are three isoelectric variants of SM22 in bovine aorta, in a ratio of α:β:γ= 14:5:1 (Lees-Miller et al, 1987); however, others have not reported the presence of SM22 variants (Nishida et al, 1991; Prinjha et al, 1994). The complete amino acid sequence of SM22α has been determined (Pearlstone et al, 1987), but the function remains unknown. Although SM22α can be phosphorylated, the physiological significance is in question because it occurs at a ratio of only 0.5moles/mole of protein (Lees-Miller et al, 1987) and is not seen in vivo (Gimona et al, 1992). SM22α cDNA sequenced from chicken gizzard smooth muscle (Nishida et al, 1991) reveal that it is a 22.6KDa polypeptide that is the avian homologue of transgelin, the two proteins being 93.6% identical (Prinjha et al, 1994).

P27 is a mouse serum-induced cytoplasmic protein associated to the actin cytoskeleton and present in primary fibroblasts (Alemendral et al, 1989). Mouse p27 cDNA (Alemendral et al, 1989) shows that it is 96% identical to transgelin, and is therefore the mouse homologue (Prinjha et al, 1994). cDNA of WS3-10 has been isolated from senescent human fibroblasts from a patient with Werner Syndrome (WS) of premature ageing (Thweat et al, 1992). WS3-10 cDNA is 99% identical to transgelin, identifying it as the human homologue (Prinjha et al, 1994).

The rat neuronal protein, **NP25** (unpublished), a protein of 219 residues, is placed within the transgelin-related multigene family by virtue of the fact that its cDNA is 64.2% identical to transgelin (Prinjha et al, 1994).

Mp20 represents a cytoplasmic protein of 184 amino acids confined to synchronous muscles in *Drosophila* and is proposed to have two calcium-binding sites that are presumably involved in pumping calcium into and out of the sarcoplasmic reticulum (Ayme-Southgate et al, 1989). Mp20 cDNA identifies a 20KDa protein that is 38.6% identical to transgelin (Prinjha et al, 1994; Ayme-Southgate et al, 1989).

Calponin is a basic actin-binding, 34KDa, smooth muscle protein isolated from chicken gizzard (Walsh, 1991). Co-localisation of calponin and actin is also seen in bovine platelets and mouse fibroblasts (Takeuchi et al, 1991). It is a heat-stable, calmodulin-, F-actin-, tropomyosin-binding protein (see Walsh, 1991 for review). Calponin is considered as a regulator of smooth muscle contraction because: (i) it binds smooth muscle actin and actin-tropomyosin with high affinity; (ii) it is present in smooth muscle in equimolar concentration to tropomyosin ($\approx$1 calponin:7 actin monomers); (iii) immunofluorescence microscopy of cultured aortic and chicken gizzard smooth muscle cells localise calponin to the microfilament bundles; (iv) calponin and skeletal or cardiac troponin T exhibit immunological cross-reactivity and sequence homology; (v) it inhibits the actin-activated MgATPase activity of smooth and skeletal muscle myosins (reviewed by Walsh, 1991). Further support for this idea comes from the demonstration that calponin induces conformational changes in F-actin (Noda et al, 1992), and may itself undergo conformational changes in vivo that are associated with the contractile process (Barany et al, 1992).

Two isoforms of calponin have been identified in chicken gizzard (Takahashi & Nadal-Ginard, 1991). cDNA of the two isoforms have been isolated and sequenced and reveal: α- (292 residues; 32.3KDa) and β-calponin (252 residues; 28.1KDa) (Takahashi & Nadal-Ginard, 1991). α-calponin contains a C-terminal tandem repeat segment that appears three times, and β-calponin differs to this in the deletion of 40 amino acids, and thus, contains only two of these C-terminal repeats (Takahashi & Nadal-Ginard, 1991). The calponin isoforms are ≈38% identical to transgelin and are thus classified within the transgelin gene family (Prinjha et al, 1994). The lower Mr isoform of calponin (β), also known as l-calponin in mammalian smooth muscle (Draeger et al, 1991), is down-regulated in benign smooth muscle derived tumors (Draeger et al, 1991). In mammals, a distinct variant of calponin is found (Strasser et al, 1993) that is not a product of alternative splicing of α-calponin mRNA, nor is it β-calponin (Strasser et al, 1993).

The actin-binding site in chicken calponin has been located at the C-terminal end of the N-terminal 22KDa fragment which is obtained by limited chymotryptic digestion (Vancompernolle et al, 1990). At positions 142-147, Mezgueldi et al (1992) highlight a motif, VKYAEK, that has homology to the sequence LKHAET that is found in several ABPs, and is thought to be involved in actin-binding (Vancompernolle et al, 1991).

Smooth muscle calponins can be phosphorylated in vitro, since this protein is an excellent substrate for PKC (Winder & Walsh, 1990). Phosphorylation or the presence of Ca^{2+}-calmodulin (Abe et al, 1990) abolishes the ability of calponin to bind to actin and inhibit actomyosin ATPase activity (Winder & Walsh, 1990). In contracting or resting arterial smooth muscle, calponin is not phosphorylated (Barany et al, 1992). However, in intact canine tracheal smooth muscle, calponin (and caldesmon) is phosphorylated and dephosphorylated at a rate sufficient to indicate a role in contraction and relaxation (Barany et al, 1992). Three phosphoproteins, each of which is located within the three repeating segments in α-calponin, have been identified (Nakamura et al, 1993). As briefly mentioned above, one of these repeating segments is found in all the members of the transgelin-related family. The phosphorylation of one of these phosphoproteins within calponin may regulate the binding of calponin to F-actin and tropomyosin (Nakamura et al, 1993). The functional consequence of calponin phosphorylation is to prevent it from binding actin, thus allowing phosphorylated myosin heads free access to

the actin filament with consequent activation of the MgATPase (Winder et al, 1992). A phosphatase that dephosphorylates calponin, restoring its inhibitory effect, has also been identified (Winder et al, 1992).

C4^l, the lower Mr polypeptide of protein C4, is present in all cells apart from skeletal muscle, neurones and erythrocytes, and is apparently insensitive to cell shape change and transformation (Shapland et al, 1988). It has been partially purified from human T cell lymphoma (HTCL) to greater than 90% homogeneity (this study: fig 6), and is associated with the vestigial actin bundles of SV40 transformed 3T3 fibroblasts at the immunofluorescence level (this study: fig 1). Sequencing of C4^l PCR products (derived from gene specific oligonucleotides and reversed transcribed HTCL cDNA) identify full length C4^l cDNA to encode a 199 residue polypeptide with a calculated Mr of 22.3KDa (this study: fig 11) and an estimated pI of 8.09. Northern analysis reveals that C4^l is expressed as a single message of 1.44Kb (this study: fig 23). HTCL C4^l is a previously undescribed protein of unknown function, and its cDNA is 66.8% identical to rat transgelin (this study: table 5); thus, it confirms its inclusion within the transgelin-related multigene family of actin-associated proteins.

The extensive sequence homology noted between rat smooth muscle transgelin (Prinjha et al, 1994), chick SM22α (Pearlstone et al, 1987; Nishida et al, 1991), serum induced mouse fibroblast protein, p27 (Alemendral et al, 1989), and senescent fibroblast protein, WS3-10 (Thweat et al, 1992) indicates they are likely to be homologues. The significant sequence homology noted (this study, fig 19) between rat transgelin, rat neuronal protein NP25 (unpublished sequence, EMBL accession M84725), *Drosophila* mp20 (Ayme-Southgate et al, 1989), the calponins (Takahashi & Nadal-Ginard, 1991) and C4^l (this study), suggests they could well be derived from a common transgelin-like ancestor (Prinjha et al, 1994), and can thus, be classified as members of a multigene family (Prinjha et al, 1994; this study) in common with the other ABP families presented above (intro, section F).

Many functionally important microfilament-binding proteins have been characterised at the protein level in vitro (see Pollard & Cooper, 1986; Way & Weeds, 1990). However, very little is known about their regulation, their structure or their evolutionary origins. The most effective way of answering these questions is to isolate the genes coding for them.

The protein C4 doublet has previously been identified as a microfilament associated protein, with its high Mr isoform (now known as trangelin) down-regulated in transformed mesenchymal cells and highly motile cells such as lymphocytes (Shapland et al, 1988; 1993). By contrast, the low Mr protein C4 isoform (known as $C4^{l}$) is always present, its expression apparently insensitive to oncogenic transformation (Shapland et al, 1988). Although both isoforms are recognised by the same monoclonal antibody, they were thought to be different molecules by the fact that $C4^{l}$ is the only protein C4 isoform present in transformed cells and the fact that the protein C4 doublet is unaffected by high concentrations of reducing agents or a variety of protease inhibitors (Shapland et al, 1988). Having established transgelin and $C4^{l}$ as different molecules, it is therefore of interest to investigate both molecules at the level of the gene and thus elucidate their different expression patterns upon transformation, shape change or high motility.

As both protein C4 isoforms are recognised by the same monoclonal antibody, they must share some regions of homology. Unfortunately transgelin derived oligonucleotides (Prinjha et al, 1994) do not detect $C4^{l}$ cDNA or mRNA in library screening or Northern blotting experiments. This therefore dictated I purify the $C4^{l}$ molecule. From the purification and consequent microsequencing of $C4^{l}$, gene specific primers were generated and the previously undescribed $C4^{l}$ full length coding region isolated by a unique PCR approach that avoids the problems of screening libraries with oligonucleotides.

By isolating and sequencing the gene for $C4^{l}$, it has been possible to place it within the transgelin multigene family of cytoskeletal proteins. The availability of $C4^{l}$'s primary structure will permit further experimentation (see future work, Appendix II) to determine its physiological regulation and function in normal and transformed cells.

MATERIALS AND METHODS

A. ABBREVIATIONS.

APS	Ammonium persulphate
ATP	Adenosine Tri-Phosphate
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
CHAPS	3-[(3-cholamidopropyl) dimethyl-ammonio]-1-propansulponat
DEPC	Diethylpyrocarbonate
DNA	Deoxyribosenucleic acid
DTT	Dithiothreitol
EDTA	Ethylene-diamine-tetraacetic acid
EtBr	Ethidium Bromide
HTCL	Human T-Cell Lymphoma
IMS	Industrial methylated spirits
2-ME	2-mercaptoethanol
mRNA	Messenger RNA
MOPS	Morpholino-propanesulphonic acid
OD	Optical density
PMSF	Phenylmethylsulfonylfluoride
PBSA	Phosphate Buffered Saline
RNA	Ribosenucleic acid
rRNA	Ribosomal RNA
SDS	Sodium dodecyl sulphate (sodium lauryl sulphate)
SSC	Sodium chloride, sodium citrate buffer
SSPE	Sodium chloride, sodium phosphate, EDTA buffer
TAE	Tris-acetate, EDTA buffer
TBE	Tris-borate, EDTA buffer
TE	Tris-EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
TM	Tris-magnesium
Tris	Tris-hydroxymethylamine
tRNA	Transfer RNA

B. BUFFERS AND SOLUTIONS.

All solutions were prepared using distilled-deionised water and were stored at room temperature (20°C) unless otherwise stated. Solutions are listed below, alphabetically.

SOLUTION	COMPONENTS
Acrylamide (protein gels)	30% (w/v) acrylamide (BDH); 1% (w/v) N,N'-methylene-bisacrylamide (BDH). Filtered 0.4μm (Millipore). Stored foil-wrapped at 4°C.
Agarose gel sample buffer (Type I)	40% sucrose (BDH), 0.25% bromophenol blue (Hopkin & Williams), 0.25% xylene cyanol ff (BDH). Stored at 4°C.
Ammonium persulphate (APS)	10% (w/v) APS (BioRad) freshly prepared on day of use.
Ampicillin	50mg/ml ampicillin (Sigma). Filtered 0.2μm (Millipore). Aliquoted, stored at -20°C.
Antibody Block Buffer (Immunoblotting)	3% BSA (Sigma) in PBSA with 0.1% sodium azide (Sigma).
Antibody Block Buffer (Immunofluorescence)	0.3% BSA (Sigma), 100mM lysine (Sigma) in PBSA
Antibody Wash Buffer	0.3% BSA (Sigma) in PBSA.
10x Blunt End ligation buffer	0.66M Tris. Cl (pH 7.6) (Sigma), 50mM $MgCl_2$, 50mM DTT, 1mg/ml BSA (Sigma), 10m Hexaminecolbaltchloride, 2mMATP (Pharmacia), 5mM Spermidine HCl (Sigma) Store in small aliquots at -20°C
Coomassie Blue Stain	50% (v/v) methanol (BDH), 10% (v/v) glacial acetic acid (BDH), 0.05% (w/v) coomassie Brilliant blue R (Sigma).
Denhardt's solution (50x)	1% (w/v) BSA (Sigma fraction V); 1% (w/v) Ficoll (Pharmacia); 1% (w/v) Polyvinylpyrrolidon k90 (Fluka). Filtered 0.4μm. Aliquoted, stored at -20°C.

Destain	30% (v/v) methanol (BDH), 10% (v/v) glacial acetic acid (BDH).
Diethylpyrocarbonate	DEPC (Fluka) used at 0.1% to inactivate RNAases in solutions (except solutions containing methylamines eg Tris-Cl), removed by autoclaving.
Dithiothreitol	1M DTT (Sigma) in 0.01M sodium acetate (pH5.2). Filtered 0.4µm. Aliquoted, stored at -20°C.
Dulbecco's Modified Eagle's Medium (DMEM)	Commercially purchased (500ml) and stored at 4°C (Sigma).
Electroblot Transfer Buffer	25mM Tris base (Sigma), 192mM glycine (Sigma), 20% (v/v) methanol (BDH); pH.≈8.3.
Ethylene-diamine-tetraacetic acid	0.5M EDTA (Sigma) adjusted to pH8. Sterilised by autoclaving.
Ethidium bromide	Preweighed 11mg EtBr (BioRad) tablets dissolved in 11ml dd$H_2$0, in fume hood, to 1mg/ml. Stored foil-wrapped at 4°C.
Filter Stripping Solution	0.05x SSC, 0.01M EDTA pH8 at 100°C plus 0.1% SDS (added once solution cooled to 98°C).
Formamide	Formamide (Fisons) deionised by passage over BioRad AG501-X8 mixed bed resin until neutral pH was attained. Aliquoted, stored at -20°C.
Glucose	1M glucose (BDH). Filtered 0.45µm. Stored at 4°C.

Homo Mix	5g yeast RNA (grade VI, Sigma) in 22ml ddH_20 with 3ml 5N KOH (BDH) incubated at 37°C for 24 hours. Adjusted to pH7 with 1N HCl. Extracted three times with phenol/ chloroform and once with chloroform. Precipitated, resuspended, precipitated again then resuspended to 50mg/ml. Aliquoted, stored at -20°C. Heated to 100°C for 5 minutes before use.
N-2-Hydroxy-ethyl-piperazine N'-2 ethane-sulphonic acid	2M HEPES pH7.6 (Sigma). Filtered 0.2µm. Stored at 4°C.
Laemmli Sample Buffer	80mM Tris-Cl pH6.8, 2.3% SDS (BDH), 0.002% bromophenol blue (BDH), 1.42M 2-mercaptoethanol (BDH), 20% glycerol (Fluka). Made fresh on day of use.
Laemmli Running Buffer	1x is 25mM Tris base, 192mM glycine (Sigma), 0.1% SDS (BDH).(pH≈8.3)
Luria Bertani Broth	LB (pH 7.4). 1% bacto-tryptone (Difco), 0.5% bacto-yeast extract (Difco), 1% NaCl (Sigma). Aliquoted, sterilised by autoclaving.
Magnesium sulphate	1M or 10mM $MgSO_4$ (Fluka). Sterilised by autoclaving.
Morpholinopropanesulphonic acid Buffer	5x MOPS is 0.1M MOPS pH 7.0 (Fluka), 40mM sodium acetate, 5mM EDTA pH8. Filtered 0.4µm. Stored foil-wrapped at room temperature.
2-mercaptoethanol	14.2M 2-ME (BDH). Stored at 4°C in dark bottles.

Oligolabelling Buffer (OLB)	OLB is prepared from the following stock solutions. Solution O: (1.25M Tris-HCl (pH8), 0.125M $MgCl_2$); 2M HEPES (Sigma); Random hexamers (90OD unit/ml (Pharmacia)); and 2-mercaptoethanol (14.2M (BDH)). Aliquoted, stored -20°C
Phenol	Liquefied phenol (Fluka Biochemika) was saturated with DEPC'd ddH_20. Aliquots removed to 50ml tubes and made 0.1% (w/v) 8-hydroxyquinoline (BDH). Acidic phenol for use with RNA was left with ddH_20. Phenol for use with DNA was equilibrated with 1M Tris-Cl pH8 until the pH of the aqueous layer was >7.5, then covered with TE (pH 8). Both stored foil-wrapped at 4°C for up to 2 weeks.
Phenol/ Chloroform	Phenol (Fluka), chloroform (BDH) and isoamylalcohol (BDH) mixed in the ratio (50:49:1), equilibrated with TE (pH 8) and stored foil-wrapped at 4°C for up to 2 weeks.
Phosphate Buffered Saline (PBSA)	1x PBSA. 137mM NaCl (Sigma), 2.7mM KCl (BDH), 8mM NaH_2PO_4 (Sigma), 1.45mM KH_2PO_4 (Sigma); pH >7.5.
Potassium Phophate Buffer	1M potassium phosphate buffer pH 6.8. Prepared by mixing 49.7ml 1M K_2HPO_4 (Fluka) and 50.3ml 1M KH_2PO_4 (Fluka). Sterilised by autoclaving.
Prehybridisation Buffer	2x SSC, 0.1% SDS, 5x Denhardt's, (± 50-100µg/ml denatured Homo mix).
RNAase A	10mg/ml bovine pancreatic RNAase A (Worthington Enzymes) in 10mM Tris-Cl pH7.5, 15mM NaCl. Heated to 100°C for 15 minutes, cooled slowly to room temperature. Aliquoted, stored at -20°C.

RNA Extraction Buffer	6M urea (Fluka), 3M lithium chloride (Fluka). Filtered 0.2μm. Stored at 4°C.
RNA Sample Buffer	10x Sample buffer. 50% (v/v) glycerol (Fluka), 1mM EDTA pH8 (Sigma), 0.25% (w/v) bromophenol blue (Hopkin & Williams), 0.25% xylene cyanol ff (BDH). Stored at -20°C.
Sequencing Sample Buffer	95% formamide, 20mM EDTA, 0.05% (w/v) bromophenol blue, 0.05% xylene cyanol ff, (Sequenase kit, USB). Stored at -20°C.
SOB Medium	2% (w/v) Bacto-tryptone (Difco), 0.5% (w/v) Bacto-yeast extract (Difco), 1% (w/v) NaCl (Sigma), 2.5mM KCl pH 7.0. 20mM glucose (BDH) and 20mM $MgCl_2$ (Sigma) were added after autoclaving.
SOB Agar	SOB plus 1.5% (w/v) agar (Difco) sterilised by autoclaving.
Solution I	50mM glucose (BDH), 10mM EDTA (Sigma), 25 mM Tris. Cl (Sigma) pH 8.0. Autoclaved and stored at 4°c.
Solution II	0.2M NaOH (Sigma), 1% (w/v) SDS. Freshly made and 0.45μm filtered before use.
Solution III	3M potassium acetate (Fluka), 2M glacial acetic acid (BDH). Freshly made before use.
Sodium Acetate	3M sodium acetate (Fluka) was adjusted to pH5.2 with glacial acetic acid, made 0.1% DEPC then autoclaved.
Sodium chloride	5M NaCl (Sigma). Sterilised by autoclaving.
Sodium dodecyl sulphate	10% SDS (BDH Electran). Filtered 0.4μm.

SSC	20xSSC. 3M sodium chloride, 0.3M tri-sodium citrate (BDH), pH7.0. Sterilised by autoclaving.
T4 Polynucleotide Kinase Buffer	10x T4 PNK buffer. 500mM Tris-Cl (pH7.6) (Sigma), 100mM magnesium chloride (Sigma), 50mM DTT (Sigma), 1mM spermidine (Fluka), 1mM EDTA (Sigma). Filtered 0.4μm. Aliquoted, stored at -20°C.
Tris Acetate, EDTA Buffer (TAE)	1x concentration: 0.04M Tris.Acetate, 0.001M EDTA
Tris-Borate EDTA Buffer (TBE)	10x TBE (pH ≈8.3). 0.89M Tris base (Sigma), 0.89M boric acid (Sigma), 20mM EDTA (Sigma). Sterilised by autoclaving.
TE	10mM Tris-HCl (at appropriate pH 7 - 8), 1mM EDTA. Sterilised by autoclaving.
TM	10mM Tris-HCl pH7.4, 10mM $MgSO_4$. Sterilised by autoclaving.
Trypsin	ICRF media supplies; 0.8% (w/v) NaCl, 0.038% (w/v) KCl, 0.01% (w/v) disodium hydrogen orthophosphate, 0.01% (w/v) dextrose, 0.3% (v/v) Tris.HCl (pH7.7), 0.25% (w/v) trypsin, 10,000u sodium penicillin, 0.01% (w/v) streptomycin, phenol red; stored at -20°c.
Versine	ICRF media supplies; 0.02% (w/v) EDTA in PBSA.
Wash Buffer	60mM KCl, 4mM $MgCl_2$, 10mM imidazole (pH 7), 1mM NaN_3, 0.5mM EGTA, 5mM DTT, 0.1mMEDTA, 0.2mMPMSF (all Sigma).

C. TISSUE CULTURE.

All tissue culture preparation was carried out in a Microflow Pathfinder hood, using UV sterilised plastics (Falcon Labware, Oxnard, CA). DMEM (Sigma) plus 10% heat inactivated fetal calf serum (FCS) (ICN Flow Laboratories), 4000units/litre penicillin (Flow Laboratories), 4 000µg/litre streptomysin (Flow Laboratories) and 584mg/litre glutamine (Flow Laboratories) was used in all cases. All cultures were grown in a 37°c humidified incubator (LEEC) in an atmosphere of 5% CO_2. Cultures were observed by use of an inverted Olympus Tokyo microscope at x10 objective and x15 eyepiece.

1. Rat Embryo Fibroblasts (REF)

Secondary cultures of REF were obtained by sterile dissection of 15 day old rat embryo limb-buds. The limb-buds were teased apart in DMEM and incubated at 37°c for 24 hours. DMEM and unadhered tissue was removed from the culture and fresh DME medium added. REF cultures were maintained for several weeks. Regular cell passaging was carried out by removing spent DMEM and adding 10ml of trypsin:versine (3:1) (trypsinisation) and incubating for 10 minutes at 37°c. The resultant cell suspension was centrifuged in an MSE Benchtop at 150x g for 10 minutes at room temperature. The cell pellet was resuspended in 1ml of fresh DMEM and ≈50µl added to 25ml of fresh DME medium. The remainder of the resuspended pellet could be frozen in 90% FCS, 10% DMSO at -20°c overnight, with subsequent storing in liquid nitrogen.

2. Human T-Cell Lymphoma (HTCL)

A 1ml frozen culture of HTCL was obtained from ICRF cell production unit. It was thawed slowly at room temperature, immediately transferred to 10ml of fresh DMEM and centrifuged in an MSE Benchtop at 150x g for 10 minutes at room temperature. The cell pellet was resuspended in fresh DME medium, maintained at a maximum of $2x10^6$ cells/ml and processed for biochemical analysis and protein purification.

3. SV40 Transformed 3T3 Fibroblasts

A frozen culture of SV40-3T3 fibroblasts, obtained from ICRF cell production

unit, was thawed as above. Cultures were maintained until confluence was reached when the cells were passaged as described above. SV40-3T3 fibroblasts were maintained for biochemical analysis and immunofluorescence.

D. ANTIBODIES.

1. Anti-C4 Monoclonal

This antibody was generated by D. Lawson (1983) as follows:

1.1. Immunogen:

Fresh chicken gizzard was homogenised in a Waring blender with 3x 10 seconds bursts at top speed in 10 volumes of deionised (4°c) plus 0.5mM phenylmethylsulfonylfluoride (PMSF). The suspension was centrifuged at 5000x g for 10 minutes, the supernatant discarded, the pellet resuspended as above and blended for 10 seconds at low setting. The centrifugation was repeated and the pellet resuspended at 37°c in 10 volumes of 2mM Tris, 1mM EGTA, 0.5mM PMSF (pH 9 at room temperature), and stirred for 30 minutes at 37°c. The centrifugation was repeated and the supernatant adjusted to pH 7-7.2 with 0.5mM acetic acid. Molar $MgCl_2$ was added till the suspension contained 10mM $MgCl_2$. After 15 minutes of stirring at room temperature, the suspension was centrifuged as above and the supernatant used as the immunogen

1. 2. Immunisation:

BALB/c mice were immunised as follows, with 100µg of immunogen per injection; day 1: subcutaneous in complete Freund's adjuvant; days 14, 63 and 77: interperitoneal in incomplete Freund's adjuvant; day 81: tail bleed and test for antibody response; days 84, 85 and 86: intravenous injection; day 87: fusion.

1.. 3. Fusion and Hybridoma Production:

This was performed as described by Lawson (1983), the myeloma cell line was the non-secretor SP2 and the culture medium was DMEM.

1. 4. Screening and Cloning Hybridomas:

When the hybridoma wells two thirds confluent they were tested. Antibody producing cells were cloned by limiting dilution into 96 well Linbro round bottom

microtitre plates with a feeder layer of adherent BALB/c macrophages at a concentration of $5x10^3$ cells/ml. Cloning was carried out twice and cells were frozen in 5% DMSO, 95% FCS at -27°c for 10 minutes, and then stored in liquid nitrogen. Ascites was raised by injecting nu/nu mice with $10x\ 10^7$ cells and harvesting the ascitic 10 days later. The anti-C4 monoclonal was used at a concentration of 20µg/ml in biochemical analysis and immunofluorescence.

2. Anti-C4 Polyclonal

This antibody was generated by Shapland et al (1993) as follows:

Polyclonal anti-C4 was raised by cutting strips containing the protein transgelin from aorta electrophoresed on 3mm thick SDS-PAGE gels. These were then electroeluted into 1x Laemmli running buffer (25mM Tris base, 192mM glycine, 0.1% (w/v) SDS) in dialysis tubing (Spectra/por), acetone precipitated and injected into rabbits over a three month period. A DEAE 52 (Whatman) prepared IgG fraction was passed over an affinity column of purified transgelin (2mg) coupled to Affigel 10 (Biorad), eluted with 50mM diethylamine (Sigma) (pH 11.5), immediately neutralised with 1M Tris.HCl (pH7.5), dialysed into PBSA, concentrated by millipore filtration, and used at concentration of 2-10µg/ml in biochemical analysis.

3. Goat Anti-Rabbit IgG-Fab_2

Goat anti-rabbit IgG-Fab_2 was purchased from Cappel Dynatech, affinity purified as above and conjugated to ^{125}I as described by Lawson (1983). 10µg of antibody in a total volume of 25µl of 0.5M sodium phosphate (pH 7.5) in a 1.5 ml microcentrifuge tube. 500µCi of $Na^{125}I$ (Amersham) and 25µl of 2mg/ml chloromine T (Sigma) was added and incubated for 1 minute at room temperature. Following incubation 50µl of chloromine T stop buffer (2.4mg/ml sodium metabisulfite (Sigma)) was added. The iodinated antibody was separated on a G-25 prepacked column (Pharmacia) and stored at 4°c for up to 6 weeks.

4. Goat Anti-Mouse Ig Peroxidase

Rabbit anti-mouse Ig peroxidase was purchased from DAKOPATTS, Copenhagen. It was used at a concentration of 1 in 2000 in biochemical analysis.

5. IgG Specific Goat Anti-Mouse Rhodamine

IgG specific goat anti-mouse rhodamine was purchased from Cappel Laboratories, Cochanville, PA. It was used at a concentration of 1 in 100 in immunofluorescence studies.

E. IMMUNOFLUORESCENCE.

SV40 transformed 3T3 fibroblasts were adjusted to 1x 10^3 cells/ml in DMEM and plated in 1ml aliquots onto 13mm diameter sterile coverslips in 24-well Linbro plates (Flow Laboratories) for 3 days. Cells were rinsed briefly in PBSA and then permeabilised by formaldehyde fixation (3.5% (v/v) in PBSA for 12 minutes at room temperature) followed by cold methanol (10 minutes at -20°c). After brief rinsing in PBSA, cells were incubated for 10 minutes in antibody block buffer (1% (w/v) BSA, 100mM lysine in PBSA), probed with 20μg/ml anti-C4 monoclonal antibody (methods, section D. 1) for 30 minutes at room temperature and rinsed in PBSA before incubation with an IgG-specific goat anti-mouse rhodamine (methods, section D. 5) at a dilution of 1 in 100 in antibody block buffer for 30 minutes. Coverslips were further rinsed in PBSA and mounted onto glass slides using Gelvitol mount (Monsanto Polymers Ltd, USA). Once air-dried, coverslips were visualised on a Nikon Optiphot fluorescence microscope with a 60x objective and 10x eyepiece. Fluorescent images were observed by using an epifluorescence attachment. To obtain a fluorescence exposure base line for these experiments, the exposure time required for control cells was measured and used for photographing (on a Nikon FX-35 WA camera) all other preparations.

F. BIOCHEMISTRY.

1. SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE) (Laemmli, 1970)

All SDS-PAGE analysis was carried out using Biorad Miniprotean II Minigel apparatus.

Clean glass gel plates were assembled according to manufacturer's instructions and the gel mould volume determined. The appropriate volume of solution containing the desired concentration of acrylamide (from a stock solution of 30% (w/v) acrylamide

(Biorad), 0.8% (w/v) NN'-methylenebisacrylamide (BDH)) for desired resolution (typically: 12% (v/v) acrylamide, 377mM Tris.HCl (pH 8.8) in distilled H_2O), was degassed, cross-linked with 0.1% (w/v) SDS, 0.033% (w/v) ammonium persulphate (APS), 0.00033% (v/v) NNN'N'-tetramethylenediamine (TEMED) and carefully poured into the gel mould. Gel was poured leaving space for a stacking gel (depth of the teeth of the comb plus 1cm), overlayed with isobutanol and allowed to polymerise in a vertical position at room temperature for at two hours. After polymerisation overlay was removed and the top of the gel washed with deionised water to remove unpolymerised acrylamide. A 5.5% acrylamide, 140mM Tris.HCl.(pH 6.8) stacking gel was degassed, cross-linked with 0.1% (w/v) SDS, 0.056% (w/v) APS, 0.00056% (v/v) TEMED and poured directly onto the surface of the polymerised resolving gel. A Teflon 10 well comb was inserted immediately into the stacking gel solution (no air bubbles), and allowed to polymerise for 30 minutes at room temperature. While the stacking gel was polymerising samples (typically 1x 10^6 cells or the equivalent amount of tissue) were heated at 100°c for 3 minutes in 20µl of Laemmli sample buffer (0.8mM Tris.HCl (pH 6.8), 2.3% (w/v) SDS, 0.003% (w/v) bromophenol blue, 1.42m 2-Mercaptoethanol and 20% (v/v) glycerol). After stack gel polymerisation the Teflon comb was removed and the wells immediately washed in 1x Laemmli running buffer (25mM Tris base, 192mM glycine, 0.1% (w/v), pH 8.3). The gel was mounted into the electrophoresis apparatus, with 1x Laemmli running buffer added to the top and the bottom reservoirs. 15µl of each sample was loaded in a predetermined order along with 5µg of dye-coupled molecular mass markers (BRL Life Technologies) to the bottom of the wells. Electrophoretic apparatus were attached to an electrical power supply (+ve electrode connected to the bottom buffer reservoir) and run at 50volts until the dye front reached the resolving gel when running was continued at 150volts. Following electrophoresis, the gel was removed and resolved proteins stained in 5 volumes of coomassie blue stain (50% (v/v) methanol, 10% glacial acetic acid, 0.05% (w/v) Coomassie Brilliant blue R (all BDH)) for several hours at room temperature. Gels were destained in 5 volumes of destain (30% (v/v) methanol, 10% glacial acetic acid) for the desired time and stored in 7% (v/v) acetic acid.

2. Immunoblotting (Towbin et al, 1979)

Gloves were worn at all stages of immunoblotting. 2 pieces of 3mm-CHR Whatman paper, 4 pieces of porous pad and one piece of 0.45μM pore size nitrocellulose filter (Schleicher & Schuell GmbH Dassel, FRG) were cut to the exact size of the electrophoresed SDS polyacrylamide gel under analysis. The filter, 3mm-CHR paper and porous pads were soaked in 1x electroblot transfer buffer (25mM Tris.base, 192mM glycine, 20% (v/v) methanol (pH 8.3)) for a few minutes. Onto the cathode of the Biorad Miniprotean II transfer system, 2 layers of porous pad and one sheet of 3mm-CHR paper were placed. The gel was removed from the glass plates, transferred to a tray of deionised water for a few minutes then placed exactly on top of the 3mm-CHR paper. By avoiding air bubbles, the gel was overlayed with the nitrocellulose filter (corner marked for orientation of choice). The other sheet of 3mm-CHR paper and then the other layers of porous pad were placed onto the nitrocellulose filter. The transfer apparatus were assembled and a voltage of 100v applied for 1 hour in 1x transfer buffer. Filters were removed and placed directly into antibody block buffer (3% (w/v) BSA, 0.1% (w/v) NaN_3 in PBSA) for 12 hours at 4°c. Following blocking, filters were washed briefly in antibody wash buffer (0.3% (w/v) BSA in PBSA) and incubated with the primary antibody (either 20μg/ml C4 monoclonal (methods, section D. 1), or 2μg/ml C4 polyclonal (methods, section D. 2)) for 2 hours at room temperature. Filters were washed 6x, 5 minutes each, in antibody wash buffer and then incubated with the secondary antibody (either goat anti-mouse Ig peroxidase (methods, section D. 4), or ^{125}I conjugated goat anti-rabbit IgG-Fab_2 (methods, section D. 3) for 2 hours at room temperature. Filters were washed as above and either peroxidase developed or exposed to autoradiographic film depending on which antibody system was used. Peroxidase developing was carried out by incubating filters in a solution of 16.6% (v/v) cold methanol, 0.015% (v/v) H_2O_2 and 0.05% (w/v) 4-Chloronapthol in PBSA for 15 minutes at room temperature. Developing reactions were stopped by flushing filters in PBSA, after which filters were air-dried and stored in the dark. Filters immunoblotted with ^{125}I conjugated goat anti-rabbit IgG-Fab_2 were washed in antibody wash buffer, as above, until they registered approximately 50 cps, air-dried, further dried at 80°c for one hour, and then exposed to fast Hyperfilm-MP (Amersham) at -70°c for the desired length of

time.

G. PURIFICATION OF C4[1].

1. Concentrations of Protein Solutions (Wessel & Flugge, 1984)

Fractions generated from chromatographic purification often contained low concentrations of protein. In order to visualise these fractions on SDS-PAGE (methods, section F. 1), they were concentrated by the Wessel/Flugge method of protein extraction. To 150µl of dilute protein solution, in a microcentrifuge tube, 0.75ml of methanol:chloroform (1:1) was added and vortexed vigorously. To this, 0.45ml of distilled water was added and again vortexed before centrifugation at 12000x g for 1 minute in a microcentrifuge (MSE). The aqueous phase was removed (leaving precipitated protein at the organic phase interface). A further 0.45ml of methanol was added, vortexed and microcentrifuged at 12000x g for 5 minutes. The supernatant was completely removed and the protein pellet air-dried before resuspension and boiling in Laemmli sample buffer and SDS-PAGE visualisation (methods, section F. 1).

2. Cell Harvesting

Three litres of HTCL cell culture at a concentration of 2x 10^6 cells/ml were collected by centrifugation in 250ml buckets using a JA2-21 Beckman centrifuge (Beckman Instruments Inc, Palo Alto, CA) and a JA14 angle rotor at 2460x g for 10 minutes at 4°c. The cell pellet was carefully resuspended in 100ml of ice cold PBSA containing 0.2mM PMSF. Resuspended pellets were again centrifuged using a JA2-21 Beckman and a JA17 angle rotor at 2200x g for 10 minutes at 4°c to ensure total removal of FCS.

3. Cell Lysis

All subsequent purification steps were carried out at 4°c, in the presence of 2µg/ml leupeptin, chymostatin and pepstatin A (all Sigma). To every 7x 10^7 HTCL cells 0.5ml of ice cold wash buffer (60mM KCl, 4mM $MgCl_2$, 10mM imidazole (pH 7.0), 1mM NaN_3, 0.5mM EGTA, 5mM DTT, 0.1mM EDTA, 0.2mM PMSF) plus 1% (w/v) CHAPS (Boehringer Mannheim, FRG) was added and incubated on ice for 40 minutes to allow 100% lysis. Detergent insoluble material was removed by centrifugation at 23,000x g for 30 minutes at 4°c using a 6x38 swing-out rotor in a Prepspin 50 centrifuge (MSE).

The supernatant was filtered over glass wool, millipored (0.45μM) and dialysed (in Spectra/Phor Molecularporous Membrane Tubing) for 16 hours at 4°c against 2 litres of chromatofocusing start buffer (25mM ethanolamine-HCl, pH 9.6 (Sigma)). Fractionation of HTCL lysate was carried out on a chromatofocusing column between the ranges pH 7 to pH 9.6.

4. Chromatofocusing

4. 1. Column Packing:

A C16/70 Pharmacia column (1.6cm diameter, 70cm length, 2.0cm^2 cross-sectional area) was packed with an appropriate volume of PBE 94 gel matrix (Pharmacia). The chosen amount of PBE 94 was dispersed in a small amount of start buffer (approximately 50% of settled bed volume) to make a slurry and was degassed prior to packing into C16/70 columns. Columns were mounted vertically and air bubbles flushed out from underneath the bed support net and the column outlet closed. Degassed slurry was carefully poured into columns after 2-3ml of start buffer had been added. The column outlet tube was opened and PBE 94 matrix allowed to settle. The top of column was sealed up and start buffer pumped through the packing PBE 94 by use of a2132 microperpex® peristaltic pump (LKB Bromma) at a superficial linear flowrate (volume flowrate, mls per hour / cross-sectional area, cm^2) of 100cmh^{-1}. If PBE 94 was not equilibrated to pH 9.6 prior to packing, then approximately 10-15 column bed volumes of start buffer was required for equilibration. Equilibration of the column was reached when the pH and conductivity of the eluent matched that of the start buffer. A level 'base line' was reached (as seen on a LKB Bromma 2210 2-channel chart recorder) before sample application.

4. 2. Sample Application:

Polybuffer 96-HCl (Pharmacia) was adjusted to pH 7.0 by the addition of 0.5M HCl. 5ml of Polybuffer 96-HCl was added to the top of the column bed immediately prior to sample application. Dialysed and 0.45μM millipored HTCL lysate (typically 40ml in total) was carefully loaded onto the top of the column bed at a superficial flowrate of 30mlh^{-1}.

4. 3. Elution:

Polybuffer 96-HCl, adjusted to pH 7.0 was used as the chromofocusing eluent once the sample had been fully applied. Elution was carried out at a linear flowrate of 30cmh^{-1} and continued until the column eluent was at pH 7.0 (approximately 14 bed volumes; ie. 2 litres). Eluent was analysed by a 2238 Uvicord SII (LKB Bromma), monitoring at 280nM, and charted via an LKB Bromma 2210 2-channel chart recorder. 0.5ml fractions were collected on a 2111 Multirac fraction collector (LKB Bromma). Peaks (as seen on chart tracings) were analysed by SDS-PAGE (methods, section f. 1), and the fractions containing significant levels of C4^{l} pooled and dialysed against 1 litre of 10mM potassium phosphate buffer (pH 6.8) for 16 hours at 4^{o}c.

5. Hydroxylapatite

5. 1 Resuspending:

Bio-Gel HT hydroxylapatite (Biorad) was shipped in 10mM sodium phosphate buffer, pH 6.8. It tends to pack in the bottle during shipping and requires resuspension before column pouring. The bottle was swirled gently until the gel was in suspension (no stirring rods used). After suspension the desired volume of gel was poured into a beaker and allowed to settle for 30 minutes. The cloudy upper level (fines) were decanted from the settled bed which was equilibrated using 10mM potassium phosphate buffer, pH 6.8.

5. 2. Column Packing:

A 0.75cm diameter, 10cm length glass barrel. Econo-column® (Biorad) was packed with 3 ml of Bio-gel HT hydroxylapatite slurry. The slurry was gently poured into the column, with the column outlet closed. The slurry was allowed to settle under gravity, and the column outlet opened. When the bed was stable, at least 5 column bed volumes of 10mM potassium phosphate buffer (pH 6.8) was pumped (by use of a LKB Bromma 2132 Microperpex® peristaltic pump) through the bed at a linear flowrate of 6.8cmh^{-1} to equilibrate the gel and allow the construction of a 'base line' (as seen on an LKB Bromma 2210 2-channel chart recorder) before sample application.

5. 3. Sample Application:

Pooled, C4^{l} positive, chromatofocus fractions (10ml in total) was applied carefully to the top of the column bed at a linear flowrate of 6.8cmh^{-1}. The column outlet led to a 2238 Uvicord SII (LKB Bromma) and proteins not binding to hydroxylapatite

recorded as a peak on the chart recorder. Once the sample had been loaded, the column was washed with 10mM potassium phosphate buffer, pH 6.8, to bring the chart recorded trace back down to 'base line'.

5. 4. Elution:

Hydroxylapatite columns were eluted using 4ml steps of 2x increasing concentrations of potassium phosphate buffer, pH 6.8 (10-100mM). Eluent was monitored, recorded and 0.5ml fractions were collected on a 2111 multirac fraction collector (LKB Bromma). Protein purity was assessed by SDS-PAGE (methods, section F. 1).

Fractions positive for $C4^{l}$ (approximately 2ml in total) were pooled. Further concentration and buffer exchange into 10mM Tris.HCl, pH 8.0 was carried out by use of an Ultra Free-MC 10,000 NMWL filter unit (millipore) centrifuged at 6500x g in a microcentrifuge (MSE) until a volume of 100µl was reached.

H. AMINO ACID SEQUENCING.

All amino acid sequencing was carried by Drs. J. Hsuan and N. Totty (Ludwig Institute, London).

Purified $C4^{l}$ was applied to an automated protein sequencer. Material was reduced and alkylated with 4-Vinylpyridine as described (Andrews & Dixon, 1987) prior to digestion with trypsin. The resultant peptides were separated by reverse phase HPLC using an Applied Biosystems OD. 300 3.1x 100mm column employing a linear gradient of acetonitrile and monitored by a diode array detector (HP. 1090/1040A). Peptides were sequenced using an Applied Biosystems 477A using modified chemistry and analysis cycles (Totty et al, 1992).

I. RNA PREPARATION.

1. Total RNA Isolation (Auffray & Rougeon, 1980)

All solutions were treated with DEPC before use in RNA isolation.

To tissue (approximately 1g) or pelleted tissue culture cells (approximately 6x 10^{8}) 10ml of RNA extraction buffer (6M urea, 3M LiCl (both Fluka)) was added.

Samples were homogenised by use of an Ultra-Burrax homogeniser set at high speed for 1 minute on ice. Samples were allowed to precipitate for 16 hours at 4°c. Sample homogenate was transferred to a 15ml corex tube and centrifuged at 16,700x g for 30 minutes at 0°c in a Beckman JA17 rotor mounted in a Beckman J2-21 centrifuge. Supernatant was discarded and pellets resuspended by vortexing in 8ml of RNA extraction buffer and centrifuged as above. Supernatant was removed and the tube drained well to leave a dry pellet. Pellets were resuspended in 6ml Tris.HCl (pH 7.5), 0.5% (w/v) SDS and incubated at room temperature for at least 2 hours, with occasional vortexing. Preparations were transferred to a 14 ml polypropylene tube (Falcon Labware) and extracted twice with an equal volume of phenol. Aqueous layers were further extracted with an equal volume of phenol:chloroform (1:1), followed by an extraction with an equal volume of chloroform. To extracted aqueous layers 1/10 volume of 3M sodium acetate (pH 5.5) was added and transferred to a 30ml corex tube where 2.5 volumes of absolute ethanol was added and incubated at -70°c for 16 hours. Precipitated samples were centrifuged at 16,700x g for 45 minutes at 0°c in a Beckman JA17 rotor mounted in a Beckman J2-21 centrifuge. Supernatants were removed, pellets completely air-dried and washed in 80% (v/v) ethanol. Further centrifugation (as above) and 80% (v/v) ethanol washing was carried out. Final pellets were completely air-dried and dissolved in 0.5ml DEPC treated distilled water and stored in a sterile 1.5ml microcentrifuge tube at -70°c. Concentration of samples were determined by optical density (see methods, section J).

2. Poly A$^+$ Selection of mRNA (Sambrook et al, 1989)

All solutions used in polyA$^+$ selection of mRNA were treated with DEPC. Selection was carried out in batch fashion, rather than use of columns.

1mg of total RNA from each sample was made up to a volume of 0.5ml in distilled water. Samples were heated to 65°c for 5 minutes and then quenched on ice to quickly cool them to room temperature. To the samples, an equal volume of 2x column loading buffer (40mM Tris.HCl, pH 7.6, 1M NaCl, 2mM EDTA and 0.1% (w/v) N-lauryl sarcosine) was added and loaded to 100µl of oligo dT (Sigma) slurry. Slurry was initially prepared by swelling 0.5g oligo dT in 5ml 1x column loading buffer (20mM Tris.HCl, pH 7.6, 0.5M NaCl, 1mM EDTA and 0.1% (w/v) N-lauryl sarcosine) for 12 hours at 4°c.

