GENOMICS 36, 440–448 (1996)
ARTICLE NO. 0489

Molecular Cloning and Domain Structure of Human Myosin-VIIa, the Gene Product Defective in Usher Syndrome 1B

ZHENG-YI CHEN,* TAMA HASSON,† PHILIP M. KELLEY,‡ BRIAN J. SCHWENDER,*,§ MARC F. SCHWARTZ,† MEENA RAMAKRISHNAN,† WILLIAM J. KIMBERLING,‡ MARK S. MOOSEKER,† AND DAVID P. COREY*,§,[1]

*Departments of Neurobiology and Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114; †Departments of Biology, Cell Biology, and Pathology, Yale University, New Haven, Connecticut 06520; ‡Center for Hereditary Communication Disorders, Boys Town National Research Hospital, Omaha, Nebraska 68131; and §Howard Hughes Medical Institute

Received February 14, 1996; accepted July 4, 1996

Myosin-VIIa is an unconventional myosin with relatively restricted expression. Cloned first from an intestinal epithelium cell line, it occurs most notably in the testis, in the receptor cells of the inner ear, and in the pigment epithelium of the retina. Defects in myosin-VIIa cause the *shaker-1* phenotype in mice and Usher syndrome 1B in human, which are characterized by deafness, lack of vestibular function, and (in human) progressive retinal degeneration. Because the described cDNAs encode less than half of the protein predicted from immunoblots, we have cloned cDNAs encoding the rest of human myosin-VIIa. Two transcripts were found, one encoding the predicted 250-kDa protein and another encoding a shorter form. Both transcripts were found in highest abundance in testis, although the shorter transcript was much less abundant. Both could be detected in lymphocytes by RT-PCR. The myosin tail encoded by the long transcript includes a long repeat of ~460 amino acids. Each repeat contains a novel "MyTH4" domain similar to domains in three other myosins, and a domain similar to the membrane-associated portion of talin and other members of the band-4.1 family.

INTRODUCTION

Myosin-VII is a member of the myosin superfamily, which now includes 11 or more different branches (Cheney and Mooseker, 1992; Hasson and Mooseker, 1995). It is considered an "unconventional" myosin, distinct from the conventional myosin-II family originally characterized in muscle. Two forms of this myosin, VIIa and VIIb, were first identified in a human intestinal epithelium cell line, on the basis of partial amino acid sequence of the conserved head domain (Bement *et al.*, 1994). A portion of the myosin-VIIa head domain was also cloned from the sensory epithelium of the bullfrog inner ear (Solc *et al.*, 1994). Larger portions of the mouse and human myosin-VIIa sequences were described when it was found that defects in myosin-VIIa cause the mouse *shaker-1* phenotype and the human Usher Syndrome 1B, which are characterized by deafness, lack of vestibular function, and (in human) progressive retinal degeneration (Gibson *et al.*, 1995; Weil *et al.*, 1995). A still larger portion of the human sequence was described by Hasson *et al.* (1995), who also used a specific antibody to the tail domain to characterize the protein. Immunoblots indicated that myosin-VIIa is a protein of molecular weight 230–250 kDa which is expressed most highly in testis but also in the retina and the cochlea. Immunocytochemical labeling showed that myosin-VIIa is expressed in the receptor cells of the inner ear and the pigment epithelium cells of the retina, consistent with the sensory deficits of Usher syndrome 1B (Hasson *et al.*, 1995).

The published sequence still accounts for less than half of the predicted coding sequence of myosin-VIIa. Whereas the "motor" head domain is like other members of the superfamily, containing the conserved actin- and ATP-binding regions, the uncharacterized portion includes most of the unique tail domain that gives the molecule its specificity. To understand the interaction of myosin-VIIa with other proteins, specific probes must be generated from unique segments of the tail. In addition, gene defects in the head domain have been found in fewer than half of the families with Usher 1B. Additional defects in the tail may be expected in the

[1] To whom correspondence should be addressed at WEL414, Massachusetts General Hospital, Boston, MA 02114. Telephone: (617) 726-6147. Fax: (617) 726-5256.

0888-7543/96 $18.00

remaining families and must be identified for accurate diagnosis.

We have used probes based on the myosin-VIIa head domain (Hasson *et al.*, 1995) and on the genomic structure of the Usher 1B locus (Kelley *et al.*, in preparation) to clone the rest of the cDNA encoding myosin-VIIa from a human testis cDNA library. Myosin-VIIa is expressed in at least two alternatively spliced transcripts, with the longer form predicting a protein of 250 kDa. The tail domain includes a repeated element of ~460 amino acids. Each element contains two domains: one (~100 aa) that is similar to domains occurring in the tails of three other myosin families and a second (~300 aa) similar to the membrane-associated portion of talin and other members of the band-4.1 family.

METHODS

Library screening. Two different probes were used to screen a random- and oligo(dT)-primed human testis cDNA library, constructed in the Lambda ZAP Express vector (Stratagene, La Jolla, CA) and plated at 2×10^6 plaques. Hybond-N+ (Amersham) was used for the phage lifts. Hybridization and isolation of phagemid clones were essentially as described (Chen *et al.*, 1992a,b).

The first probe came from an 85-kb P1 clone (P-8079) obtained by screening the Dupont P1 library (Genome Systems, St. Louis MO) with a probe derived from a cosmid subcloned from the YAC 965F10, which mapped to this region (Kelley *et al.*, in preparation). A 10-kb *Eco*RI fragment of the P1 clone that mapped just past the 3′ end of U34227 was then identified and isolated by gel purification. The *Eco*RI fragment was preassociated with human Cot-1 DNA (GIBCO-BRL) before hybridization to eliminate background from repetitive DNA, and two overlapping clones (Nos. 4.4 and 1.13) were derived from the *Eco*RI fragment screen. These encompass nt 3985–7368 of myosin-VIIa transcript 1 (GenBank Accession No. U55208). RT-PCR was used to obtain an additional clone that linked clone 4.4 and U34227 (see below).

The second probe was a PCR product of 382 bp synthesized from bases 3010–3391 of our previous myosin-VIIa sequence (Accession No. U34227; Hasson *et al.*, 1995). Two overlapping clones were derived from this screen, encompassing nt 488–4060 of myosin-VIIa transcript 2 (GenBank Accession No. U55209).

The database of expressed sequence tags was screened with the sequence derived from these clones, using the BLAST network service at the NCBI, and 11 positives were identified. One of these (Accession No. R27660, clone 134502) was requested from the IMAGE consortium and sequenced. Clone 134502 overlapped with nt 6954–7368 of myosin-VIIa transcript 1.

