The mouse filensin gene: structure and evolutionary relation to other intermediate filament genes

Fotini Gounari[a,*], Niki Karagianni[a], Antoaneta Mincheva[b], Peter Lichter[b], Spyros D. Georgatos[c], Volker Schirrmacher[a]

[a] *Abteilung 710, Schwerpunkt Tumorimmunologie, Deutsches Krebsforschungzentrum, Im Neuenheimerfeld 280, Heidelberg 69120, Germany*
[b] *Abteilung 845, Organization Komplexer Genome, Deutsches Krebsforschungzentrum, Im Neuenheimerfeld 280, Heidelberg 69120, Germany*
[c] *Department of basic Sciences, Division of Medicine, School of Health Sciences, University of Crete, P.O. Box 1393, 71110 Heraklion, Crete, Greece*

Received 5 June 1997; revised version received 14 July 1997

Abstract Filensin and phakinin are two lens-specific members of the intermediate filament (IF) superfamily of proteins. They coassemble to form a beaded submembraneous filamentous network, the beaded filaments (BFs). The low sequence homology and differences in assembly compared to other IF proteins do not allow their classification in any of the five IF subgroups. The organization of the phakinin gene exon/intron boundaries provides evidence that this partner may be sharing a common origin with type I cytokeratin genes. Here we report the molecular cloning, sequence and characterization of the mouse filensin gene. The filensin gene consists of 8 exons and 7 introns, with 6 introns interrupting its rod domain in a highly conserved manner characteristic of type III IF genes, like vimentin, desmin, or peripherin. Of the two tail domain exons the one adjacent to the rod domain, compares to exon 7 of the non-neuronal cytoplasmic IF gene of helix aspersa and to the lamin region bridging the end of the rod domain to the nuclear localization signal. Altogether, these observations indicate that the lens beaded filaments form an independent class of IF.

Key words: Intermediate filaments; Beaded filaments; Mouse filensin gene

1. Introduction

The cytoskeleton of the differentiated lens fiber cells (LFCs) consists of several ubiquitous components as well as a lens-specific submembraneous beaded filamentous network, designated as beaded filaments (BFs). BFs are morphologically distinct from the other filamentous structures present in the lens. Earlier biochemical studies showed that they consist of two protein subunits, a heavier [CP94 (chicken), CP95 (rat), CP115 (bovine)] and a lighter [CP49 (chicken) CP47 (bovine)] [1–5]. Both subunits of the beaded filaments filensin (formerly CP94, CP95, or CP115) and phakinin (formerly CP49 or CP47) were cloned and shown to fulfil the structural principles of intermediate filament proteins [6–12]. However, they also possess several peculiarities which clearly distinguish them from ordinary IF subunits. Thus, filensin has a naturally 'truncated' rod domain, due to a 29 residue deletion in the coil 2 region and it possesses a very long COOH-terminal tail domain (438 residues) [6–8]. Filensin seems to be involved in interactions with components of the lens plasma membrane [13] as well as vimentin [14]. Phakinin, filensin's partner, contains a 114 residue NH_2-terminal head domain, a 'classical' rod domain of 311 amino acids, but completely lacks a COOH-terminal tail domain.

The peculiarities of filensin and phakinin triggered studies concerned with the architecture of beaded filaments, and the mechanism of partner recognition [15,16]. In vitro polymerization studies combined with scanning transmission electron microscopy and mass per length measurements shed light into the molecular arrangement of the BF. The filament is proposed to consist of two, topologically and compositionally distinct filament moieties: (a) an inner filament built of 8 phakinin homodimers; and (b) a peripherally-disposed subfilament consisting of 8 filensin/phakinin heterodimers representing the beads on the filaments [15]. The bead-like morphology of the peripherally-disposed filensin/phakinin oligomers was attributed to the long COOH-terminal domain of filensin which, similar to the COOH-terminal tails of the lamins, may have a compact, globular structure [17]. The presence of filensin in the filament is thought to prevent the unravelling of core phakinin protofilaments and to attenuate the tendency of phakinin to grow laterally into thick bundles [15]. The overall, molecular architecture of the beaded filaments may be paralleled to that of neurofilaments which are also thought to contain an NF-L core and a peripheral shell consisting of NF-L/NF-M and NF-L/NF-H heterotetramers [18]. However, instead of having a globular COOH-terminal tail as with filensin, NF-M and NF-H have extended COOH-terminal tail domains ('sidearms') [19–21]. Both neurofilaments and beaded filaments differ from the homogeneous keratin filaments assembled by type I and II obligate heterodimers, and the homopolymeric type III filaments.

In contrast to filensin, which shows diffuse homology to a variety of IF proteins [7,8], phakinin possesses a clear structural kinship to type I cytokeratins [9–12]. However, both the low level of sequence identity and the lack of complementation with other cytokeratins in in vitro polymerization assays [10] complicates the immediate classification of phakinin into the cytokeratin group. The problem of low level primary sequence similarity between the different IF subgroups is one of the characteristics of this rather diverge protein family. This phenomenon appears exaggerated in the case of phakinin and filensin. Analysis of the gene organization has proved instrumental in establishing the evolutionary relation between the different groups of IF proteins (reviewed by [22,23]). The human gene encoding for phakinin has recently been cloned and shown to have an identical intron/exon organization with the type I cytokeratin genes [12]. This suggests that phakinin shares the same evolutionary origin with the type I cytokeratin genes. It becomes now even more triggering to

*Corresponding author. Fax: (49) (6221) 423702.
E-mail: f.gounari@dkfz-heidelberg.de

PII S0014-5793(97)00937-X

establish the evolutionary origin of the partner of phakinin, filensin. In an attempt to address more rigorously the relation of the beaded filaments to the other IFs we proceeded to the molecular cloning and characterization of the mouse filensin gene.

