A STRUCTURAL AND FUNCTIONAL INVESTIGATION INTO THE FILAMIN A G2593E MUTATION: IMPLICATIONS FOR NEUROLOGICAL DISEASE

Submitted in partial fulfillment of a Master of Philosophy (Science) (Biochemistry) at the Institute of Molecular BioSciences, Massey University, Pamerston North

MANICKAVASAGAR ANANDASAYANAN

2012

ACKNOWLEDGEMENTS

I wish to thank my supervisor Dr. Andrew Sutherland-Smith at the Center for Structural Biology, Institute of Molecular BioSciences, Massey University for his advice and encouragement throughout the course of this study and in the preparation of the thesis.

I thank Professor Stephen Robertson (University of Otago) for the kind gift of *filamin A* cDNA clone.

I acknowledge the generosity of my employers HortResearch Ltd, New Zealand for providing the necessary funds to conduct my research.

ABSTRACT

The dimeric F-actin cross-linking protein human filamin a (hsFLNa) is an actin binding protein, which modulates the properties of the actin cytoskeleton and plays a major role in maintaining the integrity of plasma membrane associated actin, the viscoelastic properties of the cytoplasm, endocytosis, cytoplasmic streaming, cell division and cell motility [McGough *et al*, 1998; Small *et al*, 2002; Popowicz *et al*, 2006]. Dimerization is crucial for actin cross-linking functions of filamins [Davies *et al*, 1980] and the most C-terminal repeat of hsFLNa (hsFLNa24) is sufficient for hsFLNa dimerization [Himmel *et al*, 2003].

Several mutations in *hsflnA* are associated with pathologies such as periventricular nodular heterotopia (PVNH) [Robertson, 2005]. In this study we examined the possible cause and effect of a G2593E mutation in hsFLNa24 in a male patient diagnosed with PVNH on protein functionality. This was done by comparing relevant biochemical properties of wildtype and mutant hsFLNa24 proteins. For this purpose, recombinant proteins were expressed from cloned *hsflnA*24 coding sequence.

Full length wildtype *hsflnA24* (wt *hsflnA24*) was amplified from a cDNA library prepared from the PVNH patient. wt *hsflnA24* was used as the template to generate mutant *hsflnA24* (mt *hsflna24*) by site directed mutagenesis. The amplified wildtype and mutant sequences were cloned and over-expressed in an *Eschericia coli* system. hsFLNa24 proteins were isolated from crude protein extracts by immobilized metal affinity chromatography, purified by gel filtration and concentrated.

Several features of the mutant protein indicate that it has a high-entropy, disordered structure. Thermal stability of the two proteins determined by melt curve analyses showed that mt hsFLNa24 is less stable than wt hsFLNa24, their respective melting temperatures being < 303 K and 317 K, respectively. The mutant protein also tended to aggregate during concentration and was prone to precipitation during low speed centrifugation.

The two proteins displayed different elution volumes in size exclusion chromatography; wt hsFLNa24 eluted at a volume characteristic of a dimer indicating that in its native

form wt hsFLNa24 exists as a dimer, while the mutant's elution volume suggests that in most probabilities it exists as a monomer with a slightly larger molecular size. The oligomerisation status of the proteins in solution was further confirmed in crosslinking assays using the chemical crosslinker ethylene glycol succinimido succinate ester (EGS). In assays where the proteins at a final concentration of 4 μM were reacted with 1.3 mM EGS, the major form of wt FLNa24 was shown to be a high molecular weight dimeric species, while dimerisation was inhibited in mt hsFLNa24 and the protein exists predominantly in a monomeric state.

Both proteins were subject to various crystallizing conditions. Only wt hsFLNa24 produced diffracting crystals which have immunoglobulin (Ig)-like folds of the E-set superfamily [Murzin *et al*, 1995] with a predominantly β–sheet structure, where seven β-strands organized as two anti-parallel β–sheets of four (ABED) and three (CFG) strands, respectively, are arranged as a β-sandwich. The asymmetric unit consists of a dimer.

The dimer interface is formed by β-strands C and D of the monomers and the G2593 residue occurs in β-strand-C; it resides within the hydrophobic core of the dimer interface, where it is involved in a putative hydrophobic stacking. This glycine residue is highly conserved in most vertebrate filamins. Ablation of the G residue in conjunction with substitution by a polar residue is predicted to play a major role in disrupting the hydrophobic nature of the interface thereby inhibiting dimerisation. The inhibition effect arose probably more through unfavourable entropy change induced by substitution of the native G by an E residue than a reorganization of the dimer interface.

The actin cross-linking ability of filamins is ascribed to its dimerization mediated by the C-terminal repeats and the loss of the latter function may have serious implications to the patient harbouring the G2593E mutation. The relatively mild symptoms exhibited by the patient leads one to believe that there may be compensatory mechanisms in operation or that the basic premise that the most C-terminal repeat is important for filamin dimerisation is questionable.

CONTENTS

List of figures

CHAPTER I - INTRODUCTION

1.1. ACTIN CROSS-LINKING PROTEINS

The internal and external shape of cells is determined by the cell cytoskeleton composed of actin filaments, microtubules and intermediate filaments. The actin cytoskeleton, which may be organized into various discrete structures such as filapodia, lamellipodia and stress fibres, consists of filamentous actin (F-actin) formed by reversible polymerization of globular monomeric actin (G-actin) [Winder & Ayscough, 2006]. In addition to the proteins that comprise the fibers themselves, several accessory proteins are required to connect the fibres to one another and to the plasma membrane, or other cellular structures. These actin binding proteins (ABP) integrate actin organization in response to cellular signals and orchestrate the polymerization (and depolymerization) of actin filaments into bundles and their attachment to membranes [Annesley *et al*, 2007].

F-actin cross-linking proteins modulate the properties of actin cytoskeleton and thereby play a major role in maintaining the integrity of plasma membrane associated actin, the viscoelastic properties of the cytoplasm, endocytosis, cytoplasmic streaming, cell division and cell motility [McGough *et al*, 1998; Small *et al*, 2002; Popowicz *et al*, 2006]. Cross-linking of actin filaments is achieved by binding of different actin filaments by at least two actin binding domains (ABD) of a cross-linking protein, occurring either in tandem on the same polypeptide chain and separated by a variable linker, or singly on two dimerized polypeptide chains, [Puius *et al*, 1998].

Filamins represent one of three classes of actin cross-linking proteins; another subclass consists of fimbrins and plastins which are characterized by the presence of two tandem ABDs on the same polypeptide chain. A third subclass comprises α-actinin, spectrin, and plakin families which form non-covalent dimers through their spectrin domain composed of a variable number of triple helical coiled-coil domains [Gimona *et al*, 2002; Korenbaum & Rivero, 2002].

Filamins (FLN), the first non-muscle actin crosslinking proteins to be recognized [reviewed by Stossel *et al*, 2001], are a family of high molecular weight, homodimeric, scaffolding phosphoproteins first identified in rabbit

macrophage cells [Hartwig & Stossel, 1975]. Their name refers to their ability to form filamentous arrays along stress fibres as observed in fibroblasts [Wang *et al*, 1975]. Filamins organize actin filaments into parallel arrays, or orthogonal, isotropic, three-dimensional webs in a concentration dependant manner [Niederman *et al,* 1983] and link them to cellular membranes through a variety of transmembrane proteins [Hou *et al*, 1990; van der Flier & Sonnenberg, 2001]. Although filamins reside and execute most of their functions in the cytoplasm, they are also involved in nuclear-translocation functions and transcription regulation [Loy et al, 2003]. Filamins and filamin-like proteins are found in almost all eukaryotes, from the higher vertebrates to lower forms such as the slime mould *Dictylostelium discoideum*; the exceptions are yeasts and plants [Stossel *et al*, 2001].

1.2. FILAMIN STRUCTURE

Homo sapiens filamin A (hsFLNa, ABP-280, α-filamin, FLN1) [Gorlin et al, 1990] encoded by the X-linked gene *flnA* typifies *H. sapiens* filamins, is an elongated protein, which appear as a V-shaped homodimer (figure 1) in electron micrographs of rotary shadowed molecules [Tyler *et al*, 1980] as a result of two filamin subunits linked to each other at their C-terminus through dimerisation of their C-terminal tail.

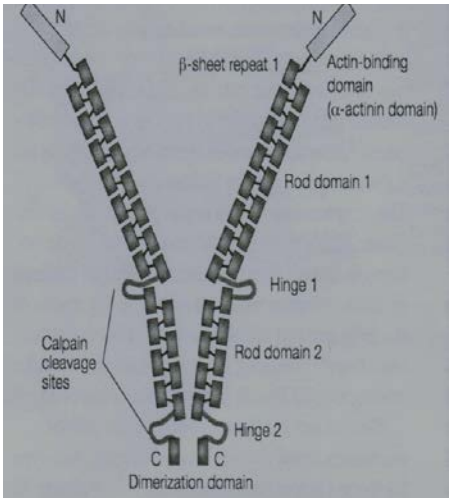

Figure 1: Schematic of the human filamin A (hsFLNa) dimer. The N-terminal end (N) contains the ABD. This is followed by 24 repeats of antiparallel β-pleated sheet Ig-like domains which form the rod domain. Dimerization occurs through the 24th repeat. Two hinge regions occur between filamin repeats 15 and 16 (hinge 1) and 23 and 24 (hinge 2) respectively [Stossel *et al*, 2001].

The hsFLNa monomer is approximately 5 nm wide and 80 nm long [Matsudaira, 1991] and has a molecular weight of 280 kDa, hence also referred to as ABP280. It contains an ABD (actin binding domain) in its N-terminal region which is followed by a semi-flexible rod region of 24 tandem immunoglobulin (Ig)-like domains.

Analyses of various α-actinin-like ABD family members using deletion and point mutations, peptide binding, immunochemistry and sequence alignments suggests that the hsFLN ABD comprises amino acids 1-269 containing two calponin homology (CH) subdomains [Stradal *et al*, 1998], CH1 (40-152) and CH2 (170-267), separated by a 17 aa linker. The length of the linker determines the geometry of actin cross-linking [Stossel *et al*, 2001]. Each CH domain contains four main and two shorter α-helices connected by variable loops; three of the main helices form a triple helical bundle with the N-terminal α–helix which lies perpendicular to the bundle as observed in spectrin and dystrophin [Banuelos *et al*, 1998; Norwood *et al*, 2000]. Deletion analyses and antibody binding studies show that three potential actin binding sites, ABS1 and -2 in CH1 and ABS3 in CH2 reside in the ABD [Clark *et al*, 2009]; of these, a hydrophobic sequence corresponding to the last 15 aa C-terminal residues in the last α–helix of CH1, first recognized in β-spectrin,, dystrophin and α-actinin and conserved in all filamins and actin binding proteins such as fimbrin, calponin, plectin and utrophin [Bresnick *et al*, 1991; Levine *et al* 1992; Corrado *et a*l, 1994; Lebart et al, 1994; Morris *et al*, 1999] is crucial for actin binding. The CH2 domain by itself does not bind actin, but enhances binding by CH1 which alone can bind actin [Gimona & Mittal, 1998; Banuelos *et al*, 1998; van der Flier & Sonnenberg, 2001]. Conformational movements between CH1 and CH2 subdomains may be of potential importance for actin binding as in fimbrin and plectin [Garcia-Alvarez *et al*, 2003; Brower *et al*, 1995; Klein *et al*, 2004].The binding site in actin isoforms resides between residues 105-120 and 360-372 in actin subdomain-1 [Mejean *et al*, 1992; Lebart *et al*, 1993; Robertson *et al*, 2003]. Hydrophobic interactions dominate the binding interface, while hydrophilic interactions are also involved since binding is susceptible to ionic strength [Lebart *et al*, 1994]. Binding is facilitated by Ca^{++} and calmodulin, the latter which binds to a cryptic binding site located between residues 50-96 of CH1 that is probably

uncovered following ABD binding to actin [Nakamura *et al*, 2005]. A 70 kDa proteolytic fragment that includes filamin ABD was shown to bind actin with a dissociation constant of 0.1 μM.

The tandem, filamin repeats (FR) of the rod region, measure 17Åx22Åx48Å each as shown by Fucini *et al* [1997] and others. The repeats, which are capable of reversible unfolding [Yamazaki *et al*, 2003] consist of ~96 (82-98) aa residues, in which short zones rich in turn predictors such as proline, glycine, aspartic acid and asparagine alternate with zones with a high β–potential. The repeat domains have immunoglobulin (Ig)-like folds of the E-set superfamily [Murzin *et al*, 1995] with a predominantly β–sheet structure, where seven β-strands organized as two anti-parallel β–sheets of four (ABED or 1-2-5-4) and three (CFG or 3-6-7) strands, respectively, are arranged as a β-sandwich (figure 2).

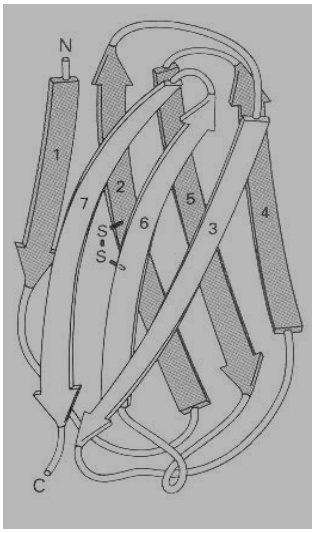

Figure 2: The immunoglobulin fold. The Ig β-barrel fold consists of a pair of β-sheets linked by loops, a disulfide bond and hydrophobic interactions. The two antiparallel β–sheets are formed by strands 1, 2, 5, 4 and 3, 6, 7, respectively.

The crystal structure of an hsFLNa fragment containing repeats 19-21 shows that repeats do not necessarily fold as distinct domains [Lad *et al* 2007]. The crystal structure revealed that repeat Ig 20 is partially unfolded but contributes one β-strand to domain 21 and that domain 21 is positioned between repeats Ig19 and Ig20. The strand from repeat 20 occupies the binding site for integrin. The structures of repeat domain 17 of hsFLNa in complex with the GPIbα peptide corresponding to amino acids 556-577 [Nakamura *et al* 2006], hsFLNa domain 21 and integrin β7 peptide [Kiema *et al*, 2006] and hsFLNc repeat 23 [Sjekloca *et al*, 2007] support this view. The

generic structure of filamin repeats comprise a β-sandwich immunoglobulin-like confomation consisting of a three-stranded and four-stranded β-sheet

The rod region is interrupted by two flexible, non-modular hinge regions (figure 1) that are susceptible to proteolysis by the calcium dependant protease calpain; they are known to bind to multiple proteins [Stossel *et al*, 2001; van der Flier & Sonnenberg, 2001]. These swivel regions are believed to confer flexibility to the otherwise rigid rod region [Gorlin *et al*, 1990; Robertson *et al*, 2003]. Hinge 1 (32 residues), positioned between repeats 15-16 and hinge 2 (38 residues) positioned between repeats 23-24 divide the rod region into rod 1 and rod 2 regions (figure 1).

The linkers between the other filamin repeats are short in comparison and are rich in proline including a Pro-Ala-Pro stabilizing sequence which lends rigidity to the rod region [Popowicz *et* al, 2004]. Salt bridges exist between neighbouring repeats [Popowicz *et* al, 2006]. There is little interaction between repeat domains of the two dimer chains, except at the dimerizing domain [Tyler *et al*, 1980].

1.3. FILAMIN GENES

Sequencing of cloned genes and genome database searches reveal that filamins exist as small multigene families in several organisms. Alternative gene splicing, or multiple gene promoters [Barry *et al*, 1993; Li *et al*, 1999] result in many filamin isoforms that often outnumber the filamin genes. The human genome contains three functional filamin genes *flnA*, *flnB* and *flnC* located at chromosomal loci Xq28 [Gorlin *et al*, 1993; Maestrini *et al*, 1993], 3p14.3 [Gariboldi *et al*, 1998; Takafutu *et al*, 1998] and 7q32-q35 [Maestrini *et al*, 1993], respectively. Genomic organization and intron-exon boundaries are highly conserved among the genes. X-linked *flnA* is composed of 47 exons which span nearly 26kb. The actin binding domain is encoded by exons 2-5 and the 24 filamin repeats forming the rod backbone by the remaining 42 exons. While the boundaries of exons do not correlate with the filamin repeats, the two hinge regions are each encoded by single exons (31 and 49, respectively), which suggests that they may be insertions of recent origin [Patrosso *et al*, 1994]. Alternative splicing possibly occurs in an eight aa segment (aa 1649-1956) in repeat 15 encoded by exon 29 to give rise to

hsFLNa splice variants [Maestrini *et al*, 1993]. In *flnB*, the 72 bp exon which gives rise to hinge 1 is characterized by exon-intron junctions showing the invariant GT:AG splice site which provides for alternative splicing [Xu *et al*, 1998]. Likewise, hinge 1 of *flnC* is encoded by a 99 bp exon flanked by two introns of ~500 and 550 bp. Differential splicing of this exon generates two transcripts, where one of them lacks the hinge region [Xie *et al*, 1998].

A homology search of other vertebrate genomes shows that most species have three filamin genes [Birney *et al*, 2004]. The *fln* gene of *Gallus gallus* retina filamin (ggFLNb) also shows a 10-nucleotide pyrimidine-rich sequence at position 961 bp, which is a potential source of alternative splicing leading to the production of truncated isoforms. Two promoters situated at least 8kb apart within this gene may be responsible for encoding at least two translation products [Barry *et al*, 1993].

Two filamin genes showing varying levels of homology to human filamins are observed in *D. melanogaster*. Screening of a P1 genomic library of *D. melanogaster* with filamin cDNAs obtained by yeast two-hybrid screen led to the cloning of the *cherio* (*cher*) gene at locus 89E10-89F4 on the right arm of chromosome 3 [Guo *et al*, 2000], which encodes ddFLN1-20 (filamin240) and ddFLN1-9 (filamin90). The gene is made up of 17 exons spanning ~27kb; five splice sites clustered around hinge 1 and 2 [Sokol & Cooley, 1999] are also conserved in human *fln*. An internal promoter, probably located in the intron following the exon coding for hinge 1 is responsible for the smaller 3 kb *cheerio* transcript that encodes ddFLN1-9 [Li *et al*, 1999]. Additional filamin and filamin-like genes have been mapped to the right arms of chromosomes 2 and 3 at loci 58F8-59A2 and 97F4, respectively [van der Flier & Sonnenberg, 2001].

The nematode *Caenorhabditis elegans* has two filamin genes and several X-linked filamin-like genes. The genome database of *C. elegans* contains a filamin gene with an ORF arrangement similar to that of *cher* [Pudas, 2006].

The human pathogen *Entamoeba histolytica* responsible for amoebic dysentery, appears to contain a single filamin gene [Stossel *et al*, 2001].

1.4. FILAMIN PROTEINS

1.4.1. HUMAN FILAMIN ISOFORMS

The three human filamin paralogs share high protein sequence homology and the same molecular structure [Seo *et al*, 2009]. hsFLNa, -b and –c show 70-80% sequence homology at the amino acid level along their whole sequence, except at the less homologous hinge regions [Feng & Walsh, 2004]. They and their splice variants maintain the essential features of wildtype hsFLNa, but differ in certain respects. Alternative splicing of *flnA* results in two isoforms containing eight and 41 amino acid deletions, respectively, in FR 15 and in the linker between FR 19 and 20 (A_{var-1}) [Gorlin *et al*, 1990; Maestrini *et al*, 1993]. A naturally occurring C-terminal isoform of hsFLNa containing repeats 16 to 24 generated by proteolytic cleavage at the protease cleavage site in hinge 1 translocates to the nucleus and downregulates the androgen receptor [Loy *et al*, 2003].

hsFLNb (Filamin B, ABP-278, β-filamin, Filamin-3), which shows 70% aa identity to hsFLNa and possesses a hinge 1 region of 24 amino acid residues with no apparent homology to hsFLNa hinge 1; however, hinge 2 shows 43% identity with its counterpart in hsFLNa. hsFLNb splice variant B_{var-1} bears a 41 residue deletion between rod repeats 19 and 20, while hinge 1 region is absent in a second splice variant ΔH1 [Xu *et al*, 1998]. In two cardiac-specific variants B_{var-2} and B_{var-3}, four carboxy-terminal repeats including the dimerization zone are absent [van der Flier & Sonnenberg, 2001].

hsFLNc (Filamin C, ABPL, γ-filamin, Filamin-2), which is expressed preferentially in skeletal muscle and heart possesses both hinge regions and bears 70-72% sequence identity to hsFLNa and -b. Hinge 1 is 33 aa residue long and its sequence is entirely different to those of hinge 1 of hsFLNa and hsFLNb; hinge 2 shows 45% and 51% identity to corresponding regions of hsFLNa and hsFLNb, respectively. hsFLNc also contains a unique 82 aa insertion in FR 20. A truncated isoform of hsFLNc, which lacks the N-terminal ABD and therefore cannot bind actin, can still modulate cytoskeleton formation by competing with other paralogs for associated receptor proteins. Hinge 1 is deleted in the hsFLNc splice variant ΔH1 [Xie *et al*, 1998].

1.4.2. FILAMIN ORTHOLOGS

Paralogs and orthologs of filamin and filamin-like proteins from different taxonomic species share the basic structural features of hsFLN, but exhibit varying degrees of divergence in protein architecture and aa sequence. The CH domains of the ABD exhibit a considerable degree of conservation (20-60% sequence identity) [Popowicz *et al*, 2006]. The ABD domains of filamins and actin-binding proteins are thought to have evolved from an ancestral ABD containing gene by gene duplication, exon shuffling and mutation. [Patrosso *et al*, 1994; Korenbaum & Rivero, 2002]. Thus, CH domain sequences are widely used to discern evolutionary relationships among filamins and other actin binding proteins which contain single or multiple calponin homology domains (figure 3).

