Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this dissertation. I retain all ownership rights to the copyright of the dissertation. I also retain the right to use in future works (such as articles or books) all or part of this dissertation.

Signature:

_____ _____

Ge Xiong Date

Molecular determinants of the localization and assembly of the giant protein UNC-89 (obscurin) in *C. elegans* muscle

By

Ge Xiong
Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

_____
Guy M. Benian, M.D.
Advisor

_____
Anthony W.S. Chan, DVM, Ph.D.
Committee Member

_____
Ping Chen, Ph.D.
Committee Member

_____
Victor Faundez, M.D., Ph.D.
Committee Member

_____
William G. Kelly, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Molecular determinants of the localization and assembly of the giant protein UNC-89 (obscurin) in *C. elegans* muscle

By

Ge Xiong
B.S. and M.D., Xiangya School of Medicine, Central South University, P. R. China, 2003

Advisor: Guy Benian, M.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2011

Abstract

Molecular determinants of the localization and assembly of the giant protein UNC-89 (obscurin) in *C. elegans* muscle

By Ge Xiong

Sarcomeres, highly ordered assemblages of hundreds of proteins, perform the work of muscle contraction. Despite increasing knowledge of sarcomere components and their functions, we still do not have a clear understanding about how sarcomeres are assembled and maintained during muscle contraction. Our laboratory studies sarcomere assembly in the model genetic organism *C. elegans*, focusing on the function of the giant muscle protein UNC-89. To understand the molecular mechanisms by which UNC-89 is assembled at the M-line, and how UNC-89 performs its functions, our laboratory is identifying its binding partners, and studying their functions.

My thesis focused on finding binding partners for two regions of UNC-89: the C-terminal protein kinase region, and the N-terminal Ig domains 1-5. I discovered that the PK1 protein kinase domain and interkinase region interact with LIM-9 (FHL in humans), and that LIM-9 interacts with SCPL-1, a CTD type phosphatase, previously identified by the lab as interacting with the kinase domains of UNC-89. I propose two structural models for the function of these interactions at the M-line.

I have discovered a new binding partner for the N-terminal region of UNC-89, called CPNA-1. CPNA-1 contains a conserved "copine domain" with weak homology to the extracellular portions of integrins, and largely unknown function. I have defined a new category of copine domain containing proteins, the "atypical" ones. CPNA-1 specific antibodies localize to integrin adhesion complexes (M-lines and dense bodies) of nematode body wall muscle. I found that CPNA-1 binds to the M-line proteins UNC-89, LIM-9 (FHL), SCPL-1, UNC-96, and a protein common to the M-line and dense body, PAT-6 (actopaxin). A genomic deletion for *cpna-1, gk266*, displays the typical Pat (Paralyzed arrested at two-fold) phenotype. By localizing previously characterized muscle adhesion complex proteins in *cpna-1* mutant embryos, and localizing CPNA-1 in other Pat mutants, I have placed CPNA-1 in the M-line/dense body assembly pathway of embryonic muscle. By using RNAi and mutants, I have begun to define a role for CPNA-1 in adult muscle. I conclude with a model that PAT-6 recruits CPNA-1, and in turn, CPNA-1 recruits additional proteins (UNC-89, LIM-9, SCPL-1, UNC-96) to the M-line.

Molecular determinants of the localization and assembly of the giant protein UNC-89 (obscurin) in *C. elegans* muscle

By

Ge Xiong
B.S. and MD, Xiangya School of Medicine, Central South University, P. R. China, 2003

Advisor: Guy Benian, M.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
Graduate Division of Biological and Biomedical Sciences
Genetics and Molecular Biology

2011

Acknowledgements

I want to thank everyone who was a part of this work, and a part of my success in general. To all members of the Benian's lab, I enjoyed the time I spent with you. In particular, I want to acknowledge Dr. Guy Benian, who was a wonderful scientific mentor. Without his guidance and supervision, I can not finish my thesis projects successfully. I deeply appreciate the trust and confidence he has in me as well as the countless hours he spent helping me, an international student, to be a scientist. I also want to acknowledge Dr. Hiroshi Qadota for his help on my experimental skills. I have learned a lot from working with him.

I would like to extend my high appreciation to my committee members, Dr. Anthony Chan, Dr. Ping Chen, Dr. Victor Faundez and Dr. William. G. Kelly. Their advices have been very helpful in my project progress and future career development.

I would like to express my gratitude to my family, who has offered me unconditional love all the time. My success would have been impossible without tremendous support and encouragements from my parents and brother.

Thanks to my colleagues in the GMB program, good luck to all. To all friends who have shared happiness and stress together with me over the past several years, I am grateful. Furthermore, I would like to thank Dr. Bo Xiao and his team at the Department of Neurology, Xiangya Hospital of Central South University in China. The experience on working in this famous hospital benefits me a lot.

Table of Contents

List of Figures

1. Chapter 1: Introduction

1.1 Muscle Organization

Muscle is a contractile tissue that produces forces that permits voluntary and involuntary motion. There are three types of muscles: skeletal muscle, heart muscle and smooth muscle. Skeletal muscle is attached to the skeleton and contracts voluntarily. Cardiac muscle is the main component of the heart and the presence of intercalated discs differentiates them from skeletal muscle. Smooth muscle is located in blood vessels as well as the digestive and reproductive systems and it lacks the ordered organization of myofibrils. However, skeletal muscle and heart muscle are packed into many highly ordered bundles so they are classified as striated muscles. I will discuss skeletal muscle in this thesis. As is shown in Fig. 1-1, a piece of skeletal muscle has many bundles called fascicules and every fasciculus is composed of individual muscle cells. A muscle cell also known as a muscle fiber, is packed full of myofibrils. Each myofibril is composed of a chain of sarcomeres, the most fundamental contraction unit. A sarcomere occupies from one Z disc to the next Z disc and has a notable banded pattern under light microscopy. In human muscle, such a sarcomere (Fig. 1-1, bottom) is about 2.2 μm in length. Within a sarcomere, the dark portion is called the A band (dark band) and it is the region overlapped by thick filaments and thin filaments; the light portion is called the I band (light band) and it only includes thin filaments. Interestingly, under polarized light, A bands appear bright and I bands look dark. Thin filaments are anchored at the Z disc and thick filaments are crosslinked at the M line. The M line lies in the middle of the A band.

Myosin is the major component of thick filaments. A conventional myosin molecule

(myosin II) consists of six polypeptides (Fig. 1-2A): two myosin heavy chains (MHC), two essential light chains (ELC) and two regulatory light chains (RLC). The polypeptides form a hexamer that includes two globular heads, two necks and a rod tail. Myosin heads have both actin binding sites and ATP binding site. The single rod consists of two myosin heavy chains intertwisted into an α helical coiled coil structure (Fig. 1-2B). The rods of individual myosin molecules assemble into the thick filament shaft (reviewed in Vikstrom et al., 1997). An anti parallel overlap of rods forms the central bare zone at first and myosins are added later on either side of the bare zone to form overlapped rods in parallel. Myosin heads project from the surface of the thick filament shaft with a defined geometry which is conserved amongst all thick filaments and all animal species. The amino acid sequence of myosin heavy chain can be used to explain the formation of the α helical coiled coil rod, and the parallel assembly of individual myosin rods into the thick filament. As with most α helical coiled coils, the sequence of the myosin rod region is predicted to be α helix has a "heptad repeat" -an inexact repeat of seven amino acids. The first fourth amino acid residues (a & d position) are hydrophobic, and these residues are close to each other on the surface α helix cylinder. In a coiled coil rod, these hydrophobic residues of one polypeptide interact with the hydrophobic residues of the other polypeptide twist into a coiled coil (Fig. 1-2B). Hydrophilic residues are displayed on the surface of the coiled coil. It turns out that there is an even larger repeat in the myosin heavy chain sequence-28 residues long and there are a total of about 40 of these repeats. McLachlan and Karn (1982) proposed a model in which, in order to maximize favorable electrostatic interactions that would between adjacent parallel myosin rods, each one should be offset by approximately 3.5 of these 28 residue repeats. In a coiled coil rod,

these about 98 residues span 14.6 nm. Since this number is very close to 14.3 nm measured for the closest distance between myosin heads as the surface of the thick filament shaft, it is an excellent model. However, we still can not explain how the initial anti-parallel assembly of myosin rods occurs, or how it switches from anti-parallel to parallel assembly. Another fundamental observation that cannot yet be explained is the length of thick filaments is uniform (although it varies from 1.6 μm in human striated muscle, to about 10 μm in nematode muscle). Under appropriate in vitro conditions, purified myosin has the ability to form into filaments, which by EM, appear very much like native thick filaments. However, they are "heterodisperse", i.e., of variable length. Clearly, inside the muscle cell, this assembly is regulated.

Thin filaments mainly contain actin. Monomeric actins form two helical, interwined strands known as F-actin. The helical repeat spans F-actin monomers in one strand and is 37 nm in length. The plus end of F-actin is at Z disc and the minus end is in the middle of A band. When muscle contracts, myosin heads move from minus end to plus end of F-actin. There are other components of thin filaments. Tropomyosin which is also an α helical coiled coil rod sits in the groove of F-actin, spinning one F-actin helical repeat. It is an actin binding protein and it regulates the interaction between actin and myosin together with the troponin complex. Troponin is composed of three subunits: TN-I, TN-T and TN-C (reviewed in Phillips et al., 1986). TN-I binds to F-actin and block the actin-myosin interaction. TN-T tethers TN-C and TN-I to tropomyosin. TN-C binds calcium to remove the TN-I inhibition of interaction between actin and myosin. Actin strands surround the myosin fiber but they do not directly contact the myosin when the muscle is relaxed. Free calcium regulates the interactions between myosin and actin in both striated

muscles and smooth muscles. In the presence of high intracellular calcium concentration, calcium binds the TN-C subunit of troponin. The binding causes a conformational change of the entire tropomysin-troponin complex such that tropomysin is moved away from the position where it blocks sites on the actin monomers that could bind myosin heads. Then myosin crossbridges interact with actins and drag along the actin strands to make the striated muscles contract (Fig. 1-3). In smooth muscles, calcium regulation is different. The key protein involved is myosin light chain kinase, a protein kinase whose activity is normally blocked by having a protein of its own sequence, the "autoinhibitory sequence", sitting in its catalytic pocket preventing access to substrates. A rise in intracellular calcium results in binding to calmodulin. This calcium-calmodulin binds to the autoinhibitory the sequence of MLCK and results in the autoinhibitory sequence moving sufficiently out of the catalytic pocket on that substrate can enter. Phosphorylation of the regulatory myosin light chain of myosin results in activation of the myosin head in a still mysterious manner.

Finally, it should be pointed out that in any particular muscle cell, thin filaments are of uniform length ranging from 1.0-1.3 μm long. How this is achieved and maintained is another fundamental mystery of muscle biology. In fact, it is a fundamental unanswered question for the more general topic of actin cytoskeleton. We also don't understand how the uniform lengths of F-actin bundles comprising microvilli or stereocilia are assembled/ maintained. In fact, the heights of stereocilia in the inner ear hair cells are varied in a uniform way to achieve the best response to vibrations of sound waves. We still do not have answers for these fascinating questions. Perhaps the only satisfactory explanation for length regulation of F-actin filaments has been obtained from the short actin filaments

of the red blood cell cytoskeleton (Fowler, 1996).

Contraction is the most basic function of sarcomeres. It is now well established that the sliding filament mechanism explains how muscle contracts (Huxley and Hansen, 1954). Thick filaments and thin filaments maintain constant lengths in both contraction and relaxation states. However, the length of sarcomere becomes shorter during muscle contraction. The length of A band is constant whereas the I band shrinks during contraction. Actually thick filaments slide towards thin filaments to shorten the sarcomere and muscle contracts as a result. How the sliding is achieved? There is a lot of data to indicate that cyclical attachment and detachment between myosin heads (crossbridges) and thin filaments produce the force for contraction (Warrick and Spudich, 1987). As I mentioned before, the myosin head has an ATP binding site. In the absence of ATP, the myosin head binds to actin rigidly (the “rigor state”). In the presence of ATP, the myosin head binds ATP and dissociate from actin. When the ATP is hydrolyzed, the myosin head undergoes a conformational change and binds to a new actin monomer within the same thin filament. The phosphate is released immediately and the neck of myosin swings to pull the actin filament. Now the ADP is released and the myosin head binds to an actin monomer further down the thin filament and the cycle can start again (Fig. 1-4).

Muscles have very important physiological functions, such as movements of the animal, heart pumping and peristalsis. The highly organized structure of sarcomeres is necessary to perform optimal functions of muscles. What keeps thick filaments in lateral registration? It is suggested that the M line plays a very important role in maintaining the A band pattern. The M line is in the middle of the A band, about 100 nm wide (Fig. 1-

5A). Under the electromicroscope, a section perpendicular to M lines showed that thick filaments are cross-linked. The M line is a complex network of many proteins (Fig. 1-5B). Similarly, the Z disc is an interwoven network of many proteins that are involved in thin filament attachment/anchorage (Fig. 1-6 A&B). In fact, some proteins are localized to both M lines and Z discs and some others have the ability to shuttle in the nucleus to affect transcription. The components of M lines and Z discs not only play important roles in sarcomere assembly and anchorage but also take part in many signaling pathways including those that maintain or build myofibrils in response to muscle activity.

It has been demonstrated that hundreds of proteins are involved in the assembly and maintenance of sarcomeres. Indeed, each year new sarcomeric components are discovered. Nevertheless, we still do not have a clear picture about how sarcomeres are assembled and maintained in a highly ordered and precise way and how these structures are maintained when muscle contracts.

1.2 Use of *C. elegans* to Study Muscle Sarcomere Organization and Assembly

Caenorhabditis elegans is a free-living roundworm that can be isolated from temperate zones throughout the world, in which it lives in the soil and eats bacteria. This nematode has an unsegmented, bilaterally symmetrical cylinder body structure and it is covered by a cuticle outlayer. This multicellular organism has a very simple anatomy and it is widely used as a model system to study genetics, development and cell biology (Fig. 1-7).

In early 1970s, Sydney Brenner began to use *C. elegans* as a model organism (Brenner, 1974) because of various reasons. The worm can be fed on bacteria and easy to breed. The lab strains may be stored in liquid nitrogen and upon thawing, many can be revived.

In crowded living conditions or lack of food, worms will enter a specific larval stage called dauer state to resist the unfavorable environment. When the living conditions improve, dauer larvae can grow into normal L4 larvae and ultimately adult worms. *C. elegans* has two sexes: hermaphrodites and males. An individual hermaphrodite can self fertilize and lay about three hundred eggs. After hatching from the egg, a larva only takes 36 hours to grow into adult worms at 20 degree. Therefore, the large brood size and short life cycle make *C. elegans* a great model system to study genetics (Brenner, 1974). In 1998, the complete sequence of *C.elegans* was published (*C.elegans* Sequencing Consortium, 1998). This represented the first complete genome sequence of any metazoan. Sequence analysis predicts approximately 19,000 protein coding genes. About 60% of these genes have been characterized by mutation or intragenic deletion. The loss of function phenotype for most of the other genes can be characterized by RNAi. About half of these protein coding genes have homologs in the human genome.

The tiny worm is optically transparent facilitating microscopic analysis. This advantage facilitates cell biological studies. Every somatic cell has an invariable developmental fate so this pattern of cell lineage is useful for the study of development.

Because of the features mentioned above, *C. elegans* is a model system widely used in biomedical research, and our lab is studying the questions of sarcomere assembly and maintenance by making use of this model organism. It is very easy to carry out mutational analysis in *C. elegans*, through both forward and reverse genetics. The sarcomere structure can be evaluated under polarized light because of its optical transparency. Moreover, we can localize the sarcomeric components in fixed animals by using antibodies or GFP tagged proteins in live animals. In addition, the self-fertilization

of *C. elegans* allows propagation of muscle mutants that would be unable to mate. Most importantly, the sarcomeric structure is highly conserved from *C. elegans* to vertebrates (Fig. 1-8). The worm has several types of muscle in different locations: striated body wall muscle and single sarcomere muscle in the vulva, pharynx and male tail. In the major striated muscle of *C. elegans* , which is found in the body wall and used for locomotion, the myofibrils are restricted to a narrow ~1.5 µm zone adjacent to the cell membrane along the outer side of the muscle cell (Moerman and Fire, 1997; Qadota and Benian, 2010). All the dense bodies (Z disc analogs) and M-lines are anchored to the muscle cell membrane and extracellular matrix (ECM), which is attached to the hypodermis and cuticle. This allows the force of muscle contraction to be transmitted directly to the cuticle and allows movement of the whole animal. Worm muscle M-lines and dense bodies serve the function of analogous structures in vertebrate muscle: the dense bodies anchor thin filaments and the M-lines anchor thick filaments. In addition, because of their membrane anchorage and protein composition, they are also similar to costameres of vertebrate muscle and focal adhesions of non-muscle cells. Costameres are muscle specific focal adhesions and consist of protein assemblies aligned circumferentially in the region of Z disc or M lines of peripheral myofibrils in vertebrate muscle. Focal adhesions are multi-protein assemblages, which are crucial for cells to attach to the extracellular matrix. They are composed of integrin, alpha actinin, vinculin and many other proteins. These components are also highly conserved from *C.elegans* to vertebrates so that the M-lines and dense bodies of the body wall muscle of *C. elegans* offer a good model for studying focal adhesions. Through the studies primarily conducted by Hiroshi Qadota in our lab, we have discovered a set of interacting proteins that form a physical linkage

between the ECM, muscle cell membrane, membrane associated proteins and thick filaments at the M-line and to thin filaments at the dense body (Fig. 1-9).

There are some differences between the body wall muscle of *C.elegans* and skeletal muscle of vertebrates. As I mentioned before, all the myofibrils attach to the cell membrane in worms. However, in vertebrate skeletal muscles, only the peripheral myofibrils are anchored to the membrane directly. The body wall muscle cells are mononucleated whereas individual skeletal muscle cell contains many nuclei in vertebrates. Although vertebral muscle contains stem cells ("satellite cells"), there are no such cells in *C. elegans* muscle. (In fact, all somatic tissues of the adult nematode lack stem cells.) There is a slight difference in the space organization between vertebrate skeletal muscles vs. worm body wall muscles. The sarcomeres are parallel to the longitudinal axis of the muscle in vertebrate skeletal muscles. But the sarcomeres are oblique to the longitudinal axis of the muscle in *C.elegans*. Although there are some difference between body wall muscle of *C. elegans* and skeletal muscles of vertebrates, the sarcomeric structure and assembly are quite conserved from *C. elegans* to vertebrates. Therefore, *C. elegans* is an excellent model organism to study the assembly and maintenance of myofibrils.

By using *C. elegans* as a model system, researchers have performed large scale genetic screens to uncover many mutants, mostly recessive, with defective locomotion. Those are called "uncoordinated" (abbreviated for "Unc") and are either slow moving, paralyzed or move in an abnormal manner, as adults. There are 129 genes that cause "unc" phenotype when mutated. Among them, about 90 genes specifically affect neurons, whereas about 40 genes specifically affect body wall muscle. Our lab is interested in "muscle unc" genes

such as *unc-89*, the main focus of my thesis. "Pat" (paralyzed arrested at two fold stage) is another phenotype with muscle defects. During normal embryonic development there are both an expansion of cell number and morphogenesis in which the initially football shaped embryo elongates 4-fold, going through stages that are named according to length, 1.5-fold, 2-fold and 3-fold, before hatching from the eggshell. Normally, the embryo moves at the 1.5 fold stage, and then move vigorously within the eggshell at the 2-fold stage (Fig. 1-10). However, "Pat" mutants embryos do not move in the eggshell at the 1.5-fold stage and arrest development at the 2-fold stage. Other aspects of development continue and the embryo hatches as an abnormal L1 larva and dies. So far, 20 genes, when mutated, have been shown to give the Pat phenotype (William and Waterson, 1994; Meissener et al., 2009). A few genes (e.g., *unc-45, unc-97 and unc-112*) are Pat as nulls and Unc as hypomorphs. I will demonstrate in Chapter 3, that the loss of function phenotype for *cpna-1* is Pat.

1.3 Giant Polypeptides in Muscle

There is a preponderance of giant polypeptides in muscle. The calcium release channel of the sarcoplasmic reticulum consists of a homotetramer, of a polypeptide of a 450,000 Da. Dystrophin, mutated in Duchenne's and Becker's muscular dystrophy exists in several different isoforms, the main one being ~ 400, 000 Da. The largest polypeptides even described in any cell type are the giant polypeptides in muscle that range in size from 700, 000 Da to 4 MDa and are composed of multiple immunoglobulin (Ig) and fibronectin type 3 (Fn3) domains, one or two protein kinase domains, and in some cases, elastic regions. Although the calcium release channel and dystrophin are highly

conserved amongst nematodes and humans, there is more diversity in these Ig-domain containing giants. Nematodes have twitchin and TTN-1. Insects have projectin and D-titin, whereas vertebrates have titin. I will discuss *C. elegans* twitchin, human titin, *C. elegans* UNC-89 and its human homologue, obscurin in this section.

1.3.1 C.elegans Giants

Twitchin

Twitchin, encoded by the *unc-22* gene, is named for its characteristic loss of function phenotype: adults are either slow moving or even paralyzed, with superimposed alternating extra bending of the body at one of a few places simultaneously and at a frequency of 1-2 × per second. There are two isoforms of twitchin, twithin A (6, 839 aa, MW=753,499) (Benian et al., 1989; Benian et al., 1993) and twitchin B (7,158 aa; MW=788,952). Twitchin A has 30 Ig domains, 31 Fn3 domains and a single protein kinase domain (Fig. 1-11). The only difference between the two isoforms is the non-modular sequence between Ig domain 3 and 4; there are 25 residues in this non-modular region in isoform A whereas isoform B consists of 344 residues in the non-modular region (Ferrara et al., 2005). Between Ig domain 4 and 5, there is a glycine rich segment and it was predicated that this segment might form a flexible hinge (Benian et al., 1993). Interestingly, it is possible that the non-modular sequences between Ig domain 2 and 3 and between Ig domain 3 and 4 are structurally flexible and may function as molecular springs (Benian et al., 1993). Antibodies localize twitchin to the outer portion of body wall muscle A bands, co-localizing with MHC B (Moreman et al., 1988). Based on its "twitching" phenotype and the close similarity of its protein kinase domain to myosin

light chain kinase (a known regulator of smooth muscle contraction in vertebrates), twitchin was hypothesized to be an important regulator of muscle contraction (Benian et al., 1989). Later, its role in inhibiting the rate of relaxation was identified from studies of twitchin in Aplysia (Probst et al., 1994) and Mytilus (Siegman et al., 1998).

The Mytilus twitchin is quite similar to twitchin in *C. elegans* in its size and domain organization (Funabara et al, 2003). The release of the "catch" state of Mytilus muscle is correlated in time with the cAMP dependent kinase phosphorylation of twitchin. "Catch" is a state of muscle contraction in which high tension is maintained with little expenditure of energy. In the mussel Mytilus and other bivalves like oysters and clams, "catch" is what keeps the two shells tightly held together. Funabara et al. (2007) showed that the C terminal phosphorylation site of twitchin with its flanking Ig domain interacts with actin and myosin to form a phosphorylation-sensitive complex. Interestingly, researchers also found that an N terminal portion including two phosphorylation sites, and their flanking Ig domains 7 and 8, interact with thick and thin filaments in Mytilus anterior byssus retractor muscles (Butler et al., 2010). The interaction with thin and thick filaments at multiple sites within twitchin suggests that twitchin acts as a tight linkage between thick and thin filaments during "catch". Whether there is a similar function for nematode twitchin, and whether there is a state comparable to "catch" in nematode muscle is not currently known.

Twitchin has a protein kinase domain and this kinase domain is most homologous to vertebrate smooth muscle myosin light chain kinase (MLCK). The identity between *C. elegans* twitchin and MLCK is as high as 52% (Benian et al., 1989). Previously, our lab reported that bacterially expressed twitchin kinase domain has kinase activity *in vitro* and

it is autoinhibited by a segment of 60 residues just C terminal portion of the catalytic core (Lei the et al., 1994). The crystal structure of twitchin kinase domain has been solved and it demonstrated for the first time an intrasteric mechanism for the regulation of protein kinases that is common for the regulation of many protein kinases (Hu et al., 1994). Moreover, the crystal structure of a larger segment of twitchin showed that the Ig domain just C terminal of the kinase domain contacts the kinase domain on a surface opposite from the catalytic pocket (Kobe et al., 1996). It is hypothesized that this Ig domain might bind to myosin and direct the twitchin kinase to the substrates. Interestingly, all of the twitchin/titin/MLCK family members share a conserved pattern for kinase domain and its adjacent Ig domains: Ig-Ig-Fn3-kinase-Ig (Southgate and Southgate 2001). Such a conserved pattern may serve very important functional roles.

Recently, by using the yeast two hybrid method, our lab uncovered a new binding partner for the kinase domain of twitchin-K08F8.1, a conserved protein kinase called MAP kinase activated protein kinase 2 (MAPKAP kinase 2) (M. Furakawa, H. Qadota, G. Benian, unpub data). Domain mapping shows that only the kinase domain of twitchin is required for interaction and the interaction is quite specific because of two experimental results. First, the kinase domains of TTN-1 and UNC-89 do not interact with K08F8.1. Second, the kinase domains of the two other MAPKAP kinase homologs in the nematode (MAK-2 and MNK-1) do not interact with twitchin. As I mentioned above, *unc-22* mutants show the “twitching” phenotype. When wild type worms are exposed to the acetylcholine agonist nicotine, they become paralyzed, whereas *unc-22* mutants twitch violently. Interestingly, mutants with intragenic deletions of K08F8.1 become paralyzed at a faster rate than wild type when exposed to nicotine. Knocking down *unc-22* by RNAi

results in animals that continue to move in nicotine regardless of whether or not K08F8.1 is mutant. Therefore, *unc-22* (twitchin) is epistatic to K08F8.1, and this suggests that twitchin is downstream of K08F8.1. Based on these findings, we hypothesize that K08F8.1 might phosphorylate twitchin kinase and this might affect the kinase activity of twitchin.

TTN-1

C. elegans TTN-1 has a predicted molecular mass of 2.2 MDa (Flaherty et al., 2002) and such a protein has been detected by western blot (Forbes et al., 2010). *ttn-1* is expressed in body wall and vulval muscle. TTN-1 is localized to I bands by immunofluorescent staining with specific antibodies. It has been shown that TTN-1 can be detected as early as the L2-L3 larva stages by immunostaining and the staining pattern continues to the adult stage (Flaherty et al., 2002). At 1.5 fold stage, embryonic myofibrils have already formed and the embryo starts to move within the eggshell; therefore, TTN-1 may not be involved in the embryonic muscle development.

In addition to 56 Ig domains, 11 Fn3 domains, and a protein kinase domain near its C-terminus, TTN-1 has two regions consisting of tandem repeats (Flaherty et al. 2002), called PEVT and AAPLE (Fig. 1-12). The largest of these is PEVT which is ~2400 residues in length, ~56% of which is just four amino acids (P, E, V, T (or K)) and exists in mostly tandem repeats of 39 residues in length. The amino acid composition and tandem repeat structure is similar (but not linearly homologous) to the main elastic region of vertebrate titin called PEVK (70% P, E, V, K; tandem repeats of 28 residues). Thus, this region of TTN-1 is hypothesized to be elastic. Circular dichroism on peptides

representing one or several copies of these repeats in TTN-1 indicate that they have the same conformations (PPII helix and random coil) and conformational malleability as the PEVK region of vertebrate titin (Forbes et al. 2010), consistent with the idea that indeed, much of TTN-1 is elastic.

The protein kinase domain of TTN-1 at its C terminus is most homologous to twitchin MLCK kinase (~52%). Compared to twitchin, the identity between the kinase domain of TTN-1 and vertebrate titin kinase domain is lower, about 39%. Thus, we view TTN-1 as a twitchin/titin hybrid: the presence of multiple Ig and Fn3 domains, and a kinase most similar to twitchin, and multiple long elastic region similar to the I-band portion of vertebrate titin. Bacterially-expressed TTN-1 kinase domain has kinase activity in vitro (Flaherty et al., 2002). Similar to twitchin kinase, TTN-1 can be autoinhibited by a sequence C-terminal to its catalytic core. Our lab found that there are two autoinhibitory sequences with different lengths (60 and 75 residues) of TTN-1 generated by alternative splicing. The change in length of autoinhibitory sequence results in different protein kinase activitities. Thus, it is speculated that this represents a novel regulation mechanism of protein kinase activity.

A key fascinating question in the field of the giant kinases is how the normally autoinhibited kinases become activated. For human titin kinase, in developing muscle, this occurs through 2 steps—calmodulin binding and phosphorylation of a tyrosine in the catalytic core (Mayans et al. 1998). However, (1) this tyrosine kinase activity is not present in mature human muscle, and (2) this tyrosine is not conserved in nematode twitchin or insect projectin kinases. Calmodulin can bind to twitchin kinase, but this only increases the activity by 2 fold (Lei et al. 1994). Recently, it has been speculated that

these protein kinases are activated by small mechanical forces that occur repeatedly as muscle undergoes contraction and relaxation (Grater et al. 2005). Forces are small enough to cause the autoinhibitory segment to be removed from the catalytic pocket to allow access to the substrate, but not enough to unwind the catalytic core. Through a collaboration of our lab with the lab of Andres Oberhauser at UTMB, we have provided the first single molecule evidence for this hypothesis by using atomic force microscopy (Greene et al. 2008), and this was confirmed and extended for human titin kinase (Puchner et al. 2008).

1.3.2 Human Titin

In vertebrate striated muscle, titin functions both in myofibril assembly and in providing passive tension for muscle. Single titin polypeptides are 3-4 MDa, 1.2 μm long at rest, and span half of a sarcomere, with N-termini at the Z-disk and C-termini at the M-line (reviewed in Kontrogianni-Konstantopoulos et al. 2009, and in Linke et al. 2008). The complete sequence of one human titin isoform contains 166 Ig domains and 132 Fn3 domains, a single protein kinase domain, and a PEVK region (Labeit and Kolmerer, 1995). The Ig and Fn3 domains are arranged into different patterns or super-repeats in different regions of the sarcomere (Fig. 1-13). Indeed, the main types of domains, Ig and Fn3, which are found in titin, are the same as the Benian lab first described earlier for nematode twitchin (Benian et al. 1989; Benian et al. 1993). In fact, there is a conservation of arrangement of the combination of 16 Ig and Fn3 domains that surround the kinase domains in each of the human and nematode proteins. The A-band segment of titin is tightly associated with the shaft of the thick filament, and specific regions of titin interact

with myosin, thick filament accessory proteins (such as myosin binding protein C), and M-line proteins (such as myomesin). Differential splicing of the titin gene yields many isoforms ranging in size from 700 to 3700 kDa (Bang et al. 2001). Most of the variation is found in the I-band portion created by varying the numbers of tandem Ig domains and the length of the PEVK region. Most of the passive tension of muscle arises from the reversible extension of the I-band segment of titin. Both poly-Ig and PEVK regions are considered distinct spring elements. For skeletal muscle titin, the poly-Ig region straightens at modest sarcomere stretch (without unfolding of Ig domains), and the PEVK region extend at higher physiological stretch. In cardiac muscle titins, there is a third spring element comprised of the "N2B" sequence, which extends together with the PEVK region.

There is also convincing evidence that titin is involved in several signaling pathways. At least three regions of titin form complexes with other proteins that are implicated in signaling. In the Z-disk region of titin, repeats Z1 and Z2 interact with telethonin (T-cap), which itself interacts with a potassium channel, myostatin (an inhibitor of muscle growth), and muscle LIM protein (Mues et al. 1998; Furukawa et al. 2001; Nicholas et al. 2002; Schallus et al. 2009). Z-line repeat Z4 and the 700 kDa "novex-3 titin" interact with obscurin, which at least partly activates rhoA (see above). The M-line region of titin, specifically the protein kinase domain, interacts with a complex of proteins (nbr1, p62, MuRF2, SRF) and as a result of phosphorylation of nbr1 and p62, less MuRF2 and more SRF appear in the nucleus, and this stimulates transcription of myofibrillar genes (Lange et al. 2005). As noted above, an attractive model is that muscle activity promotes the activation of the protein kinase domain of titin. Thus, in this way, muscle activity

increases titin kinase activity, and through increased SRF in the nucleus, more transcription of myofibrillar genes occurs. This mechanism, at least partly explains how exercise promotes the building of muscle (increasing numbers of myofibrils per muscle cell), and how muscle disuse promotes muscle atrophy (decreasing the numbers of myofibrils per muscle cell). Finally, the PEVK region of titin contains many tandem repeats of Src homology 3 (SH3) binding motifs and is thought to be a stress-sensitive scaffolding adaptor for SH3-containing signaling proteins (Ma et al. 2006).

Autosomal dominant mutations in human titin result in various forms of cardiomyopathy or muscular dystrophy: some cases of dilated cardiomyopathy, some cases of limb girdle muscular dystrophy, tibial muscular dystrophy (a late-onset distal myopathy of skeletal muscle without cardiac involvement), and hereditary myopathy with early respiratory failure ("HMERF") (Linke 2008; Lange et al. 2005). Mutations in several proteins that interact with titin also cause various types of limb girdle muscular dystrophies, including mutations in telethonin (T-cap), myotilin, and the muscle-specific protease calpain-3 (Udd 2008). Additionally, patients with dilated cardiomyopathy, coronary artery disease and heart failure show altered titin isoform expression patterns in their cardiac tissue (Kruger and Linke, 2009). These different isoforms, with different mechanical properties (e.g. more elastic) are likely to result in the titin kinase experiencing different levels of force. Thus, a better understanding of the activation of titin kinase may lead eventually to new treatments for these more common heart diseases.

1.3.3 Obscurin (Human UNC-89)

Obscurin is the human homologue of UNC-89 (Young et al., 2001; Bang et al., 2001; Small et al., 2004). It is composed of the same domains but the arrangement is somewhat different (Fig. 1-14). The trio of SH3, DH and PH domains is located near the C terminus of obscurin while it is located at the N terminus of UNC-89. Although UNC-89 is located only at the M-line, various obscurin isoforms are located at either the M line, the A/I junction or at the Z-disc and Z/I junction (Bowman, et al., 2007).

As with UNC-89 (see below), binding partners have been found for obscurin: Ig48 and Ig49 within the C-terminal 1/3 of obscurin interact with specific Ig domains within the Z-disk portion of titin (Young et al. 2001) and with a C-terminal portion of the 700 kDa "novex-3 titin" (Bang et al. 2001). A sequence of 25 residues located near the C terminus of obscurin interacts with a small isoform of ankyrin 1, and this is postulated to provide a link between myofibrils and the sarcoplasmic reticulum (Kontrogianni-Konstantopoulos et al., 2003; Bagnato et al., 2003; Kontrogianni-Konstantopoulos et al., 2006). Overexpression of part of the C terminus in primary skeletal myotubes results in disorganization of the A-band, and co-immunoprecipitation experiments suggest that obscurin interacts with myosin (Kontrogianni-Konstantopoulos et al. 2004). This indicates that obscurin, like UNC-89 (see below), has a role in the assembly or maintenance of A-bands. The DH domain of obscurin has been shown to interact with the scaffolding protein "Ran binding protein 9" (RanBP9), and both obscurin's DH domain and RanBP9 interact with titin to regulate titin's incorporation into Z-disks (Bowman et al., 2008). However, interaction to RanBP9 may be an artifact: RanBP9 is commonly isolated as a binding partner when using the yeast 2-hybrid system employed by the

authors (H. Qadota, pers. communication). A more likely binding partner for the obscruin DH domain is rhoA (Ford-Speelman et al., 2009), as this agrees with data from our lab on UNC-89 (Qadota et al. 2008b). These authors found that the DH-PH region (a putative rhoGEF domain) binds selectively to rhoA and that rhoA co-localizes with obscurin at the M-line in skeletal muscles. Fukuzawa et al. (2008) have shown that the most N-terminal Ig1-3 of obscurin interact with the C-terminal most Ig domain of titin, and also with the M-line protein myomesin. Both titin and myomesin also interact with the N-terminal domains of obscurin-like 1 (Obsl1), a small homolog of obscurin that is not found in *C. elegans*. When the expression of myomesin is either knocked down or increased in neonatal cardiomyocytes, the assembly of obscurin and Obsl1 at M-lines is disrupted. Interestingly, mutations in titin that have been associated with limb-girdle muscular dystrophy type 2J, result in decreased interaction of titin with obscurin and with Obsl1. This decreased binding results in obscurin mis-localization, and might contribute to the pathogenesis of this form of muscular dystrophy. Finally, Ig2 of obscurin interacts with a variant of myosin binding protein C at the M-line (Ackermann et al. 2009). Given the high degree of conservation in domain organization between obscurin and UNC-89, we would expect that many of the binding partners will also be conserved. Indeed, all 8 of the UNC-89 interactors that we have identified (see below) have human homologs.

