An Unconventional Myosin Heavy Chain Gene from
Drosophila melanogaster

Kathryn A. Kellerman and Kathryn G. Miller

Department of Biology, Washington University, St. Louis, Missouri 63130

Abstract. As part of a study of cytoskeletal proteins involved in *Drosophila* embryonic development, we have undertaken the molecular analysis of a 140-kD ATP-sensitive actin-binding protein (Miller, K. G., C. M. Field, and B. M. Alberts. 1989. *J. Cell Biol.* 109:2963–2975). Analysis of cDNA clones encoding this protein revealed that it represents a new class of unconventional myosin heavy chains. The amino-terminal two thirds of the protein comprises a head domain that is 29–33% identical (60–65% similar) to other myosin heads, and contains ATP-binding, actin-binding and calmodulin/myosin light chain-binding motifs. The carboxy-terminal tail has no significant similarity to other known myosin tails, but does contain a ~100–amino acid region that is predicted to form an α-helical coiled-coil. Since the unique gene that encodes this protein maps to the polytene map position 95F, we have named the new gene *Drosophila* 95F myosin heavy chain (95F MHC).

The expression profile of the 95F MHC gene is complex. Examination of multiple cDNAs reveals that transcripts are alternatively spliced and encode at least three protein isoforms; in addition, a fourth isoform is detected on Western blots. Developmental Northern and Western blots show that transcripts and protein are present throughout the life cycle, with peak expression occurring during mid-embryogenesis and adulthood. Immunolocalization in early embryos demonstrates that the protein is primarily located in a punctate pattern throughout the peripheral cytoplasm. Most cells maintain a low level of protein expression throughout embryogenesis, but specific tissues appear to contain more protein. We speculate that the 95F MHC protein isoforms are involved in multiple dynamic processes during *Drosophila* development.

MYOSINS are actin-activated ATPases that convert chemical energy into mechanical movement along actin filaments. All myosins have NH₂-terminal head domains, which share ATP-binding, actin-binding and calmodulin/myosin light chain-binding motifs, and tail domains, which are highly divergent. Conventional double-headed (type II) myosins are hexamers of four light chains bound to two central heavy chains that assemble into bipolar filaments through interactions in their coiled-coil tails. Unconventional type I myosins that contain a single myosin heavy chain (MHC)[1] with associated light chains have been identified in *Acanthamoeba* and *Dictyostelium* (Pollard and Korn, 1973; Cote et al., 1985). Additional unconventional myosins have been identified (mouse dilute MHC, chicken p190 MHC; Mercer et al., 1991; Espreafico et al., 1992) that are thought to be double-headed like conventional type II MHCs, but apparently do not form bipolar filaments. Most of the unconventional MHCs are thought to have lipid-binding and/or ATP-insensitive actin-binding regions in their tails. Although the biological functions of the unconventional MHCs are as yet unknown, they are likely to be involved in vesicular movement, microfilament-membrane crosslinking, and cell locomotion.

In *Drosophila*, several genes of the myosin family have been identified. All muscle MHC isoforms appear to be encoded by a single gene (polytene map position 36B) with alternatively spliced transcripts (Bernstein et al., 1986; Rosek and Davidson, 1986). The nonmuscle MHC gene (map position 60E9) encodes another conventional myosin, which is involved in embryonic cellularization and morphogenesis (Kiehart et al., 1989). In addition, the product of an unusual myosin-related gene, *ninaC*, is required for photoreceptor cell function in the adult eye (Montell and Rubin, 1988). Another sequence with MHC homology (map position 35B-C) has been reported, but is not yet well characterized (Dr. D. P. Kiehart, personal communication). In this report, we identify and characterize a novel MHC gene at polytene map position 95F.

In previous work, F-actin column chromatography (Miller and Alberts, 1989) was used too systematically isolate and identify novel actin-binding proteins from embryonic extracts (Miller et al., 1989). Among the 15 major and at least 30 minor proteins recovered was a 140-kD actin-binding protein which eluted from F-actin in the presence of ATP, a characteristic of known myosins. This and other properties

---

1. *Abbreviations used in this paper*: MHC, myosin heavy chain; PCR, polymerase chain reaction.

(see below) suggested that this molecule might be a myosin and, as has been suggested for unconventional myosins in other systems, might be involved in cytoplasmic transport. In this work, we describe the cloning and sequencing of the gene that encodes this protein. We have determined that it is a member of the myosin family, but apparently represents a new class of unconventional myosins. The gene encodes multiple protein isoforms that are expressed in a developmentally regulated manner. In addition, immunolocalization studies suggest tissue-specific regulation and participation in cytoplasmic transport. Our data suggest that the products of this gene are involved in a variety of dynamic processes required throughout the life of the fly.

Materials and Methods

Molecular Cloning of Drosophila 95F MHC

A *Drosophila* ovarian cDNA λgt11 expression library (Steinhaur et al., 1989) was screened using the 3C7 mAb and a mouse polyclonal antibody produced previously (Miller et al., 1989). The library was plated (Snyder et al., 1987) and induced by standard techniques. Reactive clones were detected using alkaline phosphatase-conjugated goat anti–mouse secondary antibody (Boehringer-Mannheim Corp., Indianapolis, IN) and standard color reaction (Bio-Rad Laboratories, Richmond, CA). To isolate cDNAs with complete protein coding regions, a polymerase chain reaction (PCR)–generated radiolabeled probe made from the 5′ end of the Ov-3.5 partial cDNA (see Fig. 1 *a*) was used to screen an 8–12-h embryo cDNA library (complexity = 4 × 10⁵) (Brown and Kafatos, 1988) and a 3–12-h cDNA library (complexity = 5 × 10⁵) (Poole et al., 1985). The clones recovered from this screen did not have complete coding regions, so a second round of library screening was done using a radiolabeled PCR-generated probe from the 5′ end of the Em-2 cDNA clone (see Fig. 1 *b*) that was isolated from the 3–12-h embryonic cDNA library.

Fly Culture

Canton S wild-type flies incubated at 18°C on yeasted corn-meal agar medium were used for salivary gland preparations for in situ hybridizations to polytene chromosomes. RNA, genomic DNA, and embryos for antibody staining were obtained from Oregon R wild-type flies. Fly stocks carrying chromosomal deletions were maintained over balancer chromosome TM3 (Lindsley and Grell, 1968). The deletion chromosomes S87-4 and S87-5 were made by Elizabeth Knust and Jose Campos-Ortega and others, and were obtained, with their cytological descriptions, from Eric Weischaus.

In Situ Hybridization to Polytene Chromosomes

Salivary gland squashes were performed according to the protocols of Pardue (1986) and Spierer et al. (1983). Biotinylated probes were made from cDNA clones Ov-2.4 and Ov-3.5 using random hexamer primed synthesis (Feinberg and Vogelstein, 1983).

Southern Blot Hybridizations

Genomic DNAs were prepared by standard methods (Bender et al., 1983). DNA was digested to completion with restriction endonucleases (HindIII, PstI, and PvuII) and electrophoretically fractionated on duplicate 0.7% agarose gels in TAE (40 mM Tris-Acetate/1 mM EDTA) running buffer. 10 μg of DNA was loaded in each lane. The DNAs were transferred to Nytran (Schleicher & Schuell, Inc., Keene, NH), fixed, hybridized, and washed by standard techniques (Sambrook et al., 1989). One blot of Oregon R wild-type genomic DNA was hybridized with a random hexamer radiolabeled probe made from a gel-isolated 830-bp HindIII fragment of the Em-3 cDNA clone. A duplicate blot was hybridized simultaneously with the same Em-3 probe and a control probe made from a gel-isolated 745-bp PstI–SalI fragment of the ftz gene (Laughon and Scott, 1984) labeled to the same specific activity. PvuII-cut DNAs from deletion fly stocks were electrophoretically separated, blotted, and probed simultaneously with the Em-3 and *ftz* probes, as above.

Northern Blot Hybridizations

Total RNA preparations from embryos, larvae and pupae were obtained by the hot phenol method (Jowett, 1986). Total RNA preparations from adult flies were made using the guanidine hydrochloride method (Chirgwin et al., 1979). Polyadenylated RNAs were isolated using oligo-dT cellulose (Collaborative Research Inc., Lexington, MA). 3 μg of each total RNA sample was denatured and size-fractionated on a 0.9% agarose formaldehyde gel (Sambrook et al., 1989). The RNAs were transferred to Nytran, fixed, hybridized, and washed by standard techniques. Samples of the total RNA from different developmental time points were used for the developmental Northern blot and this blot was probed with a radiolabeled PCR-generated probe from the Em-3 cDNA clone. The two primers used for PCR amplification were: (*a*) a 24-mer in the SP6 RNA polymerase promotor of the pNB40 vector (Brown and Kafatos, 1989); and (*b*) the 18-mer CAAATC-GCGAGGAGTTGT which hybridizes to nucleotide positions 963–946 in the 95F MHC nucleotide sequence (see Fig. 4 *a*). The Northern blot of poly A selected RNAs was probed with an antisense radiolabeled ribroprobe starting at the T7 RNA polymerase promoter in the pNB40 vector and extending to the BamHI site at position 1266 in the 95F MHC nucleotide sequence.

DNA Sequencing

The longest cDNA encoding the complete protein, Em-3, was sequenced on both strands by the dideoxy chain-termination method (Sanger et al., 1977) using a combination of subcloning and internal priming with synthetic oligonucleotides.

Antibody Production and Purification

The Ov-3.5 partial cDNA was subcloned into the unique EcoRI site of the pGEX.1 expression vector, transformed into bacterial host HB101, and induced with IPTG to generate a glutathione S-transferase-95F myosin fusion protein (Smith and Johnson, 1988). This protein was insoluble and was purified by SDS-PAGE. The expressed protein was cut from the gel, electroeluted, TCA precipitated and resuspended in PBS. A rabbit was injected with this immunogen using standard immunization techniques and antiserum was obtained that reacted with the fusion protein on immunoblots. A full-length *Drosophila* 95F MHC protein unfused to bacterial sequences was expressed using the pET 5a vector (cloning will be described elsewhere) and was used to affinity purify the polyclonal antiserum using standard techniques (Harlow and Lane, 1989). The affinity purified antibody was used at an approximate concentration of 3 μg/ml on immunoblots. Culture supernatant of 3C7 hybridoma cells was diluted 1:10 in PBS for incubation with blots or embryos.