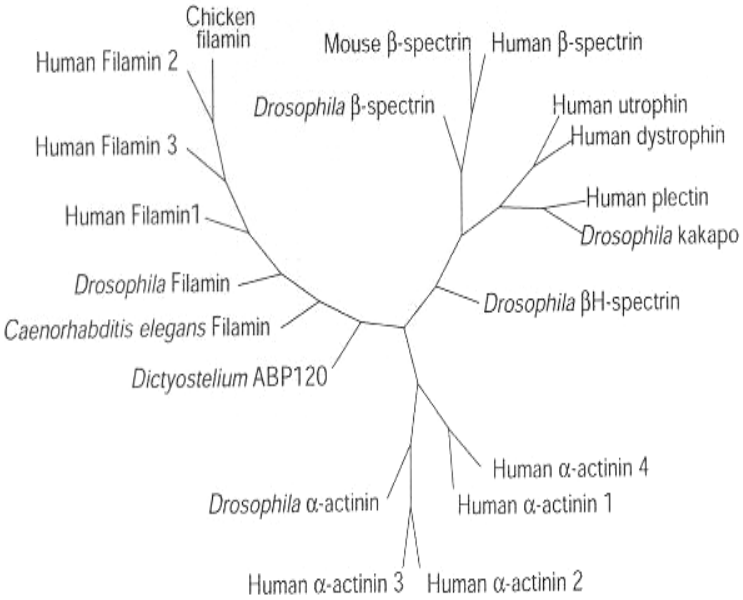

Figure 3: Phylogenetic tree of CH domains. The unrooted tree shows relationship among filamins and other actin binding proteins [Korenbaum & Rivero, 2002].

Variability in protein sequence is more pronounced in the rod repeat domains as exemplified by a comparison of repeat sequences of various species (figure 4).

Figure 4: Multiple alignment of hsFLNa repeat 24 sequence and FR sequences of 10 species using GeneIous® alignment software. GenBank accessions *Bos Taurus*-GI:18377580; *Canis Familiaris*-GI:74008821; *Danio rerio*-GI:125814244; *Homo sapiens* muscle isoform-GI:5419655; *Pan troglodytes*-GI:114615844; *Rattus norvegicus*-GI:109473203; *Gallus gallus*-GI:15341204; *Dictyostelium discoideum*-GI:66826629; *Xenopus tropicalis*-GI:111306161; *hsFLNa*- GI:57284203; *Mus musculus*-GI:123228989; *Sus scrofa*-GI:82465299. Conserved residues are shown in yellow.

Filamin orthologs in several vertebrates, from the amphibian *Xenopus tropicalis* to the mammal *Bos taurus*, contain two N-terminal calponin homology domains and a rod region of 24 filamin repeats, as in *H. sapiens* filamins. [Barry *et* al, 1993; Birney *et al*, 2004]. The gizzard- (ggFLNa), retina- (ggFLNb) and cgABP260 (ggFLNc) filamin isoforms of *G.gallus* [Wang *et al*, 1975; Barry *et al*, 1993; Tachikawa *et al*, 1997], which are smooth muscle-, nonmuscle- and pan muscle-type filamins, respectively, correspond to hsFLNa, -b and –c [Ohashi *et al*, 2005]. Amino acid identities between the corresponding chicken and human orthologs are 76%, 81% and 70%, respectively. The three isoforms contain hinge 2, but not hinge 1; ggFLNa has the longest hinge region (46 aa residues) of any vertebrate filamins. ggFLNc has a 76 aa insertion in repeat 20, similar to the 82 residue insertion in hsFLNc. A high level of aa identity (72-81%) is observed between the ABDs of chicken filamins, while the rod regions show lower homology (48-67% aa identity). Immunoblotting of selected tissues with highly specific antibodies shows that ggFLNa is expressed in the gizzard, ggFLNb in the gizzard and brain and to a lesser extent in the heart and ggFLNc only in muscular organs [Ohashi *et al*, 2005].

Among arthropods, filamin analogs have been reported in the mosquito *Anopheles gambia* and the fruitfly *Drosophila melanogaster*. The putative 260

kDa ortholog of *A. gambia* has two CH domains separated by a 23 aa linker at the N-terminus followed by a rod region of 22 filamin repeats [Hammond, 2007]. In *D. melanogaster*, based on available EST information, there are more than four filamin isoforms [Rus *et al*, 2006]. dmFLN1-20 (filamin-240), which is abundantly expressed in the outer and inner rims of the ovarian ring canals, shows 60% identity to *G.gallus* filamins and 46% identity to full length hsFLNa with a higher identity (70%) observed in the ABD regions [Li *et al*, 1999]. It has two hinge regions which show 34-52% identity to the corresponding human domains and a shorter rod region comprising 20 repeats in which repeats 6-9 of *H. sapiens* are deleted. This full length protein suppresses the development of lamellocytes associated with immune response [Rus *et al*, 2006] and possibly plays a role in in bringing together the structural and regulatory components required for recruitment of actin filaments in the ring canal [Sokol & Cooley, 1999]. The amino-terminal truncated isoform dmFLN1-9 transcribed by *cher* from an internal promoter and showing 54% aa identity to the carboxy terminal one third of hsFLNa, contains only hinge 2 and rod 2 consisting of nine C-terminal repeats. This 90 kDa protein is expressed in the somatic muscle sheath. Another carboxy-terminal truncated variant encoded by *cher* has only the ABD and rod 1 domains [Sokol & Cooley, 1998]. *Drososphila* jbug filamin isoform contains three calponin homology domains [Pudas, 2006].

Chromosome IV encoded filamin of *C. elegans* [Wilson *et al*, 1994] shows high identity to filamins of *H. sapiens*. However, other X-linked filamin-like proteins show more similarity to filamin jbug of *D. melanogaster* [Pudas, 2006].

Filamin and filamin-like proteins have been identified in three protozoans. ddFLN (gelation factor) isolated from cytoplasmic extracts of *Dictyostelium discoideum* [Condeelis *et al*, 1981] is regarded as the prototype filamin from which animal filamins arose by multiplication of the rod repeats [Fucini *et al*, 1997; McCoy *et al*, 1999]; the assignment of filamin repeats are now adjusted after the structure of ddFLN repeats. It is ~35 nm long and contains two calponin homology domains followed by a short rod region built of six ~100 residue repeats [Noegel *et al*, 1989]. In *ddfln*$^{-}$ mutants where the gene is disrupted, the actin cytoskeleton collapses and cells exhibit defects in

the rate and extent of pseudopod formation; this suggests that ddFLN plays a role in crosslinking actin filaments during pseudopod formation [Cox *et al*, 1995].

Photosensory- and thermosensory responses were also affected in null mutants. Another 240 kDa actin binding protein isolated from vegetative amoebae of *D.discoideum* cross reacts strongly with affinity purified anti-ggFLNa antibodies. The flexible rod shaped protein monomer measuring 70 nm, dimerizes when covalently crosslinked with dimethyl suberimidate. These and other properties suggest that it is a filamin-like protein [Hock & Condeelis, 1987]. The rod region of filamin ehFLN (ehABP-120) of *Entamoeba histolytica* contains four quasi-repetitive domains made up of 72 residues in repeats 1 and 2, 73 residues in repeat 3 and 43 residues in a part of repeat 4. The fourth repeat is interrupted at aa residue 615 by a 101-residue acidic-basic region rich in lysine and glutamic acid; this is followed by ~30 residues which terminate in a 100-residue C-terminal region known as END, which is homologous to domain 6 of *D.discoideum*. The protein localizes to the pseudopod and uroid regions during parasite movement implying a role for ehFLN in motility and capping of surface receptors [Vargas *et al*, 1996]; furthermore, over-expression of NH2- and –COOH domains of ehFLN in trophozites increased their motility and chemotactic response to serum added culture medium [Diaz-Valencia *et al*, 2007]. The 95 kDa filamin immune-analog apABP-F1 of *Amoeba proteus* which is preferentially expressed in the cortical layer, the nucleus and within the perinuclear area possesses a single N-terminal CH domain (residues 1-104) followed by one FR (residues 209-314) and a putative coiled coil region (residues 344-577) [Sobczak *et al*, 2007]. The minimal requirement for a filamin family member is the presence of filamin rod repeats [Stossel *et al*, 2001] and it remains to be seen whether apABP-F1 with a single FR would be considered a filamin ortholog.

1.5. HUMAN *fln* EXPRESSION

In early studies of filamin expression, no distinction was made between the isoforms and they were collectively called filamin; in cultured cells, filamins localized along stress fibres, to the cortical actin cytoskeleton, membrane ruffles of migrating cells and the cleavage furrow in dividing cells [Nunnally *et*

al, 1980; Langanger *et al*, 1984; Vadlamudi *et al*, 2002]. Filamins were first recognized as a protein family only after 10 years of their discovery [Pollard & Cooper, 1986]; since then, gene specific PCR primers and highly specific antibodies have been employed to untangle the expression patterns of individual isoforms.

The three *fln* paralogs are almost constitutively expressed to various degrees in human tissues [van der Flier & Sonnenberg, 2001]. hsFLNa is expressed in tissues of the prostate, uterus, placenta, bone marrow, liver, thyroid, Daudi cells, stomach, small intestine and skeletal muscle and is highly expressed in heart, lung, blood vessel and haematopoietic cells [Xu *et al*, 1998; Sheen *et al*, 2002; Sheen *et al*, 2005]. *flnA* is also widely expressed in all cortical layers of the brain during development; hsFLNa is highly expressed in leading processes and somata of migratory neurons during corticogenesis [Sheen *et al*, 2002].

hsFLNb is preferentially expressed in the kidney and pancreas [Takafuta *et al*, 1998]. The wildtype and splice variants of FLNb show different expression patterns; while they are both found in the liver, thyroid, spleen, placenta, bone marrow, brain, umbilical vein endothelial cells, retina and skeletal muscles, the wildtype is preferentially expressed in prostate, uterus, lung, liver, thyroid, stomach, lymph node, small intestine and spleen and the ΔH1 variant is mostly found in Daudi cells, fetal brain and the spinal cord. The B_{var-1} variant is detected in the heart, lung and skeletal muscle [Takafuta *et al*, 1998; Xu *et al*, 1998; van der Flier & Sonnenberg, 2002].

hsFLNc, which is not present in the central nervous system unlike hsFLNa and hsFLNb, is preferentially expressed in skeletal and cardiac muscles [Maestrini *et al*, 1993]; FLNc is the only isoform expressed in mature myotubes and in adult muscle fibres. It is enriched in Z-disks of the striated muscles, intercalated disks of cardiac muscles and in myotendinous junctions of skeletal muscles. Wildtype hsFLNc and its ΔH1 splice variants are also found in the thyroid, stomach, uterus, prostate, retina, spinal cord and bone marrow [Xie *et al*, 1998; Van der Ven *et al*, 2000].

1.6. PHYSIOLOGICAL FUNCTIONS OF HUMAN FILAMINS

The actin crosslinking property of filamin often overshadows its multifunctional nature [Clark *et al*, 2009]; filamins act as organizers of the actin cytoskeleton, integrators of cellular architectural signals, modulators of receptor binding and regulators of transcriptional activity (figure 5).

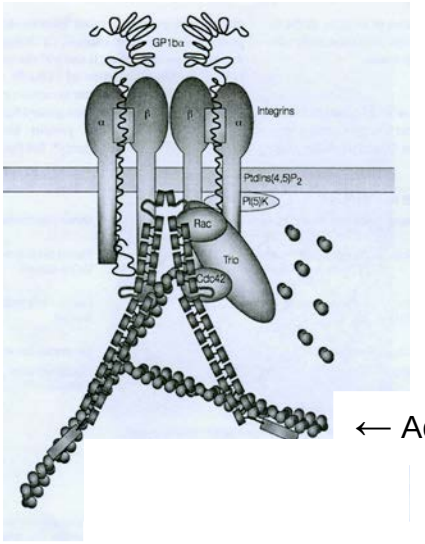

Figure 5. The multifunctional nature of filamins. hsFLNa is shown crosslinking actin filaments and engaging the transmembrane proteins GP1bα and β3 integrin as well as the small GTPases Cdc42 and Rac and their guanine nucleotide exchange factor Trio. Activated Rac stimulates phosphatidylinositol-4,5 bisulfate leading to promotion of G-actin polymerization [Stossel *et al*, 2001].

Deficiency of hsFLNa in M2 melanoma cell lines results in loss of ability to generate cytoplasmic extensions and migrate in response to environmental cues and blebbing of plasma membrane in melanoma cells [Cunningham *et al*, 1992; Cunningham, 1995]. FLNb knockout mice have shortened distal limbs, fused ribs and vertebrae and dysmorphic facial or calvarial bones, similar to human autosomal recessive spondylocarpotarsal phenotype associated with hsFLNb mutations [Lu *et al*, 2007].

Filamin functions are regulated post-translationally at many levels by proteolysis within the hinge regions [Gorlin *et al*, 1990;], phosphorylation [Wallach *et* al, 1978] by serine threonine kinases (protein kinase A, protein kinase C, calmodulin dependant protein kinase II), receptor occupancy by binding of phospholipids [Ohta *et* al, 1991; Furuhashi *et al*, 1992] and organellar compartmentalization [Liu *et al*, 1997; van der Flier & Sonnenberg, 2001; Feng & Walsh, 2004; Woo *et al*, 2004]. A small degree of redundancy

may exist among some filamin isoforms with respect to their functions [Feng & Walsh, 2004].

1.6.1. LIGAND BINDING

In cells, F-actin is in a dynamic state where it is remodeled with the help of proteins such as filamins in response to signal cascades; such events change actin polymerization kinetics or stabilize existing networks resulting in alterations to the viscosity, elasticity and mechanical resistance of cells necessary for processes such as contraction, adhesion, motility, organelle transport, phagocystis and cell division. By binding cortical actin by their N-terminal region and transmembrane receptors mainly by their C-terminal region (figure 5), filamins tether the two, thereby providing membrane surface stabilization. The association of membrane embedded proteins with actin cytoskeleton is believed to be important in regulating cell adhesion and extension and transmembrane signalling.

Gelation [Brotschi *et al*, 1978] occurs due to the crosslinking of actin filaments into orthogonal or parallel networks. Human FLNa is the most potent actin cross-linking protein [Bennet *et al*, 1984]. Small amounts of filamin cause large amounts of purified muscle actin to precipitate; one hsFLNa dimer per actin filament is sufficient to induce gelation [Ito *et al*, 1992]. Thus, it is suggested that filamins act as actin-saving proteins since much more actin would be needed to build actin networks in the absence of filamins. The high gelling ability of hsFLNa is due to the geometry of actin filament branching imposed by the protein. hsFLNa is able to crosslink F-actin chains over distances of more than 1000 Å [Popowicz *et al*, 2004].

While spectrin, fimbrin and α-actinin form primarily parallel actin bundles, filamins can also form orthogonal networks. The type of actin filament organization depends on the filamin:actin ratio. Filamin at 1:10-50 stoichiometry promotes parallel actin bundles and tighter networks, while a filamin:actin ratio of 1:150-740 leads to the formation of orthogonal networks [Brotschi *et al*, 1978; Dabrowska *et al*, 1985]. FLNa crosslinked actin shows high-angle branching and the interbranch distances are inversely proportional to filamin concentration. At branching points, the pointed ends of actin

filaments intersect with the sides of the other filament to form T-, X- or Y-junctions [Stossel *et al*, 2001].

Thus, filamin in low concentrations at the edge of cells crosslinks actin into a flexible, three-dimensional, high branching, cortical network, which presents a degree of resistance to mechanical stresses and also provides for fast remodeling. In a different context, filamin at higher concentrations, provides stretchability to actin contractile bands in the ventral side of adherent cells.

Actin filament organization also depends on the filamin type. The hinge regions, along with the filamin repeats which are able to undergo reversible unfolding under stress as observed by atomic force microscopy [Furuike *et al*, 2001] may be the structural basis for the intrinsic flexibility of actin networks generated by filamins [Popowicz *et al*, 2006]. Natural variants lacking hinge 1 result in tight actin fibres instead of orthogonal networks [Gorlin *et al*, 1990; Xu *et al*, 1998]; *in vitro* experiments demonstrate that actin crosslinked by H1 deleted recombinant filamin shows linear stiffening and breaking at much lower stresses [Gardel *et al*, 2006]

1.6.2. RECEPTOR BINDING

Filamins not only bind actin, but also interact with a multiplicity of transmembrane receptors and peripheral membrane proteins. Over 45 functionally diverse proteins are known to bind to filamins and these include presenilins, sarcoglycans, integrins, caveolin, FILIP, Rho family GTP-binding proteins and several other receptors. [Feng & Walsh, 2004; Robertson, 2005]. Remarkably, most of the interacting proteins bind to filamin between repeats 16-24 (figure 6).

By integrating membrane receptors such as β-integrins, α and δ sarcoglycans and Gp1bα and intra-cellular signaling proteins such as the small GTPases Rho, Rac, Ra1 and Cdc42 with the submembrane actin network and intracellular signaling components, filamins can activate local cellular processes. For instance, hsFLNA binds by its C-terminal 19 and 21 rod repeats to the cytoplasmic domain of type 1 transmembrane protein integrin [Kiema *et al*, 1998]. Binding of signal molecules in the extracellular

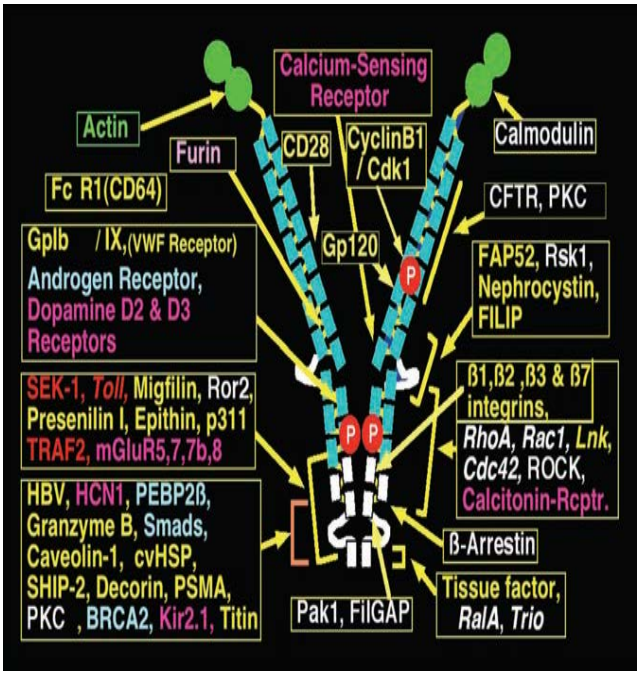

Figure 6: Filamin A binding partners and their location on its subunits [Stossel, 2010]

matrix to the extracellular ligand-binding domains of the receptor is transmitted across the membrane to initiate signaling. Mig-2 is a component of adhesion complexes where cells attach to the extracellular matrix and is an activating intermediate for matrix receptors such as integrins (figure 7). Using RNA interference, Tu *et al* [2003] showed that Mig-2 recruits migfilin to adhesion complexes. Migfilin binds hsFLNa through FR 21 which then

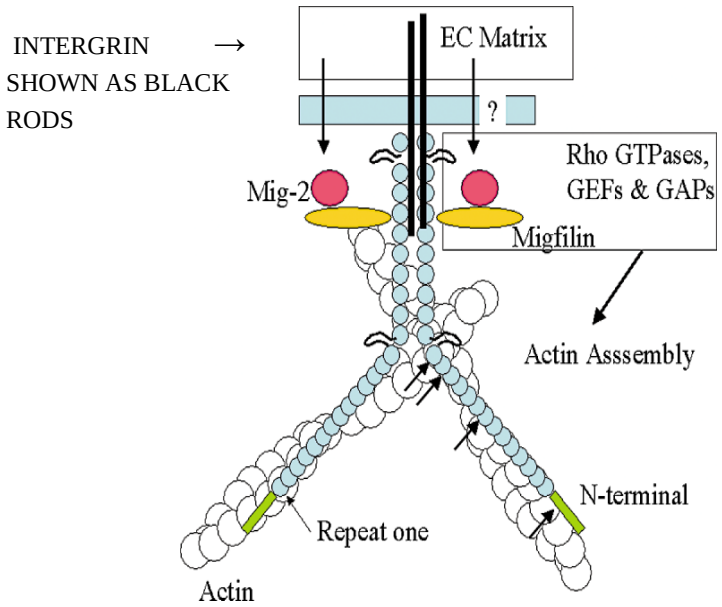

Figure 7: Schematic of filamin A scaffolding leading from Mig-2 to actin remodeling. hsFLNa orients with its C-terminal dimerization domain towards the plasma membrane where it binds transmembrane matrix receptors such as integrins. Receptor engagement by extracellular signals localizes Mig-2, which in turn recruits migfilin leading to Rho GTPase activation and actin assembly [Stossel & Hartwig, 2003].

promotes net actin assembly during cell spreading, either through direct or indirect activation of GTPases

Thus filamins act as scaffolding proteins that coordinate a variety of functions including signal transduction, receptor translocation and gene transcription regulation [Garcia & Jay, 2006]. Filamin A interacting protein FILIP regulates cortical cell migration out of the brain ventricular zone by regulating the level of hsFLNa in neurons within the subventricular zone [Nagano *et al*, 2002]. ddFLN of *D. discoideum* is thought to act as a scaffold in a photosensory signaling complex with RasD, ErkB, GRP125 and PKB. Since filamin contains no recognizable Ras binding domain, this interaction could be mediated by other proteins bound to actin [Khaire *et al*, 2007].