Following swelling, slurry was centrifuged in an MSE Benchtop at 500x g for 4 minutes. Pellets were resuspended in 5ml of 0.1N NaOH, 5mM EDTA and centrifugation repeated as above. Resultant pellets were resuspended in 15ml distilled water, recentrifuged as above and washed in 1x column loading buffer. Pellets were continually washed until the pH was less than 8.0. Final pellets were resupended and stored in 1ml of 1x column loading buffer at 4°c. Samples were incubated with oligo dT slurry in a 14ml polypropylene tube (Falcon Labware) for 40 minutes at room temperature with constant shaking. Following incubation, samples were centrifuged at 1000x g for 5 minutes in an MSE Benchtop. The supernatant was removed from pelleted oligo dT, heated to 65°c for 5 minutes and rapidly cooled to room temperature on ice. Oligo dT was kept constantly moist in 1x column loading buffer. Cooled supernatant was reapplied to oligo dT and incubated for 15 minutes at room temperature with constant shaking. Centrifugation was repeated as above and supernatant removed, stored at -20°c and known as polyA$^+$ 'flow-through'. Oligo dT (and bound mRNA) was washed with 5ml of 1x column loading buffer and centrifuged as above and supernatant discarded. The washing procedure was repeated 5 times. PolyA$^+$ mRNA was eluted into 4 times 1ml changes of elution buffer (10mM Tris.HCl, 7.6, 0.05% (w/v) SDS and 10mM EDTA, pH 8.0) by resuspending oligo dT and collecting the resultant supernatant after centrifugation as above. Fractions were pooled and heated to 65°c for 3 minutes and rapidly cooled to room temperature on ice. To pooled fractions 0.45ml of 5M NaCl was added before reapplication to moist oligo dT. Incubation with oligo dT was carried for 45 minutes at room temperature with constant shaking. After centrifugation (as above) supernatant was removed and oligo dT washed 3 times with 1x column loading buffer. mRNA was again eluted from oligo dT into 4x 1ml changes of elution buffer (as above). To pooled elution fractions 1/10 volume of 3M sodium acetate, pH 5.2, was added and transferred to 30ml corex tubes where 2.5 volumes of absolute ethanol was added. PolyA$^+$ selected mRNA was precipitated for 16 hours at -70°c. After precipitation, samples were centrifuged at 16,700x g for 30 minutes at 0°c in a Beckman JA17 rotor mounted in a Beckman J2-21 centrifuge. Supernatant was removed, pellets air-dried and washed in 80% (v/v) ethanol and recentrifuged as above. Resultant pellets were gently air-dried and dissolved in 100µl of distilled water.

PolyA$^+$ selected mRNA samples were stored at -70^{o}c and concentration determined as described below (methods, section J).

J. DETERMINATION OF NUCLEIC ACID CONCENTRATION.

All concentrations of nucleic acid were determined by measuring optical density at 260nm using a SP8-400 uv/vis spectrophotometer (Pye/Unicam) and a matched pair of quartz cuvettes.

Nucleic sample was diluted into distilled water and optical density measured. Concentration could be determined by using the following equation:

Concentration= OD_{260} x Cuvette cell length x dilution factor x Coefficient for nucleic acid under examination

[Coefficients used: 33μg/ml for single stranded DNA, 40μg/ml for RNA and 50μg/ml for double stranded DNA].

K. OLIGONUCLEOTIDES AND PRIMERS.

The following section lists the oligonucleotides and primers used in all polymerase chain reactions (PCR) and sequencing reactions.

1. A 64 fold degenerate 23mer sense oligonucleotide was derived from the C4[1] peptide ENFQNWLK (table 2) and designated oligonucleotide 1 (5'- GGA/GAACTTTC/TCAA/GAAC/TGGC/TC/TTIAA -3'). Oligonucleotide 1 was synthesised by Genosys, Cambridge, UK and was utilised in PCR amplification reactions.

2. A 2 fold degenerate 23mer antisense oligonucleotide was derived from the last 9 residues of C4[1] peptide GASQAGMTGGYGMPR (table 2) and designated oligonucleotide 2 (5'- GGCATICCA/GTAICCIGTGATICC -3'). Oligonucleotide 2 was synthesised by Genosys, Cambridge, UK and was utilised in PCR amplification and sequencing reactions.

3. A 35mer sense primer (designated oligonucleotide 3; 5'- CTATCTAGAGAGCG AGATGATGGGCTCTTCTCTGG -3') and a 35mer antisense strand primer (designated

oligonucleotide 4; 5'- CTATCTAGAGAGCCAGAGAAGAGCCCATCATCTCG -3') were directed against the sequenced peptide region RDDGLFS of $C4^1$ (residues 139-145, see fig 11 B). XbaI restriction sites at the 5' termini were employed to clone into pUC19. Oligonucleotides 3 and 4 were utilised in PCR amplification reactions and were synthesised by Cruachem Ltd, Glasgow, UK.

4. A 35mer $dC_{(13)}$ anchor primer (5'- CTATCTAGAGAGCTCGCGGCCGCCCCC CCCCCCCC -3') directed against homopolymeric dGTP tails and containing an XbaI restriction site at the 5' terminus, was synthesised by the ICRF oligonucleotide facility and was used in all 5'- end anchor PCR amplification reactions.

5. A universal adapter primer (UAP; 5'- CTACTACTAGGCCACGCGTCGACTAC -3') directed against the adapter primer (AP; 5'- GGCCACGCGTCGACTAGTACT$_{(17)}$ -3') of the 3'- end rapid amplification of cDNA ends (RACE) kit (Gibco BRL Life Technologies), containing a 5'- termini SalI restriction site, was utilised in 3'- end RACE PCR reactions.

6. pUC19 specific primers -40 (5'- GTTTTCCCAGTCACGAC -3') and M13 reverse (5'- AACAGCTATGACCATG -3') were employed to sequence inserts cloned into pUC19 in double stranded sequencing reactions (methods, section Q. 1).

L. RT--PCR AMPLIFICATIONS. (Rappolee et al, 1991)

Internal coding region for $C4^1$ was obtained from HTCL by using RT-PCR amplifications.

Total RNA isolated from HTCL (methods, section I. 1) was batch purified on oligo dT (methods, section I. 2) and 400ng of polyA$^+$ mRNA was reversed transcribed according to manufacturer's instructions (Strategene First Strand Synthesis Kit). Centrifuged and dry precipitated polyA$^+$ mRNA (400ng) was resuspended in 32µl of DEPC treated water. 3µl of oligo dT primer (100ng/µl) was added and mixed gently. The mixture was incubated at 65^{o}c for 5 minutes and then cooled to room temperature to allow annealing of the primer to polyA$^+$ mRNA. The following reagents were then added: 5µl of 10x first strand buffer (50mM Tris.HCl (pH 8.3), 75mM KCl, 3mM $MgCl_2$), 5µl of 0.1M DTT, 1µl RNase Block I (40 units/µl), 2µl of 25mM dNTPs and 1µl of Moloney Murine Virus reverse transcriptase (20 units/µl). The components were

mixed gently and incubated at 37°c for 1 hour. The completed reaction was placed on ice or stored at -70°c. 1µl (≈8ng) of first strand synthesis product was used in 100µl reactions containing 1x PCR buffer (10mM Tris.HCl (pH 8.3),50mM KCl, 25mM $MgCl_2$, 0.01% (w/v) gelatin), 0.5mM dNTPs (dATP, dTTP, dGTP and dCTP. Boeringher Mannheim), 1µg each of primers (oligonucleotides 1 and 2; see methods, section K), 2.5units of AmpliTaq® (Perkin Elmer/Cetus) (Saiki et al, 1988 and Erlich, 1989) and overlayed with 100µl of paraffin oil (BDH). Amplification was carried out in a Hybaid Thermal Cycler, with an initial denaturation step at 97°c for 3 minutes, followed by 30 cycles of primer annealing at 47°c for 1 minute, extension at 72°c for 1 minute and denaturation at 95°c for 1 minute. Control reactions without DNA were always included. 10% of the product was visualised on an a 1% agarose gel (methods, section O), while the remainder was cloned into pUC19 (methods, section P) and sequenced (methods, section Q).

M. 5'- END ANCHOR PCR AMPLIFICATIONS.

The 5'- end of the coding region of $C4^l$ was obtained from HTCL was based on methods described by Moss et al (1991) Frohman et al (1988), and Dr. L. Wedderburn (ICRF, London; personal comm).

First strand synthesis of HTCL was carried out as described above in section L. Excess oligo dT primer was removed from first strand synthesis products: 50µl of DEPC treated water was added to give a final volume of 100µl. To this, 100µl of 4M ammonium acetate and 3 volumes of absolute ethanol was added, mixed vigorously and allowed to precipitate at room temperature for 10 minutes. The mixture was centrifuged at 12000x g for 30 minutes at room temperature in a microcentrifuge (MSE). Resultant supernatants were discarded and pellets air-dried for several minutes. Once dry, pellets were dissolved in 100µl of distilled water. Dissolved cDNA was then ammonium acetate/ethanol precipitated (as above) a further 3 times. After the final precipitation pellets were washed in 70% (v/v) ethanol and air-dried. Resultant cDNA was dissolved in 10µl of distilled water and allowed to dissolve for several minutes at room temperature. Homopolymeric dGTP tailing of the above cDNA was carried out by adding 4µl of 5x tailing buffer (1M potassium cacodylate (Boeringher Mannheim), 125mM Tris.HCl, 1.25mg/ml BSA, pH 6.6), 100pMol of freshly opened dGTP (Boeringher

Mannheim) and 25 units of terminal transferase (Boeringher Mannheim). Tailing reactions were incubated at 37°c for 15 minutes. 80µl of distilled water was added and then extracted with an equal volume of phenol:chloroform (1:1) and then an equal volume of chloroform. To the aqueous layer, 100µl of distilled water was added. 16µl of the product of a tailing reaction was used in 100µl reactions with 1x PCR buffer (10mM Tris.HCl (pH 8.3), 50mM KCl, 2.5mM $MgCl_2$, 0.01% (w/v) gelatin), 0.5mM dNTPs (dATP, dTTP, dGTP and dCTP. Boeringher Mannheim), 25µM each of primers (oligonucleotide 4 and $d(C)_{13}$ anchor primer, see methods, section K), 2.5 units of AmpliTaq® and overlayed with 100µl of paraffin oil (BDH). Amplification was carried out in a Hybaid Thermal Cycler. An initial cycle of denaturation at 95°c for 5 minutes, primer annealing at 56°c for 45 seconds and extension at 72°c for 5 minutes was carried out. This was followed by a further 30 cycles of primer annealing at 56°c for 45 seconds, extension at 72°c for 1.5 minutes and denaturation at 95°c for 1 minute. Control reactions without DNA or non-tailed cDNA were always included. 10% of PCR products were visualised on 1% agarose gels (methods, section O), whilst the remainder was cloned into pUC19 (methods, section P) and sequenced (methos, section Q).

N. RAPID AMPLIFICATION OF cDNA ENDS (RACE). (Frohman et al, 1988)

The 3'- end of the coding region for $C4^l$ was obtained from HTCL by use of a 3'-RACE system (Gibco BRL Life Technologies). A control RNA was included in the 3'-RACE system as an aid in verifying that the system performed correctly. 400ng of HTCL $polyA^+$ selected mRNA (methods, section G) was diluted into 13µl of DEPC treated distilled water, in a 0.5ml mirocentrifuge tube. 500nM of adapter primer (AP) (methods, section K) was added, mixture heated to 65°c for 10 minutes and chilled on ice for 2 minutes. Tube contents were collected by brief microcentrifugation and 2µl of 10x synthesis buffer (200mM Tris.HCl (pH 8.4), 500mM KCl, 25mM $MgCl_2$, 1mg/ml BSA), 1µl of 10mM dNTP mix (dATP, dCTP, dGTP, dTTP) and 2µl of 0.1M DTT were added. Mixture was gently vortexed and equilibrated to 42°c for 2 minutes. 200 units of Superscript™ reverse transcriptase was added mixture incubated at 42°c for 30 minutes. Reactions were placed on ice, 2 units of *E. coli* RNase H added and incubated for 10 minutes at 42°c. First strand cDNA obtained above was amplified directly by PCR. To a

50µl PCR reaction mix (final concentrations: 20mM Tris.HCl (pH 8.4), 50mM KCl, 2.5mM $MgCl_2$, 100µg/ml, 200µM each of four dNTPs, 0.5µg of oligonucleotide 1 and 200nM of UAP (methods, section K), 2.5units of AmpliTaq®), 2µl of first strand cDNA was added in a 0.5ml microcentrifuge tube. Reactions were overlayed with 75µl of paraffin oil (BDH). Amplifications were carried out in a Hybaid Thermal Cycler, with an initial inactivation/denaturation step of 94°c for 7 minutes, followed by 30 cycles of primer annealing at 56°c for 45 seconds, extension at 72°c for 2 minutes and denaturation at 94°c for 1 minute. A further extension of 72°c for 15 minutes was carried out after the last cycle of PCR. 10% of PCR product was visualised on a 1% agarose gel (methods, section O).

AP was designed to synthesise first strand cDNA from all polyadenylated mRNA; the sequence specificity in PCR amplifications was therefore derived solely from the gene specific oligonucleotide 1 (methods, section K). A second 'nested' gene specific oligonucleotide 3 (methods, section K) was therefore utilised in conjunction with UAP in a second amplification reaction to give the 3'- RACE procedure the specificity of a second primer. 1% of the above PCR product obtained from oligonucleotide 1 and UAP was added to an identical PCR reaction mix as above, except 0.5µg of oligonucleotide 1 was substituted with 200nM of oligonucleotide 3. Identical reaction conditions were performed as above. Control reactions without DNA were always included. 10% of product was visualised on 1% agarose gels (methods, section O), whilst the remainder was cloned into pUC19 (methods, section P) and sequenced (methods, section Q).

O. AGAROSE GELS.

1% agarose (FMC) gels were utilised for visualisation of nucleic acid. 40ml gels (0.4g of agarose melted in 40 ml 1x TAE (0.04M Tris.Acetate, 0.001M EDTA) and 5µg/ml ethidium bromide) were cast at approximately 1cm thick into Pharmacia GNA-100 gel electrophoresis apparatus. 1x TAE completely covered the gel which was typically prerun at 40volts before sample loading. Samples were loaded after mixing with 1/6 volume of type I loading buffer (6x: 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 40% (w/v) sucrose in distilled water). Gels were typically run at 40

volts until the first marker dye had migrated at least 8cm from the wells. 1µg of 123 ladder or 1kb ladder molecular weight markers (both Gibco BRL Life Technologies) was loaded in all cases. Larger 1% agarose gels were cast (100ml in volume) in a Horizontal Gel H5 (Gibco BRL Life Technologies) inorder to visualise many samples. Visualisation was carried out on a uv source and photographed using a Polaroid DS34 camera.

P. CLONING OF PCR PRODUCTS.

Unless specified, molecular cloning methods were performed essentially as described in Sambrook et al (1989).

1. Glassmilk Purification of Double Stranded DNA

This method of DNA purification was carried out by use of the GENECLEAN II® Kit, BIO 101 Inc.

To a solution of DNA in a 1.5ml mirocentrifuge tube, or an excised band from an agarose gel, 3 volumes of 6M sodium iodide solution was added (excised bands were incubated for 5 minutes at 50°c to melt the agarose). GLASSMILK® suspension was added (generally 5µl for DNA solutions containing 5µg or less) and incubated for 5 minutes on ice. Glassmilk/ DNA complex was pelleted for 5 seconds in a microcentrifuge at 12000x g. Pelleted glassmilk/DNA complex was thoroughly 3x with NEWWASH (NaCl/ethanol). DNA was eluted into the desired volume of water by incubating washed glassmilk/DNA complex with eluent for 5 minutes at 50°. Glassmilk was pelleted and eluted DNA removed carefully avoiding any contamination from glassmilk. DNA purification by this method completely removed excess primers, enzyme and low molecular double stranded DNA (glassmilk purification was employed for DNA above 0.4Kb without loss of efficiency).

2. Ligations

PCR products and pUC19 vector (Gibco BRL Life Technologies) were digested with appropriate restriction enzymes.

Inorder to blunt-end PCR products, to the insert DNA (4µl of one glassmilk purified PCR product), 1.5µl of 10x T4 polynucleotide buffer (500mM Tris.HCl (pH 7.6), 100mM $MgCl_2$, 50mM DTT, 1mM spermidine, 1mM EDTA), 5 units of Klenow fragment of *E. coli* DNA polymerase I (Amersham), 1.5µl of 10x nucleotide mix

(450μM dGTP, 2mM dCTP, 120μM dATP, 1.4mM dTTP. Boeringher Mannheim) and distilled water up to 15μl was added. After incubation for 1 hour at 37°c, 5 units of T4 polynucleotide kinase (Pharmacia) was added with 1μl of 10mM ATP (Pharmacia). Mixture was further incubated for 1 hour at 37°c. Following incubation mixture was extracted in an equal volume of phenol:chloroform (1:1) and precipitated at -20°c for 16 hours by adding 1/10 volume 5M sodium acetate, pH 5.2, and 2.5 volumes of absolute ethanol. Precipitated insert was collected by microcentrifugation for 20 minutes, air-dried pellet dissolved into 6μl of distilled water and used in all blunt-ended ligation reactions. To 1μg of pUC19 vector, 10 units of SmaI (Gibco BRL Life Technologies), 1μl of Gibco BRL Life Technologies 10x REACT 4 (200mM Tris.HCl (pH 7.4), 50mM $MgCl_2$, 500mM KCl) and distilled water up to a total of 10μl was added and mixture incubated at 37°c for 4 hours. Resultant digested pUC19 was glassmilk purified and a small fraction visualised on a 1% agarose gel (methods, section O) to identify that pUC19 had completely digested. 500ng of SmaI digested pUC19 was phosphatased by adding 0.1 units of calf alkaline phosphatase (Pharmacia), 5μl of One-Phor-All Plus buffer (Pharmacia) and distilled water up to a volume of 50μl. Mixture was incubated for 30 minutes at 37°c and then enzyme heat-inactivated for 15 minutes at 85°c. 100ng of SmaI digested, phosphatased and glassmilk purified (methods, section P. 1) pUC19 was used in all blunt-ended ligation reactions.

Cohesive ends were generated on PCR products by digesting with either XbaI (Gibco BRL Life Technologies) or double digesting with XbaI/SalI (Gibco BRL Life Technologies) restriction enzymes. To insert DNA (4μl of one glassmilk purified PCR product) 10 units of XbaI, 1μl of Gibco BRL Life Technologies 10x REACT 2 (500mM Tris.HCl (pH 8.0), 100mM $MgCl_2$, 1M NaCl) and distilled water up to a volume of 10μl was added. After incubation at 37°c for 4 hours glassmilk purified (methods, section P. 1) product was either stored at -20°c or further digested in 10 units of SalI, 1μl of Gibco BRL Life Technologies 10x REACT 10 (1M Tris.HCl (pH 7.6), 100mM $MgCl_2$, 1.5M NaCl) and distilled water up to volume of 10μl at 37°c for 4 hours, glassmilk purified (methods, section P. 1) and 6μl used in cohesive ligation reactions. 1μg of XbaI or

XbaI/SalI cohesive pUC19 was generated and 500ng phosphatased as above and 100ng used in ligation reactions.

In ligation reactions, 2µl of either blunt or cohesive pUC19 (100ng) was incubated with 6µl of either blunt or cohesive glassmilk purified (methods, section P. 1) PCR insert. To vector/insert mixtures 1µl of 10x ligation buffer (either blunt-end: 0.66M Tris.HCl (pH 7.6), 50mM $MgCl_2$, 50mM DTT, 1mg/ml BSA, 10mM hexamineacolbalt, 2mM ATP, 5mM spermidine-HCl; or cohesive: Pharmacia One-Phor-All Plus buffer) was added. In addition, 1mM ATP was present in all cohesive ligation reactions. Distilled water was added to a final volume of 10µl. Ligation reactions were heated at 45°c for 10 minutes and quenched on ice for 2 minutes. 10 units of T4 ligase (Pharmacia) was added to blunt-end ligations, whilst 4 units of T4 ligase was added to cohesive ligations. All ligation reactions were incubated at 10°c for 20 hours. Completed ligation reactions were stored at -20°c for up to 1 month. In all cases control ligation reactions of unphosphatased pUC19 and no DNA insert present were performed.

3. Transformations

Transformations were carried out using Subcloning Efficiency DH5α™ Competent Cells (Gibco BRL Life Technologies, Lot Nº DKU 101).

Competent cells were removed from -70°c freezer and thawed on wet ice. The required number of autoclaved 1.5ml microcentrifuge tubes were also placed on wet ice. The competent cells were gently mixed and 50µl aliquoted into chilled microcentrifuge tubes. To determine transformation efficiency 5µl (0.5ng) control pUC19 was added to one tube containing 50µl competent cells and gently mixed. 3µl (≈10ng) of each DNA ligation reaction was added directly to an appropriately labelled tube containing 50µl competent cells and mixed gently and incubated on ice for 30 minutes. Cells were heat-shocked for 20 seconds at 37°c with no shaking and then returned to ice for 2 minutes. 0.8ml of room temperature SOB medium (2% (w/v) bacto-tryptone (Difco), 0.5% (w/v) bacto-yeast extract (Difco), 1% (w/v) NaCl, 2.5mM KCl, pH 7.0, with 20mM glucose and 20mM $MgCl_2$ added after autoclaving) was added and incubated for 45 minutes at 37°c in a Gallenkamp orbital shaker set at 225 rpm. 50µl and 200µl of each undiluted transformed culture was spread onto SOB agar plates (SOB as above, plus 1.5% (w/v) agar (Difco)) containing 50µg/ml ampicillin and incubated overnight at 37°c in a Genlab

Midi 2/55 incubator. A control transformation where no DNA was added to competent cells was performed to check sterility. Transformation efficiency (CFU/µg) could be calculated using the following equation:

$$\text{CFU/µg} = \frac{\text{CFU in control plate}}{\text{ng pUC19 used in transformation}} \times \frac{1 \times 10^6 \text{ ng}}{\text{µg}} \times \text{Dilution factor}$$

4. Mini Preparation of Plasmid DNA

Mini-preparations of plasmid DNA was carried out on 3ml overnight LB (1% bacto-tryptone, 0.5% bacto-yeast extract (both Difco), 1% (w/v) NaCl) liquid cultures containing 50µg/ml ampicillin of single colonies generated from transformations. 1.5ml of overnight cultures were centrifuged at 12000x g for 30 seconds in a microcentrifuge. Pelleted cells were resuspended in 100µl of solution I (50mM glucose, 10mM EDTA, 25mM Tris.HCl, pH 8.0) and incubated for 5 minutes at room temperature. 200µl of solution II (0.2M NaOH, 1% (w/v) SDS) was quickly added, mixed by gentle inversion and incubated on ice. Ice-cold solution III (3M potassium acetate, 2M glacial acetic acid, pH 4.5) was added, quickly vortexed and placed on ice for 5 minutes. Mixture was centrifuged at 12000x g for 10 minutes at room temperature in a microcentrifuge and supernatant extracted with an equal volume of phenol:chloroform (1:1). To aqueous phase 1ml of 95% (v/v) ethanol was added, vigorously and incubated for 1 minute at room temperature. Precipitated plasmid DNA was pelleted by microcentrifugation at 12000x g for 2 minutes at room temperature. Air-dried pellet was washed in 70% (v/v) ethanol and after a final air-drying pellets were dissolved in 50µl of TE (10mM Tris.HCl (pH 8.0), 1mM EDTA (pH 8.0)) plus 20µg/ml RNAase (Sigma). 5µl of dissolved plasmid DNA was digested with appropriate restriction enzymes and inserts visualised on 1% (w/v) agarose gels (methods, section O). Insert containing cultures were frozen in 7% (v/v) DMSO by plunging into liquid nitrogen and storing at -70°c.

5. Midi Preparation of Plasmid DNA

Midi-preparations of plasmid DNA was carried out on 50ml overnight LB liquid cultures containing 50µg/ml ampicillin of an insert-positive single colony, derived from transformations. Cells were pelleted by centrifugation at 2200x g in a JA17 rotor mounted within a J2-21 Beckman centrifuge for 10 minutes at 4°c. Pelleted cells were

resuspended in 2ml of solution I and 4ml of solution II. Cells were mixed gently and stored on ice for 5-10 minutes. 3ml of solution III was added and after gentle vortexing incubated on ice for 10 minutes. Mixture was centrifuged at 16,700x g for 15 minutes at 4°c in a JA17 rotor mounted within a J2-21 Beckman centrifuge. Nucleic acid was precipitated in 0.6 volumes of isopropanol (BDH) at room temperature for 10 minutes. Centrifugation at 16,700x g was repeated as above. Nucleic acid pellet generated by centrifugation was dissolved in 2ml of TE. Dissolved nucleic acid was transferred to a 15ml corex tube and 3ml of ice-cold 5M LiCl added. Mixture was centrifuged at 16,700x g for 10 minutes at 4°c (as above) and resultant supernatant transferred to a fresh 15ml corex tube, to which an equal volume of isopropanol was added. Resultant pellet generated after repeated centrifugation, was washed in 70% (v/v) ethanol, dissolved in 0.5ml TE containing 20µg/ml RNAase, transferred to a 1.5ml microcentrifuge tube and incubated at room temperature for 30 minutes. An equal volume of 1.6M NaCl containing 13% (w/v) polyethylene glycol (PEG) 8000 (Sigma) was added and vortexes extensively. Plasmid DNA was recovered by microcentrifugation at 12000x g for 5 minutes at room temperature. Pellets were dissolved in TE and extracted in an equal volume of phenol, followed by phenol:chloroform (1:1) and finally chloroform. To final aqueous layer of 4M ammonium acetate and 2 volumes of absolute ethanol was added and incubated for 10 minutes at room temperature. Nucleic acid was recovered by microcentrifugation at 1200x g for 5 minutes. Pellets were washed in 70% (w/v) ethanol, air-dried and dissolved in TE. Concentration of dissolved plasmid DNA was determined by optical density at (methods, section J).

Q. DIRECT DOUBLE-STRANDED SEQUENCING. (Sanger et al, 1977)

1. Reactions

All sequencing reactions were carried out according to manufacturer's instructions for the Sequenase™ Version 2.0 DNA Sequencing Kit (USB). 5µg of insert-containing plasmid was purified on glassmilk (methods, section P. 1) , and eluted into 9µl of of distilled water plus 1µl (0.1µg) of sequencing primer (-40, M13 reverse or oligonucleotide 2, see methods, section K), heated to 100°c for 5 minutes then placed directly into liquid nitrogen. Tubes were thawed to room temperature (for 5 minutes) as

required and 1µl of 0.1M DTT, 2µl of diluted GTP labelling mix (diluted 1:3 in distilled water), 0.5µl of [^{35} S] dATP (25µCi/µl, Amersham), 2µl of Sequenase buffer (5x concentrated: 200mM Tris.HCl (pH 7.5), 100mM $MgCl_2$, 250mM NaCl) and 9 units of Sequenase polymerase added. The resultant labelling reactions were extensively mixed and incubated at room temperature for no longer than 5 minutes. Four 0.5ml microcentrifuge tubes, labelled G, A, T and C were on hand. 2.5µl of ddGTP termination mix (80µM dGTP, 80µM dATP, 80µM dCTP, 80µM dTTP, 8µM ddGTP and 50mM NaCl) was added to the tube labelled G. Similarly, 2.5µl of ddATP, ddTTP, ddCTP termination mixes were placed into tubes A, T and C respectively. Tubes were capped to prevent evaporation and prewarmed at 40°c for at least 1 minute. When the labelling reaction was completed, 4µl was removed an transferred to the tube labelled G. Contents were mixed and incubation continued at 40°c for 10 minutes. Similarly 4µl of the labelling reaction was added to the A, T and C tubes, mixed and returned to 40°c incubation for 10 minutes. Solutions were collected to the bottom of the tubes by brief microcentrifugation. 4µl of stop solution (95% (v/v) formamide, 20mM EDTA, 0.05% (w/v) bromophenol blue, 0.05% (w/v) xylene cyanol FF) was added to completed termination reactions and mixed thoroughly and stored at -20°c until ready to load onto sequencing gels.

2. Denaturing Gel Electrophoresis

To 120ml of 6% gel mix (Sequagel 6, National Diagnostics), 0.96ml of 10% APS was added. Gel mix was carefully poured between taped together, IMS cleaned, gel plates (30x 40cm, using 0.4-1.2mm wedge spacers and 0.4mm thick, 24 lane sharkstooth combs), air bubbles were avoided. Polymerising gels were typically matured for 16 hours. Tape was removed from the bottom of the plates to allow the bottom of the gel to be in complete contact with sequencing running buffer (1x TBE: 0.089M Tris base, 0.089M boric acid (Sigma), 2mM EDTA) when assembled into a Gibco BRL Life Technologies Model S2 sequencing apparatus. Gels were prerun for 1 hour at 60 watts. When gels were ready for loading, samples were heated to 85°c for 2 minutes and 3µl of each termination mix loaded into the corresponding lane. Each sample was loaded in a pattern (A, G, C and T) that allowed ease of reading closely spaced bands. 6% gel runs for 6 hours at 50 watts were necessary for reading in the 80-350bp range.

Following electrophoresis for the desired time, gels were fixed in 5% acetic acid, 15% methanol for 60 minutes to remove urea. Fixed gels were pressed onto 3mm CHR Whatman paper, covered in Saran-wrap (DOW) and vacuum dried at a moderate temperature (80°c) for 1 hour in a Biorad Model 483 Slab Dryer to preserve resolution. Exposure of gels was performed with direct contact between dried gels and fast Hyperfilm-MP (Amersham) film for approximately 36 hours. Sequence was read from developed film by starting at the bottom of the gel and moving upwards recording the order of occurrence of bands in the A, G, C and T columns. Gels read in this manner yielded sequence in 5'- to -3' orientation from the priming site.

R. COMPUTER ANALYSIS OF SEQUENCE DATA.

The nucleotide and amino acid sequences of $C4^1$ were analysed using DNA Strider and Gene Jockey (Apple) and the Intelligenetics suite of programs on a VAX workstation. Nucleotide sequences were analysed and manipulated using SEQ, amino acid sequences were handled using PEP. Database searches (Swiss Prot release 24; EMBL release 33; UGenBank 75_33 and PIR release 31) were performed using FASTDB and IFIND (Wilbur &Lipman, 1983; Gribskov et al, 1987). Multiprediction program (Zvelebil et al, 1987) was employed to calculate secondary structure (Garnier et al, 1978).

S. RANDOM PRIMER LABELLING OF cDNA (Feinberg & Vogelstein, 1983).

Restriction fragments were separated on a 1% (w/v) low melting point agarose (FMC) gel in 1x TAE (methods, section O). The desired fragment was excised cleanly with a razor blade and weighed in a 1.5ml microcentrifuge tube and glassmilk purified (methods, section P. 1). eluted into 32µl of distilled water and stored at -20°c until it was required.

For labelling, DNA fragment was boiled for 5 minutes and cooled on ice for 2 minutes. The following reaction mixture was set up: 10µl of oligolabelling buffer (OLB), 2µl of BSA (10mg/ml), 5µl α-^{32} P- dATP (2000-4000 Ci/mM) (Amersham), 32µl of eluted DNA fragment, distilled water to a final volume of 50µl and 2 units of Klenow fragment of *E. coli* polymerase I (Amersham). Incubation was carried out at room

temperature for at least 5 hours for 30-50ng of template. Reaction was terminated by addition of 0.2ml of 20mM NaCl, 20mM Tris.HCl (pH 7.5), 2mM EDTA, 0.25% (w/v) SDS, 1μM dATP. Unincorporated labelled nucleotide was removed by use of a Biogel P6 spun column (Biorad). P6 column was firstly drained and the entire terminated labelling reaction loaded to the top of the column. Loaded P6 column was centrifuged at 1100x g for 4 minutes. Labelled, centrifuged, fragment constituted a probe, and was boiled for 5 minutes before use.

T. GEL ELECTROPHORESIS OF RNA, NORTHERN BLOTTING AND HYBRIDISATION.

Gel (usually 1% (w/v)) was prepared by melting high gelling temperature agarose (FMC) in distilled water and cooling to 55°c before adding Morpholinopropanesulphonic acid buffer (MOPS buffer) to 1x and formaldehyde (BDH) to 2.2M final concentrations. RNA samples (all polyA$^+$ selected mRNA; methods, section G. 2) up to a concentration of 0.1μg/μl in 1x MOPS buffer, 50% (v/v) deionised formamide (BDH) and 2.2M Formaldehyde. Samples were incubated at 55°c for 15 minutes, RNA sample buffer was added to a final concentration of 1x and then samples were loaded onto a gel (usually 22μl; ie. ≈2μg of mRNA). Electrophoresis was performed at 30-40 volts for 5 hours in 1x MOPS buffer using a Gibco BRL Life Technologies Horizontal gel, model H5. Following electrophoresis the gel was washed for 5 minutes in distilled water. After washing in distilled water, RNA was transferred to 0.45μM pore size nitrocellulose (Schleicher & Schuell GmbH Dassel, FRG). Nucleic acid transfer was carried out using the Hybaid Vacuum Transfer System. Into the apparatus, 3mm-CHR Whatman paper presoaked in 20x SSC was overlayed with similarly presoaked nitrocellulose cut to the same size as the gel. A rubber mask was placed over the nitrocellulose, and a section was removed so as to expose the nitrocellulose to the gel when carefully placed over it. Onto the gel several mls of 20x SSC was layered, and this represented the transfer solution. Upon vacuum generation, transfer solution was continuously added to the gel as levels dropped. Transfer was carried out for at least 1 hour. Following transfer the nitrocellulose filter was air-dried and bake-fixed (80°c) in an EM scope Laboratories Ltd oven for 1 hour.

Bake-fixed nitrocellulose filters (blots) were presoaked in 2x SSC and then transferred to prehybridisation buffer (2x SSC, 0.1% (w/v) SDS, 5x Denhardt's solution, 50μg/ml denatured Homo Mix), usually 0.5ml of prehybridisation buffer was used per $1cm^2$ of nitrocellulose. Prehybridisation was carried out for 2-16 hours at 65°c. Fresh prehybridisation buffer was added to the blots and prepared, boiled probe (methods, section S) added. Addition of labelled probe to prehybridisation buffer constituted hybridisation. Hybridisation was carried out for 24 hours at 65°c in a Gallenkamp shaking water bath. Blots were then washed at 65°c with increasing stringency from 2x SSC/0.1% (w/v) SDS to 0.5x SSC/0.1% (w/v) SDS. Following washing, blots were covered in Saran-wrap (DOW) and exposed to fast Hyperfilm-MP (Amersham) at -70°c. Varying autoradiographic exposure times were carried out to maximise visualisation of bands. After autoradiography, blots were stripped by gently boiling them in stripping solution (0.05x SSC/0.1% (w/v) SDS) for 15 minutes. Stripped blots were hybridised to an α–^{32}P labelled (methods, section S) actin cDNA insert (30ng) common to all actin messages (Sambrook et al, 1989) as described above.

RESULTS

A. BACKGROUND

Protein C4 is an actin associated polypeptide doublet, originally designated $C4^h$ and $C4^l$, which is present in all cells and tissues apart from skeletal muscle and erythrocytes (Shapland et al, 1988). The higher molecular weight polypeptide ($C4^h$) is absent in motile cells such as lymphocytes and is down-regulated when normal mesenchymal cells are (a) transformed by DNA and RNA viruses, or (b) switched to non-adherent culture conditions (Shapland et al, 1988). This protein has now been purified and shown to induce the rapid gelation of actin filaments (Shapland et al, 1993). The cDNA encoding $C4^h$ has been sequenced (Prinjha et al, 1994). Since this molecule is transformation sensitive and gels actin it has been named transgelin (Shapland et al, 1993). In contrast, the lower molecular weight polypeptide is apparently unaffected by transformation and is always present in motile cells such as lymphocytes (Shapland et al, 1988).

Since both protein C4 polypeptides are recognised by a monoclonal antibody they must contain at least one homologous region. Initial attempts to isolate the gene encoding for the $C4^l$ polypeptide were carried using a spectra of oligonucleotides derived from transgelin. Since these did not recognise cDNA from tissue which only expresses $C4^l$, I therefore obtained oligonucleotides based directly on amino acid sequence data from partially purified $C4^l$.

B. IMMUNOFLUORESCENCE

Immunofluorescence studies using the anti-C4 monoclonal antibody (methods, section D. 1) were carried out on SV40 transformed 3T3 fibroblasts, cells known to express only the $C4^l$ polypeptide (results section C; fig 2 g). Studies showed uniformly stained linear actin filament stress fibre bundles (fig 1 b and d). Both cell processes and diffuse actin networks, which are present in these cells, are also stained with anti-C4 (fig 1 a and c). Nuclear staining can also be detected in many of the cells observed (fig 1 d).

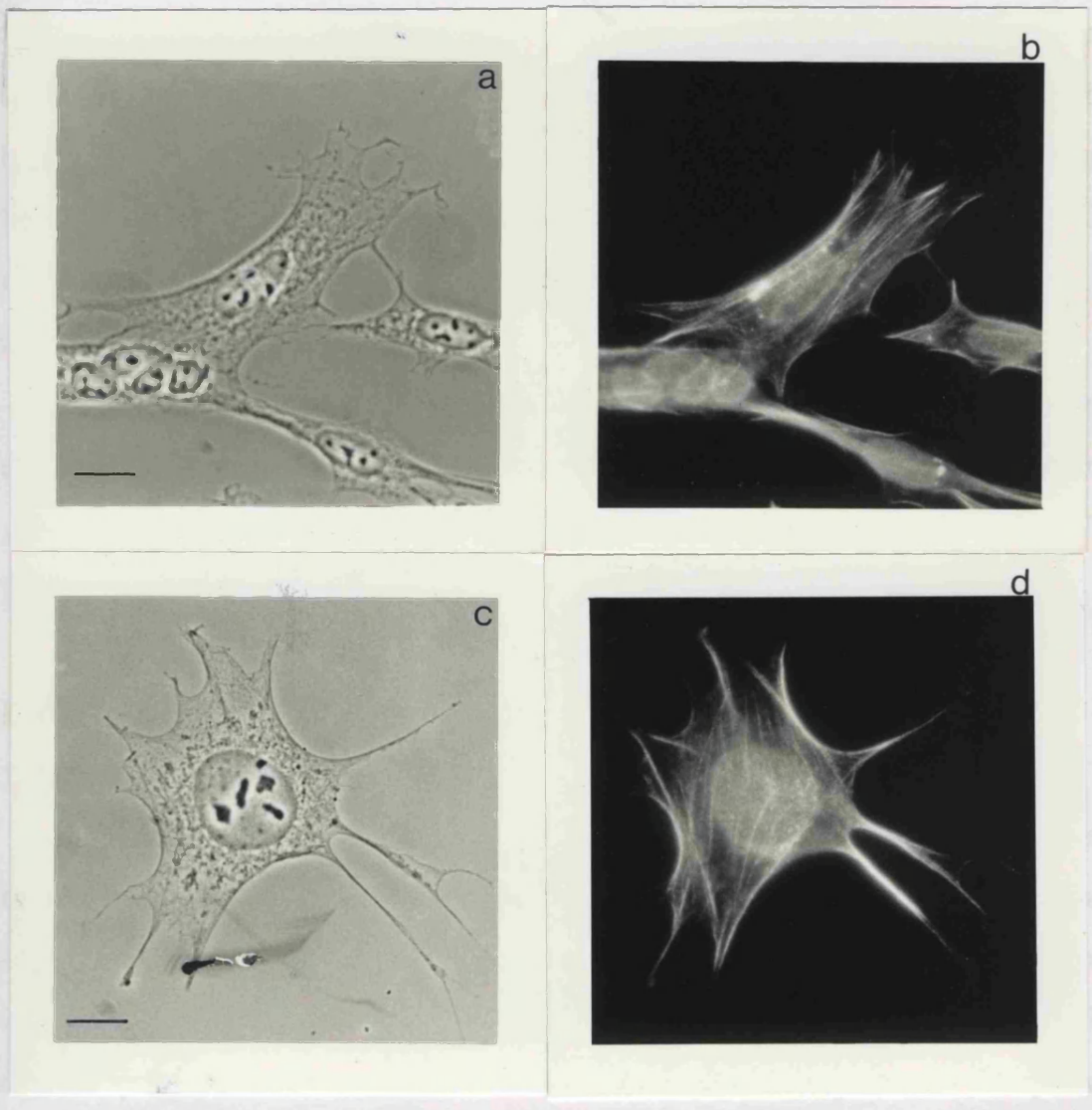


FIGURE 1. Immunofluorescence localisation in SV40 transformed 3T3 fibroblasts of C4[1]. Cells are shown under phase constrast (**a** and **c**). Immunofluorescence of the same cells shows the superposition of anti-C4 monoclonal antibody (**b** and **d**). Uniformly stained linear actin filament stress fibre bundles can be seen (**b**) and in many cells, nuclear staining is observed (**d**). Bar= 10μm.

C. TISSUE SELECTION FOR THE PURIFICATION OF C4l

SDS-PAGE and immunoblot analysis of human T cell lymphoma (HTCL) and SV40 transformed 3T3 fibroblasts using an affinity purified rabbit anti-C4 polyclonal antibody (Shapland et al, 1993; methods, section D. 2) revealed that C4l was present at high levels at a molecular weight of approximately 21KDa (fig 2 c, e and g). Immunoblotting of rat embryo fibroblasts (REF) indicated the molecular weights of C4l and transgelin which are present in approximately equal amounts and are separated by 0.5-1 Kd in these cells (fig 2 c).

From these data HTCL were chosen for the purification of C4l. These cells expressed higher levels of C4l than both activated T cells and thymus. They were easy to grow in culture, growing rapidly and harvested without trypsinisation (unlike SV40 transformed 3T3 fibroblasts).

D. PURIFICATION OF C4l

1. Detergent Solubility of C4l

Figure 3 shows C4l solubility in the zwitterionic detergent CHAPS. HTCL cells lysed in wash buffer + 1% CHAPS (fig 3 a), and the subsequent pellet (fig 3 c) and supernatant (fig 3 e), derived after centrifugation were analysed by SDS-PAGE and immunoblotting with monoclonal anti-C4 antibody. Millipored (0.45μM) HTCL 1% CHAPS lysate was also analysed to show loss of C4l yield during the millipore process (fig 3 g). Figure 3 shows that the majority of C4l was solubilised after HTCL lysis in 1% CHAPS.

2. Proteolysis of C4l

To minimise proteolytic degradation a cocktail of proteolytic inhibitors was used during all steps of purification. To show that these were effective in blocking C4l degradation I dialysed an HTCL 1% CHAPS lysate for 50 hours at 4°c in the presence or absence of the proteolytic inhibitors listed in table 1 and analysed C4l degradation over time by immunoblotting with the anti-C4 monoclonal antibody (fig 4).

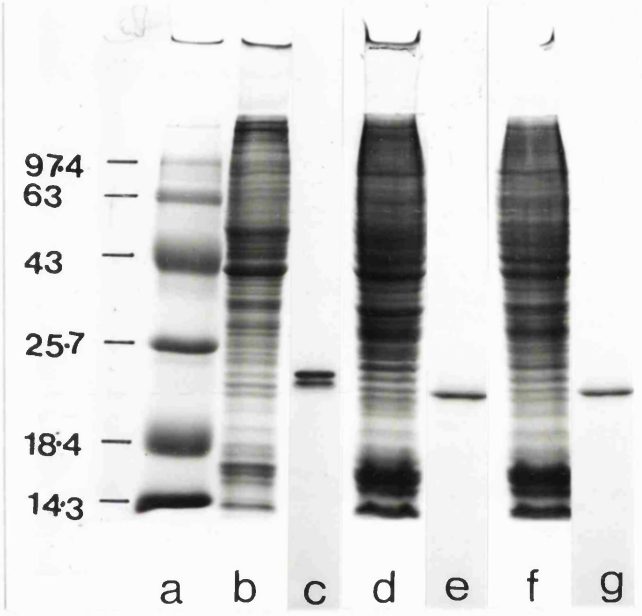


FIGURE 2. Immunoblot analysis of proteins transferred to nitrocellulose membranes labelled with a rabbit anti-C4 polyclonal antibody. Molecular weight markers are shown (**a**) with weights in kDa on the left. $C4^1$ and transgelin are present (in approximately equal amounts) when total REF lysate (**b**) is transferred to nitrocellulose (**c**). Total HTCL lysate (**d**) expresses only the lower molecular weight $C4^1$ (**e**). Identical results are seen for total SV40-3T3 cell lysate (**f**) with transgelin completely down-regulated (**g**).

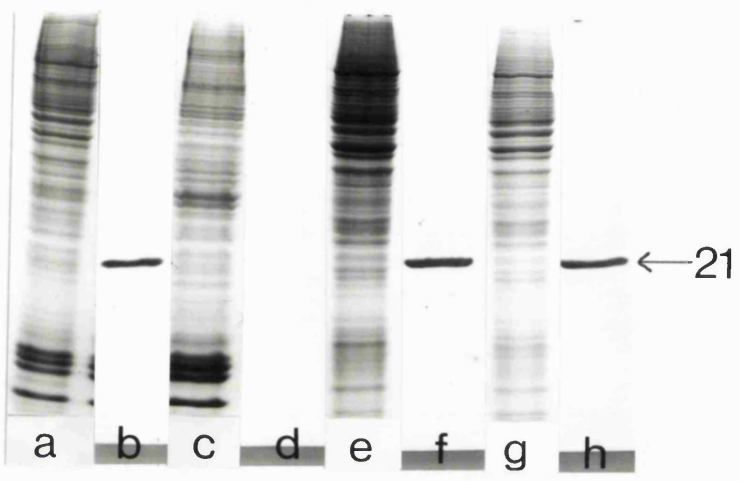


FIGURE 3. Detergent solubility of $C4^l$: Total HTCL cells (**a**) were lysed in wash buffer + 1% CHAPS and the $C4^l$ population shown when the lysate was transferred to nitrocellulose (**b**) and probed with the anti-C4 monoclonal antibody. The subsequent pellet derived from centrifugation (**c**) is shown to contain no detectable population of $C4^l$ when transferred to nitrocellulose (**d**). The resultant supernatant derived from the same centrifugation process (**e**) is seen to contain a $C4^l$ population when transferred to nitrocellulose (**f**). There appears to be no significant loss of the $C4^l$ population when a HTCL lysate centrifugation supernatant is 0.45μm millipored (**g**) and transferred to nitrocellulose (**h**). The arrow on the right indicates a molecular weight of 21kDa and the presence of $C4^l$.

Table 1. Protease Inhibitors. (From Harris & Angal, 1989)

Inhibitor	Target	Effective Working Concentration	Stock Solution
Chymostatin (Sigma)	Chymotrypsin	2ug/ml	5mg/ml in DMSO (-20°c)
E DTA (Sigma)	Metalloproteases	0.2mM	0.5M in H_2O pH 8.0 (RT)
Leupeptin (Sigma)	Serine and Thiol-proteases	2ug/ml	5mg/ml Sterile H_2O (-20°c)
Pepstatin A (Sigma)	Acid Proteases	2ug/ml	5mg/ml in methanol (-20°c)
PMSF (Sigma)	Serine Proteases	0.2mM	100mM in Isopropanol

Figure 4 illustrates that $C4^1$ maintains its integrity for up to 50 hours when the proteolytic inhibitors are present. As the proteolytic inhibitor cocktail prevented $C4^1$ degradation during cell lysis and dialysis, it was therefore utilised during all subsequent steps of purification at the same working concentration.

3. Chromatofocusing

Utilisation of a pH 7-9.6 chromatofocus column, employed directly after 1% CHAPS lysis and dialysis in start buffer, extensively separated an HTCL lysate of typically 5x 10^9 cells. This analysis revealed that $C4^1$ eluted from the chromatofocus within a pH range of 8.0-8.2 in a concentrated peak obtained in 8 ml. SDS-PAGE analysis of 0.5 ml fractions within this peak are shown in figure 5. Four fractions within this peak (fig 5 h, i, k and l), show $C4^1$ as being the major protein in a fraction containing 11 visible bands. These fractions were pooled for further purification.