Sequencing and sequence alignment. The DNA from selected clones was sequenced using an automated sequencer (ABI) and assembled in Sequencher 3.0 (Gene Codes, Ann Arbor, MI), which permitted simultaneous visualization of the chromatograms. All compiled sequence represents the agreement of at least two forward sequences and one reverse sequence.

Alignment of similar regions in myosin-VIIa and other proteins was performed with CLUSTAL W version 1.5, running at the Human Genome Center, Baylor College of Medicine.

Polymerase chain reaction. Amplification consisted of 30 cycles using 200 m*M* nucleotides, 1 m*M* primers, 1× standard buffer supplied by the manufacturer, and *Taq* DNA polymerase (Boehringer Mannheim, Indianapolis IN) in a total of 50 μl. Each cycle consisted of a 15-s denaturation at 94°C; a 30-s annealing at 67°C; and a 30-s elongation at 72°C followed by a final elongation at 72°C for 10 min. For any products larger than 2 kb, a combination of *Taq* and extender was used as recommended (Stratagene) and elongation time was increased by 1 min/kb. Products were analyzed by electrophoresis on a 1–3% agarose gel and subsequently purified using the Geneclean II kit (Bio101, Vista CA). PCR primers used had the following sequences (5′ to 3′):

my7ct34	Common	Forward	ACCAGCTGGCAGCCCTGGCGGTCT	3368→
pHM7A3431F	Common	Forward	CCAAGTACCACACAGCCATGAGTGAT	3431→
4.4jr3R	Transcript 1	Reverse	GGTGGCCTGCAGCTCCAGCCAGCT	←4023
my7c5600R	Transcript 1	Reverse	GATCTGCACATATGCCTCGTCCTT	←5604
my7c3.1t73	Common	Reverse	TTGTGCCTCACACTGCTCTTCTTCT	←3596
my7c31t71	Transcript 2	Reverse	CACCACTCAGCCACCAAAAGACTTT	←3815

RT-PCR and cloning of PCR products. PCR primers directed against the 3′ end of U34227 and the 5′ end of clone 4.4 (HM7A3431F/4.4jr3R) were used in an RT-PCR to isolate the intervening cDNA fragment. Human testis poly(A) RNA was from Clontech. Total RNA was isolated from human lymphocytes by the guanidine thiocyanate method (Chirgwin, 1979). Five micrograms of lymphocyte RNA and 1 μg of testis RNA were reverse transcribed with 200 U of M-MLV reverse transcriptase (Promega, Madison WI) using an oligo(dT) primer. The cDNA was amplified in a polymerase chain reaction as described above, and the resulting PCR products were cloned into the PCR II vector using the TA cloning kit (Clontech, Palo Alto CA).

To distinguish expression of two myosin-VIIa transcripts, they were amplified from lymphocyte cDNA (two different samples), using human testis cDNA as a positive control. For transcript 1, a primer set specific for the tail of transcript 1 (HM7A3431F/my7c5600R) was used for the first round of RT-PCR, followed by a second round of seminested PCR using primers HM7A3431F/4.4jr3R. For transcript 2, the first round of PCR was carried out with primers HM7A3431F/my7c3.1t71, followed by a second round of PCR using primers HM7A3431F/my7c3.1t73. Amplified products could be detected only after the second round of seminested PCR.

RESULTS

Identification of Two Transcripts

The apparent molecular weight of human myosin-VIIa is ~240 kDa, based on immunoblots, but we had previously cloned only the N-terminal half (Hasson *et al.*, 1995). To recover cDNAs representing the rest of the myosin-VIIa message, we used a cosmid subcloned from a YAC containing the myosin-VIIa gene to isolate a P1 genomic clone and identified a 10-kb fragment that mapped just past the 3′ end of the existing cDNA. This was used to screen a human testis cDNA library, and two overlapping clones (Nos. 4.4 and 1.13 of 1.7 and 2.3 kb, respectively) were recovered. Neither clone overlapped with U34227, however, so an RT-PCR was performed with human testis cDNA and a pair of primers derived from the 3′ end of U34227 and the 5′ end of the new clone 4.4. A product of 593 bp was amplified and subsequently subcloned. These four overlapping clones (Fig. 1A) represent 7368 bases of a transcript (transcript 1), extending to the poly(A) tail. It includes a potential polyadenylation site (CATAAA) located 20 bp upstream of the poly(A) tail. CATAAA is an uncommon form of polyadenylation signal that has been observed in less than 1% of all polyadenylation signals observed (Wickens and Stephenson, 1984).

