Assemblases and Coupling Proteins in Thick Filament Assembly

Feizhou Liu¹, Jose M. Barral², Christopher C. Bauer¹, Irving Ortiz¹, Richard G. Cook³, Michael F. Schmid², and Henry F. Epstein¹,²,⁴

¹Departments of Neurology, ²Biochemistry, and ³Microbiology and Immunology, Baylor College of Medicine, Houston, TX 77030, USA

Key words: directed assembly/myosin/paramyosin/novel thick filament proteins

ABSTRACT. Thick filaments are stable assemblies of myosin that are characteristic of specific muscle types from both vertebrates and invertebrates. In general, their structure and assembly require remarkably precise determination of lengths and diameters, structural differentiation and nonequivalence of myosins, a high degree of inelasticity and rigidity, and dynamic regulation of assembly and disassembly in response to both extracellular and intracellular signals. Directed assembly of myosin in which additional proteins function in key roles, therefore, is more likely to be significant than the simple self assembly of myosin into thick filaments. The nematode Caenorhabditis elegans permits a wide spectrum of biochemical, genetic, molecular and structural approaches to be applied to the experimental testing of this hypothesis. Biochemical analysis of C. elegans thick filaments reveals that paramyosin, a homologue of the myosin rod that is the unique product of a single genetic locus, exists as two populations which differ by post-translational modification. The major paramyosin species interacts with the two genetically specified myosin heavy chain isoforms. The minor paramyosin species is organized within the cores of the thick filaments, where it is associated stoichiometrically with three recently identified proteins P20, P28 and P30. These proteins have now been characterized molecularly and contain unique, novel amino acid sequences. Structural analysis of the core shows that seven paramyosin subfilaments are crosslinked by additional internal proteins into a highly rigid tubule. P20, P28 and P30 are proposed to couple the paramyosin subfilaments together into the core tube during filament assembly. Mutants that affect paramyosin assembly are being characterized for alterations in the core proteins. A fourth protein has been identified recently as the product of the unc-45 gene. Computational analysis of this gene’s DNA suggests that the predicted protein may exhibit protein phosphatase and chaperone activities. Genetic analysis shows that three classes of specific unc-45 mutant proteins differentially interact with the two myosins during thick filament assembly. The unc-45 protein is proposed to be a myosin assemblase, a protein catalyst of thick filament assembly.

The assembly of myosin into thick filaments has served as a paradigm for molecular self-assembly generally (25) and for molecular mechanisms of myofibril assembly more specifically (12). The precise construction of the resulting structures and their organization into highly ordered arrays within muscle cells contrasts greatly with the results of cell-free experiments using purified components (21). The paradigm of self-assembly (Fig. 1) must be replaced by alternative explanations. The model of directed assembly was first implicated in bacteriophage morphogenesis, in which specific proteins may catalyze the assembly of other proteins or serve as molecular rulers for the polymerization of other protein components (27).

In vertebrates, at least eight additional proteins have been shown to be associated with thick filaments: C-protein, H-protein (mammals) or 86 kDa protein (birds), M-protein, myomesin, M-creatine kinase, adenosine monophosphate deaminase, skelemin, and titin (12). C-protein and its isoforms H-protein and 86 kDa protein appear to be intrinsic thick filament components, whereas the other proteins appear to be interacting with the surfaces or protruding cross-bridges of the thick filaments. Titin possesses a catalytically active myosin light chain kinase-like domain in addition to its many immunoglobulin C-2 like and fibronectin III-like structural domains (22). Vertebrate thick filaments are associated with several enzymes as well as purely structural proteins. However, it is not clear which of these proteins are necessary for the proper assembly of thick filaments or what roles they play.

In the nematode Caenorhabditis elegans and other invertebrates, paramyosin and not myosin is the prominent protein of thick filament cores (38). Recent analy-
sis of the amino acid sequence of paramyosin shows that it is really an isoform of the rod domains of myosin heavy chains (26). Twitchin, encoded by the unc-22 locus in C. elegans, is a close relative of vertebrate titin with myosin light chain kinase, immunoglobulin, and fibronectin-type domains. The unc-89 locus encodes an isoform of twitchin which appears related to the M-line and associated structures (3). Specific mutants of paramyosin, twitchin, and unc-89 disrupt myofibril assembly in C. elegans body wall muscles.