1.6.3. NUCLEAR TRANSPORT AND TRANSCRIPTION REGULATION

Nuclear localization of filamin A has been described [Loy *et al*, 2003; Yuan & Sheen, 2001] and is possibly involved in transcriptional regulation. hsFLNA sequesters transcription factors in the cytoplasm thereby preventing them from being active in the nucleus [Yoshida *et al*, 2005]. It is postulated that filamins may act as a nucleating center for co-localization of nuclear functional complexes participating in nucleo-cytoplasmic transport [Berry *et al*, 2004; Meng *et al*, 2004; Yuan & Shen, 2001; Ozanne *et al*, 2000]

1.7. MUTATIONS AND PATHOLOGICALCONDITIONS

Mice lacking filamin A owing to the introduction of a stop codon in domain 22 (Y2388X) display a phenotype exhibiting skeletal defects and cardiac malformation [Hart *et al*, 2006]. Mutations in *flnA* leading to loss-of-function of one allele cause the dominant, X-linked, brain malformation periventricular nodular heterotopia (PVNH), a condition in which normal heterotopic neurons fail to undergo radial migration from the brain subventricular zone to neocortex. At least 24 PVNH mutations (figure 8), mainly frameshift- or nonsense-mutations have been identified in the ABD, rod-1 and -2 regions [Robertson, 2005]. The frameshift mutations often cause abnormal RNA splicing or premature transcription termination.

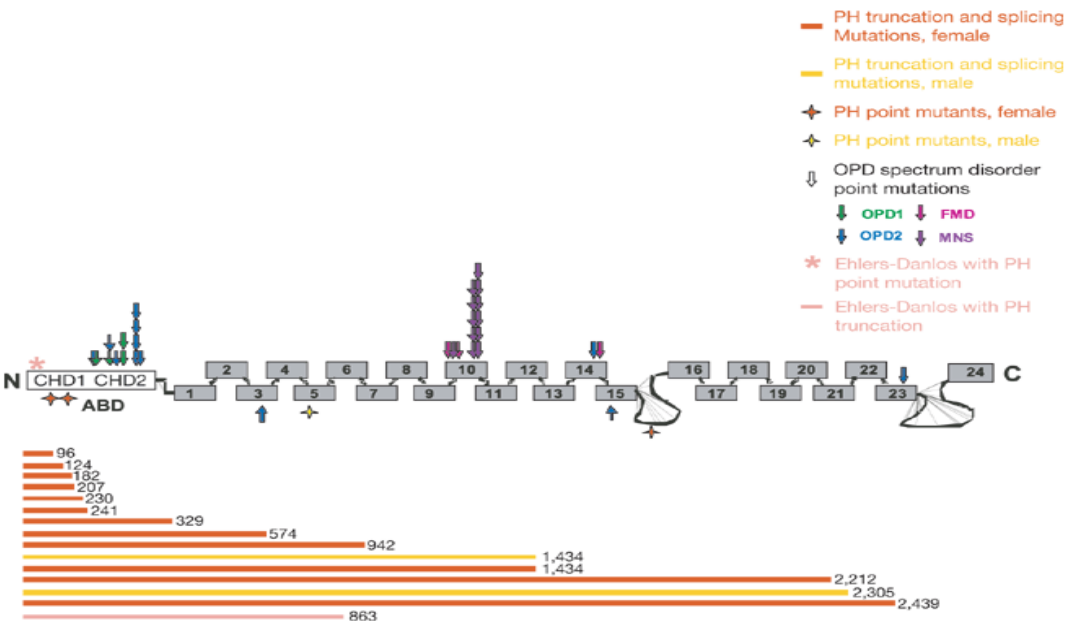

Figure 8: hs*flnA* mutations identified in various human genetic disorders. Inset legends are self explanatory; arrows, astericks and diamonds refer to the sites of point mutations in the ABD, FRs and hinge regions. Dark and light colored lines below the monomer chain indicate the truncation or splicing mutations in diseased females and males, respectively [Feng & Walsh, 2004].

An EDS (Ehlers-Danlos syndrome) variant of PVNH is found to be associated with single base deletions (ΔG2762 or ΔG4147), or an A39G change due to C116 point mutation [Sheen *et al*, 2005]. Another set of 46 mutations which cluster within a few regions of *flnA* ie. the CH2 domain, repeats 3, 9-10, 14-15 and 22-23, some of which are highly recurrent, give rise to the otopalatodigital (OPD) spectrum of disorders such as cleft palate, facial malformation and bone dysplasia. The mutations are invariably missense mutations which probably cause defective cell migration during embryonal development in these altered or gain-of-function phenotypes [Robertson *et al*, 2003].

Mutations in *flnB* cause diseases characterized by abnormal vertebral segmentation, joint formation and skeletogenesis. Mutations in the ABD cause Larsons syndrome, atelosteogenesis I and III, while point mutations in filamin repeats 5, 6, 14, 15, 20 and 22 are responsible for spondylocarpotarsal syndrome (in Feng & Walsh, 2004).

A W2710X mutation in filamin C domain 24 that prevents translation of the C-terminal 16 amino acids of the dimerisation domain underlies a human myofibrillar myopathy [Vorgerd *et al*, 2005]. This myopathy is associated with

myofibrillar degeneration and cytoplasmic aggregation of proteins normally associated with the sarcomere z-disk and subsequent studies [Lowe *et al*, 2006] revealed that the truncated FLNc24 domain is less stable than wild type and aggregates instead of forming dimers.

1.8. DIMERIZATION

Evidence provided by electron microscopy, gel filtration, ultracentrifugation, chemical cross-linking and NMR/X-ray diffraction structural studies indicate that filamins of *H. sapiens*, *D. discoideum*, *Drosophila melanogaster*, *E. histolytica* and vertebrates exist as dimers whose polypeptide chains associate at their C-termini [Noegel *et al*, 1989; Vargas *et al*, 1996; McCoy *et al*, 1999; Popowicz *et al*, 2004; Pudas 2006]. They form homodimers, although heterodimerization among isoforms is also possible. For instance, Sheen *et al* [2002] showed that hsFLN isoforms A and B which occur as homodimers may also form heterodimers *in vivo*, although this is disputed by later *in vitro* studies [Himmel *et al*, 2003]. In the latter studies, however, hsFLNb and hsFLNc were observed to form heterodimers.

1.8.1. STRUCTURAL BASIS OF ddFLN DIMERIZATION

Dimerization is crucial for actin crosslinking functions of filamins [Davies *et al*, 1980] and requires the most C-terminal Ig domain (Ig24) [Hock *et al*, 1990 in Himmel *et al*, 2003]. A dimerized filamin with its ABDs at both N-termini is perfectly suited to link actin filaments into bundles or meshworks.

The structural basis of filamin dimerization was first resolved in *Dictyostelium discoideum* filamin ABP-120, which forms antiparallel rod-like dimers *in vitro*, measuring 35 nm [Condeelis *et al*, 1984]; this suggests that dimerization occurs through complete or partial overlapping of repeats between two dimer chains [Fucini *et al*, 1997]. When ddFLN repeat domains Ig5 and Ig6 were expressed as a single construct, they dimerized in solution and gave an apparent molecular weight of of 62 kDa by gel filtration, as compared to the calculated molecular weight of 23 kDa [McCoy *et al*, 1999]. This implies that Ig-5 and -6 domains alone are sufficient for ddFLN dimerisation.

The crystal structure of the dimerized construct obtained at 2.2Å resolution shows that dimerization is mediated primarily by Ig6 and the two chains overlap only at domain 6 where two copies of Ig6 lie antiparallel to each other; Ig5 as well as the linker between Ig5 and Ig6 (FR-6) also contribute to the interaction. The main dimerizing domain Ig6 has a β-barrel structure formed by an upper β-sheet containing strands BED and a lower β-sheet containing strands CFF′GH (figure 9).

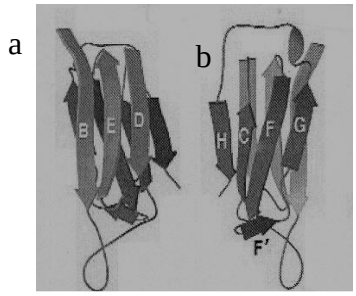

Figure 9: Topology of ddFLN rod domain 6. Ribbon representation showing (a) upper β-sheet containing antiparallel strands BED (b) lower β-sheet containing strands HCFF′G [McCoy *et al*, 1999].

It is distinct from the structurally homologous Ig5 domain, which closely resembles the I set of Ig folds and has strands ABED and A′CC′FF′G, in the corresponding β-sheets. FR-6 shows no sequence homology to the other five rod repeats, lacks strand A and contains an additional H strand at the C-terminus; in addition, 12 residues at the N-terminus and up to the middle of the β strand are missing in the B strand [McCoy *et al*, 1999].

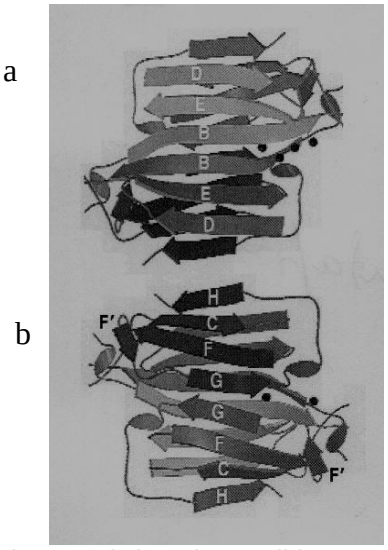

Figure 10: Dimerization of ddFLN rod domain 6. Ribbon representation showing the double extension of the β-sheet in the upper (a) and lower (b) sheets. Four water molecules that mediate β-sheet extension on one side of the interface are shown as spheres [McCoy *et al*, 1999].

FR-6 contributes 69% of the accessible area buried in the dimer interface, the linker 22% and domain Ig5 9%. Ig6 of adjacent chains interacts through β–sheet extension by interactions between their respective B and G strands which form a large β–sandwich of six strands on one side and eight on the other. The interface has 39 hydrogen bonds, which also includes water mediated hydrogen bond interactions along one side of the B strands (figure 10).

1.8.2. STRUCTURAL BASIS OF hsFLN DIMERIZATION

In the case of hsFLN, it was suggested [Gorlin *et al*, 1990] that repeat 24 (Ig24, FR24) is at least required for dimerization. Sheen *et al* [2002] later proposed that Ig24 alone is not sufficient for dimerization. However, subsequent work [Himmel *et al*, 2003] using all three hsFLN isoforms demonstrated that the Ig24 domain is necessary and sufficient for dimerization *in vitro* and that H2, which immediately precedes Ig24, significantly enhanced dimerization.

hsFLN dimerization differs from that of ddFLN in several ways. The rod regions of hsFLN homodimer chains lie parallel to each other [Gorlin *et* al, 1990] unlike the antiparallel arrangement of peptide chains in *D. discoideum*. Secondly, the structure of human filamin Ig24 is distinct from that of ddFLN dimerizing domain Ig6; the two structures show an rms (root mean square) deviation of 8.3Å of superimposed 69 Cα atoms and hsFLNa Ig24 structure more closely resembles that of the non-dimerizing ddFLN domain 5 (rms deviation of 1.6Å of the same Cα atoms). The length of the interacting strands are shorter in hsFLN and there are no buried linker at the N-terminus of Ig24 [Popowicz *et al*, 2006]. Thus it is not surprising that the dimerization mechanism in hsFLN is different to that of ddFLN.

Human filamin C Ig24 domain is similar to other hsFLNc rod repeats. The crystal structure of hsFLNc Ig24 obtained at 1.4Å resolution [Pudas *et* al, 2005] shows that the domain forms the expected Ig-like β-sandwich fold consisting of a 4-stranded antiparallel β sheet containing strands ABED and a 3-stranded sheet consisting of strands CFG. The putative dimerization interface involves strands C and D of adjacent Ig24 domains (figure 11).

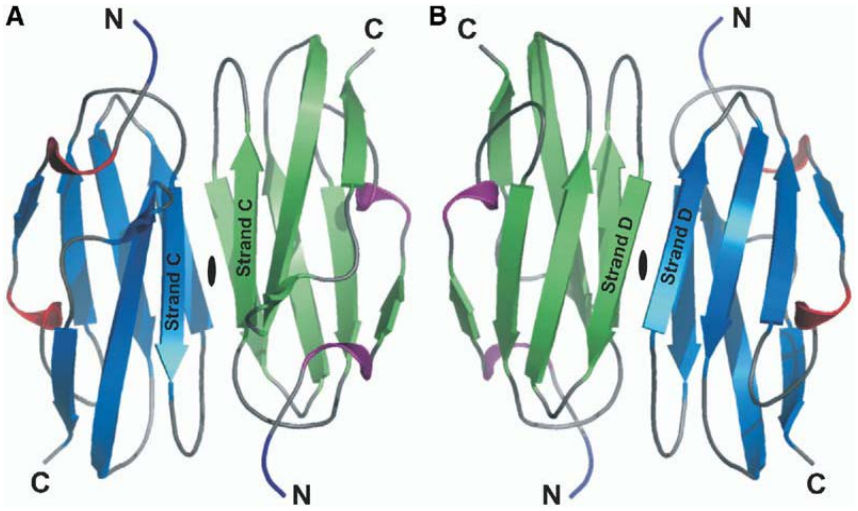

Figure 11: Crystallographic Dimer of Filamin C Domain 24 Views rotated by 180° about the vertical. Strands of the monomers are colored blue and green. Strands C and D involved in dimer contact are indicated. The extended antiparallel β sheet that forms on dimerization is shown in (B). The two-fold axis is marked. The extra residues at the N terminus that are derived from the vector are indicated in blue.

The interface is much smaller than in ddFLN. The buried surface area is 1109Å^2 (compared to 4501Å^2 in ddFLN) and takes up 19% of the molecular surface. Forty five percent of the atoms comprising the interface are polar. The D-strands of the two dimer chains are connected by six hydrogen bonds so that the 4-stranded β-sheets of each chain form a 8-stranded antiparallel β-sheet. Point mutation (M2669D in strand C) or complete strand replacement (Strand C with strand C of hsFLNa Ig20 which does not dimerise in solution) in the dimer interface disrupts the dimer and the recombinant repeat 24 has a dissociation constant in the micro-molar range [Pudas *et al*, 2005]. hsFLNa uses the same structural region of the repeat domain (the face comprising β-sheets C and D) of FR 21 to bind platelet adhesion protein GPIbα [Nakamura *et al*, 2006] and integrin β7 [Kiema *et al*, 2006]. On the basis of these structural similarities, the C-D interface is proposed as a general ligand binding area where specificity is determined by hydrogen bonds and hydrophobic interactions at the interface. The signature sequence of the dimerization interface is shared by the C-terminal domains of all vertebrates which suggests that all vertebrate filamins use the same dimerization mechanism.

1.9. AIM OF STUDY

A patient diagnosed with a mild form of PVNH displayed a G2593E mutation in the C-terminal FR 24 of hsFLNa (hsFLNa24). In this study we investigated the effect of this mutation on the dimerization functionality of Ig24 in order to determine the molecular basis of the pathology.

CHAPTER II - MATERIALS AND METHODS

2.1. PRIMERS FOR hsFLNa REPEAT DOMAIN 24 AMPLIFICATION

Restriction sites were introduced into the 5' overhangs of primers used to amplify hsFLNa24 for cloning. The nucleotide region between positions 7663 and 7944 of hsFLNa, which corresponds to amino acids 2555-2647 constituting hsFLNa repeat domain 24 (Ig24=hsFLNa24) was scanned with DNA Strider® for restriction sites present in the multiple cloning site (MCS) of cloning vector pProEX HTb (figure 11).

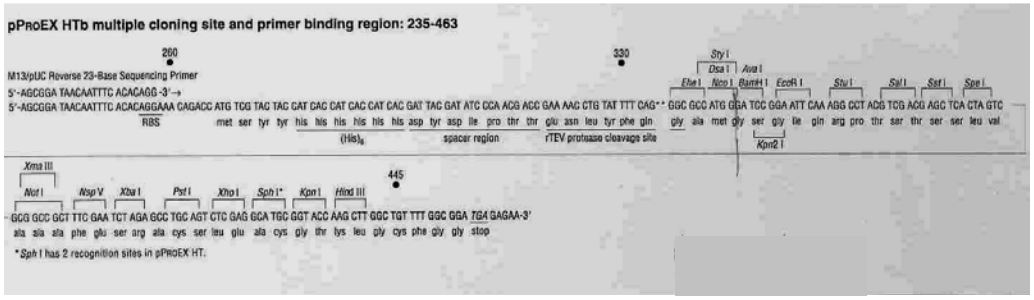

Figure 11: MCS of pProEX HTb containing *Bam*HI and *Xho*I restriction sites. The MCS is preceded by a 5´ (N-terminal) translation start site (ATG) followed by a His6-tag separated from the cleavage sequence for the tobacco etch virus protease (Tev) by a spacer region.

Restriction sites for *Bam*HI and *Xho*I were not present and were therefore included in the 5´ overhangs of primers used to amplify hsFLNa24 for cloning in pProEX HTb. Primer 3® was used to design wt forward (wt-F) primer 5´ CG*GGATCC*CGCTGACGCCAGCAAGGTG 3´ and wt reverse (wt-R) primer 5´ cc*gctcgag*tcagggcaccacaacgcg 3´ to amplify wildtype (wt) hsFLNa24.

2.2. AMPLIFICATION OF wt hsFLNa24 INSERT

cDNA clone pFLNa, containing the full-length hsFLNa was amplified with wt-F and wt-R primers using high-fidelity, recombinant *Pwo* DNA polymerase which has 3´-5´ exonuclease proof-reading activity providing 18-fold higher fidelity than *Taq* polymerase. Amplification was performed in a 50 µl reaction volume containing 1x PCR buffer containing $MgSO_4$, 200 µM dNTP, 5 pM wt-F & wt-R primers, 1 unit (0.5 µl) *Pwo* polymerase and 2 ng template cDNA. Annealing temperature was empirically derived by the algorithm of Rychlik *et al* [1991] and final PCR was conducted with an initial

denaturation at 95ºC x 2 minutes followed by 30 cycles of 94ºC x 30 seconds, 58ºC x 15 seconds, 72ºC x 45 seconds. The longer extension time was allowed to compensate for the proof reading activity of *Pwo* polymerase. Aliquots of PCRs were electrophoresed in 1.5% TAE gels to ascertain reactions which gave the expected amplicon size.

2.3. CLONING VECTOR PREPARATION

E. coli cells containing pProEX HTb were incubated on Luria-Bertani broth ampicillin (LB. Ap_{50}) plates for 16 hours at 37ºC. Single colonies were transferred to 5 ml LB.Ap_{50} broth and cultured overnight with shaking. Plasmid DNA was purified with Roche HiPure kit™ and eluted in 100 µl elution buffer. Plasmid isolates were digested with *Bam*HI and *Xho*I. Aliquots of the digests were electrophoresed in 1.5% TAE agarose gels and the size of the linearized plasmid confirmed. Undigested plasmid was used as control. The digests were purified in QiaPrep spin columns™ (Qiagen) and concentration of the linearized cloning vector determined using Nanodrop sphectrophotometer.

2.4. CLONING OF wt hsFLNA24 FRAGMENTS

Purified blunt-end wt hsFLNa24 amplicons were digested with *Bam*HI and *Xho*I and the digested fragments with cohesive ends were purified using Qiaquick™ spin columns. Inserts were quantified and ligation was conducted in a 10 µl volume containing 1x ligation buffer, 10 ng linearized pProEX HTb, 40 ng wt FLNa24 insert, 0.2 mM ATP and 1 unit ligase at 4ºC for 24 hours.

2.5. TRANSFORMATION

The ligation mixtures were transformed into chemically competent Top 10 dH5α® cells by heat shock method. Five microliters of the ligation mixtures was mixed with 50 µl competent cells, incubated at 42ºC for 2 minutes, kept on ice for 2 minutes, plated on LBAp_{50} media and incubated at 37ºC for 18 hours. Putative transformed colonies were picked and cultured in LBAp_{50} broth for 18 hours. Plasmids were purified from broth cultures by alkaline lysis method, aliquots were digested with *Bam*HI and *Xho*I and electrophoresed in 1.5% TAE gels. Clones which gave inserts of the expected size were

deemed positive and sequenced and the recombinant wt FLNa24 clone-1 was selected for further studies.

2.6. SITE DIRECTED MUTAGENESIS AND PRODUCTION OF MUTANT mt FLNa24 CLONES

Complementary primers 5´ CAAATGCTGCTGGTGG***A***GGTTCATGGCCCAAGGAC 3´ and 5´ GTCCTTGGGCCATGAACC***T***CCACCAGCAGCATGTTG 3´ containing the G2593E mutation (GGG→GAG) in the middle were designed with a melting temperature of >78ºC. Mutagenic PCR was conducted in a reaction volume of 50 µl containing 1x buffer, 150 ng forward and reverse primers, 200 µM dNTP and 3 units high fidelity *Pfu* Turbo DNA polymerase. 10 ng of wt clone-1 plasmid was used as template. Following an initial denaturation at 95ºC x 30 seconds, 16 cycles of PCR were conducted at 95ºC x 30 seconds, 55ºC x 60 seconds, 68ºC x 6minute. PCR was digested with 1 µl *Dpn*I, which cleaves the template plasmid at its methylated sites, but not the PCR product. PCR products were directly transformed into Top 10 dH5α^{TM} cells and putative transformed colonies on LBAp$_{50}$ plates were grown in LBAp$_{50}$ medium. Plasmids were purified and positive clones identified by double digestion with *Bam*HI/*Xho*I and mt FLNa24 clone-1 was sequenced to confirm the G2593E mutation.

