Invertebrate Muscles: Muscle Specific Genes and Proteins

SCOTT L. HOOPER AND JEFFREY B. THUMA

Neuroscience Program, Department of Biological Sciences, Ohio University, Athens, Ohio

I. Introduction

A. Why study invertebrate muscles, genes, and proteins?

B. Invertebrate and vertebrate phylogeny

C. Scope of review and literature database

II. Review of Vertebrate Muscle Specific Proteins

III. Invertebrate Muscle Proteins

A. Thin filament proteins

B. Thick filament proteins

C. α-Actinin and other Z line proteins

D. Ca²⁺ binding proteins and their targets

E. Giant sarcomere-associated proteins

F. Miscellaneous other proteins

G. Summary

Hooper, Scott L., and Jeffrey B. Thuma. Invertebrate Muscles: Muscle Specific Genes and Proteins. Physiol Rev 85: 1001–1060, 2005; doi:10.1152/physrev.00019.2004.—This is the first of a projected series of canonic reviews covering all invertebrate muscle literature prior to 2005 and covers muscle genes and proteins except those involved in excitation-contraction coupling (e.g., the ryanodine receptor) and those forming ligand- and voltage-dependent channels. Two themes are of primary importance. The first is the evolutionary antiquity of muscle proteins. Actin, myosin, and tropomyosin (at least, the presence of other muscle proteins in these organisms has not been examined) exist in muscle-like cells in Radiata, and almost all muscle proteins are present across Bilateria, implying that the first Bilateralian had a complete, or near-complete, complement of present-day muscle proteins. The second is the extraordinary diversity of protein isoforms and genetic mechanisms for producing them. This rich diversity suggests that studying invertebrate muscle proteins and genes can be usefully applied to resolve phylogenetic relationships and to understand protein assembly coevolution. Fully achieving these goals, however, will require examination of a much broader range of species than has been heretofore performed.

I. INTRODUCTION

A. Why Study Invertebrate Muscles, Genes, and Proteins?

Although pure research has its defenders, science is generally justified by perceived human benefit. Two arguments suggest that studying invertebrate muscle genes and proteins can reveal generally applicable principles that could benefit humans. First, the last common ancestor of vertebrates and invertebrates had muscle, and most vertebrate muscle genes and proteins have invertebrate homologs. The experimental advantages of invertebrate preparations often allow these genes and proteins to be investigated more easily, or at a greater level of detail, than is possible in vertebrates. Many human diseases result from errors in muscle protein structure, and thus invertebrate studies have the possibility of improving human health. Second, invertebrate muscle genes and proteins show great variation. Despite this variety, in all cases the proteins must functionally interact correctly. Comparative studies therefore provide a rich arena in which to investigate the relationship between protein assembly structure and activity, again an area with clear relevance to human well-being.

B. Invertebrate and Vertebrate Phylogeny

Figures 1–3 show a contemporary, molecular biology-based, tree of life (4 and sources listed in the legend to Fig. 1). Three things are of particular importance. First, Cnidaria (corals, jellyfish) are separate and equal to Bilateria. All Bilatera are thus equally distant from all Cnidaria. Second (as has been long known), echinoderms, tunicates, and amphioxus are more closely related to
vertebrates than they are to other invertebrates. Third, Ecdysozoa, Lophotrochozoa, and Deuterostomia, each of which contains invertebrates, are presently coequal branches. This tripartite split will presumably be eventually resolved into two bipartite branchings, but whether the Ecdysozoa and Lophotrochozoa, or one of them and the Deuterostomia, will end up being most closely related is as yet unclear. This issue has profound implications for comparative research since, depending on the ultimate resolution of bilaterian relationships, it may be that lobsters are more closely related to humans than they are to leeches. Although research on muscle genes may help resolve this issue, it is so far insufficient to do so. We have therefore organized the data presented here in simple concordance with the relationships shown in Figures 1–3.

C. Scope of Review and Literature Database

A comprehensive review of invertebrate muscle is unavailable. However, our invertebrate muscle database contains over 6,700 references, and this massive literature cannot be covered in a single review. We therefore intend to produce a series of canonic reviews covering all journal articles (due to their limited availability, books and book chapters are not included) on invertebrate muscle written before 2005. To that end, the field has been divided into subsets, of which this review covers the first. Subsequent reviews will cover 1) thick filament, thin filament, and sarcomere structure; the molecular basis of contraction and its regulation; asynchronous muscle and catch; 2) muscle and synaptic ultrastructure and excitation/contraction coupling; 3) voltage- and ligand-gated ionotropic channels; 4) metabotropic channels (modulation); and 5) integrative properties and production of behavior. Even with this broad net, some boundaries had to be drawn. In particular, papers dealing with metabolic pathways are generally not included, and no attempt to cover molting and muscle proteases, regeneration, or muscle development has been made (papers that identify regulatory regions in muscle protein genes are included, but papers...
further “upstream” are not). The database is available at http:\crab-lab.zool.ohiou.edu\invert. Every effort will be made to maintain this site for at least 10 years from publication date.

II. REVIEW OF VERTEBRATE MUSCLE SPECIFIC PROTEINS

Due to the number of references in this review, this section is only sparsely referenced; References 6 and 128 are excellent reviews of vertebrate muscle. All muscles contain thick and thin filaments. Thick filaments are composed of myosin. Myosin is composed of three pairs of proteins: the heavy chain and the essential and regulatory light chains (Fig. 4). The tails of the heavy chains form an α-helical coiled-coil tail. The other end of each heavy chain and one essential and one regulatory chain form one of the combined molecule’s two heads, each of which contains an ATPase activity and can independently bind to the thin (actin) filament. Trypsin severs the myosin tail, resulting in light meromyosin, which contains only tail (rod) sequences and heavy meromyosin (HMM), which contains part of the tail and the two head regions (30, 42, 525, 1135, 1138). Further digestion of HMM results in the S1 and S2 fragments, S1 consisting of the heads and S2 of the HMM tail portion. Thin filaments are a double helix of polymerized actin monomers. Tropomyosin and troponin are two thin filament-associated proteins involved in contraction regulation in striated muscle (muscles with well-organized sarcomeres, Fig. 5). The other type of vertebrate muscle, smooth muscle, does not have well-organized sarcomeres. Vertebrate smooth muscle contraction is regulated both by myosin light-chain phosphorylation by myosin light-chain kinase and a thin filament-based regulatory system based on the actin-binding proteins calponin and caldesmon.

All vertebrate striated muscle sarcomeres are very similar (Fig. 5). Two Z lines, composed largely of α-actinin, define the sarcomere edges. The thin filaments attach to each Z line and extend toward the center of the sarcomere. The region adjacent to each Z line containing only thin filaments is the I band. One-half of each I band

FIG. 2. Phylogeny of Ecdysozoa (except Insecta, see Fig. 3).
therefore belongs to one sarcomere and the other half to the adjacent sarcomere. The thick filaments are located at the center of the sarcomere. The region of the sarcomere with only thick filaments is the H band, and the region defined by their extent is the A band. At the very center of the sarcomere there is often also a line (due to the presence of additional proteins) called the M line.

Sarcomeres consisting of only thick and thin filaments and Z lines would be inherently unstable. To appreciate this, consider a muscle fiber stimulated to contract while maintained in an isometric (constant length) condition. If the thick filaments were exactly centered in the sarcomere, equal numbers of myosin heads would engage the thin filaments on the two sides of the M line, the thick filament would feel equal force in both the right and left directions, and the thick filaments would therefore remain centered in the sarcomere. However, if a thick filament was even slightly uncentered, it would experience greater force in one direction and would therefore slide in that direction. The force imbalance on the thick filament would now be even greater, and it would thus continue to slide until it reached the Z line. Solving this difficulty requires a mechanism that develops...
a centering force if the thick filaments became uncentered. For example, if the thick filaments were attached to the Z lines by springs, any thick filament movement away from the center would decrease force in the shortened spring, and increase force in the stretched spring, which would recenter the thick filament.

Electron microscopic evidence of filaments linking the thick filaments to the Z line (and which could thus function as springs) was early obtained in both vertebrates and invertebrates (781). However, these filaments are not composed of polymerized smaller subunits but are instead enormous single proteins, and it took almost 20 years to characterize them chemically. Intriguingly, many of them contain large numbers of immunoglobulin and fibronectin III repeats. The largest of these proteins is titin (3 MDa, ~30,000 amino acids, Fig. 5). Titin can be 1 μm in length, and single titin molecules connect the M and Z lines. Titin's NH₂ terminus extends through the Z line in close association with the thin filament. At the A/I junction it leaves the thin filament and joins the thick filament, with which it runs until reaching the M line, where it overlaps the titin filaments from the other half-sarcomere. The M line-associated proteins M-protein, myomesin, and skelemin (which is generated from the myomesin gene by alternative splicing) are other members of this family, as are the A band myosin binding proteins C and H (also called C-protein and 86-kDa protein) (Fig. 5). These proteins are much smaller than titin (a few hundred kDa), and unlike titin lack a serine/threonine kinase activity. Myosin light-chain kinase and telokin are two other members of this protein family, found in smooth muscle.

For a recent general review of sarcomere structure and proteins (primarily vertebrate, but includes some invertebrate work), see Reference 182; Reference 933 reviews all aspects of both vertebrate and invertebrate muscle. Reference 1136 reviews the early history of muscle protein isolation and analysis. Reference 188 is a dated, primarily vertebrate, review of calcium binding proteins, troponin C, and myosin light chains. References 112, 276, 330–332, 451, 1223; 17, 276, 1261; and 437 review, respectively, Drosophila, Caenorhabditis elegans, and amphioxus muscle. References 236, 237, 440–443, 571, 815, 869 and 43, 378, 673, 987, 1266, 1282, respectively, describe phenotypic mutant derivations in Drosophila and C. elegans. Reference 333 reviews basic methods in Drosophila muscle biology. Gene expression profiles that included muscle specific proteins have been performed in jellyfish (Cynnea capillata) tentacle (1307); oyster (Crassostrea gigas) mantle (799); Mytilus galloprovincialis (1210); the platyhelminths Clonorchis sinensis (652) and Schistosoma japonicum (291, 1214, 1215); the nematodes Brugia malayi (118, 119), C. elegans (719, 769, 774, 798, 1259), Globodera species (951), Haemonchus contortus (436), Meloidogyne incognita (773), Onchocerca volvulus (697), and Strongyloides stercoralis (798) [for nematodes, multiple expressed sequence tag databases are now available on-line (926, 927, 1293); the mites Psoroptes ovis (555) and Boophilus microplus (220); and amphioxus notochord (1124) and Ciona intestinalis embryo (1048, 1049), larva (632, 1048), and adult (173, 1048), and cDNA clones covering nearly 85% of C. intestinalis mRNA species are available (1050). References 1060 and 281, 814, 1159 show two-dimensional electrophoresis profiles of, respectively, C. elegans and Drosophila proteins. References 814 and 1159 show that anatomically different Drosophila muscles (fibrillar vs. tubular) have different protein compositions. Also uncategorized are papers describing muscle post mortem changes, food-related properties, and calorimetric measurements of muscle proteins (this list is not comprehensive) (18, 19, 65, 170, 267, 269, 308, 367–372, 462, 463, 467, 493, 518–520, 542, 543, 546, 565, 588, 589, 620, 687, 701, 753, 764, 765, 770, 807, 808, 839,
Actin and myosin heavy chain have been extensively studied, and these sections are therefore subordered by phylogenetic group. For some groups not all the articles associated with it are explicitly covered in the text (e.g., descriptions of isolation techniques). For completeness, in these cases all references dealing with that group are listed immediately after the relevant subtitle. References about purification of myosin as an oligomer (i.e., heavy and light chains together) are listed in this manner in the myosin heavy chain section. Table 1 provides actin and myosin heavy chain data for groups for which only limited information is available (for actin, Pterobranchia, Annelida, Gastropoda, Brachiopoda, Chaetognatha, Chelicerata; for myosin heavy chain, Cephalochordata, Urochordata, Annelida, Chaetognatha, Chelicerata).

### A. Thin Filament Proteins

#### 1. Actin

Mammals have six (two striated muscle, two smooth muscle, and two cytoplasmic) (1203) and teleost fish have nine (1211) actin isoforms. References 303 and 418 are general (vertebrate and invertebrate) reviews of actin molecular genetics, References 557 and 1025 review vertebrate and invertebrate actin isoforms, and Reference 499 reviews ascidian actin. Reference 238 shows that muscle actins isolated from a variety of invertebrates all have the same molecular weight and coelectrofocus with the β-form of vertebrate smooth muscle actin, but are immunologically distinct from each another.

**A) Cnidaria.** Two coral actin cDNA clones have been identified, one of which is expressed only in adults, the only stage with muscles. Although not verified by in situ hybridization, cnidaria may thus have a muscle specific actin. The putative muscle actin gene showed greatest homology to metazoan cytoplasmic actins (which metazoans are not clear from the article) (321). However, invertebrate muscle actins are typically most similar to vertebrate cytoplasmic actins (see below), and thus, depending on which metazoans the authors used for comparison, this similarity is not strong evidence that the adult actin gene is not a muscle gene.