In parallel experiments, the testis cDNA library was

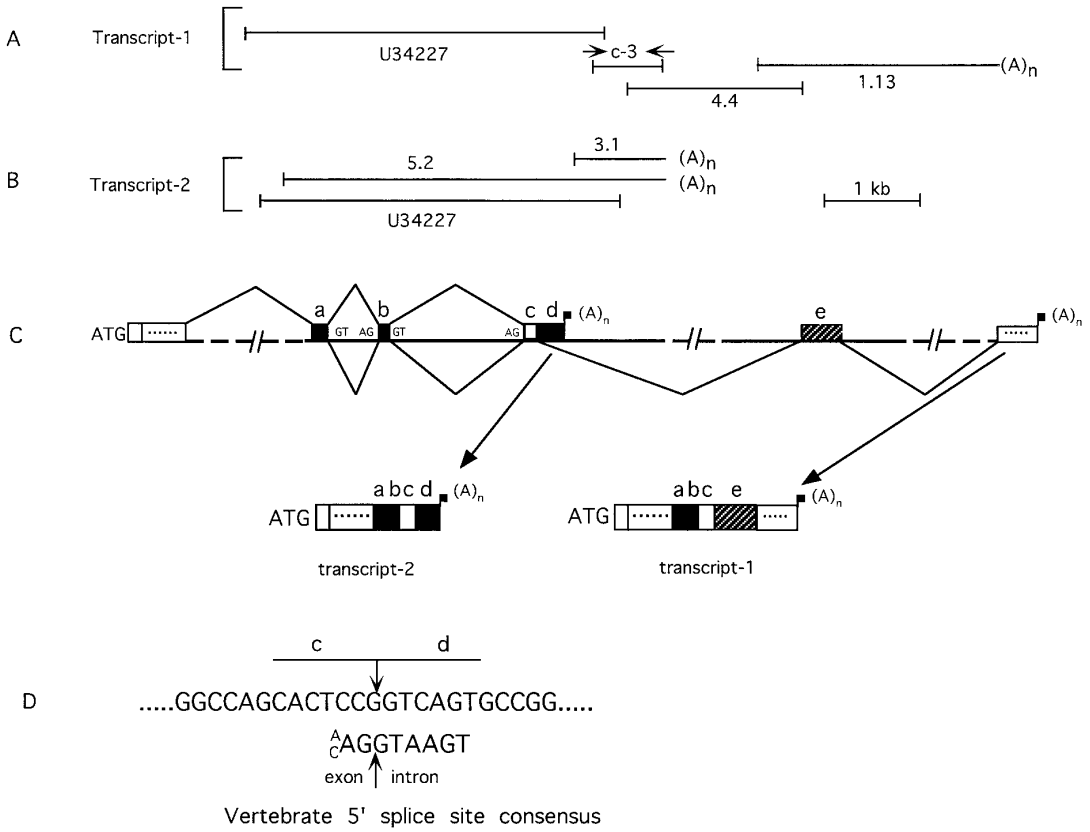


FIG. 1. Sequence assembly and splice boundaries of two transcripts. (A) cDNA clones. For transcript 1, the positions of two cDNA clones (4.4 and 1.13) and a subcloned PCR product (C-3) are shown. (B) For transcript 2, the positions of two positive cDNA clones (clones 3.1 and 5.2) are shown relative to our previous sequence (Accession No. U34227). (C) Genomic organization in the splice region. Splicing for transcript 2 (top) includes upstream exons encoding the head domain, and exons *a, b,* and *cd.* Transcript 1 (bottom) includes exons *a, b,* the "*c*" part of exon *cd,* exon *e,* and others downstream. (D) The division between the two parts of exon *cd* is marked by a vertebrate splice site consensus. GenBank Accession Nos. are U55208 (transcript 1) and U55209 (transcript 2).

also screened with a PCR product of 382 bp (3010–3391) constructed from the 3′ end of U34227. Two positive clones had 5′ ends that matched the existing myosin-VIIa sequence at different points, and both matched each other at their 3′ ends, which included a poly(A) tail and two potential polyadenylation signals (AATAAA) located 14 and 26 bases upstream of the poly(A) tail. However, these clones did not match transcript 1 at their 3′ ends; instead, they represented a second transcript that was much shorter, extending 4060 bp from the start of U34277 (Fig. 1B; transcript 2). Exhaustive screening of this library with a probe derived from U34227 never produced a clone corresponding to the longer transcript 1, perhaps due to insufficient internal priming in library construction.

Transcript 1 has an open reading frame of 6525 bp (bases 268–6792, counting from the 5′ end of U34227) which encodes a protein of 2175 amino acids with a predicted molecular weight of 250 kDa. The predicted amino acid sequence is shown in Fig. 2A and is most likely the protein studied in Hasson *et al.* (1995). Transcript 2 is smaller; it has an open reading frame of 3609 bp (bases 268–3876) encoding a protein of 138 kDa. The predicted amino acid sequence of transcript 2 is identical to that of transcript 1 over most of its length. At amino acid 1171, however, transcript 2 diverges entirely (indicated by an asterisk in Fig. 2) and ends 32 aa later in a stop (Fig. 2B). One clone representing transcript 2 also had a 9-bp insert, encoding 3 additional amino acids at position 1095, so that this clone read GEGEVLQAQL, where transcript 1 had GEGEAQL (insert underlined). Other clones from this library representing transcript 2 do not have the VLQ insert. Thus, there appear to be at least 3 alternative splice forms of myosin-VIIa.

Boundaries for Alternative Splice Forms

To understand the splice events leading to the expression of these two major transcripts, PCR was used to amplify genomic DNA between U34277 and the divergent part of transcript 2, using primers my7ct34 and my7c31t71. A single PCR product of 2.5 kb was amplified and cloned into the PCR II vector for sequencing. Figure 1B illustrates the arrangement of exons in the region of the alternative splicing, deduced from the genomic PCR product. Exons *a,b, cd,* and *e* follow immediately in the genome from exons in the U34227