The genetic approach afforded by C. elegans, Drosophila, and potentially mice is likely to be necessary in addition to molecular and structural analyses for the identification and characterization of additional proteins active in thick filament assembly. In C. elegans, the myo-3 and unc-54 encode the two body wall myosin heavy chains; the unc-15 locus encodes paramyosin; and, as mentioned, the unc-22 and unc-89 loci encode twitchin and its M-line associated isoform (3). Other genes have been shown to encode the myosin light chains. Two other loci, unc-82, and unc-45 have been implicated in thick filament assembly but are clearly not myosin genes (Fig. 2). Specific mutants in unc-82 show functional interaction at the level of thick filament assembly, or epistasis with specific alleles of paramyosin (44). Similarly, specific mutants in unc-45 show epistatic interactions with specific mutants of myo-3 and of unc-54 myosin heavy chains (39). These epistatic interactions suggest that either the proteins encoded by unc-45 and unc-82, or their actions, directly interact with myosin and paramyosin during assembly. A more complex mode in which additional proteins connect the actions of unc-45 and unc-82 with the assembly of myosin and paramyosin cannot be excluded. In either case, the action of additional proteins is necessary for myosin assembly into the thick filaments of C. elegans.

In this report, we will describe biochemical and structural studies of C. elegans thick filaments which lead to conclusions consistent with their assembly being complex and directed. We present the identification and characterization of three novel proteins associated with paramyosin in the filament cores. Finally, we review the genetics of the unc-45 and unc-82 loci from the view-

![Fig. 1. Myosin Self-Assembly](image)

**Fig. 1.** Myosin Self-Assembly. Schematic diagram of the self-association of myosin molecules into nascent filaments. Note that the filaments are bipolar. The myosins of the medial region pack together in antiparallel; the myosins of the lateral regions pack together in parallel. No other types of protein are indicated.

**Fig. 2.** Thick Filament Genes. Genetic map of genetic loci encoding body-wall muscle myosin heavy chains (unc-54, myo-3) paramyosin (unc-15), putative core coupling proteins (P20, P28, P30) and putative myosin assemblase (unc-45) in C. elegans.

point of recent molecular and structural work. We propose that the unc-45 protein may function as a myosin assemblase, a protein catalyzing myosin assembly, and that the unc-82 protein may function as a paramyosin coupling protein, cross-linking paramyosin subfilaments to form the core.

**MATERIALS AND METHODS**

The procedures required for the experiments described below have been documented in previous reports most notably (1, 7, 8, 10, 11, 13, 15, 16, 30, 32).

The ACeDB computer database on C. elegans genetics, physical DNA mapping, and DNA sequencing was used extensively in analysis of the P20, P28, and P30 sequences and of the unc-45 sequence. Additional software for sequence analysis includes the Wisconsin GCG Peptide Structure, BLASTP, and TBLASTN programs.

**RESULTS**

**Commonality and Diversity in Thick Filament Structures.** Thick filaments of both vertebrate and invertebrate striated muscles appear to be bipolar structures with medial bare zones representing the interaction of myosin rods packed antiparallel to one another (25). The more lateral regions of myosins are packed in parallel and give rise to a helical arrangement of myosin heads (28, 37). The number of genetic helices, the exact repeat of myosin heads, and the diameters of thick filaments vary with the type of muscle.

The lengths of thick filaments within any muscle appear very homogeneous (Fig. 3). For example, the thick filaments in the A bands of rabbit psoas muscle, are
Directed Assembly in *C. elegans*

**Fig. 3.** Precision and Diversity in Thick Filament Assembly. Thick filaments from rabbit psoas, blowfly flight, and nematode body-wall muscle were isolated and purified (10). Note the homogeneity of length and diameter within any species and the differences between species. All of the filaments appear relatively inflexible with at most only slight curvatures. Bar is 1.0 μm. Reprinted from "Cellular and Molecular Biology of Development" ed. by L.H. Kedes and F.A. Stockdale, and with permission from Alan R. Liss, Inc.

15,500±200 Å long (6). The mean lengths vary between muscle types from 1.55 μm of rabbit psoas and other vertebrate striated muscles to the 9.7 μm of nematode body-wall muscle or even longer in some molluscan and crustacean muscles. One correlate of the variation in length of invertebrate thick filaments is the relative content of paramyosin with respect to myosin (29). Testing as to whether specific myosins and paramyosins are sufficient to explain the properties of their parent thick filaments has been difficult because of the great tendency of purified paramyosin to form paracrystals under the conditions favoring myosin polymerization (21).