2. Materials and methods

2.1. RT-PCR amplification

Total lens RNA was isolated from lenses of newborn mice. Typically 8 to 10 newborn mice were sacrificed and their lenses isolated. Total lens RNA was prepared using the acid phenol method described in [24]. After controling the quality, 1 μg of total RNA was reverse transcribed. Reverse transcription was carried out with the SuperScript Preamplification System for first strand cDNA synthesis from GibcoBRL (Gaithersburg) using either oligo dT or the 'ratr4' filensin-specific primer (CGATGATACCGGTCTAGCTCGTTCCTCAGGGTCTGCT) to prime the reaction. The first strand cDNA was PCR amplified using the 'ratr4' downstream primer in combination with the 'ratd1' upstream primer (GGCTCGACGAGTTCCGCAGCAAATACGAAAATGAGTGTG). Since no obvious specific amplification product was observed with this first PCR reaction, a second amplification step was applied using the nested upstream primer 'ratd2' (GTGAATGTCAATTGGTGCTAAAGGAAATGCTGGAAC) in combination with the downstream primer 'ratr3' (CCTTCCGGAGGTTTTCGATCTGCCCATTATAAAGCTG). The 433-bp mouse-specific amplification product was subcloned into the pGEM vector and characterized.

2.2. Screening of the genomic library

A total of 10^6 plaques of a lambda FixII mouse 129SV library purchased from Stratagene (La Jolla) were plated with LE392 cells onto five 22×22 cm bacteriological plates according to standard procedures. Two replica Hybond Amersham (Buckinghamshire) nylon membrane filters were lifted from these plates. Probes of the 433-bp mouse cDNA-specific probe were radiolabeled with ^{32}P using the standard random priming procedure to approximate specific activities of 10^9 cpm/μg. The hybridization was carried out according to [25] in a volume of 10 ml per filter and a probe concentration of 3×10^6 cpm/ml. Positives were considered the clones that generated signal in both replica membranes. The initial isolates were purified by three cycles of plating and hybridization. In the last cycle of screening the phage lysate represented a pure population of the positive lambda isolate.

2.3. Southern and Northern blot analysis

Tail genomic DNA of 129SV mice was used in Southern blots. Approximately 10 μg of DNA were independently digested with restriction enzymes and run in 0.8% agarose gels. Gels were treated and blotted to Hybond membranes Amersham (Buckinghamshire) according to [26]. For Northern blots, 10 μg of total lens RNA prepared as described above were electrophoresed in 1% agarose gels which were treated and blotted according to [26]. Both Southern and Northern blots were hybridized under the conditions described in [25], using ^{32}P-radiolabeled probes. A 2-kb *Hin*dIII fragment comprising 750 bp of upstream sequences, the first filensin exon and 750 bp of the first intron was used for the Southern blots and a 433-bp mouse-specific cDNA fragment was used for the Northern blots.

2.4. Fluorescence in situ hybridization

For fluorescence in situ hybridization the mouse filensin bacteriophage clone 18 consisting of a 16-kb genomic DNA fragment containing coding information for the filensin gene was digoxigenin-labeled and used as a probe for chromosomal in situ hybridization as described [27]. Mouse metaphase cell spreads were prepared from the spleen cells of a female Balb/c animal following an established protocol [28]. Seventy ng of digoxigenin-labeled probe DNA were combined with 3 μg of mouse Cot 1-DNA and 8 μg salmon sperm DNA in a 10 μl hybridization cocktail, and hybridized to mouse metaphase chromosomes. Specific hybridization was detected in the rhodamine channel, while chromosomes were banded by DAPI. Digitized images were acquired separately for DAPI and rhodamine fluorescence, using a cooled CCD camera (Photometrics). These images were aligned and electronically overlayed using the NIH image 1.41 software. Pictures were taken directly from the monitor.

2.5. Sequence analysis and multiple alignment

Database searches were performed with FASTA version 1.6c [29]. The SWISSPROT data base release 34 was used. A multiple alignment of the 4 filensin protein sequences was constructed using the PILEUP of the GCG sequence analysis package version 7.1 [30]. The default parameters for gap open (3.0) and gap elongation (0.1) were used. The amino acid comparison matrix was the normalized Dayhoff matrix [31].