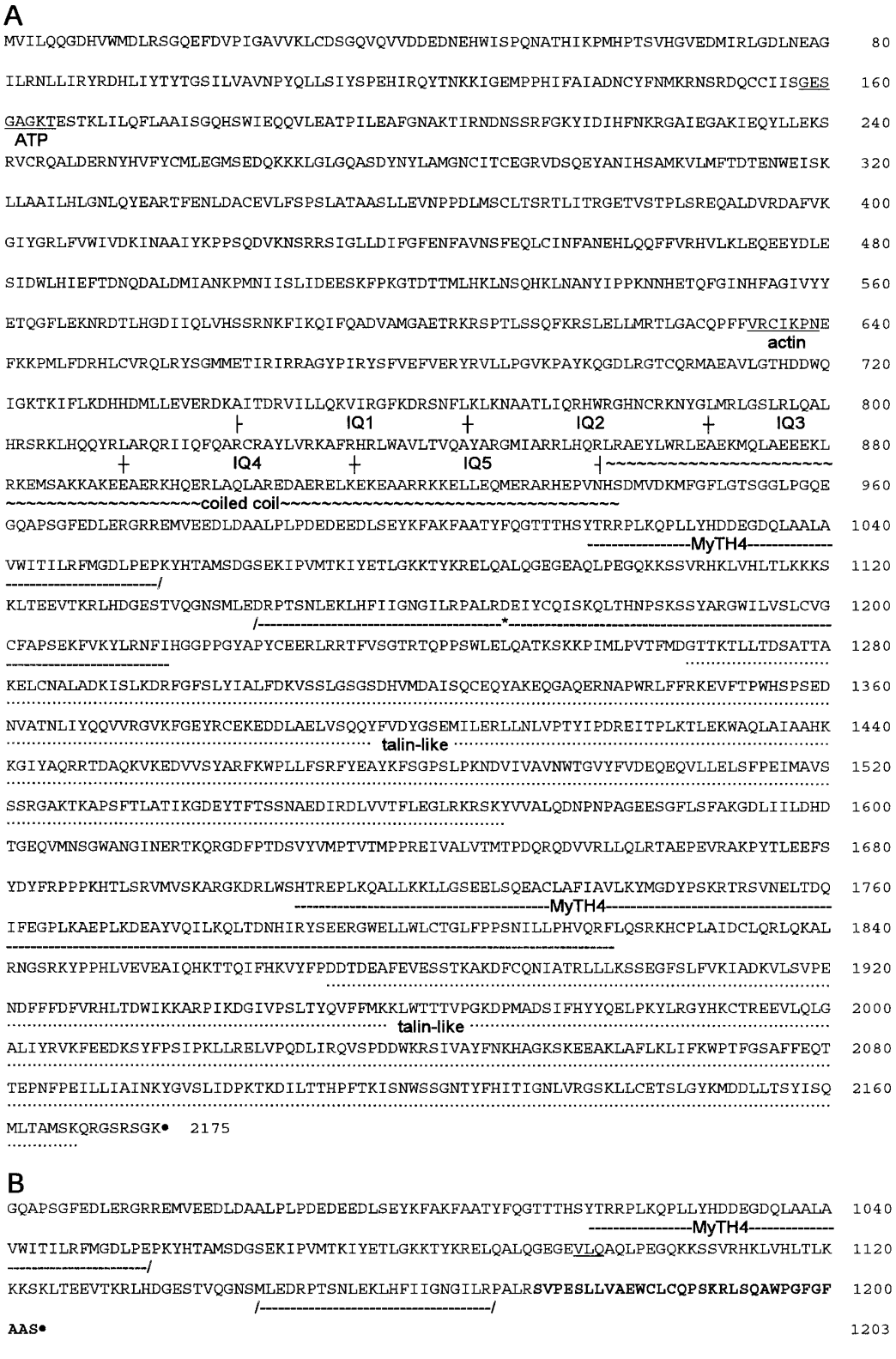


FIG. 2. (A) Amino acid sequence predicted from transcript 1. The head domain (aa 1–742) includes the ATP- and actin-binding sites (underlined), followed by five light-chain-binding IQ motifs (IQ 1–5), and a predicted coiled-coil region (aa 859–940). Two MyTH4 domains (aa 1016–1054//1145–1217 and 1710–1821) are indicated by dashes, and two talin-like domains (aa 1267–1568 and 1872–2167) are indicated by dots. The point of divergence of transcript 2 at aa 1171 is indicated by an asterisk. (B) the C-terminal 243 amino acids of transcript 2. The divergent portion is indicated in boldface; the alternatively spliced amino acids (VLQ) are underlined.

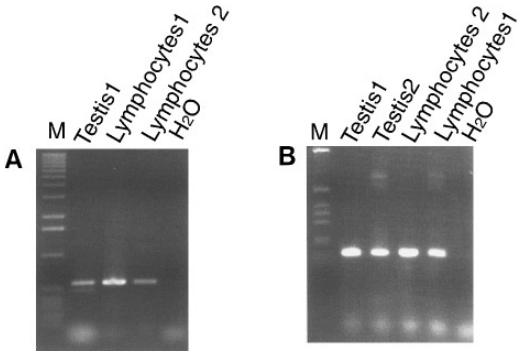


FIG. 3. RT-PCR of myosin-VIIa transcripts 1 and 2. (A) Transcript 1. PCR primers unique to transcript 1 (HM7A3431F/my7c5600R + 4.4jr3R) were used to amplify a fragment of ~600 bp from testis cDNA and two samples of lymphocyte cDNA. (B) Transcript 2. With primers unique to transcript 2 (HM7A3431F/my7c31t71 + my7c31t73), a fragment of ~170 bp was amplified from two testis samples and two lymphocyte samples. Arrows in Fig. 1 indicate the position of PCR primers used for RT-PCR. M indicates 1-kb ladder.

cDNA. In addition to upstream exons, transcript 2 includes exons *a*, *b*, and *cd*; transcript 1 includes exons *a*, *b*, part *c*, exon *e*, and others downstream. The last exon of transcript 2 (exon *cd*) represents bases 3652–4060. However, the point of divergence is at base position 3780; thus the splice occurs in the middle of an exon, and the divergent part of transcript 2 (*d*) is in the same exon as a common part (*c*). The apparent splice point is next to a potential 5′ splice donor site (GT), and the sequence around the splice junction also fits the vertebrate 5′ splice site consensus (Fig. 1C). Moreover, this type of splice event has been documented for the immunoglobulin μ constant region, where it results in either the secreted or membrane-bound form of the protein (Roger *et al.*, 1980). A detailed analysis of intron/exon boundaries of the entire myosin VIIa gene will be presented separately (Kelley *et al.*, submitted for publication).

Expression of Myosin-VIIa

RT-PCR was also used to study the expression of the myosin-VIIa in testis and lymphocytes. With a forward primer derived from U34227 and nested reverse primers unique to transcript 1 or transcript 2, PCR products of the correct size (~600 and ~170 bp) were amplified in both lymphocytes and testis (Figs. 3A and 3B). With a different forward primer from U34227 and a single reverse primer from transcript 2 (my7ct34/my7c31t71), a product of the correct size (~440 bp) was amplified from human testis cDNA and the control transcript 2 clones, confirming that this transcript is expressed in testis (data not shown).

Domain Structure of the 250-kDa Form

Figure 2 shows the predicted amino acid sequence of the larger and most abundant form of myosin-VIIa, and Fig. 4 illustrates the notable domains of both transcripts schematically. The predicted molecular weight of the larger form is 250 kDa and the pI is 8.5. The myosin head domain, IQ motifs, and coiled-coil domain have been reported previously (Gibson *et al.*, 1995; Weil *et al.*, 1995; Hasson *et al.*, 1995)

An internal repeat. Myosin-VIIa has a long repeated element of about 460 amino acids (aa 1016–1568 and 1708–2167). The first element is interrupted by an insertion of about 90 amino acids (aa 1055–1145), but the elements are otherwise 28% identical at the amino acid level. Each element has two domains that are similar to domains in other proteins: the MyTH4 domain and the talin homology domain.

The MyTH4 domain. The first ~110 amino acids of each repeat element constitute a highly conserved domain found in four different branches of the myosin superfamily (Fig. 5). We have named this motif the MyTH4 domain, for myosin tail homology, and numbered it to distinguish it from homologies TH1-3 found in the tails of myosins-I (Hammer, 1994). An *Acanthamoeba* myosin heavy chain that has been assigned to class IV contains one domain toward the C-terminal of the tail (Horowitz and Hammer, 1990). Myosin-VIIa has two copies of the domain, as does a novel *Caenorhabditis elegans* myosin, encoded by a gene on cosmid F46C3 and likely to be named myosin-XII (Accession No. Z66563). Within the domain, each myosin is about 35% identical to a consensus sequence.

The talin homology domain. The last ~300 amino acids of each myosin-VIIa element are similar to a conserved domain in the band-4.1 superfamily of proteins (Algrain *et al.*, 1993b). These regions (amino acids 1267–1568 and 1872–2167) are most similar to talin and filopodin, being about 20% identical at the amino acid level (Fig. 6). They are slightly less similar (18% identical) to the same domain in the ERM branch of the band-4.1 family, which includes ezrin, radixin, moesin, and merlin, and they have limited similarity to band-4.1 itself (11%). Although the identity is not strong, the similarity is significant and extends over a long distance. These two domains in myosin-VIIa appear to be equally distant from the talins and from the ERMs: the amino acids conserved between myosin-VIIa and the talins are not consistently conserved between myosin-VIIa and the ERMs, but other amino acids are conserved between myosin-VIIa and the ERMs. Identities between each group are illustrated in Fig. 6.