**B) Cephalochordata.** Cephalochordate (Branchiostoma only) actin gene number and expression are confusing. Work using antibodies specific for vertebrate smooth muscle, striated muscle, and cytoplasmic actins shows that *B. lanceolatum* has at least three actin isoforms, with the antismooth and antistriated antibodies staining separate sets of muscles and the anticytoplasmic antibody staining most other cells (note that Ref. 1204 is wrong in stating cephalochordates express only one muscle actin isoform). Despite early confusion on this point (373, 946, 1234), in these animals the notochord is an innervated

<table>
<thead>
<tr>
<th>Group/Protein</th>
<th>Species</th>
<th>Identification Method</th>
<th>Muscle Versus Cytoplasmic Isoforms?</th>
<th>Muscle or Developmental Stage Specific Expression?</th>
<th>Reference Nos.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pterobranchia actin</td>
<td><em>Saccoglossus kowalevskii</em></td>
<td>Two cDNAs</td>
<td>Unknown</td>
<td>Unknown</td>
<td>127</td>
</tr>
<tr>
<td>Annelida actin</td>
<td><em>Glycera</em>, <em>Nereis</em>, species unreported</td>
<td>Protein purification</td>
<td>Unknown</td>
<td>Unknown</td>
<td>658</td>
</tr>
<tr>
<td></td>
<td><em>Helobdella triserialis</em></td>
<td>One cDNA, apparent family</td>
<td>Many tissues including muscle, but not segmental ganglia</td>
<td>Unknown</td>
<td>1269</td>
</tr>
<tr>
<td>Gastropod actin</td>
<td><em>Aphysia californica</em></td>
<td>Protein purification</td>
<td>Unknown</td>
<td>Unknown</td>
<td>401</td>
</tr>
<tr>
<td></td>
<td><em>Aphysia californica</em></td>
<td>One cDNA, 3-5 gene family</td>
<td>Unknown</td>
<td>Unknown</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td><em>Zebra mussel</em></td>
<td>One cDNA</td>
<td>Cloning of 5’-flanking and part of coding region of one gene</td>
<td>Unknown</td>
<td>643</td>
</tr>
<tr>
<td></td>
<td><em>Haliotis rufescens</em></td>
<td><em>Glottidea</em>, unreported species</td>
<td>Protein purification</td>
<td>Unknown</td>
<td>658</td>
</tr>
<tr>
<td></td>
<td><em>Paraspadella gotoi</em></td>
<td>Immunohistochemistry</td>
<td>Unknown</td>
<td>Unknown</td>
<td>1011</td>
</tr>
<tr>
<td></td>
<td><em>Paraspadella gotoi</em></td>
<td>Three genes</td>
<td>Differentially expressed in different muscles</td>
<td>Unknown</td>
<td>1308</td>
</tr>
<tr>
<td>Chelicerata actin</td>
<td>Various</td>
<td>Protein purification or immunohistochemistry</td>
<td>Unknown</td>
<td>Unknown</td>
<td>247, 658, 1021, 1204</td>
</tr>
<tr>
<td>Cephalochordata myosin</td>
<td><em>Amphioxus</em> (notochord)</td>
<td>Expressed sequence tag</td>
<td>Unknown</td>
<td>Unknown</td>
<td>1124</td>
</tr>
</tbody>
</table>
muscle (310, 311, 511, 831, 1124, 1273). Interestingly, this tissue nonetheless stains with the anticytoplasmic antibody (285, 1096).

Cloning in *B. lanceolatum* identified a cytoplasmic and a muscle actin on the basis of “diagnostic” amino acids (see below) (127). However, no tests for tissue specific expression to verify these assignments were made. Cloning in *B. floridiae* identified one cytoplasmic and two muscle actins on the basis of diagnostic amino acids (623). In situ hybridization showed that staining for one muscle actin was less strong in gill slit muscle and that the cytoplasmic form was weakly expressed in axial muscles in embryos, which suggests muscle and stage specific variation in actin gene expression. Cytoplasmic actin was present in notochord during early development, but none of the three forms was present in adult notochord. Cloning in *B. belcheri* identified only one cytoplasmic and one muscle actin gene (on the basis of diagnostic amino acids alone), but Southern blot analysis suggested the presence of “a number” of additional actin genes (624).

Expressed sequence tag analysis of a *B. belcheri* notochord cDNA library suggests that notochord has three actin genes, each of which codes for an identical actin that differs from the cytoplasmic and muscle clones already identified in this species (1124). The notochord, muscle, and cytoplasmic actin amino acid sequences are identical in the regions used to construct the oligonucleotide primers used to clone the muscle and cytoplasmic actins. It is thus unclear why the notochord genes were not identified in the earlier work (one possibility being a difference in codon usage, although codon variation is small for all amino acids in question, two of the seven amino acids in each primer have unique codons, and for the other five, only third nucleotide use varies). PCR analysis of notochord, muscle, and ovary libraries shows that one notochord actin gene is expressed only in notochord but some of the others are also expressed in muscle (1124). The most conservative interpretation of these data is that in Cephalochordata there are three actin groups (cytoplasmic, muscle, and notochord), each of which may contain more than one actin gene, but the number of these genes, and which tissues each is expressed in, is not clear.

**C. UROCHORDATA** Urochordate (87, 173, 176, 428, 429, 500, 605, 623, 625–632, 858, 875, 1047, 1049, 1096, 1126, 1181–1183, 1204, 1280, 1281) actin genes have been studied almost exclusively in Phlebobranchia (*Ciona intestinalis, Ascidia ceradotes*) and Stolidobranchia (*Styela clava, S. plicata, Halocynthia roretzi, Molgula oculata, M. occulta*). *Ciona* has eight nonmuscle and six muscle actin genes (classified by diagnostic amino acid position, not verified by in situ hybridization) (note Ref. 1204 is wrong in stating urochordates express only one muscle actin). Three muscle actin genes encode identical proteins (173). All six muscle actins have amino acid positions diagnostic of vertebrate striated muscle actin. *Ciona* thus appears to have no counterpart to vertebrate smooth muscle actin, a conclusion supported by the failure of antisMOOTH muscle antibodies to stain *Ascidia* muscles (1096). Gene expression profiles show that muscle actin is expressed in *Ciona* embryos (1049), larvae (632), and adults (173). Five muscle actins are expressed only in embryos and larvae, while the sixth is expressed solely in adults (173). In neither embryos nor larvae was an actin gene showing notochord specific expression mentioned (632, 1049).

Two-dimensional gel electrophoresis indicates that *S. clava* embryos, tadpoles, and adults contain three major and two minor actin isoforms. Two of the major isoforms are likely cytoplasmic and the third a muscle actin (1181). Cloning indicates four to seven muscle actin genes (according to diagnostic amino acid position). The four well-described genes encode identical proteins, but show different temporal and spatial expression patterns. In particular, one is expressed in a wide variety of tissues (including nonmuscle), but only in larva and younger animals, whereas another is expressed primarily in muscle cells in embryos but in nonmuscle tissue in adults. None of the genes has been shown to be expressed at high levels in adult muscle, which may suggest an adult-specific actin gene remains to be found. None of the clones stained the notochord (87). A muscle actin gene (by diagnostic amino acid position) that shows some amino acid variation (5.6%) from the *S. clava* muscle actins (and which might thus be the “missing” adult muscle actin gene) has been isolated from an adult *S. plicata* muscle cDNA library (605). In embryo and larva, this gene is expressed only in muscle cell lineages or functional muscle cells. Although the clone’s origin shows the gene is expressed in adult muscle, whether it is expressed only in muscle is unknown (1182).

*H. roretzi* adult body wall muscle contains two actins differing in isoelectric point (875). The gene(s) coding for these actins are not identified, but differ from those coding for larval muscle actin (625, 626, 628–630, 1047). *H. roretzi* has seven larval muscle (verified by in situ hybridization (625, 630)) actin genes. The genes are arranged in two clusters, one of which has five tandemly repeated genes in the same orientation and the other of which has two genes, arranged head to head on opposite DNA strands, that share a common interposed promoter (626, 628, 629). Two of the actins are identical, and all are very similar. Genes in each cluster have similar regulatory elements and are believed to be coordinately controlled (626, 628, 629, 1047). The sequences responsible for this regulation are beginning to be identified (428). When reporter genes with a *H. roretzi* 5’-upstream muscle actin gene (from the 5 gene cluster) flanking region are introduced into *C. savignyi*, reporter gene function is ob-
erved only in larval muscle cells, suggesting that larval muscle actin regulatory processes have been conserved in the two groups (429).

Two works in Molgula species show that changing actin gene expression can affect morphological development. The first (627, 631) involves M. oculata, which has typical tailed larvae, and M. occulta, which has tailless larvae. M. oculata has two actin genes (classified by diagnostic amino acids, but verified in the larva by in situ hybridization), one expressed in larva and the other in adults. M. occulta has two actin genes orthologous to the M. oculata larva gene, but in situ hybridization with M. oculata-derived probes do not detect any muscle actin production in M. occulta. However, when M. occulta gene 5'-flanking regions are attached to a reporter gene, reporter gene product is observed in M. occulta embryo vestigial muscle cells. Thus both actin gene promoter functionality and proper spatial production of the trans-acting factors that activate the genes are present in M. occulta. The coding regions of the M. occulta genes, however, contain insertions, deletions, and codon substitutions that would result in their producing nonfunctional actin. M. occulta taillessness thus appears to be due, at least in part, not to changes muscle actin gene activation, but to changes in gene coding regions such that the activated genes produce no functional actin.

The second work involves precocious development in M. citrina larvae (1126). In most Molgula species, mesenchyme cells, believed to be adult muscle progenitors, remain undifferentiated in larvae, but in M. citrina they begin to differentiate during the larval stage. A M. citrina actin gene has been identified that is expressed in juveniles and adults (it is not known if it is exclusively expressed in muscle at these stages), but not in larval tail muscle, suggesting that it codes for an adult muscle actin. In situ hybridization shows that the gene is expressed in the precociously differentiating mesenchyme cells in larva and (at least) early after larval metamorphosis. Precocious development of adult features in this species is thus likely also associated with precocious expression of an adult muscle actin gene.

Larvacea muscle (classification confirmed by in situ hybridization) actin genes have been investigated only in Oikopleura longicauda (858). This work identified one actin gene expressed in larval and adult tail muscle (larvaceans retain their tails as adults), but not in adult heart. Undiscovered muscle actin genes thus presumably exist in this species.

d) Echinodermata. Actin genes have been studied in echinoids (Stronglylocentrotus purpuratus, S. franciscanus, Lytechinus pictus, Heliocidaris erythrogramma, H. tuberculata) and asteroids (Pisaster ochraceus, Dermasterias imbricata) (199, 215, 218, 258, 265, 292, 293, 342, 508, 513, 568, 569, 602–604, 606, 650, 651, 909, 948, 1055, 1062, 1063, 1085, 1204, 1280, 1281). Echinoids have 6–10 cytoplasmic actins, but all possess only one muscle actin gene [early studies suggesting as many as 20 actin genes in Stronglylocentrotus and Lytechinus species apparently being mistaken (265, 508)] (215, 218, 265, 292, 293, 342, 568, 569, 650, 651, 909, 1055, 1085). In S. purpuratus, the muscle actin gene is expressed at high levels only in postpluteus muscle (1085). Dermasterias has eight actin genes, but which are cytoplasmic and which muscle is unknown (603). Pisaster has five actin genes, one believed to be cytoplasmic, two muscle specific, and two unspecified (identifications on the basis of cDNA library source tissue, not verified by in situ hybridization) (602–604, 606).

e) Mollusca. For bivalves, see References 153, 230, 296, 556, 558, 658, 738–740, 859, 930, 1000, 1125, 1137, 1204, 1210, 1281, 1309; for gastropods, see Refs. 246, 366, 401, 643, 826, 1280, 1281; for cephalopods, see Refs. 161, 448, 595, 1106, 1204. Although one of the first techniques to isolate large quantities of invertebrate thin filaments was developed in bivalves (1137), and regulation of bivalve actomyosin has been extensively studied (see second review), relatively little is known about bivalve actin genes. Early electrophoresis work showed no difference in actin across a wide range of bivalve species (740) or between different tissues within one examined species (Spisula) (739). Actin cDNA clones have been isolated from scallop (Placopecten magellanicus) (930) and oyster (Crassostrea gigas) (153). The scallop actin gene is likely a muscle actin and appears to be the primary actin expressed in adductor muscle. Southern blot analysis suggests 12–15 actin genes in the genome. Scallop actin polymerizes more slowly than rabbit actin, and once polymerized, the cleft between actin subdomains 2 and 4 is larger in scallop than in rabbit actin (556, 558). The oyster cDNA was used to locate the gene and its upstream region, but nothing is known about temporal or tissue expression. Actin gene polymorphisms can be used to identify clam species muscle (to prevent fraudulent use of less desirable species in consumer products) (296) and to study interspecies hybridization (230). Bivalve actin levels change in characteristic ways in response to various pollutants (1000).

The information available about gastropod muscle actin is presented in Table 1. Cephalopods are only slightly better investigated. Southern blot analysis suggests coleoid (all cephalopods but Nautilus) have at least three actin loci. Phylogenetic analysis suggests three actin isoforms, two of which were identified as muscle or cytoplasmic on the basis of being, respectively, “related to the mollusk muscle type” and “clustered among the other mollusk cytoplasmic” actins (161). However, almost all these sequences were obtained from GenBank, not from published work showing tissue specificity. As such, although the sequence comparisons showing three actin isoforms are likely valid, the identification of the isoforms...
as to type needs experimental confirmation. A capillary sodium dodecyl sulfate gel electrophoresis technique for actin separation in squid has been developed (1106).

F) PLATYHELMINTHS. Cestoda is represented by Taenia solium (15, 155), Diphyllobothrium dendriticum (1237–1241), and Echinococcus granulosus (229). Taenia has seven actin isoforms (15). Two genes, with identical coding sequences, have been cloned (155). Diphyllobothrium has five actin genes (1239, 1241), three muscle specific (1237). Two actin genes have been isolated from Echinococcus, and Southern blotting indicates as many as eight (229). Trematoda is represented by Schistosoma mansoni (1, 234, 706, 901, 1342) and Fasciola hepatica (1119). Two-dimensional gel electrophoresis identifies seven Schistosoma actin isoforms, and two actin genes have been cloned. Fasciola has three actin isoforms, one of which is specific to tegumental spines. Turbellaria is represented by Dugesia lugubris, which has at least two major actin isoforms, one muscle specific (928, 929).

G) NEMATODA. All information comes from Caenorhabditis elegans and Onchocerca volvulus (300, 608, 609, 646, 718, 902, 1052, 1204, 1235, 1264, 1280, 1281). In C. elegans four actin genes have been identified, two of which produce identical proteins and none which differs by more than three amino acids (300, 609). Three of the genes are clustered (10, 300). All four genes are transcribed, and three acquire a 22-nucleotide leader sequence via RNA trans-splicing (608). Genetic evidence suggests that two of the genes are involved in muscle thin filaments (646, 1264). However, the actins are so similar that they migrate as a single species by isoelectric focusing (1052), no isoform-specific antibodies have been generated, and thus other evidence of tissue specific isoform expression is lacking. One reference (902) states that all four genes code for muscle specific actins (and refers to a personal communication about a putative cytoplasmic actin gene), but on what basis is unclear. Expression of total actin varies during development (718). C. elegans actin can be efficiently extracted at high purity (902). Some biochemical differences between C. elegans and rabbit actin have been identified (902), and the structure of C. elegans actin resolved at 1.75 Å resolution (1235). There is suggestive evidence that C. elegans actin folding may be chaperoned (667). O. volvulus has four actin genes that can be divided into two classes on the basis of EcoR1 digestions. The genes are arranged in two clusters, each of which contains one copy of each class (1328). Nothing is known about tissue or stage specific expression.