Cleveland Peptide Digests

Proteolytic peptide digests were generated according to the method of Cleveland et al. (1977). Protein encoded by the cloned gene was bacterially expressed from a pET 5a vector containing the complete protein-coding region from the Em-3 canonical cDNA. *Drosophila* 95F protein was enriched by F-actin chromatography of embryonic extracts. Both bacterially expressed and *Drosophila* proteins were size fractionated by 7.5% SDS-PAGE and the 140-kD bands were cut from the gels. The 205-kD band of nonmuscle myosin from embryonic extracts was also excised. Gel slices containing equal amounts of protein (as judged by Coomassie staining) were placed in the wells of a 12.5% SDS-polyacrylamide gel, and were overlaid with 0.1 or 0.02 μg of chymotrypsin. Samples were run to the stacking gel-resolving gel junction and permitted to sit without current for 30 min before electrophoresis was resumed. The gel was Coomassie stained, then silver stained (Merril et al., 1981).

Immunoblots and Immunostaining of Embryos

Protein samples were fractionated on 7.5% SDS–polyacrylamide gels and electroblotted onto nitrocellulose in 50 mM Tris/380 mM Glycine/0.1% SDS/20% methanol. Blots were blocked with PBS + 1% Tween 20 + 1% BSA and incubated in primary antibody (1:10 dilution of 3C7 mAb culture supernatant or ∼3 μg/ml of affinity-purified polyclonal antibody). Immunoreactive bands were visualized using alkaline phosphatase-conjugated goat anti–mouse secondary antibody (with the 3C7 mAb primary; Boehringer-Mannheim Corp.), or using alkaline phosphatase-conjugated goat anti–rabbit secondary antibody (with the polyclonal antibody;

Promega Corp., Madison, WI), and standard color development reagents (Bio-Rad Laboratories, Richmond, CA). Embryos were collected, fixed, and stained with antibodies as previously described (Miller et al., 1989).

Sequence Analysis

Most of the sequence analysis was done using the programs in the GCG software package, versions 5.1 and 7.1 (Devereux et al., 1984). The nucleotide sequence was entered into the computer using the SEQED program and translated using both the MAP and TRANSLATE programs. The DNA sequence was analyzed using the TESTCODE and CODONPREFERENCE programs to confirm that the codon usage in the 95F MHC open reading frame is both nonrandom and exhibits Drosophila codon bias. Database searches were done using the TFASTA program, and the sequences of similar molecules were obtained with FETCH. Common motifs, such as phosphorylation sites, were found using the MOTIFS program. Searches for sequence motifs found in other myosins were performed using the BESTFIT program. Hydrophobicity and secondary structure predictions were made using PEPTIDESTRUCTURE and PLOTSTRUCTURE. Comparisons to other myosins were accomplished using the COMPARE and DOTPLOT programs, counting conserved amino acid substitutions using the PAM matrix, as well as identical matches. The dendrogram representing the degree of similarity between different MHC heads was obtained using PILEUP. Prediction of α-helical coiled-coils was done using the algorithm of Lupas et al. (1991).

Results

Cloning and Sequencing of Drosophila 95F Myosin

To clone Drosophila 95F MHC, we screened a λgt11 Drosophila ovary cDNA expression library with the 3C7 mAb (Miller et al., 1989). This antibody was previously shown to react with the single major 140-kD protein in whole embryo extracts that bound to F-actin affinity columns and eluted with ATP. Seven clones that reacted with 3C7 antibody were isolated; these seven were also recognized by a polyclonal mouse antiserum raised against the SDS-PAGE purified 140-kD protein band (Miller et al., 1989). Six of the seven clones appeared to be closely related, both by restriction mapping (Fig. 1 a) and by Southern blot hybridization (data not shown). The seventh cDNA neither crosshybridized nor shared the restriction fragments of the other clones, and was not studied further. Additional fine-scale restriction mapping of the six related cloned DNAs revealed that they apparently differ only in the 5' extent of the cDNA (Fig. 1 a). Indeed, three of the clones, Ov-1.4 A, B and C, have identical restriction maps, and probably represent multiple isolates of the same clone.

We did two experiments to demonstrate that the cDNAs we cloned correspond to the 140-kD ATP-sensitive actin-binding protein recognized by the mAb 3C7. First, we generated a new rabbit antibody to the bacterially expressed protein encoded by the ovarian cDNA clone, Ov-3.5. This antibody and the 3C7 mAb were used to probe duplicate Western blots of Drosophila embryo extract and bacterially expressed fusion protein (Fig. 2 a). Both antibodies recognize the same size proteins and proteolysis products expressed in bacteria (Fig. 2 a, lanes 1 and 4) and in whole embryo extract (lanes 3 and 6). In addition, both antibodies recognize partially purified protein from Drosophila embryos (lanes 2 and 5; V. Mermall and K. Miller, unpublished data), confirming that the protein undergoing purification and biochemical analysis is encoded by this gene. Second, we performed Cleveland partial-proteolysis to show that the

same proteolytic cleavage pattern is generated from Drosophila 95F MHC from embryos and from the bacterially expressed product of the canonical (Em-3) cDNA (Fig. 2 b). Of the 12 visible peptide bands, all but one were the same in the two samples; a ~40-kD band seen only in the bacterially expressed protein probably represents a difference in posttranslational modification. In contrast, a similarly digested sample of Drosophila nonmuscle MHC has a distinctly different proteolytic pattern (Fig. 2 b, lane 1). No peptide bands are contributed by the chymotrypsin enzyme (Fig. 2 b, lane Ch). Thus by two criteria, we conclude that our cDNAs encode the ~140-kD ATP-sensitive actin-binding protein recognized by the 3C7 mAb.

To determine whether the 95F MHC gene is unique, we did in situ hybridizations to polytene chromosomes and genomic Southern blots. The two longest clones, Ov-2.4 and Ov-3.5, were used as probes for in situ hybridization to Drosophila polytene chromosomes. Both hybridized to a single site located at 95F on the polytene map (data not shown), suggesting that they are the products of a single-copy gene. Genomic Southern blot hybridization was done to confirm that the gene is single copy. A blot of adult Drosophila genomic DNA was probed with a DNA fragment from the carboxyl end of the head, and shows a single band for each restriction digest (Fig. 3 a). An identical blot was simultaneously hybridized with the 95F MHC probe and a similarly-sized probe from the known single-copy fushi tarazu gene (first lane, band with arrow is fushi tarazu signal); both bands are of equal intensity. We conclude that the Drosophila 95F MHC gene is single copy in the haploid genome.

To confirm the cytological position of the 95F MHC gene, genomic Southern blots of DNAs from flies carrying deficiencies were probed with the 95F MHC and ftz probes described above (Fig. 3 b). Since the deletion stocks are maintained as heterozygotes with the balancer chromosome TM3, which is wild type for the 95F gene, stocks with deletions that remove the 95F gene were expected to have half the hybridization signal intensity of a balancer stock control. The two deletion stocks are missing the 95F gene, as demonstrated by the half-intensity signal relative to the ftz control signal (Fig. 3 b, lanes 1 and 2) and relative to the intensity of signal from a balancer fly stock (Fig. 3 b, lane TM3/TM6). The chromosomal deletions have been reported to overlap in the region from 95F6 to 96A2, which is consistent with the polytene location determined by in situ hybridization to polytene chromosomes.

Since the cDNA clones obtained from the ovarian expression library were not full length, we screened additional cDNA libraries for clones encoding the complete protein. The Ov-3.5 kb clone was used to screen 8–12-h (Brown and Kafatos, 1988) and 3–12-h (Poole et al., 1985) Drosophila embryo cDNA libraries. A longer clone was isolated from each library, but neither of these were full length. A new probe was made from the 5' end of the longest clone (clone Em-2, from the 3–12-h library), and was used to screen the 8–12-h library, yielding clones that contained the complete coding region. In all, five embryonic cDNA clones were isolated (Fig. 1 b), and three of these appear to have complete coding regions. The clone that extends the farthest 5' (clone Em-3, from the 8–12-h library) was sequenced on both strands in its entirety. Although the entire coding region is
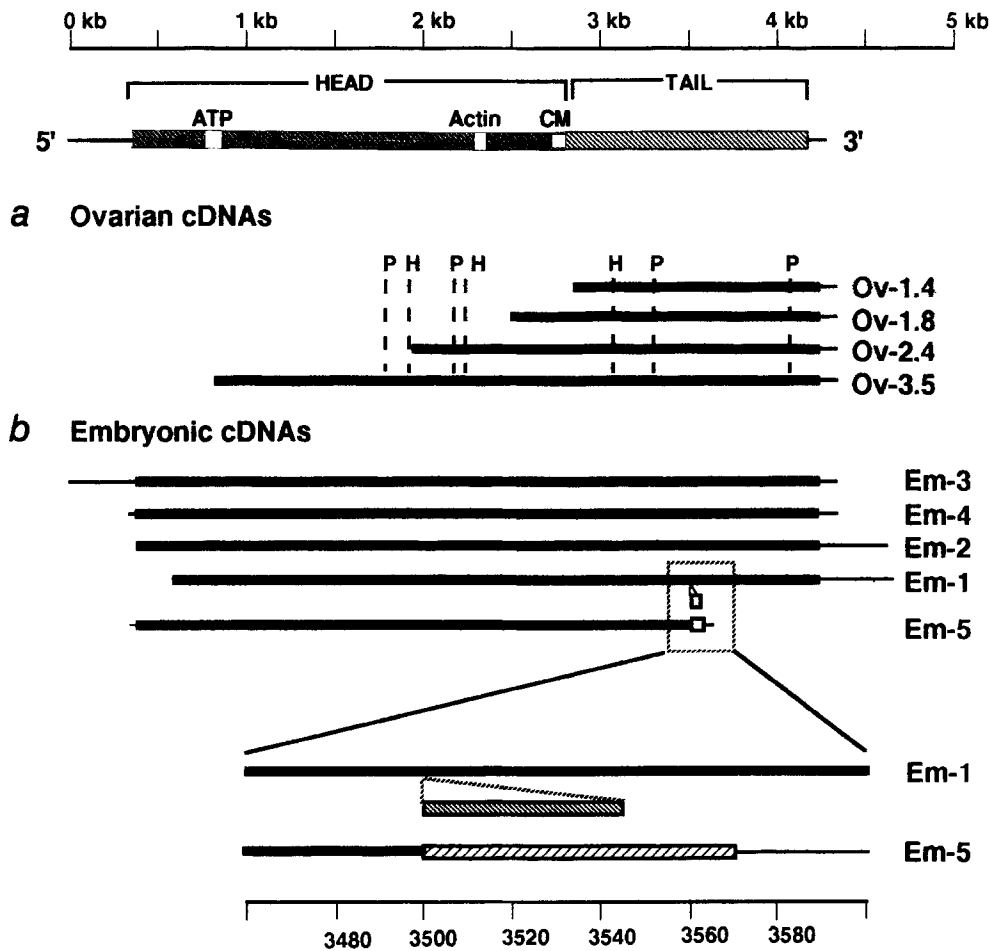
Figure 1. Schematic representation of 95F MHC cDNA clones. Coding regions are depicted as thick horizontal lines and 5' and 3' untranslated regions as thin lines. (*a*) Six 95F MHC cDNA clones were obtained by immunoscreening a *Drosophila* ovarian cDNA library. Vertical dashed lines designate sites cut by PstI (*P*) or HindIII (*H*) restriction endonucleases. (*b*) Five clones with the nearly complete (*Em-1* and *Em-2*) or complete (*Em-3, Em-4,* and *Em-5*) coding region were obtained from embryonic cDNA libraries. All ovarian cDNAs and the Em-3 and Em-4 embryonic cDNAs share the same restriction map, while the Em-1, Em-2, and Em-5 cDNAs have slightly different patterns at the 3' end. The inset shows an enlarged schematic of the divergent tail regions in Em-1 and Em-5. The 45-bp Em-1 inset is shown as a backslashed bar at position 3,500, and the 102-bp alternate tail of Em-5 is shown as a slashed bar (coding region) and thin line (3' untranslated region). Locations of predicted functional regions are abbreviated as follows: *ATP,* ATP-binding site; *Actin,* actin-binding site; *CM,* calmodulin/myosin light chain-binding site.
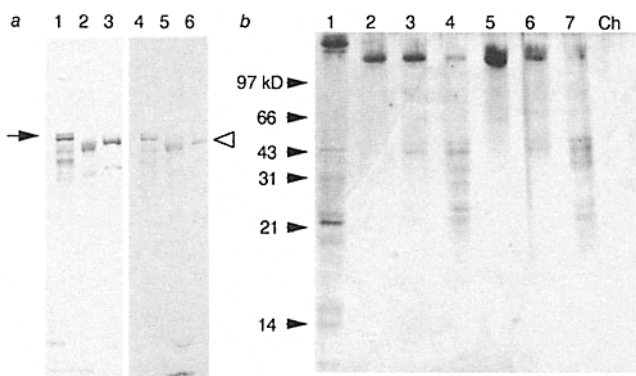