DISCUSSION

Expression of Two Myosin-VIIa Transcripts

Analysis of cDNA sequences for myosin-VIIa has identified two different transcripts, encoding different tails associated with the same motor domain. The major cDNA for myosin-VIIa encodes a 250-kDa poly-

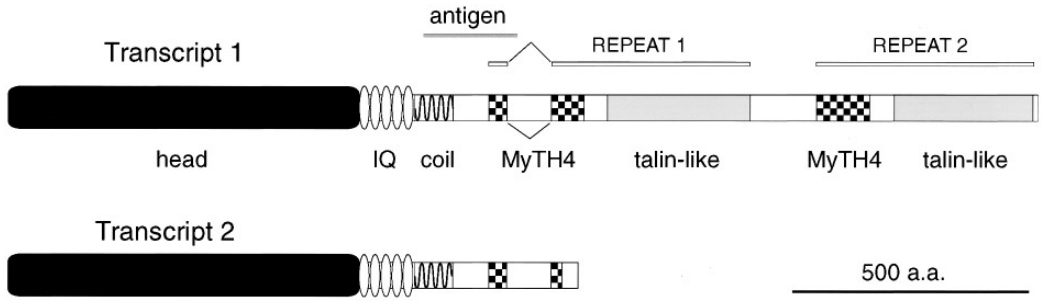
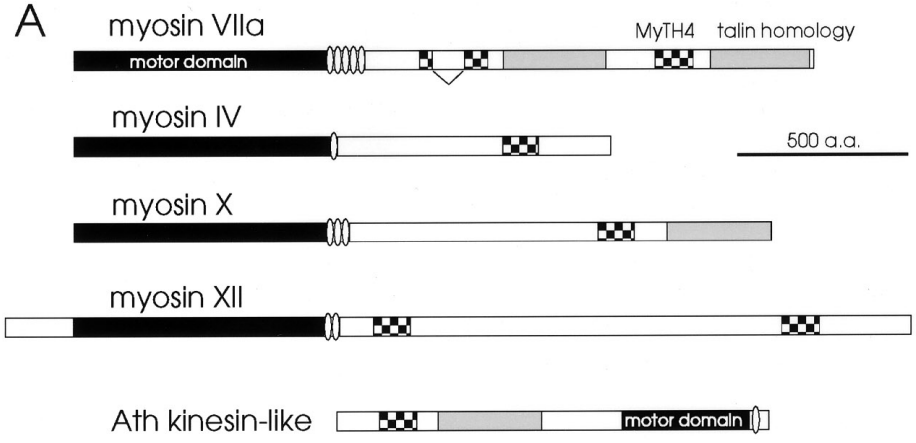


FIG. 4. Domain structure of the predicted amino acid sequences for both transcripts. The antigen used for the myosin-VIIa-tail-specific antibody (Hasson *et al.*, 1995) is common to both transcripts, so the antibody should also recognize a protein of the transcript 2 molecular weight in tissues where that transcript is expressed.

peptide, in keeping with the identification of a 240-kDa protein by immunoblot using antibodies directed against the proximal portion of the myosin-VIIa tail (Hasson *et al.*, 1995). Our RT-PCR data indicate that transcript 2, encoding the shorter 138-kDa form of myosin-VIIa, is expressed in testis. Using a probe specific to the unique 3′ end of transcript 2, however, we were unable to detect expression of this transcript in Northern blots of testis, suggesting a low level of expression. In frogs and rodents, we have not detected the expression of the 138-kDa form in either retina or cochlea as judged by immunoblot (T. Hasson, unpublished data) although the 250-kDa form can be clearly detected. Production of 138-kDa-form-



B

```
1018 TRRPLKQPLLYHDD--EGDQLAALAVWITI--LRFMGDLPEP-//-EDRPTSNLEKLHFII  Myo VIIa
1710 TREPLKQALLKKLLGSEELSQEACLAFIAV--LKYMGDYP-----SKRTRSVNELTDQIF  Myo VIIa
1262 TKSPIPTSLTTLD---QLAVKSATRLFKNV--LGFMGDRPLP--------YPNALAQDLL  Myo IV
1542 PYGDINLNLLK-DKGYTTLQDEAIKIFNSLQQLESMSD-PIP------------IIQGIL  Myo X
1085 RREPIMTPFLHKE--SDYDFRLSVEIFKLI--LKYMNDI------KLTKKQREDLGRYIV  Myo XII
2285 SEKPISQSLLA-DLGNEE-SKYAVETFHAI--MKFMGDEPL-----KKSESMTDVVFKVL  Myo XII
118  IPIPIPTSLL-KINSDLV-SR-ATKLFHLI--LKYMGVDSSDRSTPPSLDERIDLVGKLF  Ath kin

1159 GNGILRPALRDEIYCQISKQLTHN--PSKSSYARGWILVSLCVGCFAPSEKFVKYLRNFI  Myo VIIa
1763 EGPLKAEPLKDEAYVQILKQLTDN--HIRYSEERGWELLWLCTGLFPPSNILLPHVQRFL  Myo VIIa
1309 EQCLAAPELRNEVYCQIIKQLTEN--PSPQSVTKGWQLMRCCLQTFPPSEEFANCLEMFL  Myo IV
1588 QTGHDLRPLRDELYCQLIKQTNKVPHPGSVGNLCSWQILTCLSCTFLPSRGILKYLKFHL  Myo X
1135 QQGISNPCQRDEILVQTINQINKN--QDKTASDNGWKLVHMAISVFPPTENIIPMLIGFF  Myo XII
2336 LICHRQPTLRDEVYCQLIKQTTSNISQKPNSALRAWRLLTIITAYFPSSLTLKPYVLQYL  Myo XII
173  KKTLKRVELRDELFAQISKQTRHN--PDRQYLIKAWELMYLCASSMPPSKDIGGYLSEYI  Ath kin
```

FIG. 5. (A) Positions of MyTH4 domains (checkered) in myosins VIIa, IV, X, and XII, and a kinesin-like protein in plants. "Motor domain" indicates myosin heads or a kinesin motor domain; the ovals indicate calmodulin or other light-chain-binding domains. (B) Sequence comparisons. Amino acids identical to the consensus are shaded (human myosin-VIIa, Accession No. U55208; *Acanthamoeba* myosin-IV, Accession No. M60954; *C. elegans* myosin-XII, Accession No. Z66563; *Arabidopsis thaliana* kinesin-like protein, Accession No. L40358).