4. Hydroxylapatite Purification

A final purification and concentration step of the pooled chromatofocus fractions positive for $C4^1$ was carried out on a small (3ml) hydroxylapatite column. Elution produced a preparation containing a band present in three fractions of 0.5 ml each (fig 6 f, h and j). The bands appeared to be higher than 90% homogenous with no apparent degradation as shown by SDS-PAGE and immunoblotting (fig 6). These bands migrated at the same position in immunoblots as $C4^1$ (seen previously in results, section C). By pooling the three fractions approximately 10μg of > 90% homogenous $C4^1$ was obtained.

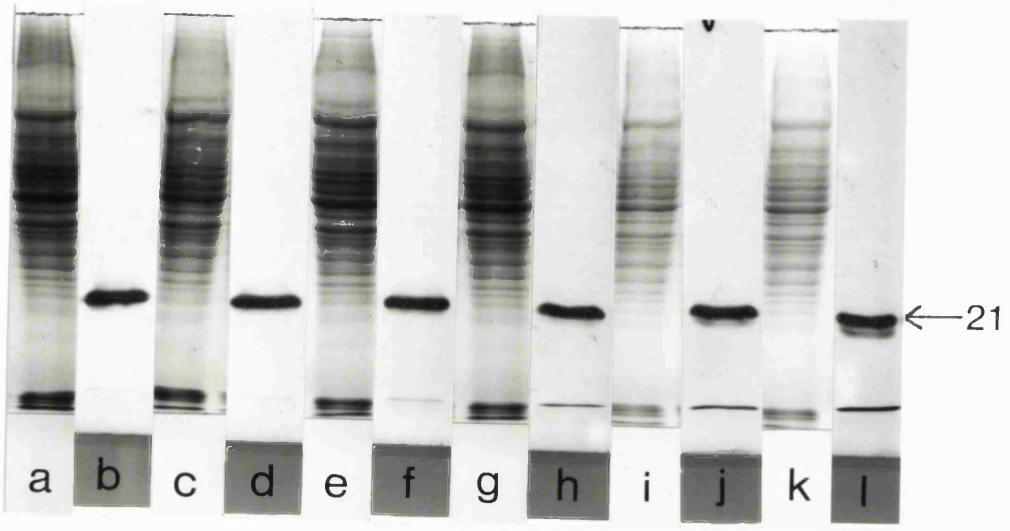


FIGURE 4. $C4^l$ proteolysis: $C4^l$ is seen to maintain its integrity when an HTCL lysate using 1% CHAPS in the presence of the protease inhibitor cocktail shown in table 1 (**a**) is transferred to nitrocellulose and probed with the anti-C4 monoclonal antibody (**b**). The integrity of $C4^l$ is seen to be maintained when the HTCL lysate was dialysed for 24 (**c**) and 50 (**e**) hours at 4°c in the presence of the inhibitor cocktail and transferred to nitrocellulose (**d** and **f** respectively). The $C4^l$ population of a similar HTCL lysate in the absence of the inhibitor cocktail (**g**) is seen to be subjected to slight proteolytic attack (illustrated by low molecular weight bands containing the epitope for the C4 monoclonal antibody) when transferred to nitrocellulose (**h**). Increasing proteolytic attack of the $C4^l$ population is seen when the HTCL lysate, dialysed for 24 (**i**) and 50 (**k**) hours at 4°c in the absence of the inhibitor cocktail is transferred to nitrocellulose (**j** and **l** respectively) An arrow on the right indicates a molecular weight of 21kDa and the presence of $C4^l$.

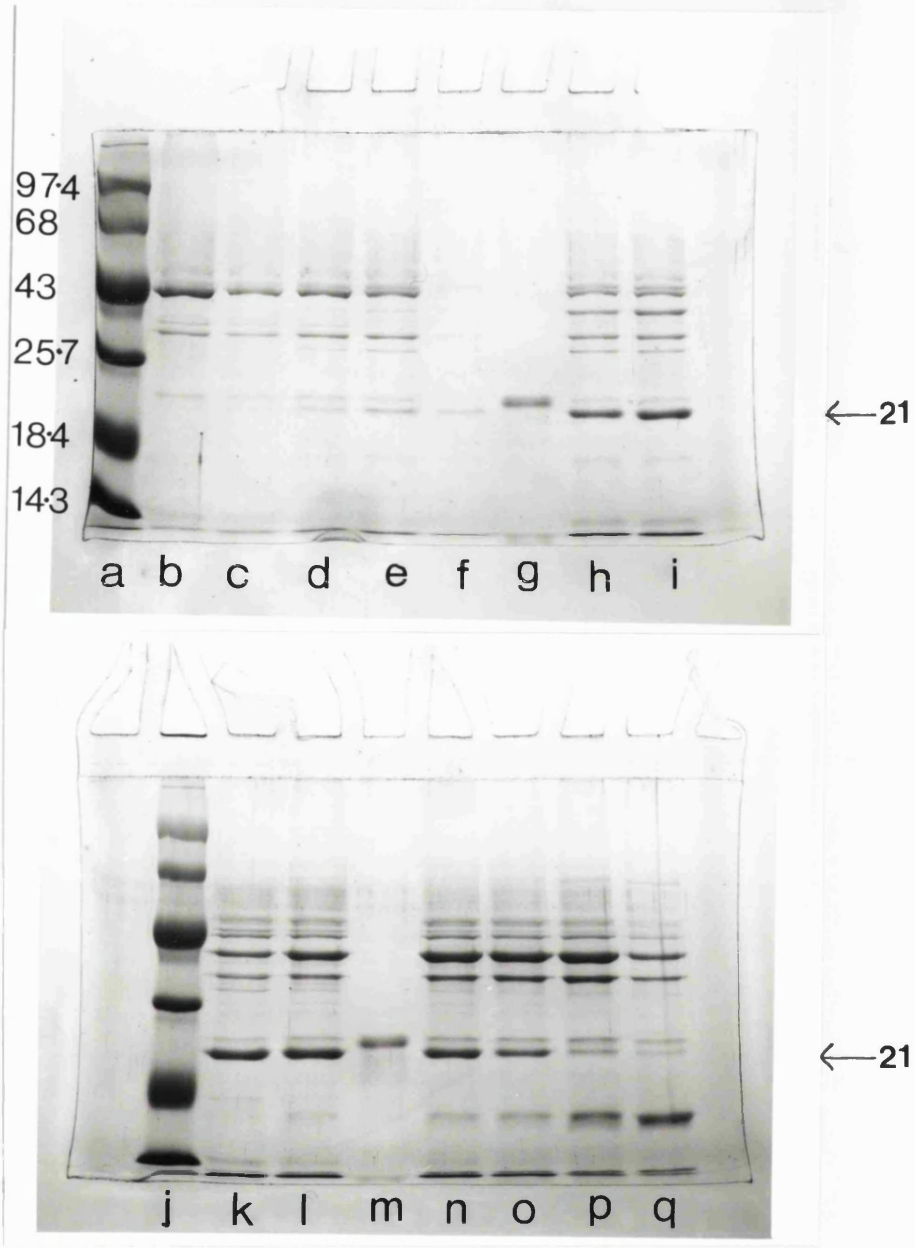


FIGURE 5. Chromatofocusing chromatography of an HTCL 1% CHAPS lysate applied to a pH 7-9.6 chromatofocus column. SDS-PAGE analysis of 0.5ml fractions within an 8ml peak that eluted from the column in the pH range 8.0-8.2 is shown. Molecular weight markers are shown (**a** and **j**) with weights indicated in kDa on the right. An arrow on the left indicates the position of 21kDa and the presence of $C4^{l}$. Purified transgelin (from C. Shapland) was loaded (**g** and **m**) as a further indication of molecular weight. $C4^{l}$ is detected as the major protein in four fractions (**h**, **i**, **k** and **l**) and were consequently pooled for further purification, whilst the remaining fractions (**b**, **c**, **d**, **e**, **f**, **n**, **o**, **p** and **q**) of this peak, discarded.

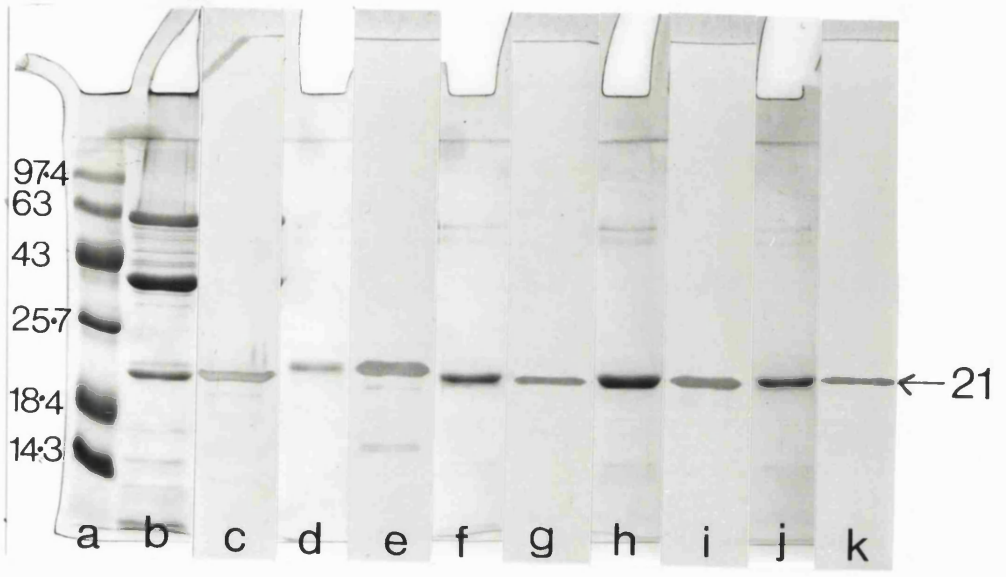


FIGURE 6. Hydroxylapatite chromatography of pooled $C4^l$ positive chromatofocus fractions. Molecular weight markers are shown (**a**) and weights indicated in kDa on the left. An arrow on the right indicates a molecular weight of 21kDa and the presence of $C4^l$. The pooled chromatofocus fractions (**b**) are shown to contain a population of $C4^l$ when transferred to nitrocellulose and probed with anti-C4 monoclonal antibody (**c**). Purified transgelin (from C. Shapland) (**d**) transferred to nitrocellulose (**e**) acts as a further molecular weight marker. Elution of adsorbed proteins produces three 0.5ml fractions (**f**, **h** and **j**) that are seen to contain >90% homogeneous $C4^l$ with no apparent degradation when transferred to nitrocellulose (**g**, **i** and **k**).

E. OLIGONUCLEOTIDE SELECTION FOR C4[1] IDENTIFICATION

C4[1] purified to approximately 90% homogeneity was fragmented, HPLC resolved peptides isolated and their amino acid sequences determined by Drs. J. Hsuan and N. Totty at the Ludwig Institute, London. This approach, in conjunction with N terminal sequencing (also by Drs. J. Hsuan and N. Totty), generated eight amino acids of N terminal data and nine major peptides giving a combined total of 106 amino acids (see table 2).

Table 2. Peptides Generated From Purified C4[1]

Peptide N°.	FRAGMENT	Codons Amino Acid	Oligonucleotides Synthesised
1	GPAYGLSR	4.5	
2	ENFQNWLK	2.38	23mer. 64fold degeneracy with 1 Inosine base.
3	KIQASTMAFK	3.0	
4	YGINTTDIFQTVDLWEGK	2.94	
5	NFSDNQLQEGK	2.91	
6	GASQAGMTGGYGMPR	3.57	23mer. 2 fold degeneracy. with 4 Inosine bases. Antisense
7	TLMNLGGLAVAR	4.25	
8	DDGLFSGDPNWFPK	3.15	
9	NVIGLQMGTNR	3.45	
N-terminal Sequence	GPAYGLSR		

Peptide number 2 (ENFQNWLK) with a codon possibility ratio of 2.38 could, when 'reversed translated' be seen to possess a region of 23 nucleotides in length (spanning eight amino acids) containing 7 regions of variability with 256 fold degeneracy. Use of the neutral base inosine (Martin et al 1985) at one of these positions reduced the degeneracy to 64 fold (designated oligonucleotide 1; 5'-GAA/GAAC/TTTC/TCAA/GAAC/TTGGC/TTIAA -3'). Peptide number 6 (GASQAGGMTGYGMPR) possessed a region 23 nucleotides in length (spanning nine amino acids) containing 5 positions of variability with 512 fold degeneracy. Use of the neutral base inosine at four of these positions reduced the degeneracy to 2 fold. An antisense oligonucleotide (designated oligonucleotide 2; 5'-GGCATICCA/GTAICC IGTGATICC-3') complementary to the last nine amino acids in peptide 6 was generated.

F. PCR AMPLIFICATION OF HTCL cDNA BY $C4^l$ SPECIFIC OLIGONUCLEOTIDES 1 AND 2

The $C4^l$ gene specific primers, oligonucleotides 1 and 2, generated a single abundant 438bp product in RT-PCR amplifications from reverse transcribed HTCL polyA$^+$ selected mRNA (the cell source of $C4^l$ purification) (fig 7). An Mg^{2+} concentration of 2.5mM per PCR reaction produced the sharpest, most abundant product (fig 7 d).

The PCR product derived from oligonucleotides 1 and 2 was blunt-end ligated into Sma I digested pUC19 plasmid (designated clone 1) (fig 10 b). Double stranded sequencing of clone 1 using primers complementary to pUC19 (-40 and M13 reverse) identified that it contained coding information with absolute homology to purified $C4^l$.

G. 5'- END ANCHOR PCR OF THE $C4^l$ CODING REGION

From the resultant sequencing of clone 1, a 35mer sense primer (designated oligonucleotide 3; 5'-CTATCTAGAGAGCGAGATGATGGGCTCTTCTCTGG-3') and a 35mer antisense strand primer (designated oligonucleotide 4; 5'-CTATCTAGAGAGC CAGAGAAGAGCCCATCATCTCG-3') were directed against the peptide region RDDGLFS (see fig 11 B). XbaI restriction sites at the 5'- termini were employed to clone into pUC19.

A $dC_{(13)}$ anchor primer (5'-CTATCTAGAGAGCTCGCGGCCGCCCCCCCCC CCCC-3') directed against homopolymeric dG tails and the $C4^l$ gene specific oligonucleotide 4 generated a single abundant 469bp product in PCR amplifications from HTCL, dGTP tailed reverse transcribed polyA$^+$ selected mRNA (fig 8).

The 469bp PCR product derived from the $dC_{(13)}$ anchor primer and oligonucleotide 4 was XbaI digested and ligated into XbaI digested pUC19 plasmid (designated clone 2) (fig 10 c). Double stranded sequencing of clone 2 using primers complementary to pUC19 (-40 and M13 reverse) identified that it contained coding information with absolute homology to the N terminal peptide derived from purified $C4^l$ (see table 2).

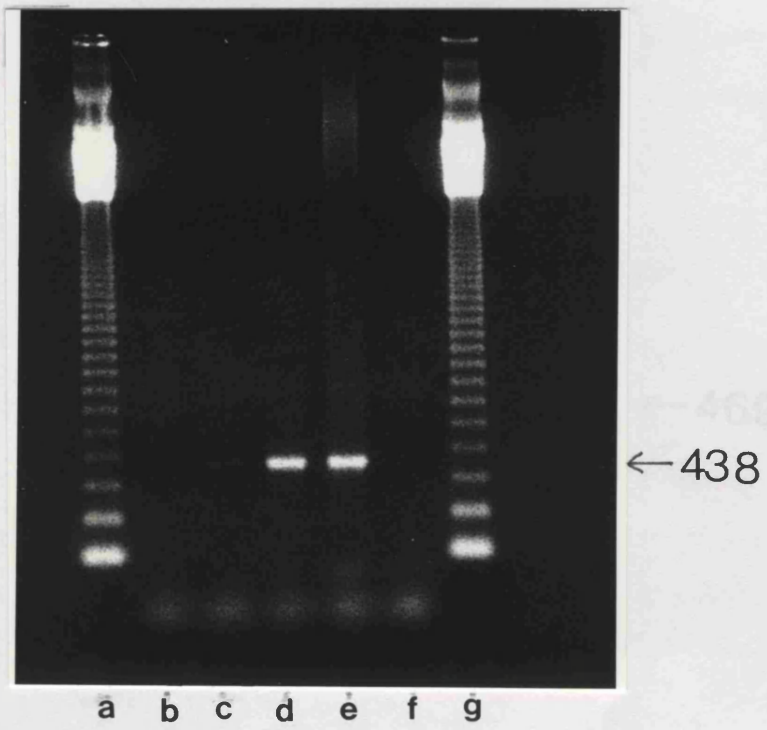


FIGURE 7. PCR amplification of HTCL cDNA by $C4^l$ specific oligonucleotides 1 and 2. 10% of all PCR products were loaded and visualised on 1% agarose gels and ethidium bromide staining. A 123 ladder (Gibco BRL) is loaded (**a** and **g**) indicating molecular weight. The arrow on the right indicates a molecular weight of 438bp. $C4^l$ gene specific primers (oligonucleotides 1 and 2) were used in RT-PCR amplifications from reverse transcribed HTCL poly A^+ selected mRNA. PCR amplifications were carried out at Mg^{2+} concentrations of 1.5mM (**b**), 2mM (**c**), 2.5mM (**d**), 3mM (**e**) and 3.5mM (**f**) per PCR reaction. The sharpest, most abundant, product is seen with a Mg^{2+} concentration of 2.5mM using a primer annealing temperature of 47°c for 30 cycles (**d**).

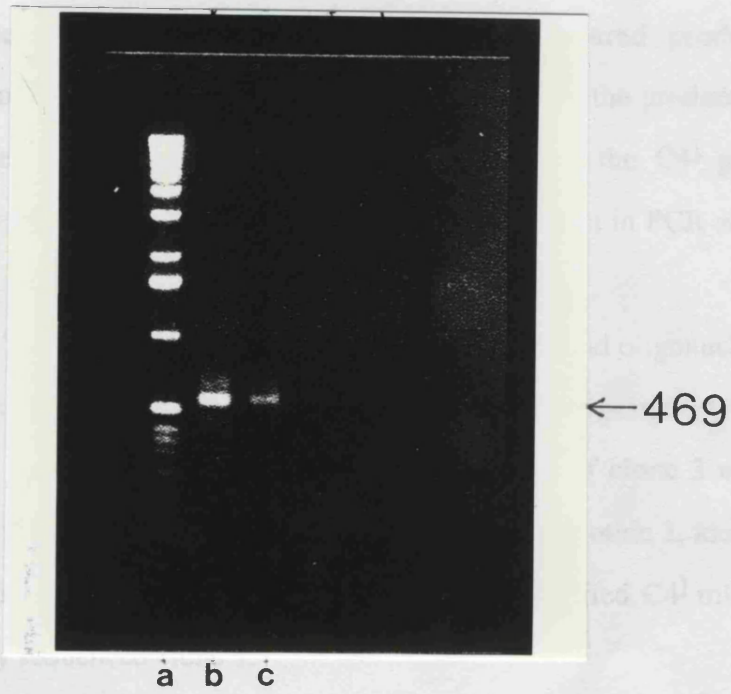


FIGURE 8. 5'-anchor PCR of the $C4^l$ coding region. 10% of all PCR products were loaded and visualised on 1% agarose gels and ethidium bromide staining. A 1Kb ladder (Gibco BRL) is loaded (**a**) to indicate molecular weight. An arrow on the right indicates a molecular weight of 469bp. A $dC_{(13)}$ anchor primer directed against homopolymeric dG tails and the $C4^l$ specific oligonucleotide 4 generates a single 469bp product in PCR amplifications from 400ng of HTCL, dGTP tailed, reverse transcribed poly A^+ selected mRNA (**b**) and 1µg of HTCL, dGTP tailed, reverse transcribed total RNA (**c**). Note the abundance of product when poly A^+ selected mRNA (**b**) is used in 5'-anchor PCR reactions at a Mg^{2+} concentration of 2.5mM and a primer annealing temperature of 56°c for 30 cycles.

H. RAPID AMPLIFICATION OF THE 3'-END OF THE C4^{l} CODING REGION

First strand cDNA from HTCL polyA^{+} selected mRNA was generated by using a 3' Rapid Amplification of cDNA Ends (RACE) kit (Gibco BRL Life Technologies). A primer (UAP) complementary to the primer used for first strand cDNA synthesis (AP) and the C4^{l} specific oligonucleotide 1 generated a smeared product in PCR amplifications from HTCL polyA^{+} mRNA (not shown). 1% of the products from these PCR reactions were further 'nested' by using the UAP and the C4^{l} gene specific oligonucleotide 3 and generated a single abundant 876bp product in PCR amplifications (fig 9).

The 876bp 'nested' PCR product derived from the UAP and oligonucleotide 3 was SalI/XbaI double digested and ligated into SalI/XbaI double digested pUC19 plasmid (designated clone 3) (fig 10 d). Double stranded sequencing of clone 3 using primers complementary to pUC19 (-40 and M13 reverse) plus oligonucleotide 2, identified that it contained coding information with absolute homology to purified C4^{l} microsequence data and previously sequenced clone 1.

I. SEQUENCING

Double stranded sequencing of the three C4^{l} clones (1, 2 and 3) allowed overall alignment, translation and comparison of the encoded protein with partially purified C4^{l}.

The complete cDNA sequence for C4^{l} is shown in figure 11. 58bp of 5' leader sequence can be seen preceding the first ATG (fig 11 A). This is followed by a 594bp opening reading frame (ORF) ending with a TGA stop codon (597-600) (fig 11 B (a)). Following on from the stop codon is approximately 746bp of 3' untranslated region (based on PCR product size and transcript size (results, section R. 1) of which 105bp have been sequenced (fig 11 C). The translated amino acid sequence for C4^{l} is shown below the open reading frame (fig 11 B (b)) and the amino acid sequence of the peptides obtained from purified C4^{l} are shown below their counterparts (fig 11 B (c)).

A diagrammatic representation of the C4^{l} clones (1, 2 and 3) showing the oligonucleotides that generated them and the C4^{l} coding regions they encode is shown in figure 12.

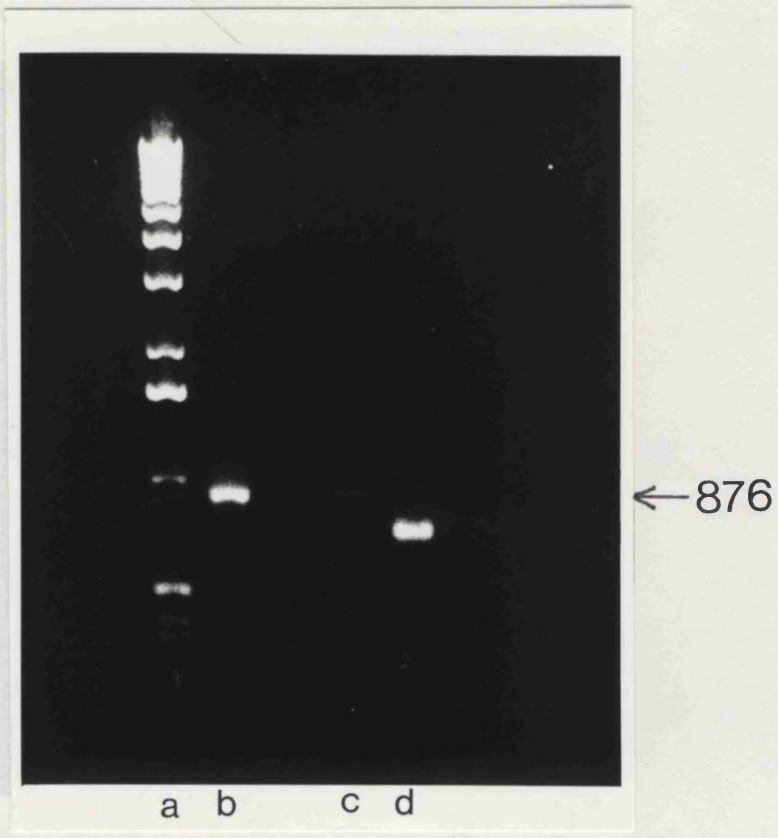


FIGURE 9. Rapid amplification of the 3'-end of the C4[l] coding region. 10% of all PCR products were loaded and visualised on 1% agarose gels and ethidium bromide staining. A 1Kb ladder (Gibco BRL) is loaded (**a**) to indicate molecular weight. An arrow on the right indicates a molecular weight of 876bp. First strand cDNA from HTCL RNA was generated using a 3'- Rapid Amplification of cDNA Ends (RACE) kit (Gibco BRL). A primer (UAP) complementary to the primer used for first strand synthesis (AP) and the C4[l] gene specific oligonucleotide 1 generated a smeared product from HTCL RNA in PCR amplifications (not shown). 1% of the products from these PCR reactions were further 'nested' using the UAP and the C4[l] gene specific primer oligonucleotide 3 and generated a single 876bp product in PCR amplifications (**b** and **c**). Note the abundance of product when 400ng of HTCL poly A^+ selected mRNA (**b**) is used in first strand synthesis compared to when 1µg of HTCL total RNA is used (**c**). A control PCR amplification reaction, with components provided in the RACE kit, produces a single abundant product of 720bp (**d**), indicating the viability of the procedure.

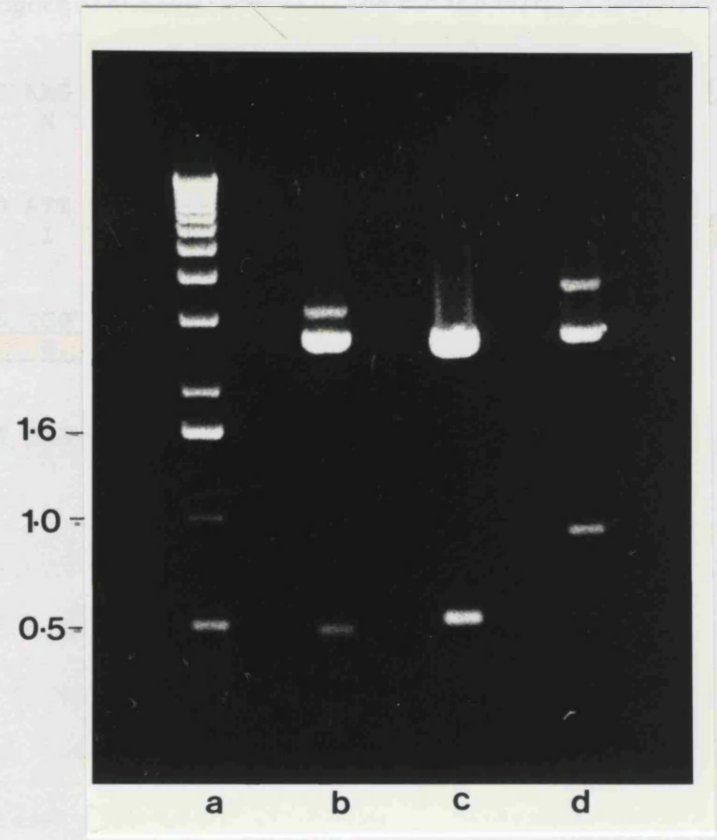
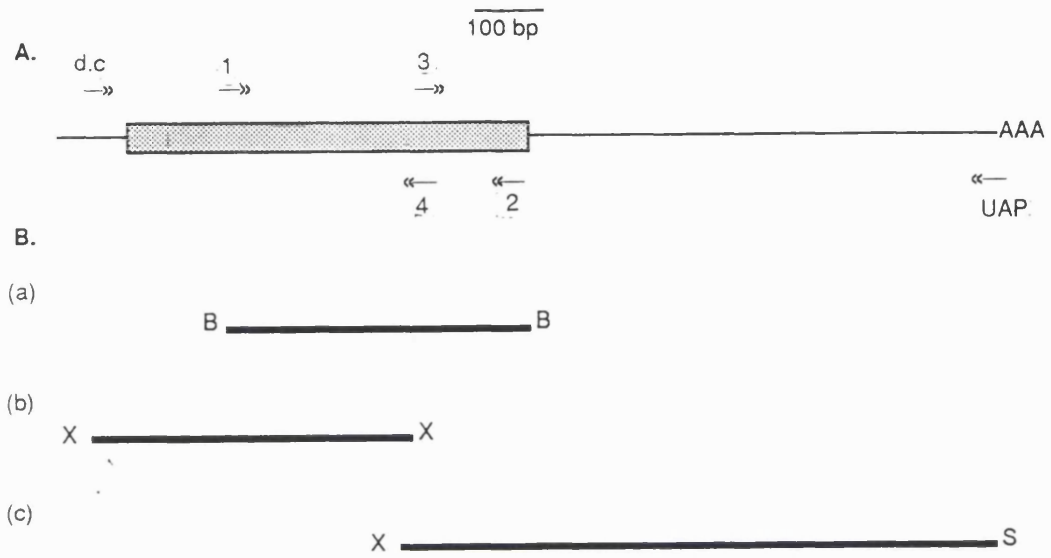


FIGURE 10. Cloned C4[1] specific PCR products. 2µg of digested miniprep products were loaded and visualised on 1% agarose gels and ethidium bromide staining. A 1Kb ladder (Gibco BRL) is loaded to indicate molecular weight (**a**), with weights indicated in Kb on the left. The PCR product derived from oligonucleotides 1 and 2 was blunt-end ligated into SmaI digested pUC19 plasmid and designated clone 1. Double digestion of clone 1 with PstI and EcoRI restriction enzymes produces the 438bp PCR insert (**b**). The PCR product derived from the $dC_{(13)}$ anchor primer and oligonucleotide 4 was XbaI digested, ligated into XbaI digested pUC19 plasmid and designated clone 2. Digestion of clone 2 with XbaI restriction enzyme produces the 469bp PCR insert (**c**). The 'nested' PCR product derived from UAP and oligonucleotide 3 was SalI/XbaI double digested, ligated into SalI/XbaI double digested pUC19 plasmid and designated clone 3. Double digestion of clone 3 with SalI and XbaI restriction enzymes produces the 876bp 'nested' PCR insert (**d**).

FIGURE 11. Complete nucleotide and derived amino acid sequence for C4[l]. **A**: The 5'-non-coding leader sequence found present in clone 2. 58 nucleotides was the largest region of 5'-leader sequence found in all clone 2 inserts sequenced. **B**: HTCL C4[l] nucleotide sequence obtained from overlapping RT-PCR, 5'-anchor and 3'-RACE amplified then cloned products (**a**) is listed above the deduced amino acid sequence (**b**) and aligned peptides from purified HTCL C4[l] (**c**). A putative actin-binding domain is marked in *pink* and the position of potential serine and theronine phosphorylation sites marked in *blue*. The location of two possible EF-Hand structures are marked in *yellow*. Identity to a consensus nuclear localisation signal is marked in *green*. Three cysteine residues at amino acid positions 38, 63 and124 are circled. **C**: 3'-untranslated region (UTR) present in clone 3. 105 nucleotides of 3'-UTR has been sequenced. Note the presence of a second in frame TGA stop codon at position 670 confirming that translation cannot extend this region in this reading frame. This indicates that the cDNA sequence presented for C4[l] encodes the full length protein.

C4l cDNA AND DERIVED PEPTIDE SEQUENCE

```
(A)
-58   cagtgcgccgctctccagcccgcttgaacgctccccgcagccaccgccacccattgga

(B)
(a)  ATG GCC AAC AGG GGA CCT GCA TAT GGC CTG AGC AGG GAG GTG CAG   (45)
(b)   M   A   N   R   G   P   A   Y   G   L   S   R   E   V   Q   (15)
(c)                     G   P   A   Y   G   L   S   R

(a)  CAG AAG ATT GAG AAA CAA TAT GAT GCA GAT CTG GAG CAG ATC CTG   (90)
(b)   Q   K   I   E   K   Q   Y   D   A   D   L   E   Q   I   L   (30)
(c)

(a)  ATC CAG TGG ATC ACC ACC CAG TGC CGA AAG GAT GTG GGC CGG CCC   (135)
(b)   I   Q   W   I   T   T   Q   C   R   K   D   V   G   R   P   (45)
(c)

(a)  CAG CCT GGA CGC GAG AAC TTC CAG AAC TGG CTC AAG GAT GGC ACG   (180)
(b)   Q   P   G   R   E   N   F   Q   N   W   L   K   D   G   T   (60)
(c)                     E   N   F   Q   N   W   L   K

(a)  GTG CTA TGT GAG CTC ATT AAT GCA CTG TAC CCC GAG GGG CAG GCC   (225)
(b)   V   L   C   E   L   I   N   A   L   Y   P   E   G   Q   A   (75)
(c)

(a)  CCA GTA AAG AAG ATC CAG GCC TCC ACC ATG GCC TTC AAG CAG ATG   (270)
(b)   P   V   K   K   I   Q   A   S   T   M   A   F   K   Q   M   (90)
(c)             K   I   Q   A   S   T   M   A   F   K

(a)  GAG CAG ATC TCT CAG TTC CTG CAA GCA GCT GAG CGC TAT GGC ATT   (315)
(b)   E   Q   I   S   Q   F   L   Q   A   A   E   R   Y   G   I   (105)
(c)                                                 Y   G   I

(a)  AAC ACC ACT GAC ATC TTC CAA ACT GTG GAC CTC TGG GAA GGA AAG   (360)
(b)   N   T   T   D   I   F   Q   T   V   D   L   W   E   G   K   (120)
(c)   N   T   T   D   I   F   Q   T   V   D   L   W   E   G   K

(a)  AAC ATG GCC TGT GTG CAG CGG ACG CTG ATG AAT CTG GGT GGG CTG   (405)
(b)   N   M   A   C   V   Q   R   T   L   M   N   L   G   G   L   (135)
(c)                         T   L   M   N   L   G   G   L

(a)  GCA GTA GCC CGA GAT GAT GGG CTC TTC TCT GGG GAT CCC AAC TGG   (450)
(b)   A   V   A   R   D   D   G   L   F   S   G   D   P   N   W   (150)
(c)   A   V   A   R   D   D   G   L   F   S   G   D   P   N   W

(a)  TTC CCT AAG AAA TCC AAG GAG AAT CCT CGG AAC TTC TCG GAT AAC   (495)
(b)   F   P   K   K   S   K   E   N   P   R   N   F   S   D   N   (165)
(c)   F   P   K                           N   F   S   D   N

(a)  CAG CTG CAA GAG GGC AAG AAC GTG ATC GGG TTA CAG ATG GGC ACC   (540)
(b)   Q   L   Q   E   G   K   N   V   I   G   L   Q   M   G   T   (180)
(c)   Q   L   Q   E   G   K   N   V   I   G   L   Q   M   G   T

(a)  AAC CGC GGG GCG TCT CAG GCA GGC ATG ACT GGC TAC GGG ATG CCA   (585)
(b)   N   R   G   A   S   Q   A   G   M   T   G   Y   G   M   P   (195)
(c)   N   R   G   A   S   Q   A   G   M   T   G   Y   G   M   P

(a)  CGC CAG ATC CTC TGA                                         (600)
(b)   R   Q   I   L   *                                          (199)

(C)
tcccaccccaggccttgccctgccctcccacgaatggttaatatatatgtagatacatatgtta   (664)
gcaagtgacattcccagagagccccagagctctcaagctct                       (705)
```



Key: d.c = Oligo $d.C_{(13)}$
1 = Oligo 1 2 = Oligo 2
U.A.P= Universal Primer
A = Poly A Tail B = Blunt-End Site
X = XbaI Restriction Site
S = SalI Restriction Site

FIGURE 12. Schematic representation of the extent of overlapping sequence in clones 1, 2 and 3. Scale is indicated by the top bar. A representation of the HTCL $C4^1$ transcript is depicted (**A**) The shaded area represents the 199 amino acid open reading frame. The thin lines represent the 5'-leader sequence and the 3'-untranslated region plus a poly A tail. The position and 5'-3' orientation of oligonucleotides 1, 2, 3 and 4, as well as the primers $dC_{(13)}$ and UAP are marked with small arrows. The positioning of clones 1, 2 and 3, with respect to the transcript, are shown as bold lines (**B**). The region of the $C4^1$ transcript encoded by clone 1 (derived from oligonucleotides 1 and 2) is illustrated (**a**). The region encoded by clone 2 (derived from oligonucleotides $dC_{(13)}$ and 4) and clone 3 (derived from oligonucleotides 3 and UAP) are also shown (**b** and **c** respectively). Letters seen flanking the clones represent restriction sites utilised during ligation reactions.

J. TRANSLATION

The first ATG is found at position 1 and occurs 23 bp upstream of a nonsense ATG codon (fig 11 B (a)). A Kozak consensus sequence can be seen around this first ATG. Bases -3 and +4 are defined as being particularly critical (Kozak, 1991). Although there is a second ATG codon, starting at position 23 this is only surrounded by one of the critical Kozak bases (fig 13).

The assigned start of C4[1] translation is therefore within 4 amino acids from the start of the N terminal sequence obtained from purified C4[1] (see fig 11 B (c)).

K. GENE PRODUCT

(1) The translated product of C4[1] open reading frame is **199** amino acids in length.

(2) The deduced polypeptide has a calculated molecular weight of **22, 381 Da.**

(3) pI estimation of C4[1] was carried out by statistically averaging the pKa values of all the charged amino acids. An estimate of the value can be made using the equation:

$$pI = \frac{1}{n} \{(X \times pK_{R1}) + (X \times pK_{R2}) + (X \times pK_{R3}) + \ldots + (X \times pK_{Ri})\}$$

where pK_R is the dissociation constant of the side chain group for amino acids 1 to i (see table 3), X is the number of such amino acids in the total protein and n is the total number of amino acid types being considered.

KOZAK TRANSLATION SEQUENCE OF C4[l]

				*		*
		-9	-6	-3	1	+4
	CONSENSUS			A		
		GCC	GCC	GCC	ATG	G
First	(-9-14)	CCC	ATT	GGA	ATG	G
Second	(14-26)	GAC	CTG	CAT	ATG	G

Key: * = Critical Residues.

FIGURE 13. Kozak translation sequence of C4[l]. A 13 nucleotide consensus Kozak initiation sequence was compiled by comparing 699 vertebrate 5'- leader sequences (Kozak, 1987). Two potential C4[l] Kozak initiation sites are presented and their base position within the C4[l] sequence shown in brackets. The A base of the start codon is assumed as position 1. Asterisks indicate bases deemed critical for efficient translation (positions -3 and +4), as shown by point mutation experiments (Kozak, 1986).

Table 3. pK_R of Charged Amino Acids Values
(From Lehninger, 1970)

Amino Acid	PK_R	N° in $C4^l$
D	3.86	10
E	4.25	11
H	6.0	0
C	8.33	3
Y	10.07	5
K	10.53	12
R	12.53	11
C-terminal average	2.185	1
N-terminal average	9.549	1
Total number	-	54

Using this equation and the values in table 3 the calculated pI of $C4^l$ is **8.09**.

(4) The amino acid composition of $C4^l$ is shown in table 4.

Table 4. Amino Acid Composition of $C4^l$

Amino Acid	Number	% of Total
Ala (A)	14	7.0
Arg (R)	11	5.5
Asp (D)	10	5.0
Asn (N)	12	6.0
Cys (C)	3	1.5
Glu (E)	11	5.5
Gln (Q)	21	10.6
Gly (G)	19	9.5
His (H)	0	0
Ile (I)	11	5.5
Leu (L)	16	8.0
Lys (K)	12	6.0
Met (M)	8	4.0
Phe (F)	7	3.5
Pro (P)	8	4.5
Ser (S)	7	3.5
Thr (T)	11	5.5
Trp (W)	4	2.0
Tyr (Y)	5	2.5
Val (V)	8	4.0

The $C4^l$ polypeptide contains 23 positively charged amino acids (12 K, 11 R and 0 H) and 21 negatively charged amino acids (10 D, 11 E), with a net charge of +2 at neutral pH. Three cysteine residues are found occurring at positions 38, 63 and 124 (see fig 11 B (b)).

L. SEQUENCE MOTIFS

1. Putative Actin Binding Domains

Homologies with the $C4^l$ peptide LQAAERY (residues 97-103) (see fig 14 B (b)) were found in transgelin (LKAAEDY) (Prinjha et al 1994) and repeated up to seven times within the α chain of various tropomyosin isoforms (Ruiz-Opazo & Nadal-Ginard, 1987) (KKAAED, LRASED, LHKAED, LLAADE, EKAADE, LKSLEA and LKEAET). Less well conserved partial homologies are found in caldesmon (KKAAEE) (Bryan et al 1989) and yeast ACT2 (LRAEER) (Lees-Miller et al 1992b). These peptides show high degrees of homology with the conserved hexapeptide motifs (LKHAET of actobindin and LKKTET of thymosin β 4) which form the major contact site with actin (Vancompernolle et al 1992). This peptide is also found conserved within a range of other actin-binding proteins (fig 14).

The cluster of positively charged amino acids found in transgelin (residues 54-61), which may be involved in actin gelation (Shapland et al 1993) are not present in $C4^l$.

No other putative actin-binding domains were identified within the $C4^l$ polypeptide sequence.

2. Phosphorylation Sites

Potential serine phosphorylation sites for calcium / calmodulin dependent protein kinase II and cAMP dependent protein kinase (Carlson et al, 1979) were found at residues 160-163 (RNFS) and residues 182-185 (RGAS) respectively. This last site is situated immediately adjacent to a potential site for serine / threonine protein kinase C (Woodjett et al, 1986) at residues 180-182 (TNR) (see fig 11 B (b)).

PROPOSED ACTIN-BINDING DOMAIN

C4[l] (96)	F**LQAAERY**GI
Transgelin (97)	F**LKAAEDY**GV
Thymosin beta 4 (16)	**LKKTET**
Actobindin I (15)	**LKHAET**
Actobindin II (51)	**LKHAET**
NP25 (96)	F**LKAAEVY**GV
Calponin α &β	F**LRAIKHY**GV
mp20 (82)	F**QKALKEY**GV
Alpha Actinin (89)	**LKHIESH**GV
Tropomyosin (TM3) (23)	**KKAAED**
Tropomyosin (TM4) (82)	**KAADE**
Caldesmon I (269)	**KKAAEE**
Caldesmon II (282)	**KKAAEE**
Caldesmon III (297)	**KRAAEE**
Caldesmon IV (312)	**RKAAEE**
Troponin C (105)	**LKAALD**
Parvalbumin alpha (80)	F**LKAADK**
Villin (Human) (249)	**LKAALK**
Villin (Chick) (773)	**NTAAED**
ACT2 (Yeast) (39)	**LRAEER**
Fast Skeletal MHC (1010)	**LQAEED**
Fragmin (210)	**VKAAEL**
Band III (137)	**FKRAEDY**
Ezrin (364)	**RKAQEE**
Yeast Profilin I (70)	**LRADDR**

FIGURE 14. Putative actin binding domain of C4[l]. The FLQAAERYGI motif within C4[l] is shown aligned against homologous sequences within other actin-binding proteins. The LKHAET motif of actobindin has been identified as its actin-binding domain (Vancompernolle et al, 1992). The numbers in brackets indicate location within the protein sequences.

3. EF-Hand Structure of C4[l]

Two regions of twelve amino acids (23-34 and 107-118) (see fig 11 B (b)) can be aligned with the core region of the sequence motif defined by Tufty & Kretsinger (1975) as an EF Hand structure. Three (site I) and four (site II) of the seven necessary residues are conserved within C4[l] (fig 15).

4. Nuclear Localisation Signal (NLS)

A seven amino acid motif within C4[l] (residues 152-158) (see fig 11 B (b)) shares identity with an NLS consensus sequence (Garcia-Bustos et al 1991). Related features are a cluster of positively charged amino acids (K, R) associated with a proline (P). Figure 16 shows this motif within C4[l] aligned to a consensus NLS and well defined NLS sequences from SV40 and polyoma virus large T antigens.

M. HYDROPATHY PLOT

The Kyte and Doolittle method (1982) for displaying the hydropathic character of proteins was used to analyse C4[l]. The 'Gene Jockey' program was set to statistically average hydropathy over a six residue window. The resultant plot revealed an essentially neutral core region with short hydrophobic N and C termini. No strongly hydrophobic regions were found (fig 17).

N. SECONDARY STRUCTURE PREDICTION

A prediction of secondary structure was carried out using a multiprediction program (Zvelebil et al 1987) by Dr. M. J. Zvelebil (on a VAX 11/750 system) at the Ludwig Institute for Cancer Research, London. The program uses information available from a family of homologous sequences (see results, section O). The approach is based both on averaging the Garnier et al (1978) secondary structure propensities for aligned residues and on the observation that high sequence variability tends to occur in loop regions between secondary structure. The output generated by the program is shown in figure 18. These data suggested that a substantial portion (40%) of C4[l] adopts α-helical configuration with <15% in the form of β-sheets, while the remainder are either regions of random structure, coils or turns.

EF-HAND ALIGNMENT

7 Critical Residues	* * * * * * *	
Consensus	D D' D'G I'D' E	
Carp Parvalbumin I	D Q E K S G F I E E D E	(7/7)
Site II	D S D G D G K I G V D E	(6/7)
Rabbit Troponin I	D A D G G G D I S V K E	(6/7)
Site II	D E D G S G T I D F E E	(7/7)
Site III	D R N A D G Y I D P E E	(7/7)
Site IV	D K N N D G R I D F D E	(7/7)
Drosophila mp20 I	D K E A Q E W I E A I I	(5/7)
Site II	D I D V F Q T V D L Y E	(5/7)
$C4^1$ Site I	D A D L E Q I L I Q W I	(3/7)
Site II	T T D I F Q T V D L W E	(4/7)
Transgelin I	D E E L E E R L V E W I	(3/7)
Site II	K T D M F Q T V D L F E	(4/7)
NP25 Site I	D A D L E N K L V D W I	(3/7)
Site II	T T D I F Q T V D L W E	(4/7)
Consensus	D D' D'G I'D' E	

Key: D=D; D'=D,N,E,Q,S or T; G=G; I'=I or V; E=E

FIGURE 15. The EF-hand like regions of $C4^1$ are shown aligned with a number of functional calcium-binding motifs in other proteins. Asterisks mark residues proposed to be critical (the first D residue is of particular importance) for the binding of the calcium ion (Tufty & Kretsinger, 1975). Figures in brackets after each line indicate the number of residues conforming to the consensus sequence. The key indicates the permissible variants which still allow calcium binding.