Lange et al. (2009) have reported knocking out obscurin in the mouse. Remarkably, animals homozygous for the knockout have sarcomeres with normal organization, and muscle function, assessed by various methods, is normal. However, there is disorganization of the sarcoplasmic reticulum architecture: the SR appears broken or not continuous, as it is in wild type muscle. The SR phenotype is consistent with the

previously reported interaction obscurin with an SR-specific small ankyrin (noted above). In addition, the obscurin knockout mice display an age-dependent occurrence of centrally placed nuclei in some of their muscle fibers (cells): this is a sign of muscle degeneration and consequent regeneration. Whether this function of obscurin to link myofibrils and SR is also conserved in nematode UNC-89, is currently unknown, but is under investigation (K. Norman, pers. commun.).

1.3.4 UNC-89

UNC-89 is one of the three giant polypeptides in *C. elegans*. Although not obviously slow moving by casual observation, *unc-89* loss of function mutants do display reduced locomotion when assayed by counting the number of times the worm moves back and forth in liquid (Fig. 1-15). By polarized light microscopy, wild type animals display alternating bright A-bands and dark I-bands in their body wall muscle cells. In *unc-89* mutants, this organization is missing, and instead a "basket weave" pattern is characteristically found (Fig. 1-15). In addition, the single sarcomere muscle of the pharynx, the neuromuscular pump that brings and grinds up bacterial food, is also disorganized in *unc-89* mutants (Fig. 1-15). By electron microscopy, in *unc-89* mutants, the entire myofilament lattice is thinner than in wild type, and for most alleles, the M-line is missing (Fig. 1-15).(Waterston et al. 1980; Benian et al. 1999). *unc-89* mutants show a disorganization of myosin thick filaments by immunostaining (Qadota et al. 2008b).

unc-89 is a complex gene: through the use of 3 promoters and alternative splicing, at least 6 major polypeptides are generated, ranging in size from 156,000 to 900,000 Da (Benian et al. 1996; Small et al., 2004; Ferrara et al., 2005). The largest of these isoforms,

UNC-89-B and UNC-89-F, which are each ~900,000 Da, consist of 53 Ig domains, 2 Fn3 domains, a triplet of SH3, DH and PH domains near their N-termini, a region containing multiple KSP motifs (perhaps phosphorylated), and two protein kinase domains (called PK1 and PK2) near their C termini (Fig. 1-16). Although the Benian lab has not yet been able to demonstrate protein kinase activity for either PK1 or PK2, molecular modeling suggests that PK2 is active, while PK1 is inactive (Small et al., 2004). By immunofluorescence microscopy, antibodies localize UNC-89 to the middle of the sarcomere, consistent with an M-line localization (Benian et al. 1996; Small et al. 2004). Moreover, immunogold labeling shows that UNC-89 is indeed localized to the M-line, a structure that can really only be visualized by EM (P. Hoppe and G. Benian, unpub. data).

To understand how UNC-89 is localized to the sarcomere and performs its functions, the Benian lab is using a yeast 2-hybrid approach to identify its binding partners. My thesis project is part of this large effort, which is also being conducted by 3 other people in the lab. The coding sequence for UNC-89-B is entirely represented as a set of 17 overlapping segments in 2-hybrid vectors (explained further in chapter 3). All 17 have been used to screen our yeast 2-hybrid bookshelf of ~23 known components of the nematode M-line. Seven of the 17 have been used to screen a yeast 2-hybrid library. Interactions are summarized in Fig. 1-17. The kinase domains, PK1 and PK2, each interact with SCPL-1 (Qadota et al. 2008a), a CTD-type protein phosphatase. Until our work was published, the CTD phosphatases were well known to be involved in the regulation of transcription either through dephosphorylation of the C-terminal domain (CTD) of the largest subunit of RNA polymerase II, or in dephosphorylating Smad transcription factors. *Our results implicate an absolutely new function for this class of*

phosphatases: in sarcomeres they are somehow involved in giant kinase signaling. Loss of function for *scpl-1* results in a mild defect in the function of egg laying muscles. In fact, *scpl-1* is epistatic to *unc-89*; thus, *scpl-1* may function downstream of *unc-89* in egg laying muscles. Although loss of function of scpl-1 does not result in a defect in sarcomere organization, it does result in increased ability of adult worms to maximally bend, interpreted as a defect in force transmission (J. Nahabedian, H. Qadota, H. Lu, and G. Benian, manuscript in prep.). As described in chapter 2, I have discovered that both the PK1 region and the "interkinase region" each interact with LIM-9 (Xiong et al. 2009), the closest worm homolog of the human protein FHL ("four and a half LIM domains" protein). In addition, the lab has shown that the DH-PH region of UNC-89 activates RHO-1 (RhoA) specifically, and attenuated RNAi for *rho-1* results in disorganization of muscle thick filaments (Qadota et al. 2008b). Again, as mentioned above, this is consistent with what was found later for human obscurin. The lab has also determined that the segment Ig49-Ig50-Ig51 interacts with HIF-1 (hypoxia inducible factor), and the neighboring segment Fn1-Ig52 interacts with HUM-6, a class VII unconventional myosin (K. Wilson, H. Qadota, G. Benian, unpub. data). Most recently, our search for binding partners is pointing to a role for UNC-89 in regulating ubiquitin-mediated degradation of sarcomeric proteins. One of the postdocs in the lab, M. Berenice Duran, has found that 3 regions of UNC-89 (Ig8-13, Ig18-23 and Ig48-Ig52) interact with CUL-1, the nematode ortholog of cullin 1. Cullins are highly conserved proteins that act as scaffolds for the assembly of the ubiquitination machinery, one of the main ways in which cells breakdown proteins. Moreover, Kristy Wilson (another postdoc in the lab) and I found that UNC-89 (Ig2-Ig3 and interkinase Ig53-Fn2) interacts with MEL-26, which is one

type of "substrate recognition protein" (and contains a BTB domain) that interacts with CUL-3, the worm ortholog of cullin 3. The lab is testing the hypothesis that the interaction of UNC-89 regulates the activity (reduces or increases) of the cullin complexes and thus protein degradation in muscle.

Finally, my main studies have revealed that UNC-89 Ig1-Ig3 interacts with CPNA-1, a new and conserved muscle focal adhesion protein. What makes this finding unique is that, among all the interactors found so far, loss of function for *cpna-1* has the strongest phenotype; it is embryonic lethal. Moreover, loss of cpna-1 results in mis-localization of UNC-89 in the sarcomere. Thus, CPNA-1 may direct assembly of UNC-89 to the M-line, whereas the other partners identified to date, may assemble upon UNC-89, acting as a scaffold, after UNC-89 is localized to the M-line. These findings and models will be fully explained in chapter 3.

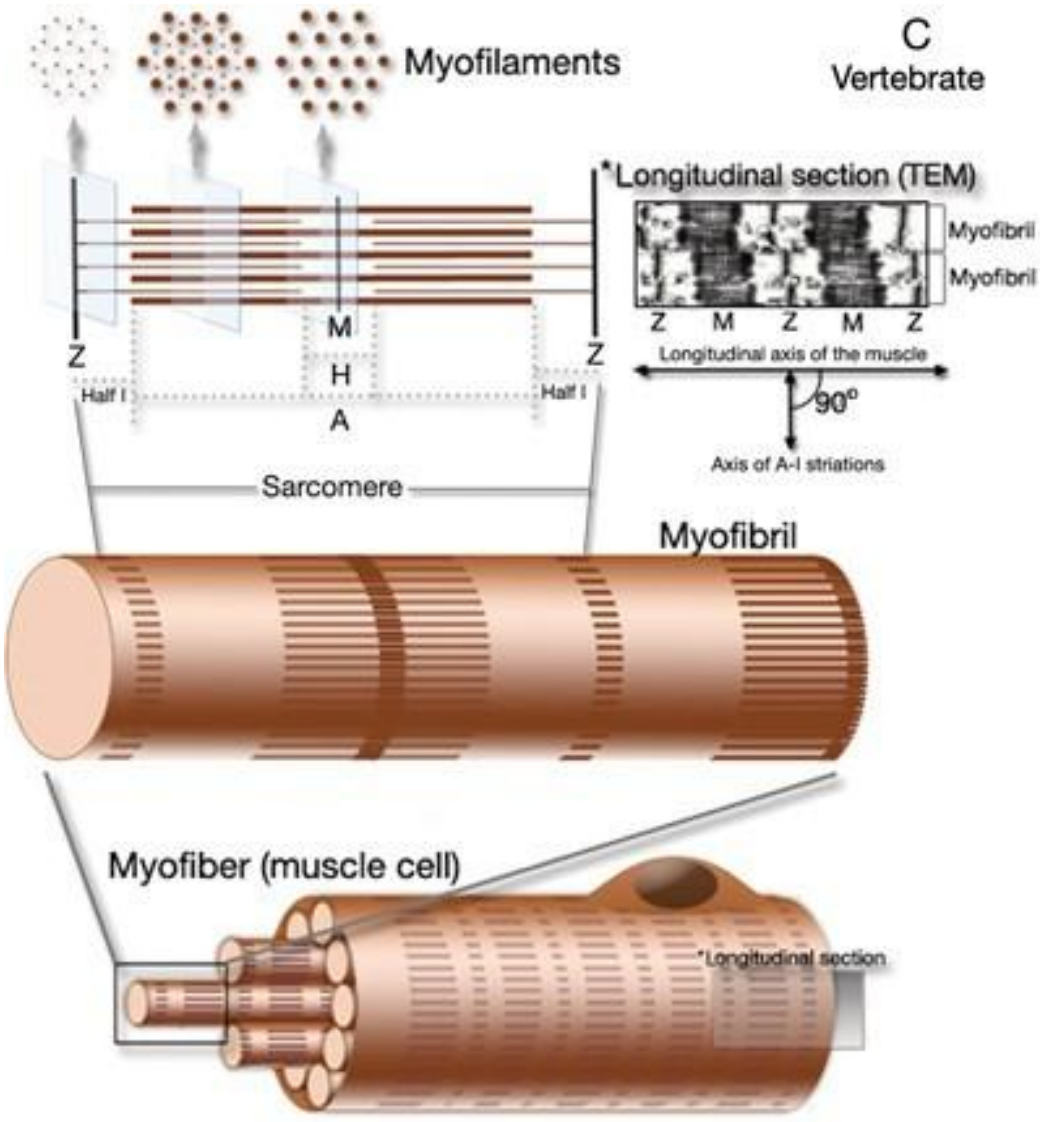
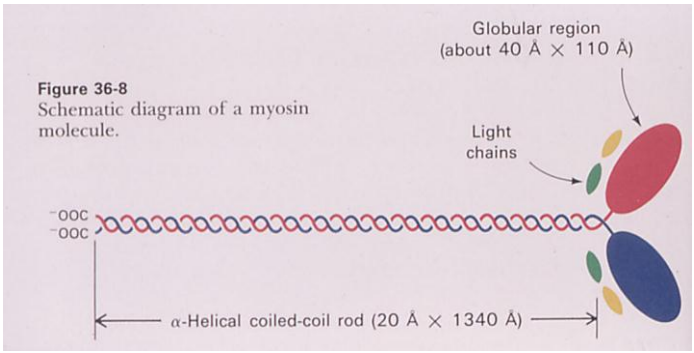


Figure 1–1. Vertebrate Skeletal Muscle.

Skeletal muscle is composed of many bundles of individual muscle cells. Each muscle cell, also known as a "muscle fiber", contains multiple nuclei resulting from fusion of multiple myoblasts during development. The muscle cell is packed full of the contractile apparatus, called the myofibrils. Each myofibril is made of a long chain of contractile units called sarcomeres. As can be seen by transmission EM, a sarcomere spans from one Z-line or Z-disk to the next Z-line or Z-disk. The middle of the sarcomere contains the A-band, and is flanked by two half I-bands. The I-band contains only thin filaments and the

A-band contains thick filaments and overlapping thin filaments. Thin filaments are anchored to Z-lines, and thick filaments are anchored to M-lines. The H zone refers to the region in the middle of the sarcomere devoid of thin filaments. The major protein of thick filaments is myosin and the major protein of thin filaments is actin (adapted from Altun, Z.F. and Hall, D.H. 2009. Muscle system, introduction. In *WormAtlas*. doi:10.3908/wormatlas.1.6, http://www.wormatlas.org/hermaphrodite/muscleintro/mainframe.htm).

A



B

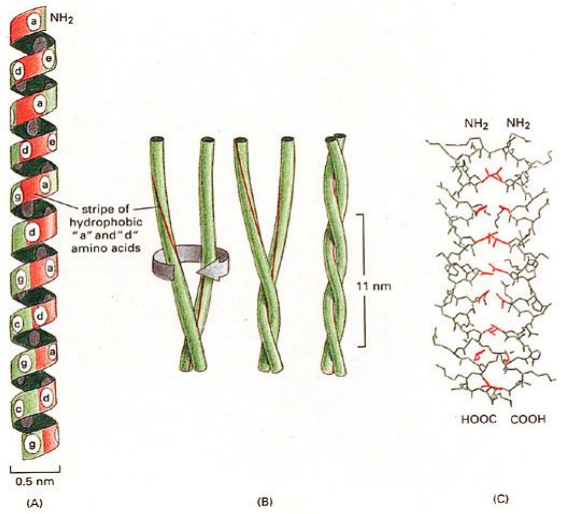


Figure 1–2. Thick filament myosin.

(A) Thick filament myosin is an asymmetric molecule comprised of two gobular heads and an a-helical coiled-coil rod. Each myosin molecule consists of six polypeptides: two myosin heavy chains, and four myosin light chains (two essential myosin light chains and two regulatory myosin light chains). (B) Explanation of how the coiled-coil rod of myosin is formed from the rod portions of two myosin heavy chain polypeptides. The rod portion of each polypeptide forms an α-helix. The sequence of this rod portion consists of

tandem (inexact) repeats of 7-amino acids (the so-called “heptad repeat”). The first and fourth residues of the heptad are hydrophobic and reside close to each other on the surface of the α-helix. These hydrophobic residues of one polypeptide interact with the hydrophobic residues of the other polypeptide, and by twisting each α-helix into a coiled coil, bury the hydrophobic residues away from the aqueous environment. This explanation is valid for the formation of any α-helical coiled coil, not just the rod portion of myosin (adapted from Alberts, B., et al., 2002, The Molecular Biology of the Cell, reproduced by permission of Garland Science/Taylor & Francis Books, Inc).

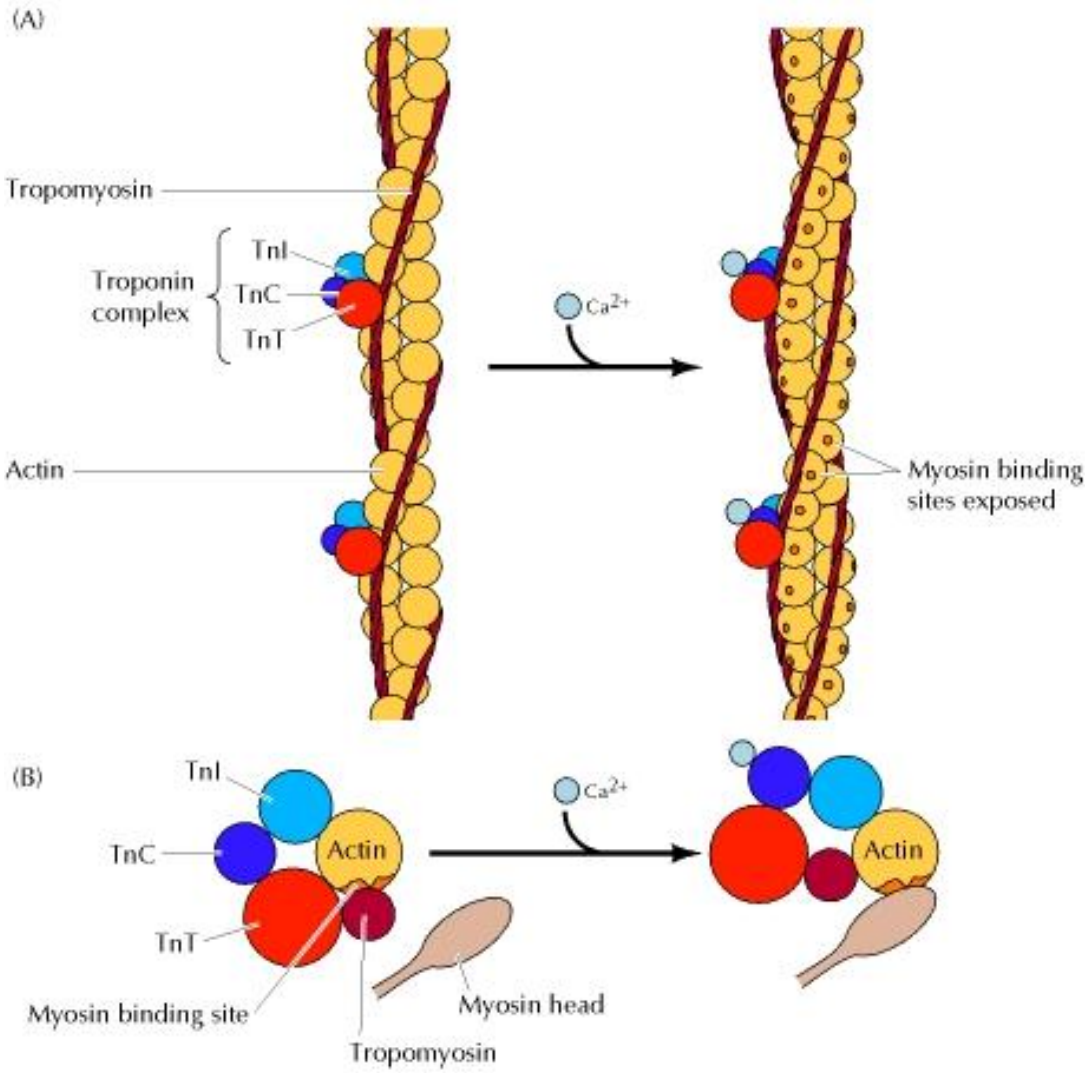


Figure 1–3. Calcium Regulates the Contraction of Striated Muscles.

(A) Tropomyosin is a coiled-coil rod like structure associated with troponins I, T and C, and resides in the groove of the F-actin helix. This tropomyosin-troponin complex blocks the myosin binding sites of F-actin when muscle is relaxed. In the presence of high intracellular calcium, calcium binds the troponin-C subunit and causes a conformational change of the entire tropomyosin-troponin complex, such that the complex is moved away and the myosin binding sites of F-actin are exposed. Myosin heads interact with actin subunits along the thin filament and pull the thin filaments, thus resulting in muscle

contraction. This is the form of calcium regulation of the myosin-actin interaction found in skeletal muscle. (B) The view of the cross section (adapted from Cooper, G.M., 2000, The Cell: A Molecular Approach).

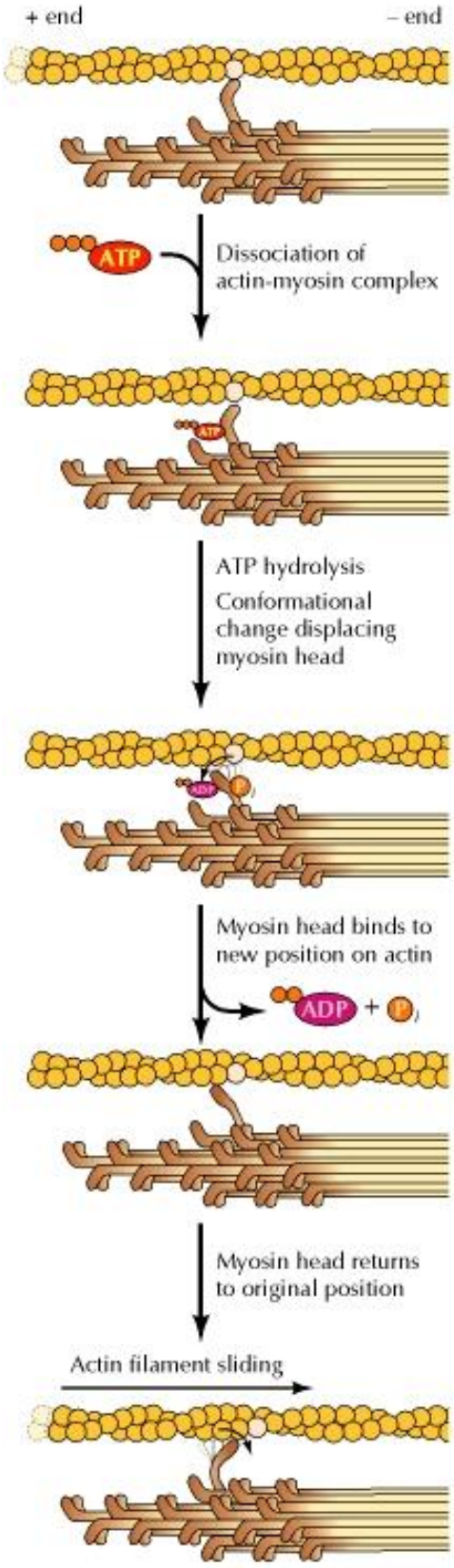

+ end
– end
ATP
Dissociation of
actin-myosin complex
ATP hydrolysis
Conformational
change displacing
myosin head
Myosin head binds to
new position on actin
ADP + Pi
Myosin head returns
to original position
Actin filament sliding

Figure 1–4. Interaction of One Myosin Head with An F-actin Filament.

In the absence of ATP, the myosin head binds to actin tightly (the "rigor state"). In the presence of ATP, the myosin head binds ATP and dissociates from F-actin. When the ATP is hydrolyzed, the myosin head undergoes a conformational change and binds to a new actin monomer within the same thin filament. The phosphate is released and the myosin head undergoes another, even larger conformational change, while still bound to actin, such that the neck of myosin swings to pull the actin filament ("power stroke"). Then, ADP is released and the myosin head binds to an actin monomer further down the thin filament and the cycle can start again (adapted from Cooper, G.M., 2000, The Cell: A Molecular Approach).

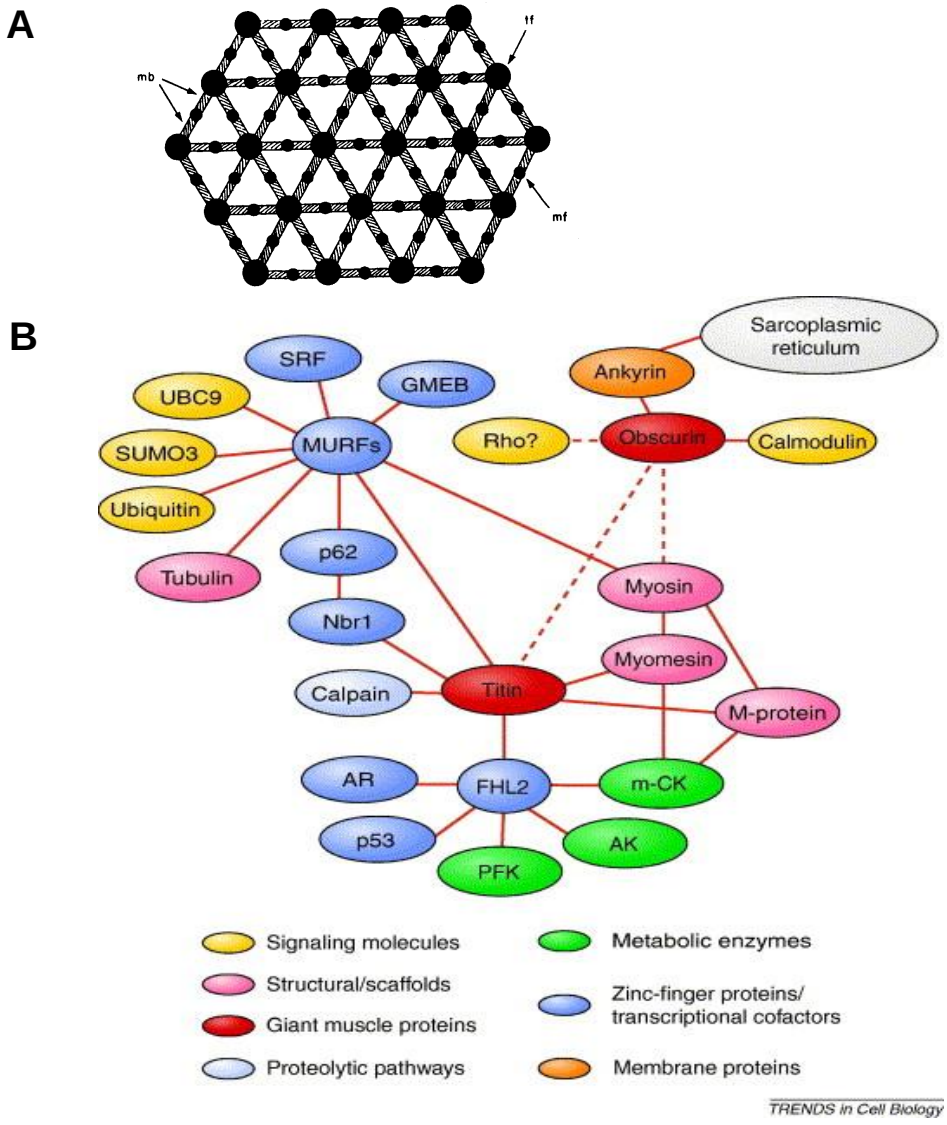


Figure 1–5. The M-line of vertebrate striated muscle.

(A) Drawing of EM of a section perpendicular to M-line is on the left; the large dots are thick filaments (tf) and appear "cross-linked" (adapted from Luther and Squire, 1978). (B) The M-line consists of a network of interacting proteins. Known or inferred functions of these proteins are indicated with the different colors (taken from Lange et al., 2006).

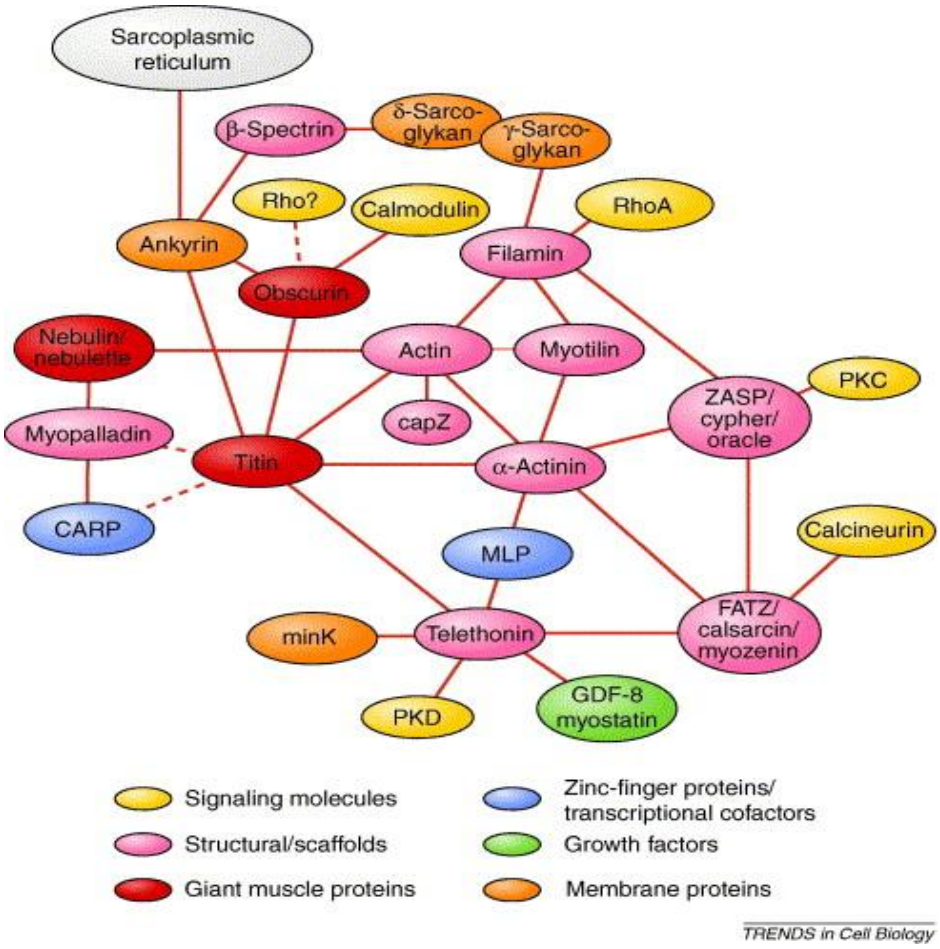


Figure 1–6. The Z-Disk of vertebrate striated muscle.

Like the M-line, the Z-disk is a hub of many interacting proteins. Colors depict known or inferred functions of these proteins. (taken from Lange et al. 2006).

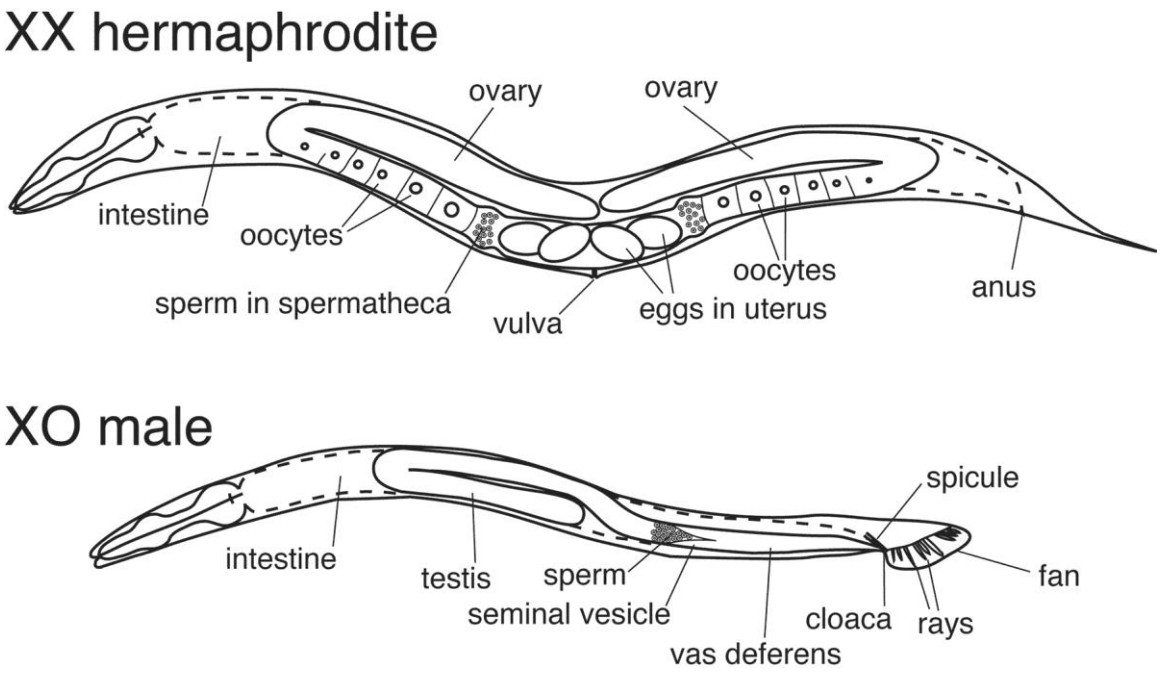


Figure 1–7. Diagram of Anatomic Features of *C. elegans*.

Although the adult *C. elegans* has only 959 somatic cells (in the hermaphrodite), these are organized into tissues such as body wall muscle, pharynx, a nervous system, digestive system and reproductive system. In a wild type population the vast majority of animals are hermaphrodites. Males, which arise by non-disjunction of the X chromosome at a frequency of approximately 1/700, primarily differ from hermaphrodites in the reproductive system and tail (adapted from Zarkower, D. Somatic sex determination (February 10, 2006), *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.84.1, www.wormbook.org).

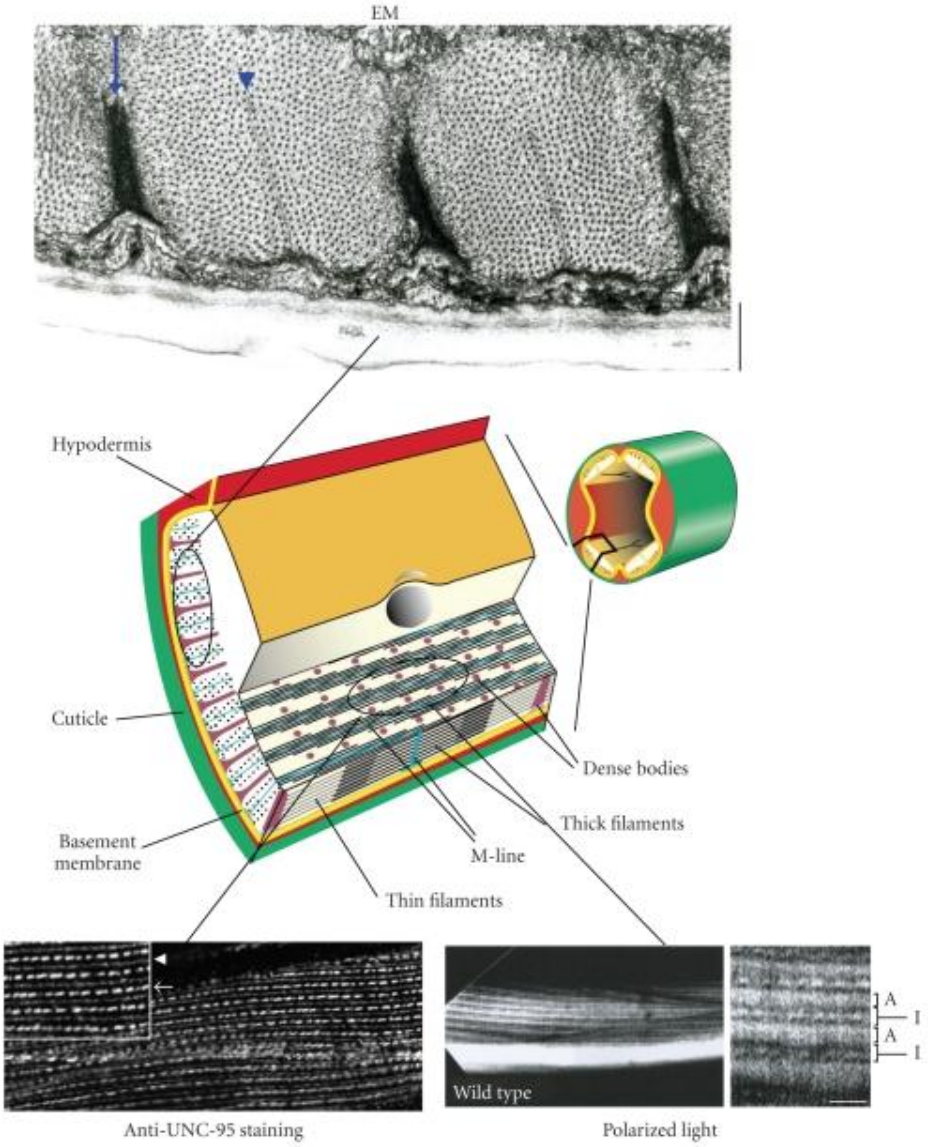


Figure 1–8. The Structure of *C. elegans* Body Wall Muscle.

The color diagram shows a blow up of a section of a quadrant of body wall muscle from *C. elegans*. On the top is the view under EM of a cross section of the worm, which is also a cross section through body wall muscle; large dots are cross sections of thick filaments; the blue arrow points to a dense body and the blue arrowhead points to an M-line. Note that dense bodies and M-lines are anchored to the muscle cell membrane. On the bottom

is the view under polarized light or immunofluoresence microscopy. Under polarized light, the dark stripes are I-bands and the light stripes are A-bands. Note that under higher magnification, shown on the right, the I-bands contain a series of dots or dashes; these are top views of the dense bodies (Z-disk analogs). On the left is an example of immunostaining of body wall muscle using antibodies to UNC-95. Note localization to M-lines (arrowhead) and dense bodies (arrow) (adapted from Qadota and Benian, 2010).