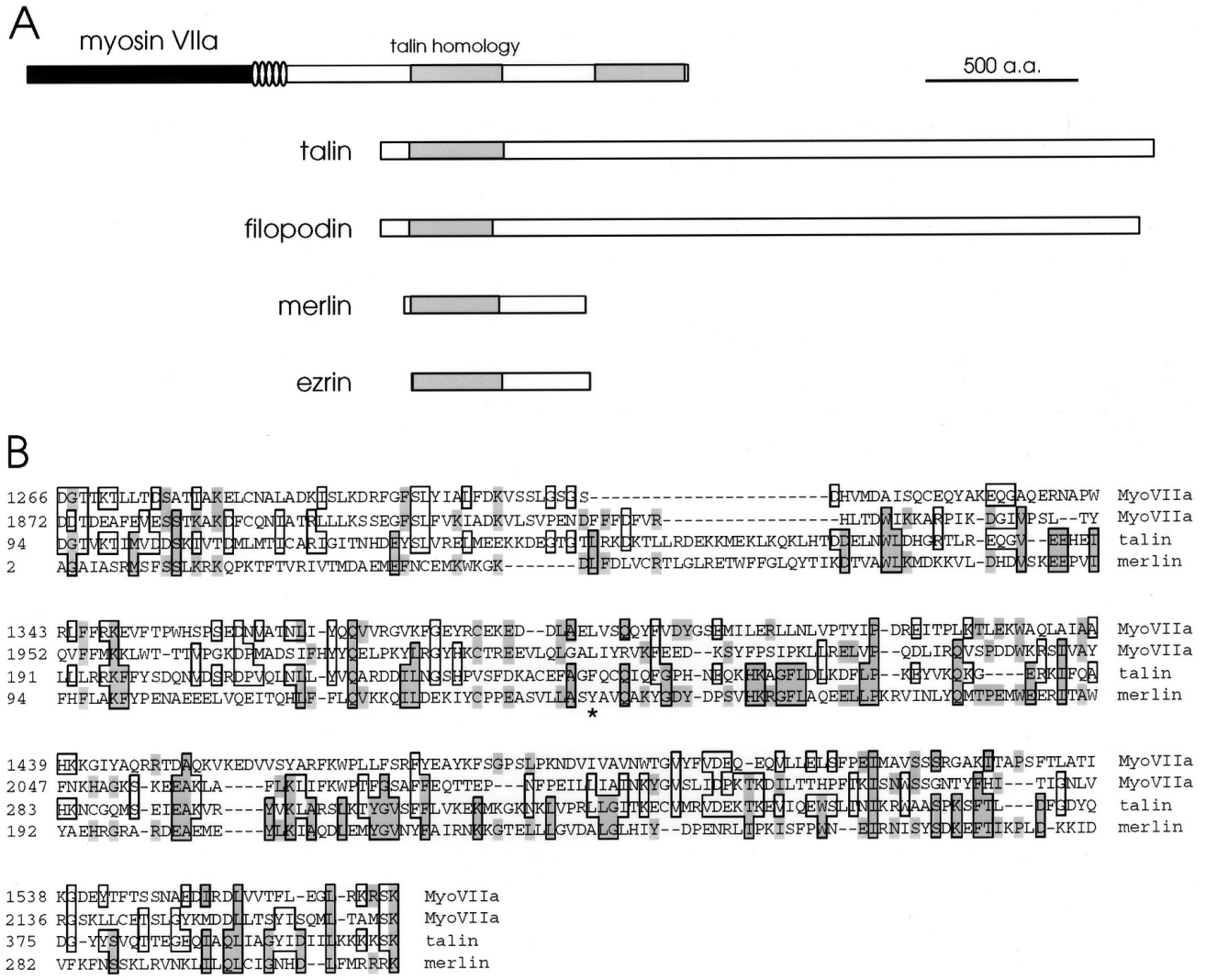


FIG. 6. (A) Location of talin-homology domains in myosin VIIa, talin, filopodin, merlin, and ezrin. (B) Sequence comparisons for selected examples (mouse talin, Accession No. P26039; human merlin, Accession No. P35240). Amino acids that are identical to talin are boxed; those identical to merlin are shaded. Note that few amino acids are shared by all four sequences. The tyrosine phosphorylated in ERM family members is indicated by an asterisk. Neither myosin-VIIa nor talin has a tyrosine at this position.

specific antibodies is in progress to address the function of this novel isoform.

RT-PCR experiments indicate that myosin-VIIa is expressed at least at low levels in lymphoblasts. This is in agreement with the work of Bement *et al.* (1994), who identified myosin-VIIa expression by RT-PCR from peripheral blood lymphocytes. Although the protein could not be detected by immunoblot, the message could be amplified. This should simplify mutation screening in those at risk for Usher 1B since the large myosin-VIIa gene would not need to be screened exon by exon, although genomic amplification will still be needed in cases where mutations involve splice sites.

Similarity to Other Myosins

As previously reported (Gibson *et al.*, 1995; Weil *et al.*, 1995; Hasson *et al.*, 1995), the head domain of myosin-VIIa is similar to that of other members of the myosin superfamily and includes the ATP-binding site and the actin-binding site (indicated on Fig. 2). This is followed by 5 IQ motifs, which are probable binding sites for regulatory light chains such as calmodulin (Cheney and Mooseker, 1992; Hasson and Mooseker, 1995). The IQ domains are followed immediately by a predicted coiled-coil domain of 80 amino acids. Although this is short in comparison to some myosins, it is probably sufficient to mediate dimerization, so that the native protein (with light chains) would occur as a complex of up to 670 kDa.

Except for coiled-coil domains, the tails of different myosin classes are thought to be quite different (Mooseker and Cheney, 1995). Thus it is especially interesting to find a large domain in the myosin-VIIa tail (MyTH4) that is shared by other myosin classes. The only other motif shared by multiple myosin classes is the SH3

domain found in myosins I and IV (Mooseker and Cheney, 1995). In contrast, the other conserved domain in the tail (talin homology) has been found primarily in the band-4.1 superfamily.

Phylogenetic analysis of myosin motor domains has shown that myosin-VIIa is most similar to the myosin-X class (D. P. Corey and R. E. Cheney, unpublished results; Mooseker and Cheney, 1995). Myosin-X has been cloned in part from bullfrog inner ear (Solc *et al.*, 1994) and in full from bovine aorta (D. P. Corey and R. E. Cheney, unpublished results). Bovine myosin-X has a single MyTH4 domain, followed by a talin homology at its C-terminal. These similarities suggest a conservation of tail function between these two classes, in addition to the motor function common to all myosin heads. Recently, a protein from *Arabidopsis* with a kinesin motor domain has been identified (Reddy *et al.*, 1996). It has also a MyTH4 domain followed by a talin homology domain (Fig. 5), suggesting that the pairing of these domains in association with a motor is functionally significant.