H) CRUSTACEA. Crustacean (202, 526, 720, 783, 826, 1206, 1280, 1281) actin genes have been relatively little studied. Artemia (species unreported) has 8–10 genes and 4 isoforms, one muscle specific (709, 905). cDNA clones have been isolated from the shrimp Marsupenaeus japonicus and crayfish Procambarus clarkii, but nothing is known about their expression (526, 720). Crab (Gecarcinus lateralis) has seven or eight actin genes, and immunocytochemistry suggests some tissue and stage specific expression (1206). In lobster (Homarus americanus), different muscles contain different amounts of total actin (783).

I) INSECTA. The references for insecta are as follows: Diptera: Drosophila (16, 20, 64, 77, 89, 92, 111, 115, 129, 130, 204, 222, 259–262, 334, 336, 339, 340, 350, 419, 430, 431, 449, 512, 536, 540, 634, 647, 658, 702, 703, 722, 733, 826, 829, 868, 897, 925, 974, 1038, 1039, 1041, 1057, 1108, 1109, 1175, 1176, 1280, 1281, 1347), Aedes aegypti (466), Mayetiola destructor (1086), Dacus dorsalis (also known as Bactrocera dorsalis) (411), Phormia regina (591); Lepidoptera: Bombyx (826–829); Coleoptera: Heliocorapis jasperus (139, 146, 658, 1037). D. melanogaster has six actin genes that are widely dispersed throughout the genome and produce three major mRNA size classes (339). The genes are similarly dispersed in other Drosophila species (702). Gene coding sequences, but not intron positions, are highly conserved (334). Four (cytogenetic positions 57B, 79B, 87E, and 88F) of these genes code muscle actin (64, 111, 340, 430, 733, 826, 1041, 1175), at least some of whose expression is muscle or developmental stage specific (64, 77, 111, 204, 340, 430, 536, 634, 868, 1041, 1175, 1176). In particular, Act88F is expressed primarily in indirect flight muscles (64, 340, 430, 722) (although it is coexpressed with other muscle actin genes, and its absence causes behavioral defects, in a small number of other muscles (868)), Act79B is primarily expressed in “tubular” muscles (an anatomically specific muscle type, see third review) (64, 204, 340, 883) [an early report (1347) that Act79B is the larval muscle actin apparently being in error], and Act57B and Act87E are expressed in embryonic and larval muscle (340, 1175) and a variety of adult nontubular, nonindirect flight muscles (340). Similar data are obtained from D. virilis (703) and, for Act88F, in D. simulans (92), except that in D. virilis gene coexpression occurs in more muscles (although this difference may stem from enhanced sensitivity of modern techniques). Regulatory regions of the Act57B, Act79B, and Act88F genes have been identified in D. melanogaster (204, 350, 431), and the Act88F gene promoter has been used to drive green fluorescent protein expression (11).

The different actins are functionally nonequivalent (336), and (although they differ by only 15 amino acids) mammalian cytoplasmic actin cannot substitute for Act88F (129). Drosophila indirect flight muscle actin requires posttranslational modification for normal polymerization (419, 722, 1057) (which presumably underlies the indirect flight muscle “actin III” reported in Ref. 449; see also Refs. 430, 647) and is the only known actin to have an unacetylated, free NH2 terminus (1057).

Multiple mutants of Drosophila indirect flight muscle actin that alter muscle force production, despite in some cases assembling into seemingly normal thin filaments,
have been obtained (16, 20, 222, 260–262, 430, 815, 897, 974, 1038, 1108, 1109), as have mutants that disrupt myofibril structure (16, 222, 260, 536, 540, 722, 1109). In a study comparing several Act88F mutants, in almost all cases protein stability was similar (261). In one mutant in which muscle force generation is altered but thin filament structure appears normal, the mutation site is distant from the myosin binding site, and actin mutations can therefore have long-range effects on force generation (262). Experiments measuring the effect of actin mutation on profilin, ATP, and DNase I binding showed similar distant effects for some mutants (259). Other mutants identified Gh1 as part of a secondary myosin binding site (974), electrostatic charge on actin domain two as critical for thin filament regulation by tropomyosin (115), and the binding site of Clostridium toxins (512). Mutants that produce no actin still produce relatively normal thick filament arrays, and experiments in which the actin-to-myosin ratio is altered suggest that filament imbalances, not lack of thin filaments per se, are responsible for the observed defects (89). Act88F epitope tagging on the COOH terminus results in flightlessness and disordered indirect flight muscle sarcomeres, but NH2 terminus tagging gives relatively normal sarcomeres and partially restores flight ability (130). Many [but not all (430, 540)] actin mutations induce heat shock protein production (260, 430, 431, 897, 925). The molecular basis of this induction is unknown but is independent of myofibril degeneration (430, 540, 1038, 1039).

Asynchronous muscle (a special type of muscle in which muscle contractions are not synchronized with motor neuron activity, see second review) in all species examined in Nepomorpha, and some species in Diptera, contain a ubiquitinated actin, arthrin (64, 93, 138, 1058). Arthrin and the tropomyosin/troponin complex are in equimolar concentration, suggesting that they may colocalize on the thin filament. This suggestion has not been verified, but it is known that the ubiquination site is on the opposite side of the thin filament from where tropomyosin binds (341). Arthrin’s function is unknown, as arthrin activates myosin ATPase, is regulated by troponin/tropomyosin in the same manner as actin, and its presence does not alter actomyosin kinetics (138, 1058). Arthrin is not required for asynchronous muscle, as mutant Drosophila in which actin ubiquination cannot occur still fly (1058), and arthrin has not been found in any Hymenoptera species with asynchronous muscle, and is absent from some Cicadellidae (a family of Cicadelloidea), Diptera, Coleoptera, and Heteroptera species, even though all animals in these groups have asynchronous muscle (Fig. 3) (937, 1058). Arthrin evolved independently at least twice, although in all cases the ubiquination site (Lys-118) is identical (148, 1058).

Actin genes from three other Diptera (A. aegypti, M. destructor, D. dorsalis) believed to code for muscle actin (on the basis of nucleotide and amino acid similarity to D. melanogaster genes) have been investigated (411, 466, 832, 1086, 1236). Southern blot analysis suggests Aedes contains at least five actin-related sequences (466) with differential expression in different muscles (832, 1236). Four muscle genes (identity confirmed by hybridization of gene specific clones to RNA extracted from different tissues) have been cloned in D. dorsalis (411). Developmental stage specific expression and differential expression in different muscles are present. The 3′- and 5′-flanking regions of these genes show very little sequence homology both among themselves and to other known actin genes. In Lepidoptera (B. mori) three actin genes have been identified, one cytoplasmic and two muscle, one of which is expressed only in adult muscle and the other in both larval and adult muscles (827, 828).

J) ACTIN AS A PHYLOGENETIC CHARACTER. Much of the above work investigates phylogenetic relationships (92, 127, 161, 218, 285, 292, 300, 334, 366, 500, 568, 605, 623, 624, 626, 627, 826, 829, 858, 903, 1085, 1204, 1239, 1269). Until recently, the generally accepted conclusion was that all invertebrate muscle actins are most closely related to vertebrate cytoplasmic actins, with insect muscle actins diverging very early from those of other invertebrates. Unfortunately, much of this work may be seriously flawed. First, much of it is based on diagnostic amino acids (623), a small number of differing amino acids that were early and successfully used to classify mammalian actins, and then (uncritically) applied to invertebrate actins. This approach has been strongly criticized because it 1) treats the other amino acids as carrying no phylogenetic information and 2) does not analyze gene evolution as a dynamic process of “descent with modification” (i.e., it is insufficient to just compare differences between two extant species; instead, enough information from multiple species must be obtained to infer the history of changes that resulted in the present differences) (1281).

Second, actin gene duplications and conversions (including between muscle and cytoplasmic forms), which complicate phylogenetic comparisons, have occurred in many lineages (218, 258, 624, 1280). Third, Drosophila actin genes show pronounced codon bias, which could bias phylogeny construction (410). Fourth, actin may not be particularly suitable as a phylogenetic marker. In most organisms actin is a relatively large proportion of total protein. Growth rate changes might therefore disproportionately affect actin evolution, and even closely related invertebrate species can grow at very different rates. Indeed, the actin gene duplications noted above may have arisen specifically to allow rapid growth in some lineages. Furthermore, different actin isoforms presumably at least partly reflect a need for muscles with different functional properties, a need that in many cases may depend on organism life-style. Closely related invertebrate species with different life-styles could thus have experienced
higher rates of actin evolution, again complicating phylogenetic analysis.

Fifth, actin-based phylogenies often have untenable relationships (Fig. 6). For instance, in phylogeny A, sea urchin and ascidia cytoplasmic actin are more closely related to mollusk cytoplasmic actin than to pufferfish cytoplasmic actin, even though sea urchin, ascidia, and pufferfish belong to Deuterostoma and mollusks to Lophotrochozoa (Fig. 1). Similarly, in phylogeny B, sea urchin cytoplasmic actin is more closely related to Cnidaria actin than to either ascidia or sea star cytoplasmic actin, even though Cnidaria belongs to Radiata and the other species to Bilateria. Again, in phylogeny C, sea star cytoplasmic actin is more closely related to urochordate cytoplasmic actin than it is to that of another echinoderm, sea urchin.

Muscle actin is part of a highly organized ensemble of proteins that might be expected to coevolve. A more fruitful approach therefore may be to build phylogenies for multiple muscle proteins (e.g., actin, troponin, myosin light and heavy chains) and superimpose the trees to arrive at the ensemble’s most likely phylogeny. When this is done (Fig. 6D), an ancestral set of muscle proteins gives rise to two branches, one of which leads to the vertebrate smooth muscle ensemble and the arthropod and vertebrate cytoplasmic ensembles. The other branch gives rise to urochordate smooth muscle and all striated muscles. Note that this phylogeny agrees with those determined by other methods (Figs. 1–3); in particular, in each branch all vertebrates are more closely related to each other than they are to nonvertebrates, and urochordates are more closely related to vertebrates than they are to arthropods.

2. Tropomyosin

Vertebrate muscle tropomyosin is a dimer of tropomyosin molecules arranged in an in-parallel, in-register coiled-coil. There are typically two isoforms, and tropo-

---

**Fig. 6.** Various molecular genetic-based phylogenies. **A–C** are based on actin alone. **D** is based on superimposition of troponin C, myosin essential and regulatory light chains, myosin heavy chain, actin, and (for the vertebrates) muscle regulatory factor trees. [A modified from Carlini et al. (161); B modified from White and Crother (1280); C modified from White and Crother (1281); D modified from Oota and Saitou (903).]
myosin can form as a hetero- or homodimer of either. Nonmuscle tropomyosin isoforms are also widely present. Multiple tropomyosin isoforms are generally present in invertebrates, and most work has concentrated on distinguishing muscle and nonmuscle forms. Whether invertebrate tropomyosin is a dimer, and if so, a hetero- or homodimer, is less investigated. However, equilibrium sedimentation work in crayfish, oyster, abalone, and blowfly indicate that the tropomyosins self-associate (1291), and optical rotary dispersion and hydrodynamic measurements suggest blowfly tropomyosin is 100% α-helical and in a two-stranded configuration (591). Given these data, invertebrate tropomyosin is presumably also a coiled-coil dimer. Muscle-type tropomyosin is in some cases expressed in nonmuscle tissues, and in some work below these tissues were used as tropomyosin source material. Reference 653 is a general (vertebrate and invertebrate) review covering muscle and nonmuscle tropomyosins. In the older literature, paramyosin is sometimes called tropomyosin A or even tropomyosin (for references, see sect. μB5 and second review); in this work tropomyosin is called tropomyosin B. In two very early articles, we are unable to determine, for the invertebrate portions, whether tropomyosin or paramyosin was being studied (938, 1174).

Tropomyosin has been identified by expressed sequence tagging, immunohistochemistry, protein isolation, or cloning in Cnidaria (53, 320, 385, 386, 699); amphioxus (1124); Ascidia (785, 1183); sea urchin (485, 486, 490–492, 1177); annelids (221, 658); mollusks: bivalves (179, 404, 469, 472, 481, 482, 484, 488, 494, 524, 658, 672, 738–740, 795, 931, 1152, 1254, 1291), gastropods (401, 480, 1291), cephalopods (448, 477, 483, 595, 806, 1197); brachipods (658); Chaeptognatha (1011); platyhelminths (283, 706, 1177); annelids (221, 658); mollusks: bivalves (179, 404, 482, 489, 740), some expressed in a muscle specific manner (488, 489, 494, 931), and multiple tropomyosin genes (404, 494, 931). NH₂-terminal blocking is important for head to tail polymerization, actin binding, and Ca²⁺ regulation in Akazara scallop tropomyosin (473). Bivalve tropomyosin levels change in characteristic ways in response to various pollutants (1000).

Platyhelminths have multiple isoforms, with some tissue-specific expression and likely multiple genes (283, 1277). C. elegans has one tropomyosin gene. Four isoforms arise by alternative splicing and show developmental stage (718) specific expression and are differentially expressed in different muscles (21, 515, 516). Limulus has four to six isoforms that are not expressed in a tissue-specific manner (804). Scorpion has four muscle isoforms that show some muscle specificity (805). Crustaceans have three or four muscle isoforms. One is uniquely expressed in cardiac muscle, and the others show considerable but not perfect muscle specific expression (487, 801, 803, 805, 837, 960, 1042). Two of the isoforms may arise by alternative splicing (837). Beetles and centipedes have three muscle tropomyosins, which show some muscle specificity (805). Locust has multiple isoforms with muscle specific expression. A cDNA clone has been sequenced, and two mRNAs are present, but whether they are from different genes is unknown (614). Drosophila has two muscle tropomyosin genes. Each produces four or five tissue or developmentally specific isoforms (including nonmuscle forms) by a combination of alternative splicing and multiple promoters (84–86, 381, 394, 395, 537–539, 1064) (two of these isoforms are the flight muscle specific tropoponin H, see below). Considerable evidence has been obtained in Drosophila on the regions of the tropomyosin gene that regulates its expression (686, 689, 690, 788, 1064).

Drosophila (538, 613, 817, 1164) and C. elegans (21, 516, 1282) tropomyosin mutations alter muscle structure and function. Some Drosophila mutants can be rescued by reprogramming correct tropomyosin sequences (337, 1163, 1164). Despite the partial tissue specificity noted above, in one case substitution of different isoforms can also rescue tropomyosin mutants (794).