Figure 2. Comparison of *Drosophila* and bacterially expressed cloned 95F MHC proteins. (*a*) Duplicate Western blots were prepared from (lanes *1* and *4*) fusion protein made from the Ov-3.5 cDNA expressed from the pGex vector; (lanes *2* and *5*) partially purified 95F myosin from embryos; and (lanes *3* and *6*) whole homogenate from 8–10-h-old embryos. The duplicate blots were probed with the 3C7 mAb which was raised against the 140-kD embryonic actin-binding protein (lanes *1, 2,* and *3*) or affinity-purified polyclonal antibody which was raised against the pGEX.Ov-3.5 fusion protein (lanes *4,5,* and *6*). The patterns are the same, indicating that the cDNA encodes a protein that shares epitopes with the 140-kD embryonic protein. (*Arrow*) Bacterially expressed fusion protein; and (*open arrowhead*) 95F MHC from embryos. (*b*) Pep-

tide patterns produced by chymotrypsin-catalysed proteolytic cleavage are compared. *Drosophila* conventional MHC protein (lane *1*), *Drosophila* 95F MHC protein (lanes *2, 3,* and *4*), and bacterially expressed cloned 95F protein (lanes *5, 6,* and *7*) were loaded without the addition of chymotrypsin (lanes *2* and *5*), with 0.02 μg chymotrypsin added (lanes *3* and *6*), or with 0.1 μg chymotrypsin added (lanes *1, 4,* and *7*). The *Drosophila* 95F MHC and the bacterially expressed protein from the cloned cDNA have nearly identical proteolytic patterns. Conventional double-headed MHC (predominantly nonmuscle MHC) has a distinctly different pattern. Lane *Ch* is a control lane of 0.1 μg chymotrypsin only.

contained on this clone, we do not know if the entire transcript is present, since S1 protection and primer extension experiments have not been done.

The sequence is shown in Fig. 4 *a.* It has a 340-bp 5' untranslated region with no significant coding potential. There are only two ATG codons before the large open reading frame. The first is in poor context for protein initiation and encodes only a 9-amino acid peptide before a termination TAA. The second is in good context for initiation of protein synthesis in *Drosophila* (CAAAATG; consensus is C/AAAA/CATG; Kozak, 1984; Cavener, 1987). The open reading frame following this ATG extends for 3.8 kb, and encodes a 142-kD protein. The termination codon is followed by a
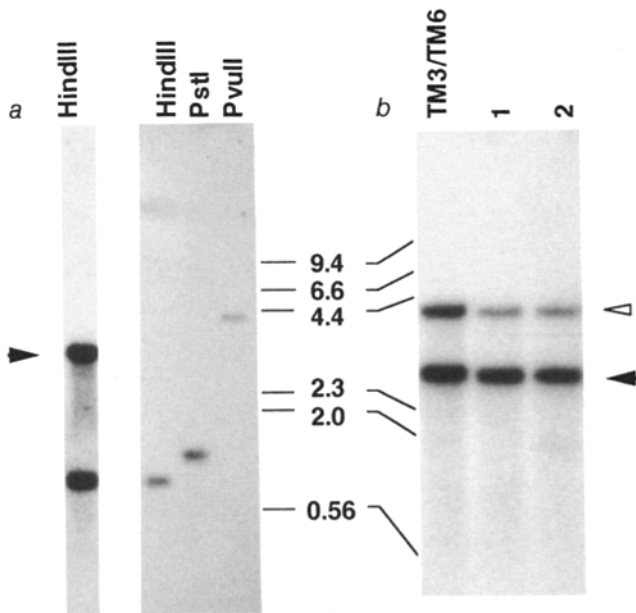
Figure 3. Genomic Southern analysis indicates that the 95F MHC gene is unique. (*a*) A Southern blot of restriction endonuclease digested genomic *Drosophila* DNA was hybridized to an 830-bp radiolabeled probe from the head–tail junction region of a 95F MHC cDNA (*right panel*). Each lane has a single hybridizing band. On a duplicate gel, a size-matched control probe from a known single-copy gene (*fushi tarazu*) was hybridized with the 95F MHC probe to compare signal intensity (*left panel*). The 95F MHC band is similar in intensity to the *fushi tarazu* band (marked by *arrow*), confirming that this is a single copy gene. (*b*) A Southern blot of PvuII-digested genomic DNAs from two fly stocks bearing different deficiency chromosomes heterozygous to balancer chromosome TM3 was hybridized with both 95F MHC and *ftz* probes, as above. The chromosome S87-5 is deleted for the region 95F7 to 96A18 (lane *1*); S87-4 is deleted for 95E8 to 96A2 (lane 2); together they place the 95F MHC gene in the region 95F7 to 96A2.

short 3′untranslated region and a stretch of the polyadenylated tail which was preserved during library construction. Since this cDNA has the same restriction map as all the ovarian cDNAs and one other embryonic cDNA (see below), it was designated the canonical form and we have compared all the other cDNAs we obtained to this complete clone.

In total, nine independent cDNA clones representing four variant forms were isolated; they are related to each other as shown in Fig. 1, *a* and *b*. All four clones from the ovarian cDNA library and two of the clones isolated from the 8–12-h-old embryo cDNA library (Em-4 and the canonical clone Em-3) are of the same form, differing only in the 5′ extent of DNA sequence. Two other embryonic cDNAs are polyadenylated either 297 (clone Em-1) or 263 bp (clone Em-2) further downstream (Fig. 4 *d*). One of these, Em-1, also has a 45-bp insert in the coding region that maintains the open reading frame. This insert alters one codon and adds 15 novel codons in the middle of the tail (Fig. 4 *b*). The final clone, Em-5, diverges in the tail at the same position as the Em-1 insertion, substituting 102 novel nucleotides before the poly A tract. This encodes a truncated protein with 26 novel amino acids at the carboxyl terminus (Fig. 4 *c*; and see Fig. 1 *b*).

Sequence Analysis

The canonical Em-3 cDNA includes the entire open reading frame for the *Drosophila* 95F MHC. The single long open reading frame has *Drosophila* codon bias and encodes a 142-kD protein that shows amino acid conservation with MHC proteins. Database searches with both the nucleotide and deduced amino acid sequences (Pearson and Lipman, 1988) revealed substantial similarity to all members of the MHC gene family, but no exact homolog of this gene in any system. In all cases, similar regions are confined to the head; no sequence in the database was significantly related to the tail. Comparison to other MHCs shows similarity throughout the head, with a high degree of conservation in the ATP-binding and actin-binding regions (Fig. 5, *a* and *b*), as well as other regions of unknown function. The sequence DALAK, which is conserved in many other unconventional myosins, is also present (Fig. 4, underlined amino acids 421–425). Two reactive thiols, SH1 and SH2, are found in many muscle MHC heads (Warrick and Spudich, 1987); only SH2 is present in 95F MHC (amino acid 686). A putative calmodulin/myosin light chain-binding site (Fig. 5 *c*) exists near the junction with the tail. Comparison of the amino acid sequences of the 95F MHC and another unconventional myosin, the mouse *dilute* MHC, illustrates the difference in sequence conservation between the head and tail regions (Fig. 6). Most of the head region exceeds the threshold of 14–amino acid matches in each window of 25 amino acids (counting conservative substitutions), while very little of the tail does. Conservation between three of the *dilute* calmodulin-binding motifs and the single motif in 95F MHC is seen as three short diagonals near the head-tail junction (Fig. 6, *arrows*).