The MyTH4 Domain

The best characterized of the four myosins carrying a MyTH4 domain is *Acanthamoeba* myosin-IV. Preliminary work has found that myosin-IV is a monomer with a globular myosin motor head and a 50-nm tail (Repezza *et al.*, 1994). This suggests that the MyTH4 domain does not facilitate dimerization. Although myosin-VIIa and myosin-X both harbor the talin homology, myosin-IV and myosin-XII do not, suggesting that the MyTH4 and talin domains in the repeat element can also serve distinct functions.

The Talin Homology Domain

The two 300-amino-acid domains found in myosin-VIIa are homologous to a portion of talin that has been well characterized biochemically. Talin was first identified as a component of adhesion plaques (Burridge and Connell, 1983) and was found to be capable of interacting with membranes, actin, and the adhesion-plaque component, vinculin (Nuckolls *et al.*, 1990; Niggli *et al.*, 1994). The membrane binding portion of talin has been mapped to the N-terminal 47 kDa (434 aa), of which aa 94–406 constitute the talin homology region in myosin-VIIa. The actin and vinculin binding sites have been mapped to the C-terminal 220-kDa fragment and are distinct from the N-terminal domain (Nuckolls *et al.*, 1990; Niggli *et al.*, 1994). The membrane-binding characteristics of the 47-kDa N-terminal talin peptide have been studied extensively *in vitro* and have been found by a number of methods to bind acidic phospholipids specifically (for review see Nuckolls *et al.*, 1990).

The N-terminus of talin is also homologous to the N-termini of band-4.1 and members of the ERM family of actin-binding proteins (ezrin, radixin, moesin, and merlin). Band-4.1 has also been found to bind preferentially to acidic phospholipids via its N-terminal domain (Cohen *et al.*, 1988). Biochemical analysis of the N-terminal domain of ERM family members has not been performed, but all ERM proteins are associated with plasma membrane compartments (Algrain *et al.*, 1993a; Sato *et al.*, 1992; Berryman *et al.*, 1993). These shared functional roles suggest that the two talin homology domains found in the tail of myosin-VIIa target this molecular motor to plasma membrane compartments and, in particular, to acidic phospholipid bilayers.

In addition to membrane binding activities, the talin homology domains may also serve a role as protein-binding sites. The N-terminus of band-4.1, for example, has been shown also to serve as the binding site for both glycophorin C and p55 (Marfatia *et al.*, 1994) as well as calmodulin (Kelly *et al.*, 1991). The N-terminus of ezrin serves a self-association role, as it can bind to the C-terminus of another ezrin molecule (Gary and Bretscher, 1995; Bretscher *et al.*, 1995). Therefore, we cannot rule out that the talin homology domains in the myosin-VIIa tail serve a role in protein–protein interaction or dimerization.

The N-terminal domain of ERM family members has also been shown to serve a regulatory function. Upon treatment with growth factors, human epidermoid carcinoma A431 cells are stimulated to form microvilli and exhibit membrane ruffling. This activity is correlated with the phosphorylation of ezrin by a tyrosine kinase (Brescher, 1989). One of the sites, Tyr145, lies in the N-terminal domain homologous to the talin homology of myosin-VIIa (Krieg and Hunter, 1992). This amino acid is conserved in other family members, including radixin and merlin (Fig. 6), but is not conserved in talin, band-4.1, or myosin-VIIa. Therefore, it is likely that myosin-VIIa is not regulated at the level of tyrosine phosphorylation of the talin homology domains.

Preliminary biochemical analysis of the 250-kDa myosin-VIIa polypeptide shows it to be tightly associated with membrane in both testis and kidney (T. Hasson and M. Mooseker, unpublished observations). In both tissues, it is also found in close association with polarized, bundled actin filaments. This type of location is also seen in retina and cochlea where myosin-VIIa is associated with the apical actin-rich villi of the retinal pigmented epithelium and the stereocilia of the cochlear hair cells (Hasson *et al.*, 1995). Given the domain structure of myosin-VIIa, it is possible that the tail of myosin-VIIa serves to target the motor to membranes or membrane proteins while the motor domain targets the protein to polarized actin filaments. Myosin-VIIa may therefore serve a structural role tethering the bundled, polarized actin filaments of both villi and stereocilia to the plasma membrane.

Since this paper was submitted, Weil *et al.* (1996) have also described the cloning of the full-length myosin-VIIa from retinal pigment epithelial cells. In addition to two small inserts, which may represent splice differences between retina and testis, their sequence

has 10 single-base differences resulting in 7 amino acid differences, which might be polymorphisms.

ACKNOWLEDGMENTS

We thank Xandra O. Breakefield (Massachusetts General Hospital) for helpful discussions and for the extended use of her laboratory, and Bruce Derfler for assistance with sequence assembly. This work was supported by the National Institutes of Health (DC002281 to D.P.C., DC00677 to W.J.K., DK38979 to J. Morrow for T.H. and M.S.M., and DK25387 and DK34989 to M.S.M.), by the Foundation Fighting Blindness (89-198 to W.J.K.), by the Deafness Research Foundation (to Z-Y.C.), by a Muscular Dystrophy Association Grant to M.S.M., and by the Howard Hughes Medical Institute. D.P.C. is an Associate Investigator of the Howard Hughes Medical Institute.

REFERENCES

Algrain, M., Arpin, M., and Louvard, D. (1993a). Wizardry at the cell cortex. *Curr. Biol.* 3: 451–454.

Algrain, M., Turunen, O., Valheri, A., Louvard, D., and Arpin, M. (1993b). Ezrin contains cytoskeleton and membrane binding domains accounting for its proposed role as a membrane-cytoskeletal linker. *J. Cell Biol.* 120: 129–139.

Bement, W. M., Hasson, T., Wirth, J. A., Cheney, R. E., and Mooseker, M. S. (1994). Identification and overlapping expression of multiple unconventional myosin genes in vertebrate cell types. *Proc. Natl. Acad. Sci. USA* 91: 6549–6553.

Berryman, M., Franck, Z., and Bretscher, A. (1993). Ezrin is concentrated in the apical microvilli of a wide variety of epithelial cells whereas moesin is found primarily in endothelial cells. *J. Cell Sci.* 105: 1025–1043.