Tropomyosin is an important component of immune and allergic reactions (7, 25, 29, 33, 37, 38, 40, 41, 50, 51, 150, 157, 158, 179, 233, 255, 295, 362, 403, 479–484, 501–
3. Troponin

Vertebrate troponin contains three subunits: troponins T, I, and C. Ascidia (273, 882, 1183); Glycera, Lumbricus, and Nereis (annelids) (658); scallop (363, 861, 885, 888–891) and squid (477, 595, 892, 1081, 1193); nematode (461, 564); Limulus (658, 659); shrimp (895, 960), lobster (202, 790, 852, 860, 985, 1081), and crayfish (103); and Drosophila (139, 143) and Lethocerus (658) troponin also contains three subunits (although with considerable molecular weight variation). Troponin T is present in cross and obliquely striated filament (192), fire ant (945), and Drosophila (68, 88, 1012, 1017, 1193, 1277, 1278, 1285, 1295, 1296, 1312).

Troponin-based phylogenies agree with Figures 1–3 (184, 187, 471, 1149, 1320). Multiple isoforms, some developmental stage or muscle specific, are present in ascidia (271, 272, 714, 882, 1321, 1323); scallop (888, 1325); C. elegans (843); shrimp, crayfish, and lobster (202, 343, 581, 783, 790, 834–836, 838, 852, 860, 895, 985, 1083, 1286); barnacle (35, 192); Lethocerus (959), Anopheles (959), dragonfly (305, 735–737); and Drosophila (68, 69, 101, 257, 328, 335, 395, 423, 538, 539, 959). Stage-specific isoform changes are associated with increased Ca2+ sensitivity and altered twitch contraction kinetics in aging dragonflies (305, 736), and expression of the asynchronous muscle isoform with flight ability in bee (257).

Halocynthia (ascidia) has one adult (which produces two isoforms by alternative splicing) and at least three larval troponin I genes; sequences regulating gene expression have been identified (1321, 1323). Ciona, alternatively, has only one troponin I gene (which again produces two isoforms), and no homologs to the Halocynthia larval genes (714, 1324). Only one troponin I gene has been identified in Drosophila (68, 69, 88, 244); sequences regulating gene expression have been identified (741, 760). Ascidia, amphioxus, sandworm, and scallop have one troponin C gene (1319, 1322, 1325, 1326), C. elegans and barnacle two (192, 1168), lobster three (343), and Drosophila five (328, 423). The fourth intron of invertebrate troponin C genes shows great variability, suggesting that the ancestral gene may not have possessed it (1325, 1326). Ascidia has two troponin T genes (271, 272) and Drosophila and dragonfly one (101, 735). Troponin I, C, T, or H mutants can disrupt muscle function or development (68, 69, 88, 335, 516, 538, 771, 833, 866, 1168), although some polymorphism is tolerated (224). Drosophila troponin I mutants can be suppressed by troponymosin (841) or myosin heavy chain (299, 615, 866, 867) mutations or troponin I second site mutations (952).

4. Calponin/caldesmon

Calponin-like proteins have been identified in Eisenia (1017); Mytilus (325); Helix (1012); Echinococcus [called myophilin (748–751)] and Schistosoma (510, 3106); and Onchocerca (where it is highly immunogenic) (475), but not in crustacean or Drosophila muscle (1012, 1017). Calponin-like cDNAs in a nematode (Meloidogyne incognita) (162), Echinococcus (751), and Schistosoma (1306) have been sequenced. Echinococcus has only one gene, but multiple isoforms are expressed due to post-translational modification, including phosphorylation by protein kinase C (750). In some Schistosoma species, several copies are present, and multiple isoforms are expressed (1306). Caldesmon-like proteins have been
identified in Eisenia (1017) and mollusks (76, 100, 225, 1012).

5. C. elegans unc-87

The unc-87 mutants have almost normal embryonic muscles that become severely disorganized as the animals mature (1266). unc-87 codes for a 40-kDa protein that is located in the I band, bundles actin filaments (but not monomers) in the absence of tropomyosin or α-actinin, and has the same actin binding sequence as calponin (361, 607). When unc-87 mutants are expressed in animals with decreased myosin heavy chain activity, the age-dependent disorganization is less, suggesting that it arises from contraction force. unc-87 is therefore not believed to be required for sarcomere formation, but instead to serve a structural role (360).

B. Thick Filament Proteins

1. Myosin heavy chain

There are at least 13 classes of myosin; all muscle (and some nonmuscle) myosins belong to class II (61, 104, 172, 200, 374, 433, 1069, 1198, 1302). We cover here only muscle myosins. References 26, 66, 400, 776, and 1134 review vertebrate and invertebrate muscle and nonmuscle myosins, and Reference 42 reviews vertebrate nonmuscle (but still class II) myosins early split from the myosins leading to vertebrate striated muscle and all invertebrate muscle myosins. A second split resulted in two branches, one of which contains all striated (both skeletal and cardiac) vertebrate and invertebrate myosins and the other of which contains nematode muscle myosin (which has a special type of striation called obliquely striated, see third review). However, given the paucity of invertebrate species contained in this work and in particular the lack of smooth invertebrate muscles in it, these groupings must be considered preliminary.

A) CNIDARIA. A cDNA clone of a striated muscle specific myosin heavy chain has been sequenced and is more similar to striated muscle heavy chain isoforms than to smooth or nonmuscle isoforms in either vertebrates or invertebrates. Because cnidarians also have smooth muscle, at least one other muscle isoform is presumably present. Southern blotting suggests this/these isoform(s) would be from other genes, not alternative splicing (5, 1061).

B) MOLLUSCA. For bivalves, see References 67, 163–165, 405, 452, 453, 496, 567, 594, 597, 599, 622, 738, 740, 862, 864, 873, 874, 941, 1101, 1246); gastropods, References 32, 52; cephalopods, References 563, 580, 768. Myosin heavy chain cDNAs or genes have been sequenced in scallops Argopecten (also known as Aequipecten) (873, 874), Patinopecten (405), Pecten (496), and Placopecten (941) and squid Loligo (768). Scallop (Placopecten, Argopecten) has three or four isoforms alternatively spliced from a single gene and expressed in a muscle-specific fashion (874, 941). Mollusk striated and catch (a special muscle type that maintains contraction with very little energy expenditure, see second review) muscle myosin heavy chains differ almost exclusively in only one region (surface loop 1). The two isoforms have different ATPase activities (941) and ADP affinities and dissociation rates (622). The sequence responsible for light-chain positioning on the heavy chain has been identified in the scallop (452). Regulatory light-chain kinase phosphorylates scallop myosin heavy chain (1101), and Mytilus myosin contains a tightly bound endogenous kinase (163). Which residues were phosphorylated in these works is unknown, but phosphorylation of specific serines in Mytilus tailpiece increases myosin solubility and favors molecular folding (164, 165). Fish light chains bind to scallop heavy chain (567, 1101). Squid has at least two isoforms, alternatively spliced from one gene (768).

C) PLATYHELMINTHS. See References 14, 387, 572, 904, 1028, and 1276 for platyhelminth data. Planaria have two muscle myosin heavy chain genes expressed in nonoverlapping sets of muscle (572, 904). Although all planarian muscles are nonstriated, planaria heavy chain myosins most closely resemble the striated muscle myosin heavy chains of other organisms (572). Phylogenetic analysis using myosin heavy chain genes suggests that the platyhelmints are polyphyletic and that two platyhelmint groups, the Acoela and Nemertodermatida, are the earliest extant bilaterians (1028). Myosin may be useful in platyhelminth vaccine development (13, 108, 114, 241, 254, 268, 282, 509, 682, 849, 935, 1103–1105, 1338–1340).

D) NEMATODA. See References 24, 94, 250, 251, 275, 278, 280, 302, 402, 445–447, 497, 541, 693, 711, 715–717, 792, 793, 812, 844, 855, 898, 956, 1053, 1161, 1260, 1265, 1274, 1275, 1279, and 1329 for nematode data. C. elegans has
four myosin heavy chain isoforms (MHC A, B, C, D), coded by widely separated genes (myo-3, unc 54, myo-2, myo-1, respectively), some on different chromosomes (10, 251, 280, 541, 715, 717, 793, 1053). MHC A and B are used only in body wall muscle and MCH C and D only in pharyngeal muscle (24, 280, 711). The molecular basis of this differential expression is beginning to be investigated (497, 898). MHC A is present only in the center of the thick filament and MHC B only at the ends (792), possibly because MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812). MHC B overexpression disrupts muscle structure (302). MHC A’s more hydrophobic rod surface binds better to paramyosin (which forms the core of nematode thick filaments, see second review) (447). Depending on the mutation, MHC B mutants (280, 716, 717) are paralyzed with a reduced number of thick filaments composed entirely of MCH A (94, 250, 278) or very slowly moving with normally assembled thick filaments (250, 812).
show different dependencies on the presence of phosphate, suggesting that the difference between them is that in the embryonic form the rate-limiting step of the crossbridge cycle is ADP release, whereas in the indirect flight muscle form the rate-limiting step is phosphate release (during the cross-bridge cycle, myosin head rotation is associated first with phosphate and then ADP release, see second review) (1131). Substituting only the embryonic S1 exon (one of four in this region that differ) decreases ATPase activity, but does not affect actin sliding velocity, flight muscle ultrastructure, or flight ability (1128). However, switching of indirect flight muscle exons into the embryonic isoform (when the embryonic isoform is the one being expressed in the flight muscles), or embryonic exons into the flight muscle isoform, alters wingbeat frequency and increases or decreases, respectively, the muscle’s maximum power generation and optimal wingbeat frequency for power generation (1130). Unlike Drosophila, two myosin heavy isoforms are present in fleshfly (Phormia) asynchronous and locust synchronous flight muscle (1344).

2. Catchin/myorod

An alternative splice of the myosin heavy chain gene codes for catchin/myorod (1075–1077, 1299), a protein found in smooth muscle thick filaments in a variety of bivalves. The protein contains the COOH-terminal rod of myosin heavy chain but has a unique NH2-terminal head (1076, 1299) and, although most highly expressed in catch muscles (1299), is not necessary for catch (1298). Despite the identity of the tail portions, pure catchin/myorod polymerizes very differently than myosin or myosin rod. Adding myosin heavy chain to myorod preparations dramatically alters myorod polymerization (1075). A similar isoform, myosin rod protein, is encoded in Drosophila by a gene internal to the myosin heavy chain gene. Myosin rod protein is present in a variety of larval, embryonic, and adult cardiac, visceral, and somatic (including direct flight) muscles and localizes with myosin in the A band (950, 1114). Myosin rod protein may be associated with the relatively disordered thin and thick filament packing, variable thin-to-thick filament ratio, and bent thin filaments observed in some direct flight muscles (1114). In muscles expressing high levels of myosin rod protein, thick filaments and sarcomeres still form in animals lacking myosin heavy chain, suggesting that sarcomere formation does not require actin-myosin heavy chain interaction (950).

3. Myosin light chains

Vertebrate myosin possesses two light chains, the essential and regulatory, in equimolar amounts (one each per myosin molecule, and hence two each per myosin heavy chain dimer). All known invertebrate myosins similarly contain two essential (also called SH, catalytic, myosin light chain 1, and alkali light chain) and two regulatory (also called myosin light chain 2) light chains, both of which have an ellipsoidal shape (Fig. 4) in a variety of species (1111). [Caution: in early scallop papers, the regulatory light chain is also called the EDTA chain because EDTA exposure removes one or both (depending on concentration and temperature) regulatory light chains (32, 169, 553, 596, 1139), presumably due to the enhanced regulatory light chain binding to the heavy chain that Ca2+ and Mg2+ induce (31, 1139). EDTA sensitivity is not present in many other species (553, 654), and the term is not used in modern parlance, but could be confusing when reading the older literature.]

Whether all invertebrate myosins contain equimolar amounts of the two light chains, however, is not as clear. Early work in the scallop suggested that there were two moles of essential light chain but only one of regulatory light chain per myosin heavy chain dimer (1139). More recent work, however, shows that scallop myosin does contain two moles of each light chain per dimer (553, 1294). The situation in Crustacea remains unclear. Lobster and crayfish myosin contains two light chain types, α and β, each with multiple variants, but which type is essential and which is regulatory is unknown (570, 834, 984). In some lobster muscle fibers, α-chain immunohistochemical staining is higher in the center of the thick filaments, whereas β-chain staining is always uniformly distributed (133). Although these data are suggestive, they do not prove nonuniform chain distribution because they could result from nonuniform masking of the α-chain in different thick filament regions. No difference is seen by two-dimensional gel electrophoresis in lobster light chains from slow versus fast muscle fibers (684), although different isoforms are present in the two muscle types in crayfish (1040). In earthworms, different light chain isoforms are expressed in a muscle-specific fashion, but which are regulatory and which are essential is unknown (160). Ascaris has two light chains in an approximate 1:1 molar ratio, but which are regulatory and which are essential is unknown (844). In vertebrates the light chains are associated with the myosin heads near the neck region (Fig. 4A). Scallop light chains are similarly situated, and the regulatory and essential chains extensively interact (34, 217, 309, 399, 598, 1112, 1140, 1217, 1219, 1244, 1245, 1247, 1283, 1294). Bivalve light-chain levels show characteristic changes in response to various pollutants (1000).

Essential light chains have been investigated in amphioxus (438), ascidia (1148), annelid (249a), several bivalves (70, 190, 375, 496, 723, 739, 740, 862, 940, 1139, 1246), abalone (32), squid (595, 596, 1258), and Drosophila (181, 288–290, 665, 1153, 1155, 1156). In amphioxus, scallop, and Drosophila only one gene was found. Early work with gel electrophoresis showed that essential light
chains vary across bivalve species (740) but were identical in all muscle types in the one species (Spisula) in which this question was examined (739). Later radioimmunoassay work using scallop antibodies showed that the essential and regulatory light chains differ immunologically, that the antibodies did not cross-react with other invertebrate essential light chains, and that striated and smooth, but not heart, essential light chains were immunologically identical (1246). Sequencing work showed that several essential light-chain transcripts are present in scallop (375), and identical isoforms are present in striped and catch muscle (940) (cardiac muscle was not examined in this work). Essential light-chain isoforms, generated by alternative splicing and protein phosphorylation, are expressed in a muscle-specific fashion (indirect flight muscles vs. all others) in Drosophila (288, 289, 1155, 1156, 1160). The splice patterns are maintained across several Drosophila species (665).