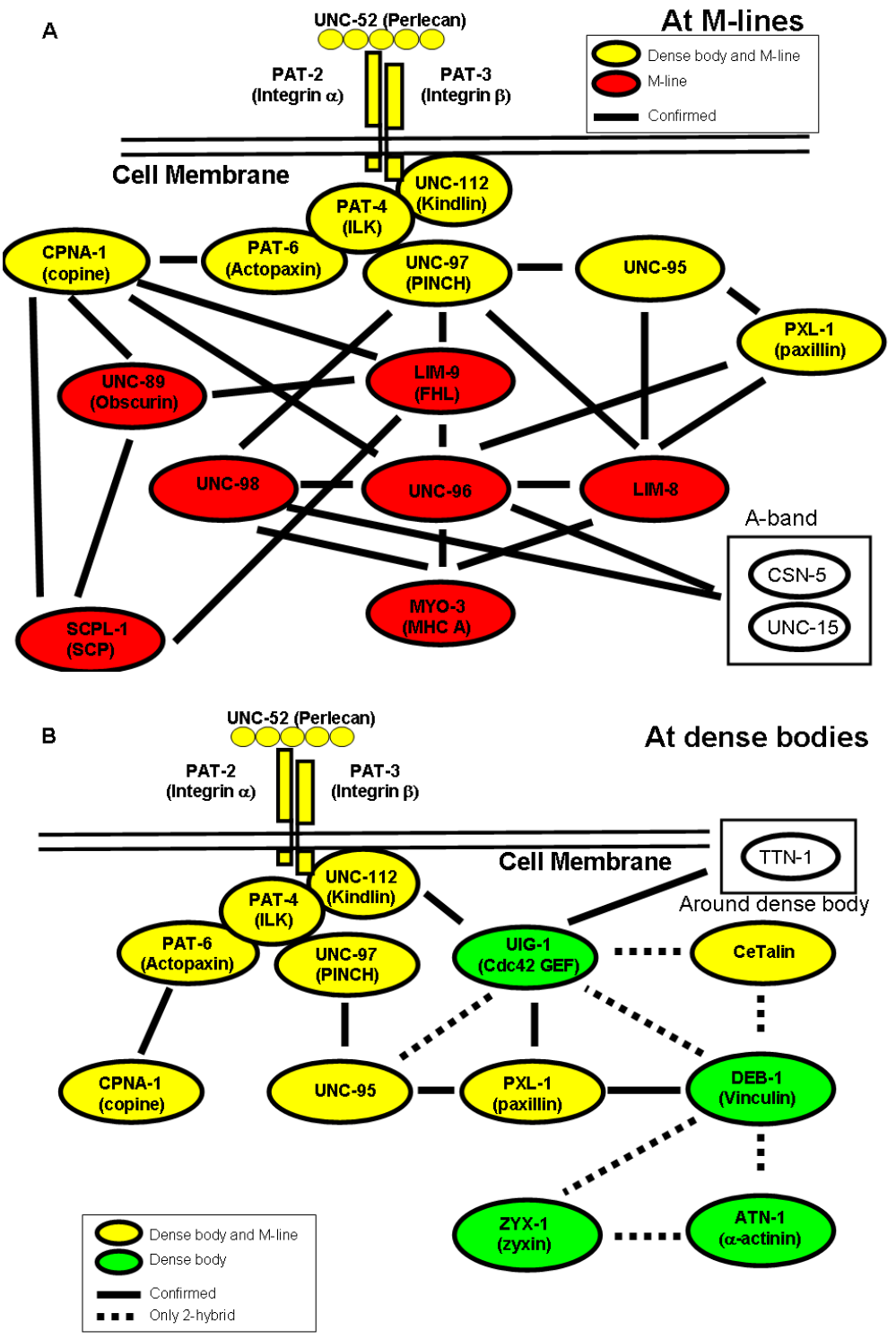


Figure 1–9. Network of Interacting Proteins at M-lines and Dense Bodies of *C. elegans* Striated Muscle.

(A) At M-lines there is a physical linkage of proteins beginning in the ECM (UNC-52 (perlecan)), through the muscle cell membrane (PAT-2 and PAT-3 (integrins)), through membrane-associated proteins (UNC-112, PAT-4, PAT-6, UNC-97), adaptors (LIM-9, UNC-98 etc.), and ending finally in myosin of the thick filaments. Most of these interactions were discovered by Hiroshi Qadota in the Benian lab, first through yeast 2-hybrid screens, confirmation by biochemical methods, and localization using antibodies or GFP fusions. (B) At dense bodies there is a physical linkage of proteins from the ECM ultimately to actin of thin filaments. Note that the same ECM, membrane and membrane-associated proteins are used at the base of both M-lines and dense bodies (shown in yellow). Vertebrate or human homologs of each protein are indicated in parentheses.

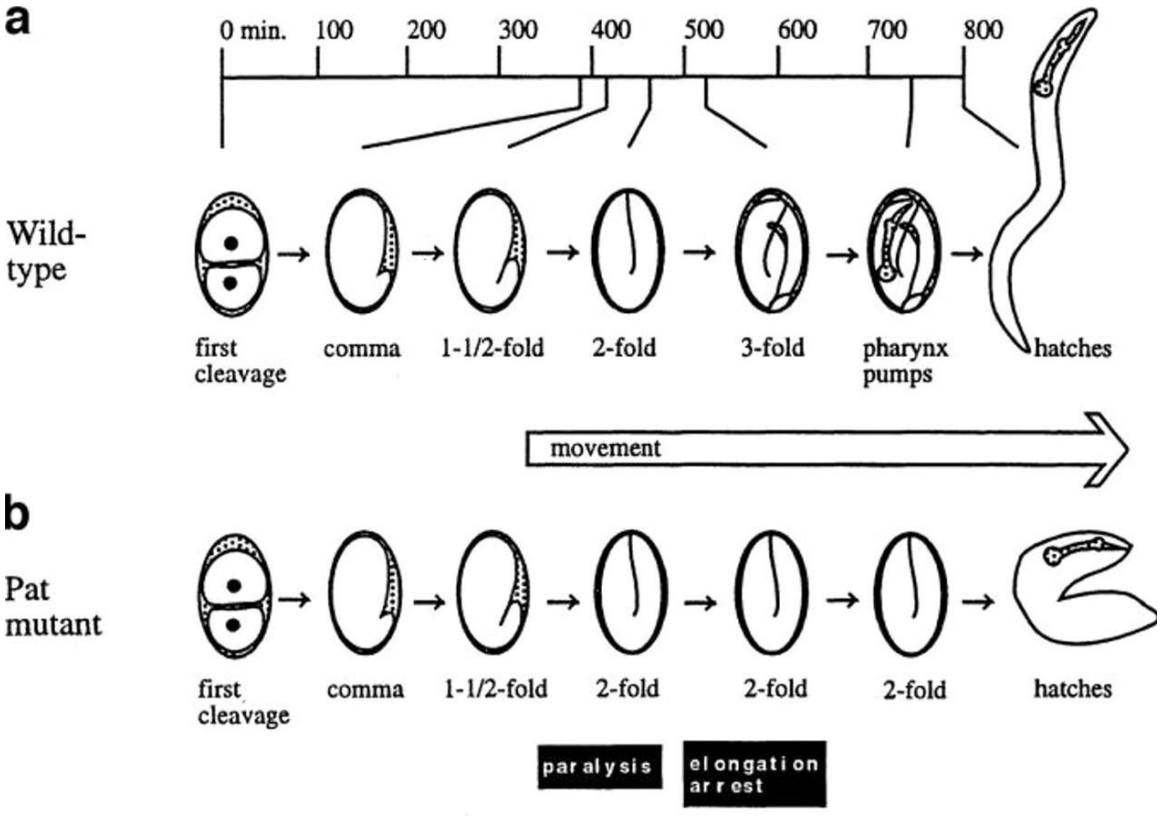


Figure 1–10. Embryonic Development of Wild Type and Pat Mutants.

The figure on the top shows the normal embryonic development of *C. elegans*. The initial football shaped embryo undergoes a dramatic shape change and elongation into something that looks more and more like a tube-like worm, progressing through stages descriptively labeled comma, 1.5-fold, 2-fold and 3-fold, before hatching from the eggshell. Normally, the embryo begins to move at the 1.5 fold stage, and then moves vigorously within the eggshell at the 2-fold stage. The bottom shows the embryonic development of pat mutants. "Pat" mutant embryos do not move in the eggshell at the 1.5-fold stage and arrest development at the 2-fold stage. Other aspects of development continue and the embryo hatches as an abnormal L1 larva and dies (Adapted from Williams et al., 1994).

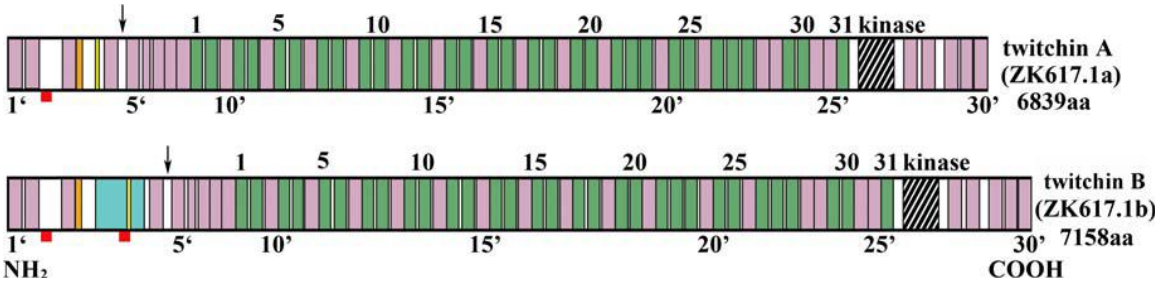


Figure 1–11. Schematic Representation of Protein Domain Organization in *C. elegans* Twitchin.

There are two isoforms of twitchin (A and B). Purple bars represents Ig domains. Green bars represents Fn3 domains. Orange region is predicted to be coiled coil and kinase domains are represented by striped white boxes. Yellow boxes represents a sequence of 25 amino acid residues and there is a blue box on either side of this sequence that represents a unique region in isoform B. The red bars represent the predicted NORS regions. The arrow indicates the glycine-rich segment between Ig domains 4’ and 5’ (adapted from Ferrara TM et al., 2005).

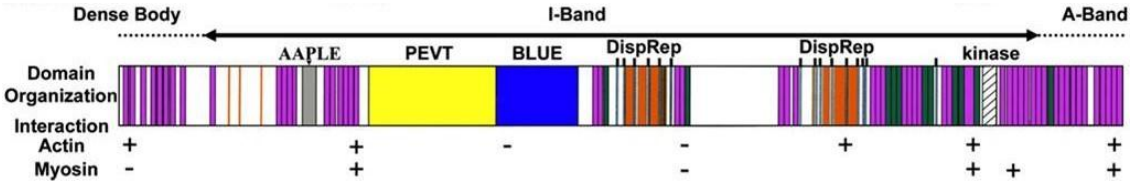


Figure 1–12. Schematic Representation of Domains in the 2.2 MDa TTN-1 Polypeptide of *C. elegans.*

Purple bars represent Ig domains, green represent Fn3 domains, yellow bar represents the 2400 amino acid PEVT/K domain, the blue box represents a 1500 amino acid region composed of tandem repeats and predicted coiled-coil structure, gray bar represents the 250 amino acid region made up of AAPLE repeats, orange bar represents sequence predicted to form coiled-coil structure, and the projecting lines represent the 30 amino acid dispersed repeat (15 copies). There is a protein kinase domain (represented by a rectangle with black stripes) near the C-terminus. As depicted, immunofluorescence microscopy indicates that a single TTN-1 polypeptide is anchored at the dense body, extends the length of the I-band and extends into the outer edge of the A-band. Immediately below the schematic are indicated whether or not (+ or −) a given region was able to bind to F-actin or myosin *in vitro* (Forbes et al. 2010).

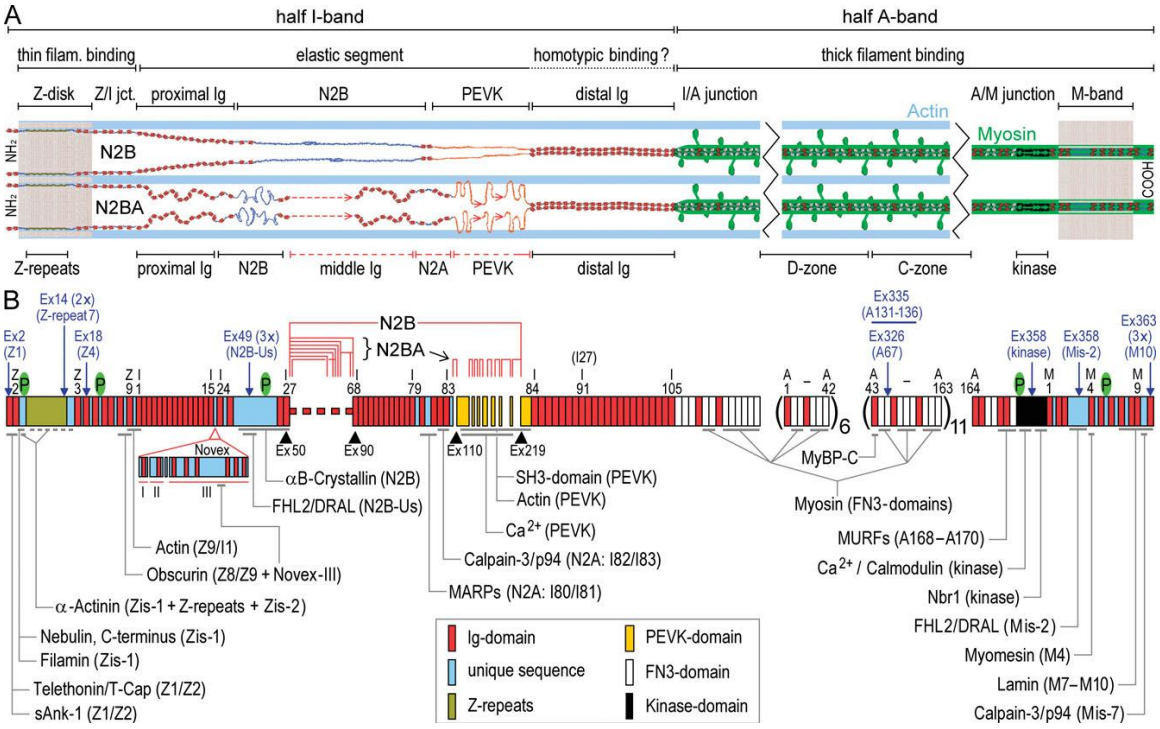


Figure 1–13. Diagram of Human Titin Domain Structure and Binding Partners.

(A) The picture shows the localization of cardiac titin isoforms (N2B and N2BA) in the half sarcomere. Dashed red lines indicate differentially spliced segments in I-band portion of titin. (B) The domain arrangement of titin. (Red rectangles are Ig domains, clear rectangles are Fn3 domains; black represents the protein kinase domain near its C-terminus; yellow represents the main elastic element in titin called PEVK.) The binding sites are shown in parentheses with each binding partner. Red lines show the splice pathways between exon 50 and exon 219. Blue arrows point to the mutations found in titin. P with green highlight represents the phosphorylation site in titin (taken from Linke, 2008).

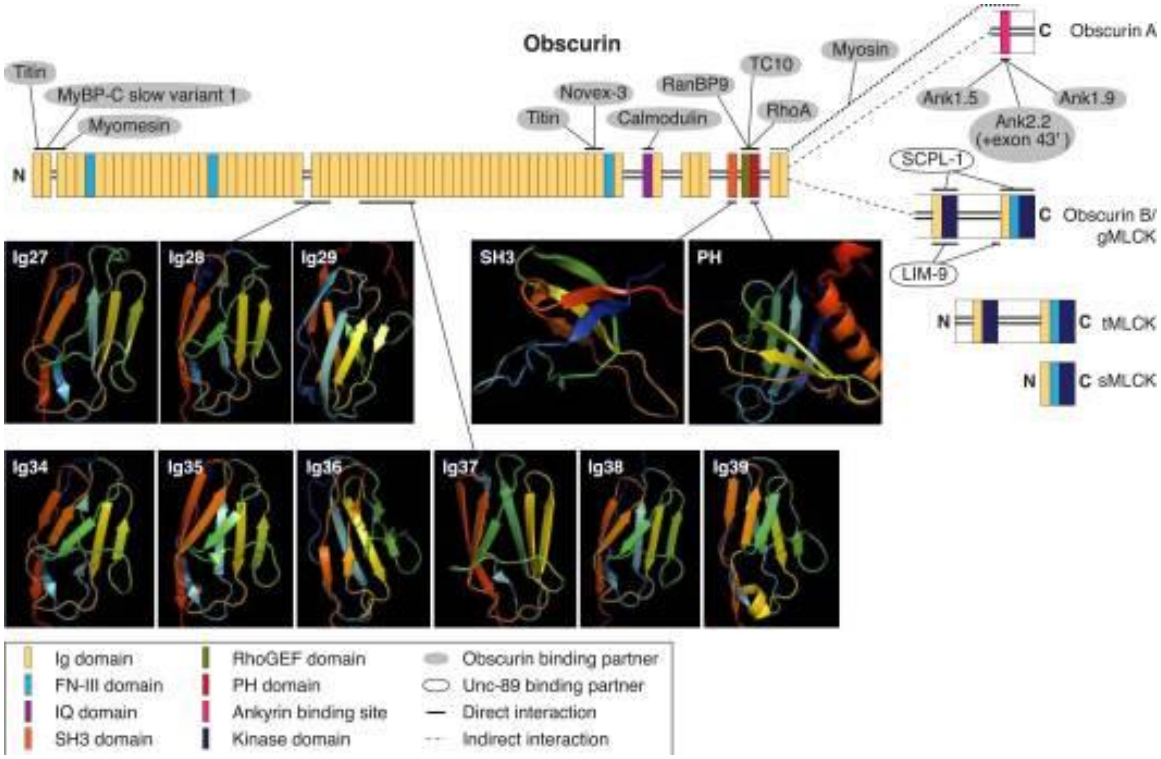


Figure 1–14. Diagram of Obscurin Domain Structure and Binding Partners

There are four isoforms of Obscurin : Obscurin A, Obscurin B/gMLCK (giant MLCK), tMLCK (tandem MLCK) and sMLCK (single MLCK). The C terminus of (Kontrogianni-Konstantopoulos et al., 2009) Obscurin A binds to ankyrin and the C terminus of Obscurin B has two kinase domains. There are 11 domains within obscurin that have been characterized by either NMR or X ray crystallography (Ig 27, 28, 29, 34, 35, 36, 37, 38 and 39, SH3 and PH). The binding partners are highlighted with grey ovals. The solid lines represent well defined interactions whereas the dot line refers to the unconfirmed interaction with myosin. Note that these authors show our lab's results demonstrating interaction of UNC-89, the nematode homolog of obscurin, with SCPL-1 and LIM-9. (copied from Kontrogianni-Konstantopoulos et al., Physiological Reviews, 2009, Am Physiol Soc, used with permission).

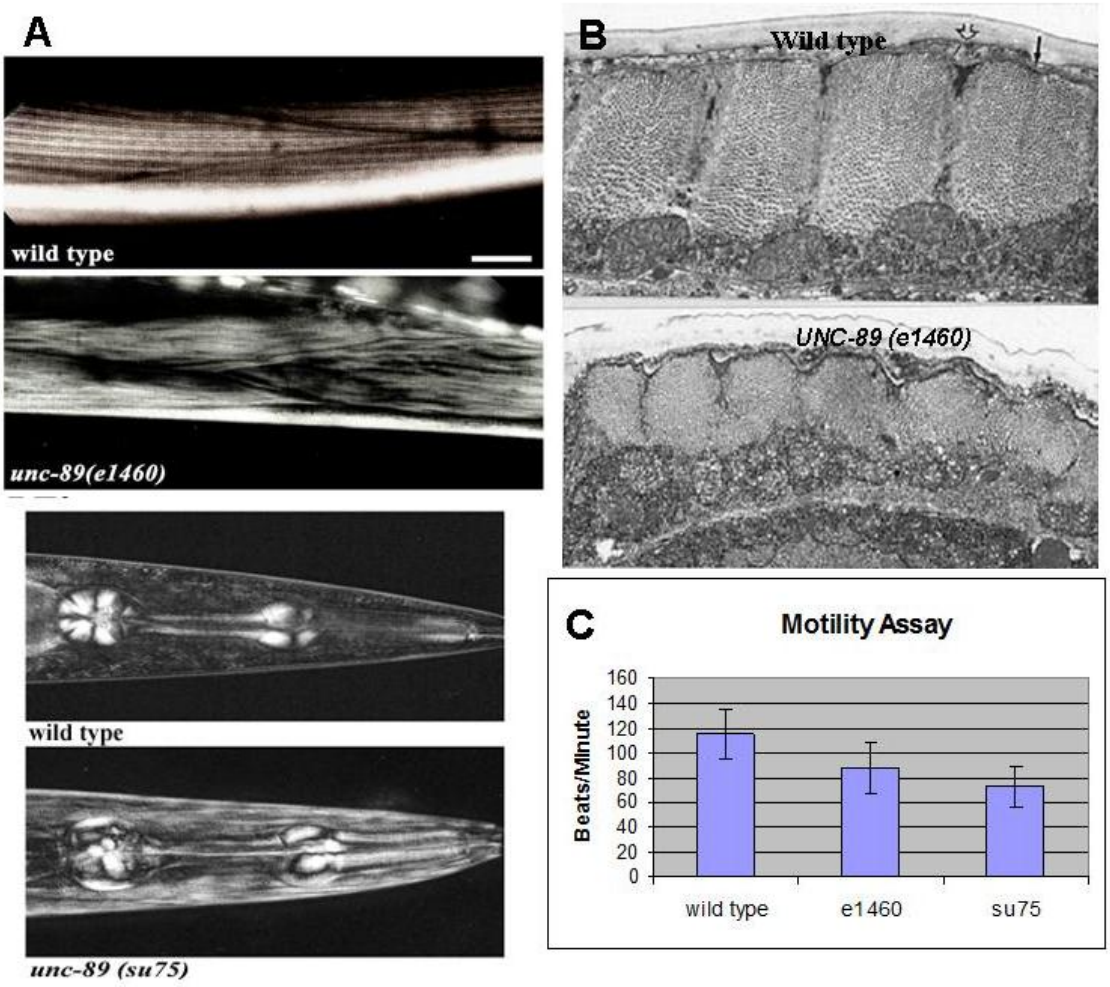


Figure 1–15. The phenotype of *unc-89* mutants in *C. elegans*.

(A) Images of wild type and *unc-89* mutant muscle obtained with polarized light microscopy. The top panel shows the disorganization of the myofilament lattice in the body wall muscle of *unc-89 (e1460)* mutants. Instead of the highly organized pattern of alternating A- and I-bands of wild type muscle, *unc-89(e1460)* shows a "basket weave" pattern. The bottom panel shows the disorganization of the single sarcomere pharyngeal muscle in *unc-89 (su75)* mutants. Instead of the radially oriented muscle filaments of wild type, unc-89(su75) shows misplaced muscle filaments at the circumference of the

pharyngeal terminal bulb. (B) Transmission EM images of cross sections of body wall muscle from wild type and unc-89(e1460). In unc-89(e1460), the entire myofilament lattice is thinner than normal, and M-lines are missing. The solid arrow points to an M-line and the hollow arrow points to a dense body in wild type muscle. (C) Motility assay shows that *unc-89* mutants move somewhat slower than wild type worms. These data were obtained by previous Benian lab students, Tina Tinley and Tracey (Small) Ferrara.

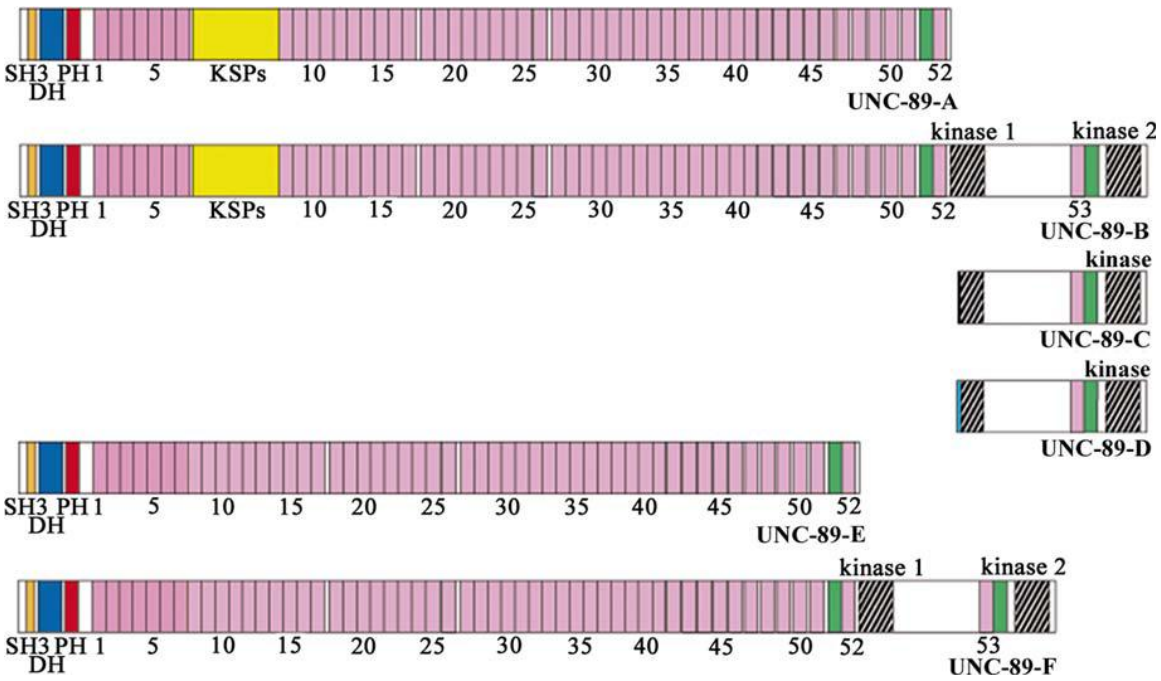


Figure 1–16. Schematic Representation of UNC-89 Isoforms in *C. elegans*.

Ig domains are purple; Fn3 domains are green; kinase domains or partial kinase domains are striped boxes; SH3, DH and PH domains are marked; the KSP region is yellow. There are six isoforms of UNC-89. Isoform A and E do not contain kinase domains whereas isoforms C and D only contain kinase domains, interkinase region and one Ig and one Fn3 domain (from Ferrara et al., 2005).

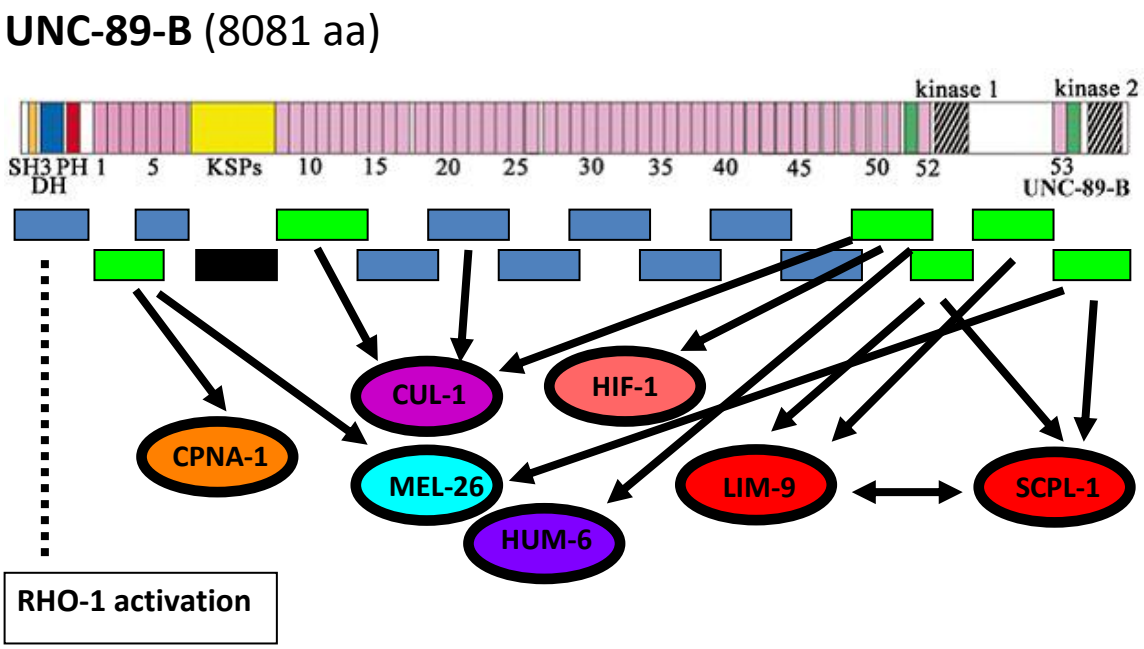


Figure 1–17. Summary of the Benian Lab Progress in Identifying Proteins that Interact with UNC-89.

Schematic presentation of domain organization of the largest UNC-89 isoform is shown on the top. The rectangles below the schematic represent 16 consecutive regions of UNC-89 placed into bait and prey vectors to perform yeast two-hybrid screens. Seven binding partners have been discovered through these screens (CPNA-1, MEL-26, CUL-1, HIF-1, HUM-6, LIM-9 and SCPL-1). RHO-1 was discovered by other approaches. All eight proteins are evolutionarily conserved. All proteins except HUM-6 have been localized by antibodies to sarcomeric M-lines, co-localizing with UNC-89. All proteins except for HIF-1 and HUM-6 have clear loss of function mutant phenotypes in muscle. The black bar represents the KSP domain, which upon screening yielded no binding partners. The green bars represent regions that have been used to screen yeast 2-hybrid libraries. The following individuals have contributed to this large-scale project: Hiroshi Qadota, Guy Benian, myself, Kristy Wilson and M. Berenice Duran.

2. Chapter 2: A LIM-9 (FHL) / SCPL-1 (SCP) Complex Interacts with the C-terminal Protein Kinase Regions of UNC-89 (Obscurin) in *C. elegans* Muscle

Ge Xiong[1,2], Hiroshi Qadota[1], Kristina B. Mercer[1], Lee Anne McGaha[1,2], Andres F. Oberhauser[3], and Guy M. Benian[1*]

[1]Department of Pathology, Emory University, Atlanta, Georgia 30322

[2]Graduate Division of Biological and Biomedical Sciences, Emory University, Atlanta, Georgia 30322

[3]Department of Neuroscience and Cell Biology, University of Texas Medical Branch, Galveston, Texas 77555

[*]Corresponding author

e-mail address of the corresponding author: pathgb@emory.edu

Running head: LIM-9 and SCPL-1 interact with UNC-89

Key words: *C. elegans*, obscurin, FHL, CTD phosphatase, muscle

Published in J Mol Biol, 386(4), 2009, 976-88

(abstract)

The *C. elegans* gene *unc-89* encodes a set of mostly giant polypeptides (up to 900 kDa) that contain multiple Ig and Fn3, a triplet of SH3-DH-PH, and two protein kinase domains. The loss of function mutant phenotype and localization of antibodies to UNC-89 proteins, indicate that the function of UNC-89 is to help organize sarcomeric A-bands, especially M-lines. Recently, we reported that each of the protein kinase domains interact with SCPL-1, which contains a CTD type protein phosphatase domain. Here, we report that SCPL-1 interacts with LIM-9 (FHL), a protein that we first discovered as an interactor of UNC-97 (PINCH) and UNC-96, components of an M-line costamere in nematode muscle. We also show that LIM-9 can interact with UNC-89 through its first kinase domain and a portion of unique sequence lying between the two kinase domains. All the interactions were confirmed by biochemical methods. A yeast three-hybrid assay demonstrates a ternary complex between the two protein kinase regions and SCPL-1. Evidence that the UNC-89/SCPL-1 interaction occurs in vivo was provided by showing that overexpression of SCPL-1 results in disorganization of UNC-89 at M-lines. We suggest two structural models for the interactions of SCPL-1 and LIM-9 with UNC-89 at the M-line.

2.1 Introduction

The fundamental unit of muscle contraction, the sarcomere, is an exquisitely organized assemblage of hundreds of different proteins. To gain insight into the functions of individual sarcomeric proteins, their interactions and the assembly and maintenance of the sarcomere, a number of laboratories profitably utilize the roundworm *C. elegans*

(Moerman 1997, 2006). *C. elegans* is an excellent model genetic organism due to its simple anatomy, sequenced genome and powerful forward and reverse genetics. Since sarcomere architecture and components are largely conserved, results from *C. elegans* can be extrapolated to other organisms, including man. Sarcomeres contain a number of unusually large polypeptides (0.7—4 MDa in molecular mass) related to mammalian titin (Tskhovrebova 2003). Members of the titin family consist primarily of multiple copies of immunoglobulin (Ig) and fibronectin type 3 (Fn3) domains, and one or even two protein kinase domains. *C. elegans* striated muscle of the body wall contains three such polypeptides (Ferrara 2005): twitchin (754,000 Da) located in the A-band (Benian 1989, 1993; Moerman 1988), TTN-1 (2.2 MDa) located in the I-band (Flaherty 2002), and UNC-89 (up to 900,000 Da) located at the M-line (Benian 1996; Small 2004). Loss of function *unc-89* mutants display disorganized myofibrils, especially at the A-band, and usually lack M-lines (Benian 1999; Waterston 1980). *unc-89* is a complex gene: through the use of three promoters and alternative splicing, at least 6 major polypeptides are generated, ranging in size from 156,000 to 900,000 Da (Ferrara 2005; Small 2004). The largest of these isoforms, UNC-89-B and UNC-89-F, which are each ~900,000 Da, consist of 52 Ig domains, 2 Fn3 domains, a triplet of SH3, DH and PH domains near their N-termini, and two protein kinase domains (called PK1 and PK2) near their C termini. Most recently, we have demonstrated that the DH/PH region of UNC-89 has guanine nucleotide exchange factor activity towards RHO-1 (RhoA in *C. elegans*) and this is important for organization of myosin thick filaments in nematode body wall muscle cells (Qadota 2008b). Although neither kinase domain has thus far demonstrated catalytic

activity, molecular modeling suggests that PK2 is active, while PK1 is inactive (Small 2004).

For several reasons, we are very interested in exploring the structures and functions of the kinase domains of the giant titin-like proteins. First, a mutation in the human titin protein kinase domain causes hereditary muscle disease (Lange 2005). Second, in *C. elegans*, two mutant alleles of *unc-89* that lack the kinase domain region result in disorganization of myofibrils (Ferrara 2005). Third, the normally autoinhibited protein kinase domains of these giant proteins may be activated by mechanical forces that occur during muscle contraction (Grater 2005; Greene 2008; Puchner 2008). One way to gain insight into the function of a protein is to identify its binding partners. We reported recently (Qadota 2008a) that the two protein kinase domains of UNC-89 interact with SCPL-1 (small CTD phosphatase-like-1), which contains a CTD type phosphatase domain. By antibody staining SCPL-1 co-localizes with UNC-89 at the M-line. Knockdown of *scpl-1* mRNA results in no detectable phenotype in body wall muscle, but it does result in a defect in the function of egg-laying muscles.