**Fig. 4.** Differential Assembly of Myosin Isoforms. Schematic diagram of *myo-3* (myosin A) and *unc-54* (myosin B) proteins localized to the medial and lateral regions of *C. elegans* thick filaments, respectively. Compare to Figure 1.

**Myosin A & B Myosins A & B**

157 and R.H. Waterston, personal communication). It is clear that the two chemically and genetically distinct myosins assemble into nonequivalent positions. The nonequivalence is partially due to the intrinsic properties of the two myosins and partially due to differential interactions with at least one other protein, paramyosin.

An analogous situation exists in thick filaments of vertebrate striated muscles where the position of C and H or 86 kDa proteins produce nonequivalence in a single species of myosin (4). Nonequivalence of myosin has also been implicated in thick filaments from scallop striated adductor muscle (41).

**Nematode Thick Filaments are Complex Tubules.** The first indications that muscle thick filaments of *C. elegans* are tubular came from inspection of electron micrographs of cross-sections of positively stained and fixed nematodes. The thick filaments of the pharyngeal muscles resemble those of insect flight muscles in appearing highly electronlucent or hollow through most of their lengths (18). The body-wall thick filaments represent a potentially more complex situation in which the medial zones appear to have highly electron dense centers, the flanking zones less dense centers and the most polar zones hollow centers (14). These qualitative changes in central density correlate with quantitative decreases in diameter from the medial to polar zones.

Body-wall muscle thick filaments can be isolated and purified from *C. elegans* (7, 10, 30, 32). Careful titration of ionic strength reveals that 90% of the myosin (predominantly *unc-54*) and 70% of the paramyosin dissociate together from the thick filaments. Core structures containing 30% of the paramyosin and about 10% of the myosin (predominantly *myo-3*) remain (7). These results are consistent with the filaments being comprised of three coaxial tubular domains: cortical myosin, intermediate paramyosin, and paramyosin-containing core domains (Fig. 5). The paramyosin of the core and the more dissociable fraction show electrophoretic species distinguishable from one another in the isoelectric focusing dimension of two-dimensional gel separations (Fig. 6). Since there is only one paramyosin gene, RNA, and polypeptide in *C. elegans* (23, 30, 43), the most likely explanation for these differences is post-translational modification. Amino acid analysis has
showed that purified paramyosin is partially phosphorylated at three sites in its N terminal region (36). Cell-free experiments indicate that there are activities that readily phosphorylate these sites (8, 36). The only other known modification is the complete acetylation of the N-terminal α-amino group. Therefore, an intriguing possibility is that the differential assembly of the paramyosin populations is regulated by their differential phosphorylation-dephosphorylation.

Three Dimensional Model of Core Structure. Attention was focused on the cores because of the reasonable assumption that their assembly precedes that of myosin and its associated paramyosin and that the cores might serve as templates for the assembly of the outer protein. Furthermore, very little direct experimental information was available on the backbone structure of any thick filament (40). Our ability to isolate and purify thick filament cores with myosin removed from most of its surface offered a potentially unique opportunity.

Direct evidence that the cores actually contain paramyosin is shown by their labeling with monoclonal anti-paramyosin antibody and examination by electron microscopy (7). The labeled epitope shows a regular repeat of 72 nm, consistent with the polar P1 packing of paramyosin in synthetic paracrystals (5). The 72 nm repeat arises from the stagger between adjacent linear arrays of paramyosin molecules and is predicted by optimization of charge neutralization of the known C. elegans paramyosin amino acid sequence (26). The labeling suggests that the paramyosin arrays are in register across the cores as they are in paracrystals.

Electron microscopy of negatively stained cores reveals an intrinsic P1-like 72 nm repeat consistent with the previous results. However, within each 72 nm period there are four sub-bands, alternating between dark (less protein) and light (more protein). This pattern is more complicated than what had either been observed with or predicted for paramyosin alone. Examples of these images were digitized, analyzed, and modeled (13). Fourier transform analysis suggests that the paramyosin is organized into seven subfilaments which re-
repeat helically every 144 nm. Additional protein is more centrally located than the paramyosin subfilaments and repeats in phase with the paramyosin stagger of 72 nm as a 54 nm long tubule (Fig. 7).