Since the 95F MHC tail sequence is unlike any entry in sequence databases, we used computer algorithms to look for regions predicted to have specific secondary structures. Two regions were predicted to have notable features. First, the algorithms of Chou-Fasman and Garnier-Osguthorpe-Robson predicted that a ∼110–amino acid stretch near the start of the tail has strong potential to form an α-helix. Portions of the α-helical region are further predicted to form a coiled-coil (Lupas et al., 1991). A region of 13 successive heptad repeats (amino acids 914–1,008) has >80% predicted probability of forming a coiled-coil (underlined in Fig. 4 *a*). This potential coiled-coil region is much shorter than in conventional type II or *dilute* class MHCs, but it is long enough to include the 4–5 heptads which are theoretically sufficient to form a stable coiled-coil (Cohen and Parry, 1990). Regions of predicted α-helical coiled-coil structure (see below) are characterized by multiple weak diagonals in the comparison with *dilute* (Fig. 6); at reduced stringency, the number and length of the diagonals increase, but do not exceed similar comparisons with other proteins containing α-helical regions. For example, a COMPARE dotplot (data not shown) of 95F MHC vs human α-spectrin (Moon and McMahon, 1990) yields a pattern of weak diagonals that is indistinguishable from 95F MHC vs *dilute* in the regions with predicted α-helical structure. Second, Hopp-Woods hydrophilicity prediction reveals a 7–amino acid region (PILLVAG, amino acids 1,187–1,193) with significantly more hydrophobic character than the rest of the tail. This region is not present in the Em-5 cDNA, but the novel 26 amino acids at the COOH terminus include a 13–amino acid region with even stronger

a   1 gagttcgactcgactcatccaacgcgacggtacgcgattcgagacgctcgtgctggtcgccttccgatcgccgcttgaatcttaaatctaaactgttaaaaccgaggcaaacaaacacaa

    121 cacccagctgaaaatacttcaagtgttacaaacttttggaaataggtgtatcgtggctataaattgcaagttctaaattcgaaccaaacaaaatccacgcaggaagatcgattccaagtg

    241 gattgcgccttcatcaatgtgttcgtgtgcgtccccggaaacataatccccatttgccagtaacttacgcacgtacagatttccagccgtcttaagtgcgcaaaatgttggaggacaccc
                                                                                                             M L E D T Q   6

    361 aactggtgtgggtgcgagatgccgccgagggctatatacagggtcggatcacogagatcggcgccaaggagttcgaggttacgcccaccgatcgcaagtatccgaagcgcacgtgtcact
            L V W V R D A A E G Y I Q G R I T E I G A K E F E V T P T D R K Y P K R T C H F   46

    481 tcgatgatatccactcttcgtgcgatggacccaggatcacgatgacaactgcgaactcatgctgctcaacgaggccacatttctggacaatttgaaaacgcgttactacaaagacaaga
            D D I H S S C D G P Q D H D D N C E L M L L N E A T F L D N L K T R Y Y K D K I   86

    601 tctacacatatgtggccaacatactaatcgccgtgaatccgtatcgtgagataaaggagctctacgcacccgataccatcaagaagtacaatggtcgttccctgggtgaactgcctcctc
            Y T Y V A N I L I A V N P Y R E I K E L Y A P D T I K K Y N G R S L G E L P P H   126

    721 atgtctttgctattgcggataaagccatacgcgatatgcgggtatacaaattgtcccagtcgatcatcgtgtccggagagtccggcgccggcaagacggagtccaccaaatacctgctca
            V F A I A D K A I R D M R V Y K L S Q S I I V S G E S G A G K T E S T K Y L L K   166

    841 aatacctgtgctattcgcacgacagtgccggtcccatagagaccaagatattggatgccaatccogtgctggaagcctttggcaatgccaagacaaccgaaacaacaactcctcgcgat
            Y L C Y S H D S A G P I E T K I L D A N P V L E A F G N A K T T R N N N S S R F   206

    961 ttggaaagttcattgaggtgcactacgatgccaagtgccagatggtggggtacatatcgcactacttgctggagaagcgtcgcatctgtaogcagcgccgaggagcggaaactatc
            G K F I E V H Y D A K C Q M V G G Y I S H Y L L E K S R I C T Q S A E E R N Y H   246

   1081 atgtttctacatgctcctcgccggagctccgcagcagctgcgcgataagctgagcttgggcaagccagatgattacaggtatctctctggctgcacacagtactttgcgaacgccaaga
            V F Y M L L A G A P Q Q L R D K L S L G K P D D Y R Y L S G C T Q Y F A N A K T   286

   1201 cggagcagctaataccgggctcacaaaagtcgaagaatcaccagcagaagggtccacttaaggatccgattatcgatgattatcagcacttccacaaccttgacaaggccttgggtcgtc
            E Q L I P G S Q K S K N H Q Q K G P L K D P I I D D Y Q H F H N L D K A L G R L   326

   1321 tgggactctggatacggagaagctgggcatctattcgctggtggcagctgtgctccacctgggcaacattgcttcgaggagataccgatgatgtgcgcggtggctgccaggtttcgg
            G L S D T E K L G I Y S L V A A V L H L G N I A F E E I P D D V R G G C Q V S E   366

   1441 aggcttcggagcagtcgctgactatcaccagtggcctgttgggtgtagatcaaactgagctgcgcactgcctggtatcccgagtgatgcagagcaaggagcggcttaaggcacgg
            A S E Q S L T I T S G L L G V D Q T E L R T A L V S R V M Q S K G G F K G T V   406

   1561 ttatcatggttcctctaaagatctacgaggcaagcaatgctcgggatgcgttggccaaggccatctacagtcgactattcgatcgcattgtgggcctgatcaaccagagcattcccttcc
            I M V P L K I Y E A S N A R D A L A K A I Y S R L F D R I V G L I N Q S I P F Q   446

   1681 aggcttcaaacttttacatcggcgtgctcgatattgctggatttgagtacttcacggtgaactccttcgagcagttctgcatcaactactgcaatgagaagctgcagaagttttcaacg
            A S N F Y I G V L D I A G F E Y F T V N S F E Q F C I N Y C N E K L Q K F F N D   486

   1801 ataatatattgaaaaacgagcaggagctgtacaaacgcgaaggcctcaatgtgcccgagatcacttcacggataaccaggacatcatcgagctgatcgagcgcaagtcgaatggcatct
            N I L K N E Q E L Y K R E G L N V P E I T F T D N Q D I I E L I E A K S N G I F   526

   1921 ttacgctgctcgacgaggagtccaagcttccgaagccctcgtactcgcattcaccgccgaggtgcacaagtcctgggcgaatcactatcgcctgggtctgccgagatcatcgcgactaa
            T L L D E E S K L P K P S Y S H F T A E V H K S W A N H Y R L G L P R S S R L K   566

   2041 aggcccaccgcactctgcgcgatgaggaggggcttcctggtgcgtcactttgccggagctgtgtgctacaacacggagcagttcattgagaagaacaacgacgccctgcatgcttcgctgg
            A H R T L R D E E G F L V R H F A G A V C Y N T E Q F I E K N N D A L H A S L E   606

   2161 agggtttggtccaggagtgcgataatccccctgctgcagaccctttccttcgggaagtagtacttcggttcgtggaaagcttaactttatatccgtgggatctaagttcaagactcagc
            G L V Q E C D N P L L Q T L F P S G S S T S V R G K L N F I S V G S K F K T Q L   646

   2281 tgggcgagctaatggagaaactggagcagaacggcaccaacttcatacgctgcatcaagccaaacgcaagatgatcgaccgacagtttgagggcagcctggcgctggccaactgaagt
            Q E L M E K L E Q N G T N F I R C I K P N S K M I D R Q F E G S L A L A Q L K C   696

   2401 gttcgggcacaatttcggtgctggaactcatggagcacggttatccatcgcgtgttctctttgccgatctctacagcatgtacaaatcggtgctgccgccggaacttgtttcactgcgg
            S G T I S V L E L M E H G Y P S R V L F A D L Y S M Y K S V L P P E L V S L P A   726

   2521 cgcgcaccttctgcgaagccatgttccagtccctcaatctgagcgccaaggacttcaagttcggtatcaccaaggtcttcttccggcctggtaaattcgtggagttcgatcgcatcatgc
            R T F C E A M F Q S L N L S A K D F K F G I T K V F F R P G K F V E F D R I M R   766

   2641 gctctgatccggagaatatgctggccatagtggccaaggtcaagaagtggctgatccgatcgcgtgggtcaagtccgctctgggagccctctgcgtgatcaagctgcgtaatcgaatca
            S D P E N M L A I V A K V K K W L I R S R W V K S A L G A L C V I K L R N R I I   806

   2761 tctaccgcaacaagtgtgttctaatagcccagcgtattgcacgtggttcctggccggaagcagcatcgcccacgctaccagggtattggcaagatcaacaagatccggactaacaccc
            Y R N K C V L I A Q R I A R G F L A R K Q H R P R Y Q G I G K I N K I R T N T L   846

   2881 tcaagacaatcgaaatagcaagtggactgaagatgggtcgtgaagagattattagcggggtaaacgatatctataggcagatcgatgatgccatcaagaaatcaagatgaatccacgca
            K T I E I A S G L K M G R E E I I S G V N D I Y R Q I D D A I K K I K M N P R I   886

   3001 ttactcaacgggaaatggactcgatgtacaccgtggtcatggccaacatgaacaagctaacagttgatcgtgaacaccaagcttaaggagcagcaacaggccgaggagcaggagcgtctgc
            T Q R E M D S M Y T V V M A N M N K L T V D L N T K L K E Q Q Q A E E Q E R L R   926

   3121 gcaagatccaagaggccctggaggccgagcgggccgccaaggaggcggaggagcagcgccagcgcgaggaaatcgaaaacaagcgactaaaagcggagatggaacacgacgaaaggcag
            K I Q E A L E A E R A A K E A E E Q R Q R E E I E N K R L K A E M E T R R K A A   966

   3241 cggaagctcagcgcctgcgtcaagaggaggaggatagacgtgccgccttggcgctgcaggagcagttggaaaaggaagccaaggatgatgccaaataccggcagcagctcgaacaggaac
            E A Q R L R Q E E E D R R A A L A L Q E Q L E K E A K D D A K Y R Q Q L E Q E R   1006

   3361 gtcgcgatcacgagttggccctgcgtttggccaacgagtccaatggccaggtggaggatagtccaccagttatacgcaatggtgtcaatgacgcgtctcccatgggcccaaacaaattga
            R D H E L A L R L A N E S N G Q V E D S P P V I R N G V N D A S P M G P N K L I   1046

   3481 tcagatccgaaaatcttcgagcccaacaacaggccctgggcaagcagaagtacgacttgtccaagtggaagtactcggagctgcgtgatgctatcaacacgtcctgtgatattgagctgc
            R S E N L R A Q Q Q A L G K Q K Y D L S K W K Y S E L R D A I N T S C D I E L L   1096

   3601 tggaggcctgccgtcaggagttccatcgccgcttgaaggtgtaccacgcctggaaggcaaagaaccgcaagcgcaccaccatggatgaaaacgagcgtgcgccacagcgttatgcagag
            E A C R Q E F H R R L K V Y H A W K A K N R K R T T M D E N E R A P P S V M E A   1126