2.7. hsFLNa PROTEIN EXPRESSION

His-tagged wildtype (wt FLNA24) and mutant (mt FLNA24) proteins were over-expressed in *E. coli* BL21 (DE3); wt FLNa24 clone-1 (see section 5) and mt FLNa24 clone-1 (see section 6) plasmids were transformed separately into 50 µl chemically competent cells. Single colonies of positive transformants were grown in 500 ml LB broth supplemented with 100 µg ml^{-1} ampicillin with shaking and induced with 0.2 mM isopropyl-β-D-thiogalactoside (IPTG) when A_{600} reached 0.6-0.8. The cultures were transferred to 25°C and grown for another 16 hours in an orbital shaker, centrifuged at 5000x g for 20' in GS3 tubes and the cell pellets stored at -20°C.

2.8. PROTEIN PURIFICATION AND IMMOBILIZED METAL ION AFFINITY CHROMATOGRAPHY (IMAC)

All procedures were performed at ≤4°C and frothing was minimized since many proteins are denatured by contact with the air-water interface. Cell pellets were washed

and resuspended by vortexing in 10 ml lysis buffer (20 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole) containing a ¼ tablet of EDTA-free protease inhibitor cocktail (Roche™). The cell suspension was sheared in a French Press twice at 5,000 psi and the lysate sonicated in three bursts of 15 seconds each to shear nucleic acids. Lysate was fractionated by centrifugation at 30,000 x g for 20 minutes at 4°C to remove cell debris. The soluble fraction was clarified using 22µm pore size filters prior to preparative purification by affinity chromatography on an automated AKTA system.

All buffers and reagents used in IMAC were filtered through 0.22 or 0.45µm filters. A 5 ml Ni^{2+}-NTA HisTrap™ column (GE Healthcare) packed with highly cross-linked sepharose resin with an immobilized chelating group and pre-charged with Ni^{2+} was equilibrated with 5 column volumes (CV) of binding buffer (20 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole). The clear cell lysate supernatant was injected into the column and washed with binding buffer till absorbance reached a steady baseline. Washing was continued with a stepwise imidazole gradient in the binding buffer consisting of 5 CV of 100, 200 and 500 mM imidazole, respectively, at each step. The flow through as well as wash eluting from the column were collected in 0.5 ml fractions and fraction purity was confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Peak fractions containing the target protein were pooled and processed.

2.9. wt- AND mt-hsFLNa PROTEIN ISOLATION

Protein profiles of IMAC fractions were studied by SDS-PAGE in stacked gels consisting of 15% total monomer resolving gel and 4% total monomer stacking gel. 6 µl samples were boiled with loading buffer for 3 minutes and electrophoresed at 100V for 60 minutes in a BioRad minigel apparatus. Pre-stained precision plus protein™ standards were run alongside samples to determine the size of protein fractions. Protein bands were visualised by Coomassie blue staining and images captured by a Biorad digitizer.

Fractions enriched for the protein were pooled and protein concentration determined by the Beer-Lambert law $A = \varepsilon cl$ where A = absorbance at 280 nm wavelength in a 1 cm cell, ε = molar extinction coefficient of protein, c = molar

concentration and *l* = length of light path in cm. The estimated extinction coefficient of 12950$M^{-1}cm^{-1}$ (1AU=0.78mg/ml) was calculated using the ProtParam tool from the ExPASy Server (Gasteiger *et al*, 2003).

The $(His)_6$ tag was cleaved from the proteins using a recombinant tobacco etch virus protease (rTEV), which itself carries a $(His)_6$ tag. Buffer exchange was performed prior to protease digestion by dialysing the His-tagged proteins, wt- and mt $(His)_6$-FLNa24, against three changes of TEV digestion buffer consisting of 50 mM Tris-HCl (pH 8.0), 20 mM NaCl, 0.5mM EDTA and 1mM DTT at 4°C. The dialysate was concentrated in Vivaspin 20™ columns (10 kDa cutoff) at 4000x g for 15' and incubated with TEV protease (10 µg ml^{-1}) at a ratio of 1:100 (w/w) overnight at 10°C. The cleaved hsFLNa proteins were recovered from the digest mix by buffer exchange with IMAC binding buffer followed by eluting the mixture through a fresh Ni^{2+} HisTrap column; the flow through contains the cleaved protein while uncleaved $(His)_6$-FLNa24, residual $(His)_6$-TEV protease and the cleaved $(His)_6$ tags remain bound to the Ni^{2+} column. Aggregated material was removed by centrifugation at 174,000 x g for 15 minutes.

The isolated proteins were further purified and sized by gel filtration, which fractionates molecules based on the diffusion of molecules into pores of the gel resin: as a result, molecules elute from the column in order of decreasing molecular weight. Molecular weight is determined by comparing the V_E/V_O ratio (V_O = void volume, V_E = elution volume) of a novel protein with those of standards of known molecular weight. A Superdex 75 HR 10/30 size exclusion column was equilibrated overnight with 1.5 CV of size exclusion buffer containing 20 mM Tris, pH 8, 50 mM NaCl and 5 mM DTT. Flow rate of the AKTA machine was set to 0.5 ml min^{-1} and the column was calibrated by injecting 100 µl of a mixture containing hsFLNB ABD (28 kDa) and BioRad™ filtration standards cytochrome C (12.4 kDa), ovalbumin (43 kDa), BSA (66 kDa), and dextran (2000 kDa); blue dextran was used to determine void volume, V_O. The linear equation $y = -5.029x + 5.338$ ($R^2 = 0.999$) derived from the calibration curve was used to estimate molecular weights of FLNa proteins. Protein samples for purification were injected as 100 µl volumes into a 200 µl loop.

2.10. ETHYLENE GLYCOL BIS(SUCCINIMIDO SUCCINATE) CROSSLINKING ASSAY

Crosslinkers are bifunctional molecules containing two or more reactive groups at either end of their spacer arm, which may be identical or different and capable of chemically attaching to functional groups residing on proteins, hsFLNa24, as in our case. The joining of two or more molecules by a covalent bond during crosslinking (conjugation) depends on the availability of specific functional groups found on the exposed surface of proteins such as primary amines ($-NH_2$) found at the N-terminus of polypeptides and in the side chain of lysine and carboxyls (-COOH) found at the C-terminus of polypeptides and the side chains of aspartic and glutamic acids and also sulfhydryls (-SH) groups. Conditions for crosslinking of filamins were based on the protocol of Himmel *et al* [2003]; 4µM FLNa24 final concentration was added to HGNED crosslinking buffer containing 25 mM HEPES, pH 8, 100 mM NaCl, 0.6 M KCl, 1 mM DTT, 0.2 mM EDTA, 0.05%NP40 and 10% glycerol and incubated on ice for 1 hour. Ethylene glycol bis(succinimido succinate). EGS was added to a final concentration of 1.3 mM and incubation was continued at 37°C for 20 minutes and the reaction stopped by TCA precipitation (10% final concentration) on ice for 30 minutes. After centrifugation at 14000x g the pellets were washed with 200 µl ice-cold acetone which was then removed by drying in air. The protein pellets were resuspended in SDS-PAGE loading buffer and electrophoresed in PAGE gels.

2.11. WESTERN BLOTTING

His-tagged FLNa24 proteins were electrophoresed in SDS-PAGE gels and electro-blotted onto nitrocellulose membrane overnight in 20 mM Tris, pH 7.6, 25 mM glycine, 7.6 mM SDS and 20% methanol. Transfer was confirmed by Ponceau staining and the membrane was blocked with 5% skim milk in blocking buffer containing 20 mM Tris, pH 7.5, 150 mM NaCl and 0.05% Tween20. The blot was incubated with 1:5000 diluted anti-6xHis mouse monoclonal antibodies (Sigma®) at room temperature for 1 hour, washed in TBS-Tween and incubated with 1:6000 diluted anti-mouse secondary antibody (Sigma®). Hybridizing bands were visualized by chemiluminiscence.

2.12. PROTEIN MELTING TEMPERATURE T_m

The thermal stability of wt- and mt FLNa24 proteins was compared by melt curve analysis using Corbett® Rotogene PCR machine. The assay was conducted in 20 µl melting buffer containing 20 mM Tris, pH 8, 50 mM NaCl and 5 mM DTT at a protein concentration of 1 mg ml^{-1}; 0.1µl SYPRO Orange™, which is non-fluorescent in a hydrophilic environment and fluorescent in a hydrophobic environment was used as the reporter dye. Identical samples containing no dye or no protein were used as control. The sample tubes were placed in a 36-well rotor and thermal profiles monitored at excitation and emission wavelengths of 470 and 510 nm, respectively. Assays were performed over a temperature range of 303-368 Kelvin (K) with a temperature ramp rate of 2 K min^{-1}. T_m of the proteins was calculated from the first derivative of the unfolding curves.

2.13. PROTEIN CRYSTALLIZATION AND DATA COLLECTION

X-ray crystallographic determination of protein crystal structure requires well-ordered crystals suitable for X-ray diffraction, usually single crystals with a high degree of homogeneity, long range 3-dimensional order (periodicity) and a size of 10-200 µm. Optimal crystal formation is dependant on the crystallization space, which refers to several variables such as the nature and concentration of the protein, crystallization method (vapour diffusion, free interface diffusion, batch under oil, etc) and physical/chemical parameters (temperature, pH, buffer,additives, precipitants, etc). Preliminary sparse matrix crystal screening was performed by mixing 1 µl of the protein at a concentration of 7 mg/ml with 1 µl mother liquor in sitting drops using 96-well Crystalquick™ plates (Greiner Bio-one®) on ice, containing 100 µl well volumes of Crystal Screen 1™ and Crystal Screen 2 ™ (Hampton Research®) precipitant solutions. Crystals were grown at room temperature and plates were monitored for crystallization using a stereo-microscope after 24 hours and thereafter daily for a week and then once a week. Optimization of crystallisation conditions proceeded by hanging-drop vapour diffusion at 4°C with the crystallisation mother liquor solution. The protein was mixed in a 1:1 ratio with the mother liquor on siliconised glass coverslips, which were then inverted on VDX plate wells containing 500 µl of mother liquor and the edges sealed

with Vaseline. The drop, which has a lower precipitant concentration equilibrates with the well solution through evaporation of water and during this process the concentrations of protein and the precipitant in the drop occurs slowly and gradually, favouring crystallization over precipitation. Crystals were picked with cryoloop™ crystal wands and stored in liquid nitrogen,

Protein crystals contain 40-60% solvent by weight and are therefore, fragile and sensitive to dehydration. Therefore, the crystal in a nylon loop was briefly soaked in cryoprotectant solution containing mother liquor and 20% glycerol and flash frozen in liquid nitrogen before mounting on the goniometer and gradually rotated to centre it vertically in the X-ray beam. Diffraction data was collected at a resolution of 2.5 Å in a RIGAKU MicroMax007 diffractometer with a RAXIS-IV detector system at a cryogenic temperature of 120 K in an X-ray beam diameter of 10-400 µm and wavelength 1.5418. The crystal-to-camera distance was set at 180 mm, oscillation 1º and exposure 10' per frame. Data collection was stopped when MOSFLM and SCALA analyses indicated data was sufficiently complete (85%).

2.14. PROTEIN MODELLING

The CCP4 (collaborative computational project, number 4, 1994) suite of programs was used to derive the initial electron density map. The exact crystal orientation with respect to the X-ray beam, the symmetry space group and unit cell size are essential for diffraction data analyses. Constraints on cell parameters determine the way in which regular 3-dimensional crystals of the seven crystals system belonging to 14 Bravais lattices are constructed. Dimensions of a unit cell, which is the smallest repeating unit that can generate the crystal with only translational operations, are described by the lengths of the three axes, a, b, and c, and the three interaxial angles, α, β, and γ. These and other parameters were determined by autoindexing of reflections.

MOSFLM was used to index the first few frames before auto-indexing and integrating all diffraction frames by assigning each reflection a spot intensity and fitting the predicted image with the (observed) diffracted image. The X and Y beam coordinates were set as 150.1 and 149.7, preliminary space group to P1 with lowest

penalty (as suggested by the program), reflection spot radius to 0.3 mm, mosaicity (angular measure of the degree of long range order of unit cell) 0.4º and backstop radius 4. After collecting the first few images two images separated by nearly 90º in phi were indexed and integrated, since the unit cell derived from only one image may be inaccurate. Iterative cycles of autoindexing (to check crystal quality), strategy (to determine phi range for completeness), cell parameter refinement and integration were performed by adjusting the various parameters; changes in spot selection and values of spot radius, spot separation, mosaicity and Min I/sigma (I) were made during these iterative cycles. The starting phi angle and phi range necessary to obtain a complete data set was calculated from the preliminary crystal symmetry and orientation; the crystal orientation and unit cell size were refined and a preliminary space group was assigned.

The software first determines the Laue class and then possible systematic absences to identify the appropriate cell centring conditions, glide planes, and screw axes, if any. Iterative indexing was performed till an optimal solution was achieved (good fit between Bragg spots and predictions with rms positional residual of 0.12 mm) at a resolution of 2.2 Å. Following this, cell parameters were refined with 2 segments of images and the complete data set was integrated through several cycles of integration and scaling (including cell refinement, mosaicity setting) culminating in accurate prediction of reflection spots and cataloguing of their intensities.

The refined optimal cell statistics derived with MOSFLM were substituted in SCALA to scale and merge equivalent intensity data. The data of individual films were scaled and merged using interactive cycles of SCALA, which uses Bayesian statistics and Wilson plots to derive structure factors and B-factors (atomic temperature factors). Structure factor amplitude calculated from the merged intensity data (proportional to the square root of the diffraction intensity) gives the arrangement of atoms in the unit cell. The calculation of electron density map requires knowledge of both the amplitude and phase of the diffracted beams. The phase is determined directly by multiple isomorphous replacement (MIR), or indirectly by molecular replacement [McCoy *et al*, 2007]. The latter method approximates the scattering of an unknown molecule from the calculated scattering of a search model of known atomic structure, which shares high

sequence similarity with the query molecule. Molecular replacement studies of hsFLNa24 were performed with PHASER [McCoy *et al*, 2007; CCP4, 1994] using hsFLNc24 [Pudas *et al*, 2005] as search model. Extra residues in FLNc24 were removed and mismatches changed to Alanine, barring a few exceptions. R_{factor} and the coefficient of correlation values were used to determine the number of molecules in the asymmetric unit.

The model was rebuilt with COOT [Emsley & Cowtan, 2004] by rigid-body refinement of the electron density maps obtained from PHASER. Further refinements including mutating mismatched residues, manipulating side chain conformation was performed with PHENIX and REFMAC [Afonine *et al*, 2005; CCP4, 1994] to a final R_{test} of 20% and R_{free} of 25%. A Ramachandran diagram was invoked to assess the stereochemical quality of the atomic model. Pymol [DeLano, 2002] was used to generate structural figures. *In silico* mutation of G2593 to E was performed using the mutate feature of Accelrys® and the structural homology between mutant and wildtype structures was estimated using the align feature of Pymol.

CHAPTER III – RESULTS

3.1. VECTOR CONSTRUCTION

3.1.1. pWT-FLNA24 VECTOR PREPARATION

The full length hsFLNa CDS contained in pFLNa (kind gift of Professor Stephen Robertson, University of Otago) was amplified with primers wt-F and R. Amplification resulted in a PCR fragment ~300 bp long (figure 12).

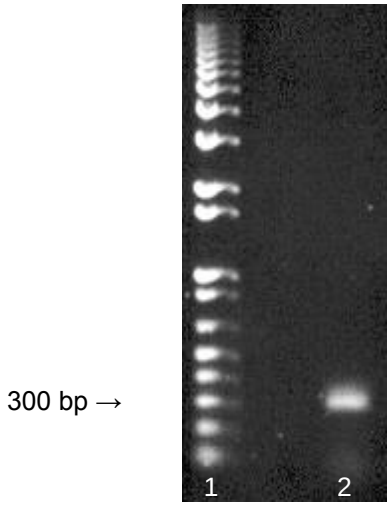

FIGURE 12: Agarose gel image of PCR. Lane 1 – 1 kb Plus™ ladder, Lane 2 – PCR reaction. PCR consisted of an initial denaturation at 95ºC x 2 minutes followed by 30 cycles of 94ºC x 30 seconds, 58ºC x 15 seconds, 72ºC x 45 seconds. A 4 µl aliquot of the 50 µl PCR reaction in gel loading buffer containing bromphenol blue dye was electrophoresed in 1.5% TAE agarose gel at a constant voltage of 100 V and stained in ethidium bromide and imaged under UV.

The fragment consists of 282 bp constituting hsFLNa24 coding sequence, flanked by 8 and 9 bp primer overhangs containing *Bam*HI and *Xho*I restriction sites, respectively. The purified, double-digested amplicon was cloned into complementary sites in the linearized expression vector pProEX HTb at a molar ratio of 60:1 and transformed into chemically competent *E. coli* DH5-α cells. Plasmids were prepared from putative transformants and digested with *Bam*HI and *Xho*I (figure 13). From among the plasmids which contained the insert, recombinant - pWt FLNa24 clone-1 was selected. Sequencing confirmed the presence of codon GGG at the expected position corresponding to glycine residue 2593 in wild type hsFLNa.

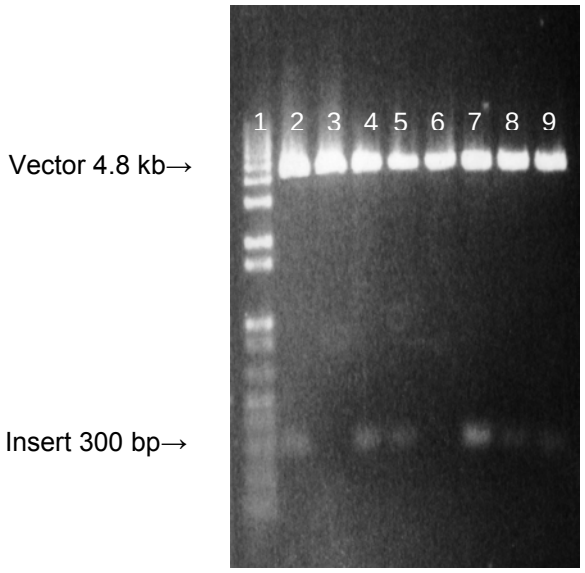

Figure 13: Agarose gel image of plasmid digests. Lane 1 – 1 kb Plus™ ladder, 2 – 9 Plasmid digests. Lanes 3 and 6 represent empty vectors. 2 µl aliquots of the digests were boiled in gel loading buffer containing bromphenol blue dye and electrophoresed in 1.5% TAE agarose gel at a constant voltage of 100 V, stained in ethidium bromide and imaged under UV.

3.1.2. pWt-FLNa24 VECTOR VALIDATION

An assay was performed to determine the ability of the recombinant plasmid to direct hsFLNa protein expression in the soluble fraction of the host bacterium. Ideally, the recombinant proteins should partition preferentially into the soluble cell fraction. However, when recombinant proteins are over-expressed in bacteria the foreign proteins may be sequestered as insoluble, denatured inclusion bodies. In such cases it becomes necessary to subject the inclusion bodies to extensive renaturation, or switch to a different expression vector altogether to produce soluble proteins.

Briefly, the recombinant pWt hsFLNa construct was expressed in *E. coli* BL21 (DE3) by IPTG induction in 10 ml LB.Ap_{100}. The cell pellet obtained by centrifugation at 4000 rpm for 10 minutes was disrupted by sonication and the soluble and insoluble cell fractions were fractionated by high speed centrifugation at 13000 rpm for 20 minutes. The insoluble fraction was diluted 100x as it was too viscous to electrophorese. One µl of the diluted pellet and 4 µl of the supernatant were examined by SDS-PAGE. A protein

of apparent molecular weight (M_w) ~13 kDa, as determined by gel calibration, was highly expressed in the soluble fraction (figure 14).

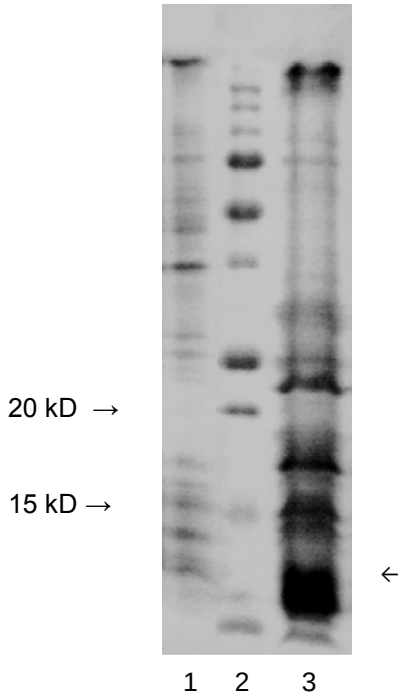

Figure 14: SDS-PAGE gel image of cell extracts from IPTG induced culture. Lane 1 – Resuspended pellet fraction, Lane 2 – Size standards, Lane 3 – Supernatant fraction. One µl 500x diluted pellet and 4 µl supernatant fractions of induced cells were boiled in gel loading buffer and electrophoresed in 15% PA stacked gels at 200 volts. Protein bands were visualized by Coomassie blue staining.

As Mw of the protein is the same as the estimated molecular weight of His-tagged hsFLNa24 (13.2 kDa) calculated from its amino acid composition (section 2.9) and since the protein was not present in extracts of non-induced cells (data not shown), it was presumed that the 13 kDa protein is indeed the recombinant hsFLNa24.

3.1.3. pMt-FLNa24 VECTOR PREPARATION

Using site directed mutagenesis, mutant FLNa24 (pMt FLNa24) plasmids containing the G2593E mutation were produced from recombinant plasmid pWt FLNa24 clone-1. Following restriction enzyme analyses, pMt FLNa24 clone-1 was selected from among the positive clones and the G2593E mutation confirmed by sequencing (data not shown).