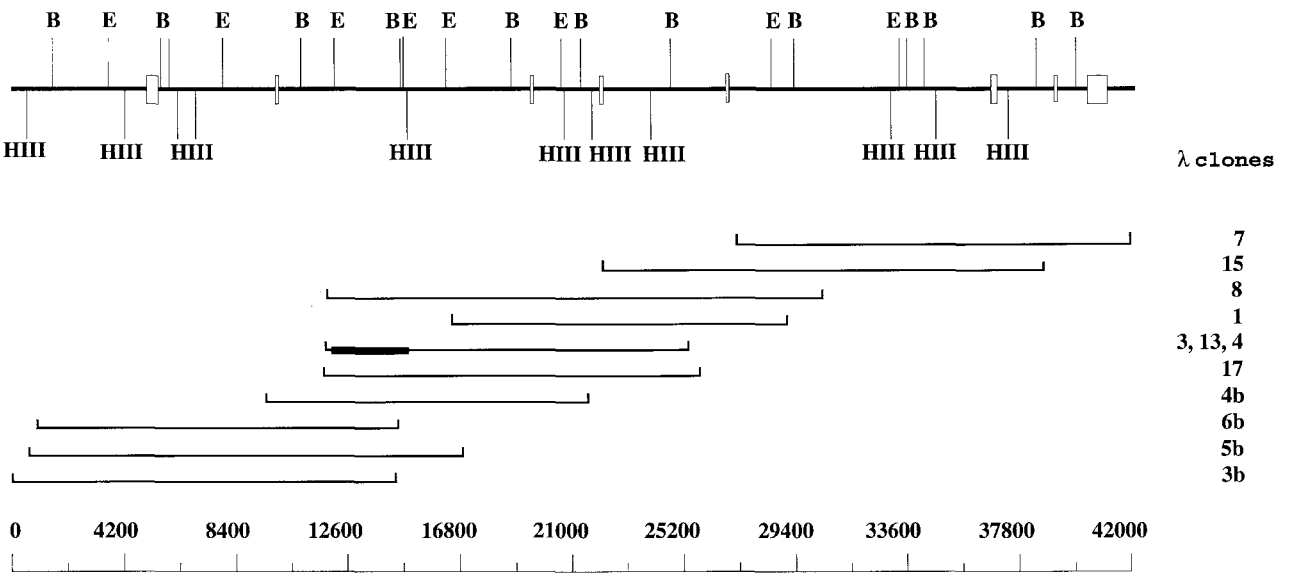

Fig. 1. Cloning and chromosomal organization of the mouse filensin gene. The 42-kb mouse filensin gene is diagrammatically represented as a line. The 8 exons are indicated by empty boxes, the width of the boxes corresponds to the relative size of the exons. The restriction map for *Eco*RI (E), *Hin*dIII (HIII) and *Bam*HI (B) is also indicated in the diagram. Bellow the gene representation, the individual lambda isolates and their relative location with respect to the gene are indicated. The thicker lane in lambda clones 3, 13 and 4, indicates the probe used for chromosome walking to obtain the 5′ part of the filensin gene.

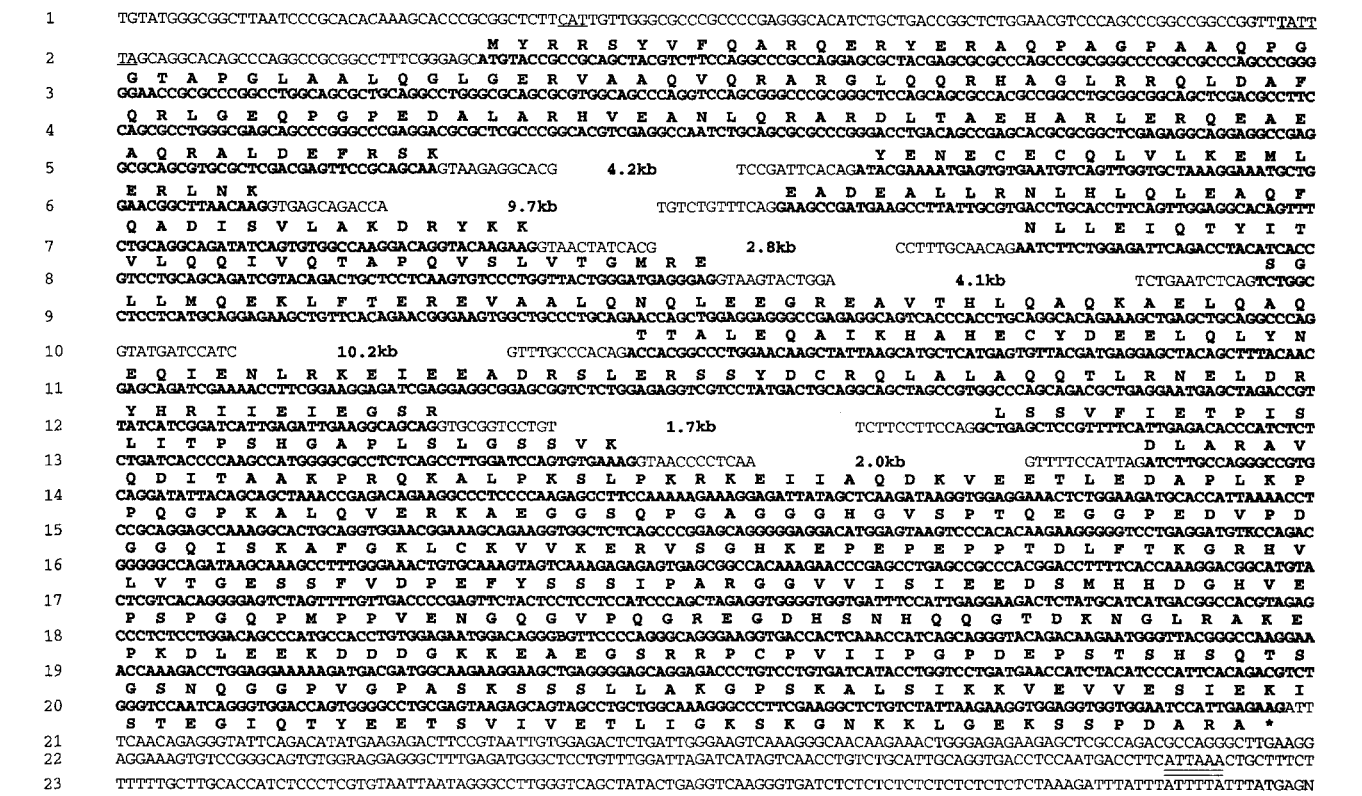

Fig. 2. Sequence of the mouse filensin gene. Documentation of approximately 17 kb of sequencing information from the mouse filensin gene is available from the EMBL data bank under the accession numbers Y13602, Y13603, Y13604, Y13605, Y13606. In bold is the cDNA sequence, the predicted protein sequence is written above the cDNA in the single letter code. The coding amino acid is placed above the second codon position. The size of the intronic sequences is indicated. Upstream the translation initiation potential TATA and CAT motifs are underlined. Double underlined is the potential polyadenylation signal downstream the termination codon which is indicated with a star.