The human homolog of UNC-89 is called obscurin (Fukuzawa 2005; Young 2001)(Young, Ehler et al. 2001; Fukuzawa, Idowu et al. 2005). Although obscurin contains all the same domains as UNC-89, the SH3, DH and PH domains are located near the C terminus rather than near the N terminus as they are in UNC-89. Another difference is that although UNC-89 is located only at the M-line, various obscurin isoforms are located at either the M-line, the A/I junction or at the Z-disk and Z/I junction (Bowman 2007). Progress has also been made on finding binding partners for obscurin: Ig48 and Ig49 within the C-terminal third of obscurin interact with specific Ig domains within the

Z-disk portion of titin (Young 2001) and with a C-terminal portion of the 700 kDa isoform of titin called “novex-3 titin”(Bang 2001). A sequence of 25 residues located near the C terminus of obscurin interacts with a small isoform of ankyrin 1, and this is postulated to provide a link between myofibrils and the sarcoplasmic reticulum (Bagnato 2003; Kontrogianni-Konstantopoulos 2006, 2003). Overexpression of part of the C terminus of obscurin in primary skeletal myotubes results in disorganization of the A-band, and co-immunoprecipitation experiments suggest that obscurin interacts with myosin (Kontrogianni-Konstantopoulos 2004). This indicates that obscurin, like UNC-89, has a role in the assembly or maintenance of A-bands. The DH domain of obscurin has been shown to interact with the scaffolding protein called Ran binding protein 9 (RanBP9), and both obscurin’s DH domain and RanBP9 interact with titin to regulate titin’s incorporation into Z-disks (Bowman 2008). Most recently, the N-terminal most Ig domains (1-3) of obscruin have been shown to interact with the C-terminal most Ig domain of titin, and also with the M-line protein myomesin (Fukuzawa 2008). Whether nematode UNC-89 and mammalian obscurin have the same or even overlapping interacting partners is unknown.

To understand how nematode SCPL-1 carries out its function, we attempted to find additional proteins that SCPL-1 interacts with, and we discovered that SCPL-1 interacts with LIM-9 (FHL in mammals). We also found that LIM-9 can interact with UNC-89 directly, via the first kinase domain region (Fn3-Ig-PK1) and a portion of the interkinase region. We present a model that at the M-line, both SCPL-1 and LIM-9 interact with two portions of UNC-89 and function as bridges to stabilize the interaction of two UNC-89

polypeptides of opposite polarity, and/or to stabilize a loop created by interaction of the two kinase regions within one UNC-89 polypeptide.

2.2 Results

Given that we had determined that UNC-89 interacts with SCPL-1, we wished to determine which additional proteins SCPL-1 might interact with to carry out its function. Therefore, we used SCPL-1 to screen the 16-protein two-hybrid bookshelf of proteins known to reside at the nematode M-line (Qadota 2007). We discovered a 2-hybrid interaction between either SCPL-1a or -1b as bait and LIM-9 as prey (Fig 2-1A). LIM-9 contains one PET domain and six LIM domains, and is the closest worm homolog of mammalian FHL2. LIM-9 was characterized initially as an interactor of UNC-97 (PINCH) and UNC-96, and resides at least partially at M-lines (Qadota 2007). Domain mapping indicates that the minimum region of LIM-9 required for interaction with either SCPL-1a or -1b consists of the 6 LIM domains of LIM-9 (Fig2-1A). Further, we can show that to obtain interaction with LIM-9, only the N-terminal, non-phosphatase region of SCPL-1b is sufficient, whereas the entire SCPL-1a protein is required (Fig 2-1B).

We confirmed the interaction between SCPL-1 and LIM-9 using two methods. First, we showed that yeast-expressed myc-tagged SCPL-1 can interact with a bacterially-expressed maltose binding protein (MBP) fusion of LIM-9 (LIM domains only). A total protein lysate from yeast expressing myc-SCPL-1a or –SCPL-1b was prepared and incubated with agarose beads coated with antibodies to myc, pelleted, washed, and incubated with MBP-LIM-9 or MBP. After pelleting and washing the beads, the proteins were eluted in Laemmli buffer and portions were separated on 3 gels, blotted and reacted

with either antibodies to myc or antibodies to MBP. As shown in Fig2-2A, as expected, both myc-SCPL-1a (top panel) and myc-SCPL-1b (middle panel) were precipitated. Reaction to anti-MBP (bottom panel) revealed that both myc-SCPL-1a and -1b co-precipitated MBP-LIM-9 but not MBP. Second, we showed that purified SCPL-1a interacts with purified LIM-9 (LIM domains only) by a far western assay (Figure2-2B). GST and GST-SCPL-1a was separated on SDS-PAGE, transferred to membrane, and incubated with either MBP, or MBP fused to LIM-9 (LIM only). The binding reactions were visualized using antibodies to MBP followed by ECL. As shown in Figure 2B, MBP-LIM-9, but not MBP, interacts with SCPL-1a. Consistent with this direct interaction between SCPL-1 and LIM-9, we had found previously that each protein resides at M-lines (Qadota 2008a, 2007).

Given that LIM-9 is located at the M-line, we wondered if LIM-9 might interact directly with UNC-89. Consistent with this expectation, screening of the M-line bookshelf with Fn3-Ig-PK1 and interkinase regions of UNC-89 revealed that each interacts with LIM-9. We utilized 2-hybrid analysis to determine which portions of each protein that were involved in these interactions. As shown in Fig 2-3A, the smallest region of LIM-9 required to interact with either the Fn3-Ig-PK1 or interkinase regions of UNC-89 contains all 6 LIM domains, the identical region of LIM-9 required for interaction with SCPL-1 (Fig 2-1A). Note that full-length LIM-9 does not interact with SCPL-1, PK1 or interkinase regions. As shown in Fig 2-3B, in order to interact with LIM-9, the kinase domain of PK1 is not sufficient (segment FB): the upstream Fn3 and Ig domains are also required. This larger segment of UNC-89, Fn3-Ig-PK1 kinase domain, had also been shown (Qadota 2008a) to be required for interaction with SCPL-1 (Fig 2-

9A). As indicated in Fig 2-3C, a smaller segment of the interkinase region of UNC-89 is required for interaction with LIM-9: as little as the C-terminal 200 residues of interkinase are sufficient (segment LD).

To confirm the 2-hybrid interactions of LIM-9 with the PK1 and interkinase regions of UNC-89, we used a similar assay as described above with yeast expressed PK1 or interkinase with bacterially expressed LIM-9. Yeast expressed PK1 AB or FB segments (described in Fig 2-3B) with HA tags were immunoprecipitated with agarose beads coated with antibodies to HA and tested for binding to either MBP-LIM-9 or MBP. As demonstrated in Figure 4A, consistent with the 2-hybrid results, AB fragment (Fn3-Ig-PK1 kinase) brought down much more MBP-LIM-9 than FB fragment (PK1 kinase alone), and neither segment brought down any detectable MBP. A similar strategy was used to verify interaction between LIM-9 and the interkinase region of UNC-89. In this case, yeast was used to express an HA tagged segment of the interkinase region (segment KD indicated in Fig 2-3C), which also interacts with LIM-9 by 2-hybrid. As shown in Figure 4B (left panels), the full-length KD fragment could be expressed and immunoprecipitated from yeast (upper band detected with anti-HA) although some apparent degradation bands were also present. Notably, this HA-KD protein was able to bring down MBP-LIM-9 but not MBP.

Given that the C-terminal 200 residues of the total 647 residues of the interkinase sequence were sufficient for binding to LIM-9, we wondered if there was anything special about the amino acid sequence or predicted structure of this binding region. As shown in Figure 5, the C-terminal 144 residues are different in several respects from the N-terminal 503 residues. In particular, residues 1-503 (italicized in Fig 2-5) is predicted

to be a “NORS” region (Liu 2003), that is, a sequence with little to no secondary structure and possibly structurally flexible (Liu 2002). Residues 1-503 are also predicted to have high solvent accessibility (80%), and have a high probability for assigning neither a-helix, nor b-strand (85%). Residues 1-470 contain 5 segments ranging from 10 to 21 residues predicted (Wootton 1996)(Wootton and Federhen 1996) to be “SEG low-complexity regions” (double underlined in Fig 2-5). Moreover, the amino acid composition of residues 1-503 and 504-647 are quite different: in 1-503, there is especially increased frequency of P, E, Q and D; consecutive Ps (up to 3) or Qs (up to 4) are found only in the N-terminal 453 residues (especially in residues 1-250). Therefore, sequence analysis suggests that the N-terminal 450-500 residues of the interkinase might be highly elastic, whereas the C terminal 150-200 residues are different, possibly forming a domain that is suited to interact with LIM-9.

If LIM-9 interacts with UNC-89 in vivo, we would expect that the proteins co-localize in the sarcomere. Previously, we had shown that antibodies to LIM-9 localize to M-lines and to the I-band around and between dense bodies (Qadota 2007). We have also shown that antibodies to three different regions of UNC-89, including the interkinase region, localize to M-lines (Small 2004). Thus, we co-stained adult worms with anti-LIM-9 and anti-interkinase region of UNC-89. As demonstrated in Figure 6, there is partial co-localization between the interkinase region of UNC-89 and LIM-9 at the M-line (note the white dots at the bottom of the figure indicate co-localization).

Previously, we reported that SCPL-1 interacts with either Ig-Fn3-PK2 or Fn3-Ig-PK1 (Qadota 2008a). As described here, LIM-9 interacts with either Fn3-Ig-PK1, the C-terminus of the interkinase region, or SCPL-1 (Fig 2-9A). Additionally, we could not find

direct interactions among Fn3-Ig-PK1, interkinase and Ig-Fn3-PK2 (data not shown). One hypothesis for the function of SCPL-1 binding to the two kinase domain regions is that in the sarcomere, SCPL-1 acts as a molecular bridge to bring together or stabilize two UNC-89 molecules or to stabilize a loop created by interaction of the two kinase regions within one UNC-89 molecule (Fig 2-9 B and C). This idea was first tested by a yeast 3-hybrid assay. As shown in Fig 2-7, when Fn-3-Ig-PK1 bait was presented to Ig-Fn3-PK2 prey, no growth of the yeast could be detected on adenine deficient media, indicating no detectable interaction (top row in Fig 2-7). However, when this bait and prey combination was expressed in the same yeast cell with myc-tagged SCPL-1 (either -1a or -1b isoforms), growth was detected on media missing adenine, indicating a three-way interaction (middle and bottom rows in Fig 2-7). Thus, the three hybrid results are consistent with SCPL-1 acting as a molecular bridge to bring together the PK1 and PK2 regions of UNC-89.

To obtain *in vivo* evidence for SCPL-1 functioning as a molecular bridge, we examined the muscle phenotype of worms overexpressing SCPL-1. We created several independent lines of wild type worms carrying extrachromosomal arrays containing a plasmid consisting of myc-tagged SCPL-1b full-length cDNA under the control of a heat shock promoter. After inducing expression by heat shock, worms were fixed and dually stained for anti-myc and anti-IK, the previously-described (Small 2004) antibody generated to the interkinase region of UNC-89. As shown in Fig 2-8, in body wall muscle cells not expressing the transgenic myc-SCPL-1b, the pattern of anti-IK staining is normal: M-lines in parallel. However, in body wall muscle cells in which SCPL-1b is mildly or highly overexpressed, there are accumulations of myc-SCPL-1b, and near these areas of

the myofilament lattice, the M-lines detected by anti-IK are broken, diffuse or even missing. This result is consistent with our model that SCPL-1 acts as a linker or bridge between neighboring UNC-89 polypeptides at their kinase domain-containing C-terminal portions (Fig 2-9 B and C).

2.3 Discussion

In this study, we provide evidence that a complex of two proteins, LIM-9 (FHL) and SCPL-1 (SCP) interact with the C-terminal protein kinase domain-containing portion of the largest isoforms of UNC-89 (obscurin). Interactions were identified by the two-hybrid method and then confirmed by showing interaction between purified components in vitro, and demonstrating co-localization in sarcomeres. Both LIM-9 and SCPL-1 were shown to interact with two portions of UNC-89 in ternary complexes. Moreover, support for the in vivo relevance of these interactions was provided by showing that overexpression of SCPL-1 results in abnormal localization of UNC-89, a key component of sarcomeric M-lines.

We previously reported that SCPL-1 interacts with the two protein kinase domains of UNC-89 (Qadota 2008a), and reported first identification of LIM-9 as a component of the integrin-associated protein complex at M-lines, and that the closest homolog for LIM- 9 in mammals is “four and one-half LIM protein”, FHL (Qadota 2007). Here, we show that SCPL-1 interacts with LIM-9, and that LIM-9 interacts with UNC-89. In fact, our finding that UNC-89 interacts with LIM-9, now allows us to link the giant multi-domain protein UNC-89 to the integrin associated protein complex at the M-line (Fig 2-10)—a continuous linkage of proteins from the extracellular matrix to the thick filaments

(Qadota 2007; Lin 2003; Mackinnon 2002; Miller 2006; Norman 2007). Moreover, although the largest isoform of UNC-89 (UNC-89B) is enormous and contains many domains, it is likely that LIM-9 and SCPL-1 only interact with the protein kinase domain containing regions of UNC-89: Using LIM-9, SCPL-1a or SCPL-1b as either 2-hybrid bait or prey to screen 16 fragments that comprise the largest UNC-89 isoform, we failed to identify interaction with any other portions of UNC-89 (results to be described elsewhere).

What is the function of LIM-9 in nematode muscle? Although two of the mammalian homologs of LIM-9 called FHL-2 and FHL-3 have been shown to interact with the cytoplasmic tails of several α- and β-integrins (Samson 2004; Wixler 2000), we have reported that LIM-9 failed to interact with the cytoplasmic tail of β-integrin (Qadota 2007). Nevertheless, LIM-9 was shown to interact with UNC-97, which is part of a 4 protein complex (including PAT-6 (actopaxin), PAT-4 (integrin linked kinase) and UNC-112(Mig-2)) that does associate with the tail of β-integrin. Thus, in both mammals and in nematodes, FHL proteins are linked to integrins, either directly or indirectly through integrin associated proteins.

Although knockout alleles of *lim-9* have no obvious phenotype (Qadota 2007), mutations in one of the mammalian homologs of LIM-9 called FHL1 has been associated with three types of X-linked human myopathies (Quinzii 2008; Windpassinger 2008; Schessl 2008). FHL1 is just one member of a family of four mammalian FHL proteins. Whereas FHL1 is ubiquitously expressed, FHL2 is primarily expressed in the heart (Scholl 2000) and localized to Z-disks and also to M-lines. FHL3 is localized to Z-disks in skeletal muscles (Bowman 2007). Recently, Cowling et al. have shown that

overexpression of FHL1 in transgenic mice results in both skeletal muscle hypertrophy and enhancement of myoblast fusion *in vitro*; in addition, these authors showed that FHL1 interacts with NFATc1 and stimulates the transcription factor activity of NFATc1, probably leading to increased transcription of genes that result in increases in myofibrillar mass (Cowling 2008). Interestingly, our finding that LIM-9/FHL interacts with a giant protein kinase (in our case, UNC-89) is not unprecedented: FHL2 was found primarily to interact with 2 regions of mammalian titin—the N2B region, a heart-specific sequence of the I-band portion of titin, and the is2 region in the M-line portion of titin (Lange 2002). These authors also showed that several muscle specific isoforms of metabolic enzymes such as creatine kinase, adenylate kinase and phosphofructokinase also interact with FHL2. Thus, our finding that LIM-9 interacts with both the first kinase domain of UNC-89 (PK1) and the SCPL-1 phosphatase adds two more enzymes to those that can associate with an FHL protein.

What is the function of the associations of LIM-9 and SCPL-1 with UNC-89? Previously, we reported that SCPL-1, under the conditions tested, was not a substrate for phosphorylation by the UNC-89 PK2 kinase domain (Qadota 2008a). We note here that LIM-9 also could not be phosphorylated by UNC-89 PK2 (data not shown). Nevertheless, we propose two alternative models (Figure 9B and C) for the function of a LIM-9 / SCPL-1 complex in stabilizing interactions between two UNC-89 molecules, or between regions of one UNC-89 molecule, to form a loop at the C-terminal part of UNC-89. Support for the first model, that is, that an SCPL-1 / LIM-9 complex acts as a bridge between two adjacent UNC-89 polypeptides was provided by our finding that overexpression of SCPL-1 disrupts the organization of UNC-89 at the M-line (Fig 2-8).

Perhaps excessive levels of SCPL-1 interfere with a ternary protein complex composed of SCPL-1 / LIM-9 / UNC-89, and this leads to disruption of the normal linkages of UNC-89 molecules through their kinase domain regions.

As shown in Fig 2-11, RNAi knockdown for either *lim-9*, *scpl-1*, or both *lim-9* and *scpl-1* together, have no affect on the localization or organization of UNC-89. Given our model, we expected that such a "double mutant" might have a phenotype similar to *unc-89* mutant alleles that lack the kinase domain-containing isoforms, UNC-89-B and UNC-89-F (alleles *st79* and *tm752*, Ferrara 2005). Because we observed normal localization of UNC-89 in the *lim-9; scpl-1* double RNAi, it suggests yet an additional protein or proteins might link the kinase domain-containing portions of UNC-89. However, screening of the two-hybrid library with the interkinase domain of UNC-89 revealed no new interactors (data not shown). In the future we will attempt co-immunoprecipitation of UNC-89 with its binding partners followed by mass spectrometry to perhaps reveal these missing proteins.

The sequence of the interkinase region, in particular, the N-terminal 450-500 residues, suggests physical properties that are consistent with either model. This 450-500 amino acid region is predicted to have very little secondary structure, low sequence complexity and high proline content (Fig 2-5). These features suggest that this portion of the interkinase region is structurally flexible or even elastic. Flexibility would facilitate an intramolecular loop (the second model). It is also possible that elastic interkinase regions of two oppositely oriented UNC-89 molecules linked together by SCPL-1 and LIM-9 might allow a certain degree of structural expansion and contraction of the M-line. By two hybrid, the LIM-9 binding site has been mapped to the C-terminal 200 residues of the

interkinase region (Fig 2-3). Again, sequence analysis is compatible with this result in that this sequence is different from the proceeding 450-500 residues, in that it is predicted to have some secondary structure, a lack of low sequence complexity regions and near normal proline content.

2.4 Materials and Methods

Worm Strains and Culture

Bristol N2 was the wild-type strain, grown at 20°C on NGM plates with *Escherichia coli* strain OP50 as food source (Brenner 1974).

Yeast Two-hybrid and Three-hybrid Screens and Assays

The yeast two-hybrid method used to screen the nematode M-line bookshelf and to test for protein-protein interactions was performed as described (Qadota 2007; Mackinnon 2002). The DNA sequences corresponding to fragments of a given protein were amplified by using synthesized primers and cDNA library template or plasmid templates isolated from the library during two-hybrid screening as described previously (Qadota 2007). The yeast three-hybrid assay was performed also as previously described (Lin 2003).

Plasmid Construction

To construct yeast 2-hybrid plasmids for the interkinase region of UNC-89, we amplified the various portions of this region as depicted in Figure 3C using primers YIKC, YIKD, YIKJ, YIKK and YIKL as shown in Table 1. After cloning into pBluescript and verifying error-free sequences, the fragments were re-cloned into pGBDUC1 and pGADC1. The

KD fragment from the pBluescript clone was also re-cloned into pGEXKK1 for expression of a GST-KD fusion protein. The method used to construct plasmids for yeast expression of myc-tagged SCPL-1a and b was described (Qadota 2008a). To express hemagglutinin (HA)-tagged UNC-89 fragments in yeast, PCR was performed to amplify UNC-89 fragments by using the primers YIKK and YIKD for interkinase region (fragment KD in Figure 3C). The cDNA sequence was cloned into pKS-HA8 (Nhex2) (three HA-tagged vectors) by using SmaI and XhoI sites. HA-tagged UNC-89 interkinase region (KD fragment) was excised from this construct using NheI and inserted into pGAP-C-Nhe (yeast expression vector, TRP1 marker) by using the NheI site of the vector. Construction of the plasmid for expression of HA-tagged AB fragment of the PK1 region of UNC-89 was described (Qadota 2008a). The plasmid used for expression of HA-tagged FB fragment of the PK1 region of UNC-89 was made by using primers U89-PK1-F and U89-PK1-B to PCR amplify and clone into pKS-HA8(Nhex2) (three HA-tagged vector) by using EcoRV and SalI sites.

Purification of Bacterially expressed proteins

The procedures for purification of GST or MBP fusion proteins were described previously (Mercer 2006).

Assays to Confirm Interaction between Proteins

To confirm a protein-protein interaction detected by the yeast two-hybrid assay, yeast-expressed myc- or HA-tagged proteins were immunoprecipitated and tested for binding with bacterially expressed MBP-tagged proteins. To verify the interaction between

SCPL-1 and LIM-9, myc-tagged SCPL-1a or -1b was expressed in yeast, and lysates were prepared as described (Qadota 2008a). Myc-tagged SCPL-1a or -1b was immunoprecipitated from yeast lysates containing 300 mg of total protein in a volume of 500 μl for 1 h with shaking at 4^{o}C in immunoprecipitation (IP) buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 1 mM EGTA, 1 mM EDTA, 0.25% gelatin, complete mini protease inhibitors, and 0.1% NP-40) using 50 μl of a 1:1 slurry of agarose beads conjugated to rabbit anti-myc antibodies (Sigma-Aldrich, cat. no. A7470). The beads were pelleted and washed 3X with IP buffer, and then incubated at 4^{o} C for 1 h with shaking in a total volume of 500 μl with 10 μgs of either bacterially-expressed MBP-LIM-9 or MBP. The beads were washed 3X with IP buffer, transferred to clean Eppendorf tubes, pelleted, and the proteins eluted in 30 μl of Laemmli buffer. Two 10% SDS-PAGE were run and transferred to nitrocellulose membranes—one gel contained 5 μl of eluted protein per lane, and the other gel contained 25 μl per lane. The blot that had the 5 μl samples was reacted with mouse monoclonal to myc (Sigma-Aldrich cat. no. M4439) at 1:1000 followed by anti-mouse conjugated to horseradish peroxidase (HRP; Amersham) at 1:10,000, and detection by ECL. The blot containing 25 μl samples was incubated with HRP-conjugated anti-MBP (New England Biolabs) at 1:5000 and detected by ECL.

To confirm the interactions between the PK1 or interkinase regions of UNC-89 with LIM-9, we expressed in yeast HA-tagged PK1 AB or FB fragments, or HA-tagged interkinase region-KD, immunoprecipitated them with anti-HA bound agarose beads and conducted a binding assay with bacterially-expressed MBP or MBP-LIM-9, using methods described (Qadota 2008a).

A far western assay was also used to confirm the interaction between SCPL-1a and LIM-9 (LIM domains only), performed essentially as previously described (Qadota 2007).

Immunostaining and Microscopy

Wild type adult worms were immunostained after fixation by the method described by Nonet et al. (Nonet 1993). Anti-interkinase region of UNC-89 (anti-IK Small 2004) was used at 1:400 dilution, and anti-LIM-9 (Benian-9) was used at 1:100 dilution. Secondary antibodies and confocal microscopy were as described (Qadota 2007)(Qadota, Mercer et al. 2007).

Amino acid sequence analysis

The interkinase sequence was analyzed using a combination of the Protean programs in the Lasergene package for tabulating amino acid composition and predicting secondary structure, and the website www.predictprotein.org for predicting SEG low-complexity and NORS regions. The average frequency of each amino acid in the *C. elegans* proteome was calculated by M. Borodovsky (pers. comm.).

RNAi for lim-9 and scpl-1

RNAi was performed by a feeding method (Simmer 2002). For feeding RNAi, the following plasmids were used: pPD129.36-lim-9 for *lim-9*, pPD129.36-scpl-1a for *scpl-1*, and pPD129.36-lim-9/scpl-1a for *lim-9* and *scpl-1* double RNAi. To construct pPD129.36-scpl-1a, the XhoI fragment of 25B (originally isolated from yeast 2-hybrid screening with UNC-89 PK2 as bait (Qadota 2008a) containing full length of *scpl-1a*)

was cloned into pPD129.36 (Timmons 2001) using its XhoI site. pPD129.36-lim-9 was made by cloning of a PCR-amplified LIM-9 N-terminal half fragment (Qadota 2007) into pPD129.36 using BamHI and XhoI sites. For making pPD129.36-lim-9/scpl-1a double RNAi plasmid, the XhoI fragment of 25B (*scpl-1a*) was cloned into the XhoI site of pPD129.36-lim-9.

For immunostaining of RNAi worms, sufficient numbers of worms were obtained by the following protocol. Feeding bacteria carrying a specific plasmid were cultured in liquid 2xYT and induced to produce dsRNA by adding IPTG to a final concentration of 0.4 mM for 4 hrs. Induced bacteria cells were spotted onto NGM containing ampicillin (50 μg/ml) and tetracycline (15 μg/ml). Worms (the RNAi hypersensitive strain *rrf-3* Simmer 2002) were picked onto five 6 cm RNAi plates (each containing 3 spots of induced bacteria) at 10 worms per plate and these plates were incubated at 20°C overnight to eliminate RNAi non-affected eggs. The next day, 10 worms were transferred from the 6 cm plates to 10 cm plates (11 spots of induced bacteria) and incubated at 20°C for 8 hrs. After 8 hrs, 10 worms were removed from the 10 cm plates. After 3 days incubation at 20°C, the F1 generation reached a good stage for immunostaining using the Nonet method (Nonet 1993).

Overexpression of myc-SCPL-1b

To construct a *C. elegans* expression plasmid of myc-SCPL-1b under the control of heat shock promoter, myc-tagged full length SCPL-1b fragment (Qadota 2008a) was cloned into pPD49.78 and pPD49.83 using the NheI site. Mixtures of pPD49.78/83-myc-SPCL-

1b(FL) with pTG96 (SUR-5::GFP, as a transformation marker) (Yochem 1998)were injected into wild type (N2) for generating transgenic worms. Transgenic animals were identified by finding animals in which SUR-5::NLS::GFP was expressed in nearly all cells. Several independent lines were examined. Transgenic worms were placed at 30°C for 2 hours to induce expression of myc tagged full length SCPL-1b. After an additional 24 hours incubation at 20°C, worms were collected, and used for preparation of fixation by the Nonet method (Nonet 1993). Fixed worms were stained with anti-myc rabbit antibodies (Sigma Aldrich, 1:500 dilution) and anti-IK (1:500 dilution Small 2004).

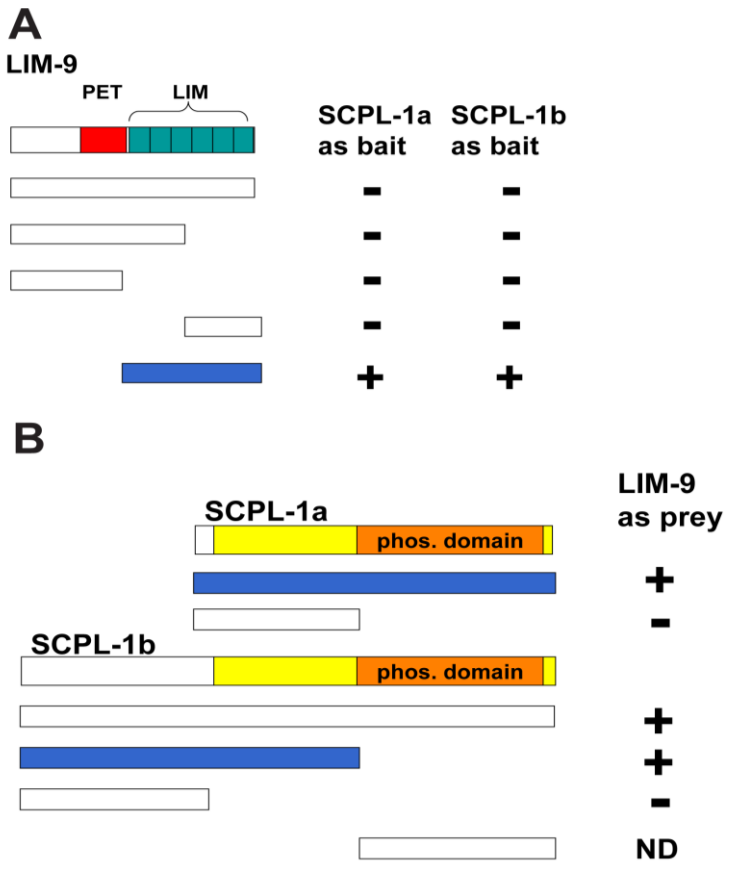


Figure 2–1 Yeast Two-Hybrid Assays Demonstrate Interaction Between SCPL-1 and LIM-9.

(A) LIM-9 contains a PET domain (red bar) and 6 LIM domains (green boxes). Full-length and deletion derivatives of LIM-9 as preys were tested for interaction with SCPL-1 or SCPL-1b as baits. The minimal interacting region of LIM-9 consists of the 6 LIM domains. (B) Both SCPL-1a and SCPL-1b have a protein phosphatase domain (orange bar) and share additional sequences (yellow). Full length SCPL-1a interacts with LIM-9. However, the N-terminal, non-phosphatase region of SCPL-1b is sufficient for interaction with LIM-9. + : growth on –Ade plates; -: no growth on –Ade plates. Blue bars represent the minimal portions required for interaction. ND: cannot be determined due to high background.

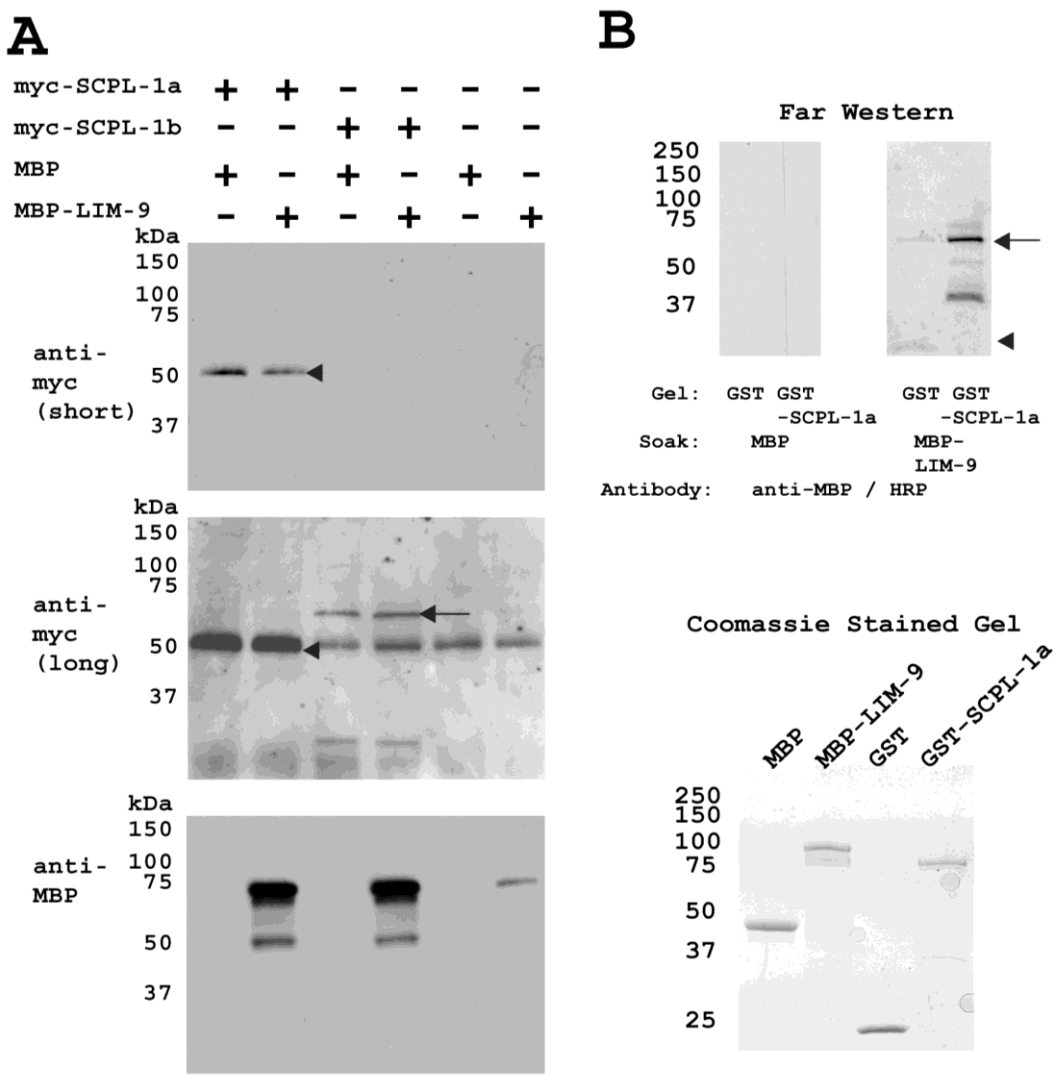


Figure 2–2 Two Methods Confirm the Interaction Between SCPL-1 and LIM-9.

(A) Yeast expressed myc tagged full length SCPL-1a or SCPL-1b were immunoprecipitated by anti-myc monoclonal antibody coated agarose beads, incubated with bacterially expressed MBP-LIM-9 (LIM only) or MBP, extensively washed, bound proteins were eluted and separated on SDS-PAGE gels, blotted and reacted with anti-myc and anti-MBP antibodies. myc-SCPL-1a (molecular weight ~50 kDa) was detected after short exposure, as shown by the arrow head. Long exposure revealed myc-SCPL-1b (molecular weight ~70 kDa), as shown by the arrow. (The 50 kDa band in the right four

lanes of this central blot is likely to be rabbit immunoglobulin heavy chain resulting from cross reactivity with the anti-mouse secondary antibody used to detect the anti-myc monoclonal.) The bottom panel shows the results of reaction with anti-MBP and demonstrates that both SCPL-1a and SCPL-1b interact with MBP-LIM-9 (prominent bands at ~75 kDa) but not MBP. (B) SCPL-1a interacts with LIM-9 (LIM only) by far western assay. The bottom shows a Coomassie-stained SDS-PAGE of the starting materials: MBP, MBP fused to the 6 LIM domains of LIM-9, GST and GST fused to SCPL-1a. On the top is the far western: a gel similar to the one on the bottom was used to separate GST or GST-SCPL-1a, and the proteins transferred to a membrane. One half of the blot was incubated with MBP, the other with MBP-LIM-9 (LIM only), washed, and each half was incubated with antibodies to MBP conjugated to horseradish peroxidase. Reactions were visualized by ECL. Note that MBP-LIM-9 but not MBP interacts with SCPL-1a. The arrowhead indicates the position of GST; the arrow indicates the position of GST-SCPL-1a.

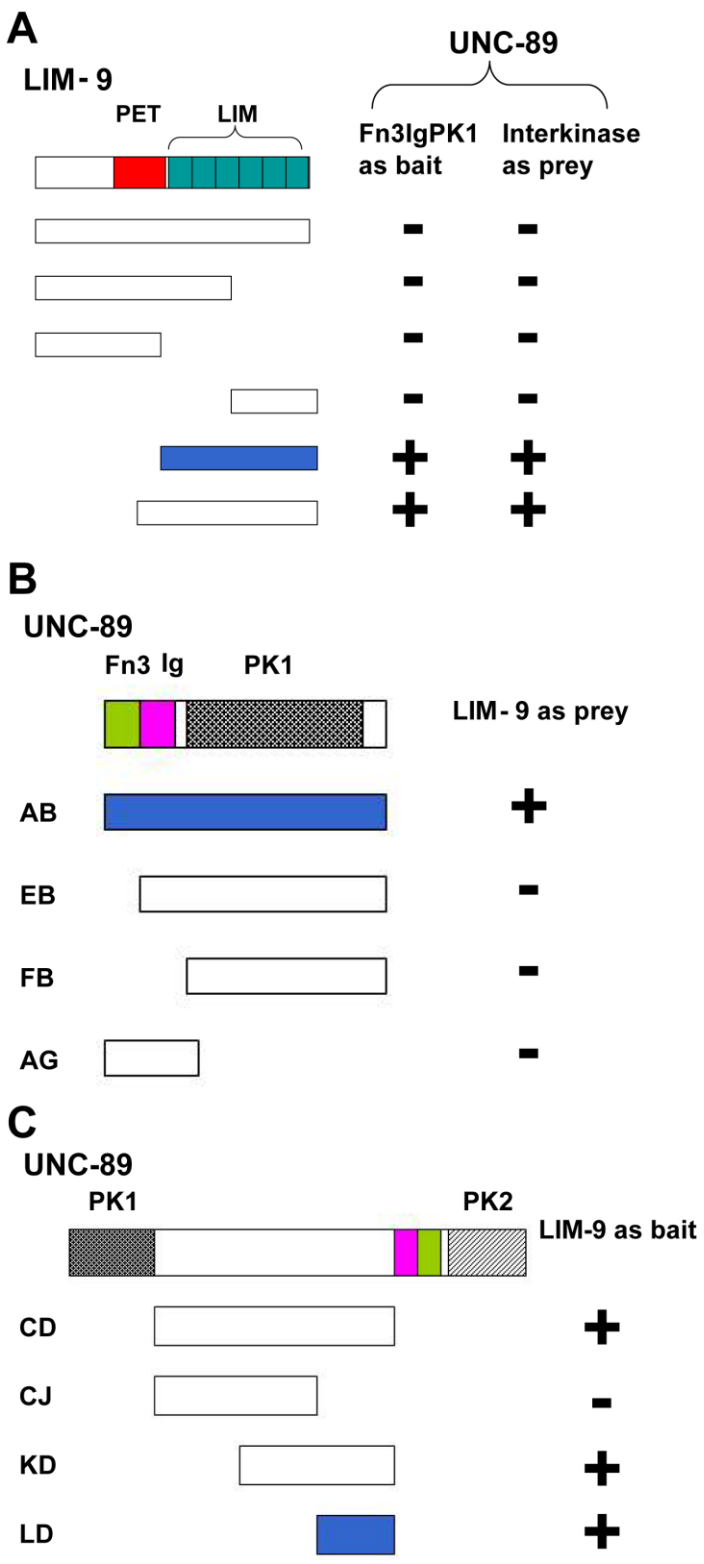


Figure 2–3 Yeast Two-Hybrid Assays Demonstrate Interaction of LIM-9 with either the PK1 or Interkinase Regions of UNC-89.