   3721 ctgccttcaagcagccgccctttggttcagccaatccaggagattgtgaccgcccagcatcgctacttccggattccttcatgcgagccaatgcgccggacaacaccaagcgcggcctgt
            A F K Q P P L V Q P I Q E I V T A Q H R Y F R I P F M R A N A P D N T K R G L W   1166

   3841 ggtatgcccactttgatggccagtggatcgcacgccagatggaactgcatgcggacaagccgccaattctactggtagccgaacggacgacatgcagatgtgcgaactgagcctggagg
            Y A H F D G Q W I A R Q M E L H A D K P P I L L V A G T D D M Q M C E L S L E E   1206

   3961 agacgggtctgacgcgcaagcgcggcgccgagatcctggagcacgagttcaatgcgaatgggagcgcaacggggcaaggcttataagaatttgggtcggcaaaaccgaatgggccgg
            T G L T R K R G A E I L E H E F N R E W E R N G G K A Y K N L G A A K P N G P A   1246

   4081 ctgcagcaatgcagaaacaacagtagacccccaatacgacaaacaggcactgcattgatcaaatacgttgccgtttcgatggtcataggtcaaaattggtaactaacgaaactcgaaaa
            A A M Q K Q Q   1286

   4201 taagttttttgtcagtcaaatttttacggaaagcgaccacgtccataaaacatttttagtaaatataactatatttgattag 4280

b 3491 tcagttttcacaagttgtgtcaaacattgcttcgcggtacttgaataaatccgaaaatcttcgagcccaacaacaggccctgggcaagcagaagtacgacttgtccaag
  1047 S F S Q V V S N I A S R Y L N K S E N L R A Q Q Q A L G K Q K Y D L S K

c 3491 tcaggtatagcacgctttatgaattacctatgtctaccactcttgtaaattttgtcaatcttttcttgctttcccaaaaacactaattggataatgaatgaataccat
  1047 R Y S T L Y E L P M S T T L V N F V N L F L L S Q K H 1078

d 4291 aaatgtgatttaatttggattccgatcccattgcccaatcatttgtgcgcgcccggactgttttggggccttatacacttagctatacaaatatatatctctaattaagtgaatcttatg

  4401 atgtaaaaggaaacggcaagttgtatttagtatttaactaattaccatttaaattgtgtatttagttcactgcgatgtatttacataattgaatatatatatatatatatttttacagtat

  4521 tttaaaaataaatatttgaagtgaatatttgaatgaaatatatcacaacttaaacgt 4578

Figure 4. Nucleotide and deduced amino acid sequence of *Drosophila* 95F MHC. The sequence of the canonical Em-3 cDNA and divergent regions from the Em-1, Em-2, and Em-5 cDNAs are shown. The nucleotide sequence is numbered on the left and the amino acid sequence is numbered on the right. (*a*) The complete canonical Em-3 cDNA sequence was obtained. Conserved portions of the ATP-binding site (amino acids 137–166), DALAK motif (421–425), actin-binding region (647–666) and calmodulin/myosin light chain-binding region (815–826) are underlined. The tail region which is predicted to form an α-helical coiled-coil (probability >80%) (Lupas et al., 1991) is underlined. The site of sequence insertion (Em-1) and divergence (Em-5) is indicated by an asterisk at nucleotide position 3,485. (*b*) The altered sequence in the coding region of the Em-1 cDNA tail is indicated by a bar. (*c*) The divergent tail of the Em-5 cDNA is indicated by a bar. (*d*) The 3' untranslated region of Em-1 and Em-2 extends beyond the end of the canonical terminus. The end of the Em-2 cDNA is shown by a vertical arrowhead at position 4,543. These sequence data are available from EMBL/GenBank/DDBJ under accession number X67077.

hydrophobicity (STTLVNFVNLFLL, amino acids 1,057–1,069).

Additional sequence motif searches show 44 possible phosphorylation sites in the canonical protein; none are added or altered in the Em-1 isoform, but nine are lost and two new sites added in the Em-5 isoform. Preliminary assays on partially purified protein indicates that the 95F MHC is phosphorylated, but the sites and specificities are unknown

a ATP-Binding Site

b Actin-Binding Site

c Calmodulin/Myosin Light Chain-Binding Site

Figure 5. Comparison of sequence motifs that are conserved between *Drosophila* 95F MHC and other MHCs. Conserved portions of functional domains in the head sequences of the *Drosophila* 95F MHC and other MHCs are aligned. (*a*) A conserved portion of the putative ATP-binding site of 95F MHC is compared with those of *Drosophila ninaC* myosin-related protein (Montell and Rubin, 1988), *Drosophila* cytoplasmic MHC (Ketchum et al., 1990), *Drosophila* muscle MHC (Bernstein et al., 1983; Rozek and Davidson, 1983), and the consensus sequences for conventional and unconventional myosins (Pollard et al., 1991). (*b*) A conserved portion of the putative actin-binding region is compared to the same sequences as above. In *a* and *b* boxed letters occur in positions where the 95F MHC sequence is the same as one or more other sequences. (*c*) Calcium-independent calmodulin/myosin light chain-binding motifs (also termed IQ motifs; Cheney and Mooseker, 1992) are compared in the same *Drosophila* sequences as above and to the six repeat

motifs found in the mouse *dilute* MHC gene (Mercer et al., 1991). Boxed letters indicate the I/LQXXXRGXXXR calmodulin/myosin light chain-binding motif consensus sequence.
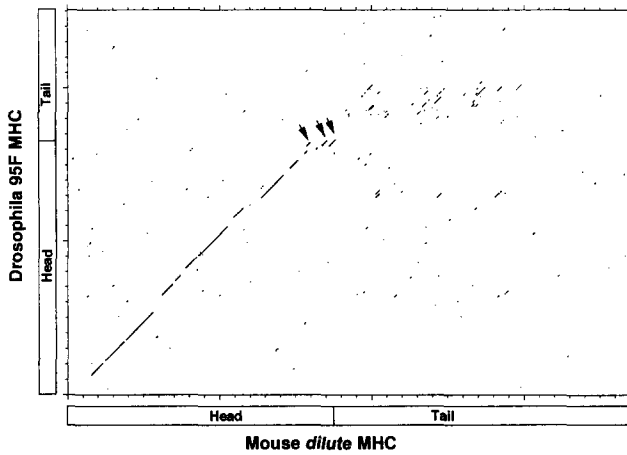
Figure 6. Comparison of *Drosophila* 95F MHC and mouse *dilute* MHC. The 95F MHC and mouse *dilute* MHC amino acid sequences were compared using the COMPARE program from the GCG software package with a stringency of 14 amino acids in a window of 25 amino acids, counting conservative substitutions (using the PAM matrix). The dot matrix representation shows the high level of conservation in the NH2-terminal heads, the lack of conservation in the tails, and three matching diagonals in the calmodulin-binding region near the head–tail junction (marked by *arrows*).

(V. Mermall, personal communication). There is no sequence element similar to the SH3 motif which has been implicated in ligand binding in other unconventional myosins and actin-binding proteins (Drubin et al., 1990). We can find no second actin-binding region, GPA-enriched region (as found in *Acanthamoeba* type I MHCs), or sequence similarity to the tail region conserved between *Acanthamoeba* type I and brush border MHCs (Pollard et al., 1991).

Expression of Drosophila 95F Myosin mRNA and Protein

Although *Drosophila* 95F MHC transcripts are present throughout the life cycle of the organism, their abundance and size vary (Fig. 7 *a*). 4.5-kb transcripts predominate in adults and early (not yet transcriptionally active) embryos. The maternally contributed transcripts in early embryos are apparently degraded, since transcripts are barely detectable from 2–4 h of development. Zygotic transcription of new 95F MHC transcripts begins at ~4 h, and these transcripts are present in increasing abundance through mid-embryogenesis. The embryonic transcripts are of varied size and are not clearly resolved by Northern analysis, but most fall within the 4.5–4.7-kb range. The diversity of transcripts is due at least in part to the use of multiple polyadenylation sites and alternate splicing, as seen in the collection of cDNAs. It is not known whether use of different promoters during embryonic development also contributes to this heterogeneity. Comparison of polyadenylated RNA from 0–3-h and 8–11-h-old embryos demonstrates that the predominant embryonic transcripts are significantly larger than those that are maternally inherited (Fig. 7 *b*).

The 95F MHC protein expression profile parallels mRNA expression, although, unlike the mRNA, the protein is present in 2–4-h-old embryos at a comparable level with younger embryos (Fig. 8 *a*). The 140-kD protein is detected at all stages examined, and the expression level appears highest
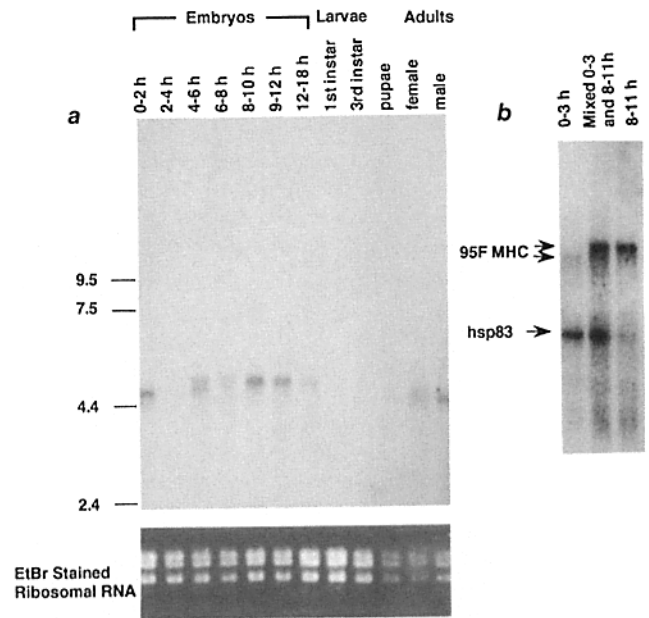

Figure 7. Transcript expression pattern. The 95F MHC transcript appears to be present throughout the life cycle, but both transcript size and abundance are developmentally controlled. (*a*) Total RNA from developmentally staged embryos, pupae, and adult female and male flies was blotted and hybridized with a 950-bp radiolabeled riboprobe from the 5' end of the Em-3 cDNA. A 4.5-kb transcript predominates in adults and 0–2-h-old embryos. This transcript nearly disappears during blastoderm formation, but is followed by the appearance of slightly longer transcripts in mid to late embryogenesis. The lower panel demonstrates that RNA, as visualized by ethidium bromide staining of ribosomal RNAs, is present in all lanes. (*b*) To resolve the size difference between very early and mid embryonic transcripts, polyadenylated RNAs from 0–3- and 8–11-h-old embryos were run mixed or alone and probed with both the 95F MHC riboprobe and an hsp 83 control riboprobe. The hsp 83 control has the same mobility in all three lanes, while the 95F MHC transcripts have different mobility.

during mid-embryogenesis and adulthood. An additional 120-kD protein band appears in 12–24 h embryos and adults when blots are probed with the affinity-purified polyclonal antibody (Fig. 8 *a*, designated by *arrow*), but is not apparent in identical samples incubated with the 3C7 mAb (data not shown). The size of this protein is consistent with the size of the protein encoded by the Em-5 cDNA clone (122 kD); the failure of the 3C7 mAb to detect this protein suggests that the epitope recognized by the 3C7 mAb is absent in the Em-5 encoded isoform.