3. Results

3.1. Isolation of a mouse filensin gene

To clone the mouse filensin gene we first generated by means of RT-PCR amplification a mouse filensin cDNA probe encoding for the second half of the coil 1b to the helix 'stutter' present in the coil2 of the rod domain. RT-PCR was carried out on RNA prepared from lenses of newborn mice using two pairs of 36 bp long nested primers (see Section 2). The primers were derived from the rat filensin sequence [6] and they were selected on the basis of their high degree of cross species identity. The amplified 433-bp mouse cDNA fragment was sequenced to establish its identity and used as a probe in the screening of a mouse genomic library. Screening of 10^6 plaques of a λ FixII 129SV mouse genomic library (Stratagene) with this probe led to the isolation of 14 overlapping λ clones spanning a continuous fragment of approximately 29 kb. The coding regions within this fragment were characterized by restriction mapping, hybridization with short cDNA fragments and extensive sequencing. The 29-kb fragment contained coding information from the middle of coil 1b to the termination codon. To identify the 5′ of the mouse filensin gene one step chromosome walking was performed (Fig. 1). A 2.2-kb *Eco*RI fragment representing the most upstream part of the already characterized filensin gene sequences was used as a probe in a second screening of the genomic library. Four additional overlapping lambda clones isolated this way contained the rest of the filensin gene including approximately 5 kb upstream the initiation of transcription. The two screening steps allowed the characterization of a continuous 42-kb genomic DNA fragment spanning the entire filensin locus. From the 42-kb locus a total of 17 kb were sequenced and the sequence information is reported in Fig. 2. Of the 17-kb sequence approximately 5 kb including all the coding sequences and the intron exon barriers were sequenced from both DNA strands.

To determine the number of filensin genes present in the mouse genome we performed Southern blot analysis of mouse DNA. The blot was hybridized with a 2-kb *Hin*dIII fragment comprising 750 bp of upstream sequences, the first filensin exon and 750 bp of the first intron. This fragment was completely sequenced and shown not to contain repetitive sequences. Genomic DNA for the blots was digested with enzymes (*Bam*HI, *Eco*RI and *Hin*dIII) already mapped in the cloned DNA (Fig. 3A). The detection of single 4.2-kb, and 2-kb fragments in the *Eco*RI and *Hin*dIII lanes respectively and of a 3.8-kb double band in the *Bam*HI lane indicates that the mouse filensin is encoded by a single copy gene. Fluorescent in situ hybridization was used to localize the filensin gene (see Section 2) in the band G of the mouse chromosome 2 (Fig. 3B). This agrees with the localization previously reported by [32] based on Southern blot analyses of DNA from hybrid cells.

3.2. Mouse filensin cDNA

The characterized mouse filensin gene consists of 8 exons and 7 introns. The exons were predicted on the basis of their identity with the mouse cDNA fragment and their similarity

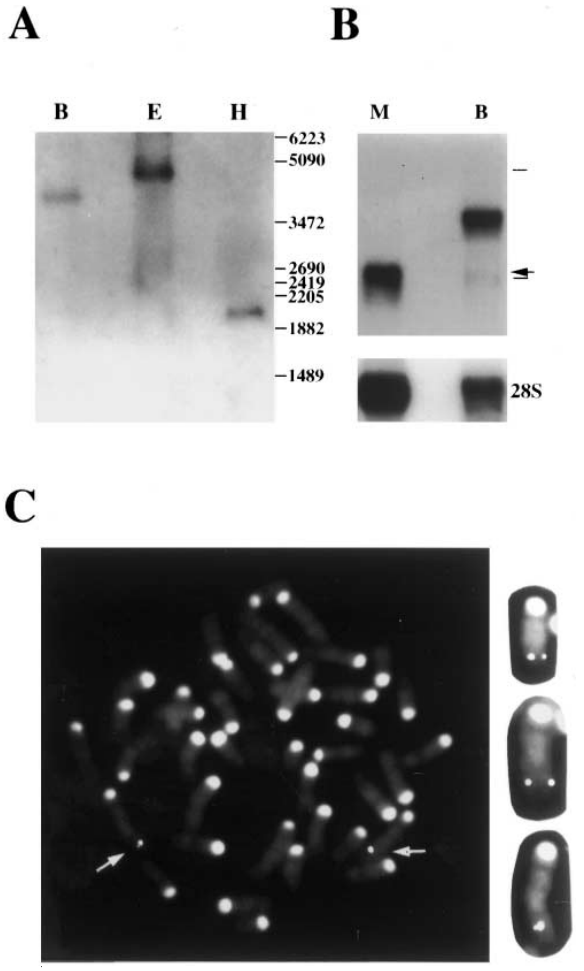

Fig. 3. (A) Southern Blot of the mouse filensin gene. Genomic DNA prepared from tails of 129SV mice was digested with *Bam*HI (lane B), *Eco*RI (lane E) and *Hin*dIII (lane H), and hybridized with radiolabeled mouse filensin probe. The sizes of the molecular weight markers are indicated on the side of the picture. (B) Northern Blot of mouse (lane M) and bovine (lane B) filensin transcripts. Blots were hybridized against a mouse cDNA-specific probe. Bars on the side of the picture indicate the position of the 28 and 16S ribosomal RNAs. The arrow depicts the position of the mouse mRNA. Control hybridization with a mouse 28S probe is shown in the panel bellow the Northern. (C) FISH chromosomal detection of the mouse filensin gene. Spread of a mouse metaphase cell after FISH with the digoxigenin-labeled probe 18, detected via rhodamine. Arrows indicate the fluorescence signals on both chromosome 2 homologs within bond G (left panel); the right panels show further examples of hybridized chromosome 2 homologs.

with the rat bovine and chicken cDNAs[6–8]. All exon/intron barriers (Fig. 5) correspond to the consensus 5′ and 3′ splice sites [33]. The 8 exons assemble an open reading frame of 2007 bp encoding for 669 amino acids. The mouse filensin coding sequence shows 90%, 75% and 66% sequence identity with the coding sequences of the rat, bovine and chicken clones respectively. A potential polyadenylation signal ATTAAA was localized 112 bp downstream the termination codon [34,35]. Thus, the mouse filensin transcript as defined by its ORF and 3′ untranslated region should be at least 2122 bp long. This estimate is in good correspondence with the single 2.2-kb band detected by northern blot analysis of total mouse lens RNA Fig. 3B. Comparison of the predicted and the apparent transcript length indicates that the mouse filensin transcript may contain a short 5′ untranslated region. Indeed a potential TATA-like element is localized at position −40 from the ATG (Fig. 2). CAAT-like motifs can be seen 45 and 72 bp upstream the potential TATA element. The mouse filensin transcript is significantly shorter than its bovine counterpart (Fig. 3B).