Bretscher, A. (1989). Rapid phosphorylation and reorganization of ezrin and spectrin accompany morphological changes induced in A-431 cells by epidermal growth factor. *J. Cell Biol.* 108: 921–930.

Bretscher, A., Gary, R., and Berryman, M. (1995). Soluble ezrin purified from placenta exists as stable monomers and elongated dimers with masked C-terminal ezrin-radixin-moesin association domains. *Biochemistry* 34: 16830–16837.

Burridge, K., and Connell, L. (1983). A new protein of adhesion plaques and ruffling membranes. *J. Cell Biol.* 97: 359–367.

Chen, Z.-Y., Hendriks, R. W., Jobling, M. A., Powell, J. F., Breakefield, X. O., Sims, K. B., and Craig, I. W. (1992a). Isolation and characterization of a candidate gene for Norrie disease. *Nature Genet.* 1: 204–208.

Chen, Z.-Y., Powell, J. F., Hsu, P., Breakefield, X. O., and Craig, I. W. (1992b). The organization of the human monoamine oxidase genes and long range physical mapping around them. *Genomics* 14: 75–82.

Cheney, R. E., and Mooseker, M. S. (1992). Unconventional myosins. *Curr. Opin. Cell Biol.* 4: 27–35.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18: 5294–5299.

Cohen, A. M., Liu, S. C., Lawler, J., Derick, L., and Palek, J. (1988). Identification of the protein 4.1 binding site to phosphatidylserine vesicles. *Biochemistry* 27: 614–619.

Gary, R., and Bretscher, A. (1995). Ezrin self-association involves binding of an N-terminal domain to a normally masked C-terminal domain that includes the F-actin binding site. *Mol. Biol. Cell* 6: 1061–1075.

Gibson, F., Walsh, J., Mburu, P., Varela, A., Brown, K. A., Antonio, M., Beisel, K. W., Steel, K. P., and Brown, S. D. M. (1995). A type VII myosin encoded by the mouse deafness gene *Shaker-1*. *Nature* 374: 62–64.

Hammer, J. A. (1994). The structure and function of unconventional myosins: A review. *J. Muscle Res. Cell Motil.* 15: 1–10.

Hasson, T., Heintzelman, M. B., Santos-Sacchi, J., Corey, D. P., and Mooseker, M. S. (1995). Expression in cochlea and retina of myosin-VIIa, the gene product defective in Usher syndrome type 1B. *Proc. Natl. Acad. Sci. USA* 92: 9815–9819.

Hasson, T., and Mooseker, M. S. (1995). Molecular motors, membrane movements and physiology: Emerging roles for myosins. *Curr. Opin. Cell Biol.* 7: 587–594.

Horowitz, J. A., and Hammer, J. A., III (1990). A new Acanthamoeba myosin heavy chain: Cloning of the gene and immunological identification of the polypeptide. *J. Biol. Chem.* 265: 20646–20652.

Kelley, P. M., Weston, M. D., Chen, Z-Y., Orten, D., Hasson, T., Overbeck, L. D., Pinnt, J., Talmadge, C., Ing, P., Mooseker, M. S., Corey, D. P., Sumegi, J., and Kimberling, W. J., The genomic structure of human myosin-VIIa, the gene product defective in Usher syndrome 1B. Submitted for publication.

Kelly, G. M., Zelus, B. D., and Moon, R. T. (1991). Identification of a calcium-dependent calmodulin-binding domain in Xenopus membrane skeleton protein 4.1. *J. Biol. Chem.* 266: 12469–12473.

Krieg, J., and Hunter, T. (1992). Identification of the two major epidermal growth factor-induced tyrosine phosphorylation sites in the microvillar core protein ezrin. *J. Biol. Chem.* 267: 921–930.

Marfatia, S. M., Lue, R. A., Branton, D., and Chishti, A. H. (1994). In vitro binding studies suggest a membrane-associated complex between erythroid p55, protein 4.1, and glycophorin C. *J. Biol. Chem.* 269: 8631–8634.

Mooseker, M. S., and Cheney, R. E. (1995). Unconventional myosins. *Annu. Rev. Cell Dev. Biol.* 11: 633–675.

Niggli, V., Kaufmann, S., Goldmann, W. G., Weber, T., and Isenberg, G. (1994). Identification of functional domains in the cytoskeletal protein talin. *Eur. J. Biochem.* 224: 951–957.

Nuckolls, G. H., Turner, C. E., and Burridge, K. (1990). Functional studies of the domains of talin. *J. Cell Biol.* 100: 1635–1644.

Reddy, A. S. N., Safadi, F., Narasimhulu, S. B., Golovkin, M., and Hu, X. (1996). A novel plant calmodulin-binding protein with a kinesin heavy chain motor domain. *J. Biol. Chem.* 271: 7052–7060.

Repezza, M., Sellers, J., Urrutia, R., and Hammer, J. (1994). Purification and characterization of high-molecular-weight myosin I (HMMI), and unconventional myosin from Acanthamoeba. *Mol. Biol. Cell (Suppl.)* 5: 277a.

Roger, J., Early, P., Carter, C., Calame, K., Bond, M., Hood, L., and Wall, R. (1980). Two mRNAs with different 3′ ends encode membrane-bound and secreted forms of immunoglobulin μ chain. *Cell* 20: 303–312.

Sato, N., Funayama, N., Nagafuchi, A., Yonemura, S., Tsukita, S., and Tsukita, S. (1992). A gene family consisting of ezrin, radixin, and moesin. Its specific localization at actin filament/plasma membrane association sites. *J. Cell Sci.* 103: 131–143.

Solc, C. K., Derfler, B. H., Duyk, G. M., and Corey, D. P. (1994). Molecular cloning of myosins from the bullfrog saccular macula: A candidate for the adaptation motor. *Auditory Neurosci.* 1: 63–75.

Weil, D., Blanchard, S., Kaplan, J., Guilford, P., Gibson, F., Walsh, J., Mburu, P., Varela, A., Levilliers, J., Weston, M. D., *et al.* (1995). Defective myosin VII gene responsible for Usher syndrome type IB. *Nature* 374: 60–61.

Weil, D., Levy, G., Sahly, I., Levi-Acobas, F., Blanchard, S., El-Amraoui, A., Crozet, F., Philippe, H., Abitol, M., and Petit, C. (1996). Human myosin VIIa responsible for the Usher 1B syndrome: A predicted membrane-associated motor protein expressed in developing sensory epithelia. *Proc. Natl. Acad. Sci. USA* 93: 3232–3237.

Wickens, M., and Stephenson, P. (1984). Role of the conserved AAUAAA sequence: Four AAUAAA points prevent messenger RNA 3′ end formation. *Science* 226: 1045–1051.