(A) The six LIM domains (green boxes) constitute the minimal portion of LIM-9 required for interaction with Fn3-Ig-PK1. The six LIM domains of LIM-9 are also sufficient for interaction with the interkinase region of UNC-89. (B) The entire PK1 region including the Fn3 and Ig domains N-terminal of the kinase domain (PK1) is required for interaction with LIM-9. (C) A C-terminal portion of the interkinase region of UNC-89 (about 200 amino acid residues) is sufficient for interaction with LIM-9. + : growth on –Ade plates; -: no growth on –Ade plates. Light green bars represent Fn3 domains and pink bars represent Ig domains. In UNC-89, the bar filled with heart dots denotes the PK1 protein kinase domain, and the bar filled in stripes denotes the PK2 protein kinase domain. Blue bars represent the minimal portions required for interaction.

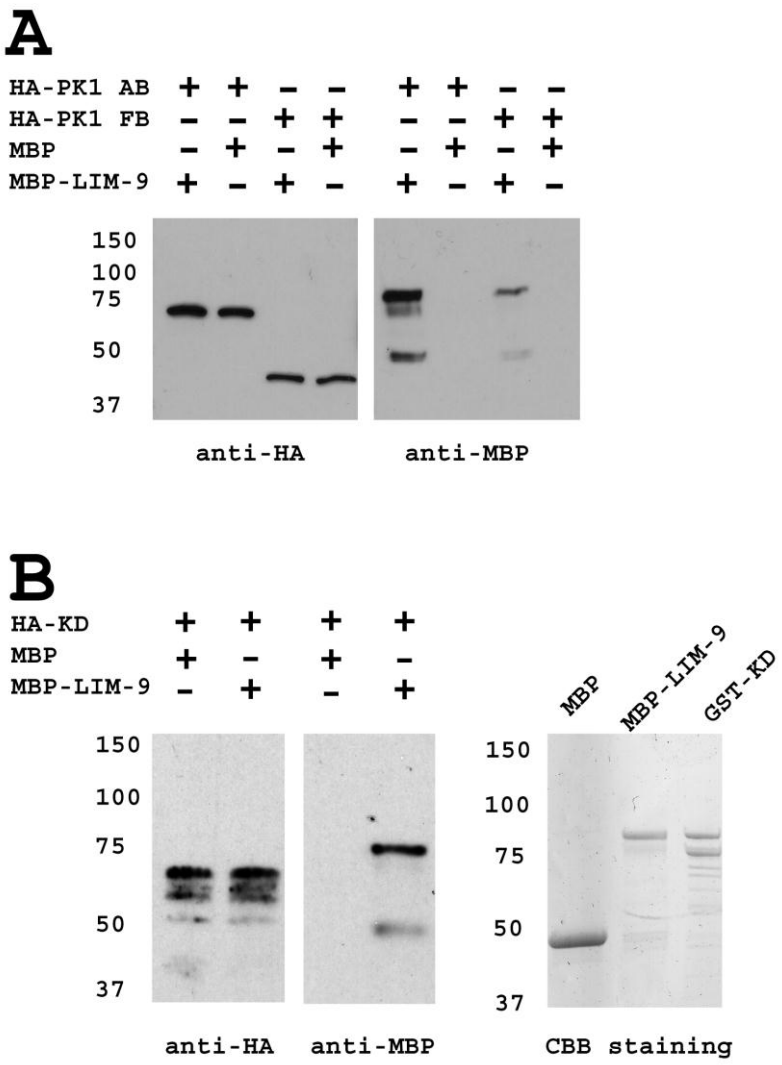


Figure 2–4 Confirmation of Interaction of LIM-9 with either the PK1 or Interkinase Regions of UNC-89.

(A) Bacterially expressed LIM-9 interacts with yeast expressed HA-PK1. Yeast expressed HA tagged PK1 (AB or FB fragments, as shown in Figure 3) were immunoprecipitated by anti-HA monoclonal antibody-coated agarose beads, incubated with bacterially expressed MBP-LIM-9 or with MBP, extensively washed, the bound proteins were eluted and separated on SDS-PAGE gels, blotted and detected with anti-HA or anti-MBP antibodies. Consistent with the two-hybrid assay, Fn3-Ig-PK1 (AB fragment) demonstrates a stronger interaction with LIM-9, than does the kinase domain

alone (FB fragment). No interaction was detected with MBP. (B) Bacterially expressed LIM-9 interacts with yeast expressed HA tagged interkinase region of UNC-89. Left, the C-terminal 2/3 of the interkinase region of UNC-89 (KD fragment, about 400 amino acid residues, as shown in Figure 3) interacts with MBP-LIM-9, but not MBP, using the same method described in (A). Right, coomassie brilliant blue (CBB) stained gel of MBP, MBP-LIM-9 and GST-KD. The top bands of MBP-LIM-9 and GST-KD are likely to be the full length fusion proteins, and the second bands may be degradation products. CBB staining demonstrates the molecular weight of MBP, MBP-LIM-9 and GST-KD. The bands of GST-KD on coomassie stained gel can be used to estimate the molecular weight of HA-KD.

```
1     NDEKLKTEPLSADTLREFKYQHKWLERRVFVQQTPSEQILEAILGPATAQ
51    AQQNAPVAPEGRRPAEIYDYLRIQPKKPPPTVEYVPQPRKEHPPFIDEFG
101   QLIDGDAFDRPEGTGFEGPHRQPPQIPPQPQRPNQAAHDSRRHEQQPQHQ
151   GQPQRIPVDQYGRPLVDPRYLNDPSHRPSSLDDAPFYVDKYGNPVHFDKY
201   GRPMAPQNLEKRKLIPQDKGETPSHSKKEKTQHPVATPILASPGGDQQQQ
251   KIPMRMIRGERREIEEEIANRILSDISEEGSIAGSLASLEDFEIPKDFQV
301   EASEPSTPTLTPEVTIRETIPKPTPSPTSPQKSPVPQPQGLLIPAKVTYS
351   DSILAGLPAADKKVLEDAENDPSIPVGAPLFLEGLHGSDLTIDTTSASGL
401   IKVTSPAINLSPNPKSPRRSTPGTKSPVVLSPRQEHSMEVLIATKRGKPG
451   FLPPGELAEDIDDEDAFMDDRKKQVKPKDHDGENDFKDEKERLEKDKNRR
501   TVNLDDLDKYRPSAFYKDDSDFGHPGYDIDATPWDSHYQIGPDTYLMAAR
551   GAAFNSRVRNYREELFGMGAPTVKQGFLGVRNRDITVRERRRYTDILRET
601   TQGLEPKSHEQSTALLQKAPSATAIERIKADIEKVTPCATKKNDDGT
```

	P	E	Q	R	D	K	A	G	T	Y
C. elegans average	5.0	6.4	4.0	5.2	5.3	6.4	6.3	5.5	5.9	3.3
1-503	***13.3***	***8.4***	***7.0***	***6.4***	***7.4***	***7.0***	6.0	5.6	5.0	2.0
504-647	***5.6***	5.6	3.5	***9.0***	***10.4***	6.3	***9.0***	***7.0***	***8.3***	***4.9***

Figure 2–5 Sequence Features of the 647 residue Interkinase Region of UNC-89. The sequence is devoid of recognizable domains. Residues 1-503 are italicized to denote that this is predicted to be a NORS region, a sequence with little to no secondary structure and possible structural flexibility. Double underlining represents the 5 segments within the first 470 residues that are predicted to be SEG low sequence complexity regions. Amino acids in color are those amino acids that are present at frequencies 0.5% above the average frequencies of a given amino acid in the *C. elegans* proteome. Note the high P,E,Q,R,D and K in residues 1-503. The binding site for LIM-9 has been mapped to residues 447-647.

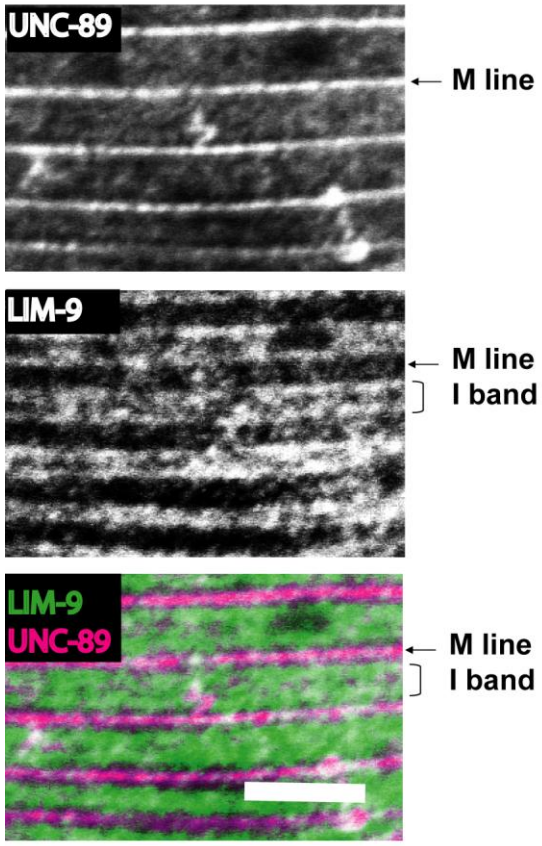


Figure 2–6 Co-localization of LIM-9 and the Interkinase Region of UNC-89 at the M lines in *C. elegans* Body Wall Muscle.

Anti-interkinase (of UNC-89) and anti-LIM-9 antibodies were co-incubated with wild type animals, and the body wall muscle was imaged by confocal microscopy. The images show a portion of one body wall muscle cell. Anti-UNC-89 is localized to M lines, anti-LIM-9 is localized to both M lines and I bands. Scale bar, 10µm.

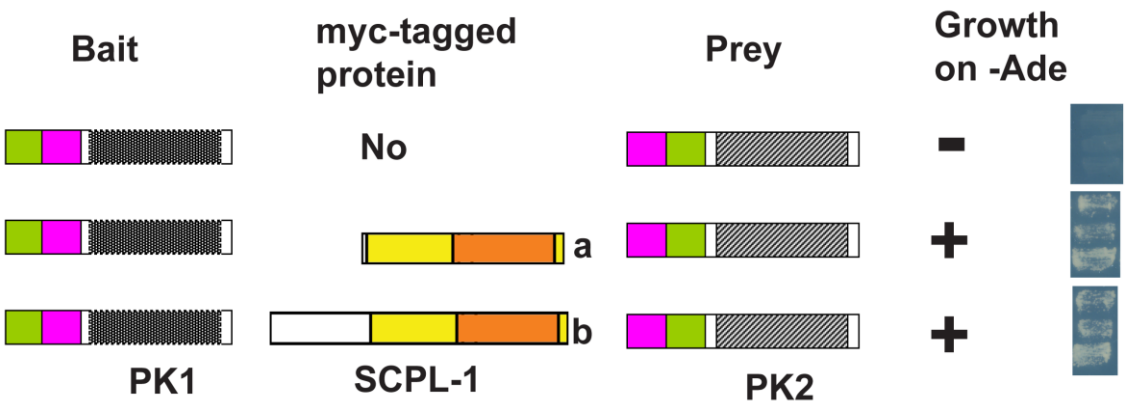


Figure 2–7 Yeast Three-Hybrid Assay Demonstrates A Ternary Complex Containing Fn3-Ig-PK1, SCPL-1 and Ig-Fn3-PK2.

Fn3-Ig-PK1 as bait was coexpressed with myc tagged SCPL-1a or SCPL-1b (or empty myc-vector as control) and Ig-Fn3-PK2 as prey. +: growth on –Ade plates; -: no growth on –Ade plates. The right panel shows the yeast growth on –Ade plates from each experiment from three independent colonies.

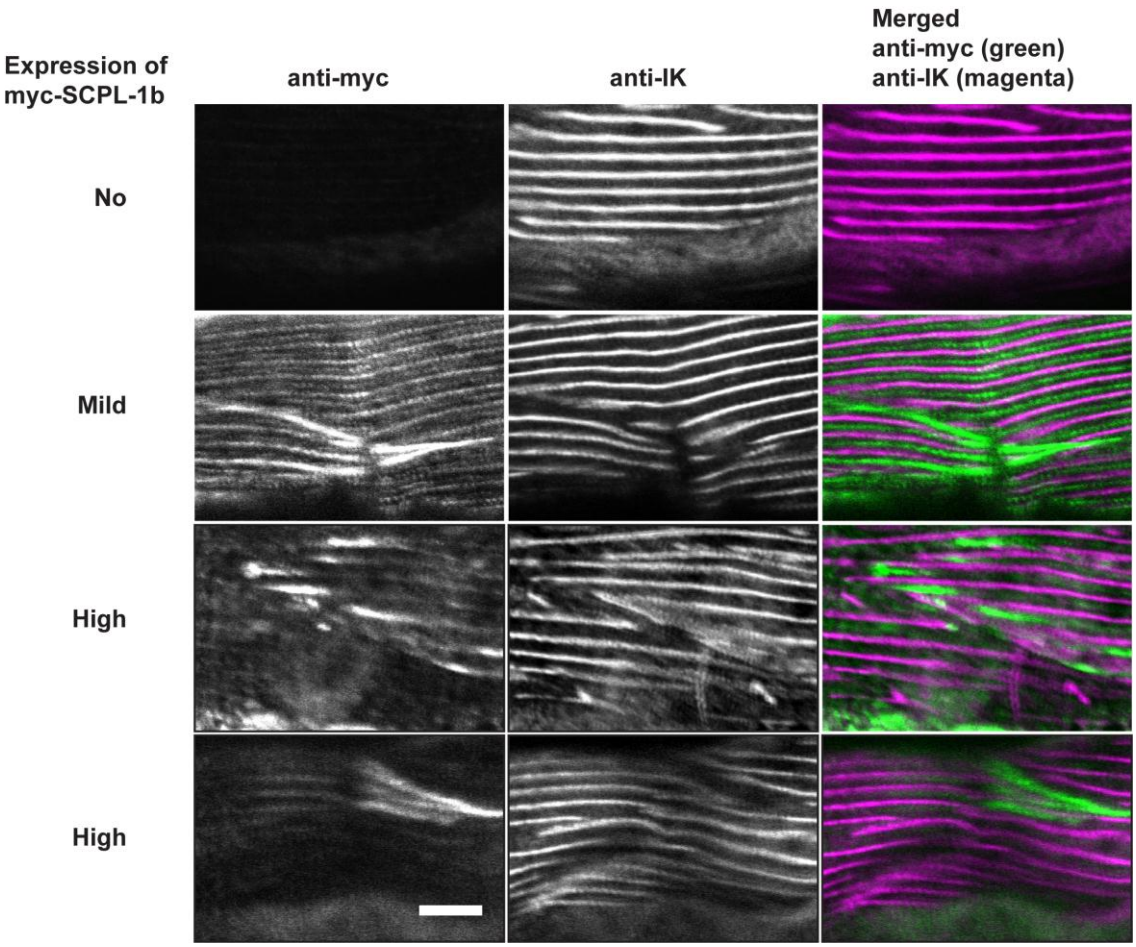


Figure 2–8 Transgenic Overexpression of SCPL-1b Results in Disorganization of UNC-89 in Body Wall Muscle.

Transgenic worms were generated that carry extrachromosomal arrays of a plasmid consisting of myc-tagged SCPL-1b full-length cDNA under the control of a heat-shock promoter. Each row shows images of body wall muscle stained with both anti-myc and antibodies to the interkinase region of UNC-89 (anti-IK). These arrays are sometimes lost during the somatic cell divisions that give rise to the mature body wall muscle cells. Moreover, of those muscle cells that retain the array, there is a variation in the level of expression from that array. The top row shows immunofluorescent staining of a muscle cell that has lost the array: anti-IK shows the normal pattern of localization of parallel M-

lines (Small 2004). The second row shows a muscle cell which retained the array but the level of myc-SCPL-1b was mild: in this cell notice that by anti-myc staining there is both normal localization to I-bands and M-lines (similar to the anti-SCPL-1 antibody staining previously reported, Qadota 2008a), and also in abnormal accumulations. Near the accumulations, anti-IK shows broken M-lines. The bottom 2 rows show muscle cells, which judging from the level of anti-myc staining have even higher expression of myc-SCPL-1b. In these cells, when stained with anti-IK, there are many broken M-lines and diffuse M-lines around accumulations of myc-SCPL-1b. Scale bar, 10 μm.

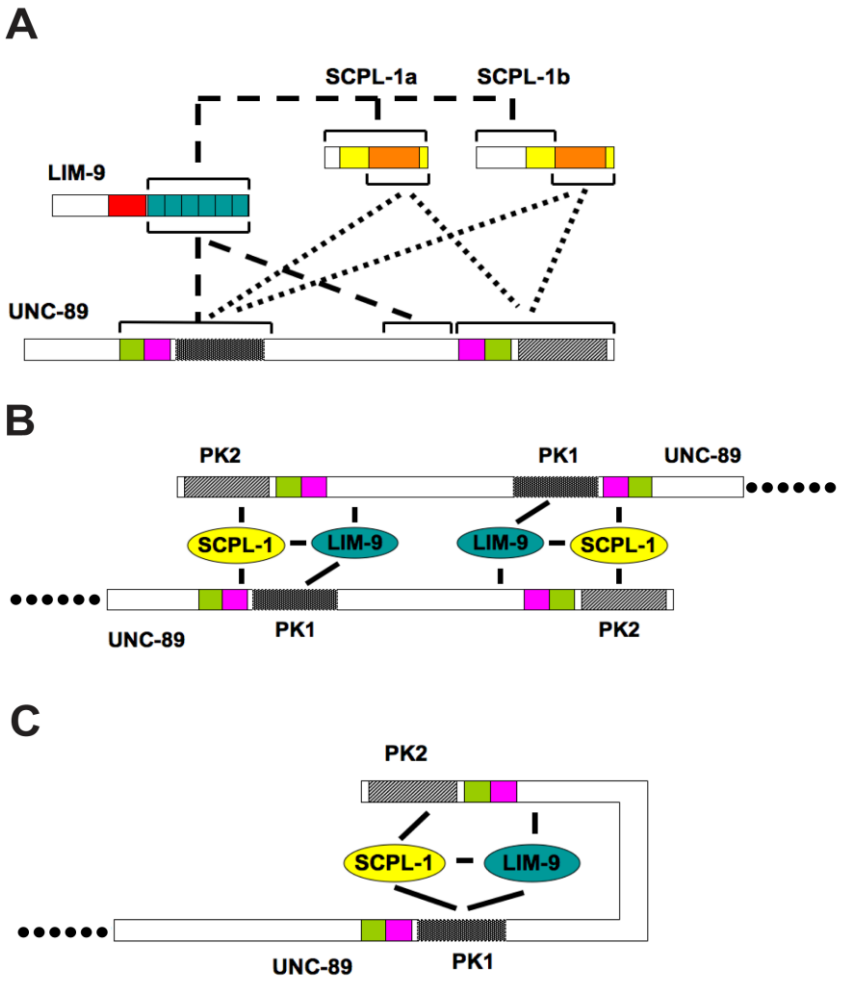


Figure 2–9 Summary of Interactions and Two Working Models.

(A) Summary of protein-protein interactions. The dotted lines show the interactions demonstrated previously (Qadota 2008a). The phosphatase domain (orange bar) of either SCPL-1a or SCPL-1b interacts with Fn3-Ig-PK1 as well as Ig-Fn3-PK2 of UNC-89. The dashed lines show the interactions newly uncovered in this study. Only the 6 LIM domains of LIM-9 are required for interaction with SCPL-1a (or SCPL-1b), Fn3-Ig-PK1, or the interkinase region of UNC-89. (B) Model 1 involves two anti-parallel UNC-89 molecules linked at their C terminal regions through interactions with LIM-9 and SCPL-1. (C) In model 2, it is assumed that the non-domain interkinase region is flexible and forms a loop that is stabilized through interactions with LIM-9 and SCPL-1.

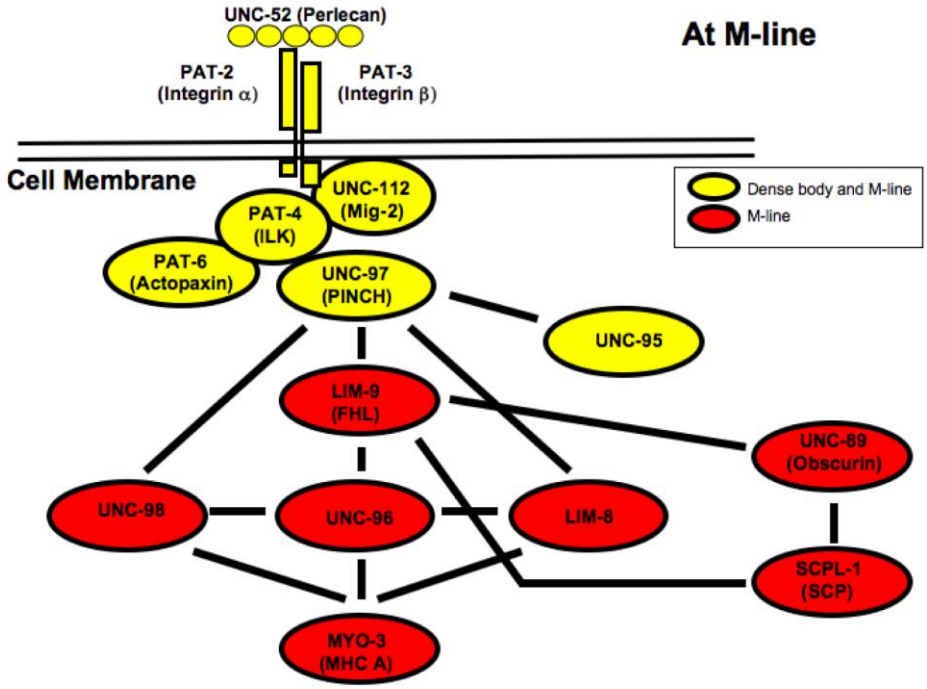


Figure 2–10 Network of Proteins Through Muscle Cell Membrane to MHC A at the M-line.

Through their interactions with LIM-9, UNC-89 and SCPL-1 can now join the complex of proteins at the M-line costamere of *C. elegans* body wall muscle. This diagram is an updated version of the network of proteins described previously (Qadota 2007) that provide a continuous link from the ECM, through the muscle cell membrane to myosin heavy chain A (MHC A) in thick filaments at the M-line.

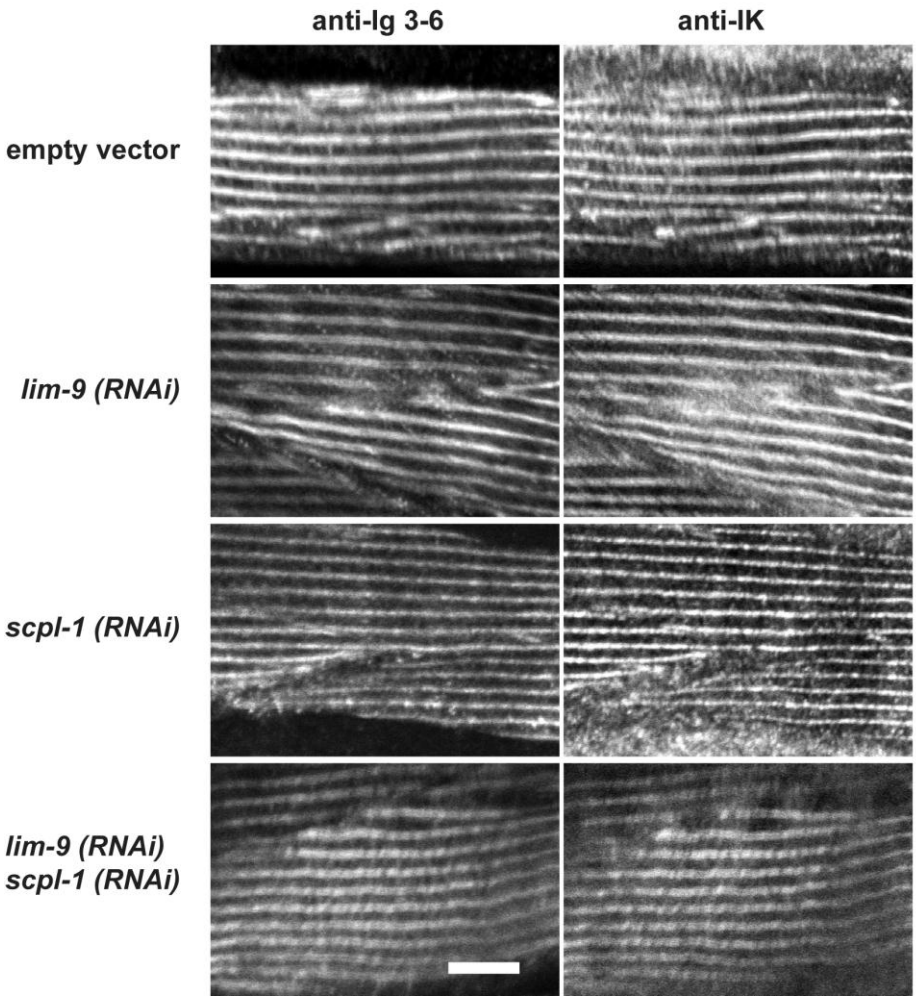


Figure 2–11 RNAi-mediated knockdown of either *lim-9, scpl-1* or *lim-9* and *scpl-1* together, does not affect the localization or organization of UNC-89.

The *C. elegans* RNAi hypersensitive strain *rrf-3* were fed E. coli carrying the empty RNAi vector, or RNAi vectors with inserts for either *lim-9*, *scpl-1*, or both together in one plasmid. The resulting F1 animals were fixed and simultaneously stained with antibodies generated to two different regions of UNC-89—either to Ig 3-6 (near the N-terminus of the large UNC-89 isoforms), or to the interkinase region (near the C-terminus of the large UNC-89 isoforms). The images show one to several body wall muscle cells. Note that in all cases, there is normal localization of UNC-89 to parallel M-lines. Bar, 10 μm.

Table 1 Primer List For Chapter 2

Primer name	Sequence
YIKC	GTACCCCGGGAACGATGAAAAGCTGAAAACTG
YIKD	CGTACTCGAGCTATCCATCATCATCATTCTTCTTTGT
YIKJ	CGTACTCGAGTCAACTCAACACAACTGGACTC
YIKK	GTACCCGGGCCACAAGACAAAGGAGAA
YIKL	GTACCCGGGTCTCCACGCCGTTCCACT
U89-PK1-B	CGTACTCGAGCTACGGACCAAGAATTGCTTCAAG
U89-PK1-F	GTACCCCGGGCGAGTGCGAGAGCTCTACGAA

3. Chapter 3: CPNA-1, A Novel Copine Containing Protein, Links Integrin Associated Protein PAT-6 (Actopaxin) to the Giant Protein UNC-89 (Obscurin) at Muscle Adhesion Site in *C. elegans*

Ge Xiong[1], Adam Warner[2], Hiroshi Qadota[1], Donald G. Moerman[2], Guy M. Benian[1,3]

[1]Department of Pathology, Emory University, Atlanta, Georgia, USA, 30322

[2]Department of Zoology, University of British Columbia, Vancouver, B.C., Canada, V6T1Z4

[3]Corresponding author: Guy Benian

Dept. of Pathology

Whitehead Biomedical Research Building, room 105E

Emory University

Atlanta, Georgia, USA, 30322

404-727-5953

pathgb@emory.edu

Will be submitted to Curr. Biol. in June, 2011.

3.1 Introduction

Sarcomeres, highly ordered assemblages of several hundred proteins, perform the work of muscle contraction. Despite ever increasing knowledge of the components and functions of sarcomeric proteins, we still do not have a clear picture about how sarcomeres are assembled, and maintained in the face of muscle contraction. The nematode *C. elegans,* is an excellent model genetic system in which to study this question (Waterston 1988; Moerman and Fire 1997; Moerman and Williams 2006). In addition to being an excellent system in which to carry out mutational analysis in a whole organism, through both forward and reverse genetics, this nematode offers several advantages for studying striated muscle. These include its optical transparency, which allows evaluation of sarcomeric structure by polarized light, and localization of GFP tagged proteins. In addition, its usual mode of self-fertilization allows propagation of muscle mutants that would be unable to mate. *C. elegans* has proven to be a superb platform for discovery of the first examples of new and conserved components of sarcomeres. For example, the myosin head chaperone, UNC-45, was first described in *C. elegans* (Barral et al. 1998; Barral et al. 2002). The first clue that Kindlins have a role in integrin adhesion complexes came from cloning and cell biological analysis of the *unc-112* gene in *C. elegans* (Rogalski et al. 2000). In addition, the first evidence for interaction of Kindlins with integrin linked kinase came from study of *C. elegans* (Mackinnon et al. 2002).

The sarcomere contains a number of extraordinarily large polypeptides (700,000 Da—4 MDa) composed of multiple copies of immunoglobulin (Ig) and fibronectin type 3 (Fn3) domains, one and even two protein kinase domains, and in some proteins, elastic regions. In general, results indicate that these giants act as scaffolds for the assembly of other

proteins into the sarcomere (Kontrogianni-Konstantopoulos et al. 2009). The giant protein for which we have the most information is vertebrate titin (Granzier and Labeit, 2004; Linke, 2008). In addition to its role in sarcomere assembly and passive tension, there is experimental support for the idea that the protein kinase domain of titin acts as a sensor of muscle activity, and transmits signals to the nucleus to increase sarcomere gene expression as a way to maintain and build muscle mass (Lange et al. 2005). Much less is currently known about the function of the other titin-like giant proteins. *C. elegans* striated muscle contains three such polypeptides: twitchin (754,000 Da) located in the outer region of the A-band (Moerman et al. 1988; Benian et al. 1989), TTN-1 (2.2 MDa) located in the I-band (Flaherty et al. 2002; Forbes et al. 2010), and UNC-89 (as large as 900,000 Da) located at the M-line (Benian et al. 1996; Small et al. 2004).

Loss of function mutations in *unc-89* result in animals that display reduced locomotion, disorganized myofibrils, especially at the A-band, and usually lack M-lines (Waterston et al., 1980; Benian et al., 1999). *unc-89* mutants show a disorganization of myosin thick filaments by immunostaining (Qadota et al. 2008b). *unc-89* is a complex gene: through the use of 3 promoters and alternative splicing, at least 6 major polypeptides are generated, ranging in size from 156,000 to 900,000 Da (Benian et al. 1996; Small et al., 2004; Ferrara et al., 2005). The largest of these isoforms, UNC-89-B and UNC-89-F, which are each ~900,000 Da, consist of 53 Ig domains, 2 Fn3 domains, a triplet of SH3, DH and PH domains near their N-termini, and two protein kinase domains (called PK1 and PK2) near their C termini. Antibodies localize UNC-89 to the M-line (Benian et al. 1996; Small et al. 2004). Although protein kinase activity for either PK1 or PK2 has yet to be demonstrated, molecular modeling suggests that PK2 is active, while PK1 is

inactive (Small et al., 2004). The human homolog of UNC-89 is obscurin (Young et al. 2001; Bang et al. 2001; Kontrogianni-Konstantopoulos et al. 2009). Although obscurin contains all the same domains as UNC-89, the SH3, DH and PH domains are located near the C terminus rather than near the N terminus as they are in UNC-89. Although UNC-89 is located only at the M-line, various obscurin isoforms are located at either the M-line, the A/I junction or at the Z-disk and Z/I junction (Bowman et al., 2007).

To understand the molecular mechanisms by which UNC-89 is assembled at the M-line, and how UNC-89 performs its functions, we are using a yeast 2-hybrid approach to identify its binding partners. We have reported that the protein kinase domains, PK1 and PK2, each interact with SCPL-1, a CTD-type protein phosphatase (Qadota et al. 2008a). Loss of function of *scpl-1* results in worms with normally organized sarcomeres, but these animals have increased maximal bending (Nahabedian, Qadota, Lu, and Benian, unpub. data). We have also reported that both the PK1 region and the 645 residue "interkinase region" each interact with LIM-9 (Xiong et al. 2009), the closest worm homolog of the human protein FHL ("four and a half LIM domains" protein). SCPL-1 and LIM-9 have also been localized at least partially to M-lines. We have also found that SCPL-1 interacts with LIM-9. Knockdown of *lim-9* by RNAi results in disorganization of thick filaments in body wall muscle (Meissner et al. 2009). Finding that UNC-89 interacts with LIM-9 suggested that this was one molecular mechanism by which UNC-89 might become assemble at M-lines: LIM-9 had been shown to interact with UNC-97 (PINCH) (Qadota et al. 2007) and be a member of an interacting network of proteins that form a physical linkage from the muscle cell membrane, beginning with integrin, ultimately to myosin in thick filaments at the M-line (Qadota and Benian, 2010).

However, loss of function of either *lim-9* or *scpl-1* does not result in major defects in the organization of UNC-89, suggesting that other proteins are involved in anchoring UNC-89 to the M-line.

Here, we report that through a 2-hybrid screen we have identified a new binding partner for the N-terminal region of UNC-89, called CPNA-1. We have determined that CPNA-1 localizes to muscle adhesion complexes (M-lines and dense bodies). Loss of function of *cpna-1* results in embryonic lethality with a phenotype similar to mutations in integrin and integrin associated proteins. In addition to interacting with UNC-89 and three other M-line proteins, we found that CPNA-1 interacts with PAT-6 (actopaxin), located at the base of M-lines and dense bodies, and is a member of a conserved four-protein complex that associates with integrin.

3.2 Results

Identification of CPNA-1 as a binding partner for UNC-89.

In an effort to identify new binding partners for UNC-89, a portion of UNC-89 comprised of Ig domains 1-5 (Ig1-5) was used as bait to screen a yeast 2-hybrid library of *C. elegans* cDNAs (Fig. 3-1A). Two positive preys representing the CPNA-1 (nomenclature described below) were recovered. As shown in Fig. 3-1A, CPNA-1, isoform b, has a predicted transmembrane domain (TM) near its N-terminus and a "copine domain" (copine) near its C-terminus. The smallest prey clone (residues 825-1058) contains essentially just the copine domain, suggesting that this region of CPNA-1 is minimally required for interaction with UNC-89. To determine how specific the interaction of CPNA-1 is for this small portion of UNC-89, CPNA-1 was used in yeast 2-

hybrid assays as both bait and prey to test for interaction with an additional 16 portions of UNC-89 that comprise the entire largest isoform of UNC-89 (UNC-89-B). As indicated, CPNA-1 (825-1058) only interacts with Ig1-5. Deletion derivatives of Ig1-5 were used in 2-hybrid assays against CPNA-1. As shown, in Fig. 3-1C, the minimal region necessary and sufficient to interact with CPNA-1 is Ig1-3. Two biochemical approaches were used to verify the interaction identified by yeast 2-hybrid. In the first approach, beads coats with GST-UNC-89 Ig1-3, but not GST, were used to successfully pull out CPNA-1 from a worm lysate (Fig. 3-1D1). Because this result could have occurred via an intermediary protein in the worm lysate, we also performed a binding assay using purified proteins. A "far western" experiment in which either MBP or MBP-CPNA-1 (825-1058) was separated on a gel, transferred to a blot, and reacted to a solution containing His-tagged UNC-89 Ig1-3, demonstrates direct interaction between CPNA-1 and UNC-89 Ig1-3 (Fig. 3-1D2).