3.3. *Amino acid sequence and domain structure of the mouse filensin*

Computer assisted analysis and comparison to the already characterized filensin clones established the tripartite organization of the mouse filensin protein. A multiple alignment of the different filensin protein sequences is shown in Fig. 4. The 33 amino acid long head domain shows, 57% and 33% sequence identity as compared to the bovine [7] and chicken [8] counterparts respectively. This domain is for the mouse clone 5 amino acids shorter than for the bovine and chicken. The sequence identity between the different clones clusters mainly in the first 18 amino acids. Similarly to the bovine and chicken, the mouse filensin contains the conserved di-arginine/aromatic residue motif (YRRSY) found with some perturbations in the head domains of vimentin, desmin, peripherin, a-internexin and other IF proteins [36].

The middle rod domain of the mouse clone shows 90% (over a length of 715 bp), 78% and 59% sequence identity to the rat, bovine and chicken rod domains respectively. The mouse rod domain encompasses 285 amino acids and it is 5 amino acids longer than the bovine and chicken rod domains. This can be attributed to an insertion in the beginning of the coil 2. Interestingly the position of this insertion exactly coincides with the 29 amino acid deletion in the rod domain of the other filensin isolates and with the break point between exons 4 and 5 (see Section 4). Areas of cross species conservation are evident between the mouse and the more distant bovine and chicken clones. For example a high degree of conservation is seen in the second half of the coil1b as well as in the carboxy-terminal part of coil 2. However, the generally conserved in IF proteins coil1a region shows less sequence identity than expected. A helix 'stutter' can be identified at amino acid position 258 in the coil 2 region of the mouse filensin rod domain.

The mouse filensin tail domain is encoded by 351 amino acids. Similarly to the rat and chicken clones the mouse filensin does not possess the neurofilament-like repeats present in the bovine clone. This peculiar insertion in the bovine sequence may be the result of exon shuffling (see Section 4).

3.4. *Mouse filensin gene resembles to the genes of type III and the nn helix aspersa cytoplasmic IF proteins*

Of the seven filensin introns six interrupt the sequences coding for the helical rod domain (Fig. 4). Fig. 6 compares the location of the filensin introns to that of other IF genes. The first 4.2-kb long intron of filensin interrupts the coding sequence at the second position of the codon for amino acid 119. This exactly coincides with the first intron of type III IF genes [37–39]. The absence of introns more amino-terminally differentiates filensin from the keratin type I and II genes [40,41]. The following introns 2 (9.7 kb) and 3 (2.8 kb) occur after the third codon position of amino acids 139 and 171 respectively. The location of the intron 2 interruption is highly conserved throughout the IF family of proteins with the ex-