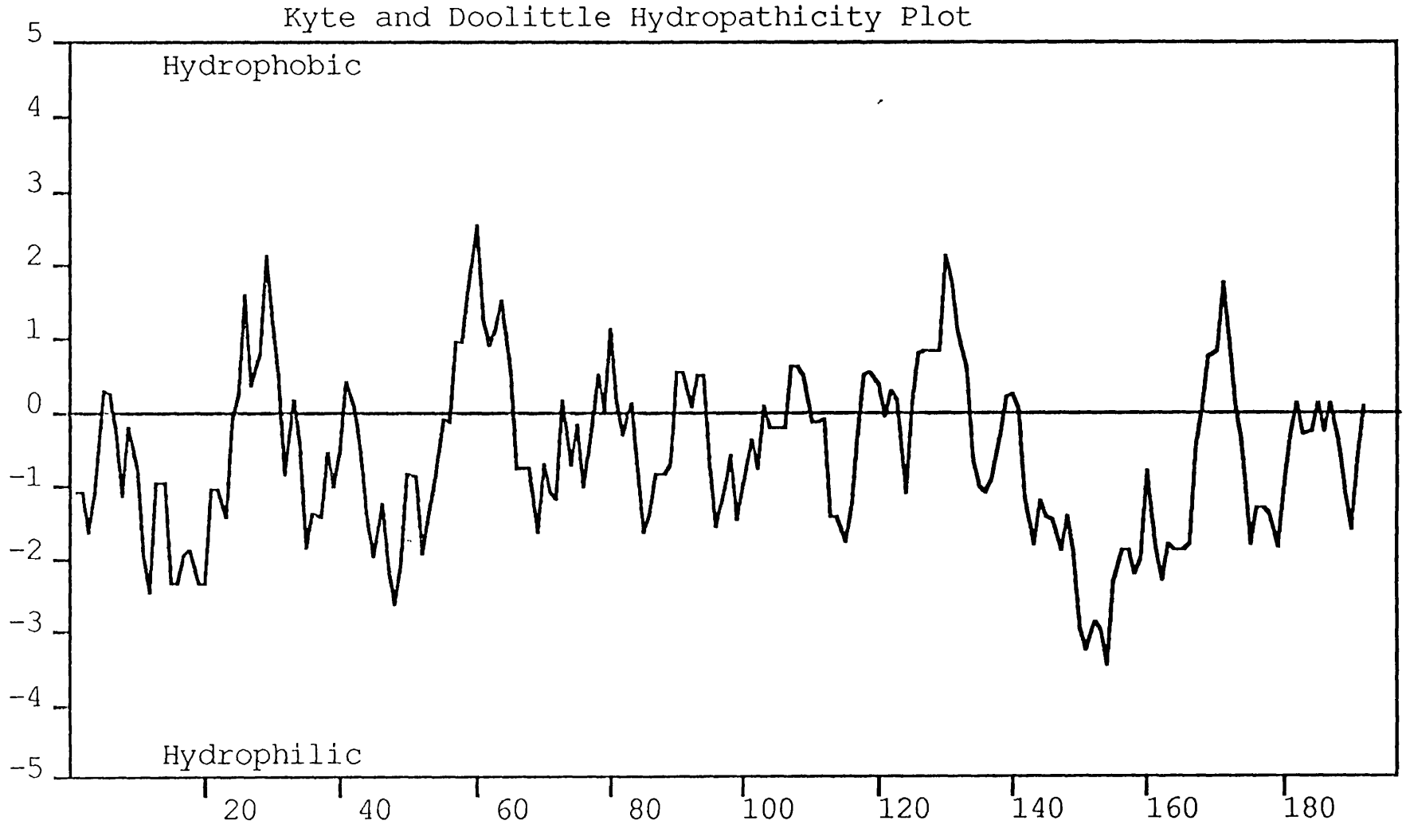
PROPOSED NUCLEAR LOCALISATION SIGNAL

Consensus		**R**			**R**		
(Associated with P)		**K**	**K**	**X**	**K**		
C4[1] (152)	P	**K**	**K**	**S**	**K**	E	N
SV40 Large T (126)	P	**K**	**K**	**K**	**R**	K	V
Polyoma Large T(189)	P	**K**	**K**	**A**	**R**	E	D

Key: X= Any residue

FIGURE 16. Proposed C4[1] nuclear localisation signal (NLS). The proposed NLS of C4[1] is shown aligned against an NLS consensus (Garcia-Bustos et al, 1991) and the known NLS sequences of SV40 large T and polyoma virus large T. Numbers in brackets indicate residue position within the proteins.

FIGURE 17. Hydropathy profile of C4[1]. The averaged hydropathic index of overlapping hexapeptides (values from the PEP program-Intelligenetics) is plotted against amino acid number.


Kyte and Doolittle Hydropathicity Plot
Hydrophobic
Hydrophilic
5
4
3
2
1
0
-1
-2
-3
-4
-5
20
40
60
80
100
120
140
160
180

SECONDARY STRUCTURE PREDICTION

```
      1           11          21          31          41
      MANRGPAYGL  SREVQQKIEK  QYDADLEQIL  IQWITTQCRK  DVGRPQPGRE
H     H        H  HHHHHHHHHH  HHHHHHHHHH  HHHHHHH
B
T      TTTTTTTT                                 TTT  TTTTTTTTTT

      51          61          71          81          91
      NFQNWLKDGT  VLCELINALY  PEGQAPVKKI  QASTMAFKQM  EQISQFLQAA
H     HHHHHHHHH                           HHHHHHHHHH  HHHHHHHHHH
B              B  BBBBBBBBB      BBBBBB
T                        TT TTTT

      101         111         121         131         141
      ERYGINTTDI  FQTVDLWEGK  NMACVQRTLM  NLGGLAVARD  DGLFSGDPTW
H                   HHHHHHHH  HHHHH    H  HHHHHHH
B       BB  BBBB  BB              BBBB
T     TT  TT                              TTT  TTTTTTTTTT

      151         161         171         181         191  199
      FPKKSKENPR  NFSDNQLQEG  KNVIGLQMGT  NRGASQAGMT  GYGMPRQIL
H                 HHH
B                             BBBBBB        BBBBBB
T     TTTTTTTTTT    TTTTTTT       TTTT  TT      TT  TTTTTTTTT
```

Key: H represents alpha helix
B represents beta sheet
T represents turns/coils

FIGURE 18. Predicted secondary structure of C4[l]. A multiprediction program (Zvelebil et al, 1987) on a VAX 11/750 system was used to calculate the regional propensity of C4[l] to adopt various structural configurations. H denotes a propensity to adopt an alpha-helical conformation. B and T denote the ability of the marked region to form beta pleated sheets or turns, respectively.

O. HOMOLOGY SEARCHES

Use of the IFIND and FASTDB programs to search available databases via the Ludwig Institute VAX 11/750 system indicated that C4^{l} had significant homology with transgelin (and its chick homologue, SM22α) (Prinjha et al 1994), the rat neuronal protein NP25 (EMBL accession N^{o} M84725), *Drosophila* synchronous flight muscle protein mp20 (Ayme-Southgate et al, 1989) and chicken gizzard calponin (Takahashi & Nadal-Ginard 1991). Figure 19 shows homology alignments (calponin β is aligned, it being the shorter isoform). Table 5 shows the percentage of identity between C4^{l} and its homologous proteins.

Table 5. Homology Comparisons

Comparison (In Descending Order of Homology).	Percentage Identical Homology.	Percentage Homology (With Conservative Substitutions).
C4^{l} Vs NP25 (over 199 amino acids)	67.0%	74.4%
C4^{l} Vs Transgelin (over 199 amino acids)	66.8%	73.9%
C4^{l} Vs Calponin β (over 199 amino acids)	39.2%	49.7%
C4^{l} Vs mp20 (over 184 amino acids)	38.0%	49.5%

NP25 and transgelin are the most closely homologous molecules to C4^{l}.

This comparison of the respective cDNA sequences indicate that:

(1) 5' leader sequences and 3' untranslated region are completely different.

(2) Differences in amino acids are distributed throughout the proteins.

Seven regions within a 379 amino acid alternatively spliced product of the *C. elegans* unc87 gene (unpublished, EMBL accession number U04711) show homology when aligned to a 21 amino acid region at the C-terminus of C4^{l} (amino acids 174-194). The fourth such site within unc87 (amino acids 198-218) displays the highest degree of homology with 66% identity (76% with conservative substitutions), and is shown aligned to C4^{l} in figure 20.

SEQUENCE ALIGNMENTS

```
1          MANRGPAYGL  SREVQQKIEK  QYDADLEQIL  IQWITTQCRK  DVGRPQPGRE   50
2          MANKGPSYGM  SREVQSKIEK  KYDEELEERL  VEWIIVQCGP  DVGRPDRGRL   50
3          MANRGPSYGL  SREVQEKIEQ  KYDADLENKL  VDWIILQCAE  DIEHPPPGRT   50
4                  MSL  ERAVRAKIAS  KRNPEMDKEA  QEWIEAIIAE  KF--PA-GQ-   39
5  MSNA    NFNRGPAYGL  SAEVKNKLAQ  KYDPQTERQL  RVWIEGATGR  RIGD------   48

1          NFQNWLKDGT  VLCELINALY  PEGQAPVKKI  QAST-MAFKQ  MEQISQFLQA   99
2          GFQVWLKNGV  ILSKLVNSLY  PDGSKPVKVP  ENPPSMVFKQ  MEQVAQFLKA  100
3          HFQKWLMDGT  VLCKLINSLY  PPGQEPIPKI  -SESKMAFKQ  MEQISQFLKA   99
4          SYEDVLKDGQ  VLCKLINVLS  PNAVPKVN--  --SSGGQFKF  MENINNFQKA   85
5          NFMDGLKDGV  ILCELINTLQ  PGSVQKVN--  --DPVQNWHK  LENIGNFLRA   94

1          AERYGINTTD  IFQTVDLWEG  KNMACVQRTL  MNLGGLAVAR  DDGLFSGDPN  149
2          AEDYGVTKTD  MFQTVDLFEG  KDMAAVQRTV  MALGSLAVTK  NDGHYRGDPN  150
3          AEVYGVRTTD  IFQTVDLWEG  KDMAAVQRTL  MALGSVAVTK  DDGCYRGEPS  149
4          LKEYGVPDID  VFQTVDLYEK  KDIANVTNTI  FALGRATYKH  ADFKGPF---  132
5          IKHYGVKPHD  IFEANDLFEN  TNHTQVQSTL  IALASQAKTK  GN-----NVG  139

1          WFPKKSKENP  RNFSDNQLQE  GKNVIGLQMG  TNRGASQAGM  TGYGMPRQIL  199
2          WFMKKAQEHK  REFTDSQLQE  GKHVIGLQMG  SNRGASQAGM  TGYGRPRQII  200
3          WFHRKAQQNR  RGFSEEQLRQ  GQNVIGLQMG  SNKGASQAGM  TGYGMPGRSC  199
4          LGPKPADECK  RDFTEEQLKA  GQTIVGLQAG  SNKGATQAG-  -NLGAGRKIL  181
5          LGVKYAEKQQ  RRFQPEKLRE  GRNIIGLQMG  TNKFASQQGM  TAYGTRRHLY  189

2          S  201
3          KMLYSECPLV  ERTDVPHHSL  219
4          LGK      184
5          DPKLGTDQPL  DQATISLQMG  TNKGASQGMT  VYGLPRQVYD  PKYCDAPGLL  239

5          GEDGLNHSFY     NSQ  252
```

FIGURE 19. Alignment of homologous sequences. The deduced C4[1] sequence (**1**) is shown aligned with rat transgelin (Prinjha et al, 1994) (**2**), rat NP25 (unpublished, EMBL accession number M84725) (**3**), *Drosophila* mp20 (Ayme-Southgate et al, 1989) (**4**), and chicken calponin β (Takahashi and Nadal-Ginard, 1991) (**5**). Hyphens represent gaps introduced to optimise alignments.

ALIGNMENT OF C4l WITH A *C. elegans* unc87 PEPTIDE

C4l	(174-194)	IGLQMGTNRGASQAGMTGYGM
Unc87	(198-218)	IPLQSGTNKFASQKGMTGFGR

FIGURE 20. Alignment of C4l with a *C. elegans* unc87 (unpublished, EMBL accession number U04711) peptide. A 21 amino acid residue peptide of C4l is shown aligned to a 21 amino acid residue peptide within an alternatively spliced product of the *C. elegans* gene, unc87. Numbers in brackets indicate residue position in the proteins. It must be noted that similarity to this C4l-like motif appears seven times within unc87, and it is the fourth of these (the most identical) that is shown aligned to C4l.

P. DOT PLOT HOMOLOGY COMPARISONS

The Dot-Plot method (Gibbs & McIntyre, 1970; Staden, 1982) can be used to produce illustrative two dimensional representations of 'long-range' homologies. In addition these plots can be used to investigate the occurrence of internal repeats. Comparison using this method indicated that $C4^l$ has no internal repeats (fig 21 a).

Comparison of $C4^l$ with homologous sequences are described below and listed in decreasing order of homology

(1) Comparison of $C4^l$ with NP25 (filter function 4-7, window 8) illustrated extensive end-end homology between the two molecules with both termini and a central region between residues 51-132 of both molecules being extremely similar (fig 21 b).

(2) Comparison of $C4^l$ with transgelin (filter function 4-7, window 8) illustrated the presence of major end-end homology between the molecules. Extensive homology is seen at the N and C termini (fig 21 c).

(3) Comparison of $C4^l$ with calponin α (filter function 4-7, window 8) illustrated N terminal homology with further 'islands' of homology along the length of $C4^l$. A C-terminal region within $C4^l$ is seen repeated three times throughout the sequence of calponin α (fig 21 d).

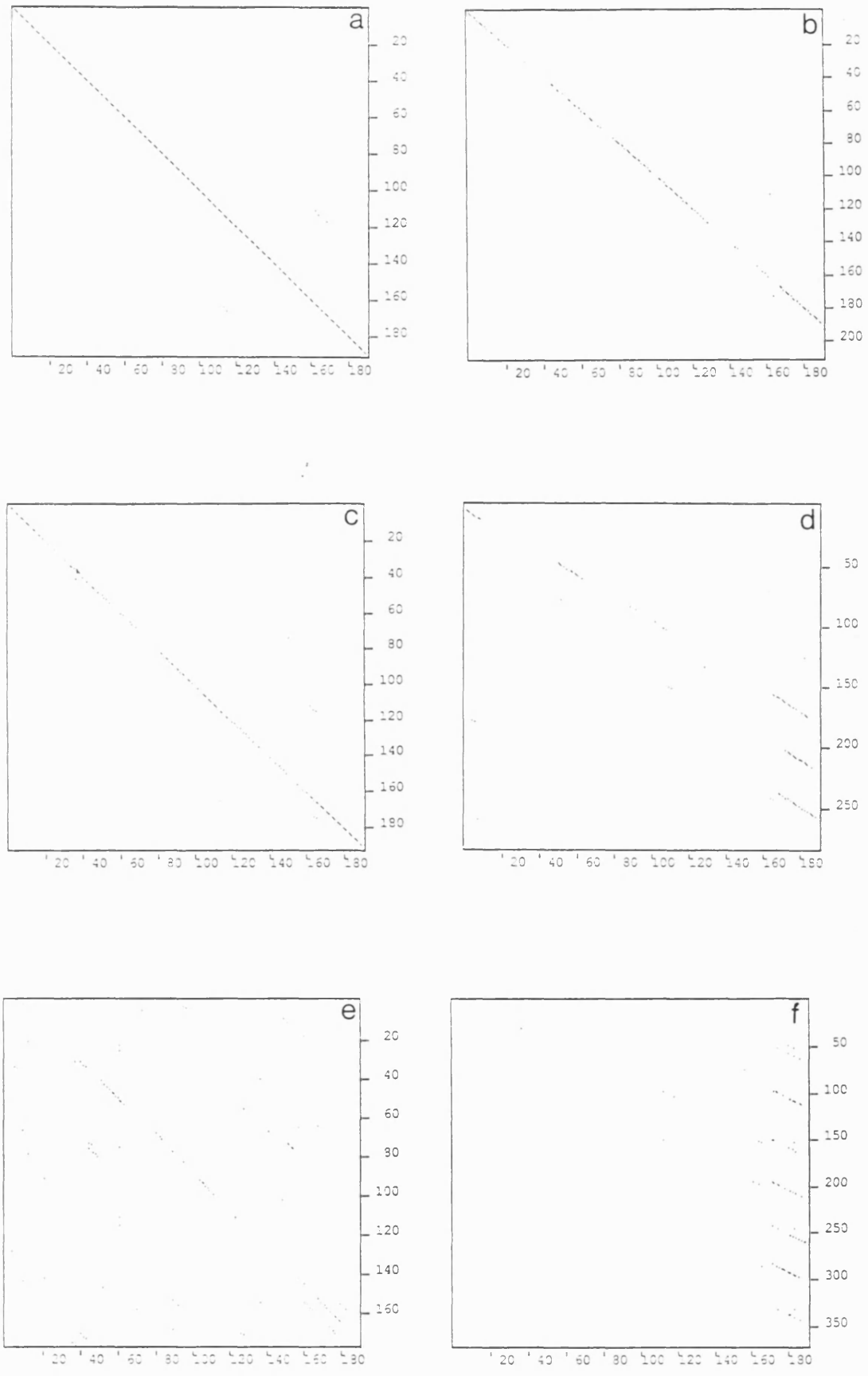
(4) Comparison of $C4^l$ with *Drosophila* mp20 (filter function 4-7, window 8) illustrated the presence of three major regions of homology. The most extensive region occurs near the N termini of both molecules. A second central region ($C4^l$ residues 91-111) of homology is seen and a third, less concentrated, region occurring near the C termini (fig 21 e).

(5) Comparison of $C4^l$ with an alternatively spliced product of the *C. elegans* unc87 gene (filter function 4-7, window 8) illustrated that a $C4^l$ C-terminal region is repeated seven times throughout the sequence of unc87 (fig 21 f).

Q. ADDITIONAL HOMOLOGOUS PEPTIDE

A short peptide present within $C4^l$ (RGASQAG; residues 182-188) (see fig 11 B (b)) can also be seen in a number of other actin-binding proteins. This peptide spans a putative protein kinase A site and is found near the N terminus of gelsolin (Kwiatkowski et al, 1986), in transgelin, NP25, mp20, three times in calponin-α, and seven

FIGURE 21. The diagonal dot-plot method (Gibbs & McIntyre, 1970; Staden, 1982) was used to graphically compare proteins using the Gene Jockey program. A window size of eight residues (filter function 4-7) was used throughout, with C4[l] residues 1-199 shown along the x-axis in all cases. Comparison of C4[l] (**a**) against itself; (**b**) against rat NP25; (**c**) against rat transgelin; (**d**) against calponin-alpha; (**e**) against *Drosophila* mp20; and (**f**) against an alternatively spliced product from the *C. elegans* gene unc87. Note the occurrence of a C4[l] C-terminal region repeated throughout calponin-alpha and unc87.


a
b
c
d
e
f

times within the 379 residue product of the unc87 gene. The seven unc87 peptides also contain an additional protein kinase C site (except site VI) and can be seen in figure 22.

R. NORTHERN BLOT ANALYSIS OF C4l

1. Samples

All blots were hybridised with the 438bp PCR product of nucleotide 1 and 2 which encodes for C4l between amino acids residues 50-196 (see fig 7 and 12).

(i) Human T Cell Lymphoma:

A single message with apparent size of 1.44 Kb was observed in blots of polyA$^+$ selected HTCL mRNA (fig 23 a).

(ii) Rat Thymus:

A weak single message migrating with apparent size of 1.44 Kb was observed in blots of polyA$^+$ selected rat thymus mRNA (fig 23 b).

(iii) SV40 Transformed 3T3 Fibroblasts

High levels of a single message migrating with apparent size of 1.44 Kb was observed in blots of polyA$^+$ selected SV40 transformed 3T3 fibroblasts (fig 23 c).

(iv) 3T3 Fibroblasts:

3T3-fibroblasts have previously been shown to express both C4l and transgelin in approximately equal quantities (Shapland et al (1988), 'Identification of New Actin-Associated Polypeptides That are Modified by Viral Transformation and Changes in Cell Shape'. J. Cell. Biol, 107: 153-161, fig 7). High levels of a single message migrating with apparent size of 1.44 Kb was observed in blots of poly A$^+$ selected 3T3-fibroblast mRNA (fig 23 d).

(v) Mature Rat Skeletal Muscle:

C4l and transgelin have previously been shown to be absent in skeletal muscle (Shapland et al 1988; Prinjha et al, 1994)). Negligible hybridisation, at any stringency, was observed with polyA$^+$ selected mature rat skeletal muscle mRNA, irrespective of lengthy (up to two weeks) autoradiographic exposure times (fig 23 e).

(vi) HTCL PolyA$^+$ selection 'Flow-Through':

Negligible hybridisation was observed, at any stringency, with 20µg of 'flow-through' (see methods, section I. 2) from HTCL polyA$^+$ selection (fig 23 f).

HOMOLOGOUS PEPTIDE

C4^{l} (182-188)	RGASQAG
Transgelin (183-189)	RGASQAG
Gelsolin (32-38)	RGASQAG
Calponinα I (172-178)	KFASQQG
Site II (212-218)	KGASQAG
Site III (251-257)	KGASQQG
NP25 (182-188)	KGASQAG
mp20 (166-172)	KGATQAG
SM22α (182-188)	KGASQAG
unc87 I (59-65)	KGDSQKL
Site II (107-113)	KYCSQRG
Site III (160-166)	KYDSQKG
Site IV (206-212)	KFASQKG
Site V (253-259)	QYASQKG
Site VI (294-300)	RFASQAG
Site VII (339-345)	KGDSQKK

FIGURE 22. Additional C4^{l} peptide alignment. The C4^{l} heptapeptide RGASQAG is shown aligned to similar heptapeptides within other actin-binding proteins. Complete identity to the C4^{l} heptapeptide is seen in transgelin (Prinjha et al, 1994) and gelsolin (Kwiatkowski et al, 1986). Homology to the heptapeptide is also noted three times within calponin α (and seven times within an alternatively spliced product of the *C. elegans* unc87 gene. Bracketed numbers indicate the residue position in the proteins.

2. Actin

By probing RNA samples with an actin cDNA insert, a method of comparing the relative loadings of different samples is achieved. The extent of trailing of the actin message can also be used as a measure of the integrity of the RNA samples. All blots were stripped as described (methods, section T) and hybridised with an α-^{32}P labelled cDNA insert of actin (fig 23 lower panel). This probe recognised all isoforms of actin (β- and γ- actins seen at 2.1 Kb, α- actins seen at 1.6 Kb).

Hybridisation with actin indicated that less HTCL mRNA was loaded than the other samples (fig 23 lower panel; a), where similar mRNA loadings were detected (fig 23 lower panel; b, c and d). Abundant α- actin message (1.6 Kb) was observed in mRNA isolated from mature rat skeletal muscle (fig 23 lower panel; e). HTCL polyA$^+$ mRNA selection 'flow through' (methods, section I. 2) has no actin message present, indicating polyA$^+$ selection was performed correctly (fig 23 lower panel; f).

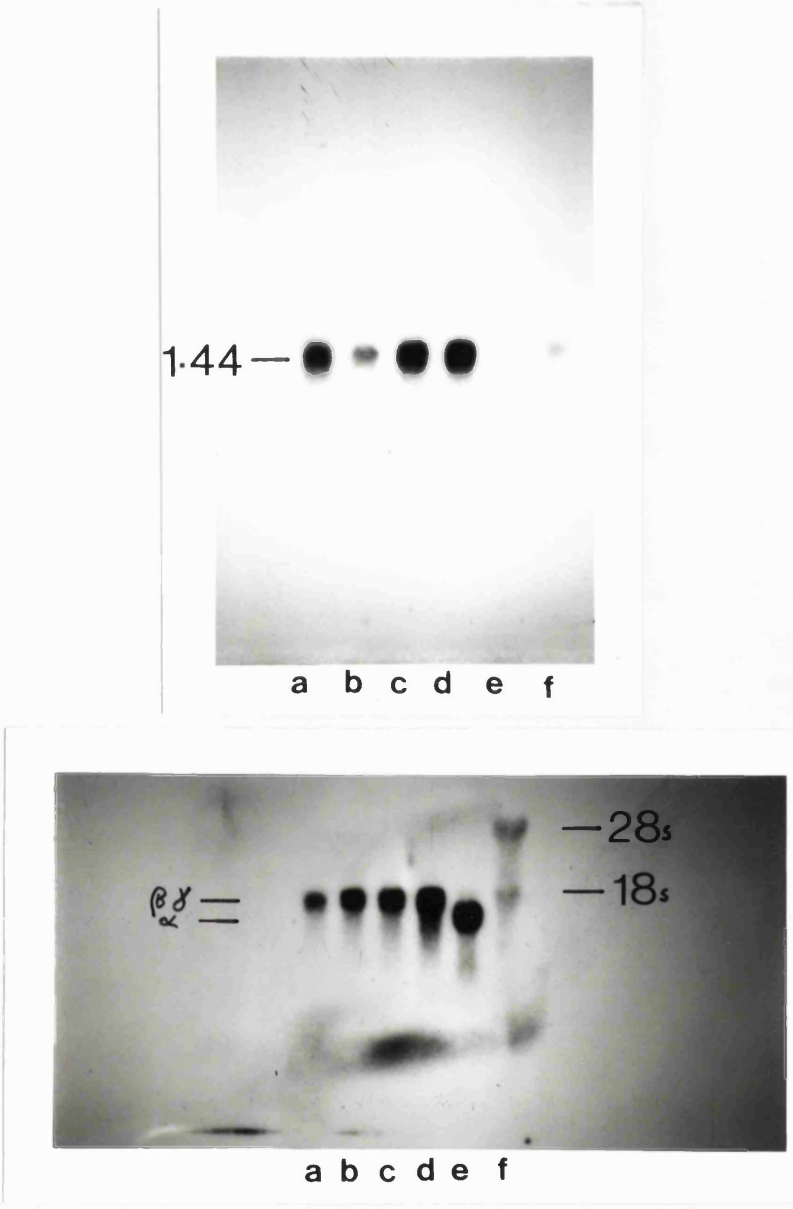


FIGURE 23. Northern blot analysis. The C4[1] 438bp PCR product derived from oligonucleotides 1 and 2 (see fig 7) was hybridised to a Northern blot containing poly A$^+$ selected mRNA from (**a**) HTCL, (**b**) rat thymus, (**c**) SV40 transformed 3T3 fibroblasts, (**d**) 3T3 fibroblasts, (**e**) rat skeletal muscle, (**f**) 'flow-through' from poly A$^+$ selected HTCL. The position of the message is shown on the left in Kb. Stripped filters were reprobed with an actin cDNA insert to determine loading accuracy (**lower**). The location of β, γ (2.1Kb), and α (1.6Kb) actins is marked on the left. The location of ribosomal RNA 28s and 18s is marked on the right.

DISCUSSION

A. BACKGROUND

A polypeptide doublet of 21KDa (designated $C4^h$ and $C4^l$) found across a range of species in a variety of tissue and cell types was shown to be localised along the entire length of actin filament bundles (Shapland et al, 1988). The higher molecular weight isoform ($C4^h$) was found to be down-regulated in mesenchymal cells after an enforced shape change, following oncogenic transformation with the RNA tumor virus RSV or the DNA tumor virus SV40 and highly motile cells such as lymphocytes (Shapland et al, 1988). This protein has been purified and since it induces the gelation of actin filaments (Shapland et al, 1993) it has been called transgelin. The cDNA encoding transgelin has been sequenced (Prinjha et al, 1994).

In contrast the lower isoform ($C4^l$) is present in all the cell types in which $C4^h$ is down-regulated, and is thus apparently insensitive to oncogenic transformation (Shapland et al, 1988). This project describes the partial purification and complete cDNA sequence encoding $C4^l$, its tissue specific expression patterns, and the relationship of $C4^l$ to other cytoskeletal proteins.

B. BIOCHEMISTRY

1. Immunofluorescence

The technique of immunofluorescence has been widely used in cytoskeletal studies (eg. Shapland et al, 1988; Tsukita et al, 1989; Quirk et al, 1993) and has the benefits of simplicity and speed, and can produce a 3D image of cytoskeletal networks.

SV40 transformed 3T3 cells were used in these studies to investigate the distribution of $C4^l$ within the cell. Classified as type 2 or 3 cells, using the criterion of Verderame et al (1980) which classes cells on the number of SFs present, SV40 3T3 cells were used because they do not express transgelin (Shapland et al, 1988; Prinjha et al, 1994; this study: fig 2). Therefore, by using these cells, any staining present after probing with the anti-C4 monoclonal antibody should indicate the distribution of only $C4^l$. The SV40 3T3 cells contain vestigial actin cables, and since they are flat cells, these are easily seen compared with the more rounded cell types that do not express transgelin

such as lymphocytes (Shapland et al, 1988) Thus, staining to F-actin structures in SV40 3T3 cells can be easily identified. At the resolution of the light microscope $C4^l$ was found to be distributed non-periodically along the lengths of the vestigial actin fibres (fig. 1). In contrast to the non-periodic distribution of $C4^l$, most microfilament-associated proteins show a periodicity along the length of the structure; for example, α-actinin (Blanchard et al, 1989) and tropomyosin (Lees-Miller & Helfman, 1991).

The intensity of anti-C4 monoclonal fluorescence is generally consistent along the length of the vestigial actin fibres with no greater staining associated with regions of filament termination such as focal contacts (as seen with α-actinin (Blanchard et al, 1989)). Staining was also noted within the nuclei of SV40 3T3 cells (fig. 1 (d)) and may possibly indicate a nuclear population of $C4^l$ within these cells.

In these experiments it is possible that the permeabilisation process routinely used (methanol) may have altered the in situ distribution of $C4^l$ by the non-specific precipitation of proteins to microfilaments. This, however, seems unlikely as cells were fixed in formaldehyde which results in lysine residues becoming cross-linked and thus the protein is immobilised.

In conclusion, the immunofluorescence experiments strongly suggest that $C4^l$, like transgelin, is a microfilament-associated protein. Its presence in cells that lack long actin cables such as transformed fibroblasts and lymphocytes (Shapland et al, 1988) indicate that $C4^l$ would bind to the F-actin structures present, although in lymphocytes (Dr. S. de Petris, personal comm) these F-actin arrays are below the limit of resolution of the light microscope (0.2μm).

2. $C4^l$ Tissue Source Selection

SDS-PAGE/immunoblotting shows that affinity purified anti-C4 polyclonal recognises a closely spaced doublet in REF (fig 2) which is unaffected by (i) rapid solubilisation of cells in SDS-sample buffer; (ii) high concentrations of reducing agents; and (iii) a variety of proteolysis inhibitors.

As a source of $C4^l$ I used transformed lymphocytes (HTCL) since (i) $C4^l$ is the only protein C4 isoform present (fig 2), (ii) this cell line grows to high confluence, and (iii) HTCL cells are larger than non-transformed lymphocytes and, I found, express more of the $C4^l$ isoform.

3. Detergent Solubility

When HTCL cells were extracted for 40 minutes in 1% CHAPS, and when both the soluble and insoluble components were analysed separately on SDS-PAGE (fig 3), it was found that all the detectable $C4^1$ was present in the soluble fraction. It is therefore concluded that after extraction with the zwitterionic detergent, no $C4^1$ remains associated with the cytoskeleton.

4. Proteolysis

Ultimately, all proteins are degraded, either intracellularly or extracellularly, and thus, under appropriate conditions, can be completely and effectively hydrolysed to amino acids by proteolytic enzymes (Harris & Angal, 1989). In vivo the degradation of a protein is progressive and subject to its contact to the degradative machinery. When a tissue or a cell is disrupted (lysed) the proteins may be brought into contact with new proteolytic enzymes from a different intracellular component, extracellular space or cell. As a result there may be peptidase attack on the protein of interest. Losses of protein or biological activity could be attributed to proteolysis during extraction and purification. Effective control of proteolysis in vitro is therefore necessary for the purification of a target, and requires at least some information about the protease(s) that are responsible. For the purposes of this study, a range of various proteolytic inhibitors were tested and the most effective utilised. Opportunities for the loss of inhibitor are many and include dialysis, gel permeation and ion exchange chromatography (Harris & Angal, 1989). It is therefore preferable to select a cocktail of inhibitors which targets the widest field of common proteases (see table 1).

The effectiveness of this protease inhibitor cocktail in minimising $C4^1$ proteolysis is shown in figure 4. $C4^1$ maintains its integrity for up to 50 hours when the proteolytic inhibitors are present. In their absence, $C4^1$ readily undergoes proteolytic cleavage (note the existence of low molecular weight bands containing the epitope for the C4 monoclonal antibody, fig 4).

5. Protein Purification

Since both protein C4 molecules are recognised by a monoclonal antibody they must share at least one common determinant. Initially, therefore, I used oligonucleotides derived from transgelin sequence to try to identify the gene encoding $C4^1$. This was unsuccessful (discussed below section G. 3) and dictated that I partially purify $C4^1$ in order to obtain amino acid sequence data from which oligonucleotides could be derived.

5.1. Extraction

The first step in the purification of a protein is the preparation of an extract containing the protein in a soluble form. $C4^1$ is shown to be completely soluble upon 1% CHAPS lysis of HTCL cells (discussed above, section B. 3). $C4^1$ extraction was carried out by detergent lysis, mechanical cell disruption was avoided as it may cause local overheating with consequent denaturation of protein.

Once released from its native environment an intracellular protein, such as $C4^1$, will be subject to many inactivating conditions. Within the cell, the pH is maintained at ≈6.5-7.5, the protein concentration [illegible] with the reducing potential also high. On disruption into buffer the protein concentration is reduced and the protein experiences an oxidising environment. Extraction will also lead to the breakdown of cellular components and thus expose the target protein to proteases. Wash buffer (methods, section G. 3), a buffer approximating to physiological conditions, was utilised for $C4^1$ extraction in an effort to minimise denaturation and/or proteolysis of intracellular $C4^1$. Wash buffer, in the presence of protease inhibitors, at 4°c and a pH of ≈7.0, ensures that $C4^1$ remains stable, with pH denaturation and proteolysis kept at a minimum. The presence of the reducing agent dithiothretol (DTT) and EDTA to chelate metal ions in wash buffer ensures that $C4^1$ is not rapidly oxidised and thus rendered inactive. In addition, by lysing HTCL cells at a ratio of every 7×10^7 cells:0.5ml of wash buffer + 1% CHAPS ensured that the resultant lysate was at a high enough protein concentration so as to avoid adsorption to surfaces and dissociation of subunits (see Harris & Angal, 1989).

HTCL lysate was clarified, with detergent insoluble material and particulates removed by centrifugation. Fat particles may be formed, which will float to the surface during centrifugation (Harris & Angal, 1989); these were removed by coarse filtration through glass wool and use of a 0.45μM millipore syringe filter.

Detergents, used to solubilise proteins so they can be more easily purified, may interfere with purification procedures. They can prevent acid or salt precipitation of the protein, or they may form a precipitate themselves. In all cases, the extraction and solubilising of $C4^1$ from cells was carried out with the zwitterionic detergent CHAPS which was removed from the soluble material by dialysis for 16 hours at 4^oc against several changes of 2 litres of 25mM ethanolamine HCl pH 9.6 (chromatofocusing start buffer) in the presence of protease inhibitors.

5.2. Chromatofocusing

The goal of a purification strategy is to obtain a maximum yield of protein with maximal purity. In order to achieve this, a choice of purification steps can be made. Each protein has a unique combination of properties which can be exploited for its purification. Thus, by combining a series of steps which exploit several of these properties, such as charge, pI, size, hydrophobicity and bioactivity, the protein can be isolated.

After 1% CHAPS extraction, clarification and dialysis into 25mM ethanolamine HCl pH 9.6, HTCL lysate was applied to a 1.6 x 70cm ($2.0cm^2$ cross-sectional area) chromatofocusing column. Chromatofocusing, which exploits a protein's specific pI, can be considered as an extension of isoelectric focusing (IEF) and ion exchange chromatography (Harris & Angal, 1989). In IEF, proteins are separated by electrophoresis in a pH gradient in a matrix produced by a current. In ion exchange, proteins are bound to the column at an initial pH and may then be eluted by a changing pH gradient. This is produced by mixing a limit buffer (of different pH) with the initial buffer and then pumping this through the column. The change in mobile phase pH partitions the protein into the mobile phase where it is eluted from the column. In chromatofocusing the pH gradient is produced inside the column by mixing an anion exchange matrix pre-adjusted (in start buffer) to one pH, with an elution buffer at a second lower pH. In the case of the purification of $C4^1$, the matrix (PBE 94) is pre-adjusted to pH 9.6 by 25mM ethanolamine HCl and a gradient formed by the eluting buffer which is polybuffer 96-HCl pH 7.0.

Upon eluent addition a gradient will eventually be generated throughout the column. When $C4^1$ reaches a pH just above its pI it will become negatively charged and

bind to the positively charged column matrix. As elution buffer flows through the column the bound protein will experience a lowering of the pH until it becomes positively charged and moves with the buffer down the column to reach a second point where the pH is just above its pI, and adsorption occurs. This sequence of events is repeated many times until the protein is eluted from the column at its isoelectric point. Thus, a focusing and peak sharpening effect operates during the separation.

Protein elution from the pH 7-9.6 chromatofocusing column was recorded and peaks analysed by SDS-PAGE (fig 5). The $C4^l$ positive peak was seen to be eluted from the column at a pH of approximately 8.0. Contaminating polybuffer can be separated from proteins eluted from a chromatofocusing column by a variety of methods (see manufacturer's specifications) and also by applying protein samples to hydroxylapatite (as discussed below, section B. 5. 3).

5.3. Hydroxylapatite Chromatography

Hydroxylapatite (HA) $[Ca_{10}(PO_4)_6(OH)_2]$, a form of calcium phosphate, is widely used as a purification and final concentration step (Harris & Angal, 1989). It also displays negligible adsorptive capacity for low molecular weight substances such as salts and polybuffer. The mechanism of protein adsorption onto HA is thought to involve both Ca^{2+} and PO_4^{3-} groups on the crystal surface. It has been suggested that acidic and neutral proteins bind to the HA calcium whilst basic proteins adsorb to surface phosphate groups.

For $C4^l$ purification a small (3ml) HA column was equilibrated in 10mM potassium phosphate buffer at pH 6.8. The column was carefully poured with all fines removed, to ensure a stable column bed, an even flow rate and a steady 'base line'. The $C4^l$ positive chromatofocusing fraction that had previously been dialysed against 10mM potassium phosphate pH 6.8 for 16 hours, was applied to the column at a slow linear flow rate so as not to disturb the column bed or induce compression. The HA column was eluted using 4ml steps of increasing concentrations of potassium phosphate buffer pH 6.8. Elution was recorded, fractions collected, and analysed by SDS-PAGE and immunoblotting (fig 6). The major band seen in figure 6 migrates at the same position in immunoblots as $C4^l$ (see fig 2) indicating that $C4^l$ had been isolated. Fractions positive for $C4^l$ were pooled (1.5-2ml in total), concentrated and buffer exchanged into 10mM

Tris pH 8.0 by use of an Ultra Free-MC 10,000 NMWL filter unit (Millipore). The purity of $C4^1$ was estimated at >90% homogeneity with a yield of ≈10μg from 5 x 10^9 HTCL cells (Dr. J. Hsuan personal comm), and thus deemed suitable for microsequencing.

6. Microsequencing

All amino acid sequencing was carried by Drs. J. Hsuan and N. Totty (Ludwig Institute for Cancer Research, London).

Purified $C4^1$ was applied to an automated protein sequencer. Material was reduced and alkylated with 4-Vinylpyridine as described (Andrews & Dixon, 1987) prior to digestion with trypsin. The resultant peptides were separated by reverse phase HPLC using an Applied Biosystems OD. 300 3.1x 100mm column employing a linear gradient of acetonitrile and monitored by a diode array detector (HP. 1090/1040A). Peptides were sequenced using an Applied Biosystems 477A using modified chemistry and analysis cycles (Totty et al, 1992). This approach, in conjunction with N-terminal sequencing, generated eight amino acids of N-terminal data and nine major peptides giving a combined total of 106 amino acids (see table 2).

The microsequence data for $C4^1$ provides 53% of the total molecule (when compared to the full length sequence; as discussed below, section C. 4), and thus, is therefore very useful information when searching for suitable sites of $C4^1$ specific degenerate oligonucleotides.

C. MOLECULAR BIOLOGY

1. Oligonucleotide Selection

The use of the triplet codons for encoding amino acids results in degeneracy (one of four nucleotides at each of the three positions of a codon allow 64 possible codons for 20 amino acids). As a result of this system, amino acids such as methionine and tryptophan are both encoded by a single triplet codon, whilst amino acids such as serine, leucine and arginine are each encoded by six triplet codons (Alberts et al, 1994).

When peptides to be used for oligonucleotide generation are selected, the codon possibilities of that peptide must be considered in order to indicate suitability. The $C4^1$ peptide ENFQNWLK has a codon:amino acid ratio of 2.38 (table 2). Reverse translation

of this peptide, and removal of the last base of the final amino acid, gives an oligonucleotide of 23 nucleotides almost spanning the eight amino acids (23 of 24 base pairs) containing seven regions of variability with 256 fold degeneracy. The inclusion of the neutral base inosine (which is able to pair with all four nucleotides with almost equal strength) at one of these positions reduces the degeneracy of the oligonucleotide (designated oligonucleotide 1) to 64 fold without destabilising a potential hybrid (Martin et al, 1985). The use of an inosine base also has the advantage of enabling the calculation of the melting temperature of the oligonucleotide/target hybrid to be more accurate, and thus, reducing the possibility of false positives when screening libraries or promiscuous priming during PCR reactions (Sambrook et al, 1989).

A second $C4^1$ peptide GASQAGMTGGYGMPR has a codons:amino acids ratio of 3.57 (table 2). A region 23 nucleotides in length (spanning the last nine residues of this peptide) contains five positions of variability with 512 degeneracy. Use of the neutral base inosine at four of these positions reduces the degeneracy to 2 fold. An antisense oligonucleotide (designated oligonucleotide 2) complementary to this region was generated. Short oligonucleotides containing inosine bases have been used in the screening of cDNA libraries (Ohtsuka et al, 1985). By generating two $C4^1$ specific degenerate oligonucleotides, one of which is antisense, the possibility to carry out PCR reactions exists. Degenerate oligonucleotides have been widely used in PCR amplifications (Erlich, 1989; Sambrook et al, 1989)). However, degenerate oligonucleotides containing inosine bases are rarely used in reverse transcription PCR (RT-PCR) amplifications (Rappolee et al, 1991; Erlich, 1989); for an example, see this study figure 7, where such oligonucleotides have been successfully utilised in PCR reactions inorder to obtain $C4^1$ specific PCR products.

2. PCR

The polymerase chain reaction can be used to amplify a segment of DNA between two regions of known sequence. Two oligonucleotides complementary to sequences on opposite strands are used as primers for a series of synthetic reactions catalysed by the thermostable *Thermus aquaticus* DNA polymerase (Taq polymerase) (Saiki et al, 1988; Erlich, 1989). Repetitive cycles of denaturation, primer annealing and

extension (with the products of each cycle acting as the template for the next) exponentially amplifies the target DNA producing a concentrated and effectively homogeneous solution of DNA, whose ends are delineated by the primers.

Degenerate oligonucleotides derived from $C4^1$ microsequence data could have been used to screen libraries in order to obtain full length cDNA. There are several problems in screening libraries , particularly with degenerate oligonucleotides (Sambrook et al, 1989). In many cases clones obtained from screening are not full length, lacking the 5'-end. This may well reflect degeneration of mRNA during cDNA synthesis, and in conjunction with low processivity reverse transcriptase enzymes and oligo-dT directed priming, commonly results in the absence of 5'-sequences in cDNA populations (Frohman et al, 1988). This problem can be overcome by rescreening the same or alternative libraries to obtain 5'-sequences (Fechheimer et al, 1991; Gu et al, 1991). Degenerate oligonucleotides are pools representing all possible codon choices for a short amino acid sequence. Although use of these probes in screening has been successful, there is considerable uncertainty in the hybridisation conditions because the binding of the oligonucleotides depends on the length of the hybrid formed and the G.C content of the probe. Empirically determined formulas allow for estimation of the oligonucleotide dissociation temperature; however, these methods can be unsatisfactory, making suitable stringent and selective hybridisation conditions difficult to find for all members of the pool. Thus, a large number of false positives can occur when screening complex libraries. In addition, only a minor fraction of cDNA gives rise to recombinant clones, of which 10^6 may have to be screened in order to find clones corresponding to rare mRNAs (Sambrook et al, 1989). Due to these problems library screening PCR techniques were utilised to obtain full length $C4^1$ cDNA.

2.1. RT-PCR Amplification

The efficiency of PCR in amplifying DNA presents an obvious need for precautions and controls to exclude the possibility of contamination by, and amplification of, extraneous DNA sequences. Thus, all PCR reagents were stored away from the area in which template stocks were handled. At least one control reaction was always included with every set of reactions to test for possible contamination of stock reagents. Reaction

conditions, particularly primer annealing temperature and magnesium ion concentration were empirically tested and optimised for each pair of oligonucleotide primers used. The final magnesium concentration used was adopted to maximise yields without compromising specificity, mispriming, visible as extraneous bands or smearing (Erlich, 1989).

HTCL (the source of $C4^l$ protein purification) RNA was used in all PCR amplifications. First strand cDNA was synthesised by use of a poly dT primer that takes advantage of the poly A sequence that is located at the 3'-end of most mRNAs (Frohman et al, 1988). Both oligonucleotides 1 and 2, being degenerate, were used at high concentrations (1μg per reaction) in all RT-PCR reactions to ensure the presence of the absolute primer within the oligonucleotide pool. A temperature of 47°c ensures sufficient primer annealing and also minimal primer promiscuity. Use of degenerate oligonucleotides containing inosine bases is rare in obtaining internal cDNA sequence in RT-PCR reactions, as it is thought the feasibility of this approach may be limited by non-specific amplification due to primer mismatch (Frohman et al, 1989). Most RT-PCR reactions use absolute primers complementary to known cDNA sequence; examples being the generation of full length transgelin cDNA (Prinjha et al, 1994) and the isolation I-plastin cDNA sequence (Lin et al, 1994). This study therefore provides an example where degenerate oligonucleotides can produce internal coding region cDNA of a target molecule (in this case $C4^l$) in RT-PCR reactions (see fig 7).

2.2. 5'-Anchor PCR

Clone 1 codes only $C4^l$ internal sequence (fig 12), therefore both coding region termini had to be achieved separately. The 5'-terminus of $C4^l$ was obtained by use of the $C4^l$ gene specific oligonucleotide 4 in 5'-anchor PCR amplifications in conjunction with a dC anchor primer (return primer) complementary to a dGTP tail added to the 3'-end of the first strand cDNA (fig 8).