3.2. CONDITIONS FOR SCALED-UP PROTEIN EXPRESSION

The yield of soluble hsFLNa24 protein is influenced by the temperature at which transformed bacteria is cultured. The optimal temperature regime for maximum expression of (soluble) recombinant wt hsFLNa24 was determined by inoculating 500 ml $LBAp_{100}$ broth with single colonies of transformed *E. coli* BL21 (DE3) (section 2.7) and growing them under different conditions. Two replicates were grown at 37ºC and after induction, one replicate continued to be cultured at 37ºC for 3 hours, while the other was transferred to 25ºC and culture continued for 16 hours. A third replicate was grown continuously at 25ºC before and after IPTG induction. The cells were spun down, aliquots were sonicated and centrifuged and the pellet and supernatant fractions were examined by PAGE (figure 15).

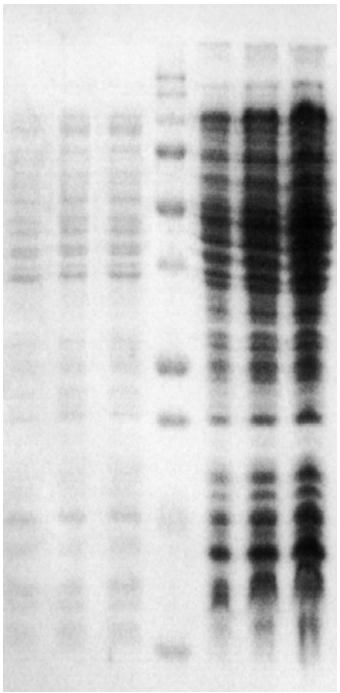

Figure 15: SDS-PAGE gel image of cell extracts of cultures grown at different temperature regimes. Lanes 1–3 Pellet fraction, Lanes 5-7 Supernatant fraction. Lanes 1 & 5 - Continuous 25ºC culture, Lanes 2 & 6 – continuous 37ºC culture, Lanes 3 & 7- 37ºC followed by 25ºC culture, Lane 4 – Size standards. One µl each of supernatant and 1000x diluted pellet of induced cells were boiled in gel loading buffer and electrophoresed in 12% PA stacked gels at 200 volts. Protein bands were visualized by Coomassie blue staining.

Expression of soluble wt hsFLNa24 was marginally higher when cells were grown at 37ºC pre-induction and 25ºC post induction. These conditions were used for scaled-up protein expression.

3.3. AFFINITY CHROMATOGRAPHY

An aliquot of soluble His-tagged wt hsFLNa24 was used in a small-scale binding assay with Ni^{2+}-NTA resin to study the recombinant protein's Ni^{2+} metal affinity characteristics. 80 µl of resin was washed in a PCR tube with water and equilibrated with 1 ml lysis buffer (20 mM Tris, pH 8, 500 mM NaCl, 10 mM imidazole). Following centrifugation, the supernatant was removed and the resin was equilibrated with hsFLNa24 diluted in 300 µl lysis buffer. The supernatant was aspirated and the resin washed thrice with lysis buffer to remove unbound protein. The immobilized protein was sequentially equilibrated with 2x1 ml aliquots of lysis buffer containing 100, 300 and 500 mM imidazole, respectively, followed by centrifugation and aspiration of the supernatant at each step. Supernatants were concentrated and examined for protein released from the resin. SDS-PAGE showed that hsFLNa24 was predominantly eluted at 500 mM imidazole concentration (data not shown).

In preliminary IMAC runs, wt hsFLNa24 was eluted at 500 mM imidazole concentration, dialysed and digested with TEV protease. However, there were early signs that the protein displayed instability (highly disordered). The protein aggregated during buffer exchange (dialysis) and concentration (in vivaspin™ columns). Most of the TEV protease digested protein was prone to precipitation during low-speed centrifugation. In addition, during the 500 mM imidazole wash stage of Ni^{2+} affinity chromatography protein elution was staggered over 25 ml elution buffer volume (figure16).

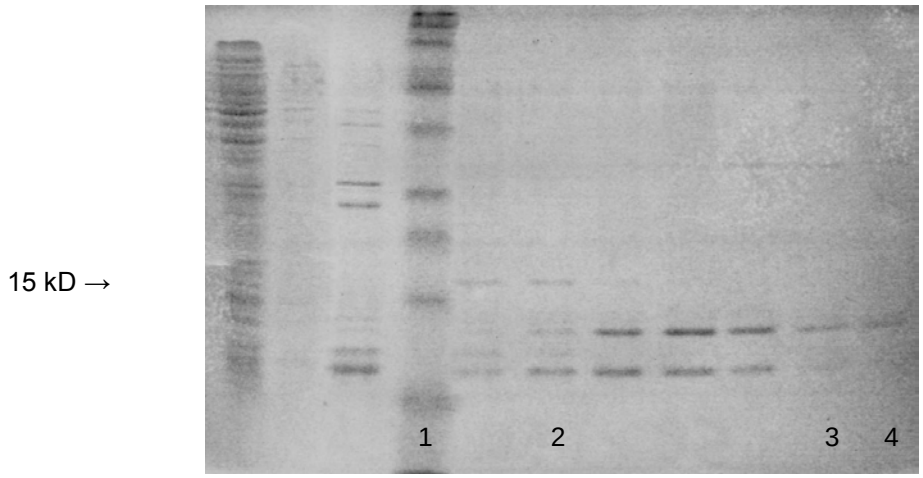

Figure 16: SDS-PAGE gel image of IMAC eluate fractions. Lane 1 – Size markers, Lanes 2-4 – Eluate fractions collected from Ni^{2+}-NTA columns washed with 500 mM imidazole. Lane 2 – Fraction after 1 ml elution, Lane 3 – Fraction after 15 ml elution, Lane 4 – Fraction after 25 ml elution. Soluble protein was injected Into the column and washed sequentially with 5 CVs of lysis buffer containing 200, 300, 400 and 500 mM imidazole. Eluates were collected as 0.5 ml fractions. Aliquots of the samples were boiled in gel loading buffer and electrophoresed in 12% PA stacked gels at 200 volts. Protein bands were visualized by Coomassie blue staining. The lower band (~11 kDa) is explained in section 3.5.1.

Several reasons which may account for protein instability were reviewed. The nucleotide sequences of pWt- and pMt FLNa24 plasmids were also re-scrutinized; it became obvious that an unwanted L2567Q mutation, 81 bases towards the N-terminal end of the target residue Gly2593 had inadvertently inserted into pWt hsFLNa24 CDS. It is presumed that this mutation originated from cDNA clone pFLNa used in wt FLNa24 amplification (section 2.2). Mutation of the conserved, nonpolar Leu residue in the putative loop region connecting beta sheets A-A' [Himmel *et al*, 2003] to a Gln residue with uncharged polar side chains may have contributed to protein structural instability. A somewhat analogous situation was reported by Vorgerd *et al* [2005], where a W2710X truncation mutation in the C-terminal end of hsFLNc24 outside the dimer interface resulted in a protein exhibiting improper folding, as indicated by lower amplitudes in circular dichroism spectrometry studies (compared to the wildtype) and a high degree of aggregation.

3.4. REDESIGNED VECTORS

Thus, pWt- and pMt hsFLNa24 plasmid constructs (sections 3.1.1 and 3.1.3) both contain a superfluous L2567Q mutation (C*T*G → C*A*G). The mutation at residue 2567 was reversed by site directed mutagenesis to revert to wild type sequence. New complementary primers 5´ GGCCAAGGGCC*T*GAGCAAGGCCTACGTAGG 3´ and 5´ CCTACGTAGGCCTTGCTCAGCCCC*A*GGCCCTTGGCC 3´ containing the Q2567L mutation (CAG→CTG) in the middle were designed. New pWt- and pMt hsFLNa24 plasmids were prepared as described in section 2.6 and the proteins expressed in *E. coli* BL21 (DE3) under optimized conditions (section 3.2).

3.5. hsFLNa PROTEIN ISOLATION

Frozen cell pellets expressing the corrected wt hsFLNa24 (section 3.2) were disrupted in lysis buffer using French Press followed by sonication to fragment genomic cell lysate supernatant obtained by ultracentrifugation was injected into Ni^{++}-NTA column

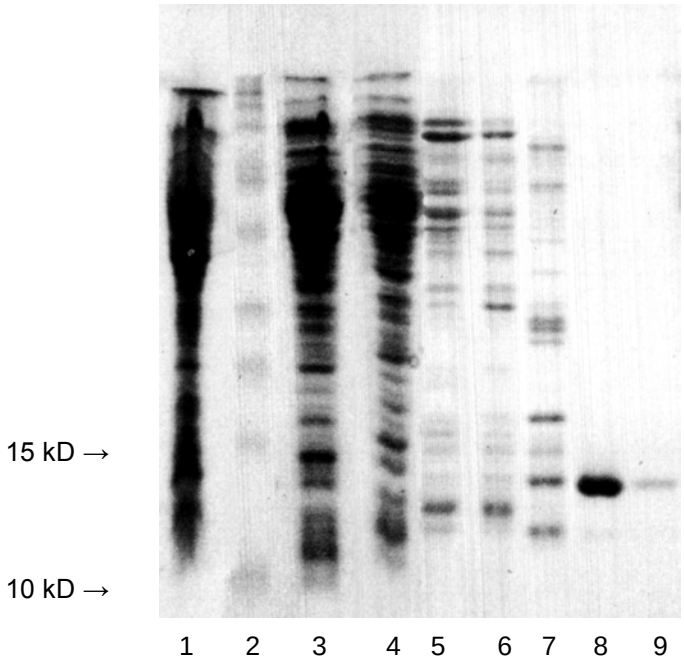

Figure 17: A 12% SDS-PAGE gel of cell lysate and column eluates of mt hsFLNa. Lane 1 – 1 µl 1000x diluted soluble cell extract, Lane 2 – Size standards, Lane 3 – 1 µl 300x diluted supernatant fraction used in injection, Lane 4 flowthrough fraction, Lane 5 – 200 mM imidazole eluate fraction, Lane 6 – 300 mM imidazole eluate fraction, Lane 7 – 400 mM imidazole eluate fraction, Lanes 8 – 500 mM imidazole eluate fraction, Lane 9 – aliquot from the fourth five ml fraction of 500 mM imidazole elution.

and washed sequentially with 5CVs of lysis buffer containing 200, 300,400 and 500 mM imidazole, respectively. Pure His-Tagged wt hsFLNa24 and mt hsFLNa24 were obtained in the 500 mM imidazole eluates (figure 17).

3.5.1. TWIN BANDS

Fractions eluted at 500 mM imidazole concentration contained the putative hsFLNa24 protein band accompanied by a faster migrating band of M_w ~11 kDa (figure 18). The lower band is more intense in wt FLNa24 eluate fractions. Several scenarios are possible in this situation. A truncated form of His-tagged hsFLNa24 could be co-expressed, or a His-rich protein or a cryptic His-tagged translation product was being co-eluted with His-tagged hsFLNa24.

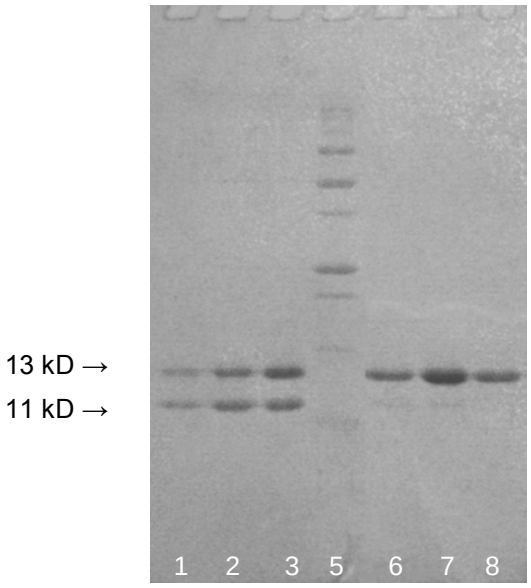

Figure 18: SDS-PAGE gel image of fractions eluted from Ni^{2+}-NTA columns with 500 mM imidazole. Lanes 1-3 wt FLNa24 elutions from 3 separate runs. Lanes 6-8 mt FLNa24 elutions from 3 separate runs. Samples were boiled in gel loading buffer and electrophoresed in 12% PA stacked gels at 200 volts. Protein bands were visualized by Coomassie blue staining.

Western hybridization (section 2.11) was performed to investigate these possibilities using wt- and mt FLNa24 eluate fractions of 500 mM imidazole washes; as control, an aliquot of mt FLNa24 protein was digested with TEV protease and the digest mixture containing the His-tagged enzyme and cleaved mt FLNa24 was used (figure

19B). The blot was screened with anti-His_6 Mabs and hybridization was visualized by chemiluminescence (figure 19A).

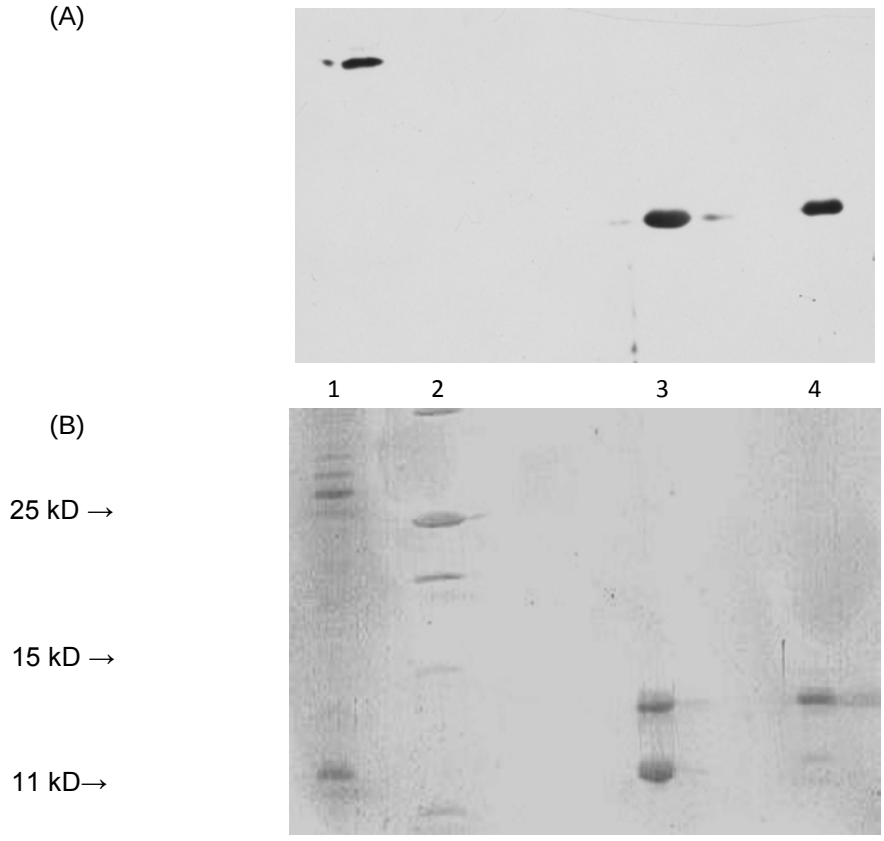

Figure 19: (A) Western blot of the gel (B) Duplicate SDS-PAGE gel image. Lane 1 – Digest mix containing TEV protease and cleaved wt FLNa24, Lane 2 – Size markers, Lane 3 – wt FLNa24 eluate fraction from 500 mM imidazole wash, Lane 4- mt FLNa24 eluate fraction from 500 mM imidazole wash. For SDS-PAGE gel (A), samples were boiled in gel loading buffer and electrophoresed in 12% PA stacked gels at 200 volts. Protein bands were visualized by Coomassie blue staining. For western blots (A) gels were run in duplicate and one of the gels was electroblotted to a nitrocellulose membrane (section 2.11) and bands hybridizing to anti-His antibodies were visualized by chemiluminescence.

Hybridizing signals obtained in lanes 1, 3 & 4 correspond to his-tagged proteins TEV protease (M_w 29kDa) wt- and mt FLNa24 (M_w 13 kDa), respectively; no signals were obtained with the lower (M_w 11 kDa) bands in the three lanes. This suggests that the fast moving bands of lanes 3 and 4 are not His-tagged proteins, or proteins whose surface presents a His-rich interface which binds with high affinity to Ni^{2+} (at high imidazole concentration); the latter supposition can be clarified by interrogating

the blot with anti-$His_{4/5}$ antibodies capable of detecting a smaller number of (sequentially or spatially) contiguous His residues.

The low molecular weight bands in lanes 3 & 4 are identical in size to the His-cleaved FLNa24 protein in lane 1; secondly, the twin bands of His-tagged wt FLNa preparation (lane 3), on digestion, were reduced to a single band (lane 1) which was collinear with the lower band of undigested protein preparations (lane 3). This leads us to presume that the smaller protein observed in the eluate fractions of 500 mM imidazole wash is indeed the His-tag cleaved form of FLNa24; anti-hsFLNa antibodies may help to confirm the identity of the smaller protein. Nevertheless, it is difficult to explain the occurrence of two forms of the proteins in the eluate fractions unless we assume that N-terminal overhang cleavage arose by *in situ* proteolysis; proteolysis was inhibited in mt FLNa24 probably due to an altered domain structure.

3.6. SIZE EXCLUSION CHROMATOGRAPY (SEC)

During gel filtration, proteins partition between the mobile phase within and without the resin support as a function of their distribution coefficient, KD defined by their molecular size. Dimensional parameters such as Stokes radius (R_S) or viscosity radius (R_η) which determine KD are used to study the molecular size of new proteins [Le Maire *et al*, 1989]. In addition, Andrews [1962] showed that within a defined molecular weight range, a linear relationship exists between Log_{10} molecular weight and the relative elution volumes (V_e/V_o) of proteins, which can be used to estimate molecular weight of proteins in a gel filtration medium composed of cross-linked dextrans [Andrews, 1964; Andrews, 1965] from a standard curve or the equation defining the best-fit regression line derived from the elution profiles of molecular weight standards..

TEV protease-digested hsFLNa24 proteins were recovered by IMAC and buffer exchanged with size exclusion chromatography buffer. Concentrated proteins were injected into a calibrated Superdex 75 HR 30/10 size exclusion column with a fractionation range 3-70 kDa globular proteins, in a FPLC system. Peak fraction contents were confirmed by PAGE and the V_e/V_o ratio was used to determine protein

molecular weight (section 2.9). wt- and mt hsFLNA24 proteins displayed a shift in their peak elution volumes with mt hsFLNa eluting later; elution volumes of wt hsFLNa24 and mt hsFLNa24 were 13.45 ml (figure 20, solid line) and 14.1 ml (figure 20, dotted line), respectively.

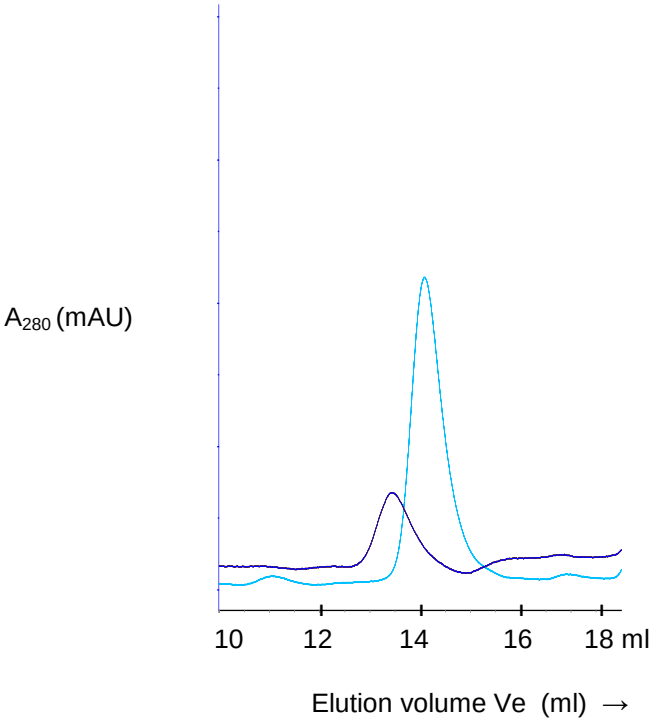

Figure 20: Elution profiles of proteins after SEC. Solid line – retention volume of wt FLNa24, dotted line – retention volume of mt FLNa24. His-tagged hsFLNa proteins were digested with TEV protease, digested proteins recovered by IMAC, buffer exchanged with SEC buffer and injected into a calibrated Superdex 75 HR 30/10 size exclusion column. Peak elution was monitored by absorbance monitor at A_{280}.

The peak elution volumes correspond to molecular weights of 19 kDa and 11 kDa, respectively, which show a nearly two-fold difference. The estimated molecular weight of mt hsFLNa24 approximates to the monomer's absolute molecular weight of 10 kDa calculated from its amino acid composition; such variations in molecular weight values can be anticipated when molecular weight is determined by methods utilizing different physical and chemical parameters. The estimated molecular weight of wt FLNa24 is nearly twice its calculated molecular weight. Molecular weight values obtained by gel filtration reveal protein-protein interactions, either by self association or with other molecules since the values are those of proteins as they exist in their native

state [Irvine, 2001]. Our SEC results suggest that in its native form, wt hsFLNa24 exists as a dimer; the mutant form probably exists as a monomer with a slightly larger molecular size as a result of structural alteration.

3.7. CROSSLINKING

Protein crosslinkers bind proteins together covalently thereby providing a means of capturing even transient protein-protein complexes to give a snapshot of molecular interaction events; these are visualized by SDS-PAGE, western hybridization, immunoprecipitation or mass spectrometry. A crosslinking assay was performed as an alternative means to verify the oligomerization status of hsFLNa24 proteins in solution. His-tag cleaved wt- and mt hsFLNa24 preparations were crosslinked with the *N*-Hydroxysuccinimide (NHS) ester EGS and reaction products were examined by SDS-PAGE; reactions without EGS addition were used as assay controls (Figure 21).