CPNA-1 is an "atypical" copine domain containing protein.

Copine domains (Pfam entry PF07002; InterPro entry IPRO10734) also known as von Willebrand A-like or A domains, are approximately 180 residues long, and have weak homology to the extracellular A domain of integrins. Copine domains have an unknown function, although they are considered to be cytoplasmic and to be involved in protein-protein interactions (Tomsig and Creutz, 2002). All previously characterized proteins containing copine domains also have two C2 domains, which are calcium dependent phospholipid binding motifs. Our BLAST search of the *C. elegans* genome revealed that *C. elegans* has 7 genes that encode proteins containing copine domains (Fig. 3-2). As

compared to CPNA-1, the sequence identity of copine domains in *C. elegans* ranges from 55-28%. Except for the copine domains, and several that also have C2 domains, there is little homology between them. Of the seven genes, only *nra-1* (Gottschalk et al. 2005) and *gem-4* (Church and Lambie, 2003), encode proteins with C2 domains. Thus, we have designated F31D5.3 as CPNA-1, and the other copine domain proteins, as "copine domain atypical", *cpna-1* through *cpna-5*. In addition, CPNA-1 is the only copine domain protein in *C. elegans* that contains a predicted transmembrane domain. When tested by 2-hybrid against copine domains from each of the *C. elegans* proteins, UNC-89 Ig1-5 interacted with only the copine domain from CPNA-1 (Fig. 3-2).

Alignment of copine domains from the *C. elegans* proteins reveals 12 positions that have the same residue in all the nematode copine containing proteins (shaded green in Fig. 3-3A). These residues are also conserved in the consensus sequence derived by PFAM for all copine proteins in all species. As indicated below, I have demonstrated that the copine domain of CPNA-1 can interact with UNC-89 Ig1-5 and with UNC-96 (201-418). To obtain insight into whether this copine domain has distinct binding sites for each of these proteins, I mutated 3 of these 12 conserved residues (indicated with asterisks in Fig. 3-3B): G922 was changed to V since V has a larger side chain and we expected that this change might result in a local conformational change. Y985 was changed to A to alter the side chain, and Y985 was changed to F, which is not a major change in structure, but would eliminate a possible phosphorylation site. Similarly, S1015 was changed to A to eliminate a possible phosphorylation site. As shown in Fig. 3-2C, the G922V mutation eliminates interaction with UNC-89, but not with UNC-96. Similarly, the S1015A

mutation eliminates interaction with UNC-96, but not with UNC-89. Therefore, the copine domain of CPNA-1 may have distinct binding sites for these two M-line proteins.

CPNA-1 localizes to muscle adhesion complexes.

To determine the localization of CPNA-1 in nematode muscle, we expressed a non-domain-containing region of CPNA-1 (176-385) as a GST fusion protein (Fig. 3-4A), and it was used to raise rabbit antibodies. After affinity purification, these antibodies detected a single band on western blot using a protein lysate from *C. elegans* (Fig. 3-4B). The size of this band, approx. 130 kDa, is very similar to the size of the predicted CPNA-1b isoform. Anti-CPNA-1 antibodies were used in immunofluorescent experiments to localize CPNA-1 in wild type adult body wall muscle. As shown in Fig. 3-4C, CPNA-1 localizes to both M-lines and dense bodies, co-localizing with UNC-89 (M-lines), and with α-actinin (dense bodies). Anti-CPNA-1 antibodies were also used to stain transgenic animals expressing GFP tagged UNC-112. UNC-112 (Kindlin) is located at M-lines and dense bodies (Rogalski et al. 2000), and has been shown to interact with the cytoplasmic tail of the transmembrane protein integrin (Qadota, Moerman and Benian, submitted). Thus, UNC-112 can be considered as a marker for proteins that are located close to the muscle cell membrane. As shown in Fig. 3C (third row), at least some CPNA-1 co-localizes with UNC-112, and therefore at least some CPNA-1 is membrane-proximal. In addition, we co-stained for CPNA-1 and one more membrane-proximal component, PAT-6 (actopaxin), which is a member of the conserved four-protein complex that associates with the cytoplasmic tail of β-integrin (Lin et al. 2003). As shown in Fig. 3-4C (last row), CPNA-1 and PAT-6 co-localize. To obtain further evidence that some CPNA-1 is located

near the muscle cell membrane, adult worms were co-stained with anti-CPNA-1 and anti-UNC-89 (MH42 monoclonal; Benian et al. 1996), and using confocal microscopy, 0.5 μm optical sections were obtained from the cuticle side to deep into the body wall muscle cell (Fig. 3-12). Although UNC-89 is located throughout all the sections, CPNA-1 is located closer to the cuticle, and thus closer to the muscle cell membrane.

CPNA-1 interacts with 4 additional adhesion complex proteins.

The interaction of CPNA-1 with UNC-89 might explain the localization of CPNA-1 to the M-lines, but not dense bodies, as UNC-89 is located exclusively to M-lines (Benian et al. 1996; Small et al. 2004). Therefore, we used CPNA-1 to conduct a yeast 2-hybrid screen of a collection of 23 known components of dense bodies and M-lines (Table 3), to identify which proteins CPNA-1 might interact with at dense bodies, and perhaps additional proteins at M-lines. As indicated in Fig. 3-5, we identified 4 additional CPNA-1 interacting proteins: SCPL-1 (a CTD-type phosphatase), LIM-9 (FHL), PAT-6 (actopaxin) and UNC-96. The minimal portions of each of these proteins required for interaction with CPNA-1 were also determined. We also determined that, interestingly, although only the copine domain of CPNA-1 is required to interact with UNC-96, nearly full-length CPNA-1 is required for interaction with SCPL-1, LIM-9 and PAT-6. As shown in Fig. 3-6, these interactions were confirmed using in vitro binding experiments with purified proteins. Yeast-expressed and HA tagged copine domain of CPNA-1(825-1058) was shown to interact with MBP tagged UNC-96 C-terminal half (201-418)(Fig. 3-6A), the minimal region of UNC-96 shown by 2-hybrid to be required for interaction. A far western blot experiment showed that His tagged CPNA-1 (173-1107) can interact

with MBP fusions of SCPL-1 (phosphatase domain), PAT-6 (full-length) and LIM-9 (LIM domains only) (Fig. 3-6B). By antibody staining, UNC-96 is localized only to M-lines (Mercer et al. 2006), and SCPL-1 (Qadota et al. 2008a), and LIM-9 (Qadota et al. 2007) are found at M-lines and I-bands (but not at dense bodies). In contrast, PAT-6 is located at both M-lines and dense bodies (Lin et al. 2003; Fig. 3-4C). Thus, our screens have revealed 5 proteins with which CPNA-1 can interact with at M-lines (UNC-89, PAT-6, SCPL-1, LIM-9 and UNC-96), and one protein with which CPNA-1 can interact with at dense bodies (PAT-6).

CPNA-1 required for sarcomere assembly.

cpna-1 was independently identified as one of four new Pat mutants during an RNAi screen of ~3300 muscle-expressed genes (Meissner et al. 2009). The Pat embryonic lethal phenotype has been found in loss of function mutants for 20 genes that are essential for embryonic muscle development and mostly encode proteins localized to muscle focal adhesions (M-lines and dense bodies)(Williams and Waterston, 1994; Meissner et al. 2009). We examined the phenotype of an intragenic deletion, and likely null mutation, for *cpna-1*, *gk266*. Data on WormBase indicates that *gk266* is a 9 bp insertion in the 3' end of intron 5, followed by a 393 bp deletion that extends into the 5' end of exon 6. As shown in Fig. 3-7B, *cpna-1(gk266)* homozygotes display the typical Pat phenotype of arrest at the two-fold stage of embryogenesis and no movement within the eggshell. For comparison, a wild type worm (Fig. 3-7A) of similar age hatches and progresses normally through the larval stages. Using anti-CPNA-1 antibodies, we examined the localization of CPNA-1 in the mutant background of a number of other Pat genes. In *pat-3* (Fig. 3-7C),

and *pat-4* (Fig. 3-7D) developmentally arrested embryos, CPNA-1 is mislocalized, clumping in the middle of muscle cells (arrows). In *pat-6* embryos (Fig. 3-7E), CPNA-1 is disorganized, but properly polarized in the muscle cell membrane of most cells. CPNA-1 shows some variability in its localization in *pat-6* arrested embryos, ranging from disorganization (Fig. 3-7F), to severe mislocalization, similar to that seen in the *pat-3* and *pat-4* backgrounds. We also looked at the localization pattern of essential body wall muscle proteins in the *cpna-1* background. For each of PAT-3 (β-integrin) (Fig. 3-7J), DEB-1 (vinculin) (Fig. 3-7L), and PAT-6 (actopaxin) (Fig. 3-7N), each protein was localized properly into adhesion complexes in bands of muscle quadrants, albeit with some minor disorganization when compared to wild type embryos (Fig. 3-7I, K, M). When comparing the tight organized arrays of thick filaments in MYO-3 stained wild type embryos (Fig. 3-7N) to those in cpna-1 arrested embryos, severe disorganization (arrows) of MYO-3 is observed (Fig. 3-7O). These results place *cpna-1* in between *pat-6* and *myo-3* in the M-line/dense body assembly pathway.

PAT-6 (actopaxin) is required for the assembly of CPNA-1 into muscle adhesion complexes.

Of all the proteins that we have found interact with CPNA-1, PAT-6 is the protein that is earliest in the M-line/dense body assembly pathway, and both PAT-6 and CPNA-1 are required for embryonic sarcomere assembly. In addition, mutant analysis indicates that CPNA-1 is downstream of PAT-6, and thus, we hypothesize that PAT-6 directs the assembly of CPNA-1 to M-lines and dense bodies in embryonic muscle. We next asked what is the relationship of PAT-6 and CPNA-1 in adult muscle. To address this question,

we used RNAi to knockdown *pat-6*, beginning at the L1 larval stage and continuing into adulthood, in order to avoid the embryonic requirement of *pat-6*. We have observed that adults resulting from this type of RNAi for another pat gene, *unc-97*, show a mosaic pattern of knockdown, with some adult body wall muscle cells showing normal expression of the gene, whereas others showing a large reduction in expression (Miller et al. 2006). As shown in Fig. 3-8A, this phenomenon was also see for *pat-6* RNAi beginning at the L1 stage. Most importantly, in muscle cells in which PAT-6 was nearly undetectable, CPNA-1 was mislocalized, concentrated at the edges of the muscle cell (arrows). RNAi was also used to knockdown *unc-97* beginning at the L1 stage, and the resulting adults were immunostained for MHC A myosin and CPNA-1. Previously we had reported that *unc-97* RNAi results in large aggregates of MHC A (Miller et al. 2006). As shown in Fig. 3-8A, the muscle cell showing the clump of MHC A, also shows mislocalization of CPNA-1 in clumps, but not localization to the edges of the cell, as was found for *pat-6* RNAi.

The localization of CPNA-1 is not affected by the deficiency of UNC-89, LIM-9, SCPL-1 and UNC-96.

We next asked what would be the effect in adult muscle of deficiency of the other proteins that interact with CPNA-1 and are located at the M-lines, namely UNC-89, SCPL-1, UNC-96 and LIM-9. Each of these four proteins is not required for embryonic muscle development, and when deficient, have adult muscle phenotypes (Waterston et al. 1980; J. Nahabedian, H. Qadota, H. Lu, G. Benian, unpub. data; Mercer et al. 2006; Meissner et al. 2009). To do this, we used anti-CPNA-1 to immunostain adult body wall

muscle from loss of function mutants or animals subjected to RNAi. *unc-89(su75)* is an *unc-89* allele that lacks all the large UNC-89 isoforms (Small et al. 2004), and thus lacks all CPNA-1 binding sites on UNC-89. *unc-96(sf18)* is a nonsense mutation with undetectable UNC-96 protein by western blot and thus is a putative null allele (Mercer et al. 2006). The *lim-9* alleles, *gk106* and *gk210*, are also likely null mutants, as they are intragenic deletions with undetectable LIM-9 protein (Qadota et al. 2007). As shown in Fig. 3-9, the localization of CPNA-1 at M-lines (indicated by arrows) is unaffected by the absence or reduced levels of UNC-89, SCPL-1, UNC-96 or LIM-9. These results are consistent with these proteins being downstream of CPNA-1 in the M-line assembly pathway.

PAT-6, CPNA-1 and UNC-89 form a ternary complex.

Based on our mutant analysis, PAT-6 is required for assembly of CPNA-1 and CPNA-1 is required for assembly of UNC-89 at the M-line. If this were the case, we could expect that we could obtain evidence for a ternary complex between PAT-6, CPNA-1 and UNC-89. To investigate this, we performed a yeast 3-hybrid assay. UNC-89 Ig1-5 as bait, was co-expressed with HA-tagged CPNA-1 (173-1107), or empty vector as control, and PAT-6 as prey. As shown in Fig. 3-10, the interaction of UNC-89 Ig1-5 with PAT-6 is dependent on the presence of CPNA-1. This result is consistent with the existence of a ternary complex containing PAT-6, CPNA-1 and UNC-89.

3.3 Discussion

Model for role of CPNA-1 in assembly of integrin adhesion complexes.

Our studies have identified CPNA-1 as a new component of integrin adhesion complexes, based on three criteria. First, CPNA-1 is localized to M-lines and dense bodies, similar to integrin and integrin associated proteins. Second, loss of function of CPNA-1 is Pat embryonic lethal, a phenotype that is characteristic of integrin and integrin associated proteins. Third, CPNA-1 binds directly to PAT-6, one of the integrin associated proteins. Our data suggests that the molecular function of CPNA-1 is to act as a linker between integrin associated proteins near the muscle cell membrane, and proteins that are found deeper inside the muscle cell (Fig. 3-11). In striated muscle of *C. elegans*, both the M-lines and dense bodies are integrin adhesion complexes containing both shared and specific protein components. At the base of each M-line and dense body, associated with the cytoplasmic tail of β-integrin, is a complex of four conserved proteins including UNC-112 (Kindlin), PAT-4 (ILK) and UNC-97 (PINCH). Lin et al. (2003) showed that PAT-4 (ILK) interacts with the C-terminal CH domain of PAT-6 (actopaxin). Here, we show that the N-terminal portion of PAT-6, including the first CH domain, interacts with CPNA-1. CPNA-1, in turn, interacts with UNC-89 (obscurin), LIM-9 (FHL), SCPL-1 (SCP) and UNC-96 at M-lines (Fig. 3-11). We have not yet identified dense body specific proteins that interact with CPNA-1, that is, proteins that assemble downstream of CPNA-1 that would be analogous to those we have found at the M-line (UNC-89, LIM-9, SCPL-1 and UNC-96). Eventually, we would like to examine the phenotype of animal which is simultaneously deficient in all proteins that interact downstream of CPNA-1; this might phenocopy the loss of function of *cpna-1*.

We have shown that loss of function of *cpna-1* is embryonic lethal, displaying the characteristic Pat phenotype common to genes that are essential for embryonic muscle

development. Indeed, many of the Pat genes encode other components of muscle adhesion complexes, including UNC-52 (perlecan), PAT-3 (β-integrin), UNC-112 (Kindlin), PAT-4 (ILK), UNC-97 (PINCH), PAT-6 (actopaxin), and DEB-1 (vinculin). By localizing Pat proteins in *cpna-1* mutant embryos, and localizing CPNA-1 in other Pat mutants, we have shown that PAT-6 directs the assembly of CPNA-1 to the adhesion complexes during embryonic development. Similarly, by knocking down *pat-6* transcripts and protein after embryonic development, we have shown that PAT-6 is required for the recruitment of CPNA-1 to adult M-lines and dense bodies. The lack of mislocalization of CPNA-1 in loss of function for UNC-89, LIM-9, SCPL-1 and UNC-96, is consistent with the model that CPNA-1 directs the assembly of these four M-line proteins in adult muscle. It would be desirable to confirm this assembly pathway in adult muscle by examining the localization of these putative downstream proteins in adult muscle when there is deficiency for CPNA-1. At this time, however, an adult Unc allele for *cpna-1* is not available, although such an allele is feasible given that some genes, like *unc-97* and *unc-112*, display null Pat, and hypomorph Unc phenotypes. We also failed to observe an adult muscle phenotype when RNAi for *cpna-1* was conducted from the L1 stage. The apparent insusceptibility of *cpna-1* RNAi could, for example, be due to either because CPNA-1, once incorporated into adhesion structures is very stable, or, although the message and protein were knocked down, it was not knocked down sufficiently to create a phenotype. Nevertheless, overall, our mutant analysis is consistent with the protein interaction data and the model presented in Fig. 3-11.

New classes of copine domain containing proteins.

Copines, containing 2 C2 domains and a copine domain, are evolutionarily conserved proteins found in plants and animals. They were first identified as Ca^{2+}-dependent phospholipid binding proteins in Paramecium (Creutz et al. 1998) and because of this property are thought to be involved in membrane-trafficking. CPNA-1 is the first characterized copine domain containing protein that lacks C2 domains. The Pfam website, as of April 2011, notes 200 sequences from diverse organisms, that have a copine domain plus 2 C2 domains, 25 sequences that have a copine domains plus a single C2 domain, and interestingly, 168 sequences containing only the copine domain. Significantly, this copine domain-only category includes a putative uncharacterized protein from the mouse (UniProt entry Q8BIJ1 MOUSE; 346 amino acids), and isoform CRA_d of copine V from humans (UniProt entry Q7Z6C8 HUMAN; 301 amino acids).

Among the copine domain containing proteins in the *C. elegans* proteome, both typical and atypical, only CPNA-1 contains a predicted transmembrane domain. In fact, our in silico analysis of copine domain containing proteins in humans, has not revealed any copine proteins that also have predicted transmembrane domains. Perhaps the function of the transmembrane domain in CPNA-1 is for membrane association, to the muscle cell membrane and/or other membrane compartments. Additional experiments will be required to address this idea. It is also tempting to speculate that the reason that CPNA-1 does not contain C2 domains is that in typical copine domain proteins, membrane association is, instead provide by the C2 domains. It is also interesting to note that CPNA-1 is the only described integrin associated protein in *C. elegans* or any other animal that contains a transmembrane domain other than the integrins themselves. Note that our domain mapping (Fig. 3-5) indicates that the transmembrane domain of CPNA-1

is not required for its association with PAT-6. Thus, CPNA-1 may have two ways to localize near the muscle cell membrane—binding to the integrin associated protein PAT-6, and direct insertion of its transmembrane domain into the muscle cell membrane, or other nearby membranes.

Genetic analysis of copine domain containing proteins.

Our genetic analysis of *cpna-1* reveals that *cpna-1* is required for embryonic muscle assembly and development. The biological roles of some additional copine proteins have been studied by genetic analysis. For example, the first copine protein to be genetically analyzed was the BON1/CPN1 gene from Arabidopsis thaliana (Hua et al. 2001; Jambunathan et al. 2001). Loss of function for BON1/CPN1, results in de-repression of defense responses and a reduction in plant growth. In Dictyostelium, knockout of one of 6 copine genes called cpnA, suggests that cpnA plays a role in cytokinesis and contractile vacuole function, and is required for later stages of development. (Damer et al. 2007). In *C. elegans*, two copine domain proteins, each with two C2 domains, have been characterized, NRA-1 and GEM-4. NRA-1 was identified as an interactor of nicotinic acetylcholine receptors (nAChR) by a tandem affinity purification procedure (Gottschalk et al. 2005). Loss of function mutants and RNAi knockdown of *nra-1* display nicotine resistance and reduced synaptic nAChR expression, suggesting a normal role as positive regulator of nAChRs. Mutations in *gem-4* (Church and Lambie, 2003) were isolated as extragenic suppressors of loss of function alleles of *gon-2*, which is required for gonadal cell divisions. GON-2 is a cation channel of the TRPM family. The copine protein GEM-4 antagonizes the channel GON-2, but the mechanism is unknown.

Copine domains as protein interacting modules.

We identified CPNA-1 as an interacting partner for the giant polypeptide UNC-89. We showed that this interaction is specific in two ways: (1) Only Ig domains 1-3, and not any of the other 50 Ig domains or any other segment of UNC-89, including the SH3-DH-PH domains, Fn3 domains or kinase domains interact with CPNA-1. (2) Only the copine domain (and 26 amino acids upstream and downstream) of CPNA-1, and not any other copine domain from the 6 other copine domain containing proteins in *C. elegans*, interacts with UNC-89. We identified four other proteins that interact with CPNA-1 at muscle focal adhesions, PAT-6 (actopaxin), LIM-9 (FHL), SCPL-1 (a CTD type phosphatase) and UNC-96. For UNC-89 and for UNC-96, the minimal portion of CPNA-1 required for interaction is essentially the copine domain. Nevertheless, analysis of mutant versions of the copine domain of CPNA-1 revealed that this domain has distinct binding sites for UNC-89 and UNC-96.

Previous studies have demonstrated that the copine domain is a protein-protein interacting domain: Using the copine domain of BON1/CPN1 as bait to screen a yeast 2-hybrid library of Arabidopsis cDNAs, Hua et al. (2001) identified BAP1 (BON1 association protein), which also contains a C2 domain. The interaction of the copine domain of BON1 and BAP1 was confirmed by an in vitro binding experiment (Yang et al. 2006). Recently, Li et al. (2010) have demonstrated that mutations in the copine domain of BON1 that reduce its interaction with BAP1, result in BON1 that fails to rescue a *bon1* loss of function mutant. Human Copine-III has been found to associate with ErbB2 in breast cancer cells (Heinrich et al. 2010). The interaction was identified using a phosphorylated peptide from ErbB2 in affinity pull-down followed by mass spectrometry.

Copine-III was also found to associate with RACK1 and to co-localize with phosphorylated focal adhesion kinase at the leading edges of cells. Interestingly, knockdown of Copine-III reduced cell migration. This study, however, did not demonstrate a direct interaction of Copine-III with either ErbB2 or RACK1, as the authors employed pulldowns with cell extracts. Tomsig et al. (2003) used the copine domains of three human copine proteins (copines I, II, and IV) to screen a mouse embryo 2-hybrid library and identified 21 interacting proteins. Some proteins interacted with only a single copine domain, whereas others were less specific. The interactions were verified using an in vitro binding assay with purified proteins. The targets fell into several functional categories, including regulators of protein phosphorylation, regulators of transcription, calcium binding proteins, regulators of ubiquitination, and cytoskeletal or structural proteins. The authors found that a majority of the interactors (14 of 21) contained sequences predicted to form coiled-coils. CPNA-1 interactors are also regulators of phosphorylation (UNC-89 has two kinase domains, SCPL-1 is phosphatase), and cytoskeletal proteins (PAT-6, LIM-9 and UNC-96). None of our interactors, however, contain predicted coiled-coil regions.

In conclusion, we have identified an unusual type of copine domain containing protein that lacks C2 domains, and thus is called “atypical”. We have shown that this protein, CPNA-1, is essential for sarcomere assembly and muscle development. At the molecular level, CPNA-1 was shown to link an integrin associated protein, PAT-6 (actopaxin) to the giant protein UNC-89 (obscurin) and to three additional proteins at the sarcomeric M-line. For this linking function of CPNA-1, the copine domain is required.

3.4 Materials and Methods

Screening of yeast two-hybrid library

Two-hybrid screening was performed as described in Miller et al. (2006). The first screen of the RB2 library of *C. elegans* random primed cDNAs (kindly provided by Robert Barstead, Oklahoma Medical Research Foundation), used the bait plasmid pGBDU-UNC-89 GX34 which contains coding sequence for UNC-89 Ig1-5. To make the Ig1-5 bait, an insert generated by PCR using primers GX3 with added BamHI site and GX4 with added XhoI site was ligated into pGBDU-C1. From a screen of 819,500 colonies, the following prey clones were isolated: 35 preys representing VIG-1, one prey from UNC-54 (residues 1-602), one prey from KETN-1 (residues 4394-4889), and one prey from CPNA-1 (called GX9-109, residues 173-1107). A second screen of the library used a bait plasmid which contains coding sequence for UNC-89 Ig2-5. This bait was generated by insertion of a PCR fragment using primers GX2/5 with added BamHI site and GX4. From a screen of 1,206,000 colonies, the following prey clones were isolated: 4 preys from VIG-1, one prey from UNC-54 (residues 1-602), one prey from MEL-26 (residues 25-343), and one prey from CPNA-1 (called B15-101, residues 825-1058). Interaction with VIG-1 was not pursued because it is part of a RISC complex (Caudy et al. 2003), and thus we believed that this did not make biological sense. We did not pursue interaction with UNC-54, also known as myosin heavy chain B, is located in the polar regions of the A-band (Miller et al. 1983), rather than the M-line where UNC-89 is located. Similarly, interaction with KETN-1 was not pursued, as KETN-1 is located in the I-band/dense body region of the sarcomere (Ono et al. 2006). Interaction of UNC-89 and MEL-26 will be described elsewhere.

Construction of 2-hybrid clones covering all of UNC-89-B and its screening with CPNA-1

Seventeen segments (Fig. 3-1B), covering all of UNC-89-B, were cloned into 2-hybrid bait and prey vectors. Three of these segments were described previously (Fn1-Ig52-PK1 and Ig53-Fn2-PK2 in Qadota et al. 2008a; Interkinase in Xiong et al. 2009). For the other segments, corresponding cDNA fragments of UNC-89-B were amplified by using PCR with primers listed in Table 1 and cloned into pBluescript. After confirming that the DNA sequence of each fragment was error-free, each fragment was cloned into pGBDU or pGAD yeast 2-hybrid plasmids. To screen for interaction with CPNA-1, first, PJ69-4A yeast host strains carrying each fragment of UNC-89-B in pGBDU were prepared, and then each of these strains were transformed with CPNA-1 prey clone B15-101 (residues 825-1058). The yeast 2-hybrid assay was performed by scoring growth on media lacking histidine or adenine.

Screening of yeast two-hybrid bookshelf of known M-line and dense body proteins

CPNA-1 was used as both bait and prey to screen a collection ("bookshelf") of 30 known components of nematode M-lines and dense bodies (Table 3). Both CPNA-1 (825-1058) and CPNA-1 (173-1107) were used.

Domain mapping of UNC-89 Ig1-5, SCPL-1, LIM-9, PAT-6, and UNC-96

To map the region of UNC-89 Ig1-5 minimally required for interaction with CPNA-1, the following deletion derivatives of Ig1-5 were generated by PCR using the following primers (sequences given in Tables 1 and 2): for Ig1-4, primers GX3 (5') and GX1/4 (3');

for Ig1-3, primers GX3 (5') and GX1/3 (3'); for Ig2-5, primers GX2/5 (5') and GX4 (3'); for Ig3-5, primers GX3/5 (5') and GX4 (3'); and for Ig2-3, primers GX2/5 (5') and GX1/3 (3'). In each case, the 5' primer has an added BamHI site, and the 3' primer has an added XhoI site for cloning each fragment into pBluescript. After identifying pBluescript clones lacking PCR induced errors, the fragments were cloned into pGBDU and then tested by 2-hybrid against GX9-109 library prey clone (CPNA-1 (173-1107). The domain mapping experiments summarized in Fig. 4 utilized deletion derivatives described previously: SCPL-1 (Qadota et al. 2008a), LIM-9 (Qadota et al. 2007), PAT-6 (Lin et al. 2003), and UNC-96 (Mercer et al. 2006; Qadota et al. 2007). Derivatives of SCPL-1, LIM-9 and PAT-6 in pGAD were tested against CPNA-1 (173-1107) in pGBDU. Derivatives of UNC-96 in pGBDU were tested against CPNA-1 (825-1058) in pACT.

Testing for interaction of UNC-89 Ig1-5 against copine domains of other C. elegans proteins

To test the specificity of interaction between UNC-89 Ig 1-5 and copine domain of CPNA-1, the copine domains of other copine domain containing proteins were generated by PCR using the primers listed in Table 2. The 5' primers have added BamHI or SmaI sites and the 3' primers have added XhoI sites. The products were ligated into pBluescript, and after identification of clones containing correct sequences, the inserts were excised and cloned into pGBDU and pGAD vectors. Yeast two hybrid assays were performed by using pGAD copine domain clones against pGBDU UNC-89 Ig 1-5 yeast strain, and pGBDU copine domain clones against the pGAD UNC-89 Ig1-5 clone respectively. The results of both assays are consistent and demonstrated in Fig. 3-2.

Testing for interaction of UNC-89 Ig1-5 and UNC-96 (201-418) against single amino acid mutants of the copine domain of CPNA-1

To test if the conserved amino acid residues within the copine domain of CPNA-1 are important for the binding with other proteins, four point mutants were generated by two rounds of PCR using the primers listed in Table 2. The first round of PCR used the following primer combinations: for mutant G922V, the first pair of primers are pGBDU-B15-5 (5') and G922V-2 (3') and the second pair of primers are G922V-1(5') and pGBDU-B15-3 (3'); for mutant Y985A, the first pair of primers are pGBDU-B15-5 (5') and Y985A-2 (3') and the second pair of primers are Y985A-1 (5') and pGBDU-b15-3(3'); for mutant Y985F, the first pair of primers are pGBDU-B15-5 (5') and Y985F-2 (3') and the second pair of primers are Y985F-2 (5') and pGBDU-B15-3(3'); and for mutant S1015A , the first pair of primers are pGBDU-B15-5 (5') and S1015A-2 (3') and the second pair of primers are S1015-1(5') and pGBDU-B15-3(3'). Each mutant primer pair has added a BamHI site for the 5' primer and an EcoRI site for the 3' primer. The products of the first round of PCR were gel purified and used as templates after 1: 10 dilution with water a the second round of PCR by using 5' primer pGBDU-B15-5 site and 3' primer pGBDU-B15-3. Then the products of the second round of PCR were gel purified and cloned into the pBluescript vector. After identification of error free clones, the mutant fragments were cloned into pGAD and tested against pGBDU UNC-89 Ig1-5 and pGBDU UNC-96 (201-418) yeast strains respectively. Mutant fragments were also cloned into pGBDU and tested against pGAD UNC-89 Ig1-5 and pGAD UNC-96 (201-418). The results of yeast two hybrid assays are demonstrated in Fig. 3-3B.

Demonstration of interactions using purified proteins

To verify the interaction of CPNA-1 with UNC-89 Ig1-3, CPNA-1 was pulled out of a worm lysate using GST-Ig1-3. To generate GST-Ig1-3, the insert from pGBDU-Ig1-3 was excised with BamHI and BglII and cloned into the BamHI site of pGEX-KK1. After identifying a clone with the proper orientation, GST fusion protein was prepared. A worm lysate was prepared as follows: a mixed population of wild type strain N2 was ground to a fine powder in a mortar and pestle in liquid nitrogen; this powder was added at a ratio of approximately 1:5 to 1 ml of a lysis buffer (20 mM Tris pH8.0, 10% glycerol, 0.5% NP-40, 2 mM EDTA, 150 mM NaCl, and protease inhibitors cocktail (Roche mini-Complete)), vortexed for 1 minute, and then debris was pelleted by spinning at top speed in a microcentrifuge at 4 degrees C for 10 minutes. Thirty μl of packed glutathione agarose beads (Sigma-Aldrich, cat. no. G4510) coated with either GST or GST-Ig1-3 (30 μgs each) were incubated with 500 μl worm lysate (approximately 5-10 mg/ml total protein) for 2 hours with mixing at 4 degrees C. The beads were then washed 3 times with cold lysis buffer, and then the bound proteins were eluted in 2X Laemmli sample buffer (vortexed, heated at 95 degrees 5 min, vortexed, spun hard in microcentrifuge, and supernatants were saved). Eluted proteins were separated by SDS-PAGE, blotted and reacted with either anti-CPNA-1 or anti-GST (Sigma-Aldrich, cat. no. A7340), and detected by ECL. An HA bead pulldown assay, performed essentially as described in Qadota et al. 2008b was used to verify the interaction between CPNA-1 (825-1058) and UNC-96 (201-418). In this assay yeast was used to express HA tagged CPNA-1 (copine), using primers noted in Table 2. Construction and purification of MBP-UNC-96 (201-418) was described in Mercer et al. 2006. Far western assays, performed essentially as

described in Mercer et al. 2006 were used to verify the interactions between CPNA-1 and UNC-89 Ig1-3, and between CPNA-1 and either SCPL-1, LIM-9 or PAT-6. To express MBP-CPNA-1 (825-1058), the B15-101 prey clone was digested with XhoI and the insert was cloned into the XhoI site of pMAL-KK2. After identifying a clone with proper orientation, MBP-CPNA-1 (825-1058) was expressed and purified. To express 6His-Ig1-3, pGBDU-Ig1-3 was digested with BamHI and BglII and the insert was ligated into pET28a. A clone with the proper orientation was used to produce 6His-Ig1-3. MBP-CPNA-1 (825-1058) was run on SDS-PAGE, transferred to a blot, incubated with 6His-Ig1-3, and then detected with anti-His and anti-rabbit-HRP. To express His-CPNA-1 (173-1107), pGBDU-CPNA-1 (173-1107) was digested with EcoRI (site in the vector) and SalI (in the insert), and this fragment, encoding the N-terminal 2/3 of CPNA-1 (173-1107) was ligated into pET28a, to create "clone a". The C-terminal 1/3 of CPNA-1 (173-1107) was generated by PCR using primers and having added HindIII and XhoI sites, and cloned into pBluescript. After finding an error free pBluescript clone, the insert was excised with HindIII and XhoI and ligated into pET28a. The resulting clone was digested with EcoRI (in the vector) and SalI (in the CPNA-1 insert), and the large fragment containing mostly vector sequence and the 3' portion of the insert was ligated to insert of "clone a" previously digested with EcoRI and SalI, to produce a plasmid that could produce His-CPNA-1 (173-1107). His-CPNA-1 (173-1107) was run on SDS-PAGE, transferred to a blot, incubated with either MBP-SCPL-1 (phosphatase domain; Qadota et al. 2008a), or MBP-PAT-6 (full length; Lin et al. 2003), or MBP-LIM-9 (LIM domains; Qadota et al. 2007), and then detected with anti-MBP-HRP. Procedures for growth of yeast for expressing HA tagged proteins, and growth of bacteria and purification of GST,

MBP and His tagged proteins were described previously (Qadota et al. 2008a; Mercer et al. 2006).

Generation of antibodies to CPNA-1

Residues 176-385 of CPNA-1b were expressed and purified in *E. coli* as a GST fusion protein. To do this, primers tag-149-1 and tag-149-2 with added EcoRI and XhoI sites were used to create a PCR fragment using cDNA GX9-109 as template. This fragment was cloned into pBluescript, and a clone without PCR-induced errors was used for excising the insert. This insert was ligated into pGEX-KK1 using the same enzyme sites. Using methods described in Mercer et al. (2006), GST-CPNA-1(176-385) was expressed, purified and shipped to Spring Valley Laboratories (Woodbine, MD) for generation of rabbit polyclonal antibodies. Most of the anti-GST antibodies were removed by immunoprecipitation using GST, and anti-CPNA-1 was affinity purified using Affigel conjugated with GST-CPNA-1(176-385), as described previously (Mercer et al. 2003).

Western blots

The procedure of Hannak et al. (2002) was used to prepare total protein lysates from wild type and from worms that had undergone pat-6 (RNAi) by L1 feeding (Miller et al. 2006). Approx. 50-100 μgs total protein were separated by a 10% SDS-PAGE, transferred to nitrocellulose membrane, and reacted with affinity purified and OP50 E. coli absorbed anti-CPNA-1 at 1:100 or 1:500 dilution, or affinity purified and OP50 E. coli absorbed anti-PAT-6 at 1:200 dilution, followed by reaction with appropriate HRP-

conjugated secondary antibodies and visualization using enhanced chemiluminescence (Pierce, cat. no. 32106).