Several strategies have been devised to add a determined sequence at the 3'-end of the cDNAs generated from first strand synthesis. Moss et al (1991) has described the addition of a homopolymeric dG tail, Frohman et al (1988) describes the use of a homopolymeric dA tail. Both use terminal deoxyribonucleotide transferase (TdT). Although tailing reactions were carried out successfully in this study, one must consider

that this method has important limitations. Firstly, the TdT reaction is difficult to control and has low efficiency; use of a cobalt buffer and short reaction times are necessary. A 15 minute reaction time was found to generate a $dC_{(13)}$ tail in my hands. Secondly, the return primers containing a homopolymeric (dC or dT) region generates non-specific amplification, and thus determination of an appropriately high annealing temperature is essential (56°c in the case of this study). In addition, the presence of an oligo dG tail may after a significant number of cycles, exhaust the complementary primer which would leave the cDNA in a single stranded state (Belyavsky et al, 1989). For this reason amplifications were carried for no longer than 30 cycles.

Finally, 5'-anchor PCR has been widely used in obtaining the 5'-termini of ABPs. For example, I-plastin (Lin et al, 1994), and the α-subunit of CapZ (Casella et al, 1989), where the technique was utilised to provide information that was missing in clones gained from cDNA libraries.

2.3. 3'-End RACE

The 3'-termini of cDNAs can readily gained by use of the rapid amplification of cDNA ends (RACE) system (Frohman et al, 1989). This system has been successfully utilised by Rana et al (1993) to obtain the missing 3'-sequence of dematin. The 3'-terminus of $C4^l$ was therefore attained by this method (fig 9). Amplification was performed using the $C4^l$ gene specific oligonucleotide 1 that anneals to a site located within the cDNA molecule and the universal amplification primer (UAP) which is complementary to a site within the first strand synthesis primer (AP). Since the 3'-RACE system utilises the poly A tail region as an initial priming site, multiple amplification products may be synthesised, depending on the degree of specificity conferred by oligonucleotide 1. Specific amplification products can be obtained by 'nesting' with a second gene specific primer (in this case oligonucleotide 3).

A potential problem associated with the use of PCR amplified DNA in sequencing can occur because of the intrinsic error rate of Taq polymerase. This enzyme unlike *E. coli* DNA dependent DNA polymerase has no proof reading exonuclease activity. The error rate for Taq polymerase is estimated at $\approx 2 \times 10^{-4}$ nucleotides per cycle and results in an overall error frequency of 0.2% in a 30 cycle amplification (Saiki et al, 1988). Although this fidelity should be quite acceptable even for construction of

expression libraries (provided the clones to be isolated are not very long), for exact sequence information several cDNA clones should be analysed. For the purposes of this study, each type of PCR reaction was carried out separately three times with the products cloned into separate pUC19 plasmids. Analysis of the three separate clones indicates that no sequence error was seen in any case.

3. Cloning and Sequencing

It is possible to sequence double stranded DNA directly from PCR reactions (eg. Prinjha et al, 1994). However, the sequencing ladders produced by this technique are not consistently satisfactory due to the high background, possibly caused by minor products of the amplification. This can be overcome by using an increased amount of purified template DNA. For the purposes of this study, all PCR products were therefore cloned into pUC19 plasmid prior to sequencing. Cloning can produce a homogeneous population of plasmid DNA containing PCR inserts in large quantities that facilitates optimal sequencing reactions.

Blunt-end ligation was necessary for PCR products utilising oligonucleotides 1 and 2 (fig 7), as these oligonucleotides are degenerate. SmaI digestion results in the blunt-end formation of the polycloning region of pUC19 (Sambrook et al, 1989). Blunt-end ligation is less efficient than cohesive ligation, but efficiency can be increased by 'filling-in' the ends of the PCR product (ie. effectively removing the single deoxyadenosine at the 3'- end of duplex molecules that is generated by Taq polymerase). The use of hexamineacolbalt in the ligation buffer and increased DNA ligase concentrations will also extend the efficiency of blunt-end ligations (Sambrook et al, 1989).

Oligonucleotide directed double-stranded sequencing using dideoxynucleotide chain terminating components was the method of choice in this study (Sanger et al, 1977). Such a system allows the full characterisation of the $C4^1$ cDNA clones 1, 2 and 3 derived from PCR reactions. The quality of sequence data was found to be impaired by incomplete denaturation of the two strands of the purified template. This may serve to impede the progress of the polymerase enzyme causing it to pause or dissociate, resulting in extraneous bands within the sequencing ladder. These problems can be reduced by a procedure in which the linear insert DNA is denatured (in the presence of a single

oligonucleotide) by heating to 100°c and rapidly freezing the mixture in liquid nitrogen. The primer is able to rapidly anneal to its target as the tube is thawed to room temperature and diluted for use in the labelling reaction without any significant annealing of the opposite strands.

Identical sequence data in the overlapping regions of clones 1, 2 and 3 suggests they are all derived from transcripts of the same mRNA population. Comparison of the translated cDNA of clones 1, 2 and 3 to that of $C4^1$ microsequence data illustrates they are in 100% agreement (fig 11 B), suggesting that the three $C4^1$ clones encode for the full length HTCL $C4^1$ molecule.

4. Complete $C4^1$ cDNA Sequence

The complete cDNA sequence for $C4^1$ is presented in figure 11. A 5'-leader sequence of 58bps is seen preceding position 1, the first ATG start codon (fig 11. A), and analysis of the 5'- regions of 699 published mRNA sequences reveal that more than 75% of these leader sequences fall within the range of 20-100 nucleotides (Kozak, 1987). There is no upstream ATG start site from position 1 and this is consistent with the fact that in ≈90% of mRNAs examined, there are no extraneous ATG triplets upstream of the functional initiator codon (Kozak, 1986). A 597bp open reading frame (ORF) encodes for a 199 residue molecule, and ends in a TGA stop codon that is followed by an extensive 3'-untranslated region (3'-UTR).

5. Initiation of $C4^1$ Translation

Kozak (see Kozak, 1991 for review) has identified a sequence of 13 base pairs around an ATG start site that influences the efficiency of initiation. If the A base of the ATG start site is designated position 1, it has been found that positions -3 (either purine base) and +4 (G base) are especially critical for efficient translation, as indicated by point mutation within the protein proinsulin (Kozak, 1986). From figure 13 it can be seen that these critical residues are present correctly around the first ATG of $C4^1$ and this sequence conforms to seven of the thirteen positions (54%) of the wider sequence as defined by Kozak (1986). There is another ATG codon starting at base position 23 (fig 11 B), this codon is only surrounded by one of the critical Kozak bases (and only 6/13 of the wider consensus sequence). Also, if this were the correct start site for $C4^1$ the subsequent derived amino acids that would result from this ATG, and the following bases that would

have to be aligned in frame with it, would not match the amino acid data provided from microsequencing (fig 11 B). Therefore the ATG site at position 1 is likely to be the initiation site for C4^{l}; it is surrounded by a Kozak sequence and the derived amino acids resulting from this position exactly match microsequence data (fig 11 B). The 40s ribosome binds at the capped 5'- end of mRNA and scans the sequence until an ATG codon is reached (Kozak, 1986), where it stops and allows the other components of the ribosome to bind before translation starts (Kozak, 1987). This scanning model agrees with experimental evidence that ribosomes initiate exclusively at the 5'-proximal ATG codon when it lies in a favourable optimal context (Kozak, 1986) as it does in C4^{l} cDNA.

N-terminal microsequencing of C4^{l} partially purified protein identifies the peptide starting with GPAY residues (table 2). Alignment of this peptide with derived amino acid sequence from the C4^{l} cDNA ORF indicate they both start within four amino acids of each other (fig 11 B) and before the second ATG at base position 23. These data strongly suggest that this first ATG codon represents the in vivo initiation site for C4^{l} protein.

While protein sequences are invariably start with a methionine residue, N-terminal sequencing of purified mature, fully processed, proteins indicates that methionine is cleaved and the terminus acetylated in a number of cases; for example, in thymosin β_4 (Wodnar-Filipowicz, 1984), and in class I and II isoactins the N-terminal methionine and cysteine residues are cleaved (Herman, 1993). The removal of the first four amino acids (MANR), which are not present in the N-terminal peptide derived from microsequence data (table 2), may represent the effects of post-translational N-terminal processing and maturation of the C4^{l} gene product in HTCL.

6. Termination of C4^{l} Translation

The ORF of C4^{l} cDNA extends 597bps from the initiator codon to a TGA stop codon (598-600). In all 3'-end PCR clones sequenced, this stop codon is followed by a second, in-frame, TGA codon (670-672). This therefore confirms that translation cannot extend beyond this region in this particular reading frame and that the cDNA sequence for C4^{l} presented in figure 11 encodes for the full length protein.

7. 3'-Untranslated Region of C4l

105bp of C4l 3'-UTR is presented in figure 11. C. From analysis of 699 mRNAs from vertebrates, ≈75% of the 5'- leader sequences are within the range of 20-100bps in length (Kozak, 1987). In order to estimate the length of the 3'-UTR of C4l, one can assume there is ≈70bps of 5'- non-coding sequence (58bps is the maximum found in all of the HTCL 5'- anchor PCR clones examined; discussed above, section C. 5) followed by a 597bp coding region then, as northern blot analysis indicates that C4l message has an apparent Mr of 1.44Kb (fig 23), there must be ≈773bps of 3'- sequence beyond the stop codon. Large 3'-UTRs are seen in many ABPs, for example: the message for cofilin has 846bp of 3'-UTR , whilst in ADF there is 1236 bp (Abe et al, 1990); dematin message encodes 1088bp of 3'-UTR (Rana et al, 1993); a 3'-UTR of 1602 is found in smooth muscle caldesmon message.(Bryan et al, 1989). If one assumes that 200bp of C4l 3'-UTR is a poly A tail (the average is 100-200 nucleotides, Alberts et al (1994)) then there must be 573 bp of 3'-UTR encoded by C4l mRNA. The entire 3'-UTR for C4l has not been obtained (105 bps are presented, fig 11. C) as the 3'-RACE products (see above) were sequenced so as to allow me to obtain the 3'-terminus of the coding region, and thus elucidate the complete C4l ORF, and not the entire 3'-UTR.

A mechanism for controlling protein expression is the intracellular localisation of the translation event (Hill & Gunning,1993; Hoock et al, 1991). Transport of mRNA is probably a process that involves the 3'-UTR (Hill & Gunning, 1993). The existence of extensive 3'-UTRs within C4l, and the above mentioned ABPs, may be required to control the distribution and stability of the message either as part of the primary sequence or the secondary structures these form (see Fulton, 1993). This mechanism would therefore contribute to the complex spatial and temporal regulation of the cytoskeleton that occurs during such processes as embryonic development, cell movement and wound repair (reviewed by Singer, 1992). Localisation of mRNA (due to different 3'-UTR sequences) allows intracellular sorting of protein isoforms from a multigene family that are similar in protein coding sequence (eg. β- and γ- isoactins within myoblasts (Hill & Gunning, 1993)). This suggests that cells, such as REF, could discriminate between the equally expressed (fig 2), 66.8% identical (table 4), C4l and transgelin molecules at the

mRNA level, and this may have a role in the function of each specific protein.

The 3'-UTR of mRNA can also function as a tumor suppressor, as demonstrated by the fact that a 0.2Kb RNA transcribed from the α-TM 3'-UTR in the absence of protein coding sequence can inhibit anchorage-independent growth of neoplastic NMU2 cells (Rastinejad et al, 1993). Upon constitutive expression of single cDNAs encoding ABPs such as α-actinin and vinculin, tumorigenicity is suppressed (Gluck et al, 1993; Rodriguez-Fernandez et al, 1992b). It should be noted that in both of these cases the expressed cDNA construct included both coding and non-coding sequences. C4^{l} protein expression is apparently insensitive to transformation (Shapland et al, 1988; this study: fig 2) and it is therefore unlikely that its 3'-UTR would function as a tumor suppressor.

D. PHYSICAL PROPERTIES OF THE C4^{l} GENE PRODUCT

1. Molecular Weight

The 597bp ORF of C4^{l} encodes a 199 amino acid polypeptide (fig 11. B) with a deduced molecular weight of 22381Da. This calculated Mr is slightly higher than that determined for C4^{l} protein by SDS-PAGE (Shapland et al, 1988), but is in good agreement with the observation that C4^{l} is approximately 0.5KDa less than transgelin (22594KDa) in REF cells (Shapland et al, 1988). The difference in calculated Mr of the cDNA derived C4^{l} polypeptide and that observed from cell extracts fractionated by SDS-PAGE may be due to N-terminal processing where the residues MANR could be removed (as discussed above, section C. 5) from the mature molecule in vivo. If these amino acids are removed from the calculation the Mr value becomes 21908Da and this is in reasonable agreement with the value of 21KDa determined by electrophoresis (Shapland et al, 1988). The difference in observed and calculated Mr of C4^{l} may also be due to the non-uniform migration, seen for many proteins, through SDS gels that is caused by incomplete denaturation or inconsistent association of SDS with the polypeptide chain; for example, caldesmon migrates as a doublet with an apparent Mr of 140-142KDa on SDS gels, whereas cDNA sequencing gives molecular weights of 86.9-88.7KDa (see Marston & Redwood, 1991 for review).

2. Isoelectric Point

Knowing the amino acid composition of a protein it is possible to estimate its isoelectric point (pI) by statistically averaging the pKa values of all the charged amino acids. By use of the equation:

$$pI = \frac{1}{n} \{(X \times pK_{R1}) + (X \times pK_{R2}) + (X \times pK_{R3}) + ...+ (X \times pK_{Ri})\}$$

an estimate of the pI can be calculated. pKa is the dissociation constant of the side chain group for amino acids 1 to i, X is the number of such amino acids in the total proteins and n is the total number of amino acid types being considered.

Using this equation the pI of $C4^l$ can be estimated at 8.09. This calculated value is higher than the value of 7.0 derived using non-equilibrating pH gradient gel electrophoresis (NEPHGE) (Shapland et al, 1988). This discrepancy may be due to the fact that under the experimental conditions used to purify protein C4 from REF cells (affinity column purification on anti-C4 monoclonal conjugated to affigel) there is a much greater recovery of transgelin than $C4^l$ (Shapland et al, 1988), and thus, upon NEPHGE analysis there is difficulty in observing a $C4^l$ 'spot'. The pI of 8.09 is in absolute agreement with the observation that native $C4^l$ is eluted from a pH 7-9.6 chromatofocusing column at a pH of 8.0 (fig 5).

3. Charge Distribution

The $C4^l$ polypeptide contains twelve positively charged lysine (K) residues, eleven positively charged arginine (R) residues and zero histidine (H) residues; ten negatively charged aspartic acid (D) residues and eleven negatively charged glutamic acid (E) residues (see table 4). There are therefore 23 positive groups and 21 negative groups, giving a net charge of +2 at pH 7.0. These charged residues are evenly distributed throughout the molecule. There is no appearance of a cluster of five negatively charged groups or a cluster of five positively charged groups found within the transgelin molecule (Prinjha et al, 1994). A positively charged cluster may be of significance, as such a region has been implicated in the binding of molecules to actin filaments (Shapland et al, 1993; Muhlrad, 1991).

4. Cysteine Distribution

Comparison of the migration of $C4^l$ in reducing and non-reducing conditions

indicates an absence of any intra-chain disulphide bonds (Shapland et al, 1988). Upon sequencing of C4[l] cDNA the derived polypeptide can be seen to contain three cysteine residues at positions 38, 63 and 124 (fig 11 B) and this would therefore possibly indicate the potential for C4[l] to form one intramolecular disulphide bond, or intermolecular disulphide bonds resulting in the formation of cross-linked monomers, dimers or oligomers (see Bardwell & Bekwith, 1993). The presence of a free cysteine within C4[l] raises the possibility that it could bind to the actin cytoskeleton as a higher order structure (eg. a dimer or trimer) such as dematin (Rana et al, 1993). Free cysteine residues within C4[l] may also form a disulphide bond directly with a free cysteine within actin (there are six cysteine residues in rat γ-actin (see Sheterline & Sparrow, 1994)); such disulphide cross-linking has been observed between caldesmon (contains two cysteine residues) and actin in vitro (Graceffa & Jancso, 1991). However, a disulphide cross-link is unlikely to be a major contributor to the energy of association between C4[l] and actin since REF C4[l] is completely soluble in 0.1% Triton-X100 (Shapland et al, 1988) in the absence of reducing agents.

5. Hydropathy Plot

Measuring the hydropathy of a molecule allows one to identify the polarity of different regions, ie. the hydrophilicity and the hydrophobicity. The most stable state of a protein folding in water is when the maximum number of polar amino acid groups are orientated to the surface of the molecule, whilst the maximum number of non-polar side chains are folded within the molecule away from the surface (Doolittle, 1986). Hydropathic values at each amino acid are calculated by statistically averaging the index values for six adjacent amino acids (values from the PEP program Intelligenetics). This is carried out by moving along one amino acid at a time until overlapping values are obtained for the entire protein and can then be displayed graphically (Kyte & Doolittle, 1982). The hydropathy plot for C4[l] is presented in figure 17. By use of this plot, regions of interest can be examined, for example, areas of the protein that may be involved in interactions with other proteins should be present in hydrophilic stretches.

The plot representing C4[l] displays no highly hydrophobic regions that are characteristic of transmembrane proteins. Proteins that are transported across plasma membranes from the cell contains, in many cases, a highly hydrophobic N-terminal leader sequence, that is proteolytically cleaved during or after transportation (Alberts et al, 1994). For example, the Ca^{2+}-activated, F-actin severing protein, brevin (plasma gelsolin) possesses a hydrophobic N-terminal leader sequence. This allows it to be secreted into blood plasma and circulate around the body, possibly severing actin filaments that have been released from damaged cells (Kwiatkowski et al, 1986). C4[l] possesses a hydrophilic N-terminal region that would indicate a non-secreted, nor a membrane associated protein. This prediction is in complete agreement with the non-membranous distribution of protein C4 seen in mesenchymal cells (Shapland et al, 1988) and C4[l] in SV40 transformed 3T3 cells (this study: fig 1).

A possible actin-binding domain (residues 96-103) for C4[l] (discussed below, section E. 1) exists within a hydrophilic region, and may therefore be 'available' for interactions with actin. The three cysteine residues of C4[l] (residues 38, 63 and 124) all lie within hydrophobic regions of the molecule and would therefore not be present at the surface of the folded protein. This seems to correlate with the hypothesis that C4[l] does not interact with other proteins, namely the actin cytoskeleton, via disulphide cross-links (discussed above, section D. 4). Within the actin cross-linking protein dematin, the free cysteine lies within a hydrophilic region and is therefore 'available', and is indeed known, to form disulphide cross-linkages creating a trimeric structure (Rana et al, 1993). A putative nuclear localisation signal (residues 152-158), a basic sequence that allows targeting to the nucleus (discussed below, section E. 4), lies within the most hydrophilic region of C4[l] indicating an orientation on the surface of the folded molecule.

The potential phosphorylation sites (residues 160-163 and 180-185; discussed below, section E. 2) exist within the highly hydrophilic C-terminal region of C4[l] that would be accessible to kinase enzymes. Phosphorylation of these sites may alter the hydropathy of the region and induce conformational changes in the surrounding area, potentially in the putative nuclear localisation signal (152-158), modulating the ability of C4[l] to target to the nucleus. Indeed, in the case of the ABP cofilin, a nuclear localisation

signal becomes 'unmasked' due to dephosphorylation of adjacent sites that causes a conformational change (Ohta et al, 1989). Conformational change due to protein phosphorylation has also been seen to modulate the ability of a number of ABPs to interact with actin (eg. MARCKS (Hartwig et al, 1992)).

6. Secondary Structure Prediction

Many thousands of protein sequences are known and this wealth of biological information has highlighted the need for accurate and automatic methods to predict protein conformation and function from primary structure. A prediction of the secondary structure of $C4^1$ was carried out using a multiprediction program (Zvelebil et al, 1987) by Dr. M. Zvelebil at the Ludwig Institute for Cancer Research, London. The prediction utilises information available from a family of homologous sequences (see fig 19). The approach is based both on averaging the Garnier et al (1978) secondary structure propensities for aligned residues and on the observation that insertions and high sequence variability tend to occur in loop regions between secondary structures. Accordingly, an algorithm first aligns a family of sequences and a value for the extent of sequence conservation at each position is obtained. This value modifies a Garnier et al prediction on the averaged sequence to yield the predicted secondary structure for the molecule in question. However useful they may be in providing insights into the possible structure of globular proteins, secondary structure predictions are subject to significant errors and should be treated with caution (Kabsch & Sander, 1983).

The results of the predictive program for $C4^1$ is presented in figure 18. The results show that the protein starts with a turn that leads into an α-helical region (amino acids 10-37) which turns into the first short β-sheet region (amino acids 60-69). The central portion of the protein (amino acids 81-138) is essentially characterised by long regions of α-helix. A turn is followed by a further short α-helix (amino acids 139-164), further turns and short β-sheet regions (amino acids 172-177). The molecule ends in a β-sheet, followed by a turn. While β-sheets provide considerable structural stability, most interactions between proteins often occur by the co-alignment of α–helices (eg. see hisactophilin (Habazetti et al, 1992)). It may therefore be of significance that the potentially important region of $C4^1$ that may be involved in the binding to actin (amino acids 96-103; discussed in section E. 1), occurs in a region likely to exist as α-helix.

E. C4l SEQUENCE MOTIFS

1. Actin-Binding Domain

Actin-binding domains (ABDs) are structural units able to interact with actin (Vandekerckhove & Vancompernolle, 1992).

There are several different approaches that have been used to study the binding of ABPs to actin. Definitive information comes only from atomic models derived from actin crystallised with an ABP (eg. DNaseI; Kabsch et al, 1990). These data are potentially definitive in terms of identification of the amino acids involved in bond formation, but they provide limited or no information of interactions, and thus, may only represent one of several possible structural interactions between the ABP and actin (Sheterline & Sparrow, 1994). In addition, it is possible to ask whether pairs of proteins (or protein fragments) compete for the same binding region on actin (see Vandekerckhove & Vancompernolle, 1992). Evidence can also be gained by looking for common actin-binding sequences on different ABPs, and by use of specific peptides testing them by falling ball viscometry experiments (eg. as for a hexapeptide present in cofilin (Yonezawa et al, 1991)). A list of identified ABDs can be found in Vandekerckhove & Vancompernolle (1992) and Sheterline & Sparrow (1994).

C4l has been shown to be associated with vestigial actin bundles (fig 1). If C4l is associated with actin directly (ie. it is not localised to F-actin via interaction with other proteins), it must contain an ABD, a sequence of amino acids that allows interaction with the actin filament.

Homology to the sequence within C4l, ^{96}FLQAAERY^{103}, is found in a range of ABPs (including transgelin), and is shown in figure 11. It is present in a number of tropomyosin isoforms (Ruiz-Opazo & Nadal-Ginard, 1987). Conserved homology can be found repeated through the tropomyosin molecules. This is representative of the fact that the tropomyosin molecule consists of six or seven (nonmuscle- and muscle isoforms respectively) linearly repeated ABDs that bind the sequential seven (or six) actin subunits along the major groove on each side of the actin filament in each half-turn of the helix (Kabsch et al, 1990; Milligan et al, 1990). Less well conserved homologies are found 4x in caldesmon (fig 14). Caldesmons are a family of elongated molecules that contain a

core region composed of eight thirteen amino acid repeats containing alternating basic and acidic residues and play a major role in the regulation of contraction in smooth muscle (Wang et al, 1991). Homology to this peptide is also found in yeast ACT2, a type III actin molecule (Lees-Miller et al, 1992b).

Mezgueldi et al (1992) highlight the homologous VKYAEK motif of calponin (38% identical to C4[1], see table 5), present at the C-terminus of a 22KDa N-terminal fragment that is known to bind F-actin.

Two regions very much like this C4[1] peptide are found in actobindin (fig 14). Actobindin is a small (88 residues), *Acanthamoeba*, monomer-binding protein that is composed of two near identical 33 amino acid repeats (Vandekerckhove et al, 1990b), and is capable of inhibiting an early stage of actin polymerisation by sequestering nucleation competent dimers (Vancompernolle et al, 1992). It appears to have two consensus putative actin-binding sequences, LKHAET (Vancompernolle et al, 1991). Similar sequences are seen in fimbrin, plastin, villin Tβ4, myosin and α-actinin (Vancompernolle et al, 1991; this study: fig 14). Actobindin can be cross-linked both to glutamate 100 and to the N-terminal acidic residues 1, 2 or 3 on the barbed end of actin subdomain IA (Vancompernolle et al, 1991). Interestingly, mutagenesis of one of the amino acids in the sequence LKHAET motif of either actobindin site converts an inhibition of assembly into promotion (Vancompernolle et al, 1992), suggesting an abolition of actin binding. Actobindin remains the only protein in which this motif has been closely apposed to actin, but its existence in several ABPs may be of 'great significance' (Vandekerckhove & Vancompernolle, 1992).

Falling ball viscometry experiments using a synthetic hexapeptide (DAIKKK) found at the N-terminus of tropomyosin and near the centre of cofilin indicates that this peptide could reduce the pH dependent binding of cofilin to F-actin, but could not affect the protein's depolymerising activity. However, a synthetic dodecapeptide specific to cofilin, while having no effect on the ability of cofilin to bind actin, was found in polymerisation assays to depolymerise F-actin in a pH independent manner (Yonezawa et al, 1991) suggesting the need for the cooperative involvement of distinct regions of the protein for biological function.

The N-terminus of actin contains four (class I, muscle) or three (class II, nonmuscle) acidic residues which can be cross-linked to a number of ABPs using a zero-length crosslinker (Sutoh & Yin, 1989). The N-terminus of actin has a high terminal motion and probably forms a highly mobile structure and could therefore react with a number of ABPs (Kabsch et al, 1990). A number of ABPs including transgelin (Prinjha et al, 1994), ABP30 (Fechheimer et al, 1991), fimbrin (de Arruda et al, 1990) and the caldesmon repeats (Bryan et al, 1989) possess acidic regions. These regions may either be involved in regulating the binding of these proteins with basic residues in actin (Sheterline & Sparrow, 1994), or with displacing other ABPs associated with the N-terminus of actin. It should be possible to test the functional importance of these regions by use synthetic peptides during in vitro competitive assays such as falling ball viscometry. Transgelin also contains a positive cluster of residues (Prinjha et al, 1994) that are possibly involved in the binding or gelation of actin, in an association that is likely to be electrostatic in nature (Shapland et al, 1993). Similar results have been found for the 21KDa N-terminal fragment of myosin (Muhlrad, 1991). No such acidic regions or positive residue clusters are found within the sequence of $C4^1$, therefore the only possible ABD so far identified is the FLQAAERY motif. This region may represent a complete or partial actin-binding sequence. If $C4^1$ is to cross-link and gel actin in the same manner as transgelin, which is monomeric and possibly contains more than one ABD (Shapland et al, 1993), it must contain another region which could function as an ABD; or it could cross-link actin filaments as a higher order structure such as a dimer or trimer, using cysteine residues (as discussed in section D. 4), in which the single ABD would be present more than once (as with dematin (Rana et al, 1993)).

2. Phosphorylation Sites

Protein phosphorylation is recognised as a major mechanism through which extracellular agents and events such as hormones, growth factors, chemotactic stimuli and viral infection influence intracellular events such as movement, cell growth and tumorigenesis (Woodjett et al, 1986; Kellie et al, 1991). The kinase substrate target sequences for Ser/Thr phosphorylation by protein kinase C (T/S-X-R/K; where X represents any residue), calcium/calmodulin dependent protein kinase II and cAMP

dependent protein kinase A (R/K-X-X-S/T) are well known (Woodjett et al, 1986; Carlson et al, 1979).

Within C4[l] a putative calcium/calmodulin dependent protein kinase II is found at residues 160-163 (RNFS). The sequence in C4[l], ^{180}TNRGAS^{185}, contains both a threonine protein kinase C and a cAMP dependent protein kinase A site immediately adjacent to each other. The kinases mentioned above respond to stimuli such as cAMP and calcium and have been shown to induce cytoskeletal reorganisation (Hitt & Luna, 1994; Titus, 1993a; Blackshear, 1993; Bennett, 1989).

The activity of many ABPs [illegible] to be regulated by phosphorylation, as indicated by some examples described below. Dematin, a human erythroid membrane skeleton protein, has its actin binding activity abolished upon phosphorylation by cAMP dependent kinase A, and restored again after dephosphorylation (Rana et al, 1993). Phosphorylation of myosin light chain by calcium/calmodulin dependent myosin light chain kinase regulates actomyosin interactions in smooth muscle and nonmuscle cells (Kanoh et al, 1993). Phosphorylation of MARCKS by protein kinase C results in its translocation from membrane actin structures (Hartwig et al, 1992). Phosphorylation of caldesmon during mitosis causes its dissociation from microfilaments (Yamashiro et al, 1990), and thus, phosphorylation of C4[l] at mitosis may result in dissociation from F-actin and therefore cause the cell shape changes seen during cell division (Satterwhite & Pollard, 1992). Actin-bundling activity may be regulated by the phosphorylation of l-plastin by Il1 and Il2 (Pacaud & Derancourt, 1993). Synapsin Ia phosphorylation by cAMP dependent and calcium/calmodulin dependent protein kinase II abolishes its ability to form thick bundles of actin filaments (Valtorta et al, 1992b). Phosphorylation of calponin (which is ≈39% identical to C4[l]; see table 5) by Ser/Thr protein kinase C abolishes its ability to bind to actin and inhibit actomyosin ATPase activity (Winder & Walsh, 1990; Nakamura et al, 1993). Changes in protein phosphorylation regulates the structure and function of focal contacts in vivo in response to ligand activation of integrin (see Hitt & Luna, 1994), the focal contact proteins that are phosphorylated include tensin (Davis et al, 1991) and paxillin (Glenney & Zokas, 1989). Dephosphorylation by specific phosphatases is also involved in cytoskeletal regulation. For example, type IA phosphatase induces the dephosphorylation of the myosin regulatory light chain and

alters cell morphology, due to the transient disappearance of SFs, following microinjection (Fernandez et al, 1990).

Neither C4[1] nor transgelin have been found to be phosphorylated by ^{32}P incorporation (Shapland et al, 1988); also, alkaline phosphatase digestion (an enzyme able to remove phosphate groups) of immunoblots of REF cells, probed with the anti-C4 monoclonal antibody, did not modify the characteristics of either protein C4 epitope (Shapland et al, 1988). This would indicate a complete absence of phosphorylation of C4[1]; however, the observation could be explained by the fact that the monoclonal antibody epitope is distinct from the possible phosphorylation sites within the molecule. It is also possible that putative phosphorylation of C4[1] only occurs at a certain point in the cell cycle (eg. as with caldesmon (Yamashiro et al, 1990)). If this is the case, alkaline phosphatase treatment would produce no alteration in the binding of anti-C4 monoclonal antibody to REF (a non-synchronous population of cells at various stages of the cell cycle) extracts from untreated cells in which phosphorylation was not at significant levels.

3. EF Hand Calcium Binding Domain

Calcium ion binding of an ABP is a possible regulatory mechanism that can modulate a particular protein's association with actin; for example, nonmuscle α-actinin forms are regulated by virtue of their EF hand motif, such that in the presence of calcium the level of actin cross-linking is decreased (reviewed by Hartwig & Kwiatkowski, 1991). The EF hand structure was first identified by Tufty and Kretsinger (1975) based on the conformation and sequence of regions in parvalbumin and troponin C. The liganding loop contains twelve amino acids of which seven positions are critical. Five of the twelve amino acids have carboxyl or hydroxyl groups in their side chains, that are spaced to coordinate the calcium ion (reviewed by Heizmann & Hunziker, 1991). An alignment of the parvalbumin and troponin C, which bind two and four calcium ions per mole respectively (Heizmann & Hunziker, 1991), calcium-binding sequences displaying the standard EF hand loop is shown in figure 15. The seven critical residues are indicated by an asterisk and the required amino acid residues displayed. The first aspartate (D) residue at position 1 is of particular importance in the formation of the liganding loop.

Mutation of this position abolishes the calcium binding of cardiac troponin C (Putkey et al, 1989) and yeast calmodulin (Geiser et al, 1991).

Alignment of mp20 (which is 38% identical to C4[1], see table 5) with the EF hand structure indicates that five of the seven critical residues (including the important D residue) are present twice within its sequence (Ayme-Southgate et al, 1989). Mp20 is proposed to be a calcium-binding protein, due to the fact it is *Drosophila* synchronous flight muscle specific, even though it does not contain all seven critical residues at the two sites (Ayme-Southgate et al, 1989). There are other examples of calcium binding proteins which have, at their known sites, an EF hand structure that does not conserve all seven of the critical residues, for example the F-actin bundling protein villin (Arpin et al, 1988).

Alignment of C4[1] with the EF hand structure (fig 15) indicates the existence of two sites (I and II) that possess a number of these critical residues. All seven critical residues are not necessarily required for a functional EF hand (see above), and thus, C4[1] may bind calcium. However, at site II, the first and most critical D residue exists as a T residue that would be unable to provide the oxygen atom necessary for calcium ion binding. C4[1] site I contains only three of the critical seven residues (including the important D residue) which is probably too few to be functional. This would therefore suggest that C4[1] protein would be unable to bind calcium. The two proposed EF hand structures of transgelin (Prinjha et al, 1994; this study: fig 15) are very similar to those of C 4[1] (site I of both molecules share 61.5% conservation; site II of both molecules share 92% conservation). Experiments using radioactively labelled calcium ($^{45}Ca^{2+}$) and purified transgelin blotted onto nitrocellulose confirm that is unable to bind calcium (Shapland et al, 1993); this is consistent with the observation that the actin-binding and gelling activity of transgelin are independent of calcium concentration (Shapland et al, 1993). As C4[1] shares such great similarity to the transgelin EF hand sites, it seems unlikely that its functions will be dependent on calcium binding (further confirmation will require the use of $^{45}Ca^{2+}$).

The EF hand structure is an ancient genetic motif; analysis of sequence conservation in different species indicates that the structure had been duplicated well before the divergence of the yeast and vertebrate evolutionary branches (Nakayama et al,

1992). It has yet to be established at what point in evolution, if there is any, that C4[l] possessed and lost its ability to bind calcium. The sequences found in C4[l], may however, have preceded the evolution of the EF hand and the molecule may never have required calcium binding for regulation of its function. By cloning and sequencing C4[l] homologues from organisms such as yeast, and using the ^{45}Ca approach one could test this possibility.

The annexin family, some members of which bind F-actin (Hosoya et al, 1992), contain a conserved domain of approximately 60-70 amino acid residues that includes the site for PPI association. The distinguishing feature of the annexin family is their ability to bind calcium without any significant sequence similarity to the EF hand motif, allowing them to interact with PPIs in a calcium dependent manner (Smith et al, 1990). Modulation of protein behaviour by calcium, can therefore, occur without an EF hand motif. If this is so, C4[l] may possibly bind calcium in vivo where protein interactions act synergistically; for example, a gelsolin fragment and actin can trap calcium between themselves (Way et al, 1990).

4. Nuclear Localisation Signal

Proteins can enter the nuclei either by passive diffusion or active transport across the channels of the nuclear envelope. If the molecular weight of a molecule is at the upper limit for free diffusion (≈39KDa; see Onoda et al, 1993), it may have to be actively transported after targeting with its own nuclear localisation signal (NLS) or bind to another protein with a targeting sequence (Garcia-Bustos et al, 1991).

Proteins are targeted to the nucleus by specific signals, the NLS, within the proteins' primary sequence. These signals have been delineated for an expanding body of proteins (see Garcia-Bustos et al, 1991), by two general methods. NLSs have been identified as sequences (i) that can by genetic or biochemical fusion render a cytoplasmic protein nuclear; or (ii) which when deleted or mutated, no longer promotes nuclear uptake of the protein in which they reside (such as cofilin NLS mutation; Iida et al, 1992). There is no single strict consensus NLS, but there are some general rules. NLSs are (i) typically short sequences, usually not more than 8-10 amino acids; (ii) contain a high portion of positively charged amino acids (K or R); (iii) are not located at specific sites within the protein; (iv) are not removed following localisation; and (v) can occur

more than once in a given protein (Garcia-Bustos et al, 1991). Comparison of eight homologues of the well known SV40 large T antigen NLS (Kalderon et al, 1984) has generated a four residue consensus sequence (see fig 16). In general, the most conserved features seems to be several positively charged residues associated with a proline. A putative NLS is found within $C4^1$ (amino acids 152-157) that fits this consensus sequence. The $C4^1$ NLS is shown aligned, along with two well known NLSs, to the consensus sequence in figure 16.

Rihs & Peters (1989) challange the notion that a minimal NLS is the lone determinant of T antigen import in vivo. They reported that 15 amino acid residues adjacent to the NLS sequence modulate the efficiency of the nuclear transportation. In this region pronounced phosphorylation sites were observed (Rihs & Peters, 1989). Phosphorylation of sites adjacent to an NLS could therefore act as a regulatory mechanism. A calcium/calmodulin-dependent protein kinase II site is present adjacent to the NLS of cofilin (KKRKK); a similar site is also found adjacent to the NLS of $C4^1$ (amino acids 160-163). Dephosphorylation of cofilin is seen to accompany its nuclear localisation following heat shock treatment, where it is proposed that the NLS becomes unmasked and active (Ohta et al, 1989). The same mechanism could apply for $C4^1$, although it has not been clarified whether the molecule is actually capable of being phosphorylated.

There are many studies which suggest that the nucleus contains actin (eg. Bremer et al, 1981; Amankwah & De Boni, 1994). In addition, an increasing number of ABPs have been isolated from nuclei, suggesting that actin structures are regulated. These include gCap39 (Onoda et al, 1993), cofilin and ADF (Ono et al, 1992), myosin (Milankov et al, 1991), NAB-34 (Rimm & Pollard, 1989), and members of the CapZ family (Ankenbauer et al, 1989). Although these proteins are found in the nucleus and undoubtedly have a regulatory role upon the nuclear actin network, their exact nuclear functions are unknown. The translocation of ABPs may act as intracellular effector mechanisms, regulated through transduction pathways. For example, cofilin nuclear translocation may exert an anti-apoptotic activity upon cells, where monomeric actin formed by its severing activity can inhibit a key enzyme involved in apoptosis, DNase I (Smatter et al, 1994). In cells such as T cell lymphomas, where cofilin is continuously

dephosphorylated and continuously present in the nucleus, a continual anti-apoptotic activity is effected; this may account for the prolonged life span that is often seen in these cells (Samstag et al, 1994). The presence of a possible NLS and an adjacent phosphorylation site, along with an initial observation of nuclear staining (fig 1), points to a C4[1] nuclear population; it may therefore be regulating the actin network of the nucleus along with the other the ABPs mentioned above.

F. HOMOLOGY SEARCHES

For the purposes of comparisons of C4[1] with other proteins, only the following substitutions were considered to be conservative: D-E; K-R; L-I-V-M-C; Y-F-W; S-T and A-G. By taking into account these conserved substitutions during comparison of C4[1] with other proteins, important regions within all of the considered molecules can be noted.

1. Homologous Proteins

A VAX 11/750 mainframe based suite of programs maintained by Intelligenetics and licensed to the Ludwig Institute of Cancer Research, London was used for the searches. Nucleotide sequences were analysed and manipulated using SEQ, amino acid sequences were handled using PEP. A number of programs are available for homology searches and they differ in their methodology that results in the exclusion or inclusion of slightly different homologous sequences, and therefore searches were always carried out using more than one program. The IFIND program uses an algorithm that sequentially compares overlapping segments of the proteins being investigated (Wilbur & Lipman, 1983). The FASTDB program uses an algorithm developed by Intelligenetics that uses a scoring matrix that statistically weights amino acid changes in terns of the number of codon requirements necessary to convert from one amino acid to another and thus, their evolutionary relatedness. It then selects sequences above a threshold level of similarity for more detailed analysis. The multi-stage analysis of FASTDB allows the detection of homologous and related sequences (Gribskov et al, 1987).

Analysis reveals that HTCL C4[1] is a unique cDNA that shows significant homology with four other proteins in the databases (fig 19).

C4[1] shows high homology with all the rat transgelin homologues that were identified as such by Prinjha et al (1994); namely chicken SM22α (Pearlstone et al, 1987), serum-induced mouse fibroblast protein (p27) (Alemendral et al, 1989), and human senescent fibroblast protein (WS3-10) (Thweat et al, 1992). Rat transgelin shares ≈95% conservation with its homologues (Prinjha et al, 1994). C4[1] and rat transgelin, an F-actin gelling, shape change and transformation sensitive protein (Shapland et al, 1988; 1993), have 66.8% identical amino acids (73.9% including conservative substitutions) over the 199 residues of C4[1]. There are no large regions of homology or difference between C4[1] and transgelin, they are spread randomly throughout both molecules (fig 19), indicating that the two proteins do not arise by alternative splicing of the same transcript (Andreadis et al, 1987). This is in agreement with the observation that transgelin message is not detected in cell types that only express C4[1] protein, such as SV40-3T3 cells (Prinjha et al, 1994). Therefore, although highly homologous, and obviously related, C4[1] and transgelin would appear to arise from distinct genes.

C4[1] has 67.0% identity (74.4% including conservative substitutions) at the amino acid level over 199 residues with rat neuronal protein NP25 (unpublished sequence, EMBL accession number M84725). At present NP25 has an unknown function, but its high degree of identity with C4[1] and transgelin indicates a related protein.

C4[1] has 38.6% identical amino acids (49.5% with conservative substitutions) over 184 residues with a protein of unknown function called mp20 (Ayme-Southgate et al, 1989). This *Drosophila* protein is found only in synchronous muscle (which is characterised by extensive sarcoplasmic reticulum and highly structured calcium regulatory machinery). mp20 is derived from a single gene that generates two transcripts (1.0 and 0.9Kb) with identical coding regions, the protein contains two putative calcium binding EF hands that possess five out of the seven critical residues including the important D residue (Ayme-Southgate et al, 1989) that is not found in site II within C4[1] (fig 15). Northern analysis of different *Drosophila* tissues indicates that the messages are developmentally regulated, with immunofluorescence data indicating that mp20 is present in larval gut muscle fibres and not in the parenchymal cells (Ayme-Southgate et al, 1989). The putative calcium binding of mp20, its distribution in striated muscle and

absence in smooth muscle are opposite to the known properties of C4[l] and transgelin making the interpretation of the sequence homology very complex. It may, however, be possible that there is a second C4[l] homologue in *Drosophila* that is more closely related to mammalian C4[l] that functions in nonmuscle and smooth muscle cells; highly related *Drosophila* homologues of other vertebrate ABPs have been identified (see Ayme-Southgate et al, 1989). The homology between C4[l] and mp20 may reflect their divergence from a common ancestor (with a possible calcium binding function) with the adoption of distinct functions in these two branches of evolution.

Calponin is a basic calmodulin-binding molecule that is also able to bind actin/tropomyosin and thin filaments, and inhibit myosin Mg-ATPase activity by reducing the cycling rate of the actomyosin ATPase (Walsh, 1991). C4[l] has 39.2% identity (49.7% with conservative substitutions) over 199 amino acids with calponin-β (the smaller isoform of the calponins, but it is identical in sequence to calponin-α over the range under comparison shown in fig 19). Calponin-α differs from the β-isoform by the inclusion of a 40 amino acid insert at residue 216. Calponin-α contains a 28 residue C-terminal tandem repeat segment that appears three times; calponin-β contains two repeats of this segment (Takahashi & Nadal-Ginard, 1991). Calponin mRNA tissue distribution is similar to that of transgelin with expression in smooth muscle and an absence in skeletal muscle (Prinjha et al, 1994).

It is interesting that the homologous regions between C4[l], transgelin, NP25 and mp20 span the entire length of the molecules, whereas the homologous regions between C4[l] and calponin-α are repeated three times near the C-terminal end of the latter and twice near the C-terminal end of calponin-β (the homologous regions are infact the 28 residue C-terminal tandem repeat segments found in the calponins). The functional significance of this is unclear, but the evolutionary implications may be that C4[l], transgelin, NP25 and mp20 share a common ancestor. Also, the presence of multiple C4[l] related repeats in the C-terminal domain of the calponins suggests that C4[l] may well have preceded calponin during evolution, with the latter extending at the C-terminus via non-homologous recombination events. With the presence of conserved putative EF hand motifs (see fig 15) and potential phosphorylation sites (see fig 11 B) it has been

suggested that transgelin, mp20, NP25, calponin, and thus C4^{l}, have evolved by gene duplication from a common ancestor, whose activity may have been regulated by calcium and/or phosphorylation (Prinjha et al, 1994). While transgelin and the calponins appear to have diverged functionally, a functional role for NP25, mp20 and C4^{l} remains, as yet, undefined. The findings that transgelin is an actin cross-linking/gelling protein (Shapland et al, 1993) and that C4^{l} is associated with actin bundles at the immunofluorescence level (this study: fig 1), coupled with the homology between transgelin and C4^{l}, may suggest an actin cross-linking/gelling role for the latter (with the same applied to NP25 and mp20). As the sequence discrepancies between C4^{l} and the proteins discussed above are distributed throughout the coding regions, and the 5'-leader and 3'-UTR sequences are completely different, it is highly likely that these proteins are all generated from distinct genes. However, given the fact that they all show high degrees of homology (see table 5), these proteins may be classified as members of a new multigene family, supporting the observation that all ABPs will eventually fit into a limited number of families (Vandekerckhove & Vancompernolle, 1992). By placing a unique protein, in this case C4^{l}, into a gene family, in this case the transgelin-related family (see introduction, section F. 20), one is able to identify possible significant homologous domains and therefore point to a possible function, that can be tested biochemically, of the protein of interest.

A 21 residue motif within the C-terminus (amino acids 173-193) of C4^{l} (also in transgelin, NP25 and mp20) is found three times in the C-terminus of α-calponin (as already discussed above). It also shares homology with a 21 residue region that is found seven times within a 379 amino acid alternatively spliced product from the *C. elegans* unc87 gene (unpublished, EMBL accession number U04711). The fourth such repeat within this unc87 gene product (at amino acids 198-218) has the highest degree of identity with the motif within C4^{l} (fig 20), where it is 66% identical (76% with conservative substitutions). The significance of this homology is unclear, but unc87 is known to be a cytoskeletal protein (Waterston et al, 1980) and thus could represent a *C. elegans* member of the transgelin-related family; functionally these motifs have been shown to be phosphorylated within calponin-α (Nakamura et al, 1993). The seven repeats

of this C4[1]-like motif within unc87 suggests that C4[1], transgelin, NP25 and mp20 (one repeat) and calponin (three repeats) may well have preceded unc87 during evolution.