Expression of 95F MHC was also examined in several tissues dissected from adult flies. Samples from adult heads (Fig. 8 *b*), gonads, and thoraxes (not shown) were size fractionated by SDS-PAGE, blotted, and probed with the affinity-purified rabbit polyclonal antibody to *Drosophila* 95F MHC. All samples contain the 140-kD protein, but in heads and gonads it comprises a larger proportion of total protein than in thoraxes. In heads, a protein 5 kD larger is also detected, and this larger isoform is also seen in late larval brains (data not shown). In addition to the 140-kD species, an immunoreactive 116-kD band is sometimes observed. This band is probably a proteolytic breakdown product, since it is present in variable amounts, and often increases in amount when
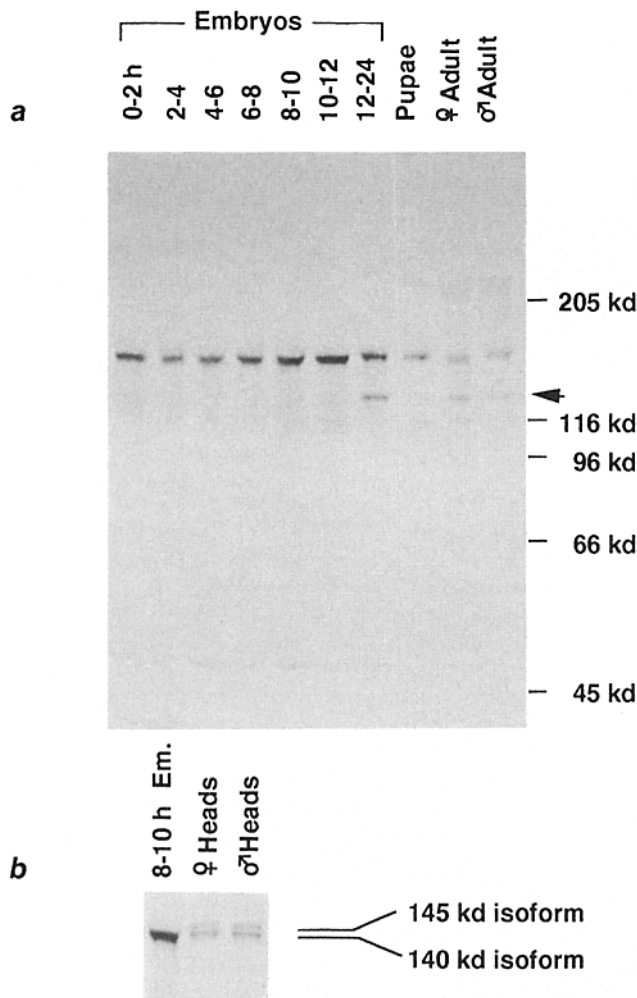
Figure 8. Drosophila 95F MHC protein expression pattern. The 140-kD 95F MHC protein isoform is expressed throughout the life cycle, while two other isoforms are seen only at particular stages or in a subset of tissues. (a) Equal amounts of protein from embryo, pupa, and adult fly homogenates were immunoblotted and probed with an affinity-purified anti-95F MHC polyclonal antibody. The major 140-kD protein isoform is seen in all lanes, with peak expression in mid to late embryogenesis and adulthood. A 120-kD isoform is seen prominently in 12–24-h embryos and weakly in older animals (arrow). (b) Equal amounts of protein from 8–10-h embryos and adult female and male heads were immunoblotted and probed with affinity purified anti-95F MHC polyclonal antibody. An ~145-kD isoform is apparent in the head samples, but absent from the embryonic sample.

Figure 9. Immunolocalization of 95F MHC in blastoderm embryos. Fluorescence micrographs of fixed embryos that were incubated with the 3C7 mAb and visualized with rhodamine-conjugated secondary antibody are shown. The 95F MHC protein is located in a punctate pattern throughout the cortical cytoplasm, but excluded from the nuclei. The panels show a cycle 11 blastoderm embryo in three focal planes; (a) at the surface, (b) 2 μm beneath the surface, and (c) 6 μm beneath the surface. Bar, 5 μm.

samples are stored (V. Mermall and K. Miller, unpublished observation).

Immunofluorescent Localization of 95F MHC in Embryos

To determine what embryonic processes this myosin may participate in, we have used immunolocalization to detect its distribution. At blastoderm, Drosophila 95F MHC is seen primarily in a punctate pattern throughout the cortical cytoplasm by immunofluorescence. The panels in Fig. 9 show different focal planes of the same syncytial blastoderm embryo during interphase of nuclear cycle 11. Punctate staining is seen throughout the cortical cytoplasm, excluded only by
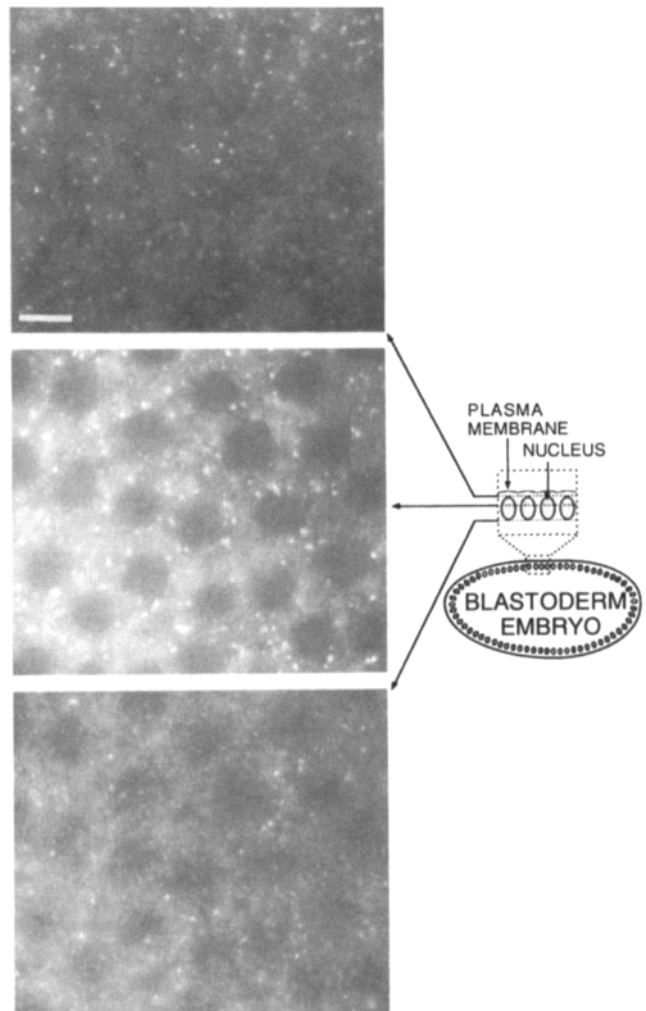
the nuclei. When a higher concentration of antibody is used, weak staining is also seen in the actin caps located apically to the nuclei (data not shown). The punctate staining is present throughout the syncytial stages, during cellularization, and persists in most cells of the later embryo. After germ band retraction is complete (~9.4 h of development), protein appears to be present at higher concentrations in a subset of tissues. One tissue in which staining is increased is the epidermal region which is undergoing dorsal closure (Fig. 10). Over the course of several hours, these dorso–lateral epidermal cells migrate dorsally to cover the amnioserosa, until they meet and complete dorsal closure. Several rows of cells at the leading edge of the epithelial sheet show more intense labeling with the 3C7 mAb.

Discussion

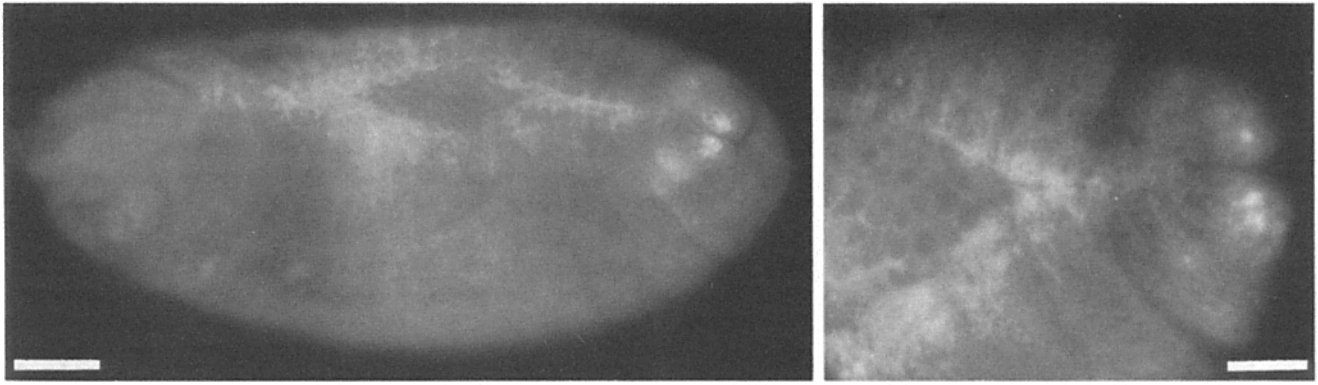We have identified a new member of the MHC gene family

Figure 10. Immunolocalization of 95F MHC in older embryos. Fluorescence micrographs showing increased staining intensity in the dorso-lateral epidermal cells that are undergoing dorsal closure and in cells forming the posterior spiracles of an embryo that has been stained with mAb 3C7 are presented. Embryos are oriented with anterior to the left. (*a*) This stage 15 embryo is laterally tilted off the dorsal midline. Three or four rows of cells along the leading edge of the dorso-lateral epidermis show relatively intense fluorescence; these cells have migrated dorsally and have met to form a seam at the anterior and posterior ends, but have not completed migration in the middle. Some staining is also seen in the cells of the amnioserosa. Regions of intense staining are seen in the cells forming the posterior spiracles at the posterior end. (*b*) This is a higher magnification view of a stage 14 embryo. Bars: (*a*) 50 μm; (*b*) 25 μm.

in *Drosophila*. This gene encodes an unconventional MHC that is not closely related to other *Drosophila* MHCs or unconventional MHCs in other organisms. The identification of this new divergent MHC illustrates the value of F-actin chromatography as an approach to finding new actin-binding proteins. Genomic Southerns and in situ hybridizations to polytene chromosomes show that this is a unique gene at polytene map position 95F, a position distinct from all other known *Drosophila* MHC genes. Comparison of the specificity of antibodies raised against the embryonic 140-kD actin-binding protein and against bacterially expressed protein from a 95F MHC cDNA demonstrates that the two proteins are likely to be products of the same gene. Cleveland peptide mapping of protein isolated from *Drosophila* and from bacteria expressing the cDNA-encoded protein confirms that we have indeed cloned the gene for the protein originally identified by the 3C7 mAb.