The stagger of 72 nm between adjacent paramyosin strands necessitates a gap of 22 nm between consecutive paramyosin molecules. Modeling shows that the optimal fit of the Fourier transforms requires that the additional central protein tubules and the gaps between paramyosin molecules be phased together with an offset of 6 nm (Fig. 8). An implication of this relationship is that the protein or proteins of the additional central tubules specifically interact with the gap regions. It should be noted that these regions contain the amino terminal sequences of paramyosin and their phosphorylatable sites. Thus, at least one independent piece of evidence suggests that the gap regions are recognized by a specific protein and the kinase activity.

**Novel Proteins Of Thick Filament Cores.** Thick filaments of greater than 90% purity by protein are obtained by sedimentation on sucrose density gradients and then rebanding (7). Similarly, cores from filament dissociations are also obtained in highly purified form by this procedure. *C. elegans* body-wall muscle thick filaments sediment at 500S, and the isolated cores sediment at 150S. Three proteins in addition to myosin and paramyosin co-sediment with thick filaments and cores. These proteins have been named provisionally P20, P28, and P30 on the basis of their mobilities on SDS/PAGE (Fig. 9).

These three proteins have now been identified and characterized in terms of their amino acid and DNA coding sequences (F. Liu, C.C. Bauer, R. Cook, and H.F. Epstein, unpublished results). Their amino acid sequences are novel and do not show significant homology to known proteins (BLASTP). They could not be derived from either myosin or paramyosin as products of degradation. The locations of the genes encoding each of these proteins are distinct from the genes encoding myosin heavy chains and paramyosin.

The full coding DNA sequences for P28 and P30 have been determined by the *C. elegans* Genome Project whereas the sequence of P20 is still being assembled. The predicted molecular weights are P20: 17,000 (tentative), P28: 24,457, and P30: 30,043. All three proteins are basic, compact, and globular. P28 is predicted
to have β-sheets as its only type of regular secondary structure; P30 is suggested to have both α-helices and β-sheets.

Antibodies to peptides of P28 and P30 selected on the basis of their predicted accessibility (Wisconsin Peptide Structure) have been produced in rabbits. A highly specific antibody to P28 have been purified by affinity chromatography (F. Liu and H.F. Epstein, unpublished results). This antibody reacts with isolated cores, thus verifying P28 as a proper component of the core structures.

Unc-82, P30, and Paramyosin Assembly. As figure 5 shows, unc-82 and the P30 locus are very closely linked on chromosome IV of C. elegans. Current experiments are testing whether cloned P30 can rescue unc-82 mutant phenotypes and whether specific unc-82 mutants exhibit alterations in their P30 coding DNA.

The known properties of strains carrying mutations in the unc-82 loci are highly suggestive of the consequences of alterations in a core protein. When double homozygous strains containing the unc-82 mutant alleles e1220 or e1323, and the unc-15 mutant allele e73, which produces defective paramyosin are constructed, the combined phenotypes are intermediate between the milder unc-82 and more severe unc-15 phenotypes (44). When doubles of the same unc-82 mutant strains are constructed with the unc-15 mutant allele e1214, which is deficient or null in paramyosin, then the phenotype of the e1220 double is a severely paralyzed but viable worm whereas the e1323 double is lethal (44). These genetic experiments indicate that the unc-82 gene product or its action must interact with paramyosin or its actions.

Figure 10 shows that the unc-82 mutant e1220 directly affects the assembly of paramyosin (15). In this and other unc-82 mutants, abnormal assemblages consisting of large central structures resembling synthetic paracrystals of paramyosin and thick filament and core-like structures extending from the poles are produced in addition to thick filaments. These assemblages are similar to structures produced by e73 and other unc-15 alleles with alterations in the paramyosin gene. The locus of these unc-15 mutations (19) is in the region of paramyosin that our structural model (13) predicts would interact with the internal protein tubule of the core (Fig. 9).

Unc-45 and Myosin Assembly. Strains carrying mutations in the unc-45 locus produce a phenotype that displays disruption of thick filament assembly and interact differentially with specific strains overexpressing myo-3 or deficient in unc-54 myosin heavy chains (17, 39). These results suggest that the unc-45 gene product or its action plays an important role in the assembly of both myosins into thick filaments. They further suggest that the structure of the unc-45 protein might be complex.