Immunolocalization in adult body wall muscle

For nearly all immunostaining of adult body wall muscle, worms were fixed by the Nonet method (Nonet et al. 1993). The exception was to obtain the image shown at the bottom of Fig. 3-4C, in which strain DM5115 (UNC-112::GFP), was fixed using the constant spring method (Benian et al. 1996). Primary antibodies were used at the following dilutions: anti-CPNA-1 at 1:100, anti-UNC-89 (monoclonal MH42 (Benian et al. 1996)) at 1:200, anti-α-actinin (MH35 (Francis and Waterston, 1991) at 1:200, anti-myosin heavy chain A (MHC A)(5-6 (Miller et al. 1983)) at 1:200, and anti-PAT-6 at 1:100. For anti-CPNA-1, the secondary was anti-rabbit conjugated to Alexa488 (the one exception was when DM5115 was stained; then we used anti-rabbit-Cy3); for the monoclonals (MH35, MH42 and 5-6) the secondary was anti-mouse-Alexa 594; and for anti-PAT-6, we used anti-rat-Alexa594. Images were captured at room temperature with a Zeiss confocal system (LSM510) equipped with an Axiovert 100M microscope using an Apochromat 63x/1.4 oil objective in 2.5x zoom mode. The color balances of the images were adjusted with Adobe Photoshop.

RNAi

RNAi of *pat-6* and *unc-97* was performed by feeding wild type strain N2 worms bacteria expressing double-stranded RNA beginning at the L1 larval stage and continuing until the animals reached adulthood. Adults were then fixed and immunostained (Miller

et al. 2006). RNAi of *cpna-1* was conducted using the method of Meissner et al. (2009); adult escapers were fixed and immunostained.

Generation and characterization of antibodies to PAT-6

Residues 1-99 of PAT-6 were expressed and purified in E. coli as a GST fusion protein. Primers PAT-6-1 and PAT-6-99 with added BamHI and XhoI sites were used to amplify a fragment using the full length clone of *pat-6* in the 2-hybrid bait vector. The resulting fragment was cloned directly into BamHI and XhoI cut pGEX-KK1. After identification of a clone with an error-free insert, the clone was used to produce GST-PAT-6(1-99). This protein was shipped to Spring Valley Laboratories for generation of polyclonal antibodies in rats. Affinity purification was performed using Affigel conjugated to GST-PAT-6(1-99). Western blot extracts, as described above, were prepared from adults that had been subjected to RNAi by feeding beginning at the L1 stage, either with bacteria containing the empty vector or *pat-6* cDNA sequence, and subjected to western blot analysis using anti-PAT-6. Wild type adult muscle was co-stained with anti-CPNA-1 and anti-PAT-6 and imaged as described above.

Localization of CPNA-1 in embryonic muscle, and affect of various Pat mutants

Images of immunostaining of wild type and mutant embryos were kindly provided by our collaborators, Adam Warner (graduate student) and Don Moerman (professor) at the University of British Columbia, Vancouver, Canada.

Yeast 3-hybrid assay

The yeast three-hybrid assay was done essentially as described in Lin et al. (2003).

Sequence analysis

BLAST, using CPNA-1 as query against the *C. elegans* translated genome was used to identify the complete set of copine containing proteins in this organism. PFAM was used to identify the copine domain boundaries, and to identify other possible domains (e.g. C2 domains. The multisequence alignment of copine domains from *C. elegans* shown in Fig. 3-3B was performed using http://www.genome.jp/tools/clustalw/.

To predict possible transmembrane helices in the copine family, we used TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/).

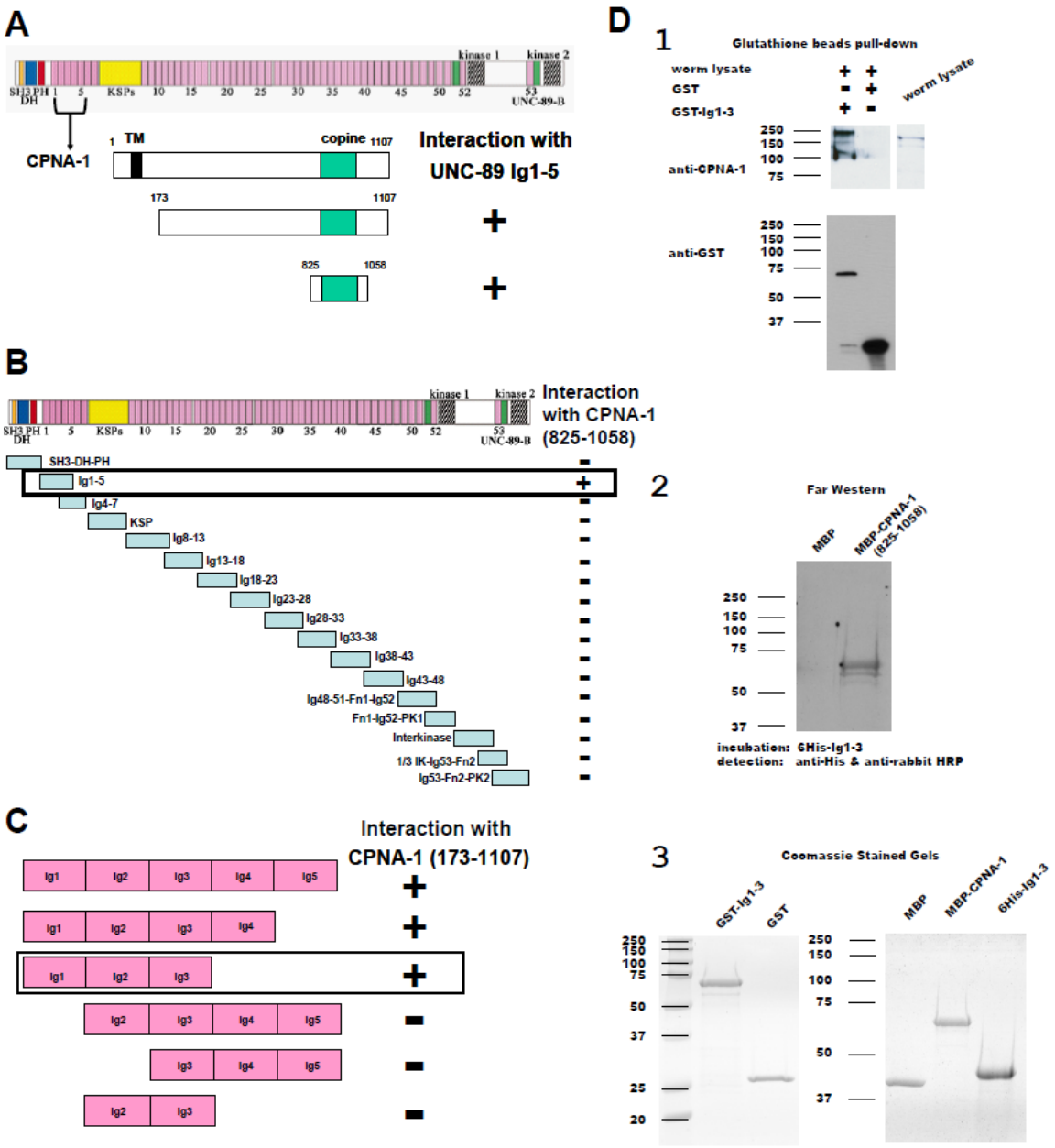


Figure 3–1 Ig domains 1-3 of UNC-89 interact with CPNA-1.

(A) Schematic representation of domains within the largest isoform of UNC-89, and indication that Ig1-5 was used as bait to screen a yeast 2-hybrid library. Two positive preys representing CPNA-1 were recovered, as indicated. CPNA-1, isoform b, has a predicted transmembrane domain (TM) and a copine domain (copine). (B) When CPNA-1 was used to test for interaction with the other 16 clones that fully cover UNC-89,

interaction was only found with Ig1-5. (C) Domain mapping of UNC-89 Ig1-5 shows that Ig1-3 are minimally required for interaction with CPNA-1 (173-1107). (D) Confirmation of the interaction using purified proteins. (1) CPNA-1 can be pulled out of a worm lysate using GST-Ig1-3. Glutathione agarose beads coated with either GST or GST-Ig1-3 were incubated with a worm lysate, washed, eluted and aliquots separated by SDS-PAGE and reacted with either anti-CPNA-1 or anti-GST. (2) 6His-Ig1-3 interacts with MBP-CPNA-1 by far western. SDS-PAGE was used to separate MBP (~40 kDa) and MBP-CPNA-1 (copine domain)(~60 kDa), transferred to membrane, incubagted with 6His-Ig1-3 in solution, washed, and detected with anti-His followed by anti-rabbit-horse radish peroxidase (HRP), and the reaction detected by ECL. (3) Coomassie stained gels of proteins used in the binding assays. The set of numbers along the left side of each blot or gel represent the position of molecular weight markers of the indicated sizes in kDa.

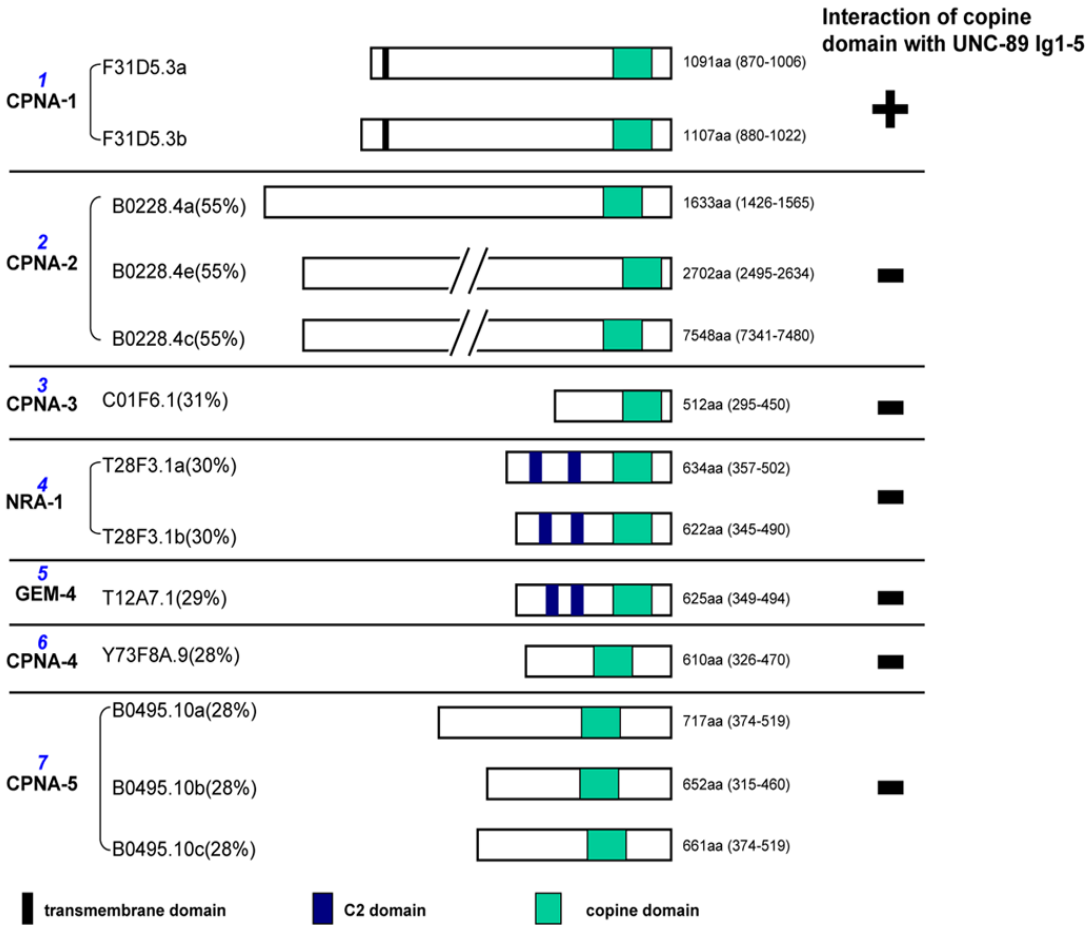


Figure 3–2 Copine family proteins in *C. elegans*.

The *C. elegans* genome has 7 genes (1-7) that encode proteins containing copine domains. The protein names shown are based on gene names (e.g. CPNA-1) and sequence names (e.g. F31D5.3). Percentages refer to the percent of identical amino acid residues in each copine domain, as compared to the copine domain of CPNA-1. The numbers in parentheses after the total number of amino acid residues in each protein, denote the positions of copine domains. Note that only NRA-1 and GEM-4 are "typical" in that they also contain C2 domains. Only CPNA-1 also has a predicted transmembrane domain. Also, some isoforms of CPNA-1 and CPNA-2 are not shown. These isoforms, predicted

on WormBase (CPNA-1c, CPNA-1d, CPNA-2b, and CPNA-2d) lack copine domains. The right-most column represents results of yeast 2-hybrid assays in which UNC-89 Ig1-5 was tested for interaction with copine domains from each of the 7 *C. elegans* genes. +, interaction; -, no interaction.

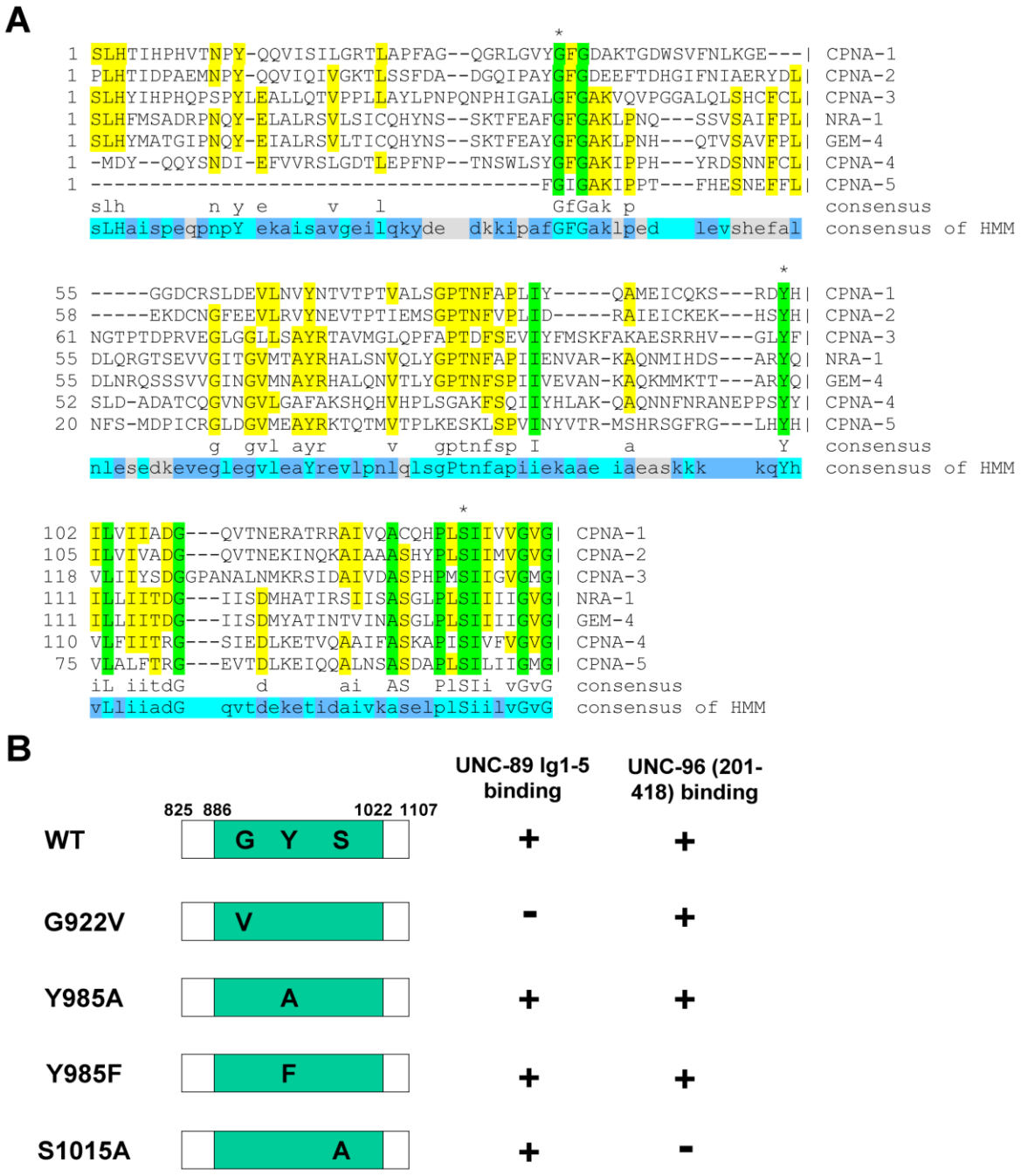


Figure 3–3 Multi-sequence alignment of copine domains in *C. elegans* copine domain containing proteins and the point mutation of copine domain affects its binding to UNC-89 or UNC-96.

(A) Multi-sequence alignment of the copine domains in the 7 *C. elegans* copine domain-containing proteins. Green shading indicates residues present in all 7, and yellow shading

indicates residues present in 4 or more proteins. “Consensus” represents the green (uppercase) and yellow (lowercase) shaded residue. Beneath this consensus for the *C. elegans* copine domains, is the consensus generated by PFAM for all copine domains across all species. Asterisks denote conserved residues that we have mutated in the copine domain of CPNA-1 and tested for binding to UNC-89 and UNC-96. (B) copine domain has distinct binding sites for UNC-89 and UNC-96. Three of these 12 conserved residues (indicated with asterisks in (A)): G922 was changed to V. Y985 was changed to A and F respectively. S1015 was changed to A. G922V mutation eliminates the binding to UNC-89 but not UNC-96. Similarly, S1015A mutation eliminates the binding to UNC-96 but not UNC-89.

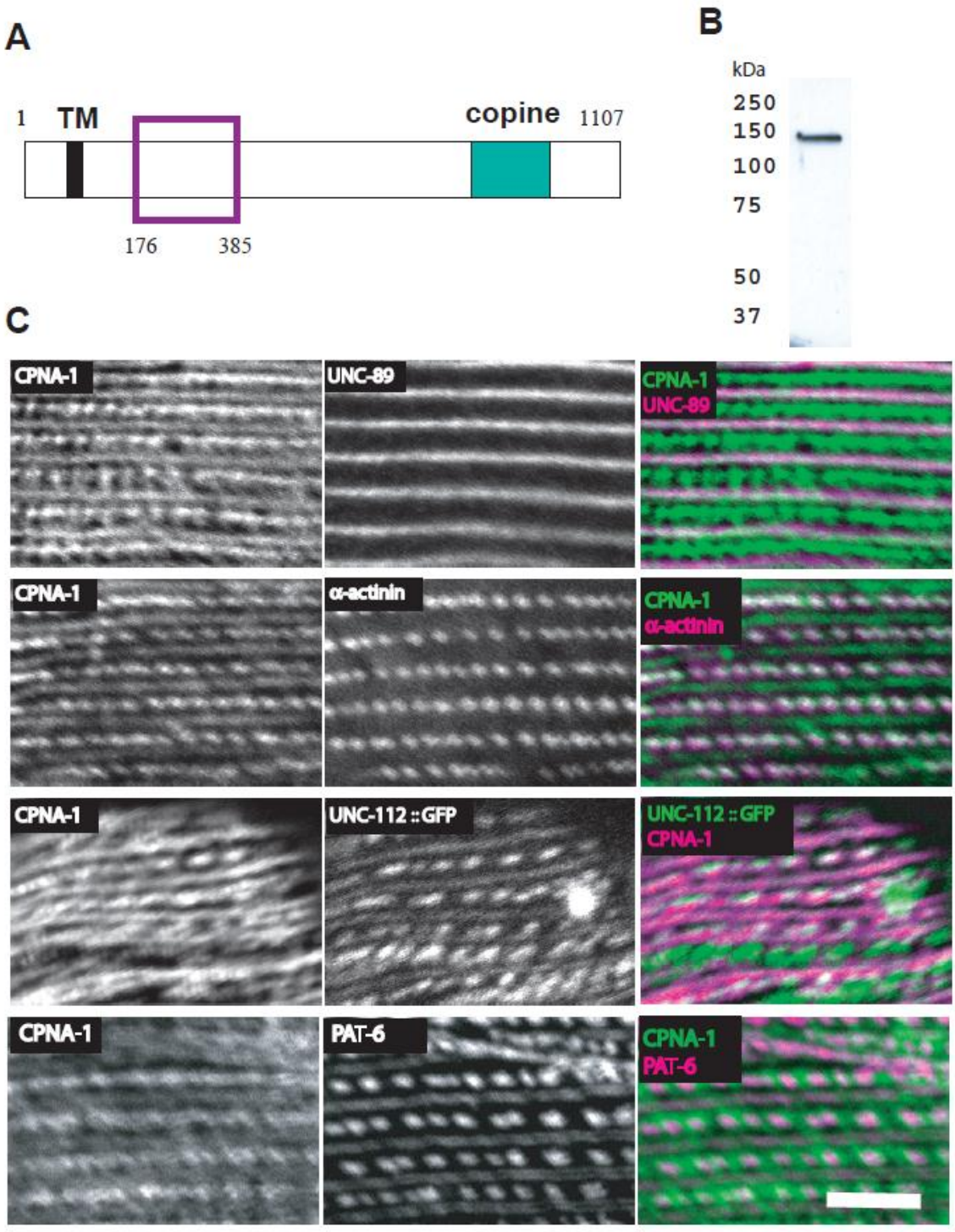


Figure 3–4 Immunolocalization of CPNA-1 to adult muscle M-lines and dense bodies.

(A) Schematic of the CPNA-1 protein indicating the region (residues 176-385; purple box) used as immunogen to generate rabbit polyclonal antibodies. (B) Immunoblot of Laemmli-soluble proteins from a mixed stage population of *C. elegans* reacted with anti-CPNA-1 and visualized by ECL. A single band of ~130 kDa, the size expected from the CPNA-1b predicted isoform, is detected. (C) Localization of anti-CPNA-1 in adult body wall muscle. Wild type adults were co-stained with anti-CPNA-1 and anti-UNC-89 (first row), or anti-CPNA-1 and anti-α-actinin (second row); a transgenic line carrying UNC-112::GFP was also stained with anti-CPNA-1 (third row); co-staining with CPNA-1 and PAT-6 is shown in the fourth row. CPNA-1 clearly localizes to both M-lines (co-localizing with the M-line protein UNC-89), and to dense bodies (co-localizing with the dense body protein α-actinin). At least some CPNA-1 localizes near the muscle cell membrane, since some CPNA-1 co-localizes with the membrane-proximal proteins UNC-112 and PAT-6. Co-localization is indicated by white (overlap of green and magenta). Bar, 10 μm.

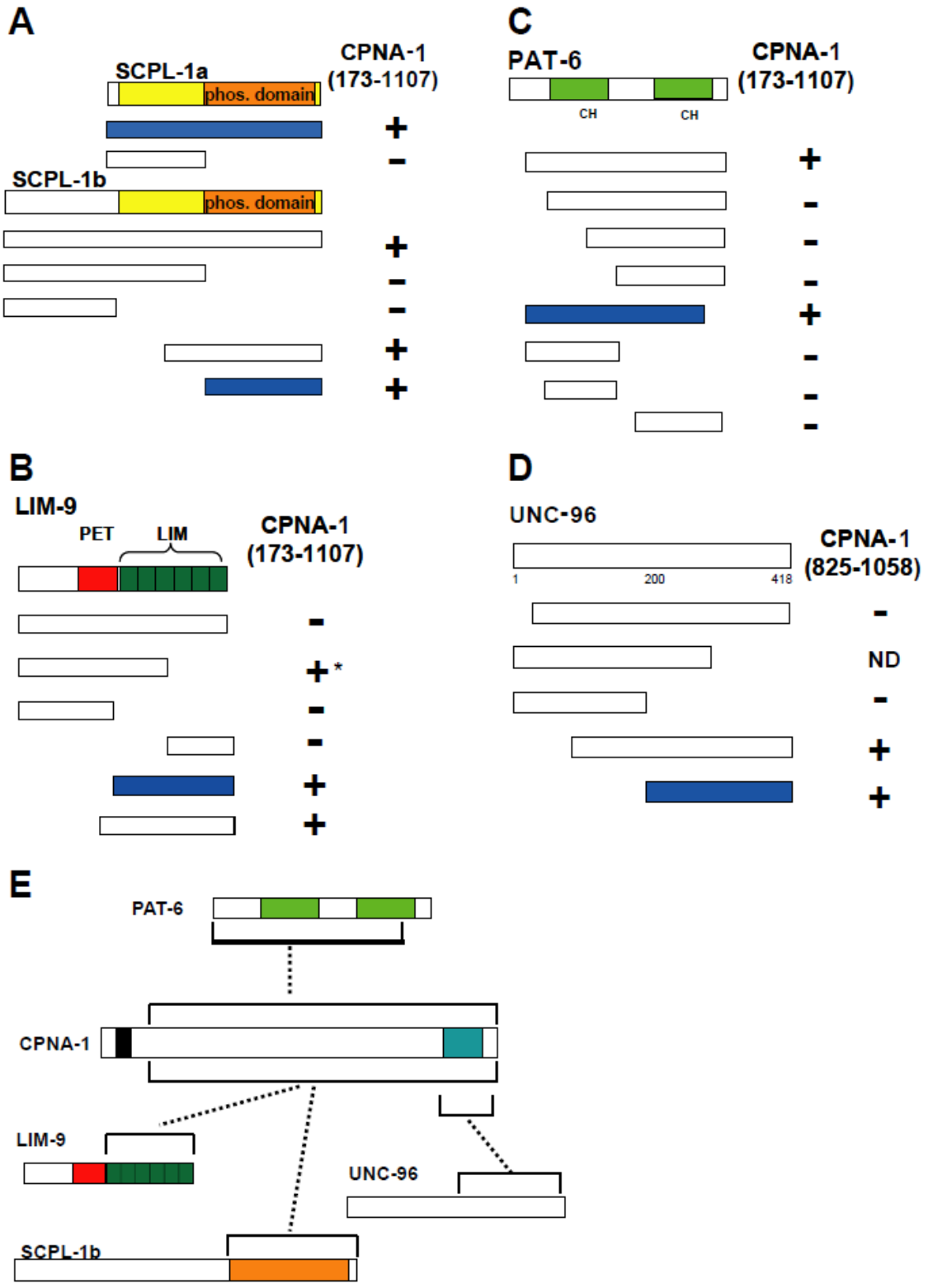


Figure 3–5 CPNA-1 interacts with several known M-line and one M-line/dense body protein.

These proteins were identified using the 2-hybrid method in which CPNA-1 was used to screen a collection of 30 known components of nematode M-lines and dense bodies. (A-D) Depiction of the results of domain mapping to determine the minimal region (indicated as a blue bar) of each protein required for interaction with the indicated regions of CPNA-1. (E) Summary of the results showing protein domains of each protein, and which regions are involved in the interactions.

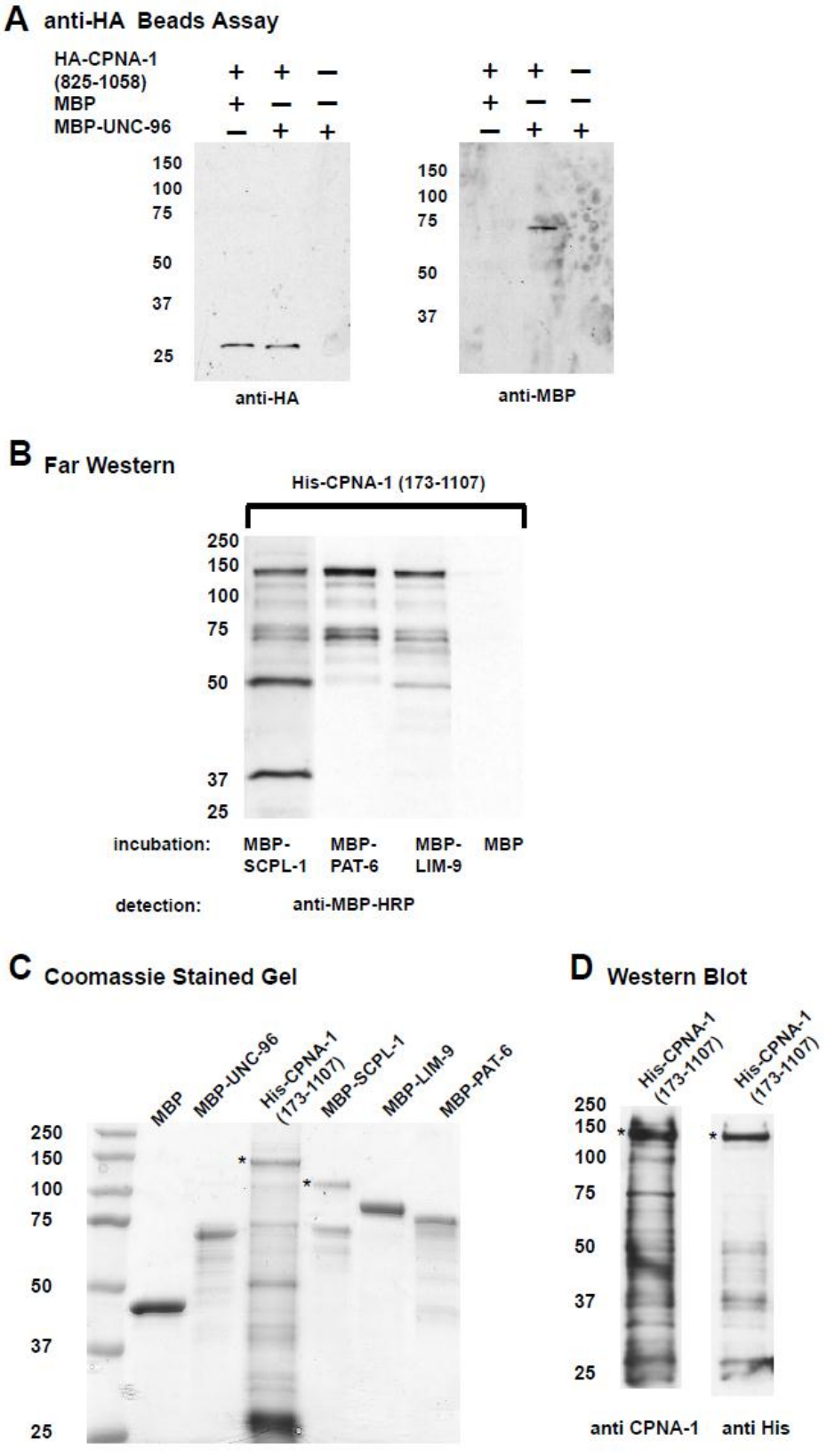


Figure 3–6 Confirmation of interactions using purified proteins.

(A) CPNA-1 (825-1058) interacts with UNC-96 (201-418) in vitro. Yeast expressed HA-tagged CPNA-1(825-1058) was precipitated with anti-HA beads, washed, incubated with bacterially expressed MBP or MBP-UNC-96 (201-418), washed, and the proteins eluted, separated on a gel, blotted, and reacted with either anti-HA or anti-MBP. (B) CPNA-1 (173-1107) interacts with SCPL-1, PAT-6 and LIM-9 in vitro. SDS-PAGE was used to separate 2 μg of His-tagged CPNA-1(173-1107) in each of 4 lanes, and the proteins were transferred to a membrane. One blot strip was incubated with MBP-SCPL-1 (phosphatase domain), one with MBP-PAT-6 (full length), one with MBP-LIM-9 (LIM domains) and the final one with MBP. After washing, each blot strip was incubated with antibodies to MBP coupled to horseradish peroxidase (anti-MBP-HRP), and reactions visualized by ECL. (C) A Coomassie-stained 10% SDS-PAGE shows 2 μg of each bacterially expressed protein used in the anti-HA bead assay and far westerns. Asterisks denote the positions of full-length fusion proteins for His-CPNA-1(173-1107) and MBP-SCPL-1. (D) The multiple protein bands detected by far western likely represent degradation products of His-CPNA-1 (173-1107). His-CPNA-1(173-1107) was separated on a gel, blotted and reacted with either anti-CPNA-1 or anti-His. Although the main reaction (indicted by asterisks) of each antibody is to full-length His-CPNA-1(173-1107), expected to be ~140 kDa, reaction to other bands are detected. For each blot or gel, the positions of molecular-weight markers in kDa are indicated.

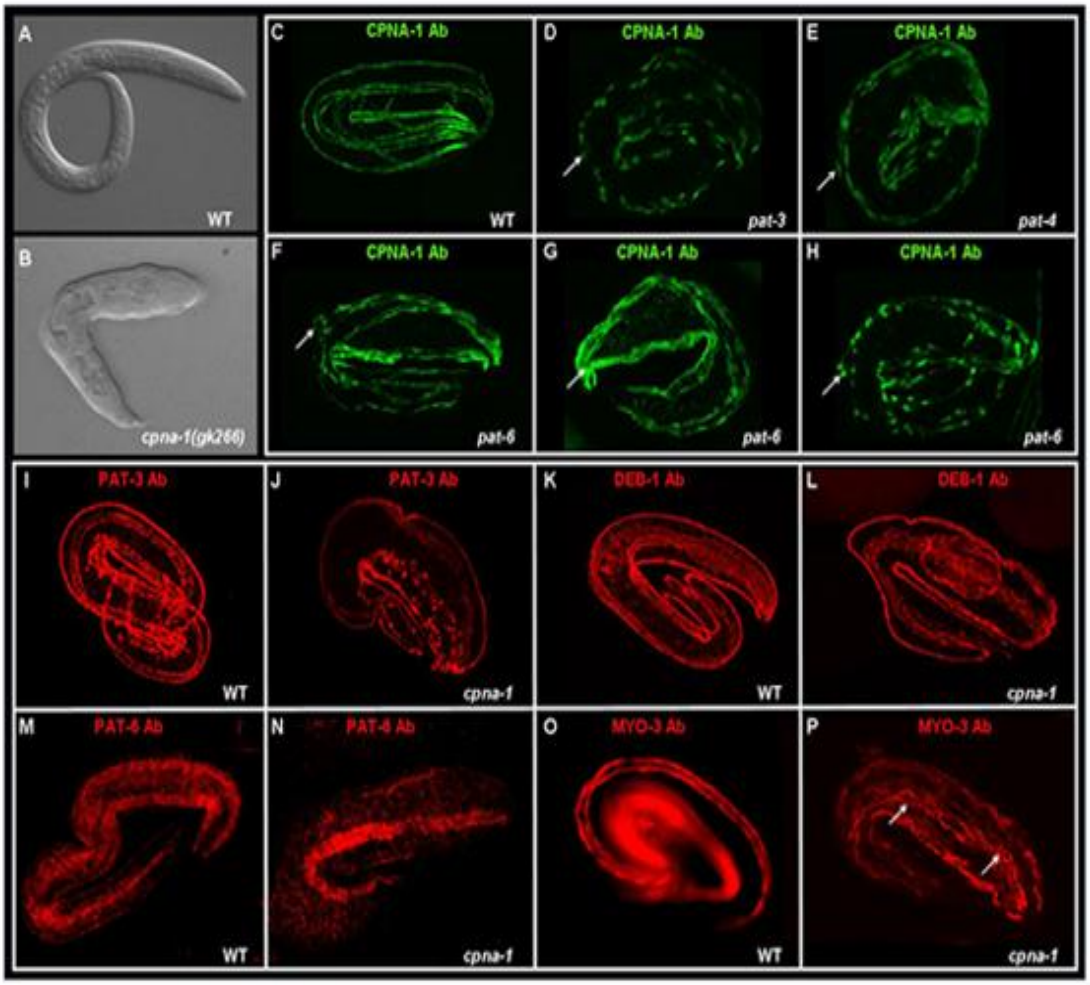


Figure 3–7 Loss of function of *cpna-1* results in a Pat embryonic lethal phenotype. Animals homozygous for the *cpna-1* g*k266* allele (B) are Paralyzed, and Arrest at the Two-fold stage of embryogenesis (PAT), a phenotype observed in loss of function mutants for numerous essential body wall muscle genes. Wild type worms (A) of similar age hatch and progress normally through the larval stages. Using a CPNA-1 polyclonal antibody, we looked at the localization of CPNA-1 in the mutant background of a number of essential muscle genes. In *pat-3* (C), and *pat-4* (D) arrested embryos, CPNA-1 is mislocalized, clumping in the middle of muscle cells (arrows). CPNA-1 shows some variability in its localization in *pat-6* arrested embryos, ranging from disorganization (F),

to severe mislocalization, similar to that seen in the *pat-3* and *pat-4* backgrounds. We also looked at the localization pattern of essential body wall muscle proteins in the *cpna-1* background. For each of PAT-3 (J), DEB-1 (L), and PAT-6 (N), each protein was localized properly into adhesion complexes in bands of muscle quadrants, albeit with some minor disorganization when compared to wild type embryos (I, K, M). When comparing the tight organized arrays of thick filaments in MYO-3 stained wild type embryos (N) to those in *cpna-1* arrested embryos, severe disorganization (arrows) of MYO-3 is observed (O). These results place *cpna-1* in between *pat-6* and *myo-3* in the dense body/M-line assembly pathway. (From Adam Warner)

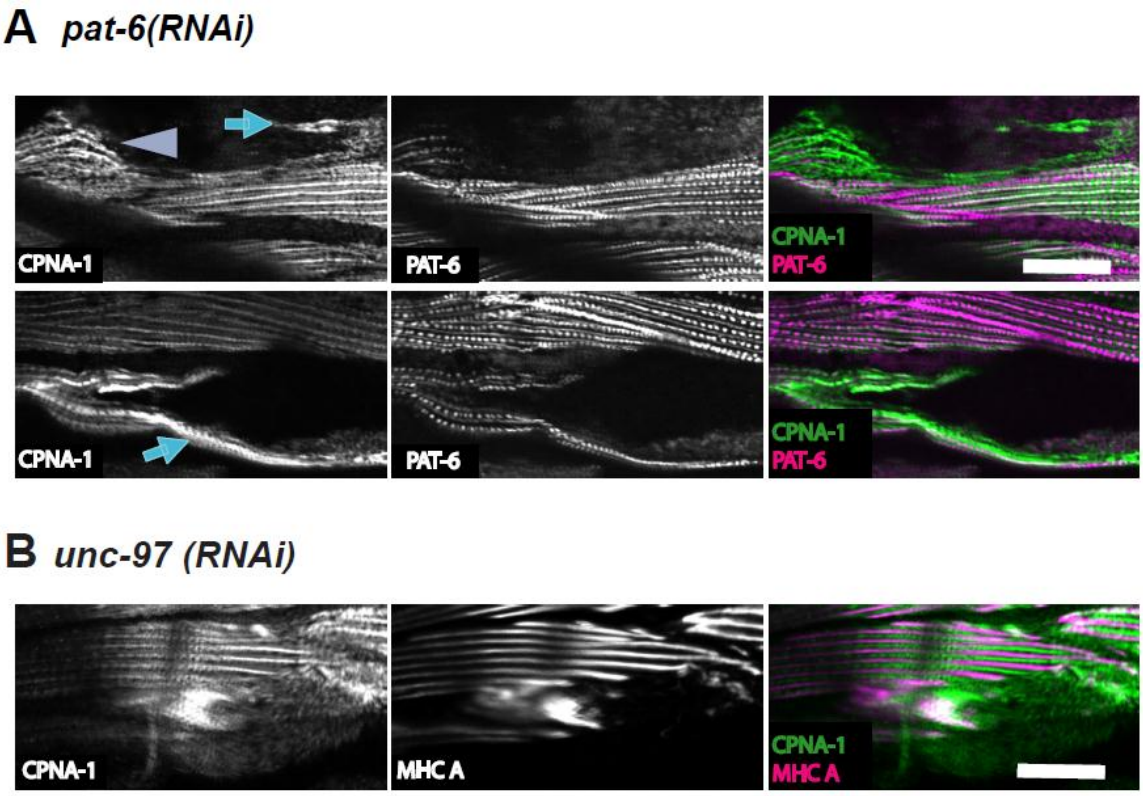


Figure 3–8 In adult muscle, PAT-6 directs the assembly CPNA-1, and CPNA-1 directs the assembly of UNC-89.