The 95F MHC probably defines a new class of unconventional myosins. In amino acid sequence comparisons of myosin heads, all are >30% identical to each other (except the myosin-related protein *ninaC*), and can be grouped into classes of >45% identity. In contrast, the 95F MHC head is only 30–33% identical to all other MHCs and thus is no more closely related to members of one class than to members of another. A dendrogram that depicts the sequence relationships of MHCs is shown in Fig. 11; the lengths of the horizontal lines are proportional to the degree of similarity between myosin heads. By this analysis, myosin genes are grouped into type I unconventional MHCs, type II MHCs, and *dilute*-type MHCs; *ninaC* is not grouped with any other member of the MHC family. The 95F MHC can not be grouped with any other MHC gene. In addition, comparisons of tail sequences show conservation of motifs among members of the same class, while 95F MHC shares no tail sequence motifs with any. Thus, by these two criteria the 95F MHC appears to represent a new class of unconventional myosin.

The developmental expression profile of the 95F MHC gene is complex. Overall, the abundance of both mRNA and protein products is regulated, with peak expression occur-

ring during mid-embryogenesis and adulthood. We have evidence for four different protein isoforms encoded by this gene, and they appear to vary in their expression patterns. The isoform encoded by the canonical cDNA is expressed both maternally and zygotically, while the Em-1 and Em-5 cDNA products are probably only expressed zygotically. The fourth isoform, a 145-kD protein detected only in larval brains and adult heads, is too large to be the product of any of the cDNAs we characterized. It appears later than any of the other isoforms, and seems to be restricted to neural tissues. We conclude that the 95F MHC gene encodes multiple protein isoforms that are regulated in abundance, time, and location of expression.

Several lines of evidence suggest that the different cDNAs from the 95F MHC gene are likely to be the result of alternate mRNA splicing. First, we identified three different embryonic cDNAs that encode different protein isoforms. They diverge at the same nucleotide and they all contain poly-adenylated tails. Second, the three nucleotides just before the point of divergence conform to the "CAG" splice donor consensus. Third, developmental Northern blots reveal multiple sizes of transcripts, which could be generated by alternate splicing. In addition, a protein band that corresponds in size to one that would be encoded by the Em-5 cDNA is detected on Western blots (Fig. 7). Thus, like muscle MHC (Bernstein et al., 1983; Rozek and Davidson, 1983), *ninaC* (Montell and Rubin, 1988), and nonmuscle MHC (Ketchum et al., 1990) in *Drosophila*, the 95F MHC gene probably undergoes alternate splicing, and like the first two, generates isoform diversity in this way.

Although the function of 95F MHC is unknown, several lines of reasoning suggest it might be involved in transport in the early embryo. Immunolocalization in early embryos shows 95F MHC in a punctate pattern, consistent with its association with some cytoplasmic structure (Fig. 9). In addition, labeling of live embryos with Rhodamine-mAb 3C7 or Rhodamine-Fab 3C7 shows a similar pattern of localization, indicating that the punctate labeling in fixed specimens is not artifactual. In vivo observation of embryos labeled in this manner has demonstrated that these structures are actively
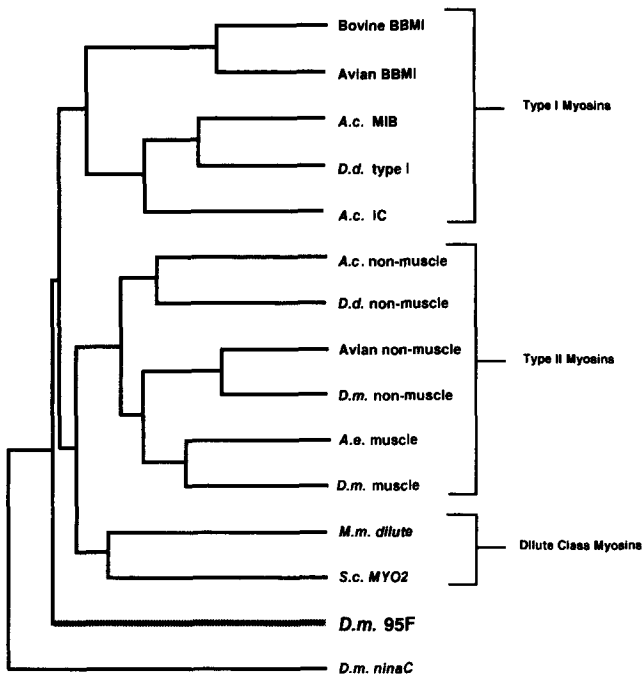
832

Figure 11. Dendrogram of myosin heavy chain head sequences. The relationships between amino acid sequences of head domains of the myosin heavy chain family members are depicted in a dendrogram. This representation was created using the PILEUP program in the GCG Software Package (Devereux et al., 1984). The distance along the horizontal axis is proportional to the amino acid sequence similarities between the MHC genes indicated, and does not imply any phylogenetic relationship. The dendrogram shows that the *Drosophila* 95F MHC head is not closely related to any other MHC, although it is not as dissimilar as the myosin-related gene *ninaC*. *Bovine BBMI*, bovine intestinal brush border myosin I (Hoshimaru and Nakanishi, 1987); *Avian BBMI*, avian brush border myosin I (Garcia et al., 1989); *A. c. MIB*, *Acanthamoeba castellanii* MIB MHC (Jung et al., 1989b); *D. d. type I*, *Dictyostelium discoideum* myosin I heavy chain (Jung et al., 1989a); *A. c. IC, Acanthamoeba castellanii* IC MHC (Jung et al., 1989b); *A. c. nonmuscle, Acanthamoeba castellanii* nonmuscle MHC (Hammer et al., 1986); *D. d. nonmuscle, Dictyostelium discoideum* nonmuscle MHC (DeLozzane et al., 1985); *Avian nonmuscle*, avian nonmuscle MHC (Shohet et al., 1989); *D. m. nonmuscle, Drosophila melanogaster* nonmuscle MHC (Ketchum et al., 1990); *A. e. muscle, A. irradians* (scallop) muscle MHC (Nyitray et al., 1990); *D. m. muscle, Drosophila melanogaster* muscle MHC (George et al., 1989); *M. m. dilute, Mus muscalus dilute* MHC (Mercer et al., 1991); *S. c. MYO2, Saccharomyces cerevisiae* MYO2 MHC (Johnson et al., 1991) and *D. m. ninaC, Drosophila melanogaster ninaC* myosin-related heavy chain (Montell and Rubin, 1988). All sequences were obtained from the Gen/EMBL database.

transported (Miller, K., V. Mermall, and J. McNally, manuscript in preparation). Other unconventional MHCs have tail sequences that are postulated to permit membrane binding and/or second-site actin binding that might mediate attachment to "cargoes." Although the 95F MHC does not appear to contain the tail domains that are implicated in the ligand binding of other unconventional MHCs, the apparent association with structures in the cytoplasm suggests that some tail sequence mediates this binding. Since we have not yet determined the identity of the labeled cytoplasmic component, we can only speculate about the binding mechanism. There are two regions which suggest possible molecular interactions. The predicted α-helical domain might interact with a complementary α-helix in a particle-associated protein to form an intermolecular coiled-coil. Alternately, the hydrophobic regions encoded by each of the cDNAs may associate directly with membrane-bound vesicles or with other proteins. Finally, the 95F MHC may associate with cytoplasmic components via an as yet unidentified binding sequence.

Punctate staining in the embryo cortex has been observed by others, using reagents specific for other components of the actin cytoskeleton. However, the punctate pattern we observe does not appear to be identical to the staining in these published reports (Warn et al., 1984; Karr and Alberts, 1986; Warn, 1986; Young et al., 1991). The punctate structures observed using anti-actin antibodies (Karr and Alberts, 1986) or phalloidin (Warn, 1986) appear to be present only within ~3 μm of the plasma membrane during the syncytial stages. The punctate staining observed with anti-nonmuscle myosin antibody is present in preblastoderm embryos, but disappears during the syncytial stages (Young et al., 1991). In contrast, we have observed punctate staining with anti-95F myosin antibodies (Fig. 9), throughout the cytoplasmic domains (including >6 μm below the surface) during the syncytial stages and persisting throughout embryonic development. Although from the published reports it is difficult to be certain that there is no overlap between these different antigens, it is likely that different cellular components are being labeled by the reagents specific for actin, nonmuscle myosin and 95F myosin. However, definitive evidence about the coincidence of these different antigens can only be obtained through double-label studies. We are currently performing such experiments, as part of a study of the distribution of 95F myosin isoforms and transcripts during embryonic development (Kellerman, K., and K. Miller, manuscript in preparation).

In addition to its possible role in cytoplasmic transport in early embryos, increased levels of immunostaining in specific subsets of tissues later in development suggest this protein might play additional specific roles in some tissues. Particularly interesting is the intense staining in the three of four rows of dorso–lateral epidermal cells which are at the leading edge of the epithelial sheet involved in dorsal closure (Fig. 10). Another myosin, nonmuscle type II MHC, is highly concentrated at the apical membrane in the dorsal-most row of epidermal cells undergoing dorsal closure (Young et al., 1992). These two MHCs may have cooperative functions in the dynamic process of dorsal closure.

We have evidence that the 95F unconventional myosin we have identified plays important roles in dynamic processes required in early embryonic development. Since the gene encodes several protein isoforms and one or more isoforms are expressed throughout the life cycle, we believe that the different molecules encoded by this gene are likely to be involved in a variety of dynamic processes at many developmental stages.