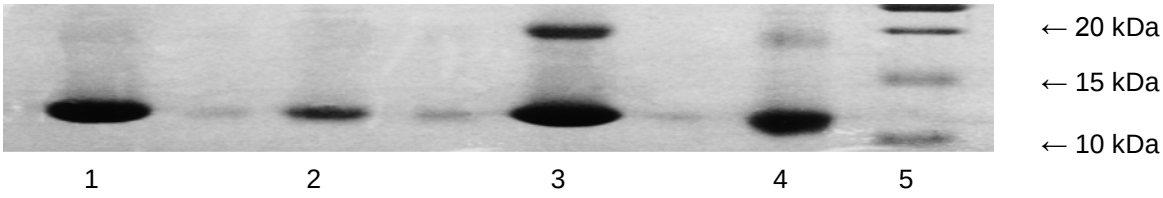

Figure 21. Crosslinking of hsFLNa24 proteins by EGS. Lane 1: wt hsFLNa24 control reaction without EGS addition, 2: Control His-tagged mt hsFLNa24 control reaction without EGS addition, 3. wt hsFLNa24 incubated with EGS, 4: mt hsFLNa24 incubated with EGS. Reaction was quenched with TCA and the air dried acetone-washed pellets were electrophoresed in 12% PA stacked gels at 200 volts. Protein bands were visualized by Coomassie blue staining.

The control reactions without EGS addition (lanes 1 & 2) contain only monomer bands of Mw 11 kDa corresponding to the His-tag cleaved protein. A prominent species, twice the size of the protein monomer is observed in reactions where wt FLNa24 was incubated with EGS (lane 3); in the case of mtFLNa24 treated with EGS, the monomer species is the major form, accompanied by a faint dimer band (lane 4).

It appears that under these crosslinking conditions, the high molecular weight dimeric species is a major form of wt FLNa24. In the case of mt FLNa24, dimerisation is significantly inhibited and the protein exists predominantly in a monomeric state.

Crosslinking is not an equilibrium measure of monomer-dimer conversion, but the observation that dimers can be efficiently trapped in dilute solutions of monomer concentrations of the order 10^{-6} M indicates that wt FLNa24 molecules exist as dimers. The rate-limiting step of a cross-linking reaction with reagents such as EGS is generally the reaction of the first reactive end of EGS with the protein. The reaction of the second end of the cross-linker is accelerated by the proximity effect of tethering the reagent to the protein surface [Staros, 1988]. The proteins should be less than 16.1 Å apart for EGS to crosslink primary amines existing on the surface of two proteins, by forming amide bonds at opposite ends of its spacer arm. Therefore, it is possible that a majority of the EGS-induced dimers trapped by cross-linking were dimeric prior to addition of the cross-linker and were not formed by collision subsequent to the addition of one end of the cross-linker to a protein monomer. Further, the presence of dimers at concentrations as low as 10^{-6} M suggests that wt hsFLNA24 dimers form at moderately high affinity.

In this assay, conditions which were previously optimized for a phosphate buffer [Himmel *et al*, 2003] were applied to a HGNED buffer (as recommended by the manufacturer); the relatively low crosslinking yields may be due to significant levels of crosslinker hydrolysis. Therefore, further optimization of reaction conditions such as cross-linker:protein ratio, salt concentration and temperature may pave the way for clearer results; in addition, immunostaining of a western blot (of the gel), or the use of a crosslinker such as DST with a shorter spacer arm length (6.4 Å) could lend more convincing evidence of FLNa24 dimerization.

3.8. THERMODYNAMIC STABILITY

The thermal stabilities of the wildtype and mutant proteins were examined to see whether the mutation caused structural instability. Protein stability can be studied by methods such as circular dichroism, nuclear magnetic resonance and differential scanning calorimetry. In this study, melt curve analyses based on fluorescent dye binding to the unfolded state was used (section 2.12). Folding of globular proteins is believed to be driven by hydrophobic collapse where the hydrophobic groups coalesce

so as to expel most of the surrounding water molecules, thereby forming a predominantly hydrophobic core and a hydrophilic exterior. In this native state its global free energy is at a minimum. An increase in entropy resulting from an increase in temperature leads to a decrease in order. As the protein unfolds more hydrophobic residues become exposed into which reporter dye SYPRO orange partitions to give rise to a fluorescence signal (section 2.12).

Using the melt curve feature of real time PCR, the T_m of wtFLNa24 was calculated to be ~317 K and that of mtFLNa24 <303 K. The wtFLNa24 thermal profile approximates to a sigmoidal curve, while that of mtFLNa24 was non-sigmoidal (melting curve electrograms not shown); in the latter case, the highest fluorescence signal was recorded at the lowest (starting) temperature and it decayed as the temperature rose. One could speculate that mtFLNa24 in its native state has a high-entropy, disordered conformation. As a consequence, all possible binding sites that would be exposed by an increase in entropy were saturated with SYPRO and hence the highest fluorescence signal registered at the lowest starting temperature of our assay. The results show that the mutant protein is of lower thermally stability, which may be a reflection of its structural disorder.

3.9. X-RAY CRYSTALLOGRAPHY

3.9.1. CRYSTALLIZATION

Protein crystallization during sparse matrix screen was monitored using a microscope (section 2.13). In plates containing wt FLNa24, some precipitant wells remained clear, while several reagents gave a granular, light brown precipitate indicating a higher than optimal protein concentration. A few others produced needle clusters, microcrystals and plates. From among the various precipitants, some that formed crystalline aggregates were,

1 M Ammonium phosphate, 0.1 M Na citrate, pH 5.6

30% PEG 400, 0..2 M Na citrate, 0.1 M Tris HCl, pH 8.5

30% PEG 4000, 0.2 M Lithium sulfate, 0.1 M Tris HCl, pH 8.5

30% PEG 4000, 0.2 M $MgCl_2$, 0.1 M HEPES, pH 7.5

1.4 M Na citrate, 0.1 M HEPES, pH 7.5

A single crystal diffracting to 1.85 Å was successfully recovered from the crystallization mother liquor containing 0.2 M lithium sulphate. 30% PEG 4K in 0.1 M Tris, pH 8.5 by hanging-drop vapour diffusion method. In plates containing mt hsFLNa24, most wells remained clear while some gave thick precipitates, indicating a need to test different precipitants or for further optimization.

3.9.2. PROTEIN MODELLING AND CRYSTAL STRUCTURE

After collecting the first few images, two images separated by nearly 90º in phi were indexed and integrated, since the unit cell derived from only one image may be inaccurate. A total of 48,967 reflections were measured of which 12,027 were unique reflections. Batch refinement solved the crystal symmetry system as orthorhombic (*oS*, cell parameters $a \neq b \neq c$; $\alpha = \beta = \gamma = 90°$). Unit cell parameters were a = 41.0 Å, b = 58.2 Å, c = 94.4 Å. The unit cell is characterized by a 222-fold rotational symmetry. Protein chiral molecules can crystallize only in one of 65 of a possible 230 unique space groups. Further processing performed to a resolution of 1.85 Å placed hsFLNa24 in the space group $P2_12_12$. The wt hsFLNa24 crystal contained two molecules in the asymmetric unit with a solvent content of 40%.

The refined optimal cell statistics derived with MOSFLM [Leslie, 1992] were substituted in SCALA [Evans, 2006] and processed through interactive cycles of SCALA, which uses Bayesian statistics and Wilson plots to derive structure factors and B-factors (atomic temperature factors). The overall R_{merge} value was 0.093 (0.42), overall I/σ(I) 7.4 (2.6) and average multiplicity ~ 4 (4) (values in paranthesis correspond to highest resolution shell – 2.32 Å).

Phases were derived by the molecular replacement method [Rossman & Blow 1962] using Phaser [McCoy *et al*, 2007]. The structure of hsFLNc (PDB ID IV05), which shows 68% amino acid sequence identity to hsFLNa (figure 22) was used as the search model after all non-identical residues were mutated to alanine.

```
FLNa24 ADASKVVAKGLGLSKAYVGQKSSFTVDCSKAGNNMLLVGVHGPRTPCEEILV
FLNc24 SDASKVVTRGPGLSQAFVGQKNSFTVDCSKAGTNMMMVGVHGPKTPCEEVYV
       :*******::* ***:*:**** .********** .**::******:*****: *
FLNa24 KHVGSRLYSVSYLLKDKGEYTLVVKWGHEHIPGSPYRVVVP
FLNc24 KHMGNRVYNVTYTVKEKGDYILIVKWGDESVPGSPFKVKVP
       **:*.*:*.*:* :*:**:* *:****.* :****::* **
```

Figure 22: Sequence alignment of hsFLNa24 and hsFLNc24 protein sequences. Alignment was performed with ClustalW. * indicates identity and : conservative substitution.

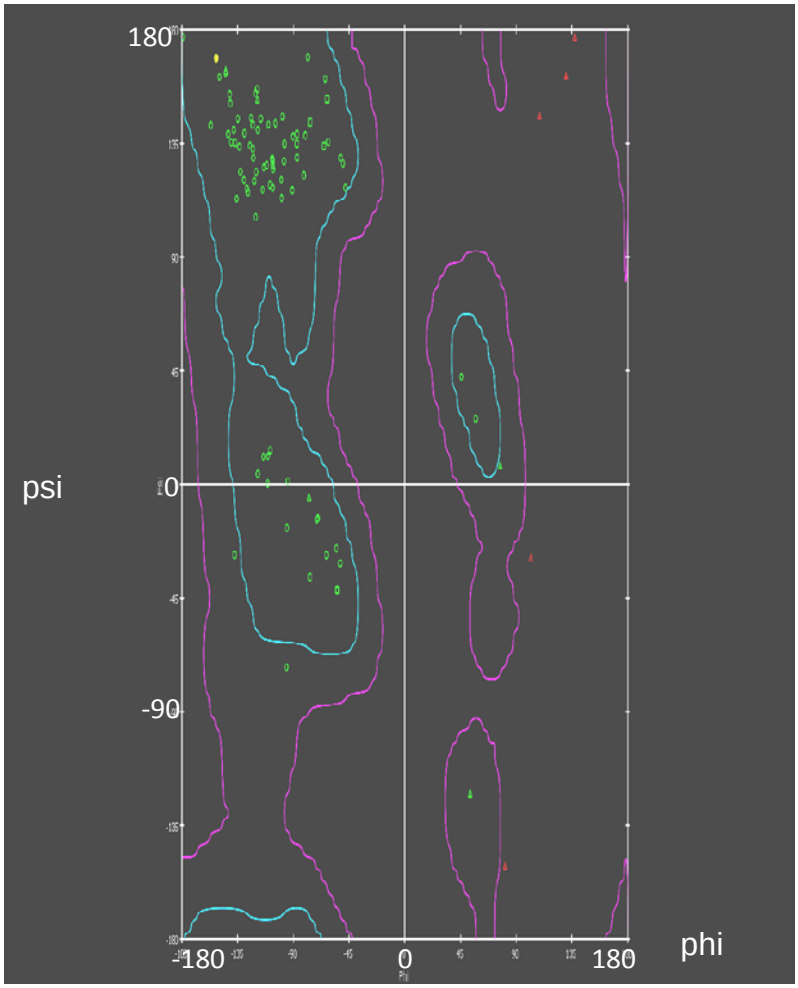

Figure 23: Ramachandran plot of hsFLNa24 residues. The pair of angles of rotation phi (N-C_α) and psi (C-C_α) of each residue is plotted against each other. Δ=glycine, □=proline and ○=other residues. Most of the residues are clustered in the sterically allowed α, β and L regions. Glycine, which has only hydrogen as a side chain, can adopt a wider range of conformation than other residues.

Rigid body refinement using REFMAC [Murshudov *et al*, 1997] gave an initial crystal structure with an R_{factor} of 44% and R_{free} of 43%; using iterative cycles of

refinement, addition of water molecules and model adjustment using COOT [Emsley & Cowtan, 2004] the final model, refined at a resolution of 2.0 Å yielded an R_{factor} of 20.0% and R_{free} of 25%. Ramachandran plot shows that 97% of the residues are in the most favoured regions, 2% in additionally allowed regions and 1% outliers, indicating good stereochemistry of the structure (figure 23).

The FLNa24 repeat domain structure (93 residues) has an immunoglobulin (Ig)-like fold of the E-set superfamily [Murzin *et al*, 1995] with a predominantly β–sheet structure. Each monomer is a β-sandwich with seven β-strands organized into two anti-parallel β–sheets, where strands A, B, E and D (or 1-2-5-4) form a four strand β–sheet and β–strands C, F and G (or 3-6-7) a three strand antiparallel β–sheet (figure 24).

Figure 24: Topology diagram of hsFLNa24 monomer. β-strands coloured yellow, loops green and helices red. N = amino terminus, C = carboxy terminus. β-strands are labeled A, B, C, D, E, F and G, from the N-terminal end to the C-terminal end.

The 7 strands displaying a right-handed twist, are 3-9 residues long and connected by loop regions of variable length. Table 1 shows the residues constituting the elements of the crystal structure.

Secondary structure elements	Residues
N-terminus end	ADASKV
Strand A	VAK
Loop AB	GLGLSKAYVGQK
Strand B	SSFTVD
Loop BC	CSKAGNNM
Strand C	**LLVGVH**
Loop CD	GPRTPC
Strand D	**EEILVKHVG**
Loop DE	SR
Strand E	LYSVSYL
Loop EF	LKDK
Strand F	GEYTLVVKW
Loop FG	GDEHIPGSP
Strand G	YRVVVV
C Terminus end	P

Table 1: Elements of hsFLNa24 secondary structure. Underlined bases are residues forming helix and residues in bold form the putative dimer interface. Cyan – hydrophilic residue, magenta – hydrophobic residue, green – turn predictor.

In a hydropathy plot of FLNa24 residues, the β-strands correspond to regions exhibiting hydrophobic properties (Figure 25A).

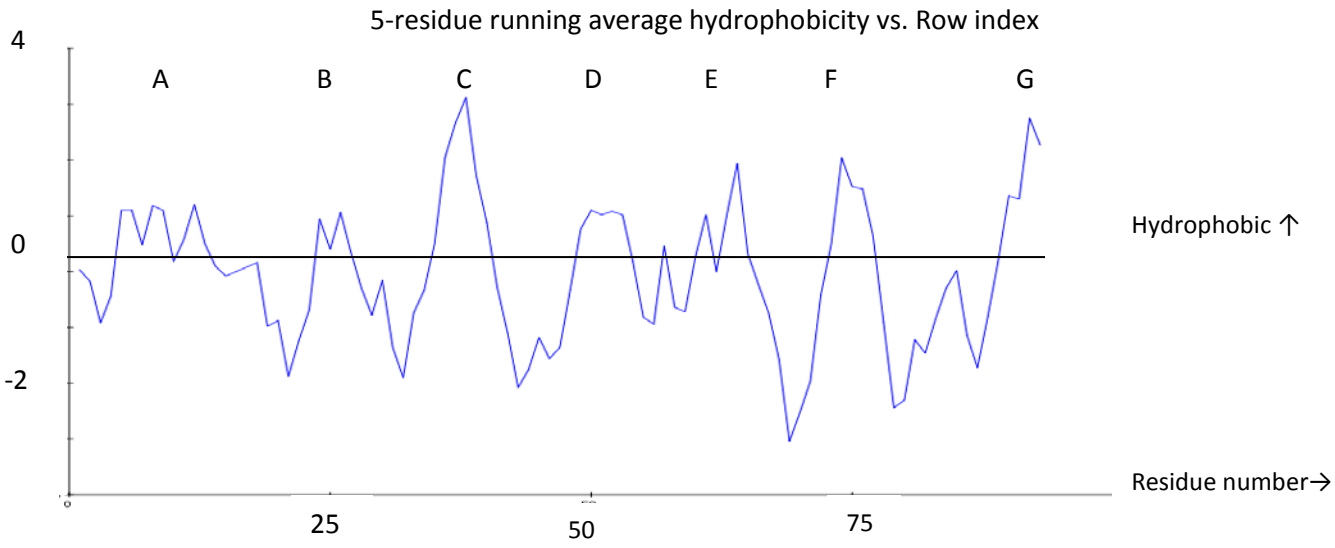

Figure 25A: Hydropathy plot of FLNa24. The sum of hydropathies of 5 consecutive residues were used with the hydrophobic scales of Kyte and Doolittle [1982] and plotted against residue number. A large positive hydropathic index indicates a hydrophobic region. Regions corresponding to the β-strands are indicated on top.

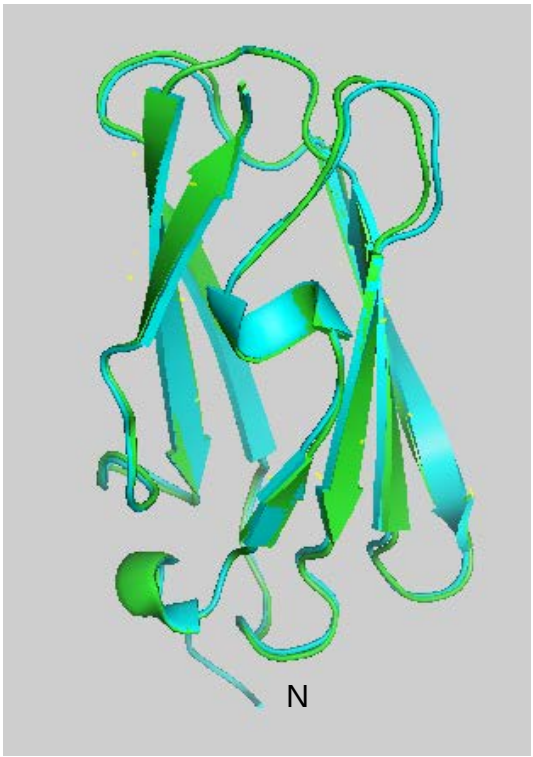

Figure 25B: Cartoon representations of FLNa24 superimposed on FLNc C-α atoms. Green – FLNa24, Blue – FLNc24. 85 atoms were used in the final alignment. N – amino terminus

The crystal structure of hsFLNa24 is similar to structures of repeat domain 24 of human filamin isoforms hsFLNc and hsFLNb, whose primary structures share 68-70% amino acid sequence identity. 90% of the amino acids constituting the β-sheets of the three hsFLN structures are invariant residues or conservative substitutions; only a single difference occurs within the putative dimerization interface at position 2605 of FLNa24 where a leucine occurs, a serine in FLNb and tyrosine in FLNc at corresponding positions. The structural homology (given by root mean square deviation - rmsd) between FLNa24 and several FLN domains was estimated using the align feature of Pymol [DeLano, 2002]. FLNa24 showed closest structural homology to FLNc24 with an rmsd value of 0.38 Å of 85 superimposed C-α atoms. The loops contribute to most of the topological differences between FLNa24 and FLNc24, of which those connecting β-strands A and B are most divergent (figure 25B).

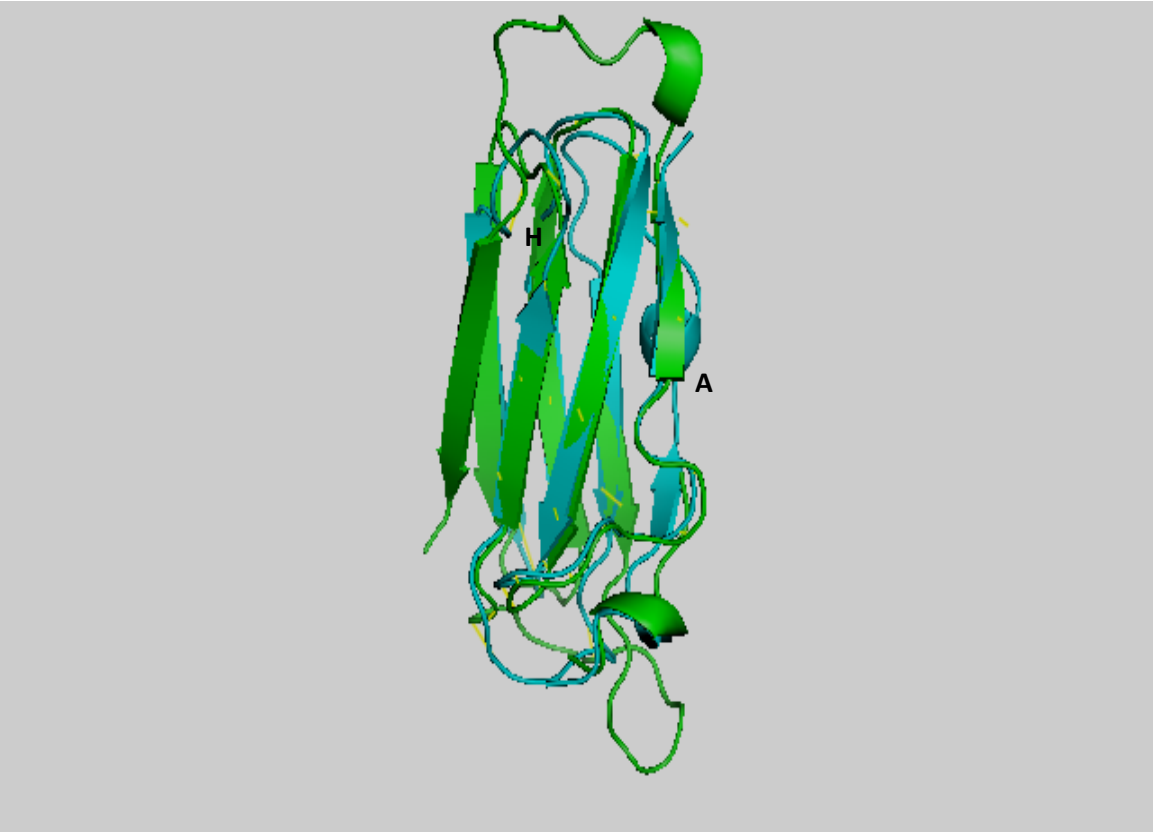

Figure 26: Cartoon representations of hsFLNa24 superimposed on ddFLN6 C-α atoms. Blue – hsFLNa24, Green – ddFLN6. 52 atoms were used in the final alignment.