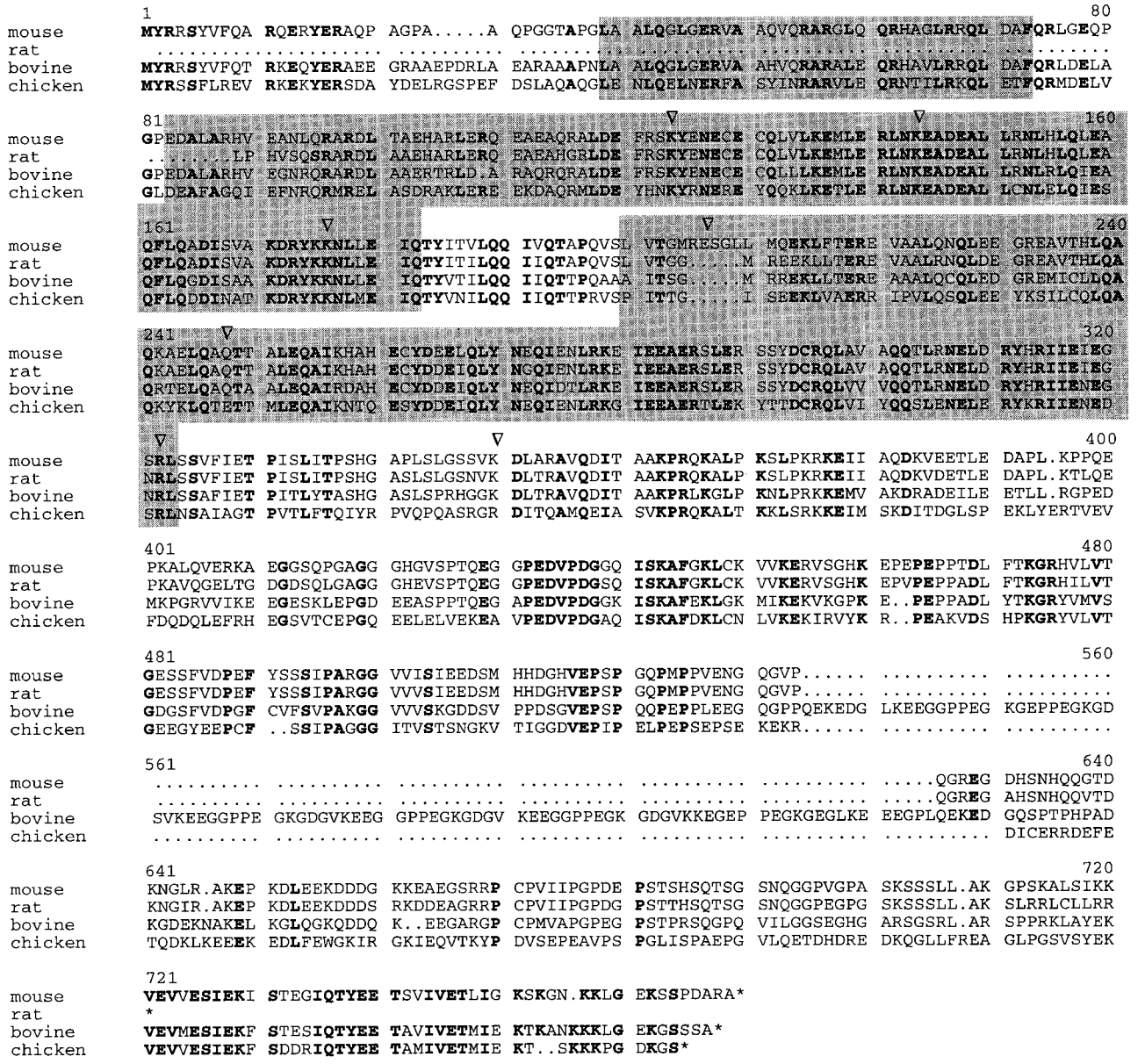

Fig. 4. Multiple alignment of the filensin protein sequence from various species. Computer assisted alignment of the predicted mouse filensin protein sequence to the already reported rat [6] bovine [7] and chicken [8] sequences. Identical amino acids in all four sequences are marked bold. The gray boxes represent the three rod helical subdomains coil1A, coil1B and coil2. Arrows over the alignment indicate positions of intron interruptions.

ception of neurofilaments [42,43]. The placing of intron 3 which is absent in type I keratins and neurofilaments provides additional evidence against a type I origin of the filensin genes. Interestingly the 4.1 kb long intron 4 starts after amino acid 201 and coincides with the 29 amino acid deletion in the rat, bovine, and chicken filensin rod domains. The equivalent rod domain deletion in the mouse clone corresponds to 24 amino acids. Mouse filensin intron 4 is placed in the exact area and codon phase with intron 4 present in all vertebrate IF genes except human keratin 8. Intron 5 with a length of 10.2 kb interrupts the coding sequence at the third codon position of amino acid 243, conserved between vertebrate IF genes. The 1.7 kb long intron 6 starts in the second codon position of the penultimate amino acid of the rod domain. Overall the intron exon organization in the rod domain of mouse filensin gene resembles this of the type III IF genes. The 2-kb long intron 7 is located 27 amino acids carboxy-terminal to the helical rod domain at the first position of the corresponding codon. The boundaries of this exon do not resemble other vertebrate IF gene. Interestingly both the length of this exon (86 bp) and the frame of the interruption compare precisely to the equivalent exon of the non-neuronal helix aspersa IF gene [44] (see Section 4). Comparison of the filensin exon 7 with the amino acid sequences of 80 IF proteins aligns this exon to the equivalent region of lamins. In this alignment the filensin exon 7 is localized exactly in the space between the end of the rod domain and the nuclear localization signal present in the lamins [45]. Finally, the last exon encodes without interruption the rest of the tail domain down to the termination codon.

Exon/Intron organization of the mouse filensin gene

Exon/Intron number	Exon size (in bp)	Sequence at exon/intron junction		
		5′ Splice donor	Intron size (in kb)	3′ Splice acceptor
1	356	TTC CGC AGC *Ly* AA gtaagaggcacg	....4.2....	tccgattcacag *s* A TAC GAA AAT
2	61	CGG CTT AAC *Lys* AAG gtgagcagacc	....9.7....	tgtctgtttcag *Glu* GAA GCC GAT GAA
3	96	AGG TAC AAG *Lys* AAG gtaactatcac	....2.8....	cctttgcaacag *Asn* AAT CTT CTG GAG
4	90	GGG ATG AGG *Glu* GAG gtaagtactgg	....4.1....	tctgaatctcag *Ser* TCT GGC CTC CTC
5	126	CTG CAG GCC *Gln* CAG gtatgatccat	...10.2...	gtttgcccacag *Thr* ACC ACG GCC CTG
6	221	GAA GGC AGC *Ar* AG gtgcggtcctgt	....1.7....	tcttccttccag *g* G CTG AGC TCC
7	86	AGT GTG AAA *A* G gtaacccctcaa	....2.0....	gttttccattag *sp* AT CTT GCC AGG
8	974			

Fig. 5. The sequences bordering the intron interruptions are depicted as well as the size of each exon and intron. The interrupted amino acids are stated according to the three-letter code to emphasize the phase of the interruption

4. Discussion

4.1. The mouse filensin

Sequence comparison of filensin gene clones from different species indicates an unusual divergence as compared to other IF proteins including its partner phakinin [6–8]. The mouse filensin clone as opposed to filensin from other species shows a 5 amino acid insertion in the beginning of the coil 2 helix overlapping with the intron 4 interruption of the coding sequence. This area of the rod domain is marked by a 44 amino acid insertion in the mouse peripherin clone [46] and the 29 amino acid deletion in the other filensin isolates. Moreover, the precise location of the intron/exon barrier in this region shows a marginal shift between type I and type II cytokeratins. Taken together, insertions, deletions and exon/intron barrier shifts indicate that this part of the rod domain may be permissive to organizational alterations. We have previously reported neurofilament-like repeats in the tail domain of the bovine filensin clone [7]. The absence of these repeats from all other filensin isolates indicates that they may have been acquired later in evolution. Insertion of the bovine repeats was localized between amino acids 527QE and 598QE, encoded by CAG GAG and CAG GAG both these sequences could represent areas of splicing events. It is therefore possible that the repeats in the bovine clone are the result of exon shuffling.