Fig. 10. Mutant unc-82 Assemblages and Thick Filaments. Structures were isolated from e1220 mutant nematodes, adsorbed to glass microscope slides and localized by fluorescein-labeled anti-paramyosin antibody (11, 15). The larger structures, so-called multifilament assemblages, contain paracrystalline domains of paramyosin very similar to structures found in unc-15 missense mutants (15, 43, 44). The thinner structures are thick filaments.

Fig. 11. Temperature-Sensitive Assembly in Unc-45 Mutants. The unc-45 allele e286 and other temperature-sensitive alleles assemble thick filaments and myofibrils when grown at 15°C (a) but produce disrupted sarcomeres and smaller numbers of abnormal thick filaments when grown at 25°C (b). The recessive and reversible character of these mutants (17) are consistent with the model of unc-45 protein as a myosin assemblase. Reprinted from Nature, with permission (17).
A number of the alleles of unc-45 are temperature-sensitive (17, 39). In such mutants thick filaments are disrupted when the nematodes are grown at 25°C, and the arrays of filaments appear normal when grown at 15°C (Fig. 11). During embryonic and larval development, changes in the growth temperature lead to reversal of phenotype (17). This reversibility and the recessive character of the temperature-sensitive mutants suggest that the unc-45 gene product might be a catalyst for myosin assembly.

These suggestions on unc-45 structure and activity are highly consistent with its newly characterized coding DNA sequence (D. Pilgrim and L. Venolia, The Worm Breeder’s Gazette, 14(3); C.C. Bauer, J.M. Barral, and H.F. Epstein, unpublished observations). The protein has a predicted molecular weight of 107,454 with significant homology to known serine-threonine protein phosphatases (BLASTP). Similarities (TBLASTN) to prolyl cis-trans isomerases, Gro-EL type chaperonins and myosin heavy chain α-helices are detected as well. The phosphatase homology is highly consistent with the predictions of unc-45 as a catalyst for assembly. The similarities of multiple types of protein binding proteins is suggestive of the ability of the unc-45 product to interact differentially with myo-3 and unc-54 myosins.

**DISCUSSION**

The experimental results with C. elegans presented in this report indicate that:

1) thick filaments are complex tubules with non-equivalent positions for myosin and other component proteins;
2) the assembly of thick filaments requires distinct assemblases and coupling proteins in addition to the major structural proteins.

Such information is necessary for understanding important questions such as the regulation of thick filament assembly, the mechanisms for assembly of precise lengths and diameters of the filaments, and the unusually rigid, inelastic nature of thick filaments. This rigidity is readily explainable by the coupling of the paramyosin subfilaments into the core tubules and its necessary consequence of additional paramyosin and myosin being assembled about the core into the outer tubular domains (M.F. Schmid and H.F. Epstein, unpublished results). Such structures are inherently stronger and more rigid than purely helical, rope-like filaments.

Clearly, self-assembly of myosin and paramyosin cannot explain the formation of thick filaments in C. elegans. An open question is whether the complex assembly can be achieved under cell-free conditions or requires other intracellular structures. Significant genetic and developmental evidence in C. elegans suggests that membrane-associated proteins including vinculin (2), β1-integrin (20), and the extracellular glycoprotein perlecans (33) are required for the assembly of thick filaments and myofibrils. Mutants that produce defective thick filaments do not affect the formation of nascent linear structures near the muscle cell peripheries (11). Other mutants which interfere with the assembly of the nascent structures, do not assemble thick filaments and are lethal (42). Certain perlecans mutants permit normal embryonic myofibril assembly but then prevent further larval assembly including the formation of additional thick filaments (31). The proper assembly of thick filaments in C. elegans body-wall muscle is not only complex in terms of its sarcomeric requirements but also requires intact transmembrane and extracellular structures (24).

**Acknowledgements.** We thank Mrs. Lois Craft and Ms. Esther Wong for their aid in preparing the manuscript. Mr. J. Hariharan provided useful discussion. The research was in part supported by grants from the Muscular Dystrophy Association and NIH (2RO1 GM33223) to H.F.E., an NSRA fellowship (AG F32 AR08374-01A1) to C.C.B., and training grant (EY07001-21) to F.L. We are grateful to the organizers of the Chiba Muscle Symposium, to Chiba University, and the Government of Japan for their kind invitation and help.

**REFERENCES**