2. Dot Plots

Dot matrix plots are an effective way of analysing amino acid sequences and showing a relationship between two proteins (Gibbs & McIntyre, 1970; Staden, 1982). One sequence is represented along the each axis (x, y) and a dot is marked at any x, y position corresponding to identity or homology. A solid line appears when identical sequences are compared. Background noise can be reduced by the addition of a filter function such that a dot is only marked if a set number of consecutive amino acids or base identities are observed (in this study 4-7), this filter value needs to vary with the window size (the number of amino acids compared at each point; in this study 8). The dot plot method is highly useful in that one can detect homologies between different sequences and also internal repeats within a single molecule.

Repeated sequences within proteins may indicate a mechanism by which a protein can elongate by tandem duplication of gene segments. The spacer domains of the α-actinin superfamily members are made up of repeated α-helical or β-sheet motifs of approximately 96 residues that are distinct from the ABD (Matsudaira, 1991). In the case of ABP120 and ABP280 (Hartwig & Kwiatkowski, 1991) the repetitive motifs serve to separate the two ABDs in the functional dimer. The length of the spacing can therefore regulate the rheological properties of the F-actin network formed by this family of cross-linking proteins (Wachsstock et al, 1993). The tandemly repeated 15KDa domains of the gelsolin superfamily members (six repeats within gelsolin and villin, with three repeats within fragmin), reflects the multifunctional ability of the family members to bind G or F-actin, and also cap or sever actin filaments (Weeds & Maciver, 1993). The repeat structure of fragmin indicates that it may have arisen from an ancestral 15KDa actin-monomer binding protein that was tandemly repeated twice to give a three domain molecule, sequence analysis indicates that it may have in turn duplicated to give a six domain molecule such as gelsolin (Hartwig & Kwiatkowski, 1991). The dot plot analysis of C4[1] (fig 21 a) was calibrated to only mark positions of identity, and it indicates that the molecule contains no internal repeats detectable above background noise. This could be

taken as an indication of a rapidly changing molecule; a suggestion that can only be established by considering the degree of conservation-evolutionary stability between $C4^l$ homologues of other species, and as yet, no such data is available. The lack of repeats within $C4^l$ may also indicate the presence of very ancient duplications that have diverged significantly and fail to be recognised (Doolittle, 1986).

Comparison of $C4^l$ with transgelin and NP25 (fig 21 b and c respectively), two molecules that are ≈67% identical to $C4^l$ (see table 5), indicates a clear but broken diagonal line representing extensive end-end homology.

Comparison of *Drosophila* mp20 and $C4^l$ (fig 21 d) set with a window size of eight amino acids indicates the existence of a partial diagonal line with 'islands' of homology. These are present in three areas at the respective N and C termini and in the central region. There is no indication of 'off-set' diagonal lines which suggests that neither protein contains homologous internal repeats. A similar comparison of $C4^l$ against calponin-α (292 amino acids) using a 4-7 filter with a window size of eight amino acids, displays the areas of maximum homology (fig 21 e). Calponin is a larger molecule and thus the scale on the y axis and hence homologous diagonals are condensed relative to those seen in mp20. Calponin-α has three islands of significant homology with $C4^l$. A striking feature of the comparison is that the third island of homology is repeated three times within the C-terminus of calponin-α. An identical comparison between $C4^l$ and an unc87 alternatively spliced gene product (379 residues) illustrates that the same $C4^l$-like motif, repeated in the calponins, is present seven times (albeit weakly) throughout the *C. elegans* molecule (fig 21 f). These regions may represent binding or regulatory sites; indeed, the three repeated domains within the C-terminus of calponin-α have been shown to be phosphorylated in vitro (Nakamura et al, 1993). The existence of a double phosphorylation site within this region of $C4^l$ (as discussed in section E. 2) allows the possibility for a regulatory system that either causes a conformational change or a dissociation from actin filaments. Such changes are seen in the case of the cell cycle specific phosphorylation of caldesmon (Yamashiro et al, 1991).

3. Homologous Peptide

The peptide ^{181}RGASQAG^{187} within $C4^l$ is also found in a number of other

ABPs. This sequence spans a cAMP dependent protein kinase A site (as discussed in section F. 2) and is found with 100% identity within the C-terminus (amino acids 183-189) of transgelin (Prinjha et al, 1994) and at the N-terminus (amino acids 32-38) of human gelsolin (Kwiatkowski et al, 1986). The peptide is also found, with almost 100% conservation, in the C-termini of the remaining members of the transgelin-related multigene family (fig 22). In the calponins, this peptide is repeated (twice in calponin-β and three times in calponin-α) corresponding to the C-terminal repeat structure of these molecules (Ruiz-Opazo & Nadal-Ginard, 1987). Interestingly, this $C4^1$ motif is found, with a high homology, seven times within the 379 residue product of *C. elegans* unc87, corresponding to the weakly repeating structure of this molecule, as illustrated by dot plot analysis (fig 21 f). The homologous peptides within unc87 contain the cAMP dependent protein kinase A site (except site V; see fig 22), and they also contain an adjacent Ser/Thr protein kinase C site (except site VI; see fig 22), and thus, a putative dual phosphorylation site is present that is similar to that seen within $C4^1$ at residues 180-185 (discussed in section F. 2) The significance of the RGASQAG heptapeptide in mature gelsolin, transgelin and $C4^1$ is currently unknown. However, this peptide is known to be phosphorylated within calponin-α (Nakamura et al, 1993), with site I (fig 22) being the preferred site of phosphorylation by protein kinase C in smooth muscle calponin (Nakamura et al, 1993). Hence, it seems likely that such sites within the molecules presented in figure 22 are also phosphorylated, although this has yet to be established biochemically.

G. NORTHERN BLOT ANALYSIS

Northern blot hybridisation of radiolabelled probes to immobilised preparations of poly A^+ selected mRNA allows the sensitive determination of the size and abundance of specific mRNA molecules (Sambrook et al, 1989).

1. Actin mRNA Content

The integrity and relative loading of RNA was always tested by probing hybridisation filters with an actin probe (methods, section T). This method allows the

approximate quantitative comparison of the C4l message in different samples and confirms equivalent loading of mRNA in lanes where no C4l message was seen. Following hybridisation β- and γ-actin isoforms are seen at 2.1Kb and α-actin isoforms are seen at 1.6Kb. The relative levels of the different isoforms in the various preparations correlated with the state of differentiation of the source. There is an alteration in actin isoform distribution in SV40 transformed 3T3 fibroblasts compared to that of normal 3T3 fibroblasts (fig 23 lower); there appears to be a down-regulation of some of the α-actin isoform in the transformed cell. This correlates with the published alterations in actin isoform distribution of transformed 3T3 cells as described by Leavitt et al (1985; 1987). The lack of any actin message in the HTCL poly A^+ selection 'flow-through' lane (fig 23 lower), indicates that the poly A^+ selection protocol was correctly carried out with ribosomal RNA removed from the recovered mRNA resulting in a more sensitive analysis. Alignment of overexpressed autoradiographs probed with C4l and actin respectively provides size markers that allows the estimate of the size of the C4l message.

2. C4l mRNA Content

Northern blot analysis indicates that a single message of 1.44Kb codes for C4l (fig 23 upper). The expression of this message in different tissue and cell types correlates exactly with the results of immunoblot experiments of the same samples providing further evidence that the 438bp PCR product (fig 7) used as a probe, encodes for the C4l molecule. The message has a restricted tissue specificity, being present in skeletal muscle in only negligible amounts at any stringency or level of loading, irrespective of autoradiographic times used (fig 23 upper). A weak C4l autoradiographic signal is seen in rat thymus as compared to the signal seen in HTCL. Actin mRNA levels also indicate that less HTCL mRNA was loaded than that of rat thymus mRNA. It would therefore appear that C4l message is up-regulated in HTCL samples. This observation may be due to the fact the C4l probe is HTCL in origin and thus in absolute homology with the HTCL C4l message. A weaker signal, due to lesser hybridisation at the stringencies used, may therefore exist between a C4l HTCL cDNA probe and C4l mRNA from rat thymus, probably caused by third base pair 'wobble' and conservative amino acid changes

(Alberts et al, 1994). However, this apparent up-regulation of an ABP in a transformed lymphocyte, is consistent with the observation that the F-actin bundling protein l-plastin, which was identified on the basis of its exclusive expression in transformed cells, was later detected in normal lymphocytes and macrophages at reduced levels (de Arruda et al, 1990).

Both SV40 transformed 3T3 cells and normal 3T3 cells express similar levels of C4[1] message, with no apparent up-regulation seen (fig 23 upper). Probing with actin indicates similar levels of mRNA loaded for each sample (fig 23 lower) even though the population of α-actin is lost in the transformed cells. It is unclear, therefore, why there appears to be less C4[1] message in rat thymus samples when compared to normal 3T3 fibroblasts. A possible explanation may be that 3T3 cells are a mouse fibroblast cell line (not non-immortalised fibroblasts such as REF) and therefore have a partially transformed phenotype. It is therefore possible that if REF (a primary fibroblast cell type with more SFs than 3T3 cells (Verderame et al, 1980)) mRNA was probed with C4[1], a similar expression level to rat thymus would be seen.

The Northern blot data presented here and those of Prinjha et al (1994) confirms, and extends, the findings of previous immunoblot analysis (Shapland et al, 1988; this study: fig 2). Neither C4[1] nor transgelin message is detectable in mature rat skeletal muscle; the molecules are also not detected, by immunoblot analysis, in erythrocytes (Shapland et al, 1988). Both cell types are highly differentiated and adapted to their specialised roles. In erythrocytes the cytoskeleton is distinct from that of other mammalian cells, erythrocytes being composed of a continuous submembrane fibrous network with few if any actin filaments traversing through the cell (Pumplin & Bloch, 1993). Bearing in mind the lack of microfilaments, the limited repertoire of proteins and the specialised function of these cells, the absence of C4[1] and transgelin is not surprising. The mouse 3T3 fibroblast cell line is widely used for cytoskeletal and oncogenic transformation studies (eg. Rodriguez-Fernandez et al, 1992a). 3T3 fibroblasts grow rapidly in culture and in common with primary fibroblasts have complex cytoskeletal networks, they also express both C4[1] and transgelin message (Prinjha et al, 1994; this study: fig 23 upper). 3T3 fibroblasts transformed with the DNA tumor virus SV40 were

found to express C4[l] in immunoblot experiments (fig 2) and northern analysis (fig 23 upper), with no transgelin expression seen in either case (Prinjha et al, 1994). Transgelin is therefore down-regulated in highly motile cells such as transformed cells and lymphocytes, with C4[l] expression apparently insensitive to the transformed and highly motile phenotypes.

C4[l] has been shown to be associated with F-actin cables (fig 1) and this, along with its high homology to transgelin and calponin (fig 19), would seem to indicate a cytoskeletal function. Whatever its function may be it is clear that a highly motile cell where stress fibres are reduced in number, whether transformed or not, requires the action of C4[l] on the F-actin network. It is known that in highly motile and weakly adhesive cells, cells that lack SFs, the functions of actin cross-linking and severing proteins are essential steps if protrusive force is to be generated. Only cross-linked actin filaments can be used to push against the membrane to form protrusion or for resisting deformation of cell shape by surface tension (Condeelis, 1993b). Proteins such as ABP120, spectrin and fimbrin, and the smaller protein ABP30, have all been identified as cross-linking and bundling proteins (Stossel, 1993; Otto, 1994) and may endow the actin cytoskeleton with properties such as rigidity (Stossel, 1993) or they may associate with polymerising filaments during network and bundle formation, thus contributing to the initial force of protrusion (Condeelis, 1993a; Cunningham et al, 1992). Severing of cross-linked filaments accompanies cellular extension (Hartwig, 1992), providing more sites for actin polymerisation by proteins such as ponticulin and profilin. Given its expression and the similarity to the actin gelling protein transgelin primary structure, it is possible that C4[l] is a cross-linking protein involved in network or bundle formation, thus contributing to the forces of protrusion in normal cells and highly motile cells such as SV40 transformed 3T3 cells where transgelin is not expressed.

3. Lack of Hybridisation Between C4[l] and Transgelin

By comparing the Northern analyses of Prinjha et al (1994) and this study (fig 23), one can see that transgelin cDNA clones fail to hybridise to C4[l] message. This observation, coupled with the fact that oligonucleotides derived from transgelin do not detect cDNA from C4[l] specific libraries (eg. rat lymphocyte) (not shown), indicated that

the two molecules may derive from two different genes. However, as the two molecules are recognised by the same monoclonal and polyclonal antibodies, it was therefore originally thought that the two molecules are highly homologous, possibly generated via alternative splicing or starting of the same transcript (Shapland et al, 1988). Subsequent purification of $C4^1$ (figs 5 and 6), microsequencing (see table 2), $C4^1$ gene specific RT-PCR, followed by cloning and sequencing (fig 11) revealed that $C4^1$ and transgelin do indeed derive from two separate genes, and therefore suggests that the C4 monoclonal antibody epitope will be present in a region of similarity. While the two molecules show a significant degree of homology, being 66.8% identical (see table 5), it is therefore highly unfortunate (but not surprising) that the original transgelin oligonucleotides did not recognise $C4^1$ cDNA since they correspond to regions of difference between the two molecules (not shown). It was surprising that larger transgelin clones (eg. clone H1 (Prinjha et al, 1994) which encodes for ≈600bp of the transgelin ORF) do not recognise either $C4^1$ cDNA or mRNA. A possible explanation could be that as the regions of difference between the two molecules are randomly spread, stable hybrids could not exist at the stringencies at which the hybridisation experiments were carried out (Sambrook et al, 1989). Similar examples are found in the literature and presented below. Human profilin II has high similarity (62-64% identity) to human profilin I (Honore et al, 1993). The levels of expression of profilin II transcripts are complementary to the level of expression of profilin I in the human tissues tested, such that in tissues with low expression of profilin I there is high expression of profilin II and *vice versa* (Honore et al, 1993). Although both isoforms are ≈62% identical in human, their full length cDNAs do not cross-hybridise to each others transcripts in Northern analysis (Honore et al, 1993). This is more than likely due to the fact that the regions of homology between the two human profilin isoforms are randomly spread throughout the molecule, resulting in unstable hybrids. The human plastins (t-, l- and I-) are ≈75% identical in amino acid sequence to each other (Lin et al, 1994). Upon Northern analysis it is seen that their respective cDNAs do not cross-hybridise to the other plastin isoform transcripts despite the high degrees of homology (Lin et al, 1994). This has proved useful, as with the study of $C4^1$ and transgelin expression, in that one can observe cell and tissue specific isoform expression. By use of a particular plastin isoform cDNA as a probe, one can examine its

specific tissue expression without fear of detecting the other plastin isoforms and observing 'false positives'. In this way Prinjha et al (1994) were able to illustrate the transcriptional regulation of transgelin, and I the apparent insensitivity of $C4^{l}$ expression, upon transformation.

APPENDIX I

Further Evidence That All C4^{l} PCR Products Are Derived From Transcripts of the Same mRNA Population.

Cloned C4^{l} PCR products, clones 1, 2 and 3 (see fig 10) have identical sequence in overlapping regions. Further evidence that these clones are derived from transcripts of the same mRNA population is provided by sequencing the full length coding region for HTCL C4^{l}.

By use of an absolute sense oligonucleotide complementary to the 5'-terminus (5'- GGAATTCTGGAATGGCCAACAGGGGACCTGCATA -3'), and an absolute antisense oligonucleotide complementary to the 3'-terminus (5'-GGAATTCGGGATCAG AGGATCTGGCGTGGCATC -3') of the C4^{l} coding region (see respective termini in fig 11), I have selectively amplified the full length coding region for C4^{l} from reverse transcribed poly A^{+} selected HTCL mRNA. EcoRI restriction sites at the 5'-termini of each oligonucleotide aid cloning into appropriate vectors; sequencing of such clones illustrates that the sequence of full length C4^{l} and those of clones 1, 2 and 3 are in absolute agreement.

APPENDIX II

Future Work

Isolation of full length C4[l] cDNA has enabled its inclusion within a multigene family of actin-binding proteins, and the availability of its primary structure will permit experimentation upon domains and regulatory sites which it contains. By comparing the sequence of C4[l] and its fellow multigene family member, transgelin, one can identify regions of difference which will serve as sites for specific monoclonal or polyclonal antibody production. Such antibodies will provide the tools for protein specific localisation in immunofluorescence studies.

The C4[l] purification protocol presented in this study produces ≈10μg of near homogeneous (>90%) protein from 5 x 10^9 HTCL cells. This result, although satisfactory for microsequence determination, may not be suitable for further biochemical analysis. A high level of homogeneous protein is often required for many analytical assays (see Quirk et al, 1993; Shapland et al, 1993). For this reason, a suitable approach would be to obtain homogeneous recombinant C4[l] from *E. coli* . By use of the two oligonucleotides mentioned in Appendix I, which are complementary to the termini of the C4[l] coding region, full length C4[l] cDNA can be selectively amplified from any source in RT-PCR reactions. Such PCR products can be ligated into a prokaryotic gene fusion vector such as pGEX. This plasmid is designed for inducible, high level intracellular expression of genes or gene fragments as fusions with *Schistasoma japonicum* glutathione S-transferase (GST). The fusion protein can thus be easily purified from bacterial lysates by affinity chromatography. This work is already underway.

Immunofluorescence microscopy studies presented in this work indicate that C4[l] is localised to actin filaments. Characterisation of the association of C4[l] with actin can be carried out by a series of tests such as viscometry, light scattering and sedimentation assays (Shapland et al, 1993). By determining the function of C4[l] as an ABP it may become clear why its expression is required in highly motile cells such as lymphocytes

and transformed mesenchymal cells, when its fellow multigene family member, transgelin, is not.

A small motif has been located within the sequence of C4[l] that shares a high degree of homology with the known ABD of actobindin (Vancompernolle et al, 1991). By use of a synthetic peptide representing this motif of C4[l] in competitive falling ball viscosity experiments, one would be able to determine whether it is a functional ABD. If C4[l] functions in vitro in the same manner as transgelin, ie. as an F-actin cross-linker, it must contain more than one ABD or bind actin filaments as a higher order structure such as a dimer, a possibility which can be elucidated by use of a sucrose density gradient.

Two regions within C4[l] share some homology with the consensus EF-Hand calcium-binding structure. By probing blotted purified C4[l] with $^{45}Ca^{2+}$ one could determine whether the molecule binds calcium. C4[l] and transgelin share a high degree of homology in these two regions and it has been shown that transgelin does not bind calcium (Shapland et al, 1993). It is therefore unlikely that these regions of C4[l] will bind calcium. By cloning and sequencing C4[l] homologues from ancient organisms, such as yeast, one can determine whether these regions were once functional or whether they preceded the evolution of EF-Hand structure.

A motif exists within C4[l] that is identical to a consensus nuclear localisation signal (NLS). This putative NLS can be delineated by two general methods. The NLSs are sequences (i) that can by genetic or biochemical fusion render cytoplasmic proteins nuclear, and (ii) which when deleted or mutated, no longer promotes nuclear uptake of the protein in which they reside. If this motif within C4[l] is confirmed as a functional NLS, then the mechanism of C4[l] nuclear translocation must be determined. A similar process to cofilin nuclear translocation may exist for C4[l], where dephosphorylation of adjacent sites cause a localised conformational change and the 'unmasking' of the NLS (Ohta et al, 1989).

REFERENCES

Abe, H; Endo, T; Yamamoto, K; Obinata, T. (1990) Sequence of cDNAs encoding actin depolymerising factor and cofilin of embryonic chicken skeletal muscle - two functionally distinct actin regulatory proteins exhibit structural homology. Biochemistry. 29:7420-7425.

Abe, M; Takahashi, K; Hiwada, K. (1990) Effects of calponin on actin-activated myosin ATPase activity. J. Biochem. (Tokyo). 108: 835-838.

Adams, A; Cooper, J; Drubin, D. (1993) Unexpected combinations of null mutations in genes encoding the actin cytoskeleton are lethal in yeast. Mol. Biol. Cell. 4: 459-468.

Adams, M; Minamide, L; Duester, G; Bamburg, J. (1990) Nucleotide sequence and expression of a cDNA encoding chicken brain actin depolymerising factor. Biochemistry. 29: 7414-7420.

Aderem, A. (1992a) The MARCKS brothers: a family of protein kinase C substrates. Cell. 71: 713-716.

Aderem, A. (1992b) Signal transduction and the actin cytoskeleton: the roles of MARCKS and profilin. TIBS. 17: 438-443.

Aguado-Velasco, C and Kuczmarski, E. R. (1993) Contraction of reconstructed Dictyostelium cytoskeletons: An apparent role for higher order associations among myosin filaments. Cell Motil. Cytoskel. 26: 103-114.

Aktories, K. (1990) ADP-ribosylation of actin. J. Musc. Res. Cell Motil. 11: 95-97.

Alberts, B; Bray, D; Lewis, J; Raff, M; Roberts, K; Watson, J. D. (1994) Molecular Biology Of The Cell. Third Edition. Garland Publishing, Inc. New York & London.

Alemendral, J. M; Santaren, J. F; Perara, J; Zerial, M; Bravo, R. (1989) Expression, cloning and cDNA sequence of a fibroblast serum-regulated gene encoding a putative actin-associated protein (p27). Exp. Cell Res. 181: 518-530.

Alexander, J. E; Hunt, D. F; Lee, M. K; Shabanowitz, J; Michel, H; Berlin, S. C; MacDonald, T. L; Sundberg, R. J; Rebhun, L. I; Frankfurter, A. (1991) Characterisation of post-translational modifications in neuron-specific class III β-tubulin by mass spectrometry. PNAS (USA). 88: 4685-4689.

Algrain, M; Arpin, M; Louvard, D. (1993) Wizardry at the cell cortex. Curr. Biol. 3: 451-454.

Amankwah, K. S and De Boni, U. (1994) Ultrastructural localisation of filamentous actin within neuronal interphase nuclei in situ. Exp. Cell Res. 210: 315-325.

Amatruda, J. F and Cooper, J. A (1992). Purification, characterisation and immunofluorescence localisation of *Saccharomyces cerevisiae* capping protein. J. Cell Biol. 117: 1067-1076.

Amatruda, J. F; Cannor, J. F; Tatchell, K; Hug, C; Cooper, J. A (1990). Disruption of the actin cytoskeleton in yeast capping protein mutants. Nature. 344: 352-354.

Ampe, C and Vandekerckhove, J. (1987) The F-actin capping proteins of physarum polycephalum: cap42(a) is very similar, if not identical, to fragmin and is structurally and functionally very homologous to gelsolin; cap42(b) is physarum actin. EMBO-J. 6: 4149-4157.

Ampe, C; Vandekerckhove, J; Brenner, S. L; Tobacman, L; Korn, E. D. (1985) The amino acid sequence of *Acanthamoeba* profilin. J. Biol. Chem. 260: 834-840.

Andre, E; Lottspeich, F; Schleicher, M; Noegal, A. (1988) Severin, gelsolin, and villin share a homologous sequence in regions presumed to contain F-actin severing domains. J. Biol. Chem. 263: 722-727.

Andreadis, A; Gallego, M. E; Nadal-Ginard, B. (1987) Generation of protein isoform diversity by alternative splicing: Mechanistic and biological implications. Ann. Rev. Cell Biol. 3: 207-242.

Andrews, P. C and Dixon, J. E. (1987) A procedure for in situ alkylation of cysteine residues prior to protein sequence analysis. Anal. Biochem. 161: 524-528.

Ankenbauer, T; Kleinschmidt, J. A; Walsh, M. J; Weiner, O. H; Fancke, W. W (1989). Identification of a widespread nuclear actin-binding protein. Nature. 342: 822-825.

Appel, L; Prout, M; Abu-Shumays, R; Hammonds, A; Garbe, J; Fristrom, D; Fristrom, J. (1993) The *Drosophila* stubble gene encodes an apparent transmembrane serine protease required for epithelial morphogenesis. PNAS (USA). 90: 4937-4941.

Arimura, C; Suzuki, T; Imamura, M; Hamada, Y; Masaki, T. (1988) Primary structure of chicken skeletal muscle and fibroblast α-actinins deduced from cDNA sequences. Eur. J. Biochem. 177: 649-655.

Arpin, M; Algrain, M; Louvard, D. (1994) Membrane-actin microfilament connections: An increasing diversity of players related to band 4.1. Curr. Opin. Cell Biol. 6: 136-141.

Arpin, M; Pringault, E; Finidori, J; Garcia, A; Jeltsch, J-M; Vandekerckhove, J; Louvard, D. (1988) Sequence of human villin: a large duplicated domain homologous with other actin-severing proteins and a unique small carboxy-terminal domain related to villin specificity. J. Cell Biol. 107: 1759-1766.

Auffray, C and Rougeon, F. (1980) Purification of mouse immunoglobulin heavy chain messenger RNA's from total myeloma tumor RNA. Eur. J. Biochem. 107: 303-314.

Ayme-Southgate, A; Lasko, P; French, C; Pardue, M. L. (1989) Characterisation of the gene for mp20: a *Drosophila* muscle protein that is not found in asynchronous oscillatory flight muscle. J. Cell Biol. 108: 521-531.

Ayme-Southgate, A; Vigoreaux, J; Benian, G; Pardeu, M. (1991) Drosophila has a twitchin/titin-related gene that appears to encode projectin. PNAS (USA). 88: 7973-7977.

Babcock, G and Rubenstein, P. A. (1993) Control of profilin and actin expression in muscle and nonmuscle cells. Cell Motil. Cytoskel. 24: 179-188.

Bahler, M; Benfenati, F, Valtorta, F; Greengard, P. (1990) The synapsins and the regulation of synaptic function. BioEssays. 12: 259-263.

Bahler, M; Benfenati, F; Valtorta, F; Czernik, A, J; Greengard, P. (1989) Characterisation of synapsin I fragments produced by cysteine-specific cleavage: A study of their interaction with F-actin. J. Cell Biol. 108: 1841-1849.

Balasubramanian, M; Helfman, D, Hemmingsen, S. (1992) A new tropomyosin essential for cytokinesis in the fission yeast *S. pombe*. Nature. 360: 84-87.

Barany, K; Polyak, E; Barany, M. (1992) Involvement of calponin and caldesmon in sustained contraction of arterial smooth muscle. Biochem. Biophys. Res. Comm. 187: 847-852.

Bardwell, J. C. A and Beckwith, J. (1993) The bonds that tie: Catalysed disulfide bond formation. Cell. 74: 769-771.

Barylko, B; Wagner, M. C; Reizes, O; Albanesi, J. P. (1992) Purification and characterisation of a mammalian myosin I. PNAS (USA). 89: 490-494.

Bazari, W; Matsudaira, P; Wallek, M; Smeal, T; Jakes, R; Ahmed, Y. (1988) The villin sequence and peptide map identifies 6 homologous domains. PNAS (USA). 85: 4986-4990.

Bearer, E. (1991) Direct observation of actin filament severing by gelsolin and binding by gCap39 and capZ. J. Cell Biol. 115: 1629-1638.

Beckerle, M and Yeh, R. (1990) Talin: role at sites of cell-substratum adhesion. Cell Motil. Cytoskel. 16: 7-13.

Belkin, A. M and Koteliansky, V. E. (1987) Interaction of iodinated vinculin, metavinculin and α–actinin with cytoskeletal proteins. FEBS Lett. 220: 291-294.

Bellas, R. E.; Bendori, R.; Farmer, S. R. (1991) Epidermal growth factor activation of vinculin and β-integrin gene transcription in quiescent Swiss 3T3 cells. J. Biol. Chem. 266: 12008-12014.

Belyavsky, A; Vinogradova, T; Rajewsky, K. (1989) PCR-based cDNA library construction: general cDNA libraries at the level of a few cells. N. A. R. 17: 2919-2932.

Bendori, R.; Salomon, D.; Geiger, B. (1987) Contact-dependent regulation of vinculin expression in cultured fibroblasts: A study with vinculin specific cDNA probes. EMBO-J. 6: 2897-2905.

Benian, G; Kiff, J; Neckelmann, N; Moerman, D; Waterston, R. (1989) Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*. Nature. 342: 45-50.

Bennett, V. (1989) The spectrin-actin junction of erythrocyte membrane skeletons. Biochim. Biophys. Acta. 988: 107-121.

Bikle, D; Munson, S; Morrison, N; Eisman, J. (1993) Zipper protein a newly described tropomyosin-like protein of the intestinal brush-border. J. Biol. Chem. 268: 620-626.

Bjorkegren, C; Rozyeki, M; Schutt, L. E; Lindberg, U; Karlsson, R. (1993) Mutagenesis of human profilin locates its poly (L-proline)- binding site to a hydrophobic patch of aromatic amino acids. FEBS Lett. 333: 123-126.

Blackshear, P. J. (1993) The MARCKS family of cellular protein kinase C substrates. J. Biol. Chem. 288: 1501-1504.

Blanchard, A; O'Hanion, V; Critchley, D. (1989) The structure and function of α-actinin J. Musc. Res. Cell Motil. 10: 280-289.

Bloom, G. S. (1992) Motor proteins for cytoplasmic microtubules. Curr. Opin. Cell Biol. 4: 66-73.

Boschek, C. B; Jockush, B. M; Friis, R. R; Back, R; Grundmann, E; Bauer, H. (1981) Early changes in the distribution and organisation of microfilament proteins during cell transformation. Cell. 24: 175-184.

Bray, D and Thomas, C. (1975) The actin content of fibroblasts. Biochem. J. 147: 221-228.

Bray, D and White, J. G. (1988) Cortical flow in animal cells. Science. 239: 883-888.

Bray, D. (1992) Cell Movements. First Edition. Garland Publishing Inc. New York & London.

Bray, D; Heath, J; Moss, D. (1986) The membrane-associated 'cortex' of animal cells: its structure and mechanical properties. J. Cell Sci. Suppl. 4: 71-88.

Bremer, J. W; Busch, H; Yeoman, C. L. (1981) Evidence for a species of nuclear actin distinct from cytoplasmic and muscle actins. Biochemistry. 20: 2013-2017.

Bresnick, A; Warren, V; Condeelis, J. (1990) Identification of a short sequence essential for actin binding by Dictyostelium ABP-120. J. Biol. Chem. 265: 9236-9240.

Bretscher, A. (1991) Microfilament structure and function in the cortical cytoskeleton. Ann. Rev. Cell Biol. 7: 337-374.

Brinkley, B. R. (1985) Microtubule organising centres. Ann. Rev. Cell Biol. 1: 145-172.

Brown, K; Zinkowski, P; Hays, S; Binder, L. (1993) Actin binding protein is a component of bovine erythrocytes. Cell Motil. Cytoskel. 24: 100-108.

Brown, S and Lucy, J. (1993) Dystrophin as a mechanical transducer in skeletal muscle. BioEssays. 15: 413-419.

Bryan, J and Kane, R. E. (1982) Actin gelation in sea urchin egg extracts. Meth. Cell Biol. 25: 175-199.

Bryan, J; Edwards, R; Matsudaira, P; Otto, J. J; Wulfkuhle, J. (1993) Fascin, an echinoid actin-bundling protein, is a homolog of the *Drosophila* singed gene product. PNAS (USA). 90: 9115-9119.

Bryan, J; Imai, M; Lee, R; Moore, P; Cook, R. G; Lin, W. G. (1989) Cloning and expression of a smooth muscle caldesmon. J. Biol. Chem. 264: 13873-13879.

Buβ, F; Temm-Grove, C; Henning, S; Jockush, B. M. (1992) Distribution of profilin in fibroblasts correlates with the presence of highly dynamic actin filaments. Cell Motil. Cytoskel. 22: 51-61.

Bubb, M. R; Lewis, M. S; Korn, E. D. (1991) The interaction of monomeric actin with two binding sites on Acanthamoeba actobindin. J. Biol. Chem. 266: 3820-3826.

Burns, R. G and Surridge, J. M. (1993) Analysis of β-tubulin sequences reveals highly conserved coordinated amino acid substitutions - evidence that these hot spots are directly involved in the conformational change required for dynamic instability. FEBS Lett. 271: 1-8.

Burridge, K. (1986) Substrate adhesion in normal and transformed fibroblasts: organisation and regulation of cytoskeletal, membrane and extracellular matrix components at focal adhesions. Cancer Rev. 4: 18-78.

Burridge, K; Fath, K; Kelly, T; Nukolls, G; Turner, C. (1988) Focal adhesions: transmembrane junctions between the extracellular matrix and the cytoskeleton. Ann. Rev. Cell Biol. 4: 487-525.

Byers, H. R and Fujiwara, K. (1986) Stress fibres in cells in situ: immunofluorescence visualisation with anti-actin, anti-myosin and anti-α-actinin. J. Cell Biol. 93: 804-811.

Byers, H. R; White, G. E; Fujiwara, K. (1984) Organisation and function of stress fibres in cells in vitro and in situ: in Cell and Muscle Motility, Vol 5 (Ed, Shay, J. W). Plenum, New York.

Caldwell, J. E; Waddle, J. A; Cooper, J. A; Hollands, J. A; Casella, S. J; Casella, J. F (1989). cDNAs encoding the beta-subunit of capZ, the actin-capping protein of the Z-line of muscle. J. Biol. Chem. 264: 12648-12652.

Cant, K; Knowles, B. A; Mooseker, M. S; Cooley, L. (1994) *Drosophila* singed, a fascin homolog, is required for actin bundle formation during oogenesis and bristle extension. J. Cell Biol. 125: 369-380.

Cao, L-G and Wang, Y-L. (1990) Mechanism of the formation of contractile ring in dividing cultured animal cells. I. Recruitment of preexisting actin filaments into the cleavage furrow. J. Cell Biol. 110: 1089-1095.

Cao, L-G; Babcock, G; Rubenstein, P. A; Wang, Y. I. (1992) Effects of profilin and profilactin on actin structure and function in living cells. J. Cell Biol. 117: 1023-1029.

Caplow, M. (1992) Microtubule dynamics. Curr. Opin. Cell Biol. 4: 58-65.

Carlier, M-F. (1987) Measurement of Pi dissociation from actin filaments following ATP hydrolysis using a linked enzyme assay. Biochem. Biophys. Res. Comm. 143: 1069-1075.

Carlier, M-F. (1991) Actin: protein structure and filament dynamics. J. Biol. Chem. 266: 1-4.

Carlier, M-F; Jean, C; Rieger, K. J; Lenfant, M; Pantaloni, D. (1993) Modulation of the interaction between G-actin and thymosin β4 by the ATP/ADP ratio: Possible implication in the regulation of actin dynamics. PNAS (USA). 90: 3034-5038.

Carlier. M-F. (1993) Dynamic actin. Curr. Biol. 3: 321-323.

Carlson, G. M; Bechtel, P. J; Graves, D. J. (1979) Chemical and regulatory properties of phosphorylase kinase cAMP dependent protein kinase. In Advances in Enzymology and Related Areas of molecular Biology. Meister, A. Ed. John Wiley and Sons Inc. New York. 41-115.

Carlsson, L; Nystrom, L. E; Sundkvist, I; Markey, F; Lindberg, U. (1977) Actin polymerisability is influenced by profilin, a low molecular weight protein in non-muscle cells. J. Mol. Biol. 115: 465-483.

Casella, J. F; Casella, S. J; Hollands, J. A; Caldwell, J. E; Cooper, J. A. (1989) Isolation and characterisation of cDNA encoding the alpha subunit of capZ (36/32), an actin-capping protein from the Z-line of skeletal muscle. PNAS (USA). 86; 5800-5804.

Casella, J. F; Craig, S; Maack, D; Brown, A. (1987) CapZ (36/32), a barbed end actin-capping protein, is a component of the Z-line of skeletal muscle. J. Cell Biol. 105: 371-379.

Cassimeris, L; Safer, D; Nachmias, V. T; Zigmond, S. H. (1992) Thymosin β4 sequesters the majority of G-actin in resting human polymorphonuclear leukocytes. J. Cell Biol. 119: 1261-1270.

Chang-Yeh, A; Mold, D. E; Huang, R. C. C. (1991) Identification of a novel murine IAP-promoted placenta-expressed gene. N. A. R. 19: 3667-3672.

Chen, M and Stracher, A. (1989) In situ phosphorylation of platelet actin binding protein by cAMP dependent protein kinase stabilizes it against proteolysis by calpain. J. Biol. Chem. 264: 14282-14289.

Chen, T. I; Edwards, K. A; Lin, R, C; Coats, L. W; Kiehart, D. P. (1991) *Drosophila* myosin heavy chain at 35B, C. J. Cell Biol. 115 (suppl): 330a.

Cheney, R. E; Riley, M. A; Mooseker, M. S. (1993) Phylogenetic Analysis of the myosin superfamily. Cell Motil. Cytoskel. 24: 215-223.

Chia, C. P; Hitt, A. L; Luna, E. J. (1991) Direct binding of F-actin to ponticulin, an integral membrane glycoprotein. Cell Motil. Cytoskel. 18: 164-179.

Chia, C.P; Shariff, A; Savage, S. A; Luna, E. J. (1993) The integral membrane protein, ponticulin, acts as a monomer in nucleating actin assembly. J. Cell Biol. 120:909-922.

Chiu, F. C; Barnes, E. A; Das, K; Haley, J; Socolow, P; Macaluso, F. P; Fant, J. (1989) Characterisation of a novel 66kd subunit of mammalian neurofilaments. Neuron. 2: 1435-1445.

Chuang, J. Z; Lin, D. C; Carter, W. L; Cunningham, M. E; Butler, J. A; Lin, S. (1991) Molecular cloning of chicken cardiac tensin and expression in cos cells. J. Cell Biol. 115: 166a.

Clark, S. W and Meyer, D. I. (1992) Centractin is an actin homologue associated with the centrosome. Nature. 359: 246-250.

Cleveland, D. W and Sullivan, K. F. (1984) Molecular biology and genetics of tubulin. Ann. Rev. Biochem. 54: 331-366.

Coleman, T: Fishkind, D; Mooseker, M; Morrow, J. (1989) Functional diversity among spectrin isoforms. Cell Motil. Cytoskel. 12: 225-247.

Conboy, J; Kan, Y; Shohet, S; Mohandas, N. (1986) Molecular cloning of protein 4.1, a major structural element of the human erythrocyte membrane cytoskeleton. PNAS (USA). 83: 9512-9516.

Condeelis, J. (1993a) Life at the leading edge: The formation of cell protrusions. Ann. Rev. Cell Biol. 9: 411-444.

Condeelis, J. (1993b) Understanding the cortex of crawling cells: Insights from Dictyostelium. Trends Cell Biol. 3: 371-376.

Cook, R. K; Blake, W. T; Rubenstein, P. A. (1992) Removal of the amino-terminal acidic residues of yeast actin: studies in vitro and in vivo. J. Biol. Chem. 267: 9430-9436.

Cooke, R. (1986) The mechanism of muscle contraction. CRC Crit. Rev. Biochem. 21: 53-118.

Cooley, L; Verheyen, E; Ayers, K. (1992) *Chickadee* encodes a profilin required for intracellular cytoplasm transport during Drosophila oogenesis. Cell. 69: 173-184.

Cooper, J. A. (1991) The role of actin polymerisation in cell motility. Ann. Rev. Physiol. 53: 585-605.

Cooper, J. A; Caldwell, J. E; Gattermeir, D. J; Torres, M. A; Amatruda, J. F; Casella, J. F (1991). Variant cDNAs encoding proteins similar to the α–subunit of chicken capZ. Cell Motil. Cytoskel. 18: 204-214.

Correas, I; Padilla, R; Avila, J. (1990) The tubulin-binding sequence of brain microtubule associated proteins, tau and MAP-2, is also involved in actin-binding. Biochem. J. 269: 61-64.

Coue, M and Korn, M. E. (1986) Interaction of plasma gelsolin with ADP-actin. J. Biol. Chem. 261: 3628-3631.

Cramer, L and Mitchison, T. J. (1993) Moving and stationary actin filaments are involved in spreading of postmitotic Ptk 2 cells. J. Cell Biol. 122: 833-843.

Crawford, A and Beckerle, C. (1991) Purification and characterisation of zyxin, an 82,000 dalton component of adherens junctions. J. Biol. Chem. 266: 5847-5853.

Crawford, A; Mitchelson, J; Beckerle, M. (1992) An interaction between zyxin and α-actinin. J. Cell Biol. 116: 1381-1393.

Crawford, A; Pino, J; Beckerle, M (1994) Biochemical and molecular characterisation of the chicken cysteine-rich protein, a developmentally regulated LIM-domain protein that is associated with the actin cytoskeleton. J. Cell Biol. 124: 117-127.

Crompton, M. R; Moss, S. E; Crumpton, M. J. (1988) Diversity in the lipocortin/calpactin family. Cell. 55: 1-3.

Crumpton, M. J and Dedman, J. R. (1990) Protein terminology tangle. Nature. 345: 212.

Cunningham, C. C; Stossel, T.P; Kwiatkowski, D. (1991) Enhanced Motility in NIH 3T3 fibroblasts that overexpress gelsolin. Science. 251: 1233-1236.

Cunningham, C; Gorlin, J; Kwiatkowski, D; Hartwig, J; Janmey, P; Byers, H; Stossel, T. (1992) Actin binding protein requirement for cortical stability and efficient locomotion. Science. 255: 325-327.

Dabiri, G; Young, C; Rosenbloom, J; Southwick, F. (1992) Molecular cloning of human macrophage capping protein cDNA. J. Biol. Chem. 267: 16545-16552.

Davis, L. H; Davis, J. K; Bennett, V. (1992) Ankyrin regulation: An alternatively spliced segement of the regulatory domain functions as a intramolecular modulator. J. Biol. Chem. 267: 18966-18972.

Davis, P. F and Tripathi, S. C. (1993) Mechanical stress mechanisms and the cell: an endothelial paradigm. Circ. Res. 72: 239-245.

Davis, S; Lu, M. L; Lo, S. H; Lin, S; Butler, J. A; Druker, B. J; Roberts, T. M; An, Q; Chen, L. B. (1991) Presence of an SH2 domain in the actin binding protein tensin. Science. 252: 712-715.

De Camilli, P; Benfenati, F; Valtorta, F; Greengard, P. (1990) The synapsins. Ann. Rev. Cell Biol. 6: 433-460.

de Arruda, M; Bazari, H; Wallek, M; Matsudaria, P. (1992) An actin footprint on villin. J. Biol. Chem. 267: 13079-13085.

de Arruda, M; Watson, S; Lin, C; Leavitt, J; Matsudaira, P. (1990) Fimbrin is a homologue of the cytoplasmic phosphoprotein plastin and has domains homologous with calmodulin and actin gelation proteins. J. Cell Biol. 111: 1069-1079.

de Hostos, E; Bradtke, B; Lottspeich, F; Gerisch, G. (1993) Coactosin, a 17kDa F-actin binding protein from *Dictyostelium discoidium.* Cell Motil. Cytoskel. 26: 181-191.

de Hostos, E; Bradtke, B; Lottspeich, F; Guggenheim, R; Gerisch, G. (1991) Coronin, an actin binding protein of Dictyostelium discoidium localised to cell surface projections, has similarities to G protein β subunits. EMBO-J. 10: 4097-4104.

de Hostos, E; Rehfueβ, C; Bradtke, B; Waddell, D; Albrecht, R; Murphy, J; Gerisch, G. (1993) Dictyostelium mutants lacking the cytoskeletal protein coronin are defective in cytokinesis and cell motility. J. Cell Biol. 120: 163-173.

Demma, M; Warren, V; Hock, R; Dharmawardhame, S; Condeelis, J. (1990) Isolation of an abundant 50,000 dalton actin filament bundling protein from Dictyostelium amoebae. J. Biol. Chem. 265: 2286-2291.

den Hartigh, J; Paul, M; Bergen, H; Verkleij, A; Boonstra, J. (1992) The EGF receptor is an actin-binding protein. J. Cell Biol. 119: 349-355.

Dharmawardhame, S; Demma, M; Yang, F; Condeelis, J. (1991) Compartmentalisation and actin binding properties of ABP-50: the elongation factor-1α of Dictyostelium. Cell Motil. Cytoskel. 20: 279-288.

Diaz-Nunoz, M; Hamilton, S. L; Kaetzel, M. A; Hazarika, P; Dedman, J. R. (1990) Modulation of Ca^{2+} release channel activity from sarcoplasmic reticulum by annexin VI (67-kDa calcimedin). J. Biol. Chem. 265: 15894-15899.

Djabali, K; Porter, M-M; Gros, F; Blobel, G; Georgatos, S. D. (1991) Network antibodies identify nuclear lamin B as a physiological attachment site for peripherin intermediate filaments. Cell. 64: 109-121.

Doolittle, R. F. (1986) Of URFs and ORFs - A Primer on How to Analyse Derived Amino Acid Sequences. University Science Books, Oxford University Press.

Draeger, A; Gimona, M; Stukert, A; Celis, J. E; Small, V. J. (1991) Calponin developmental isoforms and a low molecular weight variant. FEBS Lett. 291: 24-28.

Drenckhahn, D and Pollard, T. D. (1986) Elongation of actin filaments is a diffusion-limited reaction at the barbed end is accelerated by inert macromolecules. J. Biol. Chem. 261: 12754-12758.

Drubin, D; Mulmolard, J; Zau, Z; Botstein, D. (1990) Homology of a yeast actin-binding protein to signal transduction proteins and myosin-I. Nature. 343: 288-290.

Dustin, P. (1984) Microtubules. 2nd edition. Berlin. Springer-Verlag.

Earnshaw, W. E and Cooke, C. A. (1991) Analysis of the distribution of the INCENPs throughout mitosis reveals the existence of a pathway of structural changes in the chromosomes during metaphase and early events in cleavage furrow formation. J. Cell Sci. 98: 443-461.

Eddy, R, J; Sauterer, R. A; Condeelis, J. (1993) Aginactin, an agonist-regulated F-actin capping activity is associated with an Hsc70 in Dictyostelium. J. Biol. Chem. 268: 23267-23274.

Edwards, B. T. (1993) ABP-50: An actin-binding elongation factor 1α from *Dictyostelium discoidium*. J. Cell Biochem. 52: 134-139.

Edwards, K. A; Montague, R. A; Shephard, S; Edgar, B. A; Erikson, R. L; Kiehart, D. P. (1994) Identification of Drosophila cytoskeletal proteins by induction of abnormal cell shape in fission yeast. PNAS (USA). 91: 4589-4593.

Egelhoff, T. T; Brown, S. S; Spudich, J. A. (1991) Spatial and temporal control of nonmuscle myosin location: Identification of a domain that is necessary for myosin filament disassembly in vivo. J. Cell Biol. 112: 677-688.

Eichinger, L and Schleicher, M. (1992) Characterisation of actin and lipid-binding domains in severin, a Ca^{2+}-dependent F-actin fragmenting protein. Biochemistry. 31: 4779-4787.

Endow, S. A and Titus, M. A. (1992) Genetic approaches to molecular motors. Ann. Rev. Cell Biol. 8: 29-66.