Regulatory light chains have been investigated in amphioxus (1124); ascidia (1148); earthworm (228, 1072) and the pogonophore Riftia (971); several bivalves (70, 375, 496, 592, 593, 597, 724, 739, 740, 800, 824, 862, 940, 1139, 1157, 1246), abalone (32), and squid (595, 596, 725); C. elegans (227, 1029, 1030); Limulus (1248) and tarantula (425); and Drosophila (924, 1153, 1155, 1156, 1178). Early gel electrophoresis work showed that regulatory light chains vary across bivalve species (740). Later radioimmunoassay work showed that scallop regulatory light chain antibodies did not cross-react with other invertebrate regulatory light chains, and that striated and smooth, but not heart, regulatory light chains are immunologically identical (1246). Bivalves (375, 940) and Drosophila (924, 1178) have one gene, and C. elegans has two (227). Muscle, developmental stage, or gender specific regulatory light-chain isoforms exist in earthworm (228), C. elegans (1030), and Drosophila (1156, 1160, 1178). Muscle specific isoforms are apparently not present in all bivalves, as only one isoform is present in clam (Spisula sachalinensis) adductor mantle, leg elevator, and heart muscle (739), but two (phasic vs. catch muscle) in scallop (592, 593, 824, 940). However, the different phasic and catch muscle ATPase activities in bivalves are due to differences in the myosin heavy chains, not the regulatory light chains (940). Blocking scallop heavy chain thiol groups blocks regulatory chain binding (594, 597, 896).

In C. elegans one gene is redundant, but deletion of the other causes lethal pharyngeal muscle defects in hermaphrodites. Paramyosin/mini-paramyosin have been studied in cephalochordata (167, 437), echinoderms (876, 1195, 1256, 1284), annelids (1013, 1018, 1270, 1284), mollusks [Polyplacophora (786)], bivalves (2, 56–60, 197, 198, 205–208, 270, 392, 397, 529, 547–549, 636, 641, 642, 677, 704, 738, 740, 766, 775, 786, 789, 795–797, 944, 991, 992, 1027, 1113, 1137, 1255–1257, 1270, 1284, 1290, 1291, 1311) (called tropomyosin A in Refs. 57, 60, 124, 125, 1252), changing a single amino acid from serine to alanine reduces flight muscle power output (252). Drosophila regulatory light-chain NH2-terminal deletions reduce indirect flight muscle calcium sensitivity, maximum amplitude of oscillatory work, sarcomere length, dynamic stiffness, and elastic modulus (476, 820). The stiffness and elastic modulus changes suggest that in these muscles regulatory light chain could link the thick and thin filaments.

4. Myosin light-chain phosphorylation

Expressed sequence tag analysis shows that a light chain kinase is present in amphioxus notochord (1124). A scallop cAMP-dependent kinase that phosphorylates only one regulatory light chain isoform and a protein that modulates the kinase have been identified (1098–1100, 1102). A Limulus light chain specific kinase has been purified (1071). The Drosophila light chain kinase gene has been sequenced. Tissue- and stage-specific isoforms are present; only some forms are Ca2+-calmodulin dependent (586, 1179). Reducing Drosophila regulatory light-chain phosphorylation reduces myosin ATPase activity (1154) and, by reducing muscle stretch activation (a special property of the asynchronous muscles present in this organism, see second review) (1110, 1180), indirect flight muscle power output (252, 1110, 1180). Regulatory light-chain phosphorylation increases tarantula striated muscle contraction; phosphorylation levels vary with Ca2+ concentration (425). Ascaris cuticle muscle regulatory light chain is 48% phosphorylated in controls; GABA application reduces this phosphorylation to 18% (747). Phosphorylation of Limulus regulatory light chains in purified thick filaments changes filament structure (674). Twitchin also phosphorylates regulatory light chains (415). Light chain phosphatases have been characterized in clam (1194) and scallop; the scallop enzyme is Ca2+-dependent (474).

5. Paramyosin/mini-paramyosin

We cover here only paramyosin genes and mutants (see second review for structure of paramyosin containing thick filaments). Paramyosin/mini-paramyosin have been studied in cephalochordata (167, 437), echinoderms (876, 1195, 1256, 1284), annelids (1013, 1018, 1270, 1284), mollusks [Polyplacophora (786)], bivalves (2, 56–60, 197, 198, 205–208, 270, 392, 397, 529, 547–549, 636, 641, 642, 677, 704, 738, 740, 766, 775, 786, 789, 795–797, 944, 991, 992, 1027, 1113, 1137, 1255–1257, 1270, 1284, 1290, 1291, 1311) (called tropomyosin A in Refs. 57, 60, 124, 125, 1252), changing a single amino acid from serine to alanine reduces flight muscle power output (252). Drosophila regulatory light-chain NH2-terminal deletions reduce indirect flight muscle calcium sensitivity, maximum amplitude of oscillatory work, sarcomere length, dynamic stiffness, and elastic modulus (476, 820). The stiffness and elastic modulus changes suggest that in these muscles regulatory light chain could link the thick and thin filaments.
1291), cephalopods (56, 561, 619, 762, 1197), euchopods (1284), brachiopods (1011), platyhelminths (91, 156, 298, 359, 387, 444, 478, 521, 830, 1056, 1207, 1284, 1305), nematodes (C. elegans (24, 277, 346, 402, 514, 710, 939, 1059, 1133, 1256, 1262, 1263), Ascaris lumbricoides (1284), Brugia malayi (648, 1274), Limulus (476, 675, 676, 852, 944, 1013, 1016, 1018, 1021, 1231, 1232), Coleoptera (Heliocoris japonicus, Pachnoda ephippia) (137), Heteroptera (Lethocerus (141, 145, 270, 676, 677), and Orthoptera (Locusta migratoria, Schistocerca gregaria) (93, 590, 1284). References 270, 528, 677, 1013, 1018, 1256, and 1284 are reviews. Paramyosin is observed in amphioxus notochord (311, 437), catch (56, 60, 270, 397, 448, 766, 1027, 1127), smooth (24, 401, 876, 1013, 1018, 1027, 1195, 1284), obliquely striated (24, 396, 561, 1018, 1262, 1284), and cross-striated (24, 145, 166, 202, 248, 270, 468, 561, 675, 676, 852, 944, 1013, 1016, 1018, 1021, 1040, 1056, 1231, 1284) muscle and thus is not a marker of catch ability or muscle type (except for existing only in invertebrates). Paramyosin expression varies in mollusk cross, obliquely, and nonstriated (smooth) muscles, although which types express more is inconsistent across species (561, 1018).

Paramyosins from two bivalves, a gastropod, and a polyplacophore all have similar molecular weights, electroforethic properties, solubility, and guanidine-HCl denaturation and pepsin and trypsin susceptibilities (786). Paramyosins from two smooth and one striated molluscan muscles, and four arthropod (Limulus, Homarus, barnacle, insect) striated muscles, have similar molecular weights and are immunologically similar (270). Heteroptera and Coleoptera indirect flight muscle paramyosins also have similar intrinsic sedimentation rates and circular dichromism values (145). However, this does not mean that all paramyosins are identical; paramyosin from different species (248, 528, 529, 766, 876, 1256) and multiple isoforms within single species (2, 202, 693, 384–386, 852, 1040, 1056, 1106, 1231, 1243) show differences in antigenic structure, molecular weight, isoelectric focusing, phosphorylation ability, or amino acid composition. The two to three paramyosin isoforms present in decapod crustaceans show muscle specific expression (834–836, 852, 1040).

The sources of this variation are not completely known [for very early work, some may arise from proteolytic degradation during extraction (1311)]. Paramyosin is phosphorylatable (2, 197, 249, 1056, 1059, 1101, 1256, 1257), which could explain differences in antigenic structure, isoelectric focusing, and phosphorylation ability (since paramyosins with different phosphorylation states in vivo would vary in their ability to be further phosphorylated in vitro). The amount of phosphorylation observed depends on extraction procedure (197). In mollusks Ca^{2+}, cAMP, and calmodulin-dependent kinases phosphorylate paramyosin (1036), neuromodulators alter phosphorylation ability (2), and phosphorylation alters myofilibril ATPase activity (1257), suggesting that paramyosin phosphorylation could be functionally relevant.

Paramyosin isoforms could also arise from multiple genes. Mytilus has two paramyosin genes, but whether one is a separate miniparamyosin gene is unknown (1255). Two genes may be present in the nematode Anisakis (939), and the platyhelminth Schistosoma has more than one (444). Paramyosin cDNAs or genes have been cloned in the cestode Taenia solium (1207) (from the T. solium gene a mini-paramyosin cannot expressed); nematodes Brugia (648), Onchocerca (231, 688), and Dicrofilaria (376, 688); tick Boophilus (297); and mite Blomia (967) but whether these are single copy is unclear.

C. elegans [unc-15 (514, 993, 1263)] has a single paramyosin gene but two paramyosin isoforms, which are differentially located in the thick filaments, one being associated with the filament core and the other with the myosin heavy chains (240, 693). Reference 514 asserts that paramyosin expression differs in body wall and pharyngeal muscles, but we have been unable to find sources showing this difference. Regardless, paramyosin is expressed in all C. elegans muscle (24). C. elegans paramyosin mutants show varying degrees of locomotion impairment, and in the most extreme phenotype are paralyzed as adults (1263). The body wall muscles of these mutants show aberrant myofilament lattice and thick filament structure (277, 1263). Changing a single charge on the paramyosin can disrupt thick filament assembly (346), and paramyosin appears to be required to determine thick filament length (710) [as is also the case in Limulus (468) and locust (590)]. Many C. elegans paramyosin mutants that disrupt body wall muscle function do not alter pharyngeal muscle function, although the pharyngeal thick filaments are missing their normal paramyosin core (1263). This difference is believed to be because pharyngeal muscles have less paramyosin, use different myosin heavy chains than body wall muscles, and have only one sarcomer (514, 1263). Paramyosin degradation is increased in C. elegans mutants lacking myosin heavy chain B, although synthesis is unaffected (1279).

Drosophila (1232) has a single paramyosin gene that is required for proper myoblast fusion, myofilibril assembly, and muscle contraction (696). Paramyosin is much more highly expressed in tubular than in flight muscle in Drosophila (1231). Comparing paramyosin content of solid- versus open-cored thick filaments (see second and third reviews) across several insects shows that paramyosin...
sin is a larger percentage of filament mass in solid-cored filaments (93).


Miniparamyosin is present in Drosophila (90, 744, 1013), shrimp (1016), and mite striated (1021) muscle and Eisenia obliquely striated and smooth (1013) muscle. Whole animal extracts show it to be present in annelid, snail, mussel, and multiple arthropod species, but not C. elegans, a result supported by analysis of the C. elegans paramyosin gene (744). Drosophila miniparamyosin is encoded by the same gene as paramyosin, but its expression is transcriptionally regulated from a separate promoter (27, 90, 744). Drosophila miniparamyosin is expressed only in adults and has several isoforms (with much greater molecular weight variation than those known for paramyosin) that are expressed in a muscle specific pattern (90, 743). The sequences that regulate miniparamyosin/paramyosin expression are beginning to be described (27). Miniparamyosin overexpression results in nearly normal thick filaments and sarcomere length. However, some defects are present in indirect flight muscle with some flight impairment, and these defects increase with age (28).

6. Filagenins

C. elegans thick filaments are associated with three 30-, 28-, and 20-kDa proteins (α-, β-, and γ-filagenin, respectively) (240), whose cDNAs have been sequenced (694, 695). The proteins are believed to form the thick filament core by cross-linking paramyosin subfilaments (693). The α-form is located with myosin heavy chain A in the middle of the thick filaments while the γ-form is located more laterally with myosin heavy chain B (695). The α- and γ-forms appear at different developmental stages (695). The β-form is not present in pharyngeal muscles (694).

7. C. elegans unc-45

Temperature-sensitive unc-45 mutants have reduced body wall muscle thick filament number and are paralyzed as adults (279); homozygous, non-temperature-sensitive mutations result in arrested muscle development and are lethal (1213). The gene has been sequenced and is expressed in all C. elegans muscles, and only in muscle (1212). The protein associates with myosin heavy chain B in the thick filament. In myosin heavy chain B null mutants unc-45 muscle localization is not seen (and thick filament structure is near normal) (22). unc-45 binds myosin heads (both scallop and C. elegans, although unfortunately myosin heavy chain A and B were not separately tested) (72). Temperature-sensitive mutations in a region of unc-45 homologous to two fungal proteins involved in protein segregation result in (at the restrictive temperature) myosin heavy chain A and B being scrambled in the thick filaments (as opposed to being located in distinct thick filament regions as in wild type) (71). This work suggests that unc-45 is a chaperone that helps correctly integrate myosin heavy chain B into the thick filament (693, 1318).

8. Myonin

A 230-kDa thick filament associated protein named myonin that appears to regulate myosin heavy chain ATPase activity has been identified in clam (1310).

9. Flightin

Flightin is a 20-kDa protein that has been found only in Drosophila (1225, 1229), but efforts to identify it in other species have not been reported. It is found only in indirect flight muscles and is localized in the A band except for the H-band region (Fig. 8) (1229). The protein has multiple phosphorylation sites, and its phosphorylation state increases with age, particularly after eclosion (1227). The single amino acid myosin gene mutations noted above that cause age-dependent myofibril degeneration also fail to accumulate phosphorylated forms of flightin (618). Mutations that block thick filament assembly block all flightin phosphorylation, whereas mutations that block thin filament assembly alter flightin phosphorylation temporal sequence (1224). Flightin null mutants are viable but flightless (976). Mutant sarcomeres appear relatively normal in pupa and early adult but are ~25% longer than normal. When the animals attempt to fly, the indirect flight muscles become disrupted, and eventually site-specific thick filament cleavage occurs. Flightin binds to the myosin rod, and the flightin:myosin molar ratio is between 1:1 and 1:2 (44).

10. Zeelin

Zeelin 1 and 2 have been isolated from Lethocerus muscles (294, 1037). Zeelin 1 is present in flight and leg muscles, and zeelin 2 is only in flight muscles. In flight muscles the zeelins are present in discrete A-band regions, whereas in leg muscles the entire A band is labeled (the identification in Ref. 1037 that these proteins were in the Z disk was an artifact of fiber glycerination) (294). Both proteins are associated with the thick filament, with
zeelin 1 closer to the filament shaft than zeelin 2, and ~0.5 mol zeelin/mol myosin (294). Zeelin function is unknown. The small size and location of the zeelins are similar to *Drosophila* flightins, but antibodies to zeelin do not recognize flightin (294).