References

Bender, W., P. Spierer, and D. S. Hogness. 1983. Chromosomal walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the Bithorax Complex in *Drosophila melanogaster*. *J. Mol. Biol.* 168:17–33.

Bernstein, S. I., C. J. Hansen, K. D. Becker, D. R. Wassenberg II, E. S. Roche, J. J. Donady, and C. P. Emerson, Jr. 1986. Alternative RNA splicing generates transcripts encoding a thorax-specific isoform of *Drosophila melanogaster* myosin heavy chain. *Mol. Cell. Biol.* 6:2511–2519.

Brown, N. H., and F. C. Kafatos. 1988. Functional cDNA libraries from *Drosophila* embryos. *J. Mol. Biol.* 203:425–437.

Cavener, D. R. 1987. Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucl. Acids Res.* 15: 1353–1361.

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

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

Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli. 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.* 252:1102–1106.

Cohen, C., and D. A. D. Parry. 1990. Alpha-helical coiled-coils and bundles: how to design an alpha-helical protein. *Proteins Struct. Funct. Genet.* 7:1–15.

Cote, G. P., J. P. Albanesi, T. Ueno, J. A. Hammer III, and E. D. Korn. 1985. Purification from *Dictyostelium discoideum* of a low-molecular weight myosin that resembles myosin I from *Acanthamoeba castellanii*. *J. Biol. Chem.* 260:4543–4546.

DeLozanne, A., M. Lewis, J. A. Spudich, and L. A. Leinwand. 1985. Cloning and characterization of a nonmuscle myosin heavy chain cDNA. *Proc. Natl. Acad. Sci. USA.* 82:6807–6810.

Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387–395.

Drubin, D. G., J. Mulholland, Z. Zhu, and D. Botstein. 1990. Homology of a yeast actin-binding protein to signal transduction proteins and myosin-I. *Nature (Lond.).* 343:288–290.

Espreafico, E. M., R. E. Cheney, M. Matteoli, A. A. C. Nascimento, P. V. De Camilli, R. E. Larson, and M. S. Mooseker. 1992. Primary structure and cellular localization of chicken brain myosin-V (p190), an unconventional myosin with calmodulin light chains. *J. Cell Biol.* In press.

Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6–13 and 137:266–67.

Garcia, A., E. Coudrier, J. Carboni, J. Anderson, J. Vandekerkhove, M. Mooseker, D. Louvard, and M. Arpin. 1989. Partial deduced sequence of the 110-kD-calmodulin complex of the avian intestinal microvillus shows that this mechanoenzyme is a member of the myosin-I family. *J. Cell Biol.* 109:2895–2903.

George, E. L., M. B. Ober, and C. P. Emerson, Jr. 1989. Functional domains of the *Drosophila melanogaster* muscle myosin heavy-chain gene are encoded by alternatively spliced exons. *Mol. Cell. Biol.* 9:2957–2974.

Hammer, J. A. III, E. D. Korn, and B. M. Paterson. 1986. Isolation of a nonmuscle myosin heavy chain gene from Acanthamoeba. *J. Biol. Chem.* 261: 1949–1956.

Harlow, E., and D. P. Lane. 1988. Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory. Cold Spring Harbor, New York.

Hoshimaru, M., and S. Nakanishi. 1987. Identification of a new type of mammalian myosin heavy chain by molecular cloning: overlap of its mRNA with pre-protachykinin-B mRNA. *J. Biol. Chem.* 262:14625–14632.

Johnston, G. C., J. A. Prendergast, and R. A. Singer. 1991. The *Saccharomyces cerevisiae MYO2* gene encodes an essential myosin for vectorial transport of vesicles. *J. Cell Biol.* 113:539–551.

Jowett, T. 1986. Preparation of nucleic acids, Ch. 12. *In Drosophila*: A Practical Approach. D. B. Roberts, editor. IRL Press, Ltd., Oxford and Washington, D.C. 275–286.

Jung, G., C. L. Saxe III, A. R. Kimmel, and J. A. Hammer III. 1989a. *Dictyostelium discoideum* contains a gene encoding a myosin-I heavy chain. *Proc. Natl. Acad. Sci. USA.* 86:6186–6190.

Jung, G., C. J. Schmidt, and J. A. Hammer III. 1989b. Myosin-I heavy-chain genes of *Acanthamoeba castellanii*: cloning of a second gene and evidence for the existence of a third isoform. *Gene (Amst.).* 82:269–280.

Karr, T. L., and B. M. Alberts. 1986. Organization of the cytoskeleton in early *Drosophila* embryos. *J. Cell Biol.* 102:1494–1509.

Ketchum, A. S., C. T. Stewart, M. Stewart, and D. P. Kiehart. 1990. Complete sequence of the *Drosophila* nonmuscle myosin heavy-chain transcript: conserved sequences in the myosin tail and differential splicing in the 5′ untranslated sequence. *Proc. Natl. Acad. Sci. USA.* 87:6316–6320.

Kiehart, D. P., M. S. Lutz, D. Chan, A. S. Ketchum, R. A. Laymon, B. Nguyen, and L. S. B. Goldstein. 1989. Identification of the gene for fly nonmuscle myosin heavy chain: *Drosophila* myosin heavy chains are encoded by a gene family. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:913–922.

Kozak, M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucl. Acids Res.* 12:857–872.

Laughon, A., and M. P. Scott. 1984. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. *Nature (Lond.).* 310:25–31.

Lindsley, D. L., and E. H. Grell. 1968. Genetic variations of *Drosophila melanogaster*. Carnegie Inst. Washington Publ. No. 627.

Lupas, A., M. Van Dyke, and J. Stock. 1991. Predicting coiled coils from protein sequences. *Science (Wash. DC).* 252:1162–1164.

Mercer, J. A., P. K. Seperack, M.C. Strobel, N. G. Copeland, and N. A. Jenkins. 1991. Novel myosin heavy chain encoded by murine *dilute* coat color locus. *Nature (Lond.).* 349:709–713.

Merril, C. R., D. Goldman, S. A. Sedman, and M. H. Ebert. 1981. Ultrasensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. *Science (Wash. DC).* 211:1437–1438.

Miller, K. G., and B. M. Alberts. 1989. F-actin affinity chromatography: technique for isolating previously unidentified actin-binding proteins. *Proc. Natl. Acad. Sci. USA.* 86:4808–4812.

Miller, K. G., C. M. Field, and B. M. Alberts. 1989. Actin-binding proteins from *Drosophila* embryos: a complex network of interacting proteins detected by F-actin affinity chromatography. *J. Cell Biol.* 109:2963–2975.

Montell, C., and G. M. Rubin. 1988. The *Drosophila ninaC* locus encodes two photoreceptor cell specific proteins with domains homologous to protein kinases and the myosin heavy chain head. *Cell.* 52:757–772.

Moon, R. T., and A. P. McMahon. 1990. Generation of diversity in nonerythroid spectrins: multiple polypeptides are predicted by sequence analysis of cDNAs encompassing the coding region of human nonerythroid alpha-spectrin. *J. Biol. Chem.* 265:4427–4433.

Nyitray, L., E. B. Goodwin, and A. G. Szent-Gyorgyi. 1990. Nucleotide sequence of a full length cDNA for a scallop striated muscle myosin heavy chain. *Nucleic Acids Res.* 18:7158.

Pardue, M. L. 1986. In Situ Hybridization to DNA of Chromosomes and Nuclei. *In Drosophila*: A Practical Approach. D. B. Roberts, editor. IRL Press, Ltd., Oxford and Washington, D.C. Chapter 5, 111–137.

Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA.* 85:2444–2448.

Pollard, T. D., S. K. Doberstein, and H. G. Zot. 1991. Myosin-I. *Annu. Rev. Physiol.* 53:653–681.

Pollard, T. D., and E. D. Korn. 1973. *Acanthamoeba* Myosin. I. Isolation from *Acanthamoeba castellanii* of an enzyme similar to muscle myosin. *J. Biol. Chem.* 248:4682–4690.

Poole, S. J., L. M. Kauvar, B. Drees, and T. Kornberg. 1985. The *engrailed* locus of *Drosophila*: structural analysis of an embryonic transcript. *Cell.* 40:37–43.

Rozek, C. E., and N. Davidson. 1983. *Drosophila* has one myosin heavy-chain gene with three developmentally regulated transcripts. *Cell.* 32:23–34.

Rozek, C. E., and N. Davidson. 1986. Differential processing of RNA transcribed from the single-copy *Drosophila* myosin heavy chain gene produces four mRNAs that encode two polypeptides. *Proc. Natl. Acad. Sci. USA.* 83: 2128–2132.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

Shohet, R. V., M. A. Conti, S. Kawamoto, Y. A. Preston, D. A. Brill, and R. S. Adelstein. 1989. Cloning of the cDNA encoding the myosin heavy chain of a vertebrate cellular myosin. *Proc. Natl. Acad. Sci. USA.* 86:7726–7730.

Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene (Amst.).* 67:31–40.

Snyder, M., S. Elledge, D. Sweetser, R. A. Young, and R. W. Davis. 1987. Lambda gt11: gene isolation with antibody probes and other applications. *Methods Enzymol.* 154:107–128.

Spierer, P., A. Spierer, W. Bender, and D. S. Hogness. 1983. Molecular mapping of genetic and chromomeric units in *Drosophila melanogaster*. *J. Mol. Biol.* 168:35–50.

Steinhaur, W. R., R. C. Walsh, and L. J. Kalfayan. 1989. Sequence and structure of the *Drosophila melanogaster ovarian tumor* gene and generation of an antibody specific for the *ovarian tumor* protein. *Mol. Cell. Biol.* 9:5726–5732.

Warn, R. M. 1986. The cytoskeleton of the early *Drosophila* embryo. *J. Cell Sci. Suppl.* 5:311–328.

Warn, R. M., R. Magrath, and S. Webb. 1984. Distribution of F-actin during cleavage of the *Drosophila* syncytial blastoderm. *J. Cell Biol.* 98:156–162.

Warrick, H. M., and J. A. Spudich. 1987. Myosin structure and function in cell motility. *Annu. Rev. Cell Biol.* 3:379–421.

Young, P. E., T. C. Pesacreta, and D. P. Kiehart. 1991. Dynamic changes in the distribution of cytoplasmic myosin during *Drosophila* embryogenesis. *Development (Camb.).* 111:1–14.

Young, P. E., A. M. Richman, A. S. Ketcham, and D. P. Kiehart. 1992. Morphogenesis in *Drosophila* requires nonmuscle myosin heavy chain function. *Genes & Dev.* In press.