Erba, H. P; Gunning, P; Kedes, L. (1986) Nucleotide sequences of the human γ-cytoskeletal actin mRNA: anomalous evolution of vertebrate non-muscle actin genes. NAR. 14: 5275-5294.

Eriksson, J. E; Opal, P; Goldman, R. D. (1992) Intermediate filament dynamics. Curr. Opin. Cell Biol. 4: 99-104.

Erlich, H. A. Ed. (1989) PCR Technology. First Edition. Stockton Press. New York, London, Tokyo, Melbourne, Hong Kong.

Ervasti, J and Campbell, K. (1993) Dystrophin and the membrane skeleton. Curr. Opin Cell Biol. 5: 82-87.

Espreafico, E. M; Cheney, R. E; Matteoli, M; Nascimento, A. A. C; De Camilli, P. V; Larson, R. E; Mooseker, M. S. (1992) Primary structure and cellular localisation of chicken brain myosin-V (p190), an unconventional myosin with calmodulin light chains. J. Cell Biol. 119: 1541-1557.

Estes, J. E; Seldon, L. A; Kinosian, H. J; Gershman, L. C. (1992) Tightly-bound divalent cation of actin. J. Musc. Res. Cell Motil. 13: 272-284.

Fechheimer, M and Furukawa, R. (1993) A 27,000-D core of the Dictyostelium 34,000-D protein retains Ca^{2+}-regulated actin cross-linking but lacks bundling activity. J. Cell Biol. 120: 1169-1176.

Fechheimer, M; Murdock, D; Carney, M; Glover, C. (1991) Isolation an sequencing of cDNA clones encoding the *Dictyostelium discoidium* 30,000-dalton actin-bundling protein. J. Biol. Chem. 266: 2883-2889.

Feinberg, A. P and Vogelstein, B. (1983) A technique for radiolabelling DNA restriction fragments to high specific activity. Anal. Biochem. 132: 6-13.

Felice, G. R; Eason, P; Nermut, M. V; Kellie, S. (1990) pp60 (v-src) association with the cytoskeleton induces actin organisation without affecting polymerising status. Eur. J. Cell Biol. 52: 47-59.

Fernandez, A; Brautigan, D. L; Mumby, M; Lamb, N. J. C. (1990) Protein phosphatase type 1, not type 2A., modulates actin filament integrity and myosin light chain phosphorylation in living nonmuscle cells. J. Cell Biol. 111: 103-112.

Ferrari, S; Battini, R; Kasczmarek, L; Rittling, S; Calabretta, B; DeReil, J. K; Philiponis, V; Weifang, J; Baserga, R. (1986) Coding sequence and growth regulation of the human vimentin gene. Mol. Cell Biol. 6: 3614-3620.

Finidori, J; Friederich, E; Kwiatkowski, D; Louvard, D. (1992) In vivo analysis of functional domains from villin and gelsolin. J. Cell Biol. 116: 1145-1155.

Finkel, T; Theriot, J. A; Dire, K. R; Thomaselli, G. F; Goldschmidt-Clermont, P. J. (1994) Dynamic actin structures stabilised by profilin. PNAS (USA). 94: 1510-1514.

Fishkind, D. J and Wang, Y-L. (1993) Orientation and three dimensional organisation of actin filaments in dividing cultured cells. J. Cell Biol. 123: 837-848.

Fliegner, K. H; Ching, G. Y; Liem, R. K. H. (1990) The predicted amino acid sequence of α-internexin is that of a novel neuronal intermediate filament protein. EMBO-J. 9: 749-755.

Flynn. D. C; Len, T-H; Reynolds, A. B; Parsons, T. J. (1993) Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated $pp60^{src}$ substrate. Mol. Cell Biol. 13: 7892-7900.

Foisner, R and Wiche, G. (1991) Intermediate filament-associated proteins. Curr. Opin. Cell Biol. 3: 75-81.

Fowler, V; Sussmann, M; Miller, P; Flucher, B; Daniels, M. (1993) Tropomodulin is associated with the free (pointed) ends of the thin filaments in rat skeletal muscle. J. Cell Biol. 120: 411-420.

Friederich, E; Vancompernolle, K; Heut, C; Goethals, M; Finidori, J; Vandekerckhove, J; Louvard, D. (1992) An actin binding site containing a conserved motif of charged amino acid residues is essential for the morphogenic effect of villin. Cell. 70: 81-92.

Frohman, M. A; Dush, M. K; Martin, G. R. (1988) Rapid production of full-length cDNAs from rare gene-specific oligonucleotide primer. PNAS (USA). 85: 8998-9002.

Frosch, P; Frosch, M; Pfister, T; Schaad, V; Bitter-Suermann, D. (1991) Cloning and characterisation of an immunodominant major surface antigen of *Echinococcus multilocularis*. Mol. Biochem. Parasitol. 48: 121-130.

Fuchs, E. (1994) Intermediate filaments and disease: mutations that cripple cell strength. J. Cell Biol. 125: 511-516.

Fulton, A. B and Isaacs, W. D. (1991) Titin, a huge, elastic sarcomeric protein with a probable role in morphogenesis. BioEssays. 13: 157-161.

Fulton, A. B. (1993) Spatial organisation of the synthesis of cytoskeletal proteins. J. Cell Biochem. 52: 148-152.

Funayama, N; Nagafuchi, A; Sato, N; Tsukita, S; Tsukita, S. (1991) Radixin is a novel member of the band 4.1 family. J. Cell Biol. 115: 1039-1048.

Furst, D. O; Osborn, M; Nave, R; Weber, K. (1988) The organisation of titin filaments in the half-sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: A map of ten non-repetitive epitopes starting at the Z lines extends close to the M line. J. Cell Biol. 106: 1563-1572.

Furuhashi, K and Hatano, S. (1989) A fragmin-like protein from plasmodium of *Physarum polycephalum* that severs F-actin and caps the barbed end of F-actin in a Ca^{2+} sensitive way. J. Biochem. 106: 311-318.

Fyrberg, E; Kelly, M; Ball, E; Fyrberg, C; Reedy, M. (1990) Molecular genetics of Drosophila alpha actinin: mutant alleles disrupt Z disc integrity and muscle insertions. J. Cell Biol. 110: 1999-2011.

Gaertner, A and Wegner, A. (1991) Mechanism of the insertion of actin monomers between the barbed ends of actin filaments and barbed end-bound insertin. J. Musc. Res. Cell Motil. 12: 27-36.

Garcia-Bustos, J; Heitman, J; Hall, M. N. (1991) Nuclear protein localisation. Biochim. Biophys. Acta. 1071: 83-101.

Gardner, K and Bennett, V. (1987) Modulation of spectrin-actin assembly by erythrocyte adducin. Nature. 328: 359-362.

Garner, I; Minty, A. J; Alonso, S; Barton, P. J; Buckingham, M. E. (1986) A 5'-duplication of the α-cardiac actin gene in BALB/c mice is associated with abnormal levels of α-cardiac and α-skeletal actin mRNAs in adult cardiac tissue. EMBO-J. 5: 2559-2567.

Garnier, J; Osguthorpe, D. J; Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120: 97-120.

Geiger, B and Ginsberg, D. (1991). The cytoplasmic domain of adherens-type junctions. Cell Motil. Cytoskel. 20: 1-6.

Geiser, J. R; Van Tuinen, D; Brockerhoff, S. E; Neff, M. M; Davis, T. N. (1991) Can calmodulin function without binding calcium. Cell. 65: 949-959.

Gibbs, A. J and McIntyre, G. A. (1970) The diagram, a method for comparing sequences: Its use with amino acid and nucleotide sequences. Eur. J. Biochem. 16: 1-11.

Gieselmann, R and Mann, K. (1992) ASP-56, a new actin sequestering protein from pig platelets with homology to CAP, an adenylate cyclase-associated protein from yeast. FEBS Lett. 298: 149-153.

Gilmore, A. P; Wood, C; Ohanian, V; Jackson, P; Patel, B; Rees, D. J. G; Hynes, R. O; Critchley, D. R. (1993) The cytoskeletal protein talin contains at least two distinct vinculin binding domains. J. Cell Biol. 122: 337-347.

Gimona, M.; Small, J. V.; Moeremans, M.; VanDamme, J.; Puype, M.; Vandekerckhove, J. (1988) Porcine vinculin and meta-vinculin differ by a 68-residue insert located close to the carboxy-terminal part of the molecule. EMBO J. 7: 2329-2334

Gimona, M; Sparrow, M. P; Strasser, P; Herzog, M; Small, V. J. (1992) Calponin and SM22 isoforms in avian and mammalian smooth muscle. Eur. J. Biochem. 205: 1067-1075.

Glenney, J. R and Zokas, L. (1989) Novel tyrosine kinase substrates from Rous sarcoma virus-transformed cells are present in the membrane skeleton. J. Cell Biol. 108: 2401-2408.

Glenney, J. R; Tack, E; Powell, M. A. (1987) Calpactins: Two distinct Ca^{2+}-regulated phospholipid- and actin-binding proteins isolated from lung and placenta. J. Cell Biol. 104: 503-511.

Gluck, U; Kwiatkowski, D; Ben-Ze'ev, A. (1993) Suppression of tumorigenicity in simian virus 40 transformed 3T3 cells transfected with α-actinin cDNA. PNAS (USA). 90: 383-387.

Goedert, M and Jakes, R. (1990) Expression of separate isoforms of tau protein: correlation with the tau pattern in brain and effects on tubulin polymerisation. EMBO J. 9: 4225-4230.

Goldschmidt-Clermont, P. J and Janmey, P. A. (1991) Profilin, a weak CAP for actin and RAS. Cell. 66: 419-421.

Goldschmidt-Clermont, P. J; Furman, M. I; Wachsstock, D; Safer, D; Nachmias, V. T; Pollard , T. D. (1992) The control of actin nucleotide exchange by thymosin β4 and profilin. A potential regulatory mechanism for actin polymerisation in cells. Mol. Cell Biol. 3: 1015-1024.

Goldschmidt-Clermont, P. J; Kim, J. W; Machesky, L. M; Rhee, S. G; Pollard, T. D. (1991a) Regulation of phospholipase C-γI by profilin and tyrosine phosphorylation. Science. 251: 1231-1233.

Goldschmidt-Clermont, P. J; Machesky, L. M; Baldassare, J. J; Pollard, T. D. (1990) The actin-binding protein profilin binds to PIP_2 and inhibits its hydrolysis by phospholipase C. Science. 247: 1575-1578.

Goldschmidt-Clermont, P. J; Machesky, L. M; Doberstein, S. K; Pollard, T. D. (1991b) Mechanism of the interaction of platelet profilin with actin. J. Cell Biol. 113: 1081-1089.

Goldschmidt-Clermont, P. J; Van Alstyne, E. L; Day, J. R; Emerson, D. L; Nel, A. E; Lazarchick, J; Galbraith, R. M. (1986) Group-specific component (vitamin D binding protein) prevents the interaction between G-actin and profilin. Biochemistry. 25: 6467-6472.

Goldstein, L. S. B. (1991) The kinesin superfamily: tails of functional redundancy. Trends Cell Biol. 1: 93-98.

Goodwin, L; Lees-Miller, J; Cheley, S; Leonard, M; Helfman, D. (1990) Four rat fibroblast tropomyosin isoforms are expressed from a single gene via alternative RNA splicing and the use of two promoters. J. Biol. Chem. 262: 8408-8415.

Gorlin, J; Yamin, R; Egan, S; Stewart, M; Stossel, T; Kwiatkowski, D; Hartwig, J. (1990) Human endothelial actin binding protein (ABP-280 non-muscle filamin): a molecular leaf spring. J. Cell Biol. 111: 1089-1105.

Gould, K; Bretscher, A; Esch, F; Hunter, T. (1989) cDNA cloning and sequencing of the protein-tyrosine kinase substrate, ezrin, reveals homology to band 4.1. EMBO-J. 8: 4133-4142.

Graceffa, P and Jancso, A. (1991) Disulfide cross-linking of caldesmon to actin. J. Biol. Chem. 266: 20305-20310.

Graff, J. M; Gordon, J. I; Blackshear, P. J. (1989a) Myristoylated and nonmyristoylated forms of a protein are phosphorylated by protein kinase C. Science. 246: 503-506.

Graff, J. M; Young, T. N; Johnson, J. D; Blackshear, P. J. (1989b) Phosphorylation-regulated calmodulin binding to a prominent cellular substrate for protein kinase C. J. Biol. Chem. 264: 21818-21823.

Gribskov, M; McLachlan, A. D; Eisenberg, D. (1987) Profile analysis: detection of distantly related proteins. PNAS (USA). 84: 4355-4358.

Grieg, R. G; Koestler, T. P; Trainer, D. L; Corwin, S. P; Miles, L; Kline, T; Sweet, R; Yokoyama, S; Poste, G. (1985) Tumorigenic and metastastic properties of 'normal' and ras-transfected NIH/3T3 cells. PNAS (USA). 82: 3698-3701.

Gu, M; York, J; Warshawsky, I; Majerus, P. (1991) Identification, cloning and expression of a cytosolic megakaryocyte protein-tyrosine-phosphatase with sequence homology to cytoskeletal protein 4.1. PNAS (USA). 88: 5867-5871.

Guerriero, V; Rowley, D. R; Means, A. R. (1981) Production and characterisation of an antibody to myosin light chain kinase and intracellular localisation of the enzyme. Cell. 27: 449-458.

Haarer, B. K; Lillie, S. H; Adams, A. E. M; Magdolen, B; Bandlow, W; Brown, S. S. (1990) Purification of profilin from *Saccharomyces cerevisiae* and analysis of profilin-deficient cells. J. Cell Biol. 110: 105-114.

Haarer, B. K; Lillie, S. H; Petzold, A. S; Brown, S. S. (1992) The yeast myo4 gene encodes a type V myosin similar to the myo2 gene product. Mol. Biol. Cell. 3 (suppl): 42a.

Haarer, B. K; Petzold, A. S; Brown, S. S. (1993) Mutational analysis of yeast profilin. Mol. Cell Biol. 13: 7864-7873.

Habazetti, J; Gondol, D; Wiltscheck, R; Otlewski, J; Schleicher, M; Holak, T. A. (1992) Structure of hisactophilin is similar to interleukin-Iβ and fibroblast growth factor. Nature. 359: 855-858.

Hamagushi, M and Hanafusa, H. (1987) Association of $pp60^{src}$ with Triton X-100-resistant cellular structure correlates with morphological transformation.. PNAS (USA) 84: 2313-2316.

Hammer, J. A. (1991) Novel Myosins. Trends Cell Biol. 1: 50-56.

Harricane, M. C; Fabbrizio, E; Arpin, C; Mornet, D. (1992) Involvement of caldesmon at the actin-myosin interface. Biochem. J. 287: 633-637.

Harris, C. C. (1993) P^{53} : At the crossroads of molecular carcinogenesis and risk assessment. Science. 262: 1980-1981.

Harris, E. L. V and Angal, S. Eds. (1989) Protein Purification Methods: A Practical Approach. First Edition. IRL Press at Oxford University Press.

Hartmann, H; Noegal, A. A; Eckerskorn, C; Rapp, S; Schleicher, M (1989). Ca^{2+}-independent F-actin capping proteins. J. Biol. Chem. 264: 12639-12647.

Hartwig, J and Kwiatkowski, D. (1991) Actin binding proteins. Curr. Opin. Cell Biol. 3: 87-97.

Hartwig, J. (1992) Mechanisms of actin rearrangements mediating platelet activation. J. Cell Biol. 118: 1421-1442.

Hartwig, J. H; Chambers, K. A; Hopica, K. L; Kwiatkowski, D, J. (1989) Association of profilin with filament-free regions of human leukocyte and platelet membrane binding during platelet activation. J. Cell Biol. 109: 1571-1579.

Hartwig, J. H; Thelen, M; Rosen, A; Janmey, P. A; Nairn, A; Aderem, A. (1992) MARCKS is an actin filament cross-linking protein regulated by protein kinase C and calcium-calmodulin. Nature. 356: 618-622.

Hasson, T. B and Mooseker, M. S. (1992) Two novel unconventional myosins identified from a kidney proximal tubule cell line. Mol. Biol. Cell. 3 (suppl): 158a.

Hawkins, M; Pope, B; Maciver, S; Weeds, A. (1993) Human actin depolymerising factor mediates a pH sensitive destruction of actin filaments. Biochemistry. 32: 9985-9993.

Hayashi, K; Kanda, K; Kimizuka, F; Kato, I; Sobue, K. (1989) Primary structure and functional expression of h-caldesmon complementary DNA. Biochem. Biophys. Res. Comm. 164: 503-511.

Heacock, C. S; Eidsvoog, K. E; Bamburg, J. R. (1984) The influence of contact inhibited growth and of agents which alter cell morphology on the level of G- and F-actin in cultured cells. Exp. Cell Res. 153: 402-412.

Hedrick, L; Cho; K. R; Vogelstein, B. (1993) Cell adhesion molecues as tumor suppressors. Trends Cell Biol. 3: 36-39.

Heiss, S. G and Cooper, J. A (1991). Regulation of capZ, an actin capping protein of chicken muscle, by anionic phospholipids. Biochemistry. 30: 8753-8758.

Heizmann, C. W and Hunziker, W. (1991) Intracellular calcium binding proteins: more sights than insights. TIBS. 16: 98-103.

Hemmings, L; Kuhlman, S; Critchley, D. (1992) Analysis of the actin binding domain of α-actinin by mutagenesis and demonstration that dystrophin contains a functionally homologous domain. J. Cell Biol. 116: 1369-1380.

Hendricks, M and Weintraub, H. (1981) Tropomyosin is decreased in transformed cells. PNAS (USA). 78: 5633-5637.

Hennesey, E. S; Drummond, D. R; Sparrow, J. C. (1993) Molecular genetics of actin function. Biochem. J. 262: 657-671.

Hens, J. J. H; Benfenati, F; Nielander, H. B; Valtorta, F; Gispen, W. H; De Graan, P. N. E. (1993) B50/GAP-43 binds to actin filaments without affecting actin polymerisation and filament organisation. J. Neurochem. 61: 1530-1533.

Herman, I. M. (1993) Actin isoforms. Curr. Opin. Cell Biol. 5: 48-55.

Hill, M. A and Gunning, P. (1993) Beta and gamma actin mRNAs are differentially located within myoblasts. J. Cell Biol. 122: 825-832.

Hirano, S; Kimoto, N; Shimoyama, Y; Hirohasi, S; Takeichi, M. (1992) Identification of a neural α-catenin as a key regulator of cadherin function and multicellular organisation. Cell. 70: 293-301.

Hitt, A and Luna, E. (1994) Membrane interactions with the actin cytoskeleton. Curr. Opin. Cell Biol. 6: 120-130.

Hoffmann, A; Noegal, A; Bomblies, L; Lottspeich, F; Schleicher, M. (1993) The 100kDa F-actin capping protein of Dictyostelium amoebae is a villin prototype ('protovillin'). FEBS Lett. 328: 71-76.

Holmes, K. C; Popp, D; Gebhard, W; Kabsch, W. (1990) Atomic model of the actin filament. Nature. 347: 44-49.

Honore, B; Madsen, P; Andersen, A. H; Leffers, H. (1993) Cloning and expression of a novel human profilin variant, profilin-II. FEBS Lett. 330: 151-155.

Hoock, T. C; Newcomb, P. M; Herman, I. M. (1991) β-actin and its mRNA are localised at the plasma membrane and the regions of moving cytoplasm during the cellular response to injury. J. Cell Biol. 112: 653-664.

Horowitz, J. A and Hammer, J. A. (1990) A new *Acanthameoba* myosin heavy chain. J. Biol. Chem. 265: 20646-20652.

Hosoya, H; Kobayashi, R; Tsukita, S; Matsumura, F. (1992) Ca^{2+}-regulated actin and phospholipid binding protein (68-kD-protein) from bovine liver: Identification as a homologue for annexin VI and intracellular localisation. Cell Motil. Cytoskel. 22: 200-210.

Hug, C: Miller, T; Torres, M, A; Casella, J. F; Cooper, J. A. (1992) Identification and characterisation of an actin-binding site of capZ. J. Cell Biol. 166: 923-931.

Humphery, M. B; Herrera-Sosa, H; Gonzalez, G; Lee, R; Bryan, J. (1992) Cloning of cDNAs encoding human caldesmons. Gene. 112: 197-204.

Hynes, R. O and Lander, A. D. (1992) Contact and adhesive specificities in the associations, migrations, and targeting of cells and axons. Cell. 68: 303-322.

Hynes, R. O. (1992) Integrins: versatility, modulation, and signalling and cell adhesion. Cell. 69: 1-25.

Iida, K; Matsumoto, S; Yahara, I. (1992) The KKRKK sequence is involved in heat shock-induced nuclear translocation of the 18kDa actin-binding protein, cofilin. Cell Struct. Funct. 17: 39-46.

Imamura, M and Masaki, T. (1992) A novel nonmuscle α-actinin. J. Biol. Chem. 267: 25927-25933.

Isenberg, G. (1991) Actin binding proteins-lipid interactions. J. Musc. Res. Cell Motil. 12: 136-144.

Ishikawa, R; Yamashiro, S; Matsumura, F. (1989) Differential modulation of actin severing activity of gelsolin by multiple isoforms of cultured rat cell tropomyosin. J. Biol. Chem. 264: 7490-7497.

Itano, N and Hatano, S. (1991) F-actin bundling protein from *Physarum polycephalum:* Purification and its capacity for co-bundling of actin filaments and microtubules. Cell Motil. Cytoskel. 19: 244-254.

Janmey, P. A. (1991) Mechanical properties of cytoskeletal polymers. Curr. Opin. Cell Biol. 2: 4-11.

Janmey, P. A; Lamb, J; Allen, P; Matsudaira, P. (1992) Phosphoinositide-binding peptides derived from the sequences of gelsolin and villin. J. Biol. Chem. 267: 11818-11823.

Janmey, P. A; Stossel, T. P; Lind, S. E. (1986) Sequential binding of actin monomers to plasma gelsolin and its inhibition by vitamin D-binding protein. Biochem. Biophys. Res. Comm. 136: 72-79.

Jin, J-P and Wang, K. (1991) Nebulin as a giant actin-binding template protein in skeletal muscle sarcomere: interaction of actin cloned human nebulin fragments. FEBS Lett. 281: 93-96.

Johnston, G. C; Prendergast, J. A; Singer, R. A. (1991) The *Saccharomyces cerevisiae* MYO2 gene encodes an essential myosin for vectorial transport of vesicles. J. Cell Biol. 113: 539-551.

Jongstra, J; Tidmarch, F; Jongstra-Bilen, J; Davis, M. (1988) A new lymphocyte specific gene which encodes a putative Ca^{2+} binding protein is not expressed in transformed T-lymphocyte lines. J. Immunol. 141: 3999-4004.

Jongstra-Bilen, J; Janmey, P; Hartwig, J; Galea, S; Jongstra, J. (1992) The lymphocyte specific protein LSP1 binds to F-actin and to the cytoskeleton through its COOH-terminal basic domain. J. Cell Biol. 118: 1443-1453.

Joseph, C. K; Qureshi, S. A; Wallace, D. J; Foster, D. A. (1992) MARCKS protein is transcriptionally down-regulated in V-Src-transformed BALB/c 3T3 cells. J. Biol. Chem. 267: 1327-1330.

Joshi, R; Gilligan, D. M; Otto, E; McLaughin, T; Bennett, V. (1991) Primary structure and domain organisation of human alpha and beta adducin. J. Cell Biol. 115: 665-675.

Jove, J. A and Hanafusa, H. (1987) Cell transformation by the viral src gene. Ann. Rev. Cell Biol. 3: 31-56.

Juliano, R. L and Varner, J. A. (1993) Adhesion molecules in cancer: The role of integrins. Curr. Biol. 5: 812-818.

Kabsch, W and Sander, C. (1983) How good are predictions of protein secondary structure ? FEBS Lett. 155: 179-182.

Kabsch, W; Mannherz, H. G; Suck, D; Pai, E. F; Holmes, K. C. (1990) Atomic structure of the actin DNAaseI complex. Nature. 347: 37-44.

Kaiser, D. A; Sato, M; Ebert, R. F; Pollard, T. D. (1986) Purification and characterisation of two isoforms of *Acanthamoeba* profilin. J. Cell Biol. 102: 221-226.

Kalderon, D; Roberts, B. L; Richardson, W. D; Smith, A. E. (1984) A short amino acid sequence able to specify nuclear localisation. Cell. 39: 499-509.

Kanoh, S; Ito, M; Niwa, E; Kawano, Y; Harshorne, D. J. (1993) Actin-binding peptide from smooth muscle myosin light chain kinase. Biochemistry. 32: 8902-8907.

Karess, R. E; Chang, X-L; Edward, K. A; Kularni, S; Aguilera, I; Kiehart, D. P. (1991) The regulatory light chain of nonmuscle myosin is encoded by *spaghetti-squash* a gene required for cytokinesis in Drosophila. Cell. 65: 1177-1189.

Kellerman, K. A and Miller, K. G. (1992) An unconventional myosin heavy chain gene from *Drosophila melanogaster* . Eur. J. Cell Biol. 119: 823-834.

Kellie, S; Horvath, A. R; Elmore, M. A. (1991) Cytoskeletal targets for oncogenic tyrosine kinases. J. Cell Sci. 99: 207-211.

Kipreos, E. T and Wang, J. Y. J. (1992) Cell cycle-regulated binding of c-Abl tyrosine kinase to DNA. Science. 256: 382-385.

Klymkowsky, M. W; Bachant, J. B; Domingo, A. (1989) Functions of intermediate filaments. Cell Motil. Cytoskel. 14: 309-331.

Knight, A. E and Kendrick-Jones, J. (1993) A myosin-like protein from a higher plant. J. Mol. Biol. 231: 148-154.

Koch, C. A; Anderson, D; Moran, M. F; Elis, C; Pawson, T. (1991) SH2 and SH3 domains: elements that control interactions of cytoplasmic signal proteins. Science. 252: 668-674.

Koenig, M; Monaco, A; Kunkel, L. (1988) The complete sequence of dystrophin predicts a rod-shaped cytoskeletal protein. Cell. 53: 219-228.

Koji-Owada, M; Hakura, A; Iida, K; Yahara, I; Sobue, K; Kakiuchi, S. (1984) Occurrence of caldesmon (a calmodulin-binding protein) in cultured cells: comparison of normal and transformed cells. PNAS (USA). 81: 3133-3137.

Kojima, N; Kato, Y; Shirao, T; Obata, K. (1988) Nucleotide sequences of two embryonic debrins, developmentally regulated brain proteins, and developmental change in their mRNAs. Mol. Brain Res. 4: 207-215.

Korn, E. D; Carlier, M. F; Pantaloni, D. (1987) Actin polymerisation and ATP hydrolysis. Science. 238: 638-644.

Kozak, M. (1986) Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell. 44: 283-292.

Kozak, M. (1987) An analysis of 5'- noncoding sequences from 699 vertebrate messenger RNAs. N. A. R. 15: 8125-8148.

Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. 266: 19867-19870.

Kreig, J and Hunter, T. (1992) Identification of the two major epidermal growth factor-induced tyrosine phosphorylation sites in the microvillar core protein ezrin. J. Biol. Chem. 267: 19258-19265.

Kreis, T and Vale, R. Eds. (1993) Guidebook to the Cytoskeletal and Motor Proteins. 1st Edition. Oxford University Press.

Kruger, M; Wright, J; Wang, K. (1991) Nebulin as a length regulator of thin filaments of vertebrate skeletal muscles: correlation of thin filament length, nebulin size, and epitope profile. J. Cell Biol. 115: 97-107.

Kumar, A; Raziuddin, T; Finlay, H; Thomas, J. O; Szer, W. (1984) Isolation of a minor species of actin from the nuclei of *Acanthamoeba castellanii*. Biochemistry. 23: 6753-6757.

Kumar, C. C; Mohan, S; Chang, C; Garrels, J. I. (1989) In cytoskeletal proteins in tumor diagnosis. Osborn, M and Weber, K (Eds): 'Current Communications in Molecular Biology' Cold Spring Harbor, NY: Cold Spring Laboratory, pp 91-97.

Kuznetsov, S; Langford, G; Weiss, D. (1992) Actin-dependent organelle movement in squid axoplasm. Nature. 356: 722-725.

Kwiatkowski, D. (1988) Predominant induction in gelsolin and actin-binding protein during myeloid differentiation. J. Biol. Chem. 263: 13857-13862.

Kwiatkowski, D. J and Bruns, G. A. P. (1988) Human profilin: molecular cloning, sequence comparison, and chromosomal analysis. J. Biol. Chem. 263: 5910-5915.

Kwiatkowski, D; Stossel, T; Orkin, S; Mole, J; Cotten, H; Yin, H. (1986) Plasma and cytoplasmic gelsolins are encoded by a single gene and contain a duplicated actin-binding domain. Nature. 323: 455-458.

Kyte, J and Doolittle, R. F. (1982) A simple method for displaying the hyropathic character of a protein. J. Mol. Biol. 157: 105-132.

Labeit, S; Barlow, D; Gautel, M; Gibson, T; Holt, J; Hsieh, C; Francke, U; Leonard, K; Wardale, J; Whiting, A; Trinick, J. (1990) A regular pattern of two types of 100-residue motif in the sequence of titin. Nature. 345: 273-276.

Labeit, S; Gibson, T; Lakey, A; Leonard, K; Zeviani, M; Knight, P; Wardle, J; Trinick, J. (1991) Evidence that nebulin is a protein-ruler in muscle thin filaments. FEBS Lett. 282: 313-316.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bactriophage T4. Nature. 227: 680-685.

Lamb, N. J; Fernandez, A; Watrin, A; Labbe, J. C; Cavadore, J. C. (1990) Microinjection of $p34^{cdc2}$ kinase induces marked changes in cell shape, cytoskeletal organisation and chromatin structure in mammalian fibroblasts. Cell. 60: 151-165.

Lambert, S and Bennett, V. (1993) From anemia to cerebellar dysfunction. A review of the ankyrin gene family. Eur. J. Biochem. 211: 1-6.

Langanger, G; Moermans, M; Daneels, G; Sobieszek, A; De Brabander, M; De May, J. (1986) The molecular organisation of myosin in stress fibres of cultured cells. J. Cell Biol. 102: 200-209.

Lankes, W and Furthmayr, H. (1991) Moesin: A member of the protein 4.1-talin-ezrin family of proteins. PNAS (USA). 88: 8297-8301.

Lassing, I and Lindberg, U. (1985) Specific interaction between phosphatidylinositol 4, 5-bisphosphate and profilactin. Nature. 314: 472-474.

Lavoie, J. N; Hickey, E; Weber, L. A; Landry, J. (1993) Modulation of actin microfilament dynamics and fluid phase pinocytosis by phosphorylation of heat shock protein 27. J. Biol. Chem. 268: 24210-24214.

Lawson, D. (1983) Epinemin: a new protein associated with vimentin filaments in non-neuronal cells. J. Cell Biol. 97: 1891-1905.

Lawson, D. (1986) Myosin distribution and actin organisation in different areas of antibody labelled quick frozen fibroblasts. J. Cell Sci. Suppl. 5: 45-54.

Leavitt, J; Cunning, P; Kedes, L; Jariwalla, R. (1985) Smooth muscle α-actin is a transformed-sensitive marker for mouse NIH/3T3 and rat-2 cells. Nature. 316: 840-842.

Leavitt, J; Madhu-Varma, S-Y. N; Latter, G; Burbeck, S; Gunning, P; Kades, L. (1987) Expression of transfected mutant β-actin genes: Transitions towards the stable tumorigenic state. Mol. Cell Biol. 7: 2467-2476.

Lee, J; Ishihara, A; Jacobson, K. (1993) How do cells move along surfaces ? Trends Cell Biol. 3: 366-370.

Lee, M. G; Gowan, N. J; Kirschner, M. W. (1988) The primary structure and heterogeneity of tau protein from mouse brain. Science. 239: 285-288.

Lees-Miller, J and Helfman, D. (1991) The molecular basis for tropomyosin isoform diversity. BioEssays. 13: 429-437.

Lees-Miller, J. P; Goodwin, L; Helfman, D. (1990a) Three novel brain tropomyosin isoforms are expressed from the rat α-tropomyosin gene through the use of alternative promoters and alternative RNA processing. Mol. Cell Biol. 10: 1729-1742.

Lees-Miller, J. P; Healy, D. H; Smillie, L. B; Kay, C. M. (1987) Isolation and characterisation of an abundant and novel 22kDa protein (SM22) from chicken gizzard smooth muscle. J. Biol. Chem. 262: 2988-2993.

Lees-Miller, J. P; Helfman, D. M; Shroer, T. A. (1992a) A vertebrate actin-related protein is a component of a multisubunit complex involved in microtubule-based vesicle motility. Nature. 359: 244-246.

Lees-Miller, J. P; Henry, G; Helfman, D. M. (1992b) Identification of *act2*, an essential gene in the fission yeast *Schizosaccharomyces pombe* that encodes a protein related to actin. PNAS (USA). 89: 80-83.

Lees-Miller, J; Yan, A; Helfman, D. (1990b) Structure and complete nucleotide sequence of the gene encoding rat fibroblast tropomyosin 4. J. Mol. Biol. 213: 399-405.

Lehninger, A. L. (1970) Biochemistry: The Molecular Basis of Cell Structure. First Edition. Worth Publishers, Inc, New York.

Lewis, S; Wang, D; Cowan, N. (1988). Microtubule-associated protein MAP-2 shares a microtubule binding motif with tau protein. Science. 242: 936-939.

Libri, D; Lemonnier, M; Meinnel, T; Fiszman, M. (1989) A single codes for the β-subunits of smooth and skeletal muscle tropomyosin in the chicken. J. Biol. Chem. 264: 2935-2944.

Liem, R. K. H. (1990) Neuronal intermediate filaments. Curr. Opin. Cell Biol. 2: 86-90.

Lillie, S. H and Brown, S. S. (1992) Suppression of a myosin defect by a kinesin-related gene. Nature. 356: 358-361.

Lin, C-S; Shen, W; Chen, Z. P; Tu, Y-H; Matsudaira, P. (1994) Identification of I-plastin, a human fimbrin isoform expressed in intestine and kidney. Mol. Cell. Biol. 14: 2457-2467.

Linse, K and Mandelkow, E. M. (1988) The GTP-binding peptide of β-tubulin. Localisation by direct photoaffinity labelling and comparison with nucleotide-binding proteins. J. Biol. Chem. 263: 15205-15210.

Littauer, U. Z; Giveon, D; Thierauf, M; Ginzburg, I; Ponstingl, H. (1986) Common and distinct tubulin binding sites for microtubule- associated proteins. PNAS (USA). 83: 7162-7166.

Liu, H and Bretscher, A. (1989) Disruption of the single tropomyosin gene in yeast results in the disappearance of actin cables from the cytoskeleton. Cell. 57: 233-242.

Liu, H and Bretscher, A. (1992) Characterisation of TPM1 disrupted yeast cells indicates an involvement of tropomyosin in directed vesicular transport. J. Cell Biol. 118: 285-299.

Liu, Y; Carraway, K; Carraway, C. (1989) Isolation and characterisation of a 58kDa membrane microfilament associated protein from ascites tumor cell microvilli. J. Biol. Chem. 264: 1208-1214.

Llinas, R; McGuinness, T; Leonard, C. S; Sugimori, M; Greengard, P. (1985) Intraterminal injection of synapsin I or calcium/calmodulin-dependent protein kinase II alters neurotransmitter release at the squid giant synapse. PNAS (USA). 82: 3035-3039.

Lombardi, L; Ballinari, D; Bongarzone, I; Migliari, M; Mondellini, P; Traversari, C; Modina, S. (1990) Ultrastructural cytoskeleton alterations and modification of actin expression in the NIH/3T3 cell line after transformation with Ha-ras-activated oncogene. Cell Motil. Cytoskel. 15: 220-229.

Louvard, D. (1989) The function of the major cytoskeletal components of the brush border. Curr. Opin. Cell Biol. 1: 51-57.

Luduena, R. F; Banerjee, A; Kahn, I. A. (1992) Tubulin Structure and biochemistry. Curr. Opin. Cell Biol. 4: 53-57.

Luna, E and Hitt, A. (1992) Cytoskeleton-plasma membrane interactions. Science. 258: 955-964.

Mabuchi, I; Tsukita, S; Tsukita, S; Sawai, T. (1988) Cleavage furrow isolation from newt eggs: contraction, organisation of the actin filaments and protein components of the furrow. PNAS (USA). 85: 5966-5970.

Machesky, L. M and Pollard, T. D. (1993) Profilin as a potential mediator of membrane-cytoskeleton communication. Trends Cell Biol. 3: 381-385.

Maciver, S; Zot, H; Pollard, T. D. (1991) Characterisation of actin filament severing by actophorin from *Acanthamoeba castellanii*. J. Cell Biol. 115: 1611-1620.

MacRae, T. H. (1992) Microtubule organisation by cross-linking and bundling proteins. Biochem. Biophys. Acta. 1160: 145-155.

Maekawa, S. (1992) Purification of a membranous actin binding protein from bovine adrenal medulla. Biochem. Int. 28: 87-95.

Maekawa, S; Torijama, N; Hisanaga, S-I; Yonezawa, N; Endo, S; Hirokawa, N; Sakai, H. (1989) Purification and characterisation of a calcium dependent actin filament severing protein from bovine adrenal medulla. J. Biol. Chem. 264: 7458-7465.

Magdolen, B; Oechsner, U; Muller, G; Bandlow, W. (1988) The intron-containing gene for yeast profilin (PFY) encodes a vital function. Mol. Cell Biol. 8: 5108-5115.

Maller, J. L. (1991) Mitotic control. Curr. Opin. Cell Biol. 3: 269-275.

Mani, R. S; McCubbin, W. D; Kay, C. M. (1992) Calcium-dependent regulation of caldesmon by an 11Kda smooth muscle calcium binding protein, caltropin. Biochemistry: 31: 11896-11901.

Marchisio, P. C; Cirillo, D; Teti, A; Zambonin-Zallone, A; Tarone, G. (1987) Rous sarcoma virus-transformed fibroblasts and cells of monocyte in origin display a peculiar

dot-like organisation of cytoskeletal proteins involved in microfilament-membrane interactions. Exp. Cell Res. 169: 202-214.

Marston, B and Redwood, C. S. (1991) The molecular anatomy of caldesmon. Biochem. J. 279: 1-16.

Marston, S. (1988) Aorta caldesmon inhibits actin activation of thiophosphorylated Heavy meromyosin Mg^{2+} ATPase activity by slowing the product release. FEBS Lett. 238: 147-150.

Marston, S. (1989) What is latch? New ideas about tonic contraction in smooth muscle. J. Musc. Res. Cell Motil. 10: 97-100.

Martin, F. H; Castro, M. M; Aboul-Ela, F; Tinoco, I. J. (1985) Base pairing involving deoxyinosine: implications for probe design. N. A. R. 13: 8927-8938.

Maruyama, K; Endo, T; Kume, H; Kauamura, Y; Kanzawa, N; Nakauchi, Y; Kimura, S; Kawashima, S; Maruyama, K. (1993) A novel domain sequence of connectin localised at the I band of skeletal muscle sarcomeres: Homology to neurofilament subunits. Biochem. Biophys. Res. Comm. 194: 1288-1291.

Maruyama, K; Kurokawa, H; Oosawa, M; Shimaoka, S; Yamamoto, H; Ito. M; Maruyama, K (1990). Beta-actinin is equivalent to capZ protein. J. Biol. Chem. 265: 8712-8715.

Matsudaira, P. (1991) Modular organisation of actin crosslinking proteins. TIBS. 16: 87-92.

Matsumura, F and Yamashiro, S. (1993) Caldesmon. Curr. Opin. Cell Biol. 5: 70-76.

Matsuzaki, F; Matsumoto, S; Yahara, I; Yonezawa, N; Nishida, E; Sakai, H. (1988). Cloning and characterisation of porcine brain cofilin. J. Biol. Chem. 263: 11564-11568.

Maxfield, F. R. (1993) Regulation of leukocyte locomotion by Ca^{2+}. Trends Cell Biol. 3: 386-391.

McCrea, P; Turck, C. W; Gumbiner, B. (1991) A homolog of the *armadillo* protein in *Drosophila* (plakoglobin) associated with E-cadherin. Science. 254: 1359-1361.

McLeod, J. F; Kowalski, M. A; Haddad, J. G. (1989) Interactions among serum vitamin D binding protein, monomeric actin, profilin and profilactin. J. Biol. Chem. 264: 1260-1267.

McWhirter, J. R and Wang, J. Y. J. (1991) Activation of tyrosine kinase and microfilament-binding functions of c-Abl by bcr sequences in bcr/abl fusion protons. Mol. Cell Biol. 11: 1785-1792.

McWhirter, J. R and Wang, J. Y. J. (1993) An actin-binding function contributes to transformation by the Bcr-Abl oncoprotein of the Philadelphia chromosome-positive leukemias. EMBO-J. 12: 1533-1546.

Mercer, J. A; Seperack, P. K; Strobel, M. C; Copeland, N. G; Jenkins, N. A.(1991) Novel myosin heavy chain encoded by murine dilute coat colour locus. Nature. 349: 709-713.

Mermall, V; McNally, J. G; Miller, K. G. (1994) Transport of cytoplasmic particles catalysed by an unconventional myosin in living *Drosophila* embryos. Nature. 369: 560-562.

Messier, J; Shaw, L; Chafe, M; Matsudaira, P; Mercurio, A. (1993) Fimbrin localised to an insoluble cytoskeletal fraction is constitutively phosphorylated on its head-piece domain in adherent macrophages. Cell Motil. Cytoskel. 25: 223-233.

Metzler, W. J; Bell, A. J; Ernot, E; Lavoie, T. B; Mueller, L. (1994) Identification of the poly-L-proline-binding site on human profilin. J. Biol. Chem. 269: 4620-4625.

Mezgueldi, M; Fattoum, A; Derancourt, J; Kassab, R. (1992) Mapping of the functional domains in the Amino-terminal region of calponin. J. Biol. Chem. 267: 15943-15951.

Milankov, K and De Boni, U. (1993) Cytochemical localisation of actin and myosin aggregates in interphase nuclei in situ. Exp. Cell Res. 209: 189-199.

Milligan, R. A; Whittaker, M; Safer, D. (1990) Molecular structure of F-actin and location of surface binding sites. Nature. 348: 217-221.

Miron, T; Vancompernolle, K; Vandekerckhove, J; Wilchek, M; Geiger, B. (1991) A 25KD inhibitor of actin polymerisation is a low molecular mass heat shock protein. J. Cell Biol. 114: 255-261.

Miron, T; Wilchek, M; Geiger, B. (1988) Characterisation of an inhibitor of actin polymerisation in vinculin-rich fraction of turkey gizzard smooth muscle. Eur. J. Biochem. 178: 543-553.

Mittal, B; Sanger, J. M; Sanger, J. W. (1987) Visualisation of myosin in the cytoplasm, cleavage furrow and mitotic spindle of human cells. J. Cell Biol. 105: 1753-1760.

Montell, C and Rubin, G. (1988): The *Drosophila* ninaC locus encodes two photoreceptor cell specific proteins with domains homologous to protein kinases and the myosin heavy chain head. Cell. 52: 757-762.

Moon, A; Janmey, P; Louie, K; Drubin, D. (1993) Cofilin is an essential component of the yeast cortical cytoskeleton. J. Cell Biol. 120: 421-435.

Moon, R and McMahon, A. (1990) Generation of diversity in non-erythroid spectrins. J. Biol. Chem. 265: 4427-4433.

Mooseker, M. S. (1985) Organisation, chemistry, and assembly of the cytoskeletal apparatus of the intestinal brush border. Ann. Rev. Cell Biol. 1: 209-241.

Mooseker, M. S. (1993) A multitude of myosins. Curr. Biol. 3: 245-248.

Moraga, D. M; Nunez, P; Garrido, J; Maccioni, R. B. (1993) A τ fragment containing a repetitive sequence induces bundling of actin filaments. J. Neurochem. 61: 979-986.

Morgan, T; Lockerbie, R; Minamide, L; Browning, M; Bamburg, J. (1993) Isolation and characterisation of a regulated form of actin depolymerising factor. J. Cell Biol. 122: 623-633.

Moriyama, K; Matsumoto, S; Nishida, E; Sakai, H; Yahara, I. (1990a) Nucleotide sequence of mouse cofilin cDNA. N. A. R. 18: 3053.

Moriyama, K; Nishida, E; Yonezawa, N; Sakai, H; Matsumoto, S; Iida, K; Yahara, I. (1990b) Destrin, a mammalian actin-depolymerising protein is closely related to cofilin. J. Biol. Chem. 265: 5768-5773.

Moss, D. J; Fernyhough, P; Chapman, K; Baizer, L; Bray, D; Allsopp, T. (1990) Chicken Growth-associated protein Gap-43 is tightly bound to the actin-rich neuronal membrane skeleton. J. Neurochem. 54: 729-736.

Moss, P. A. H; Moots, R. J; Rosenberg, W. M. C; Rowland-Jones, S. J; Bodmer, H. C; McMichael, H. J; Bell, J. I. (1991) Extensive conservation of α and β chains of the human T-cell antigen receptor recognising HLA-A2 and influenza A matrix peptide. PNAS (USA). 88: 8987-8990.

Muhlrad, A. (1991) The isolated 21kDa N-terminal fragment of myosin binds to actin in an ATP and ionic strength-dependent manner. Biochim. Biophys. Acta. 1077: 308-315.

Nachmias, V. T. (1993) Small actin-binding proteins: the β-thymosin family. Curr. Opin. Cell Biol. 5: 56-62.

Nagafuchi, A; Takeichi, M; Tsukita, S. (1991) The 102Kd cadherin-associated protein: similarity to vinculin and posttranscriptional regulation of expression. Cell. 65: 849-857.

Nakamura, F; Mino, T; Yamamoto, T; Naka, M; Tanaka, T. (1993) Identification of the regulatory site in smooth muscle calponin that is phosphorylated by protein kinase C. J. Biol. Chem. 268: 6194-6201.

Nakayama, S; Moncrief, N. D; Kretsinger, R. H. (1992) Evolution of EF-Hand calcium-modulated proteins. II. Domains of several subfamilies have diverse evolutionary histories. J. Mol. Evol. 34: 416-448.

Nigg, E. A. (1992) Assembly-disassembly of the nuclear lamina. Curr. Opin. Cell Biol. 4: 105-109.

Nishida, W; Kitami, Y; Abe, M; Hiwada, K. (1991) Gene cloning and nucleotide sequence of SM22α from the chicken gizzard smooth muscle. Biochem. Int. 23: 663-668.

Noda, S; Ito, M; Watanabe, S; Takahashi, K; Maruyama, K. (1992) Conformational changes of actin induced by calponin. Biochem. Biophys. Res. Comm. 185: 481-487.