C. α-Actinin and Other Z Line Proteins

α-Actinin has been isolated or identified by immunohistochemistry in annelid (*Eisenia*) obliquely striated and smooth muscle (1010); bivalve (scallop) cross-striated, obliquely striated, and smooth muscle (559, 1074); gastropod (*Helix*) cross-striated (heart) and smooth muscle (1010); cephalopod (squid) obliquely striated muscle (1196); chaetognath (* Sagitta friderici*) body wall muscle; platyhelminth (*Schistosoma mansoni*) muscle of unidentified type (706); *C. elegans* obliquely striated muscle (316), mite cross-striated muscle (1021); and *Drosophila* (1010, 1035, 1228), honeybee (*Apis mellifera*) (1034), and giant water bug (*Lethocerus cordofanus*) (1010, 1035, 1228). Muta-
sions can be fatal or cause severe muscle defects (327, 338, 441–443). Rescue experiments show that the in-
direct/jump/leg muscle isoform can replace the cytoplas-
mic and larval/head/abdominal forms and

D. Ca²⁺ Binding Proteins and Their Targets

A wide variety of Ca²⁺ binding proteins (calmodulin, Ca²⁺-dependent Ca²⁺ binding protein, troponin C, the B subunit of phosphatase 2B, myosin light chains, parval-
bumins) have evolved from a common ancestor with four Ca²⁺ binding domains (see Refs. 574, 1144 for vertebrate references, Refs. 610, 611, 818 for extensive summaries of Ca²⁺ binding proteins and their evolutionary relationships, and Ref. 746 for a comparison of Ca²⁺ binding site amino acid complements). These proteins are divided into two types (209, 210). Proteins of the first (e.g., vertebrate parvalbumin) are present in high quantity, have high Ca²⁺ affinity, have Ca²⁺ binding regions with a stable configuration, do not interact with known targets, are believed to function as Ca²⁺ buffers and to help resquester Ca²⁺ after muscle contraction or to protect against deleterious effects of high Ca²⁺ concentrations during prolonged con-
tractions, and show relatively rapid evolutionary diver-
ge (reviewed in Ref. 421). Proteins of the second are present in lower quantity, have lower Ca²⁺ affinity, have Ca²⁺ binding regions whose conformation changes upon Ca²⁺ binding, are believed to be more directly involved in excitation-contraction coupling (examples include tropo-
nin C and calmodulin), and are relatively well conserved. We cover here only muscle specific Ca²⁺ binding proteins (in particular, calmodulin, which is present in a wide variety of tissues, is not included). Reference 348 reviews Ca²⁺ binding proteins in general.

Amphioxus (*Branchiostoma lanceolatum*) has seven sarcoplasmic Ca²⁺-binding proteins of the first type, all arising by alternative splicing from one gene. One isoform has been crystallized and its structure determined at 2.4-Å resolution (194, 196). The two isoforms for which it has been examined bind three Ca²⁺ (582, 1141, 1146, 1150). Immunohistochemistry shows the proteins are localized at the Z line (1202). One protein of the second type, Ca²⁺ vector protein, has been characterized in amphioxus and binds two Ca²⁺ (62, 63, 209, 211, 574, 1124). The mole-
cule’s three-dimensional solution structure, backbone dyn-
amics, and Ca²⁺ binding properties have been studied in great detail (62, 63, 1169–1171). Its cDNA has been se-
quenced (1320), and although clearly a Ca²⁺ binding pro-
tein, no homolog has been found in invertebrates or ver-
tebrates (1202). The protein is present in a wide variety of tissues, but in highest concentration (50–100 μM) in mus-
cle (209, 412). Immunohistochemistry shows it to be strongly localized at the Z lines and weakly localized at the M lines, and it is more difficult to extract from muscle than the sarcoplasmic Ca²⁺-binding proteins (1202). Ca²⁺ protein) 43-kDa *Drosophila* Z-disk associated protein has been identified, but its structure and function are unknown (998).
vector protein forms a high-affinity, Ca\(^{2+}\)-strengthened, complex with Ca\(^{2+}\) vector target protein (209, 942). Ca\(^{2+}\) vector target protein is also present in both muscle and nonmuscle tissues, but equally stains the Z and M lines (1202). Ca\(^{2+}\) vector target protein does not appear to have any enzymatic role, and its function is unknown. However, it has two immunoglobulin domains and thus could, as do other immunoglobulin containing muscle proteins (see sect. m6), contribute to sarcomere structure (1142). cAMP-dependent protein kinase phosphorylates four Ca\(^{2+}\) target protein sites, and phosphorylation decreases Ca\(^{2+}\) vector protein and Ca\(^{2+}\) vector target protein affinity (943).

Low-molecular-weight muscle-associated Ca\(^{2+}\) binding proteins are present in several other invertebrates. *Nereis diversicolor* and *Perinereis vancuaurica* tetradentata (annelid) sarcoplasmic Ca\(^{2+}\) binding proteins bind three Ca\(^{2+}\), but that of *Nereis virens* only binds two. The *N. diversicolor* and *Perinereis* proteins have been sequenced (189, 213, 242, 577). The *Nereis* protein’s three-dimensional structure (54, 195, 1230) and dynamics of Ca\(^{2+}\) and Mg\(^{2+}\) binding (274) have been determined, and the conformational changes associated with Ca\(^{2+}\) binding are being intensively studied (178, 216, 266, 705, 953, 1094). Earthworm has three Ca\(^{2+}\) binding protein isoforms that bind either two or three Ca\(^{2+}\), and which are present in higher concentrations in fast relaxing muscles (consistent with their being involved in Ca\(^{2+}\) buffering) (459, 460). The 20-kDa Ca\(^{2+}\) binding proteins have been isolated from two scallops. The *Patinopesten* protein has been sequenced and likely binds two Ca\(^{2+}\) (191, 1143). Clams and oysters have three Ca\(^{2+}\) binding protein isoforms (1141). Low (17 kDa)- and high (450 kDa)-molecular-weight Ca\(^{2+}\) binding proteins and a calsequestrin (the major Ca\(^{2+}\) binding protein of vertebrate muscle sarcomeres to only bind two (943)). Clams and oysters have Ca\(^{2+}\) binding proteins with sequence similarity to calcyphosine, a vertebrate nonmuscle Ca\(^{2+}\) binding protein (1051). Although early work was unable to find sarcoplasmic Ca\(^{2+}\) binding proteins in locust (212), 16- to 20-kDa Ca\(^{2+}\) binding proteins with muscle specific (tubular vs. indirect) expression were identified in *Drosophila* (46, 1160). The 23- to 24-kDa Ca\(^{2+}\) binding proteins were subsequently isolated from *Drosophila* and *Calliphora* (551, 560). The cDNA for the *Drosophila* protein has been cloned and is expressed in tubular but not indirect flight muscle (551). The gene is not single copy (551), and multiple isoforms are observed (560). A *Drosophila* calbindin homolog that is primarily expressed in neurons but is present in a few muscles has also been identified (986).

**E. Giant Sarcomere-Associated Proteins**

Like vertebrates, invertebrates also have large sarcomere proteins involved in centering the thick filaments in the sarcomere and generating muscle stiffness. Invertebrate muscles, however, show a number of variabilities not present in vertebrate sarcomeres (144, 323, 621). First, invertebrate sarcomeres cover a very wide length range (crayfish giant sarcomeres are \(\sim 8\) \(\mu\)m at rest whereas those of *Drosophila* indirect flight muscles are 3.3 \(\mu\)m). Second, invertebrate muscles have very different abilities to be stretched without damage (crayfish giant sarcomeres can stretch to 13 \(\mu\)m, *Drosophila* indirect flight muscle sarcomeres to only \(\sim 3.5\) \(\mu\)m). Third, invertebrate muscles show a very wide passive stiffness range (a tension of 8 mN/mm\(^2\) stretches crayfish sarcomeres 4 \(\mu\)m, but *Drosophila* sarcomeres only 0.2 \(\mu\)m).

Given these differences, invertebrates might be expected to have a wide variety of giant muscle proteins, and indeed, a large number were early identified on the basis of molecular weight, sarcomere position, and antibody reactivity (316, 455, 456, 544, 545, 639, 698, 730–732, 757, 845, 850, 851, 884, 1014, 1016, 1021, 1033–1035, 1037, 1218, 1220, 1221, 1228, 1343). It is now clear that these proteins belong to multiple isoform families of wide size range, that gel electrophoresis molecular weight estimates for very large proteins can be highly inaccurate (323), and that many of these proteins, even those from different families, are sufficiently similar to immunologically cross-react. Gene sequencing in *C. elegans* and *Drosophila* has begun to provide order to this field. However,
in older literature terms such as “mini-titin,” “titin/twitchin,” “titinlike,” etc., are widespread, and what protein is being described is not always clear. Even worse, vertebrate titin is also called connectin, and some authors refer to all large (>2 MDa) invertebrate sarcomere proteins as connectin (756), even though sequence data show that some of these large proteins are very different. Moreover, in an extremely unfortunate happenstance, a *Drosophila* cell adhesion molecule present on embryonic muscles is called connectin (870–872). Readers must therefore exercise caution in the older literature. Reviews covering invertebrate sarcomere giant proteins (at least in part) include References 95, 98, 142, 144, 544, 756, 758, 1184, 1185, 1190, 1343, 1345.

To provide order to these proteins, it is helpful to first examine human skeletal muscle titin (Figs. 7A and 8) (182). The protein’s first 80 kDa, with its multiple immunoglobulin domain repeats, binds to the Z disk and possibly actin. Additional immunoglobulin repeats (which can provide elasticity by unfolding under tension, Ref. 649) and an elastic PEVK region occur next and span the adjacent half I band in which the molecule moves toward the thick filaments. Next is a large region of a repeated immunoglobulin/fibronectin domain motif, which is associated with the thick filaments in the A band. The molecule ends with a 250-kDa M-line region that contains a serine/threonine kinase activity and multiple immunoglobulin repeats. Thus, in addition to being huge (one basis for calling various invertebrate giant proteins titin-like), titin also has regions in which only or primarily immunoglobulin repeats exist, areas expected to act as springs, a region with a repeated immunoglobulin/fibronectin motif, and a kinase (and hence a wide range of epitopes against which antibodies can be raised). A large variety of proteins could thus resemble one or another portion of titin but still be distinctly different from one another. Titin is believed to fulfill four functions (379, 755, 1186). First, its springlike nature helps keep the thick filaments centered in the sarcomere. Second, this same nature provides much of muscle passive resistance to stretch and elasticity. Third, it may help determine sarcomere length in development. Fourth, although its
target(s) is unknown, titin’s kinase could regulate the activity of other sarcomere proteins.

Except for amphioxus (455) and ascidia (845), in which the proteins are presumably vertebrate titin, no known invertebrate giant protein is long enough to span an entire half sarcomere (566). Proteins in the 3 MDa range (similar to titin’s molecular weight) have been identified by gel electrophoresis in some (but are clearly absent from other) crayfish muscles (732) and in the 5 MDa range in barnacle (730), but these lengths are insufficient to span half the large (10 μm) sarcomeres present in them. Very large (4-MDa) proteins that connect the Z line to the thick filaments are present in the annelid Neanthes, but how far along them the proteins extend is unknown (545). All other invertebrate giant proteins are also too small to span a half sarcomere. Present best knowledge is thus that invertebrate elasticity and sarcomere structure arise from proteins that connect Z line and thick filament proteins, but which do not extend all the way to the M line.

1. Crayfish connectin/kettin, C. elegans kettin, and Drosophila titin/kettin

Genes capable of producing 1) ~2-MDa proteins (called connectin in crayfish and titin in Drosophila) with an immunoglobulin-repeat-rich region and springlike PEVK regions but no protein kinase domain, and 2) ~550-kDa proteins (called kettin in all species) consisting of only immunoglobulin domains, have been completely or partially sequenced in C. elegans (144, 391), Drosophila (144, 587, 640, 708, 1337), Bombyx mori (584, 1123), and crayfish (323). The reported Drosophila sequences show some variation. The sequence predicted from the Drosophila genome project (144) lacks eight NH2-terminal immunoglobulin repeats (left rectangle, below sequence) present in the sequence of Reference 708, and has four COOH-terminal fibronectin repeats (right rectangle) not present in Reference 708. Drosophila and crayfish kettin are produced by alternative splicing from the titin/connectin gene. C. elegans kettin is instead produced by its own gene (see below for a large C. elegans sarcomere.
protein). In *Bombbyx*, kettin is also produced by its own, single-copy gene, and there are two titinlike genes. Given the presence in *C. elegans* (see below) of another giant sarcomere protein that is distantly, if at all, related to the crayfish connectin/insect titin/C. elegans-Drosophila-crabyfish kettin protein family, but which is nonetheless named titin, the nomenclature in this field clearly needs revision. Reference 144 suggests that the crayfish connectin/insect titin/C. elegans-Drosophila-crabyfish kettin proteins be all renamed SLS proteins (from saltimuses, the gene that codes for *Drosophila* titin/kettin), with kettin becoming just an SLS isoform.

Crayfish connectin runs from the Z line to the tips of the thick filaments and stretches with the sarcomere to lengths up to 3.5 μm (323). Crayfish connectin contains both a PEVK region and a SEK region, which is also elastic (322), and connectin proteolysis in skinned crayfish fibers abolishes resting tension (732). Antibodies to an epitope on crayfish connectin that is not present in fish fibers abolishes resting tension (732). Antibodies to elastic (322), and connectin proteolysis in skinned crayfish fibers abolish resting tension (732). Antibodies to an epitope on crayfish connectin that is not present in *Drosophila* titin recognize a large protein in barnacle ventral muscle (323), *Calliphora* larval muscle (323), and beetle (*Allomyrina*), bumblebee (*Bombus*), and waterbug (*Lethocerus*) leg muscle (323, 884). The beetle, bumblebee, and waterbug proteins are recognized by antibodies to vertebrate titin and is localized to the I band in beetle (884). The crayfish connectin antibody also recognized a 3-MDa protein in waterbug, but not beetle or bumblebee, flight muscle (884). A 540-kDa protein that localized to Z lines, elongated upon muscle stretch, bound actin, and was recognized by antibodies against vertebrate titin, and a 500-kDa *Lethocerus* muscle protein was early isolated from crayfish claw muscle (731). Given the cloning evidence for kettin in crayfish (323), this protein is presumably kettin.