FLNa24 bears close but less structural homology to FLNb24 (rmsd value of 0.89 Å of 84 superimposed C-α atoms); the divergence arise from differences in both β-strands and loops. FLNa24 is distinct and most divergent from dimerizing domain ddFLN6 of *D. discoideum* (rmsd = 2.19 Å). The large difference is attributable in main to the absence of strand A and the presence of an extra strand H in ddFLN6 (figure 26).

The asymmetric unit consists of a dimer in which two monomers lie in an antiparallel arrangement (figure 27).

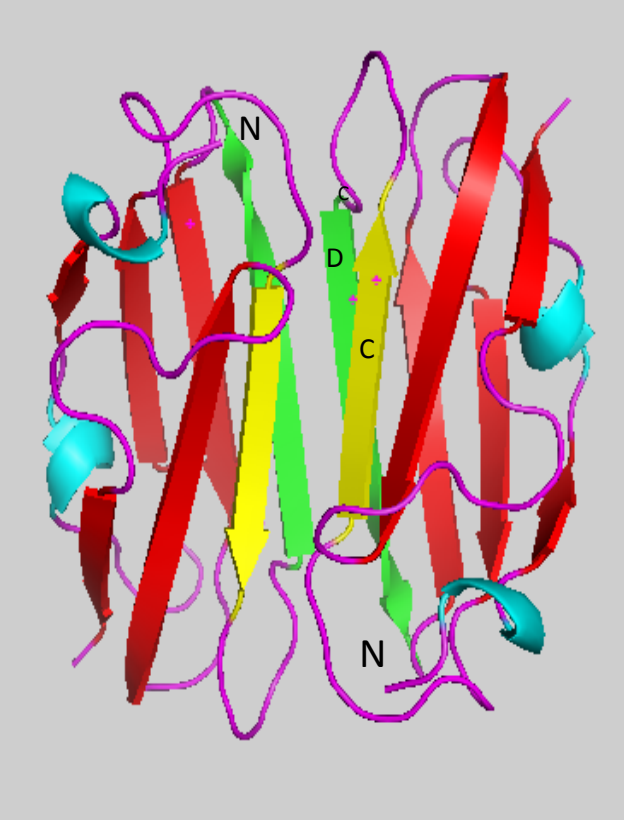

Figure 27: Crystallographic dimer of hsFLNa24. Strand C is coloured yellow and strand D green. Other β-strands are coloured red, loops magenta and helices blue. N = amino terminal end. The two monomers are arranged in antiparallel fashion.

The dimer is made up of 1424 protein atoms and 226 water atoms and displays a solvent accessible surface area of 8390 $Å^2$. The putative dimerization interface buries 1405 $Å^2$ of the monomer's solvent accessible surface area. The interface involves the opposite edges of the β sandwich mediated mainly by β-strands C and D. The largely

hydrophobic interface has a hydrophobic core consisting of hydrophobic residues leucine, valine, glycine, and isoleucine surrounded by a ring of polar and charged residues. The nonpolar side chains of residues L2591 and L2592 of the C-strand of one monomer, pack against G2593 of the C-strand of the other monomer to form hydrophobic stacking (figure 28).

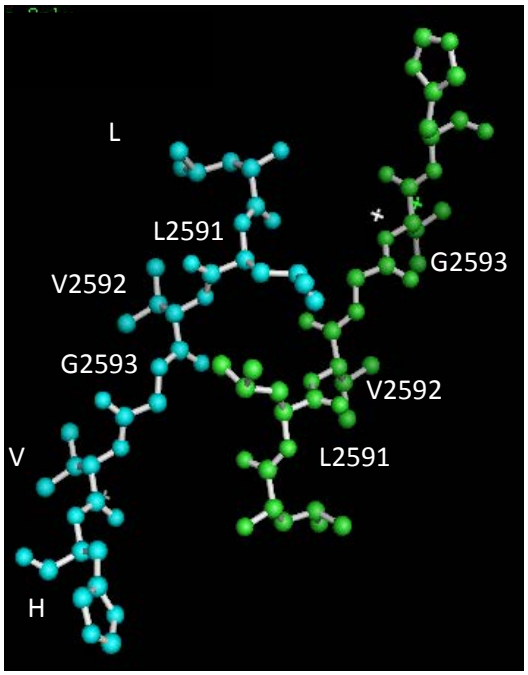

Figure 28: The antiparallel C strands of the dimer interface shown as stick-and-ball model. Hydrophobic stacking is formed by residues L2591, V2592 and G2593. The chains are colored in blue and green.

A similar hydrophobic interaction is observed between monomers of the hsFLNc dimer involving two methionines and one glycine residues of the C-strand [Pudas *et al*, 2005].

```
FLNa24 MLLVGVHGPRTPCEE
FLNb24 MLLIGVHGPTTPCEE
FLNc24 MMMVGVHGPKTPCEE
```

Figure 29: Multiple sequence alignment of hsFLNa24, hsFLNb24 and hsFLNc24 C-strand protein sequences. The hydrophobic dimerization signature is underlined Alignment was performed with ClustalW.

Glycine residue G2593 of the triumvirate forming hydrophobic packing is highly conserved in all hsFLN isoforms (figure 29). In simulation studies *in silico* (Accelrys®), G2593 was mutated to E2593 ie G2593E. The mutant structure showed no deviation from wildtype structure (rmsd 0.00 $Å^2$) or in the interface area.

Water molecules are associated with the four hydrogen bonds connecting the C strands. Also embedded within the hydrophobic core, are six hydrogen bonds (Figure 30) and five hydrogen bonds mediated by water molecules connecting the two D strands.

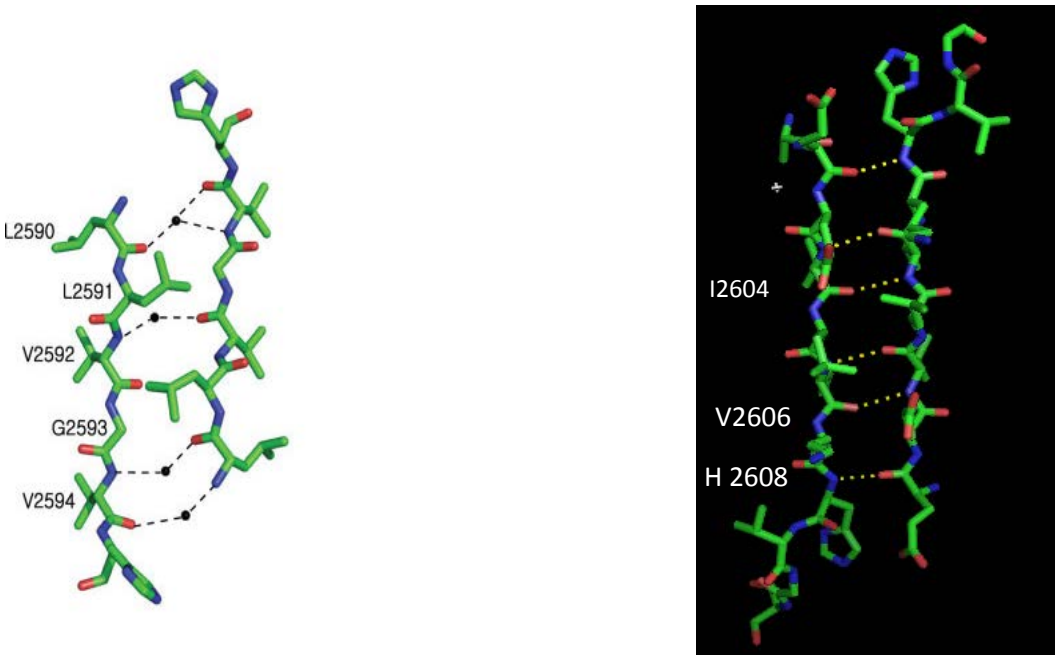

Figure 30: Antiparallel C (left) and D (right) strands of the dimer interface shown as stick-and-ball model. Black dotted lines indicate water-mediated hydrogen bonding between C-strands and yellow dotted lines polar contact between D strands.

CHAPTER IV – DISCUSSION AND CONCLUSION

4.1. hsFLNa MOLECULAR ARCHITECTURE IN THE DIMERIZATION REGION

The crystal structure of hsFLNa24 dimer (section 3.9.2) shows that the two monomers are arranged in an antiparallel somewhat similar to the arrangement in Ig6 dimer of gelation factor, whose chains appear as antiparallel rods *in vitro* [Condeelis *et al*, 1984]. In *D. discoideum*, this dimer arrangement is consistent with the observed architecture of ddFLN where the two chains of the molecule are arranged in an antiparallel manner overlapping at domain 6 (figure 10) [Popowicz *et al.* 2004].

Electron micrographs [Gorlin *et al*, 1990] of native hsFLNa dimers purified from smooth muscles on the other hand, show mostly V-shaped molecules (figure 1), where the two chains are arranged in a parallel, unidirectional formation. Pudas *et al* [2006] reconciled the anomaly between chain orientation and dimer structure in hsFLNc by proposing that the two-fold symmetry axis of the dimerization domain is orientated parallel to the long axis of hsFLNc rod region 2 (Figure 31); hinge 2 segment consisting of 38 residues, acts as swivel to confer flexibility [Gorlin *et al,* 1990;

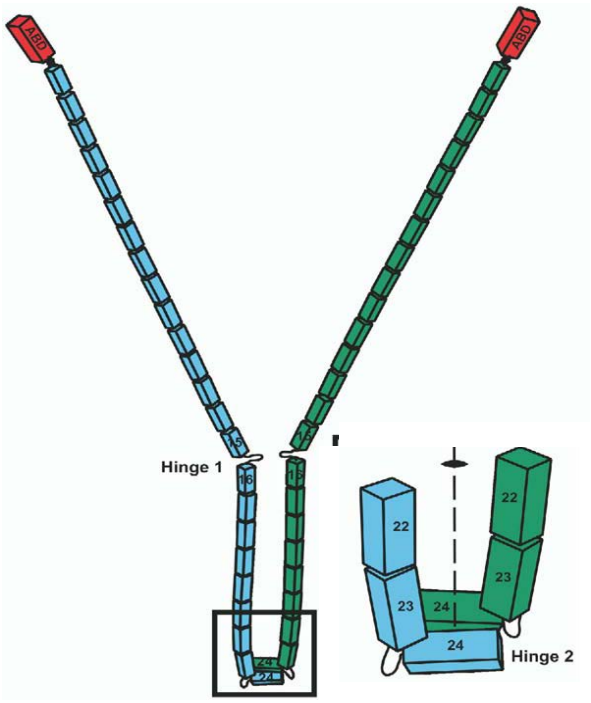

Figure 31. Proposed model of the orientation of full filamin chains in relation to FR24 dimerizing domains. The dimerization domain two-fold symmetry axis is positioned parallel to the in detail terminal rod region. Hinge 2 provides the necessary elasticity to orientate FR24. Inset shows the arrangement domains 22, 23 and 24 [Pudas *et al*, 2006].

Robertson *et al*, 2003] for rotation of the dimer dyad axis. We can speculate that this hypothesis may apply equally well to the spatial organization of hsFLNa24, since it shares high sequence and structural homology with hsFLNc24. Nevertheless, the possibility of an antiparallel arrangement of the chains cannot be totally excluded.

The study of isolated dimerising repeats alone, as in this study, is not sufficient to confirm the higher order structure of the dimerized terminal filamin repeats in relation to hsFLNa molecule. In its native state, the terminal repeats occur in association with other components of the protein chain where several inter- and intra-chain interactions are possible, which could alter their disposition. Based on early electron micrograph profiles of hsFLNa molecules, which show varying angles between the two chains, Gorlin *et al* [1990] were prompted to suggest the existence of further sites of interaction between domains of the two chains. This view is supported by evidence from the composite crystal structure of an hsFLNa fragment containing repeats 19-21, where the repeats do not necessarily fold as distinct domains [Lad *et al*, 2007]; the crystal structure reveals that Ig20 is partially unfolded and contributes one β strand to domain 21 and that domain 21 is positioned between repeats 19 and 20. Popowicz *et al* [2004] also showed the existence of salt bridges between neighboring repeats of ddFLN. Further, the dissociation constant of FLNc24 (K_d = 2.5 μM), which is of the same order as its cellular concentration implies that other interactions may also play a role in dimerisation [Pudas *et al*, 2006].

Confirmation of the packing of the Ig24 dimer within the FLNa structure is forthcoming from crystals containing contiguous repeats [Popowicsz *et al*, 2004] including repeat24, where the elbow angle between the long axes of the repeats would reveal the spatial organization of the C-terminal end domain

4.2. DETERMINANTS OF VERTEBRATE FILAMIN REPEAT DOMAIN INTERACTION

In hsFLNa24 in our studies as well as in hsFLNc24 [Pudas *et al*, 2005] and hsFLNb24, dimerization is mediated by β-sheets C & D of the monomers. Hydrophobic residues in strands C and D that form the dimerisation interface in hsFLNa24, are well conserved among vertebrate filamin species (section 1.4.2). Moreover, these sequence motifs are not found in other repeat

domains of vertebrate filamins. Based on this observation it was proposed that all vertebrate filamins dimerise through the C and D strands of the dimerizing repeats.

A similar feature is observed in interactions between the other hsFLN repeat domains and several ligands. Crystal structures of hsFLNa24 bound to the peptide from integrin β-cytoplasmic tail shows that interaction is mediated by β-strands C and D of the filamin domain [Kiema *et al*, 2006]. Similarly, strands C and D of hsFLNa17 are implicated in its binding with glycoprotein Ibα [Nakamura *et al*, 2006]. Thus, it has been suggested that the CD face is a common filamin-ligand binding interface [Kiema *et al*, 2006].

The G1897 residue in filamin A 17 β-strand C that interact ith GPIbα □s well conserved in odd-numbered domains (Ig19, Ig21, Ig23). However, C1910 located on the D strand, shows greater variability between domains, changing to F in hsFLNa21, T in hsFLNa 19, but remaining conserved in repeats 23, 22 and 18. Thus it has been generalized that while β-strands C and D are the major determinants of filamin domain interactions, particular residues in these strands are responsible for the specificity of ligand recognition.

4.3. THE G2593E MISSENSE MUTATION

Residue G2593 of hsFLNa24 β–strand C contributes to the putative hydrophobic stacking in the dimerization interface. The residue is highly conserved in most vertebrate filamins (figure 4) and possibly plays an important role in the forces associated with the packing of hydrophobic side chains in the interior of the molecule. The main driving force for folding of water soluble globular proteins is to pack hydrophobic side chains into the interior of the molecule, thus creating a predominantly hydrophobic core and a hydrophilic surface. The folded state of the protein compensates for its comparatively unfavourable entropy by bringing hydrophobic residues from the solvent front to the interior of the protein and stabilizing with specific hydrogen bonds and electrostatic interactions. Ablation of the G residue in conjunction with substitution by a polar residue is predicted to play a major role in disrupting the hydrophobic nature of the interface thereby inhibiting dimerisation.

Substitution of the G residue by the polar E residue (G2593E) compromises the molecule energetically. It would cost too much energy to have an amino acid with a charged residue in a hydrophobic environment. A main chain polar group in a hydrophobic environment must be neutralized by hydrogen bond formation (when atoms of the acceptor and donor groups are within van der Waals distance) or by forming ion pairs with a nearby residue of opposite charge ie. with K, R, or H in the case of E substitution. However, even if these bonds are formed they would not contribute significantly to the molecule's stability. Individual hydrogen bonds which are central features of protein structures make only minor contribution to protein stability; deletion of a hydrogen bond does not always lead to the overall loss of the absolute energy of a hydrogen bond, the magnitude of change depending on the type of hydrogen bond. Despite the strong electrostatic attraction between opposite charged members of an ion pair, these interactions contribute little to the stability of protein [Voet & Voet, 2004].

Thus it is predicted that the net effect of G2593E mutation is to increase molecule instability, as reflected by the lowered T_m of mt hsFLNa (section 3.8). In SEC (sections 3.6) and crosslinking assays (sections 3.7), the G2593E mutant displays a partial loss-of-function phenotype with reduced dimerising ability; the reduced affinity of mt hsFLNa24 to dimerise could have resulted from a reorganization of the dimer interface arising from a reduced buried surface area. However, this was not the case as found by *in silico* simulated mutation (section 3.9.2). In the absence of protein structural changes, the mutant's structural instability can be explained in terms of net energy changes in the dimerizing zone. For example, an A82P mutation in T4 lysozyme did not cause structural alteration, yet caused an increase in T_m of 1ºC. This indicates that the effect on T_m is due to entropy changes. [Matthews *et al*, 1987]. Methods such as surface plasmon resonance, ultracentrifugation and yeast two-hybrid analyses offer means of quantitating the difference between affinities of wildtype and mutant monomers to dimerise as a measure of the entropy changes. Pudas *et al* [2006] compared the dissociation constant of wildtype hsFLNc24 and a mutant incapable of dimerisation containing a M2669D mutation in the hydrophobic patch of β-strand C. The two proteins, which showed little structural alteration, differed significantly in

their Kd values where the mutant had a weaker Kd which was three orders lower than that of the wildtype (Kd wildtype = 2.5 µM, Kd M2669E mutant = 3 mM).

4.4. ROLE OF G2593E MUTATION IN PVNH

The mutant *flnA* cDNA clone (containing a *g/a* transition at nucleotide position 7778 resulting in a G2593E mutation) used in this study, was isolated from expression libraries of a patient who had a familial history of PVNH (section 1.9). The X-linked pathology was originally described as a neural condition arising from dominantly inherited loss-of-function mutations of hs*flnA*.; in affected families an excess of females as well as of miscarriages is observed, pointing to intrauterine lethality of male fetuses [Fox *et al*, 1998]. However, in our study the male patient and his brother who have PVNH on MRI scan and possess the G2593E mutation did not succumb to early death.

flnA mutations associated with classic PVNH are truncating mutations (section 1.7) resulting in a loss-of-function, in contrast to X-linked OPD spectrum of skeletal displasias which is thought to be gain-of-function mutant phenotypes [Robertson *et al*, 2003]. The *g**g**g>g**a**g* transition in our patient is a missense mutation and therefore allows the encryption of a full length transcript (and protein); however, it has the potential to cause a loss-of-function phenotype, which is the presumed mechanism for PVNH, with severe consequences through inhibition of FLNa24 dimerisation as demonstrated in our assays. In this respect, it is similar to the effect of a PNH truncating mutation leading to loss of the most carboxy-terminal portion of hsFLNa (Sheen *et al*, 2001; Moro *et al*, 2002].

Since the patient (and his male sibling) displays only mild symptoms of PVNH, it can be speculated that he retains enough FLNa function to avoid the usual lethality associated with loss-of-function mutations in males**.** Several possibilities can be explored to explain how the system compensates for the loss of a major function.

Based on amino acid homology and predicted β structure of repeat domains of the rod region, Gorlin *et al* [1990] proposed that all or some of repeat domains 16-24 at the carboxy terminal may be involved in intrachain pairing. Supporting this is the observation of Himmel *et al* [2002] that although

FLNa24 alone is sufficient for filamin dimerisation, the hinge region which is conserved in all filamin isoforms enhanced the stability of the dimer; thus, in the event of loss of itsFLNa24 function we can assume that it can be compensated by other rod 2 region repeat domains capable of dimerisation [Gorlin *et al*, 1990].

Another possibility is a feed-back mechanism leading to upregulation of actin binding proteins such as hsFLNb that have a similar expression profile to hsFLNa (section 1.5), thereby rescuing the lost function.

hsFLNa24 is bound by at least two tissue factors (figure 6); these, or the whole scaffold of binding proteins surrounding the chains may play a role in holding them together *in vivo*.

Hehr *et al* [2006] showed that the mild form of PVNH in a male patient carrying a splice site mutation in FLNa was due to hemizygosity, where the normal transcript compensates for the loss-of-function of a shorter aberrant transcript lacking the 3' part of exon 13. While this mechanism would not apply to the 7778 *g>a* mutation which is exonic in nature, it is possible that a concurrent mutation in a different exon, which alters protein-protein interaction of that domain offsets the loss of function of FLNa24.

4.5. CONCLUSIONS

The crystal structure of the FLNa24 dimer (figure 27) shows that the dimerisation interface involves strands C and D of adjacent chains. Strand C, containing the hydrophobic residues that form the dimerisation interface is well conserved in vertebrates (figure 4) and contains the invariant glycine residue (G2593 in FLNa24). Our study was designed to investigate the structural and biochemical implications of a G2593E mutation identified in a PVNH patient. Our (*in vitro*) assays (sections 3.6 and 3.7) demonstrate that the G2593E mutation results in the inhibition of dimerisation, which can have major repercussions leading to male lethality, although not in this case. This mutation which does not cluster with PVNH mutations described previously appears to cause only mild PVNH symptoms in the male patient. Similar cases of rare male neonatal survival where PVNH patients carry either a missense or a distal truncating mutation have been documented [Sheen *et al*, 2001; Guerrin *et al*, 2004; Parrini *et al*, 2004; Zenker *et al*, 2004; Hehr *et al*,

2006]. This leaves several unanswered questions, importantly how a putatively lethal mutation causes only mild symptoms in the carrier. Functional experiments, which are outside the scope of the present study, are needed to clarify the pathogenic mechanism of the G2593 mutation.