Noegal, A; Gerisch, G; Lottspeich, F; Schleicher, M. (1990) A protein with homology to the C-terminal repeat sequence of octopus rhodopsin and synaptophysin is a member of a multigene family in *Dictyostelium discoidium*. FEBS Lett. 266: 118-122.

Noegal, A; Rapp, S; Lottspeich, F; Schleicher, M; Stewart, M. (1989) The Dictyostelium gelation factor shares a putative actin binding site with α-actinins and dystrophin and also has a rod domain containing six 100-residue motifs that appear to have a cross-beta conformation. J. Cell Biol. 109: 607-618.

O'Reilly, G and Clarke, F. M. (1993) Identification of an actin binding region in aldolase. FEBS Lett. 321: 69-72.

Oda, A; Druker, B. J; Smith, M; Salzman, E. (1992) Association of pp60src with triton X-100-insoluble residue in human blood platelets requires platelet aggregation and actin polymerisation. J. Biol. Chem. 267: 20075-20081.

Oefner, C and Suck, D. (1986) Crystallographic refinement and structure of DNaseI at 2A resolution. J. Mol. Biol. 192: 605-632.

Ohmstede, C. A; Farrell, F. X; Reep, B. R; Clemetson, K. J; Lapentina, E. G. (1990). Rap2B: A RAS-related GTP-binding protein from platelets. PNAS (USA) 87: 6527-6531.

Ohta, Y; Nishida, E; Sakai, H; Miyamoto, E. (1989) Dephosphorylation of cofilin accompanies heat shock-induced nuclear accumulation of cofilin. J. Biol. Chem. 264: 16143-16148.

Ohtsuka, E; Matsuki, S; Ikehara, M; Takahashi, Y; Matsubara, K. (1985). An alternative approach to deoxyoligonucleotides as hybridization probes by insertion of deoxyinosine at ambiguous codon positions. J. Biol. Chem. 260: 2605-2608.

Olson, N. J; Pearson, R. B; Needleman, D. S; Hurwitz, M. Y; Kemp, B. E; Means, A. R. (1990) Regulatory and structural motifs of chicken gizzard myosin light chain kinase. PNAS (USA). 87: 2284-2288.

Ono, S; Abe, H; Nagoaka, R; Obinata, T. (1993) Co-localisation of ADF and cofilin in intranuclear actin rods of cultured muscle cells. J. Musc. Res. Cell Motil. 14: 195-204.

Onoda, K; Yu, F; Yin, H. (1993) gCap39 is a nuclear and cytoplasmic protein. Cell Motil. Cytoskel. 26: 227-238.

Otey, C; Pavalko, F; Burridge, K. (1990) An interaction between α-actinin and the β_1 integrin subunit in vitro. J. Cell. Biol. 111: 721-729.

Otto, J. J (1994) Actin bundling proteins. Curr. Opin. Cell Biol. 6: 105-109.

Otto, J. J. (1990) Vinculin. Cell Motil. Cytoskel. 16: 1-6.

Ozawa, M; Baribault, H; Kemler, R. (1989) The cytoplasmic domain of the cell adhesion molecule uvomorulin associates with three independent proteins structurally related in different species. EMBO J. 8: 1711-1717.

Ozawa, M; Ringwald, M; Kemler, R. (1990) Uvomorulin-catenin complex formation is regulated by a specific domain in the cytoplasmic region of the cell adhesion molecule. PNAS (USA). 87: 4246-4250.

Pacaud, M and Derancourt, J. (1993) Purification and further characterisation of macrophage 70kDa protein, a calcium regulated, actin binding protein identical to L-plastin. Biochemistry. 32: 3448-3455.

Pacaud, M and Harricane, M. (1993) Macrophage alpha actinin is not a calcium modulated actin binding protein. Biochemistry. 32: 363-374.

Parry, D. A. D and Steinert, P. M. (1992) Intermediate filament structure. Curr. Opin. Cell Biol. 4: 94-98.

Pataloni, D and Carlier, M-F. (1993) How profilin promotes actin filament assembly in the presence of thymosin β4. Cell. 75: 1007-1014.

Paterson, J and O'Hare, K. (1991) Structure and transcription of the *singed* locus of *Drosophila melanogaster*. Genetics. 129: 1073-1084.

Pearlstone, J. R; Weber, M; Lees-Miller, J. P; Carpenter, M. R; Smillie, L. B. (1987) Amino acid sequence of chicken gizzard smooth muscle SM22α. J. Biol. Chem. 262: 5985-5991.

Peitsch, M. C; Polzar, B; Stephan, H; Crompton, T; MacDonald, H. R; Mannherz, H. G; Tschopp, J. (1993) Characterisation of endogenous deoxyribonuclease involved in nuclear DNA degradation during apoptosis (programmed cell death). EMBO-J. 12: 371-377.

Peng, I and Fischman, D. A. (1991) Post-translational incorporation of actin into myofibrils in vitro: evidence for isoform specificity. Cell Motil. Cytoskel. 20: 158-168.

Peterson, N; Lankenau, D. H; Mitchell, H; Young, P; Corces, V. (1994) Forked proteins are components of fiber bundles present in developing bristles of *Drosophila melanogaster*. Genetics. 136: 173-182.

Petrini, M; Emerson, D. L; Galbraith, R. M. (1983) Linkage between surface immunoglobulin and cytoskeleton of B-lymphocytes may involve Gc protein. Nature. 306: 73-74.

Piperno, G; Ledviet, M; Chang, X, J. (1987) Microtubules containing acetylated α-tubulin in mammalian cells in culture. J. Cell Biol. 104: 289-302.

Pittenger, M and Helfman, D. (1992) In vitro and in vivo characterisation of four fibroblast tropomyosins produced in bacteria: TM-2, TM-3, TM5a and TM5b are colocalised in interphase fibroblasts. J. Cell Biol. 118: 841-858.

Pittenger, M; Kazzaz, J; Helfman, D. (1994) Functional properties of non-muscle tropomyosin isoforms. Curr. Opin. Cell Biol. 6: 96-104.

Pollard, T. D and Cooper, J. A (1986). Actin and actin-binding proteins. A critical evaluation of mechanisms and functions. Ann. Rev. Biochem, 55; 987-1035.

Pollard, T. D and Rimm, D. L. (1991) Analysis of cDNA clones for *Acanthamoeba* profilin-I and profilin-II shows end-to-end homology with vertebrate profilins and a small family of profilin genes. Cell Motil. Cytoskel. 20: 169-177.

Pollard, T. D. (1990) Actin. Curr. Opin. Cell Biol. 2: 33-40.

Pollard, T. D; Doberstein, S. K; Zot, H. G. (1991) Myosin I. Ann. Rev. Physiol. 53: 653-681.

Pollard, T. D; Selden, C; Maupin, P. (1984) Interaction of actin filaments with microtubules. J. Cell Biol. 99 (suppl): 33s-77s.

Pomorski, P and Grebecka, L. (1993) Is actin involved in the nuclear division in *Amoeba proteus* ? Cell Biol. Intl. 17: 521-524.

Porter, J. A and Montell, C. (1993) Distinct roles if the *Drosophila* ninaC kinase and myosin domains revealed by systematic mutagenesis. J. Cell Biol. 122: 601-612.

Porter, J. A; Hicks, J. L; Williams, D. S; Montell, C. (1992) Differential localisations of and requirements for the two ninaC kinase/myosins in photoreceptor cells. J. Cell Biol. 116: 683-694.

Prendergast, G. C and Ziff, E. B. (1991) Mbh1: A novel gelsolin/severin-related protein which binds actin in vitro and exhibits nuclear localisation in vivo. EMBO-J. 10: 757-766.

Price, G. J.; Jones, P.; Davison, M. D.; Patel, B.; Eperon, I. C.; Critchley, D. R. (1987) Isolation and characterization of a vinculin cDNA from chick-embryo fibroblasts. Biochem. J. 245: 595-603.

Pring, M; Weber, A; Bubb, M. R. (1992) Profilin-actin complexes directly elongate actin filaments at the barbed end. Biochemistry. 31: 1827-1936.

Prinjha, R. K; Shapland, C. E; Hsuan, J. J; Totty, N. F; Mason, I. J; Lawson, D. (1994) Cloning and sequencing of cDNAs encoding the actin cross-linking protein transgelin defines a new family of actin-associated proteins. Cell Motil. Cytoskel. 28: 243-255.

Pumplin, D and Bloch, R. (1993) The membrane skeleton. Trends Cell Biol. 3: 113-117.

Putkey, J. A; Sweeney, H. L; Campbell, S. T. (1989) Site-directed mutation of the trigger calcium-binding sites in cardiac troponin C. J. Biol. Chem. 264: 12370-12378.

Quaranta, V and Jones, J. C. R. (1991). The internal affairs of an integrin. Trends Cell Biol. 1: 2-4.

Quinlan, R. A and Stewart, M. (1991) Molecular interactions of intermediate filaments. BioEssays. 13: 597-600.

Quirk, S; Maciver, S. K; Ampe, C; Doberstein, S. K; Kaiser, D. A; Van Damme, J; Vandekerckhove, J. S; Pollard, T. D. (1993) Primary structure of and studies on *Acanthamoeba* actophorin. Biochemistry. 32: 8525-8533.

Raats, J. M. H; Pieper, F. R; VreeEgberts, W. T. M; Verijp, K. N; Ramaekers, F. C. S; Bloemendal, H. (1990) Assembly of amino-terminally deleted desmin in vimentin free cells. J. Cell Biol. 111: 1971-1985.

Rana, A; Ruff, P; Maalouf, G; Speicher, D; Chishti, A. (1993) Cloning of human erythroid dematin reveals another member of the villin family. PNAS (USA). 90: 6651-6655.

Rankin, S and Rozengurt, E. (1994) Platelet-derived growth factor modulation of focal adhesion kinase ($p125^{FAK}$) and paxillin tyrosine phosphorylation in Swiss 3T3 cells. J. Biol. Chem. 269: 704-710.

Rapolee, D. A; Mark, D; Banda, M. J; Werb, Z. (1991) Wound macrophages express TGF-α and other growth factors in vivo: analysis by mRNA phenotyping. Science. 241: 708-712.

Rastinejad, F; Conboy, M. J; Rando, T. A; Blau, H. M. (1993) Tumor suppression by RNA from the 3' untranslated region of α-tropomyosin. Cell. 75: 1107-1117.

Ray, K; Wang, X; Zhao, M; Cooke, N. E. (1991) The rat vitamin D binding protein (Gc-globulin) gene. J. Biol. Chem. 266: 6221-6229.

Reedy, M. C; Beall, C; Fyrberg, E. (1991) Mutations at the N-terminus of *Drosophila* actin. Biophys. J. 59: 187a.

Rees, D; Ades, S; Singer, S; Hynes, R. (1990) Sequence and domain structure of talin. Nature. 347: 685-689.

Reeves, S. A; Helman, L. J; Allison, A; Israel, M. A. (1989) Molecular cloning and primary structure of glial fibrillary acidic protein. PNAS (USA). 86: 5178-5182.

Reinhard, M; Halbrugge, M, Scheer, U; Wiegand, C; Jockusch, B; Walter, U. (1992) The 46/50 kDa phoshoprotein VASP purified from human platelets is a novel protein associated with actin filaments and focal contacts. EMBO-J. 11: 2063-2070.

Ridley, A. J and Hall, A. (1992) the small GTP-binding protein Rho regulates assembly of focal adhesions and actin stress fibres in response to growth factors. Cell. 70: 389-399.

Ridley, A. J. (1994) Membrane ruffling and signal transduction. BioEssays. 16: 321-327.

Riederer, B and Matus, A. (1985) Differential expression of distinct microtubule-associated proteins. PNAS (USA). 82: 6006-6009.

Rihs, H. P and Peters, R. (1989) Nuclear transport kinetics depend on phosphorylation-site-containing sequence flanking the karyophilic signal of the simian virus 40 T-antigen. EMBO-J. 8: 1479-1484.

Rimm, D and Pollard, T. D. (1989) Purification and characterisation of Acanthamoeba nuclear actin-binding protein. J. Cell Biol. 109: 585-591.

Risinger, M. A and Lin, S. (1988) localisation of tensin in chicken embryo fibroblasts during cell-substrate contact formation and disruption. J. Cell Biol. 107: 256a.

Rodriguez del Castillo, A; Lemaire, S; Tchakarov, L; Jeyapragasan, M; Doucet, J-P; Vitale, M; Trifaro, J. (1990) Chromaffin cell scinderin, a novel calcium dependent actin filament severing protein. EMBO-J. 9: 43-52.

Rodriguez, J. R and Paterson, B. M. (1990) Yeast heavy chain mutant: Maintenance of the cell type specific budding pattern and the normal deposition of chitin and cell wall components requires an intact myosin heavy chain gene. Cell Motil. Cytoskel. 17: 301-308.

Rodriguez-Fernandez, J. L.; Geiger, B; Salomon, D.; Ben-Ze'ev, A. (1993) Suppression of vinculin expression by antisense transfection confers changes in cell morphology, motility, and anchorage-dependent growth of 3T3 cells. J. Cell Biol. 122: 1285-1294.

Rodriguez-Fernandez, J. L.; Geiger, B; Salomon, D.; Sabanay, I.; Zoller, M.; Ben-Ze'ev, A. (1992b) Suppression of tumorigenicity in transformed cells after transfection with vinculin cDNA. J. Cell Biol. 119: 427-438.

Rodriguez-Fernandez, J. L; Geiger, B; Salomon, D.; Ben-Ze'ev, A. (1992a) Overexpression of vinculin suppresses cell motility in Balb/C 3T3 cells. Cell Motil. Cytoskel. 22: 127-134.

Rosen, A; Keenan, K. F; Thelen, M; Nairn, A; Aderem, A. (1990) Activation of protein kinase C results in the displacement of its myristoylated, alanine-rich substrate from punctate structures in macrophage filopodia. J. Exp. Med. 172: 1211-1215.

Rouleau, G; Merel, P; Lutchman, M; Samson, M; Zucman, J; Marineau, C; Hoang-Xuan, K; Demczuk, S; Desmaze, C; Plougastel, B; et al. (1993) Alteration in a new gene encoding a putative membrane-organising protein causes neurofibromatosis type 2. Nature. 363: 515-521.

Rowe, T and Kendrick-Jones, J. (1992) Chimeric myosin regulatory light chains identify the subdomain responsible for regulatory function. EMBO-J. 11: 4715-4722.

Rubenstein, P. A. (1990) The functional importance of multiple actin isoforms. BioEssays. 12: 309-315.

Ruhnau, K; Gaertner, A; Wegner, A. (1989) Kinetic evidence for insertion of actin monomers between the barbed ends of actin filaments and barbed end-bound insertin, a protein purified from smooth muscle. J. Mol. Biol. 210: 141-148.

Ruiz-Opazo, N and Nadal-Ginard, B. (1987) α-tropomyosin gene organization. J. Biol. Chem. 262: 4755-4765.

Sadler, I; Crawford, A; Mitchelson, J; Beckerle, M. (1992) Zyxin and cCrp: Two interactive LIM domain proteins associated with the cytoskeleton. J. Cell Biol. 119: 1573-1587.

Safer, D; Elzinga, M; Nachmias, V. T. (1991) Thymosin β4 and Fx, an actin-sequestering peptide, are indistinguishable. J. Biol. Chem. 268: 4029-4032.

Safer, D; Golla, R; Nachmias, V. T. (1990) Isolation of a 5 kilodalton actin sequestering peptide from human blood platelets. PNAS (USA). 87: 2536-2540.

Sahlas, J. D; Milankov, K; Park, P. C; De Boni, U. (1993) Distribution of SnRNPs, splicing factor SE35 and actin in interphase nuclei: immunocytochemical evidence for differential distribution during changes in functional states. J. Cell Sci. 105: 347-357.

Sahr, L; Laurila, P; Kotula, L; Scarpa, A; Coupal, E; Leto, T; Linnenbach, A; Winkelmann, J; Speicher, D; Marchesi, V; Curtis, P; Forget, B. (1990) The complete cDNA and polypeptide sequences of human erythroid α-spectrin. J. Biol. Chem. 265: 4434-4443.

Saiki, R. K; Gelfand, D. H; Stoffel, S; Scharf, S. J; Higuchi, R; Horn, G. T; Mullis, K. B; Erlich, H. A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science. 239: 487-491.

Sakurai, T; Kurokawa, H; Nonomura, Y. (1991) Comparison between gelsolin and adseverin domain structure. J. Biol. Chem. 266: 15979-15983.

Salmon, E. D. (1989) Cytokinesis in animal cells. Curr. Opin. Cell Biol. 1: 541-547.

Sambrook, J; Fritsch, E. F; Maniatis, T. (1989) Molecular cloning: a laboratory manual. Second Edition. Cold Spring Harbor Laboratory Press.

Samstag, Y; Eckerskorn, C; Wesselborg, S; Henning, S; Wallich, R; Meuer, S. C. (1994) Costimulatory signals for human T-Cell activation induce nuclear translocation of pp19/cofilin. PNAS (USA). 91: 4494-4498.

Sanger, F. S; Nicklen, S; Coulson, A. R. (1977) DNA sequencing with chain terminating inhibitors. PNAS (USA). 74: 5463-5467.

Sanger, J. W; Sanger, J. M; Kreis, T. E; Jockusch. B. M. (1980) Reversible translocation of cytoplasmic actin into the nucleus caused by dimethylsulfoxide. PNAS (USA). 77: 5268-5272.

Sato, N; Funayama, N; Nagafuchi, A; Yonemara, S; Tsukita, S; Tsukita, S. (1992) A gene family consisting of ezrin, radixin and moesin. Its specific localisation at actin filament/plasma membrane association sites. J. Cell Sci. 103: 131-143.

Satterwhite, L. L and Pollard, T. D. (1992) Cytokinesis. Curr. Opin. Cell Biol. 4: 43-52.

Sattilaro, R. F. (1986) Interaction of microtubule-associated protein 2 with actin filaments. Biochemistry. 25: 2003-2009.

Sauterer, R. A; Eddy, R. J; Hall, A; Condeelis, J. (1991) Purification and characterisation of aginactin, a newly identified agonist-regulated actin capping protein from Dictyostelium amoebae. J. Biol. Chem. 266: 24533-24539.

Schafer, D. A; Waddle, J. A, Cooper, J. A (1993). Localisation of capZ during myofibrillogenesis in cultured chicken muscle. Cell Motil. Cytoskel, 25: 317-335.

Schaller, M. D and Parsons, J. T. (1993) Focal adhesion kinase: an integrin-linked protein kinase. Trends Cell Biol. 3: 258-262.

Scheel, J; Ziegelbauer, K; Kupke, T; Humbel, B. M; Noegal, A. A; Gerisch, G; Schleicher, M. (1989) Hisactophilin, a histidine-rich actin-binding protein from *Dictyostelium discoidium*. J. Biol. Chem. 264: 2832-2839.

Schejter, E and Wieschaus, E. (1993) *Bottleneck* acts as a regulator of the microfilament network governing cellularisation in *Drosophila* embryo. Cell. 75: 373-385.

Schevzov, G; Lloyd, C; Gunning, P. (1992) High level expression of transfected β- and γ-actin genes differentially impact on myoblast cytoarchitecture. J. Cell Biol. 117: 775-786.

Schliwa, M and Honer, B. (1993) Microtubules, centrsomes and intermediate filaments in directed cell movement. Trends Cell Biol. 3: 377-380.

Schmid, M. F; Agris, J. M; Jakana, J; Matsudaira, P; Chiu, W. (1994) Three-dimensional structure of a single filament in the *Limulus* acrosomal bundle: scruin binds to homologous helix-loop beta motifs in actin. J. Cell Biol. 124: 341-350.

Schmid, M. F; Matsudaira, P; Jeng, T. W; Jakana, J; Towns-Andrews, E; Bordas, J; Chiu, W. (1991) Crystallographic analysis of acrosomal bundle from *Limulus* sperm. J. Mol. Biol. 221: 711-725.

Schwarz, M. A; Owaribe, K; Kartenbeck, J; Franke, W. W. (1990) Desmosomes and hemidesmosomes: Constitutive molecular components. Ann. Rev. Cell Biol. 6: 461-91.

Schwob, E and Martin, R. P. (1992) New yeast actin-like gene required late in the cell cycle. Nature. 355: 179-182.

Sellars, J. M. (1990) Regulation of cytoplasmic and smooth muscle myosins. Curr. Opin. Cell Biol. 3: 98-104.

Selve, N and Wegner, A. (1986) Rate of treadmilling of actin filaments in vitro. J. Mol. Biol. 187: 627-631.

Seykora, J. T; Ravetch, J. V; Aderem, A. (1991) Cloning and molecular characterisation of the murine macrophage 68kDa protein kinase C substrate and its regulation by bacterial lipopolysaccharide. PNAS (USA). 88: 2505-2509.

Shapland, C. E; Hsuan, J. J; Totty, N. F; Lawson, D. (1993) Purification and properties of transgelin: a transformation and shape change sensitive actin gelling protein. J. Cell Biol. 121: 1065-1073.

Shapland, C. E; Lowings, P; Lawson, D. (1988) Identification of new actin-associated polypeptides that are modified by viral transformation and changes in cell shape. J. Cell Biol. 107: 153-161.

Sheterline, P and Sparrow, J. C. (1994) Protein profile: Actin. Volume I, Issue I. Academic Press. ISSN: 1070-33667.

Shoeman, R. L and Traub, P. (1993) Assembly of intermediate filaments. BioEssays. 15: 605-611.

Simpson, L and Wieschaus, E. (1990) Zygotic activity of the *nullo* locus is required to stabilise the actin network during cellularisation in *Drosophila* embryos. Development. 110: 851-863.

Singer, R. H. (1992) The cytoskeleton and mRNA localisation. Curr. Opin. Cell Biol. 4: 15-19.

Skoufias, D. A and Scholey, J. M. (1993) Cytoplasmic microtubule-based motor proteins. Curr. Opin. Cell Biol. 5: 95-104.

Skoufias, D. A; Burgess, T. L; Wilson, L. (1990) Spatial and temporal colocalisation of the golgi apparatus and microtubules rich in detyrosinsated tubulin. J. Cell Biol. 111: 1929-1937.

Small, V. J. (1989) Microfilament-based motility in nonmuscle cells. Curr. Opin. Cell Biol. 1: 75-79.

Smith, V. L; Kaetzel, M. A; Dedman, J. R. (1990) Stimulus-response coupling: the search for intracellular calcium mediator proteins. Cell Regulation. 1: 165-172.

Southwick, F and Stossel, T. (1981) Isolation of an inhibitor of actin polymerisation from human polymorphonuclear leuckocytes. J. Biol. Chem. 256: 3030-3036.

St-Pierre, B; Couture, C; Laroche, A; Pallotta, D. (1993) Two developmentally regulated mRNAs encoding actin-binding proteins in *Physarum polycephalum.* Biochim. Biophys. Acta. 1173: 107-110.

Staden, R. (1982) An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. N. A. R. 10: 2951-2961.

Staiger, C. J; Goodbody, K. C; Hussey, P. J; Valenta, R; Drobak, B, K; Lloyd, C. W. (1993) Profilin multigene family of maize: differential expression of three isoforms. Plant-J. 4: 631-641.

Stappenbeck, T. S and Green, K. J. (1992) The desmoplakin carboxyl terminus coaligns with and specifically disrupts intermediate filament networks when expressed in cultured cells. J. Cell. Biol. 116: 1197-1209.

Stearns, T; Evans, L; Kirschner, M. (1991) γ-tubulin is a highly conserved component of the centrosome. Cell. 65: 825-836.

Steinert, P. M and Roop, D. R. (1988) Molecular and cellular biology of intermediate filaments. Ann. Rev. Biochem. 57: 593-625.

Stella, M-C; Schauerte, H; Straub, K. L; Leptin, M. (1994) Identification of secreted and cytosolic gelsolin in *Drosophila.* J. Cell Biol. 125: 607-616.

Stewart, M. (1993) Intermediate filament structure and assembly. Curr. Opin. Cell Biol. 5: 3-11.

Stossel, T. P. (1993) On the crawling of animal cells. Science. 260: 1086-1094.

Stossel, T. P; Chaponnier, C; Ezzell, R. M; Hartwig, J. H; Janmey, P. A; Kwiatkowski, D. J; Lind, S. E; Smith, D. B; Southwick, F. S; Yin, H. L; Zaner, K. S. (1985) Nonmuscle actin-binding proteins. Ann. Rev. Cell Biol. 1: 353-402.

Strasser, P; Gimona, M; Moessler, H; Herzog, M; Small, V. J. (1993) Mammalian calponins: Identification and expression of genetic variants. FEBS Lett. 330: 13-18.

Stumpo, D. J; Graff, J. M; Albert, K; Greengard, P; Blackshear, P. J. (1989) Molecular cloning characterisation and expression of a cDNA encoding the '80- to 87-kDa'

myristoylated alanine-rich C kinase substrate: A major cellular substrate for protein kinase C. PNAS (USA) 86: 4012-4016.

Sudhof, T. C; Czernik, A. J; Kao, H-T; Takei, K; Johnston, P. A; Horiuchi, A; Kanazir, S. D; Wagner, M. A; Perin, M. S; DeCamilli, P; Greengard, P; (1989) Synapsins: Mosiacs of shared and individual domains in a family of synaptic vesicle phosphoproteins. Science. 245: 1474-1480.

Sundell, C. L and Singer, R. H. (1991) Requirement of microfilaments in sorting of actin messenger RNA. Science. 253: 1275-1277.

Sutoh, K and Yin, H. L. (1989) Actin-gelsolin contact sites as revealed by end-label fingerprintings. Biochemistry. 28: 5269-5275.

Takagi, T; Konischi, K; Mabuchi, I. (1988) Amino acid sequence of starfish oocyte depactin. J. Biol. Chem. 263: 3097-3102.

Takahashi, K and Nadal-Ginard, B. (1991) Molecular cloning and sequence analysis of smooth muscle calponin. J. Biol. Chem. 266: 13284-13288.

Takeichi, M. (1990) Cadherins: A molecular family important in selective cell-cell adhesion. Ann. Rev. Biochem. 59: 327-252.

Takeichi, M. (1993) Cadherins in cancer: Implications for invasion and metastasis. Curr. Opin. Cell Biol. 5: 806-811.

Takeuchi, K; Takahashi, K; Abe, M; Nishida, W; Hiwada, K; Nabeya, T; Maruyama, K. (1991) Colocalisation of immuno-reactive forms of calponin with actin cytoskeleton in platelets, fibroblasts and vascular smooth muscle. J. Biochem (Tokyo). 109: 311-316.

Tanaka, K and Singer, S. J. (1988) Phosphotyrosine-modified proteins are connected at the membranes of epithelial and endothelial cells during tissue development in chick embryos. J. Cell Biol. 106 1757-1764.

Tang, K; Qin, Z; Leto, T; Marchesi, V; Benz, E. (1990) Heterogeneity of mRNA and protein products arising from the protein 4.1 gene in erythroid and nonerythroid tissues. J. Cell Biol. 110: 617-624.

Tatsumi, R; Hattori, A; Takahashi, K. (1993) Substructure of nebulin filaments: localisation and characterisation of subfragments produced by 0.1mM $CaCl_2$. J. Biochem. 113: 797-804.

Tellam, R. L; Morton, D. J; Clarke, F. M. (1989) A common theme in the amino acid sequences of actin and many actin-binding proteins? TIBS. 14: 130-133.

Theriot, J. A and Mitchison, T. J. (1993) The three faces of profilin. Cell. 75: 835-838.

Theriot, J. A; Rosenblatt, J; Portnoy, D. A; Goldschmidt-Clermont, P. J; Mitchison, T. J. (1994) Involvement of profilin in the actin-based motility of *L. monocytogenes* in cells and in cell-free extracts. Cell. 76: 505-517.

Theurkauf, W. (1994) Through the bottleneck. Curr. Biol. 4: 76-78.

Thierfelder, L; Watkins, H; MacRae, C; Lamas, R; McKenna, W; Vasberg, H-P; Seidman, J. G; Seidman, C. E. (1994) α-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: A disease of the sarcomere. Cell. 77: 701-712.

Thweat, R; Lumpkin, C. K; Jr; Goldstein, S. (1992) A novel gene encoding a smooth muscle protein is overexpressed in senescent human fibroblasts. Biochem. Biophys. Res. Comm. 187: 1-7.

Tinsley, J; Blake, D; Roche, A; Fairbrother, U; Rise, J; Byth, B; Knight, A; Kendrick-Jones, J; Suthers, G; Love, D; Edwards, Y; Davis, K. (1992) Primary structure of dystrophin related protein. Nature. 360: 591-593.

Titus, M. A. (1993a) Myosins. Curr. Biol. 5: 77-81.

Titus, M. A. (1993b) From fat yeast and nervous mice to brain myosin-V. Cell. 75: 9-11.

Torti, M; Ramaschi, G; Sinigaglia, F; Lapentina, E. G; Balduini, C. (1994) Glycoprotein IIb-IIIa and the translocation of Rap2B to the platelet cytoskeleton. PNAS (USA). 91: 4239-4243.

Totty, N. F; Waterfield, M. D; Hsuan, J. J. (1992) Accelerated high-sensitivity microsequencing of proteins and peptides using a miniature reaction cartridge. Protein Sci. 1:12115-12125.

Towbin, H; Stachelin, T; Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. PNAS (USA).76: 4350-4354.

Trofatter, J; MaCollin, M; Rutter, J; Murrell, J; Duyao, M; Parry, D; Eldridge, R; Kley, N; Menon, A; et al. (1993) A novel moesin, ezrin, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. Cell. 72: 791-800.

Tsukita, S; Hieda, Y; Tsukita, S. (1989) A new 82-kD barbed end-capping protein (radixin) localised in the cell-to-cell adherens junctions: purification and characterisation. J. Cell Biol. 108: 2369-2382.

Tsukita, S; Itoh, M; Nagafuchi, A; Yonemura, S; Tsukita, S. (1993) Submembraneous junctional plaque proteins include potential tumor suppressor molecules. J. Cell Biol. 123: 1049-1053.

Tsukita, S; Itoh, M; Tsukita, S. (1989) A new 400kD protein from isolated adherens junctions: its localisation at the undercoat of adherens junctions and at microfilament bundles such as stress fibres and circumferential bundles. J. Cell Biol. 109: 2905-2915.

Tsukita, S; Oishi, K; Akiyama, T; Yamanashi, Y; Yamamoto, T; Tsukita, S. (1991) Specific proto-oncogenic tyrosine kinases of Src family are enriched in cell-to-cell adherens junctions where the level of tyrosine phosphorylation is elevated. J. Cell Biol. 113: 867-879.

Tsukita, S; Oishi, K; Sato, N; Sagara, J; Kawai, A; Tsukita, S. (1994) ERM family members as molecular linkers between cell surface glycoprotein CD44 and actin-based cytoskeletons. J. Cell Biol. 126: 391-401.

Tsukita, S; Tsukita, S; Nagafuchi, A; Yonemura, S. (1992) Molecular linkage between cadherins and active filaments in cell-cell adherens junctions. Curr. Opin. Cell Biol. 4: 834-839.

Tufty, R. M and Kretsinger, R. H. (1975) Troponin and parvalbumin calcium binding regions predicted in myosin light chain and T4 lysozyme. Science. 187: 167-169.

Turner, C. E. (1991) Paxillin is a major phosphotyrosine-containing protein during embryonic development. J. Cell Biol. 115: 201-207.

Turner, C. E; Glenney, J. R; Burridge, K (1990). Paxillin: A new vinculin-binding protein present in focal adhesions. J. Cell Biol. 111: 1059-1068.

Turner, C. E; Kramarcy, N; Sealock, R; Burridge, K. (1991) Localisation of paxillin, a focal adhesion protein, to smooth muscle dense plaques, and the myotendinous and neuromuscular junctions of skeletal muscle. Exp. Cell Res. 192: 651-655.

Turner, C. E; Schaller, M. D; Parsons, J. T. (1993) Tyrosine phosphorylation of focal adhesion kinase pp125^{FAK} during development: relation to paxillin. J. Cell Sci. 105: 637-645.

Turunen. O; Winquist, R; Pakkanen, R; Grozeschik, K-H; Wahlstrom, T; Vaheri, A. (1989) Cytovillin, in microvillar, Mr 75, 000 protein. J. Biol. Chem. 264: 16727-16732.

Ullrich, A; Coussens, L; Hayflick, S; Dull, T; Gray, A; Tam, A; Lee, J; Yarden, Y; Libermann, T; Schlessinger, J; Downward, E; Mayes, E; Whittle, N; Waterfield, M; Seebing, P. (1985) Human epidermal growth factor receptor cDNA sequence and

aberrant expression of the amplified gene in A431 epidermal carcinoma cells. Nature. 309: 418-425.

Umekage, T and Kato, K. (1991) A mouse brain cDNA encodes a novel protein with the protein kinase C phosphorylation site domain common to MARCKS. FEBS Lett. 286: 147-151.

Vale, R. D; Schnapp, B. J; Reese, T. S; Sheetz, M. P. (1985) Movement of organelles along filaments dissociated from the axoplasm of the squid giant axon. Cell. 40: 449-454.

Vallee, R. (1991) Cytoplasmic dynein: advances in microtubule-based motility. Trends Cell Biol. 1: 25-29.

Valtorta, F; Benfenati, F; Greengard, P. (1992a) Structure and function of the synapsins. J. Biol. Chem. 267: 7195-7198.

Valtorta, F; Greengard, P; Fesce, R; Chieregatti, E; Benfenati, F. (1992b) Effects of the neuronal phosphoprotein synapsin I on actin polymerisation. J. Biol. Chem. 267: 11281-11288.

Van Etten, R, A; Jackson, P; Baltimore, D. (1989) the mouse type IV c-Abl gene product is a nuclear-protein, and activation of transforming ability is associated with cytoplasmic localisation. Cell. 58:669-678.

Van Etten, R. A; Jackson, P; Baltimore, D; Sanders, M. C; Matsudaira, P. T; Janmey, P. A. (1994) The COOH-terminus of the c-Abl tyrosine kinase contains distinct F- and G-actin binding domains with bundling activity. J. Cell Biol. 124: 325-340.

Van Roy, F and Mareel, M. (1992) Tumor invasion: effects of cell adhesion and motility. Trends Cell Biol. 2: 163-169.

Vancompernolle, K; Gimona, M; Herzog, M; Van Damme, J; Vandekerckhove, J; Small, V. J. (1990) Isolation and sequence of a tropomyosin-binding fragment of turkey gizzard calponin. FEBS Lett. 274: 146-150.

Vancompernolle, K; Goethals, M; Huet, C; Louvard, D; Vandekerckhove, J. (1992) G- to F-actin modulation by a single amino acid substitution in the actin binding site of actobindin and thymosin β4. EMBO-J. 11: 4739-4746.

Vancompernolle, K; Vandekerckhove, J; Bubb, M. R; Korn, E. D. (1991) The interfaces of actin and *Acanthameoba* actobindin. Identification of a new acting-binding motif. J. Biol. Chem. 266: 15427-15431.

Vandekerckhove, J and Vancompernolle, K. (1992) Structural relationships of actin binding proteins. Curr. Opin. Cell Biol. 4: 36-42.

Vandekerckhove, J and Weber, K. (1979) The complete amino acid sequence of actins from bovine aorta, bovine heart, bovine fast skeletal muscle and rabbit slow skeletal muscle. Differentiation. 14: 123-133.

Vandekerckhove, J; Bauw, G; Vancompernolle, K; Honore, B; Celis, J. (1990a) Comparative two-dimensional gel analysis and microsequencing identifies gelsolin as one of the most prominent down-regulated markers of transformed human fibroblast and epithelial cells. J. Cell Biol. 111: 95-102.

Vandekerckhove, J; Kaiser, D. A; Pollard, T. D. (1989) *Acanthameoba* actin and profilin can be cross-linked between glutamic acid 364 of actin and lysine 115 of profilin. J. Cell Biol. 109: 619-629.

Vandekerckhove, J; Van Damme, J; Vancompernolle, K, Bubb, M. R; Lambooy, P; Korn, E. D. (1990b) The covalent structure of actobindin structure. J. Biol. Chem. 265: 12801-12805.

Velaz, L; Ingraham, R. H; Chalovich, J. M. (1990) Dissociation of the effect of caldesmon on the ATPase activity and on the binding of smooth heavy meromyosin to actin by partial digestion of caldesmon. J. Biol. Chem. 265: 2929-2934.

Verderame, M; Alcorta, D; Egnor, M; Smith, K; Pollack, R. (1980) Cytoskeletal F-actin patterns quantitated with fluorescein isothiocyanate-phalloidin in normal and transformed cells. PNAS (USA). 77: 6624-6628.

Verheyen, E. M and Cooley, L. (1994) Profilin mutations disrupt multiple actin-dependent processes during *Drosophila* development. Development. 120: 717-728.

Vinson, V. K; Archer, S. J; Lattman, E. E; Pollard, T. D; Torchia, D. A. (1993) Three-dimensional solution structure of profilin-I. J. Cell Biol. 122: 1277-1283.

Vojtek, A; Haarer, B. K; Field, J; Gerst, J; Pollard, T. D; Brown, S; Wigler, M. (1991) Evidence for a functional link between profilin and CAP in the yeast *S. cerevisiae* . Cell. 66: 497-505.

Vostal, J. G. and Shulman, R. N. (1993) Vinculin is a major platelet protein that undergoes a Ca^{2+}-dependent tyrosine phosphorylation. Biochem. J. 294: 675-680.

Wachsstock, D. H.; Wilkins, J. A.; Lin, S. (1987) Specific interaction of vinculin with α-actinin. Biochem. Biophys. Res. Comm. 146: 554-560.

Wachsstock, D; Schwarz, W; Pollard, T. (1993) Affinity of α-actinin for actin determines the structure and mechanical properties of actin filament gels. Biophys. J. 65: 205-214.

Waites, G; Graham, I; Jackson, P; Millake, D; Patel, B; Blanchard, A; Weller, P; Eperon, I; Critchley, D. (1992) Mutually exclusive splicing of calcium-binding domain exons in chick α-actinin. J. Biol. Chem. 267: 6263-6271.

Walsh, M. P. (1991) Calcium-dependent mechanisms of regulation of smooth muscle contraction. Biochem. Cell Biol. 69: 771-800.

Wang, C. L. A; Wang, L. W. C; Xu, S; Lu, R. C; Saavedra-Alanis, V; Bryan, J. (1991) Localisation of the calmodulin and the actin binding sites of caldesmon. J. Biol. Chem. 266: 9166-9172.

Wang, K; McCarter, R; Wright, J; Beverly, J; Ramirez-Mitchell, R. (1991) Regulation of skeletal muscle stiffness and elasticity by titin isoforms: A test of the segmental extension model of resting tension. PNAS (USA). 88: 7101-7105.

Waseem, A and Palfrey, H. C. (1988) Erythrocyte adducin. Comparison of the alpha- and beta-subunits and multiple-site phosphorylation by protein kinase C and cAMP-dependent protein kinase. Eur. J. Biochem. 178: 563-573.

Waterston, R. H; Thomson, N. J; Brenner, S. (1980) Mutants with altered muscle structure in *Caenorhabditis elegans* . Dev. Biol. 77: 271-302.

Way, M and Weeds, A. (1990) Cytoskeletal ups and downs. Nature. 344: 292-294.

Way, M and Weeds, A. (1988) Nucleotide sequence of pig plasma gelsolin. J. Mol. Biol. 203: 1127-1133.

Way, M; Pope, B; Gooch, J; Hawkins, M; Weeds, A. (1990) Identification of a region in segment 1 of gelsolin critical for actin binding. EMBO-J. 9: 4103-4109.

Way, M; Pope, B; Weeds, A. (1992) Evidence for functional homology in the F-actin binding domains of gelsolin and α-actinin: implications for the requirements of severing and capping. J. Cell Biol. 119: 835-842.

Weber, A; Nachmias, V. T; Pennise, C. R; Pring, M; Safer, D. (1992) Interaction of thymosin β4 with muscle and platelet actin: implications for actin sequestration in resting platelets. Biochemistry. 31: 6179-6185.

Weeds, A and Maciver, S. (1993) F-actin capping proteins. Curr Opin Cell Biol. 5: 63-69.

Weigt, C; Gaertner, A; Wegner, A; Korte, H; Meyer, H. (1992) Occurrence of an actin-inserting domain in tensin. J. Mol. Biol. 227: 593-595.

Weiner, O; Murphy, J; Griffiths, G; Schleicher, M; Noegal, A. (1993) The actin-binding protein comitin (p24) is a component of the golgi apparatus. J. Cell Biol. 123: 23-34.

Weller, P. A.; Ogryzko, E. P.; Corben, E. B.; Zhidkova, N. I.; Patel, B.; Price, G. J.; Spurr, N. K.; Koteliansky, V. E.; Critchley, D. R. (1990) Complete sequence of human vinculin and assignment of the gene to chromosome 10. PNAS (USA). 87: 5667-5671.

Wessel, D and Flugge, U. I. (1984) A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Anal. Biochem. 138: 141-143.

Wilbur, W. J and Lipman, D. J. (1983) Rapid similarity searches of nucleic acid and protein data bank. PNAS (USA). 80: 726-730.

Wilkins, J. A; Risinger, M. A; Lin, S. (1986) Studies on proteins that copurify with smooth muscle vinculin: identification of immunologically related species in focal adhesions of nonmuscle and Z lines of muscle cells. J. Cell Biol. 103: 1483-1494.

Wilkins, J. A; Risinger, M. A; Lin, S. (1987) Comparison of tensin and vinculin localisation in muscle and nonmuscle cells. J. Cell Biol. 105: 130a.

Williams, R and Coluccio, L. M. (1994) Novel 130kDa rat liver myosin-I will translocate actin filaments. Cell Motil. Cytoskel. 27: 41-48.

Winder, S. J and Walsh, M. P. (1990) Smooth muscle calponin. Inhibition of actomyosin MgATPase and regulation by phosphorylation. J. Biol. Chem. 265: 10148-10155.

Winder, S. J; Pato, M. D; Walsh, M. P. (1992) Purification and characterisation of calponin phosphatase from smooth muscle. Biochem. J. 286: 197-203.

Winkelmann, J; Chang, J; Tse, W; Scarpa, A; Maechesi, V; Forget, B. (1990) Full length sequence of the cDNA for human erythroid β-spectrin. J. Biol. Chem. 265: 11827-11932.

Wodnar-Filipowicz, A; Gubler, U; Furuichi, Y; Richardson, M; Nowoswiat, E. F; Poonian, M. S; Horecker, B. L. (1984) Cloning and sequence analysis of cDNA for rat spleen thymosin β4. PNAS (USA). 81: 2295-2297.

Woodjett, J. R; Gould, K. L; Hunter, T. (1986) Substrate specificity of protein kinase C. Eur. J. Biochem. 161: 177-184.

Wu, H and Parsons, J. T. (1993) Cortactin, an 80/85-kilodalton pp60^{src} substrate, is a filamentous actin-binding protein enriched in the cell cortex. J. Cell Biol. 120: 1417-1426.

Wuestenhube, L. J and Luna, E. J. (1987) F-actin binds to the cytoplasmic surface of ponticulin a 17kDa integral glycoprotein from *Dictyostelium discoidium.* J. Cell Biol. 105: 1741-1751.

Wuestenhube, L. J; Chia, C. P; Luna, E. J. (1989) Immunofluorescence localisation of ponticulin in motile cells. Cell Motil. Cytoskel. 13: 245-263.

Wuestenhube, L. J; Speicher, D. W; Shariff, A; Luna, E. J. (1991) F-actin affinity chromatography of detergent-solubulized plasma membranes: purification and initial characterisation of ponticulin from *Dictyostelium discoidium* plasma membranes, Methods Enzymol. 196: 47-65.

Xue, F and Cooley, L. (1993) *Kelch* encodes a component of intercellular bridges in *Drosophila* egg chambers. Cell. 72: 681-693.

Yamashiro, S; Yamakita, Y; Hosoya, H; Matsumura, F. (1991) Phosphorylation of non-muscle caldesmon by $p34^{cdc2}$ during mitosis. Nature. 349: 169-172.

Yamashiro-Matsumura, S and Matsumura, F. (1985) Purification and characterisation of an F-actin-bundling 55 kilodalton protein from Hela cells. J. Biol. Chem. 260: 5087-5097.

Yamashiro-Matsumura, S and Matsumura, F. (1986) Intracellular localisation of the 55kDa actin-bundling protein in cultured cells; spatial relationships with actin, alpha-actinin, tropomyosin, and fimbrin. J. Cell Biol. 103: 631-640.

Yang, F; Demma, M; Warren, V; Dharmawardhame, S; Condeelis, J. (1990) Identification of an actin-binding protein from Dictyostelium as elongation factor 1α. Nature. 347: 494-496.

Yang, Q and Tonks, N. (1991) Isolation of a cDNA clone encoding a human protein tyrosine phosphatase with homology to the cytoskeletal-associated proteins band 4.1, ezrin and talin. PNAS (USA). 88: 5949-5953.

Yonezawa, N; Nishida, E; Iida, K; Kumajai, H; Yahara, I; Sakai, H. (1991) Inhibition of actin polymerisation by a synthetic dodecapeptide patterned on the sequence of the actin binding site of cofilin. J. Biol. Chem. 266: 10485-10489.

Yu, F-X; Johnston, P; Sudhof, T; Yin, H. L. (1990) gCap39, a calcium ion and polyphosphoinositide-regulated actin capping protein. Science. 250: 1423-1415.

Yu, F-X; Lin S-C; Morrison- Bogorad, M; Yin, H. L. (1994) Effects of thymosin β4 and thymosin β10 on actin structures in living cells. Cell Motil. Cytoskel. 27: 13-25.

Zachary, I; Sinnett-Smith, J; Rozengurt, E. (1992) Bombesin, vasopressin, and endothelin stimulation of tyrosine phosphorylation in Swiss 3T3 cells. J. Biol. Chem. 267: 19031-19034.

Zigmond, S. H. (1993) Recent quantitative studies of actin filament turnover during cell locomotion. Cell Motil. Cytoskel. 25: 309-316.

Zigmond, S; Furukawa, R; Fechheimer, M. (1992) Inhibition of actin filament depolymerisation by the Dictyostelium 30,000-D actin-bundling protein. J. Cell Biol. 119: 559-567.

Zu, Y; Kohno, M; Kubota, J; Nishida, E; Hanaoka, M; Namba, Y. (1990) Characterisation of interleukin 2 stimulated 65-kilodalton phosphoprotein in human T cells. Biochemistry: 29: 1055-1062.

Zvelebil, M. J; Barton, G. J; Taylor, W. R; Sternberg, M. J. E. (1987) Prediction of protein secondary structure and active sites using the alignment of homologous sequences. J. Mol. Biol. 195: 957-961.