*C. elegans* kettin has an immunological staining pattern consistent with dense body (equivalent to Z line) expression, all muscles were stained (144, 587). *Drosophila* kettin is present in all larval body wall and visceral muscles and in adult leg and flight muscles. Mutants of what was believed to be the kettin gene have abnormal muscles; homozygotes cannot hatch and heterozygotes cannot fly (391). However, in light of titin and kettin arising from one gene, it is now unclear if these effects are due to a lack of kettin or titin. Kettin is located at the Z line in flight muscle, with the NH₂ terminus in the line and the COOH terminus outside (140). Kettin binds to actin filaments (1205) and in vitro promotes anti-parallel association of thin filaments (the same arrangement as that present in vivo, see Fig. 5) (1205). Kettin’s immunoglobulin domains bind α-actinin and actin (640) and are consistent with a model in which each repeat and its flanking region bind to one thin filament actin monomer, each immunoglobulin/linker binding ~3 nm of the thin filament (587, 1205). With this model, each sarcomere’s kettin molecules would overlap by a small amount and extend some 60 nm outside the 80- to 120-nm Z-line width (587). Kettin competes with tropomyosin for actin binding, which could exclude tropomyosin from the Z line (1205). Kettin’s COOH terminus can bind myosin; kettin could thus link the Z line to the thick filaments (621) and is therefore believed to play a major role in muscle stiffness (140, 621). However, it also has large numbers of immunoglobulin domains, and single kettin molecules respond to tension with steplike relaxations presumably stemming from unfolding of individual immunoglobulin domains (649). Kettin may therefore also play a role in muscle elasticity.

*Drosophila* titin is required for myoblast fusion and the formation of striated muscle sarcomeres (708, 1337). Titin also plays an important role in muscle elasticity. Although immunoglobulin/fibronectin unfolding presumably can play a role in this process, it is not its only source, since muscle length changes in 2.3 nm quanta when force is applied to indirect flight muscles under conditions in which titin is the sole length-absorbing element. This distance is much less than that which would occur from immunoglobulin or fibronectin domain unfolding and may correspond to single beta-sheet unfolding (123). Multiple titin isoforms (in addition to kettin) are expressed with some muscle specificity (144). *Drosophila* titin is also present on condensed chromosomes and required for proper chromosome structure and function (707, 708). A titin-like protein is present in *Anopheles* (144).

2. *C. elegans* titin

The *C. elegans* titin gene produces a 2,200-kDa protein with multiple immunoglobulin and fibronectin repeats, an elastic region, and a kinase; a 1,200-kDa protein with the immunoglobulin and fibronectin repeats and elastic region, but no kinase; and a 301-kDa protein consisting of only the kinase and the COOH-terminal (primarily) immunoglobulin repeats (306). Immunohistochemistry shows that the 2,200- and 1,200-kDa isoforms are expressed in all muscles except in the pharynx. The 2,200- and 1,200-kDa proteins bind at their NH₂ terminal to the dense bodies (Z-line equivalents). The 2,200-kDa protein may be long enough to reach slightly into the A band, and thus its kinase could act on thick filament proteins. The position of the 301-kDa protein is unknown. The gene is not expressed until after *C. elegans* muscles are formed. No *C. elegans* titin mutants are available, and thus the functions of these proteins are unknown, as are the kinase’s targets. A survey of the *C. elegans* genome searching for immunoglobulin containing proteins identified twitchin, *unc-89*, *unc-52*, and several titinlike sequences (1167).
3. Unidentified very large sarcomere proteins

Barnacle has 5,000- and 1,200-kDa proteins that monoclonal antibodies to vertebrate titin recognize; the 5,000-kDa protein links the thick filament to the Z line (730). This could be a crustacean equivalent to \textit{C. elegans} titin, but sequence data will be required to confirm this relationship. Antibodies to vertebrate titin recognize a >2 MDa scallop protein and label scallop thick filaments along their entire length and most strongly at their tips (1221). A similarly large titinlike (by antibody binding) protein is present in squid (542). These data indicate that mollusks contain a very large sarcomere protein, but its relationship to other invertebrate muscle proteins is unknown.

4. Projectin/twitchin

A family of 600- to 1,000-kDa proteins with repeating fibronectin/immunoglobulin repeats and a serine/threonine kinase domain exists in mollusks, \textit{C. elegans}, crustaceans, and insects. The protein is called twitchin in \textit{C. elegans}, twitchin or twitchinlike in mollusks, and projectin in crustaceans and insects.

Twitchin (\textit{unc-22}) was identified from \textit{C. elegans} mutants with a constant twitch and abnormal body wall muscle sarcomeres (131, 678, 679, 809, 811, 813, 1003). Twitchin consists of several immunoglobulin repeats, multiple repeats of an immunoglobulin/fibronectin motif, a serine/threonine kinase, and several immunoglobulin repeats (96, 97, 314, 810). Up to six fibronectin/immunoglobulin repeats can be deleted without strong phenotypic effects (562). Bacterially expressed twitchin kinase autophosphorylates (664). The 60 residues COOH terminal to the kinase core inhibit kinase activity (664) by binding to the active site (intrasteric regulation) (457, 458, 578, 579). The Ca\textsuperscript{2+} binding protein S100A1 relieves the inhibition and increases kinase activity 10-fold (413, 507). Although several other Ca\textsuperscript{2+} binding proteins, including calmodulin, bind to the autoinhibitory sequence (417), they do not relieve the inhibition (664). Twitchin staining is observed at both the A band and the dense body, suggesting it links the thick filaments to the dense body (767, 810). There are likely at least two twitchin genes, as \textit{unc-22} mutants remove staining from body wall, anal, and vulval muscles, but not pharynx (810). Mutations in \textit{unc-54}, which codes for myosin heavy chain B, can suppress the \textit{unc-22} phenotype; the \textit{unc-54} mutants in isolation have normal muscle structure, but the animals are stiff (48, 329). Heterozygote muscles show reduced thick filament length (590).

Striated and smooth scallop muscle and mussel smooth catch muscle contains a 600–800 kDa (0.2 \textmu m) rod-shaped protein with a kinase domain (1220). Antibodies against the protein stain \textit{Limulus}, \textit{Lethocerus} (indirect flight), crayfish, and chicken muscle and bind to the A/I junction in all except \textit{Lethocerus}, in which they bind to the I band. Antibodies raised to the immunoglobulin repeat or kinase domains of \textit{C. elegans} twitchin recognize the protein. Although both \textit{C. elegans} twitchin and titin have immunoglobulin repeats and a kinase, the protein's molecular weight suggests it is a projectin/twitchin.

Immunohistochemistry, immunoelectron microscopy, and partial amino acid and cDNA sequencing indicate that \textit{Aplysia californica} has a twitchin-related protein that colocalizes with contractile filaments (416, 822, 954). Like \textit{C. elegans} twitchin, \textit{Aplysia} twitchin autophosphorylates (416, 822), has an autoinhibitory sequence that binds Ca\textsuperscript{2+}/calmodulin without activating the kinase (137, 413, 414, 416, 417), and is activated (in this case, >1,000-fold) by Ca\textsuperscript{2+}/S100A1 binding (413, 414, 417). However, \textit{Aplysia} twitchin phosphorylates regulatory myosin light chains well (415), whereas \textit{C. elegans} twitchin only poorly phosphorylates peptides containing the expected nematode regulatory myosin light-chain phosphorylation site (98, 664). \textit{C. elegans} and \textit{Aplysia} twitchin also differ with respect to Ca\textsuperscript{2+}/calmodulin affinity, sensitivity to naphthalene sulfonamide inhibitors, and the effect of Zn\textsuperscript{2+} on Ca\textsuperscript{2+}/S100 activation (required in \textit{Aplysia}, inhibitory in \textit{C. elegans}) (98, 414, 417). Studies on neuropeptide modulation of \textit{Aplysia} muscle relaxation indicate that cAMP-dependent protein kinase phosphorylates \textit{Aplysia} twitchin, suggesting that twitchin may regulate muscle relaxation (954), consistent with \textit{unc-22}'s phenotypic effects (809). \textit{Aplysia} twitchin's sarcomere location is unknown. Phosphorylation of a \textit{Mytilus} twitchin-like protein is involved in catch (see second review) (151, 152, 324, 326, 1088, 1089, 1298).

Early work with monoclonal antibodies against \textit{Lethocerus} muscle proteins revealed an 800-kDa protein (630). A \textit{Drosophila} protein, projectin, that is believed to be the homolog of this \textit{Lethocerus} protein, has been sequenced (47–49, 232, 329, 1107). Projectin is very similar to \textit{C. elegans} twitchin (49), and projectin antibodies cross-react with \textit{C. elegans} twitchin (1228). Several isoforms, all of which retain the kinase and which show muscle specific expression (1082, 1228), arise by alternative splicing (48, 232, 1107). The \textit{bent} locus codes \textit{Drosophila} projectin. Homozygous embryos do not hatch, and in the most severe allele no muscle contraction ever occurs (48, 329). Heterozygote muscles show reduced stretch activation (821, 1226). Projectin, paramyosin, and myosin aggregates form different diameter fibers depending on projectin's phosphorylation state (287). In locust, projectin (along with paramyosin) may help determine thick filament length (590). \textit{Drosophila} projectin binds to thin filaments (1271), and locust projectin blocks the formation in vitro of paracrystals from solutions of actin monomers, instead promoting thin filament formation (949). Locust projectin autophosphorylates (287, 1271) and phosphorylates 30-, 100-, 165-, and 200-kDa proteins (but not myosin light chain) in vitro (287, 1271). The 30-kDa protein may be troponin I, and troponin phosphor-
ylation by projectin increases troponin’s Ca$_2^+$ sensitivity (1271). Actomyosin, actin filaments, and myosin inhibit projectin autophosphorylation, and calmodulin stimulates it (287). Projectin phosphorylation of other proteins, however, is unaffected by actomyosin, actin filaments, myosin, calmodulin, Ca$_2^+$, Ca$_2^+$ and calmodulin, or cAMP (287, 1271). Calmodulin-dependent protein kinase II (CaMKII) binds to projectin, which increases CaMKII activity twofold (286).

In synchronous muscle, projectin is located in the A band similar to twitchin (47, 639, 1035, 1228). In locust, the molecule extends into the I band, and the molecule’s COOH terminus, with the kinase domain, is located in the A band (1082). In asynchronous muscle, it is instead located in the I band and forms filaments connecting the Z line and thick filament (47, 639, 1033, 1035, 1228). Neither myosin nor actin is required for projectin’s initial localization to the Z line during sarcomere assembly, but in the absence of actin projectin’s Z line localization is later lost (45). Proteins believed to be projectin have been isolated from locust, bee, beetle, and *Lethocerus* (142, 884, 1034), and staining with an antibody likely to have been raised against *Drosophila* projectin stains (as expected) *Drosophila* cross-striated muscle and annelid (*Eisenia*) obliquely striated and smooth muscle (1014).

A large sarcomere protein with immunological similarity to vertebrate titin and honeybee, beetle, and *Drosophila* projectin that contains a serine/threonine kinase and autophosphorylates was early isolated from crayfish (456, 745). This protein, crayfish projectin, has been cloned, and flexor and closer muscles express different isoforms (906).

5. *C. elegans* unc-89

The *unc-89* mutants have disorganized muscle structure and lack M lines (95, 1266). The *unc-89* gene has been cloned and is believed to be a homolog of vertebrate obscurin (98, 99, 1097, 1122). The gene produces four isoforms that are differentially expressed in different muscles. Two of the isoforms contain extensive immunoglobulin repeats. One of these isoforms also contains two protein kinase domains, although only one is believed to be functional. The other two isoforms are short versions containing only the tandem protein kinase domains. All the isoforms are localized in the center of the A band. The structure of a region of the molecule that may be involved in protein-protein interactions has been determined in detail (120, 121). *unc-98*, a protein containing four Zn finger motifs, is also located at the M line in *C. elegans* (787). A 400-kDa M-line protein has been identified in *Lethocerus* (142, 639), but without sequence data the relationship of this protein and the *unc-89* product is unknown.

F. Miscellaneous Other Proteins

1. Nebulin

Nebulin-like proteins are present in obliquely striated *Eisenia* muscle (1015) and cross-striated muscle in Antarctic mussel shrimp (1016), Antarctic mite *Halacarellus* (1021), lancelet *Branchiostoma* (cephalochordate) (313), and hagfish skeletal and heart muscle (312). Nebulin-sized proteins are present in echinoderms, an annelid, mollusks, crustaceans, and insects (698). In *Eisenia* and *Halacarellus*, anti-nebulin staining occurs along the entire sarcomere, whereas in *Branchiostoma* staining is only in the I band.

2. Nesprin

Vertebrate nesprins are giant (1 MDa) actin binding (and nuclear) proteins with extensive spectrin-like repeats (1335). Spectrin-related *Drosophila* proteins of 200–300 kDa molecular mass, located at muscle adhesion sites in early development and later at the Z line and which were originally believed to be possible dystrophin orthologs, were early identified (1007, 1233). It now appears these proteins arise from a nesprin gene ortholog that could produce isoforms as large as 11,720 amino acids (1335). A similar gene has been identified in *C. elegans* (1335).

3. Amphiphysin

Vertebrates have two amphiphysins, one involved in endocytosis and the other in striated muscle t tubules. *Drosophila* has a single amphiphysin gene whose product is involved only in muscle, as mutants have normal synaptic transmission but severely disorganized t tubule/sarcoplasmic reticulum systems (973, 1332) and defective membrane protein integration in the postsynaptic muscle membrane (761, 1327, 1332).

4. Various muscle cell attachment proteins

To transmit force to effectors, the thin filaments of the terminal sarcomeres must attach to the plasma membrane, and thence to tendons or their equivalent. This attachment is accomplished by a protein complex including integrin, talin, vinculin, and filamen. Several of these proteins, prominently integrin, also control the complex’s development. This complex is not fully understood in vertebrates and is less so in invertebrates. We here therefore simply list a number of proteins that attach thin filaments to the muscle cell membrane, or muscle cells to tendons or body wall structures, including primarily references showing that they serve a structural role. These proteins include connectin (*Drosophila*) (963), DIM-1 (*C. elegans*) (1001), integrin and integrin-associated proteins (jellyfish (975), *C. elegans* (349, 432, 450, 691, 712, 787, 1004–1006, 1054), *Drosophila* (122, 124, 147, 183, 263, 666,
713, 854, 955, 1292, 1297, 1331), kakapo (Drosophila) (380, 1120), myotactin (C. elegans) (454), spondin (Drosophila) (1201), talin (C. elegans) (219, 825), jellyfish (975), Drosophila (136), and vinculin (C. elegans) (74, 75). References 117, 125, 132, 134, 135, 149, 154, 389, 390, 465, 633, 947, 1002, 1329, 1330 are reviews or papers describing muscle adhesion-deficient mutants or antibodies against cell adhesion molecules. Analysis of talin homologs from yeast, slime mold, and nematode reveals a conserved protein domain believed to be specialized for thin filament binding (772). Early monoclonal antibody work identified three to five C. elegans muscle attachment proteins (317); which, if any, of the above proteins correspond to these antibodies is unknown.