4.2. Evolutionary relation of the mouse filensin gene to other IF genes

Unlike the gene organization of phakinin which resembles that of the type I keratins [12], the intron/exon organization of filensin does not resemble type II keratins, the obligate partners in the formation of a keratin filament. This argues against tendencies to classify filensin and phakinin to the keratin subgroups [47]. Indeed filensin lacks the coding sequence interruption in the coil 1a region, characteristic of the type II genes [40]. The location of the filensin intron interruptions is identical to the type III IF proteins, indicating that filensin may be originating from this subgroup [37–39]. The organization of filensin's large tail domain does not comparare with

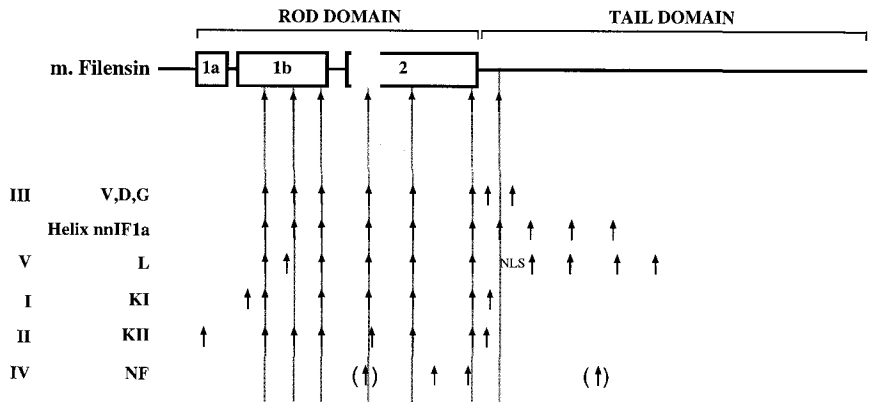

Fig. 6. A comparison of intron positions for the different types of IF genes. Intron positions (arrows) are shown with respect to the filensin protein structure. Boxes represent the coiled coil segments. Abbreviations are as follows: Helix nnIFA, helix aspersa non-neuronal IF protein A; I, KI; II, KII, neutral-basic and acidic keratins, respectively; III, V, D, G, vimentin, desmin and glial fibrillary acidic protein, respectively; IV, NF, neurofilament proteins; NLS, nuclear localisation signal. The arrows set in parentheses indicate intron positions in neurofilaments, which are only present in one of the three analyses genes. Data for the types I–IV IF genes were taken from [23] and references therein. Data for the H. aspersa IF gene are from [44] and for type V from [51].

the tail domains of vertebrate cytoplasmic IF proteins. Genes of the type I, II and III subgroups show a small exon carboxy-terminally to the rod domain which however, does not compare neither in size nor in primary structure to filensin's exon 7. Interestingly the exon 7 of the non-neuronal IF gene of the gastropod helix aspersa [44] shows similarities to filensin's exon 7. These two exons have identical length and phase of interruption by the upstream and downstream introns although their sequence identity is low (31,6% over 20 amino acids). The best analog of filensin's exon 7 is the lamin region between the end of the rod domain and the nuclear localization signal [45]. The similarity of filensin's exon 7 to the exon 7 of the gastropod nnIFa and to the equivalent area of the lamins may be a random coincidence. However, it is also possible that this filensin exon represents the organization of this area in the archetype cytoplasmic IF. Indeed filensin is to date the only vertebrate cytoplasmic IF protein with similarities in the tail domain to lamins. The lamin similarity is restricted to exon7. The large carboxy-terminal exon 8 encoding for most of the tail domain has no parallel between the tail domains of other IF proteins. Lack of conservation in the tail domain region of IF protein has led to suggestions that the variable tail domains may have been added to the IF proteins latter in evolution by exon shuffling [48].

4.3. Relation of beaded filaments to IFs

We have previously reported that filensin and phakinin assemble under conditions similar to other IF proteins to form filaments which upon negative staining/electron microscopy appear like classical IFs [10]. This was supported by mass per length measurements of the molecular architecture of beaded filament using scanning transmission electron microscopy [15]. Our findings indicated that this filament consists of two structural units, one core filament consisting of phakinin homopolymers and one peripheral subfilament consisting of filensin and phakinin heteropolymers. The non-homogeneous organization of the beaded filament is not without precedent in IFs, it parallels the architecture of neurofilaments [18]. Collectively the published in vitro assembly data argue that BFs are indeed IF structures [15], unlike earlier suggestions that they constitute a non-IF structure [12]. The molecular architecture of the BFs differs from the homogeneous keratin filaments and homopolymeric type III filaments.

Although BFs are IF structures and their constituents fulfil the structural principles of IF proteins, their low degree of sequence identity with other IF proteins complicates the classification of these proteins. Characterization of the gene structure for both phakinin and filensin added an additional complication to the classification of the BF with any of the classical types of IFs. Phakinin shows a type I cytokeratin-like gene organization and highest degree of similarity to type I cytokeratins. If BFs were simply a peculiar keratin network, then filensin should be expected to show some similarity to type II cytokeratins. However, not only the assembly data argue against BF being a keratin-like filaments, but also the gene organization of filensin resembles the organization of type III IF genes, indicating a common origin with type III genes.

Taken together, three aspects of molecular evidence lead to the postulate that the BFs are peculiar IFs which do not classify with any of the classical IF types: (1) the primary structure of the network's constituents, (2) the organization of the genes encoding for them, and (3) their assembly parameters in vivo and in vitro. The characteristics of the BFs help to appreciate the existing variability amongst IFs. New IF subunits with extreme tissue and developmental stage-specific characteristics are added continuously to the list of IFs [6–12,49,50] triggering expectations that in future we may be confronted with more highly specialized idiomorphic IF networks.

Acknowledgements: F. Gounari was supported by a 'habilitations stipendium' from the Deutsche Forschungs Gemeinshaft. The authors would like to thank Mr. E. Rezavandy and Ms. A. Lichtenauer for excelent technical assistance. The authors would like to particularly thank Dr. K. Khazaie for continuous support and encouragement.