BIBILIOGRAPHY

Andrews, P. (1962) Estimation of molecular weights of proteins by gel filtration *Nature* 196:36-9.

Andrews, P.(1964) Estimation of the molecular weights of proteins by Sephadex gel-filtration *Biochemistry Journal* 91:222-33.

Annesley, J., Bandala-Sanchez, E., Ahmed, U. and P.Fisher (2007) Filamin repeat segments required for photosensory signaling in *Dictyostelium discoideum*. *BMC Cell Biology* 8:48

Banuelos, S., Saraste, M. and K.Carugo (1998) Structural comparison of calponin homology domains:implications for actin binding. *Structure* 6:1419-31.

Barry, P., Xie, J., Lemmon, V. and P.Young (1993) Molecular characterization of a mutipromoter gene encoding a chicken filamin protein. *The Journal of Biological Chemistry* 268:25577-86.

Bennet, J., Zaner, K. and T.Stossel (1984) Isolation and some properties of macrophage α-actinin:Evidence that it is not an actin gelling protein. *Biochemistry* 23:5081-6.

Berry, B. *et al* (2005) FOXC1 transcriptional regulatory activity is impaired by PBX1 in a filamin A mediated manner. *Molecular Cell Biology* 25:1415-24.

Birney, E., Andrews, D., Bevan, P. and Y.Chen (2004) An overviewof Ensembl. *Genetics Research* 14:925-8.

Bresnick, R., Janmey, A. and J.Condeelis (1991) evidence that a 27 residue sequence is the actin binding site of ABP 120. *Journal of biological Chemistry* 266:12989-93.

Brotschi, A., Hartwig, H. and P.Stossel (1978) The gelation of actin by actin-binding protein. *Journal of biological Chemistry* 253:8988-93.

Condeelis, J., Salisbury, J. and K.Fijiwara (1981) A new protein that gels F-actin in the cell cortex of *Dictyostelum discoideum. Nature* 292:161-3.

Condeelis, J., Vahey, M., Carboni, M. and S.Ogihara (1984) Properties of the 120,000 and 95,000 dalton actin binding proteins from D.discoideum and their possible functions in assembling the cytoplasmic matrix. *Journal of Cell Biology* 99:119-126.

Corrado, K., Mills, L. and S.Chamberlain (1994) Deletion analysis of the dystrophin actin binding domain. *FEBS Letters* 344:255-60.

Cox, D., Risdale, A., Condeelis, D. and J.Hartwig (1995) Genetic disruption of the ABP-120 alters the 3-dimensional organization of actin filaments in *Dictyostellium* pseudopods. *Journal of Cell Biology* 128:819-35.

Cunningham, C., Kwiatkowski, J., Hartwig, H. and P.Stossel (1992) Actin binding protein requirement for cortical stability and efficient locomotion. *Science* 255:325-7.

Cunningham, C. (1995) Actin polymerization and intracellular solvent flow in cell surface blebbing. *Ournal of Cell Biology* 129: 1589-1599.

Davies, J., Wallach, D., Willingham, C., Yamaguchi, M. and S.Lewis (1980) Self association of chicken gizzard filamin and heavy merofilamins. *Biochemistry* 19:1366-72.

Diaz-Valencia, D., Barrera, J., Jay, D. and M.Vargas (2007) Novel structural and functional findings of the ehFLN protein from *Enatamoeba histolytica. Cell Motility and Cytoskeleton* 64:880-96.

Feng, Y. and A.Walsh (2004) The many faces of filamin. A versatile molecular scaffold for cell motility and signaling. *Nature Cell Biology* 6:1034-8.

Feng, Y. and A.Walsh (2004) The many faces of filamin:a versatile molecular scaffold for cell motility and signaling. *Nature Cell Biology* 6:1034-7.

Fox, W., Lamperti D., Eksioglu, Z., Hong, E., Feng, Y, Graham, A., Scheffer, E., Dobyns B., Hirsch, A., Radtke, A., Berkovic, F., Huttenlocher, R., A. Walsh (1998) Mutations in filamin 1 prevent migration of cerebral cortical neurons in human periventricular heterotropia. *Neuron* 21:1315-25.

Fucini, P., Koppel, B., Schleicher, M. *et al* (1997) Molecular architecture of the rod domain of *Dictiostylem* gelation factor. *Journal of Molecular Biology* 291:1017-23.

Fucini, P., Renner. C. and A.Holak (1997) The repeating segments of the F-actin cross-linking gelation factor have an immunoglobulin-like fold *Nature Structural Biology* 4:223-30.

Furuhashi, K., Inagaki, M. and T.Takenewa (1992) Inositol phospholipid induced suppression of F-actin gelation activity of smooth muscle actin. *Biochemical Biophysical Research Communication* 184:1261-5.

Furuike, S., Ito, T. and M.Yamazaki (2001) Mechanical unfolding of single filamin molecules detected by atomic force microscopy. *FEBS letters* 498:72-5.

Garcia, E. and D.Jay (2006) Filamina plaquetaria *Archivos Cardiologia de Mexico* 76:S4 67-75.

Gardel, L. *et al* (2006) Prestressed actin networks crosslinked by hinged filamins replicate mechanical properties of cells. *Proceedings of the National Academy of Sciences*103:1762-7.

Gimona, M. and R.Mital (1998) The single CH domain of calponin is neither sufficient nor necessary for F-actin binding. *Journal of Cell Science* 111:1813-21.

Gimona, M., Carugo, K., Kraniwitter, J. and J.Winder (2002) Functional plasticity of CH domains. *FEBS Letters* 513:98-106.

Gorlin, B., Yamin, R., Egan, S.and H.Hartwig (1990) Human endothelial actin-binding protein ABP-280 nonmuscle filamin: a molecular leaf spring. *Journal of Cell Biology* 111:1089-105.

Gorlin, B., Henske, E., Warren, T. and J.Kwaiarkowski (1993) Actin binding protein filamin gene maps telomeric to the color vision locus and centromeric to G6PD in Xq28. *Genomics* 17:496-8.

Guerrini, R., Mei, D., Sisodiya, S., Sicca, F., Harding, B., Takahashi, Y., Dorn, T., Yoshi, A., Campistol, J., Kramer, G., Moro, F., Dobyns, B. and E.Parrini (2004) Germline and mosaic mutations of FLN1 in men with periventricular heterotopias *Neurology* 63:51–6

Guo, Y., Zhang, X., Sokol, N. and L.Boulianne (2000) Physical and genetic interaction of filamin with presenilin in Drosophila. *Journal of Cell Science* 113:3499-3508.

Hammond, M. (2007) Genbank direct submission.

Hartwig, H. and P.Stossel (1981) Isolation and properties of actin, myosin and a new actin binding protein in rabbit alveolar macrophages. *Journal of Biological Chemistry* 250:5696-705.

Hehr, U., Hehr, A., Uyanik, G., Phelan, E., Winkler, J. and W. Reardon A filamin A splice

mutation resulting in a syndrome of facial dysmorphism, periventricular nodular heterotopia, and severe constipation reminiscent of cerebro-fronto-facial syndrome. *Journal of Medical Genetics* 2006;43:541–544

Himmel, M., Van derv en, F. and O. Furst (2003) The limits of promiscuity: isoform specific dimerization of filamins. *Biochemistry* 42:430-9.

Hou, L., Luby, K. & F.Lanni (1990) Brownian motion of inert tracer macromolecules in polymerized and spontaneously bundled mixtures of actin filaments. *Journal of Cell Biology* 110:1645-54.

Irvine, B. (2001) Determination of molecular size by size exclusion chromatography. *Current Protocols in Cell Biology*; Chapter 5:Unit 5.5.

Ito *et al* (1992) Regulation of water flow by actin binding protein induced acting elation. *Biophysics Journal* 61:1301-5.

Khaire, N., Muller, R., Schleicher, M., Rief, M. and A.Noegel (2007) Filamin regulated F-actin assembly is essential for morphogenesis and controls phototaxis in *Dictyostelium. The Journal of Biological Chemistry* 262:1948-54.

Kiema, T. *et al* (2006) The molecular basis of filamin binding to integrins and competition with talin. *Molecular Cell* 21:337-47.

Korenbaum, E. and F. Rivero (2002) Calponin homology domains at a glance. *Journal of Cell Science* 115:3543-5.

Le Maire, M., Ghazi, A., Martin, M. and F. Brochard (1989) Calibration curves for size exclusion chromatography. *Journal of Biochemistry* 106:814-7

Langanger, G., DeMay, J., Daneels, G. and V.Small (1984) Ultrastructural localization of alfa actinin and filamin in cultured cells with the immunogold staining method. *Journal of Cell Biology* 99:1324-34.

Lebart, M.C. *et al* (1993) Further characterization of the actinin:actin interface and comparison with filamin binding sites on actin. *Journal of Biological chemistry* 268:5642-8.

Lebart, M.C. *et al* (1994) Characterisation of the actin binding site on smooth muscle filamin. *Journal of Biological chemistry* 269:4279-84.

Levine, A., Moir, G. and V.Perry (1992) Binding sites involved in the interaction of actin with the N-terminal region of dystrophin. *FEBS Letters* 298:44-8.

Li, G., Serr, M., Edwards, K., *et al* (1999) Filamin is required for ring canal assembly and actin organization during Drosophila oogenesis. *Journal of cell Biology* 146:1061-74.

Liu, G., Thomas, L., Warren, A. and C.Cunningham (1997) Cytoskeleton protein ABP-280 directs the intracellular trafficking of furin and modulates proprotein processing in the endocytic pathway. *Journal of Cell Biology* 139:1719-33.

Loy, J., Sim, S. and L.Yong (2002) Filamin A fragment localizes to the nucleus to regulate androgen receptor and coactivator functions. *Proceedings of the National Academy of Sciences* 100:4562-7.

Lu, J., Lian, G., Lenkinski, R. and V.Sheen (2007) *Human Molecular Genetics* 16:1661-75.

Löwe, T.*et al* (2007) The pathomechanism of filaminopathy: altered biochemical properties explain the cellular phenotype of a protein aggregation myopathy. *Hum Mol Geneics* 1:1351-8.

Maestrini, E., Patrosso, C., Rocchi, M. *et al* (1993) Mapping of two genes encoding isoforms of the actin binding protein ABP-280, a dystrophin-like protein to Xq28 and to chromosome 7. *Human Molecular Genetics* 2:761-6.

Matthews, Nicholson, Becktel, *et al* 1987 Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfoldingProc of Natio academy of sciences84(19): 6663–6667.

Matsudaira, P. (1991) Modular organization of actin crosslinking proteins. *Trends in Biochemical Science* 16:87-92.

McGough, A. (1998) F-actin binding proteins. *Current Opinions in Structural Biology* 8:166-76.

Mejean, C., Lebart, M. and Y.Benyamin (1992) Localization and identification of actin structures involved in the filamin-actin interaction. *European Journal of Biochemistry* 209:555-62.

Meng, X. *et al* , (2004) Recovery from DNA damage induced G2 arrest requires actin binding protein filamin A. *Journal of Biological Chemistry* 279:6098-105.

McCoy, J., Grosse-Kunstleve, W., Adams, D., Winn, D., Storoni, C. and J. Read (2007) *Phaser* crystallographic software. *Journal of Applied Crystystallography* 40:658-74

Morris, E., Man, T., Huyen, N. and J.Winder (1999) Disruption of the utrophin-actin interaction by monoclonal antibodies and prediction of an actin binding surface of utrophin. *Biochemical Journal* 337:119-23.

Murzin, G., Brenner, E., Hubbard, T. and C.Chothia (1995) SCOP: A structural classification of protein database for the investigation of sequences and structures. *Journal of Molecular Biology* 247:536-40.

Nagano, T., Yoneda, T., Hatanaka, Y. and M.Sato (2002) Filamin A interacting protein regulates cortical cell migration out of the ventricular zone. *Nature Cell Biology* 4:495-501.

Nakamura, F., *et al* (2005) Ca++ and calmodulin regulate the binding of filaminA to actin filaments. *Journal of Biological Chemistry* 280:32426-33.

Nakamura, F., *et al* (2006) The structure of the GPIb-filamin A complex. *Blood* 107:1925-32.

Niederman, R., Amrein, c. and H.Haertwig (1983) Three dimensional structure of actin filaments and of and actin gelmade with actin-binding protein. *Journal of cell biology* 96:1400-13.

Noegel, A., Rapp, S., Schleicher, M. and M.Stewart (1989) The Dictystelium gelation factor shares a putative active binding site with alpha-actinins and dystrophin and also has a rod domain containing six 100 residue motifs that appear to have a cross-beta conformation. *Journal of Cell Biology* 109:607-18.

Norwood, M., Sutherland-Smith, J., Keep, H. and J.Kendrick-Jones (2000) The structure of the N-terminal actin binding domain of human dystrophin and how mutations in this domain may cause Duchenne or Becker muscular dystrophy. *Structure with Folding and Design* 8:481-91.

Nunnally, H., Dangelo, M. and W.Craig (1980) Filamin concentration in cleavage furrow and midbody region: frequency of occurrence compared with that of alpha actinin and myosin. *Journal of Cell Biology 87:219-26.*

Ohashi, K., Oshima, K., Tachikawa, M. and G.Terasaki (2005) Chicken gizzard filamin, retina filamin and cgABP260 are respectively smooth muscle-, nonmusclep and pan muscle type isoforms:distribution and localization in muscles. *Cell motility and cytoskeleton* 61:214-25.

Ohta, Y., Stossel, P. and H.Hartwig (1991) Ligand sensitive binding of actin binding protein to immunoglobulin G Fc receptor 1. *Cell* 67:275-82.

Ozanne, M. *et al* (2000) Androgen receptor nuclear translocation is facilitated by the F-actin crosslinking protein filamin. *Molecular Endocrinology* 14:1618-26.

Parrini, E., Mei, D,, Wright, M., Dorn, T. and R.Guerrini (2004) Mosaic mutations of the FLN1 gene cause a mild phenotype in patients with periventricular heterotopias. *Neurogenetics* 5:191–6.

Patrosso, C., Repetto, M., Frattini, A. *et al* (1994) The exon-intron organization of of the human X-linked gene FLN1 encoding actin binding protein 280. *Genomics* 21:71-6.

Pollard, D. and A.Cooper (1986) Actin and actin binding proteins. A critical evaluation of mechanisms and functions. *Annual review of Biochemistry* 55:987-1035.

Popowicz, M., Muller, R., Noegel, A. and A.Holak (2004) Molecular structure of the rod domain of Dictyostelium filamin.

Popowicz, G., Schleicher, M., Noegel, A. and T.A.Holak (2006) Filamins: promiscuous organizers of the cytoskeleton. *Trends in Biochemical sciences* 31:7:411-9.

Pudas, R., Kiema, T., Butler, T. and J.Ylanne (2005) Structural basis for vertebrate filamin dimerization. *Structure* 13:111-9.

Pudas, R. (2007) Structural and interaction studies on the carboxy-terminus of filamin, an actin binding protein. *A scientiae Rerum Naturalium* 477, Acta Universitatis ouluensis.

Puius, A., Mahoney, M and C.Almo (1998) The modular structure of actin regulatory proteins. *Current opinion in cell biology* 10:23-34.

Robertson, P., Twigg, R., Sutherland-Smith, J. *et al* (2003) Localized mutations in the gene encoding the cytoskeletal filamin A causes diverse malformations in humans. *Nature Genetics* 33:487-91.

Robertson, P. (2005) Filamin A:Phenotypic diversity *Current Opinions in Genetics and Development* 15:301-7.

Rus, F., Kurucz, E., Markus, R. and I.Ando (2006) Expression pattern of Filamin-240 in *Drosophila* blood cells. *Gene expression patterns* 6:928-34.

Rychlik, W., Spencer, J. and E. Rhoads (1991) Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic Acids Res* 19:698

Seo**,** M**.,** Seok**,** S**.**, Im**,** H**.**, Kwon**,** R**.**, Lee**,** J**.**, Kim**,** R**.**, Cho**,** Y**.**, Park**,** D**.** and J.Lee (2009) Crystal structure of the dimerization domain of human filamin A. *Proteins* 75:258-63

Sheen, L. *et al*, (2001) Mutations in X-linled filamin 1 gene cause periventricular nodular heterotropia in males as well as in females. *Human Molecular Genetics* 10:1775-83.

Sheen, L., Graham, D., Takafutu, S. and A.Walsh (2002) Filamin A and B are coexpressed with neurons during periods of neuronal migration and can physically interact. *Human Molecular Genetics* 11:2845-54.

Sheen, L., Jansen, A., Chen, H. *et al* (2005) Filamin A mutations cause periventricular heterotopias with Ehlers-Danlos syndrome. *Neurology* 64:254-62.

Shizuta, Y., Gallo, M., Davies, P. and I.Pastan (1976) Purification and properties of filamin, an actin binding protein from chicken gizzard. *Journal of Biological Chemistry* 251:6562-7.

Small, V., Stradal, T. and K.Rottner (2002) The lamellipodium: where motility begins. *Trends in Cell Biology* 12:112-20.

Sokol, N. and L.Cooley (1999) Drosophila filamin encoded by the cheerio locus is a component of ovarian ring canals. *Current Biology* 9:1221-30.

Staros, V. (1988) Membrane impermanent cross-linking reagents: probes of the structure and dynamics of membrane proteins. *Accounts of chemical research* 21:435-441.

Stefanova, M. *et al* (2005) A novel 9 bp deletion in filamin A gene causes an otopalotodigital-spectrum disorder with a variable intermediate phenotype. *American Journal of Medical Genetics* 132:386-90.

Stossel, T. and J.Hartwig (1975) Interaction of actin, myosin and an actin binding protein of rabbit alveolar macrophages:Macrophage myosin Mg++ adenosine triphosphatase requires a cofactor for activation by actin. *Journal of Biological Chemistry* 250:5706-12.

Stossel, P., Condeelis, J., Cooley, L., Hartwig, J., Noegell, A., Schleiker, M. and S.Shapiro (2001) Filamins as integraters of cell mechanics and signaling. *Nature Reviews* 2:138-45.

Stossel, P. and H.Hartwig (2003) Filling gaps in signaling to actin cytoskeletal remodeling. *Developmental Cell* 444-5.

Stossel, P. (2010) Filamins and the potential complexity. *Cell Cycle* 10:1463.

Stradal, T., Kranewritter, W., Winder, J. and M.Gimona (1998) CH domains revisited. *FEBS letters* 431:134-7.

Tachikawa, M., Nakagawa, H., Terasaki, G. and K.Ogashi (1997) A 260 kDa filamin related protein in chicken gizzard smooth muscle cells is a new component of the dense plaques and dense bodies of smooth muscles. *Journal of Biochemistry* 122:314-21.

Takafuta, T., Higgins, G. and S.Shapiro (1998) Human beta-filamin is a new protein that interacts with the cytoplasmic tail of Ib-alpha. *Journal of Biological Chemistry* 278:17531-8.

Tyler, J., Anderson, J. and D.Brenton (1980) Structural comparison of several actin binding macromolecules. *Journal opf cell Biology* 85:489-95.

Vadlamudi, K., Li, F., Adam, L. *et al* (2002) Filamin is essential in actin cytoskeletal assembly mediated by p21-activated kinase 1. *Nature Cell Biology* 4:681-90.

Van der Flier, A. and A.Sonnenberg (2001) Structural and functional aspects of filamins. *Biochimica Biophysica acta* 1538:99-117.

Van der Ven, M., Oberman, J., Lemke, B. and O.Furst (2000) Characterization of muscle filamin isoforms suggests a possible role for gamma filamin/APB-L in sarcomeric Z disk formation. *Cell Motility Cytoskelton* 45:149-62.

Vargas, M., Sansonnetti, P. and N.Guillen (1996) Identification and cellular localization of the actin binding protein ABP-120 from *Enatamoeba histolyca*. *Molecular Microbiology* 22:849-57.

Wallach, D., Davies, J. and I.Pastan (1978) Cyclic AMP dependant phosphorylation of filamin in mammalian smooth muscles. *Journal of Biological Chemistry* 253:4739-45.

Wang, K., Ash, J. and J.Singer (1975) Filamin a new high molecular weight protein found in smooth muscle and nonmuscle cells. *Proceedings of the national academy of sciences* 72:4483-6.

Wilson, R., Ainscough, K., Anderson, C., Cooper, J. *et al* (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.elegans. Nature* 368:32-8.

Winder, J. and R.Ayscough (2005) Actin binding proteins. *Journal of Cell Science* 118:651-4.

Woo, S., Ohta, Y., Stossel, P. and J.Blenis (2004) Ribosomal S6 Kinase regulates phosphorylation of filamin A on an important regulatory site. *Molecular Cell Biology* 24:3025-35.

Xie, Z., Xu, W., Davie, W. and W.chung (1998) Molecular cloning of human APBL an actin binding protein homolog. *Biochemichal Biophysical Research Communication* 251:914-9.

Xu, W., Xie, Z., Chung, W. and W.Davie (1998) A novel human actin binding protein that binds to platelet glycoprotein Iba. *Blood* 92:1268-76.

Yoshida, N. *et al* (2005) FilaminA bound PEBP2b/CBFb is retained in the cytoplasm and prevented from functioning as a partner of the RunX1 transcription factor. *Molecular Cell Biology* 25:1003-12.

Yuan, Y. and Z.Shen (2001) Interaction with BRAC2 suggests a role for filamin 1 in DNA damage response. *Journal of Biological Chemistry* 276:48318-24.

Zenker *et al* (2004) A Dual Phenotype of Periventricular Nodular Heterotopia and fronto -metaphyseal Dysplasia in One Patient Caused by a Single *FLNA*Mutation Leading to Two Functionally Different Aberrant Transcripts *Am. J. Hum. Genet. 74:731–737, 2004*

.