5. Dystrophin-related proteins

Mutations in the human dystrophin gene cause Duchenne muscular dystrophy. The dystrophin complex may also help transmit sarcomere force to the plasma membrane, or help maintain membrane integrity under the stress of sarcomere contraction (1031). The human dystrophin glycoprotein complex consists of dystrophin, dystrobrevin, distroglycan, sarcoglycan, sarcospan, syntrophin, and nitric oxide synthase (168). Given this close association, all these genes will be considered together here. The sea urchin dystrophin gene has been identified and shares with vertebrates a complex structure with multiple promoters and gene products (853, 1250). Immunohistochemistry and Western blot analysis show that leech (Pondtobdella) has dystrophin, dystroglycan, sarcoglycan, syntrophin, and sarcospan homologs (1020, 1022). C. elegans (113, 352, 382, 384) and Drosophila (243, 377, 853) have dystrophin, dystrobrevin, dystroglycan, sarcoglycan, and syntrophin (but not sarcospan) homologs. The C. elegans and Drosophila dystrophin genes again have multiple promoters and gene products (113, 853). C. elegans muscle has a nitric oxide synthase binding protein that can partially suppress dystrophin mutation (354).

The C. elegans proteins are also likely a complex, as dystrophin (113), dystrobrevin (352, 355), dystroglycan (384), sarcoglycan (384), and syntrophin (382, 384) mutants all have the same phenotype: hyperactivity, head bending during forward locomotion, a tendency to hypercontract, and ACh hypersensitivity. Further genetic evidence of interaction is provided by data that increased dystrobrevin expression delays muscle degeneration and locomotor defects in dystrophin mutants (353). In vitro binding assays give direct evidence of dystrophin, dystrobrevin, and syntrophin interaction and show that dystrophin/dystrobrevin binding requires the second coiled-coil dystrobrevin domain (351, 383). The basis of the mutant phenotype is not well understood, but acetylcholinesterase activity is decreased (by an unknown mechanism) in dystrophin mutants, which may explain the ACh hypersensitivity (356). Ca\(^{2+}\) channel activity is important for phenotype expression, as increased Ca\(^{2+}\) channel activity increases, and reduced Ca\(^{2+}\) channel activity decreases, muscle degeneration in dystrophin null mutants (742). References 168, 186, 226, 999, 1066, and 1067 are reviews of C. elegans and Drosophila as model systems for studying Duchenne muscular dystrophy.

6. Intermediate filaments

In vertebrates another protein that may help transmit force from the sarcomere to the cell membrane, or help maintain muscle cell integrity in the face of the stress involved, is the intermediate filament protein desmin. In mammals desmin is one of a large number of intermediate filament types (others include keratin, neurofilament protein, and nuclear lamins) that can be ordered into five or six classes on the basis of sequence homology (9, 319 are primarily vertebrate reviews). Invertebrate intermediate filaments are much less well understood, and their relationship to the mammalian classes remains unclear (256, 345, 530, 994, 996, 997, 1249, 1267, 1268). We cover here only work directly involving invertebrate muscle intermediate filaments.

Early work showed surprising variation with respect to the presence of intermediate filaments in invertebrate muscles, even between closely related species. For instance, large amounts of intermediate filaments were present in all muscle types (smooth, irregularly, and obliquely striated) in Ascaris, located throughout the sarcomere and around dense bodies, but not in C. elegans muscle (80–82). A wide survey of major phylogenetic groups showed that intermediate filaments were also present in Acanthocephala, Echiura, and Chaetognatha muscle, but absent from the other examined invertebrate groups, including Urochordata and Cephalochordata (78, 81). It was subsequently shown that some invertebrate intermediate filament proteins lack an epitope present in all vertebrate intermediate filament proteins, and thus that lack of muscle staining in some groups may be due to the lack of this epitope rather than a lack of intermediate filaments (995). Particularly in light of later work showing muscle intermediate filaments are present in some of the above “negative” groups (see below), the absence assignments made in this early work should be viewed with caution.

Molecular genetic work has shown that Styela (a urochordate) has two intermediate filament type genes, one of which is expressed in smooth body wall muscle and shows sequence similarity to desmin (997). Similarly, one of the five intermediate filament protein genes in Ciona (another urochordate) is expressed in muscle (535, 1249). Despite some confusion (994), it is now also clear that in Branchiostoma (Cephalochordata) 1 of the 10 known (530) intermediate filament proteins is present in

Physiol Rev • VOL 85 • JULY 2005 • www.prv.org
and vertebrate, ascidian, and invertebrate monophyly. The second indicates that the absence of some does cause displaced muscles and paralysis (364, 398, 531–534). Staining is also present in the muscles of several platyhelminths (Echinococcus, Schistosoma, Taenia) (1046).

Intermediate filaments composed of a protein immunologically similar to desmin are also present in several, but not all, annelids (79, 81, 239). An intermediate filament protein gene identified using a tissue sample containing body muscle and attached epidermis has been sequenced in Lumbricus terrestris, but it is unclear if this gene codes the above protein (126). Proteins that form 2- to 4-nm filaments, but which are not similar to desmin or vimentin by two-dimensional gel electrophoresis or immunological cross-reactivity, are present in sea urchin muscle (987). Arthropods appear to lack all intermediate filaments (except nuclear lamins) (81), an absence confirmed by the lack of intermediate filament genes in the Drosophila genome (365, 1026).

7. The 29-kDa ascidian protein

A body wall protein from Halocynthia roretzi of unknown function that copurifies with thin filaments is localized near the plasma membrane, has some similarity to heat shock proteins and α-cystallin, and has been isolated and sequenced (1084, 1151).

G. Summary

The most striking impression from the above is the large number and variety of different proteins, and different isoforms of individual proteins, present across invertebrates. This variety provides two great opportunities. First, because these proteins interact and thus must have evolved together, comparison of concerted changes in their sequences should provide powerful evidence with respect to evolutionary relationships. In addition to the studies on actin mentioned earlier, this work has been begun with sequence comparisons of myosin regulatory and essential light chain and troponin C (187); troponin C, myosin heavy chain, myosin regulatory and essential light chain, and actin (903); vertebrate, C. elegans, and Drosophila titin, twitchin, projectin, and unc-89 (427, 554); and vertebrate, ascidian, and Drosophila troponin I (407). The first work is in agreement with the phyla and lower assignments in Figures 1–3 but does not resolve the question of invertebrate monophyly. The second indicates that skeletal and cardiac muscle type tissues evolved before the vertebrate/arthropod split, but vertebrate smooth muscle appears to have arisen independently of other muscle types, and arthropod striated, urochordate smooth, and vertebrate muscle (except for smooth) share a common ancestry. The third suggests that titin/kettin, twitchin, projectin, and myosin light chain kinases form a related group of which unc-89 is not a member, but it was performed with too few invertebrate species to provide insight into invertebrate phylogenetic relationships. The fourth shows that a single troponin I gene (from which multiple isoforms are produced by alternative splicing), such as is present in ascidia and Drosophila, is the ancestral condition, and the gene duplications leading to the three troponin I genes present in vertebrates occurred after the ascidian/vertebrate split. Although these early observations are intriguing, much further work is necessary for the phylogenetic opportunities present here to be fully realized. Notable lacks are investigation of muscle proteins in Cnidaria, in which so far only actin, myosin, and tropomyosin have been shown to exist, and detailed study of giant muscle proteins in Lophotrochozoa.

The second opportunity the variety provides is in understanding protein ensembles. Despite their variety, the different sets of proteins present in the various invertebrates must successfully perform relatively similar functions (form thin and thick filaments and sarcomeres, contract and relax on demand). Investigating how the multiplicity of forms noted above continue to do so, and the relationship between these forms and differences in muscle structure and function, will presumably provide great insight into fundamental questions of protein interaction and function that can be successfully applied to other systems. The utility of this approach is well demonstrated by a recent analysis of myosin heavy and light chains, which showed that head, neck, and tail domains of the heavy chains, and the essential and regulatory light chains, show very similar phylogenetic trees (600). This work is beginning in Drosophila, in which large numbers of myofibril mutants are available (332, 815, 869, 1223), and null mutants of actin and myosin have been used to examine the contributions of thin and thick filaments to sarcomere assembly (89). Similar genetic opportunities are available in C. elegans, as is also the ability to use antisense RNA to inhibit expressions of specific genes (301). The increasing availability of cloned genes in ascidia should allow similar work to proceed in this “higher” invertebrate (174). The development of robotic approaches for simultaneously identifying the components
of multiprotein mixtures (proteomes) should dramatically enhance these efforts (36).

It is critical to point out, however, that to date most work has been performed in a small number of “model” systems (Drosophila, C. elegans, Bivalvia, decapod crustacea), organisms heavily studied due to their disease or economic importance (parasitic platyhelminths and nematodes), or organisms relatively closely related to vertebrates (Cephalochordata, Urochordata, Echinodermata). Although this bias was understandable in the early days in which the challenge was defining the sarcomeric muscle protein complement, and protein and gene analysis and sequencing tools were relatively primitive, neither justification exists today. These concerns about relatively small sequencing tools were relatively primitive, neither justifiable nor scientifically valid.

This work would not have been possible without, and we express our deep gratitude to, C. A. Barcroft, R. P. Harrison, R. R. Heck III, K. H. Hobbs, S. McKeen, B. Revill, J. Thuma, and R. Thuma for help in creating and maintaining the reference database and to the Ohio University library system and OHIOLINK electronic journal center for providing several thousand journal article photocopies and PDFs. This work was supported by grants from the National Institutes of Health and the University of Cologne, Cologne, Germany (to S. L. Hooper). Address for reprint requests and other correspondence: S. L. Hooper, Neuroscience Program, Dept. of Biological Sciences, Irvine Hall, Ohio University, Athens, OH 45701 (E-mail: hooper@ohio.edu).

REFERENCES


Intermediate filaments in the giant muscle cells of nematodes. 

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Evidence for two intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filaments in muscle and epithelial cells of nematodes. 

Intermediate filaments in muscle and epithelial cells of nematodes. 

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.

Intermediate filament prototypes in the invertebrate Myxicola; neurofilaments and non-neuronal intermediate filaments differ in subunit size and immunological properties.
Bloor JW and Brown NH.


Fyrborg EA, Bond BJ, Hershey ND, Mixter KS, and Davidson N. The actin genes of *Drosophila*: protein coding regions are


336. Fyrberg EA, Fyrberg CC, Biggs JR, Saville D, Beall CJ, and Kindle KL. The role of microscopy in the investigation of paramyosin as a vaccine candidate against Schistosoma japoni-


341. Go ´mez-Guille ´n MC, Solas MT, Borderı ´as AJ, and Montero P. Salt, non-


344. Gobert GN. Functional nonequivalence of Ketchum A.


354. Gieseler K, Grisoni K, Mariol MC, and Segalat L. Molecular, genetic, and physio-


356. Gobert GN. Immunolocalization of schistosome proteins. Mi-

357. Gobert GN. The role of microscopy in the investigation of paramyosin as a vaccine candidate against Schistosoma japoni-


Guimarães PM, Leal-Bertolì SCm, Curtis RH, Davis EL, and Bertoli DJ. Isolation of two cDNAs encoding a tropomyosin and an intermediate filament protein from the soybean cyst nematode Heterodera glycines. Nematropica 33: 87–95, 2003.


Hasegawa Y. Complete nucleotide sequence of a cDNA encoding a myosin heavy chain from mantle tissue of scallop Patinopecten yessoensis. Fish Sci 68: 403–415, 2002.


Herranz R, Diiz-Castillo C, Nguyen TP, Lovato TL, Cripps RM, and Marco R. Expression patterns of the whole tropoion C...


558. Khaitlina SY, Tskhovrebova LA, and Sheludko NS. *Khaitlina SY, Tskhovrebova LA, and Sheludko NS.*

559. Kissinger JC and Raff RA. *Kissinger JC and Raff RA.*


Krause M, Wild M, Rosenzweig B, and Hirsh D.

Kretsinger RH.

Korn ED.

Konno K, Kodama S, Araik K, and Watanabe S.

Kretsinger RH, Rudnick SE, Sneden DA, and Schatz VB.

Kretsinger RH.

Kovesdi I and Smith MJ.


Kronert WA, Acebes A, Ferrus A, and Bernstein SI.


high molecular weight proteins in insect flight and leg muscle. 


INVERTEBRATE MUSCLE GENES AND PROTEINS


Petrova TV, Comte M, Takagi T, and Cox JA. Thermodynamic and molecular properties of the interaction between Amphioxus...


1040. Schmitz S, Schankin CJ, Prinz H, Curwen RS, Ashton PD, Caves LSD, Fink RHA, Sparrow JC, Mayhew PJ, and Veigel...
Shanti KN, Martin BM, Nagpal S, Metcalfe DD, and Rao PV. 1074.
INVERTEBRATE MUSCLE GENES AND PROTEINS 1055


Tsuchiya T and Ehara T.

Uddin M, Ahmad MU, Jahan P, and Sanguandeekul R.

Tzolovsky G, Millo H, Pathirana S, Wood T, and Bownes M.


Valette-Talbi L, Comte M, Chaponnier C, and Cox JA.

Vargas-Parada L and Laclette JP.


Vargas-Parada L and Laclette JP.

Vasquez-Talavera J, Solis CF, Terrazas LI, and Laclette JP.

Vasquez-Talavera J, Solis CF, Terraizs LA, and Laclette JP.


Weber K, Plessmann U, and Ulrich W. Cytoplasmic intermediate filament proteins of invertebrates are closer to nuclear lamins than are vertebrate intermediate filament proteins; sequence characterization of two muscle proteins of a nematode. EMBO J 8: 3221–3227, 1989.