References

[1] FitzGerald, P.G. (1988) Curr. Eye Res. 7, 1243–1253.
[2] FitzGerald, P.G. and Gottlieb, W. (1989) Curr. Eye Res. 8, 801–811.
[3] Maisel, H. and Perry, M.M. (1972) Exp. Eye Res. 14, 7–12.
[4] Ireland, M. and Maisel, H. (1984) Exp. Eye Res. 38, 637–645.
[5] Ireland, M. and Maisel, H. (1989) Lens Eye Tox. Res. 6, 623–638.
[6] Masaki, S. and Watanabe, T. (1992) Biochem. Biophys. Res. Commun. 186, 190–198.
[7] Gounari, F., Merdes, A., Quinlan, R., Hess, J., FitzGerald, P.G., Ouzounis, C.A. and Georgatos, S.D. (1993) J. Cell Biol. 121, 847–853.
[8] Remington, S.G. (1993) J. Cell Sci. 105, 1057–1068.
[9] Hess, J.F., Casselman, J.T. and FitzGerald, P.G. (1993) Cur. Eye Res. 12, 77–88.
[10] Merdes, A., Gounari, F. and Georgatos, S.D. (1993) J. Cell Biol. 123, 1507–1516.
[11] Sawada, K., Agata, J., Eguchi, G., Quinlan, R. and Maisel, H. (1995) Curr. Eye Res. 14, 545–553.
[12] Hess, J.F., Casselman, J.T. and FitzGerald, P.G. (1996) J. Biol. Chem. 271, 6729–6735.
[13] Brunkener, M. and Georgatos, S.D. (1992) J. Cell Sci. 103, 709–718.
[14] Merdes, A., Brunkener, M., Horstmann, H. and Georgatos, S.D. (1991) J. Cell Biol. 115, 397–410.
[15] Goulielmos, G., Gounari, F., Remington, S., Muller, S., Haner, M., Aebi, U. and Georgatos, S.D. (1996) J. Cell Biol. 132, 643–655.
[16] Goulielmos, G., Remington, S., Schwesinger, F., Georgatos, S.D. and Gounari, F. (1996) J. Cell Sci. 109, 447–456.
[17] Heitlinger, E., Peter, M., Lustig, A., Villiger, W., Nigg, E.A. and Aebi, U. (1992) J. Struct. Biol. 108, 74–89.
[18] Heins, S., Wong, P.C., Muller, S., Goldie, K., Cleveland, D.W. and Aebi, U. (1993) J. Cell Biol. 123, 1517–1533.
[19] Hisanaga, S. and Hirokawa, N. (1989) J. Neurosci. 9, 959–966.
[20] Troncoso, J.C., March, J.L., Haner, M. and Aebi, U. (1990) J. Struct. Biol. 103, 2–12.
[21] Gotow, T., Takeda, M., Tanaka, T. and Hashimoto, P.H. (1992) Eur. J. Cell Biol. 58, 331–345.
[22] Osborn, M. and Weber, K. (1986) Trends Biochem. Sci. 11, 461–472.
[23] Steinert, P.M. and Roop, D.R. (1998) Annu. Rev. Biochem. 57, 593–625.
[24] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[25] Church, G.M. and Gilbert, W. (1984) Proc. Natl. Acad. Sci. USA 81, 1991–1995.
[26] Sambrook, J., Fritsch, E. and Maniatis, T. (1989) Molecular Cloning. A laboratory Manual, Cold Spring Harbor University Press, Cold Spring Harbor, NY.
[27] Lichter, P., Boyle, A.L., Cremer, T. and Ward, D.C. (1991) Genet. Anal. Tech. Appl. 8, 24–35.
[28] Sawyer, J.R., Moore, M.M. and Hozier, J.C. (1987) Chromosoma 95, 350–358.
[29] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.

[30] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[31] Gribskov, M. and Burgess, R.R. (1986) Nucleic Acids Res. 14, 6745–6763.
[32] Masaki, S. and Watanabe, T. (1994) Genomics 22, 449–450.
[33] Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Annu. Rev. Biochem. 55, 1119–1150.
[34] Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263, 211–214.
[35] Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Cell 41, 349–359.
[36] Herrmann, H., Hofmann, I. and Franke, W.W. (1992) J. Mol. Biol. 223, 637–650.
[37] Quax, W., Egberts, W.V., Hendriks, W., Quax-Jeuken, Y. and Bloemendal, H. (1983) Cell 35, 215–223.
[38] Li, Z.L., Lilienbaum, A., Butler-Browne, G. and Paulin, D. (1989) Gene 78, 243–254.
[39] Karpov, V., Landon, F., Djabali, K., Gros, F. and Portier, M.M. (1992) Biol. Cell 76, 43–48.
[40] Marchuk, D., McCrohon, S. and Fuchs, E. (1984) Cell 39, 491–498.
[41] Tyner, A.L., Eichman, M.J. and Fuchs, E. (1985) Proc. Natl. Acad. Sci. USA 82, 4683–4687.
[42] Julien, J.P., Grosveld, F., Yazdanbaksh, K., Flavell, D., Meijer, D. and Mushynski, W. (1987) Biochim. Biophys. Acta 909, 10–20.
[43] Lees, J.F., Shneidman, P.S., Skuntz, S.F., Carden, M.J. and Lazzarini, R.A. (1988) EMBO J. 7, 1947–1955.
[44] Dodemont, H., Riemer, D. and Weber, K. (1990) EMBO J. 9, 4083–4094.
[45] McKeon, F.D., Kirschner, M.W. and Caput, D. (1986) Nature 319, 463–468.
[46] Landon, F., Lemonnier, M., Benarous, R., Huc, C., Fiszman, M., Gros, F. and Portier, M.M. (1989) EMBO J. 8, 1719–1726.
[47] Quinlan, A.R., Carter, M.J., Sandilands, A. and Prescott, A.R. (1996) Trends Cell Biol. 6, 123–126.
[48] Dodemont, H., Riemer, D., Ledger, N. and Weber, K. (1994) EMBO J. 13, 2625–2638.
[49] Hemmati-Brivanlou, A., Mann, R.W. and Harland, R.M. (1992) Neuron 9, 417–428.
[50] Bilbe, G., Delabie, J., Bruggen, J., Richener, H., Asselbergs, F.A., Cerletti, N., Sorg, C., Odink, K., Tarcsay, L. and Wiesendanger, W. et al. (1992) EMBO J. 11, 2103–2113.
[51] Döring, V. and Stick, R. (1990) EMBO J. 9, 4073–4081.