(A) RNAi was used to knockdown *pat-6* beginning at the L1 larval stage, and the resulting adults were immunostained for PAT-6 and CPNA-1. Top and bottom rows show portions of body wall muscle from two such animals. In muscle cells in which PAT-6 was knocked down, CPNA-1 is found in aggregates (arrowhead) or abnormally located at the edge of the muscle cell near the muscle cell membrane (arrows). (B) RNAi was also used to knockdown *unc-97* beginning at the L1 larval stage, and the resulting adults were immunostained for MHC A (myosin) and CPNA-1. In the image, the bottom cell shows that CPNA-1, and downstream MHC A, are found in a large aggregate, but not localized to the edge of the cell. Bars, 10 μm.

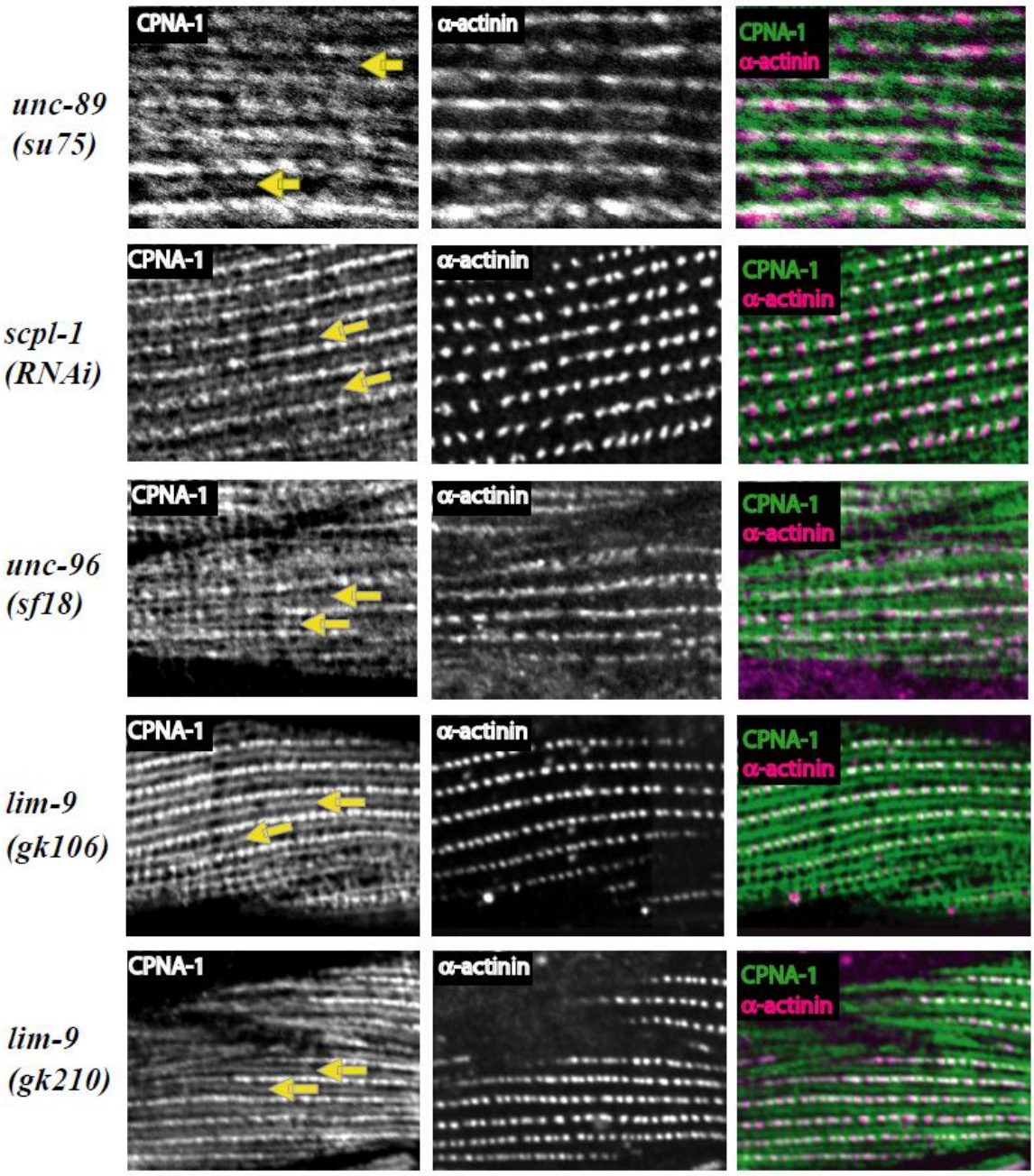


Figure 3–9 Analysis of mutants places the M-line proteins UNC-96, LIM-9 and SCPL-1 downstream of CPNA-1 in adult muscle.

The indicated adult loss-of-functions mutants or RNAi animals were co-immnunostained with anti-CPNA-1 and anti-α-actinin. The localization of CPNA-1 at M-lines (indicated by yellow arrows) is unaffected by the absence or reduced levels of UNC-89, SCPL-1,

UNC-96 or LIM-9. Note that the *unc-89* mutant allele, *su75*, which lacks all large UNC-89 isoforms, lacks CPNA-1 binding sites. Bar, 10 μm.

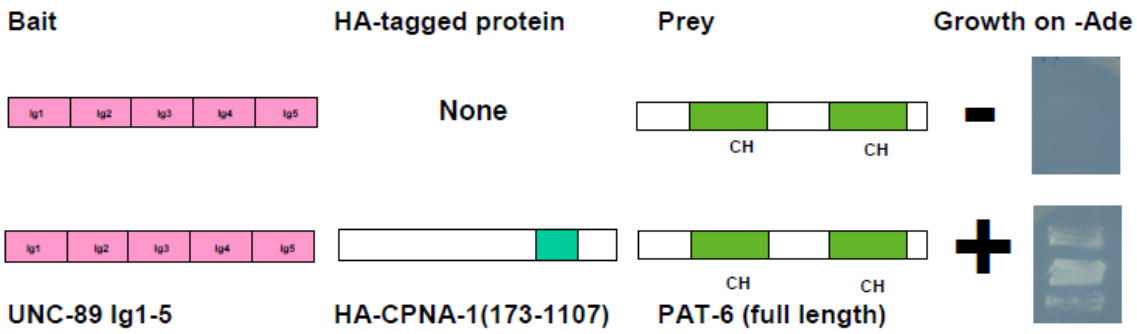


Figure 3–10 A ternary complex containing PAT-6, CPNA-1 and UNC-89, demonstrated by a yeast 3-hybrid assay.

(A) UNC-89 Ig1-5 as bait was co-expressed with HA-tagged CPNA-1 (residues 173-1107) (or empty vector as control), and PAT-6 (full length) as prey. +, growth on –Ade plates; -, no growth on –Ade plates. The right-hand panel shows the yeast growth on –Ade plates from each experiment from 3 independent colonies.

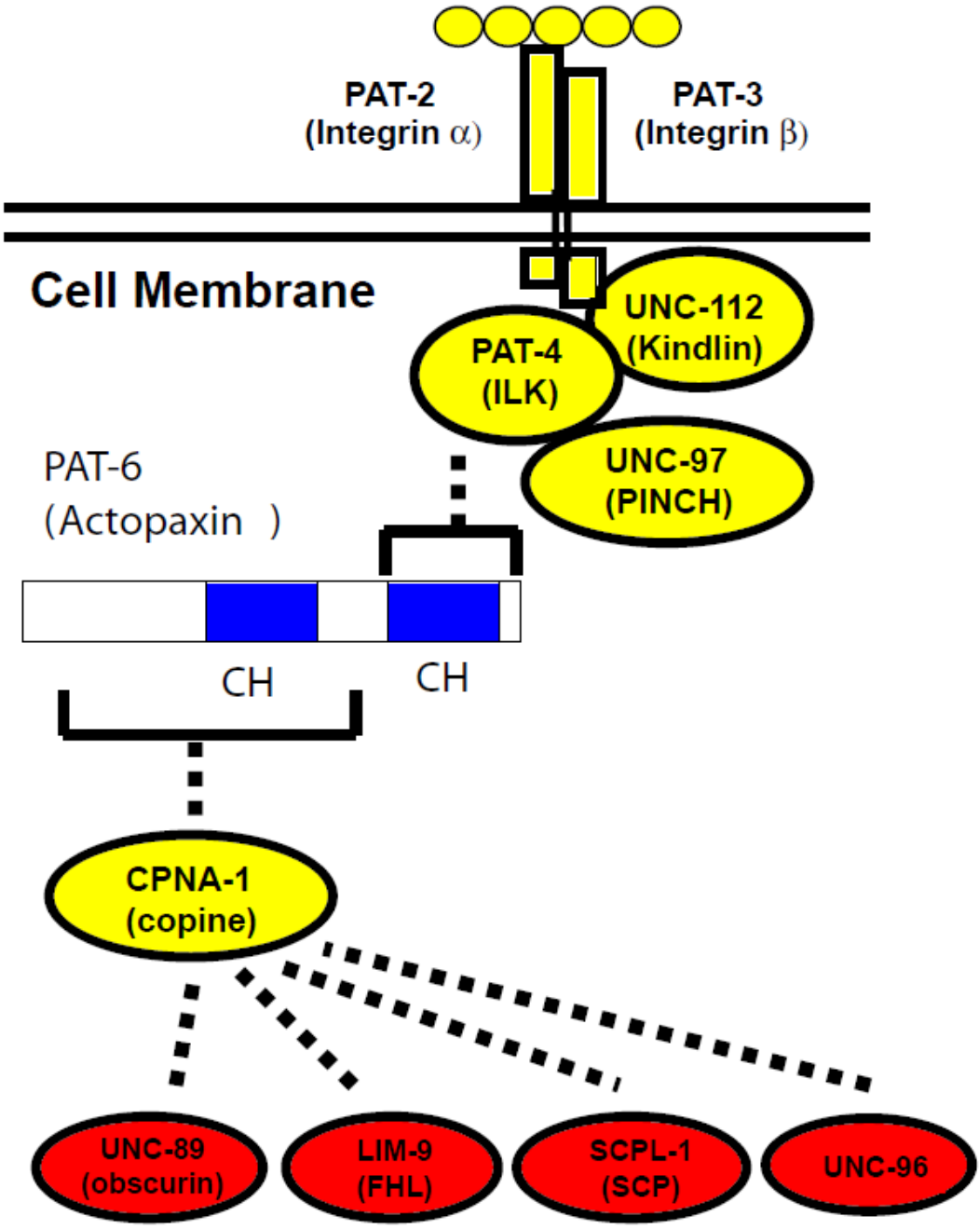


Figure 3–11 Model suggested by this study to explain the role of CPNA-1 in assembly of integrin adhesion complexes.

UNC-112 (Kindlin), PAT-4 (ILK), UNC-97 (PINCH) and PAT-6 (actopaxin) are a 4-protein complex associated with the cytoplasmic tail of β-integrin. Lin et al. (2003) showed that the 2nd CH domain of PAT-6 interacts with PAT-4. Here we show that PAT-

6 (actopaxin) directs the assembly of CPNA-1 to the muscle focal adhesions. CPNA-1, in turn, directs assembly of UNC-89 (obscurin), LIM-9 (FHL), SCPL-1 (SCP) and UNC-96 to M-lines. Yellow indicates proteins that are located at both M-lines and dense bodies, and whose loss of function result in Pat embryonic lethality. Red indicates proteins that are located at M-lines, and whose loss of function result in adult muscle phenotypes.

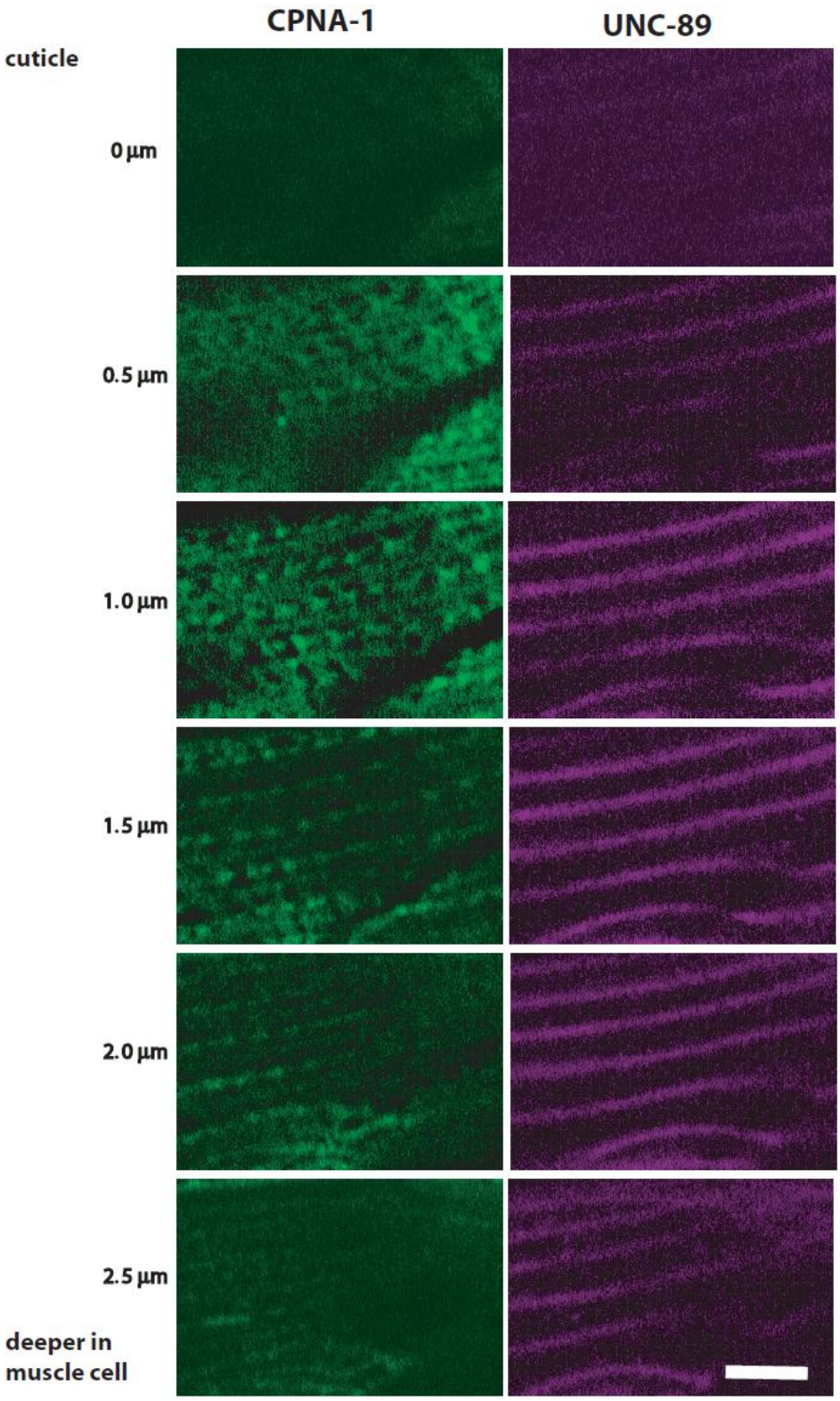


Figure 3–12 Confocal Z-sectioning suggests that some CPNA-1 is located close to the muscle cell membrane.

Adult worms were co-stained with anti-CPNA-1 and anti-UNC-89 (monoclonal MH42; Benian et al. 1996). Using confocal microscopy, 0.5 μm optical sections, from the cuticle side to deep into the body wall muscle cell, were obtained, as shown. Although UNC-89 appears to be located throughout all the sections, CPNA-1 is located closer to the cuticle, consistent with our co-localization with UNC-112 and with PAT-6, close to the muscle cell membrane.

Table 2 Primers for UNC-89 Bookshelf

Name	Sequence	UNC-89 Segments
Bam-unc-89(N)	GGCGGATCCATGGCTAGTCGACGCCAAAAG	SH3-DH-PH
unc-89(N)-Xho	CGGCTCGAGTTAACCAGGGAACAAGCTGCTCTT	SH3-DH-PH
GX3	GGACGGATCCATTGATTGGACAACAACTGGAAC	Ig1-5
GX4	CCGTCTCGAGTCACTCCACGAATTTCGGTGGTTTC	Ig1-5
GX25	GCGGATATCAAAGTCAGAGTTCCACCAGCA	Ig4-7
GX6	CGCCTCGAGTTACTTCTTCGGCTCGATCGTCAA	Ig4-7
89KSPYF	……..GTAAAAGAAGCTTCTCCTGAGG	KSP
89KSPYR	……..TGACTTCTTCTTCTTTGGTGAGC	KSP
GX7	GCGGATATCCCGGAAGCCGAAAAGCCTCCA	Ig8-13
GX8	CGCCTCGAGTTATGCATCACCGACCTTTTTAAC	Ig8-13
GX9	GCGGATATCAAGGTTGACAAAAAGACTGAA	Ig13-18
GX10	CGCCTCGAGTTACTTCTCCAATGATTCAATGAC	Ig13-18
GX11	GCGGATATCCGTACTCCAACTCCAGTTATG	Ig18-23
GX12	CGCGTCGACTTATGGAATAGAATCCCTGATGAT	Ig18-23
GX132	GCGGATATCGCAGTTGTCAAGAACTTGGTTCCA	Ig23-28
GX142	CGCCTCGAGTTAAACTTGAGCTGGTTCCTTGACTGT	Ig23-28
GX151	GCGGATATCGAAGGAGCTCCAAAGATTGAC	Ig28-33
GX161	CGCCTCGAGTTAAACAATAGCGATTCCAGGAAG	Ig28-33
GX17	GCGGATATCGACGATGGAAAGGATAAGGTG	Ig33-38
GX18	CGCGTCGACTTACAGAACATTTGGCTTGACAAC	Ig33-38
GX19	GCGGATATCAAGGGAGAGGTTGATGAGAAA	Ig38-43
GX20	CGCCTCGAGTTATTCCTGCTTCGGAACCTTGAC	Ig38-43
GX212	GCGGATATCGAGGAGAAGCGCCGAGAATATGCT	Ig43-48
GX222	CGCCTCGAGTTAAATAGTTCGGTCTCCTTCGATAAG	Ig43-48
GX23	GCGCCCGGGTCGTCCATTCGTGAAGGAAAG	Ig48-51-Fn1-Ig52
GX24	CGCCTCGAGTTAGTAGTCGCCTTCAGCTTCAAC	Ig48-51-Fn1-Ig52
YIKL	GTACCCGGGTCTCCACGCCGTTCCACT	1/3 IK-Ig53-Fn2
KJW1	CGCCTCGAGTCACTTGAACTTGAG	1/3 IK-Ig53-Fn2

Table 3 Other Primers Used in Chapter 3

Name	Sequence	Purpose
GX1/4	CCGTCTCGAGTCACAAGAAAGTCGGCGC	UNC-89 Ig1-4(3')
GX1/3	CCGTCTCGAGTCAATTAAAGGTTGGTGG	UNC-89 Ig1-3(3')
GX2/5	GGACGGATCCGCATCCACTTCTGCCTTTTTC	UNC-89 Ig2-5(5')
GX3/5	GGACGGATCCGAAAGCAAGGCTGAGCTC	UNC-89 Ig3-5(5')
TAG149-1	GTACGAATTCGAGCAAATTCCATCTATTGCACCG	CPNA-1 Antigen(5')
TAG149-2	GATGCTCGAGCTACTTCTTCAGACGGTCAGCTTC	CPNA-1 Antigen(3')
GX3-SmaI	GGACCCCGGGATTGATTGGACAACAACTGGA	UNC-89 HA-Ig1-3(5')
pGBDU-B15-5	CGATGAATTCCAGAGCTCACTTCTATCAGTA	HA-CPNA-1(COPINE,5')
pGBDU-B15-3	CGATGGATCCTCAATTTGTCGACTTCTTCAC	HA-CPNA-1(COPINE,3')
tag308-5'	GGACCCCGGGCGCCGAAACTCTCTTCTCCAG	CPNA-2 COPINE(5')
tag308-3'	CCCGTCTCGAGTCAATTTGGTGCATTGAACATTAC	CPNA-2 COPINE(3')
C01F6.1-5'	GGACCCCGGGAATTCTTACATGCTGATGGAG	CPNA-3 COPINE(5')
C01F6.1-3'	CCGTCTCGAGTCAGATCATTGCTAACACATCACT	CPNA-3 COPINE(3')
tag64-5'	GGACCCCGGGGAGACTATTATGCTCTAC	CPNA-4 COPINE(5')
tag64-3'	CCGTCTCGAGTCATTTTGTGGCGTTTACAAACTG	CPNA-4 COPINE(3')
tag178-5'	GGACCCCGGGTATTCTTTCTTGGATTATATT	CPNA-5 COPINE(5')
tag178-3'	CGTCTCGAGTCACTGAATTGGCGGTTTTGCAGC	CPNA-5 COPINE(3')
NRA-1-5'	GCATGGATCCCTCACCAACGAGAAGAAGAAG	NRA-1 COPINE(5')
NRA-1-3'	CGTAGTCGACCTACTCGCGCATCGTTACAAACTG	NRA-1 COPINE(3')
GEM-4-5'	GCCGATGGATCCCTGATTAATGAAAAAAAGAAG	GEM-4 COPINE(5')
GEM-4-3'	CGTAGTCGACCTAATCACGCATTGTCACAAATTG	GEM-4 COPINE(3')
CPNA-1Hind3 G	GACTAAGCTTCTCTTGGCAATCACCTCTCTC	HIS-CPNA(5')
CPNA-1F	CGATCTCGAGTTACAGCGGATTATGCATCAT	HIS-CPNA(3')
G922V-1	AGACTTGGAGTCTATGTATTCGGAGACGCAAAG	G922V mutant (5')
G922V-2	CTTTGCGTCTCCGAATACATAGACTCCAAGTCT	G922V mutant (3')
Y985A-1	TCAAAAGTCCAGAGATGCCCACATCCTTGTAATC	Y985A mutant (5')
Y985A-2	GATTACAAGGATGTGGGCATCTCTGGACTTTTGA	Y985A mutant (3')
Y985F-1	CAAAAGTCCAGAGATTTCCACATCCTTGTAATC	Y985F mutant (5')
Y985F-2	GATTACAAGGATGTGGAAATCTCTGGACTTTTG	Y985F mutant (3')
S1015A-1	CTGCCAGCATCCTCTCGCGATTATTGTCGTCGG	S1015A mutant (5')
S1015A-2	CCGACGACAATAATCGCGAGAGGATGCTGGCAG	S1015A mutant (3')
PAT6-1	GTCGGATCCATGTCAACACTTGGTCGTAGT	PAT-6 antigen(5')
PAT6-99	GACCTCGAGTTATGCCAACTTCTGATCGCGGGC	PAT-6 antigen(3')

4. Chapter 4: Summary and Future Directions

Sarcomeres are the most fundamental contraction units of muscle. Although actin and myosin are the main protein components of sarcomeres, there are hundreds of other proteins involved in assembly and maintenance of the sarcomere, and the regulation of muscle contraction. Indeed, new components continue to be discovered each year. The exact molecular mechanisms by which these many components are assembled into the highly ordered sarcomeres and how the sarcomeres are maintained during the stress of muscle contraction are unknown. Moreover, it is becoming increasingly apparent that mutation of many of these sarcomeric proteins is the cause of many human diseases of skeletal and cardiac muscle. For example, different mutations in the giant protein titin result in either a late onset mild myopathy (Bonnemann and Laing, 2004), or an early onset myopathy with respiratory failure (Lange et al. 2005). Familial hypertrophic cardiomyopathy results from mostly dominant mutations in sarcomeric proteins; to date, mutations in the genes that encode for at least 20 different proteins are involved (Callis et al. 2010). Despite information on which protein is deficient or mutant, in nearly all cases, we do not understand the mechanisms by which the disease manifestations occur. Thus, having a better understanding of the normal functions and interactions of sarcomeric proteins will shed light on pathogenesis and lead to specific treatments for these many diseases of muscle.

Fascinatingly, the largest polypeptides known are found in muscle sarcomeres. These proteins, all larger than 700,000 Da, are composed of a combination of multiple copies of Ig and Fn3 domains, and have one or two protein kinase domains. The founding member of this family of proteins is twitchin, from *C. elegans,* and most of this pioneering work

was done in the Benian lab. UNC-89 was more recently discovered, also first in *C. elegans*, and its human homolog, obscurin, was discovered about 5 years later. I am intrigued by the following questions about UNC-89: Do the multiple domains of UNC-89 indicate multiple functional roles for this giant polypeptide? There are 53 Ig domains in the largest isoform of UNC-89, and do many of these domains, which form the same basic β-barrel structure, have different functions? It is a fairly unique feature that UNC-89 has two protein kinase domains (PK1 and PK2), rather than one protein kinase domain? Previous molecular modeling suggests that PK1 is catalytically inactive, whereas PK2 is catalytically active. What is the function of the catalytically inactive PK1? What is the substrate for PK2? What is the purpose of the partial PK1 domains lying at the N-termini of UNC-89 isoforms C and D? My thesis has provided insight into how UNC-89 becomes localized to the sarcomeric M-line, and once assembled, how UNC-89 acts as a scaffold for the assembly of additional sarcomeric proteins.

My project was focused on two regions of UNC-89: the N terminal Ig domains 1-5 and the C terminal protein kinases and interkinase region. By conducting screens of a yeast 2-hybrid library, I discovered new binding partners for Ig domain 1-5 of UNC-89--CPNA-1 and MEL-26 and focused on the function of CPNA-1. MEL-26, a known interactor of cullin-3 (a scaffold for assembly of the ubiquitin mediated degradation machinery) is being profitably studied by a postdoc. in our lab, Kristy Wilson. Our lab previously reported that SCPL-1 (CTD type phosphatase) interacts with both protein kinase domains of UNC-89 (Qadota et al. 2008a). I discovered that LIM-9 (FHL in humans) interacts with the PK1 protein kinase domain and interkinase region of UNC-89, and that LIM-9

interacts with SCPL-1. I will summarize my research findings and future directions in the following sections.

4.1 CPNA-1

CPNA-1 is a copine domain containing protein and it has four predicted isoforms in wormbase. Besides CPNA-1, I found six other genes that encode copine domain containing proteins in *C. elegans*. I classified, for the first time, the copine domain containing proteins in *C. elegans* into two categories: typical copine containing proteins and atypical copine containing proteins. The typical copine containing proteins such as NRA-1 and GEM-4 have, in addition to a copine domain, two C2 domains, which are Ca^{2+} -dependent phospholipid binding domains and are generally involved in membrane trafficking. The atypical copine containing proteins have no C2 domains and are named as CPNA (copine protein atypical). Interestingly, according to sequence analysis, CPNA-1 may be the only protein that has a predicted N terminal transmembrane domain among CPNA members. The significance and function of the predicted transmembrane domain of CPNA-1 will be investigated in the future. I began to investigate this question by making three constructs intended to express the following fragments of CPNA-1 in transgenic worms with either HA or GFP tags: (1) essentially the copine domain; (2) a large portion of CPNA-1 including the copine domain, but lacking the transmembrane domain; and (3) the N-terminal portion of CPNA-1 that includes essentially the transmembrane domain. Our very talented super-postdoc., Hiroshi Qadota, microinjected these constructs into worms together with a transformation marker and established multiple transgenic lines. Unfortunately, by western blot, I could not obtain evidence that

any of them were expressed. The constructs were quite complicated, and it is possible that I made some errors. In the future, this line of experimentation could be repeated.

As I showed in Chapter 3, CPNA-1 is localized to both M-lines and dense bodies by using specific antibodies I developed. It has been confirmed by biochemical methods that CPNA-1 interacts with the following M-line components: UNC-89, SCPL-1, LIM-9, UNC-96 and PAT-6. In fact, all the data taken together indicate that CPNA-1 is involved in a signaling pathway of sarcomere assembly that occurs sequentially in this order: UNC-52 (perlecan) in the basement membrane, PAT-3 (β-integrin) in the muscle cell membrane, PAT-6 and then CPNA-1 near the cytoplasmic side of the muscle cell membrane. All of these just mentioned-proteins are found at the bases of both M-lines and dense bodies. My studies explain which proteins CPNA-1 recruits at M-lines (UNC-89, SCPL-1, LIM-9 and UNC-96). An unresolved and important issue is which proteins CPNA-1 recruits to the dense bodies. One way to address this question in the future is to use CPNA-1 to screen the yeast 2-hybrid library, and ultimately identify proteins that interact with CPNA-1 and are only localized to dense bodies. Another way would be to overexpress a tagged CPNA-1 in worms, immunoprecipitate with antibodies to the tag, and identify proteins in the complex by mass spectrometry.

The role of CPNA-1 in embryonic muscle development is clear from the loss of function Pat embryonic lethal phenotype of *cpna-1 (gk266)*, and examination of CPNA-1 staining patterns in other Pat gene mutants, and the staining patterns of Pat gene products in *cpna-1(gk266)* mutant embryos. Although the role of CPNA-1 in adult sarcomere assembly/maintenance has been verified by showing that RNAi by L1 feeding of *pat-6* results in loss of CPNA-1 localization, and that deficiency of UNC-89, LIM-9, SCPL-1

and UNC-96 has no effect on the localization of CPNA-1, one experiment is missing. That is, what happens to the localization or assembly of PAT-6, and UNC-89, LIM-9, SCPL-1 and UNC-96, when CPNA-1 is missing in adult muscle? We tried to conduct RNAi of *cpna-1* by L1 feeding, but there was no obvious effect. Indeed, RNAi of *cpna-1* by the standard procedure, does not reliably yield Pat embryos. An alternative method, that could be pursued in the future, is to make transgenic lines in which the embryonic lethality of *cpna-1(gk266)* is rescued by a wild type copy of *cpna-1*, and then, in adults, exploit the infrequent random loss of extrachromosomal arrays to generate mosaic adults in which some muscle cells do not express CPNA-1. Such cells would be examined by immunofluorescence microscopy for the localization of PAT-6 (expected not to be affected), and UNC-89, LIM-9, SCPL-1 and UNC-96 (some of which might be affected). A similar experiment has been done successfully by the Benian lab to demonstrate the necessity of MHC A myosin for the localization of UNC-98 (Miller et al. 2006). Another way to address this question would be to perform a F1 non-complementation screen using the *gk266* allele and looking for F1 adults that are slow moving and/or have disorganized sarcomeres. In other words, this might identify a *cpna-1* allele that is "Unc". As mentioned in the first chapter, there are several nematode genes that are Pat embryonic lethal as null alleles and adult viable Uncs as hypomorphs (e.g. *unc-97*, *unc-112*, *unc-45*, *unc-52*).

Although UNC-89 is composed of 53 Ig domains, I have demonstrated that only Ig domains 1-5 binds to CPNA-1 through yeast two-hybrid assay. Moreover, there are 6 other copine family member genes in *C. elegans* that encode proteins with copine domains that are 28-55% identical in sequence with the copine domain of CPNA-1.

However, by the yeast 2-hybrid assay, all of the other copine domains failed to interact with UNC-89 Ig1-5. Therefore, the interaction between the copine domain of CPNA-1 and UNC-89 Ig domains 1-5 is quite specific. In order to determine which amino acid residues of the copine domain of CPNA-1 determines the specificity of the interaction between CPNA-1 and UNC-89, I generated point mutations of 3 highly conserved residues in the copine domain of CPNA-1, and tested these mutants for interaction with UNC-89 Ig1-5 and with UNC-96. Results show that UNC-89 and UNC-96 bind to different sites on copine domain of CPNA-1. Since CPNA-1 interacts with multiple proteins, it will be interesting to determine in the future whether an individual CPNA-1 molecule can interact with only one or a few other proteins simultaneously. A much larger segment of CPNA-1, including all but the transmembrane domain is required for interaction with PAT-6, LIM-9 and SCPL-1; in these cases, we could more easily envision distinct binding sites for each protein. The possibility that both UNC-89 and UNC-96 could interact simultaneously with the copine domain of CPNA-1 could be addressed by performing a yeast 3-hybrid assay. Finally, I have shown that the minimal region of UNC-89 required for binding to CPNA-1 is Ig1-3, and Kirsty Wilson has shown that the minimal region for binding to MEL-26 is Ig2-3. Again, are these binding sites shared or unique? This is a question for future investigation. One way to address this question is to use error prone PCR to mutagenize the Ig1-3 region and screen by yeast 2-hybrid assay for mutants that can bind to UNC-89 but not MEL-26, and vice versa, in order to identify residues or regions that confer binding specificity.

In whole nematode immunofluorescence, CPNA-1 could also be detected in tissues beyond the body wall muscle. Preliminary data from Hiroshi Qadota shows that in *cpna*-

1(gk266) mutant embryos, the organization of pharyngeal epithelial cells and intestinal epithelial cells is disrupted. Therefore, CPNA-1 may be expressed broadly in *C. elegans* and it may have functions beyond muscles. Perhaps yeast 2-hybrid screens using CPNA-1 as bait will reveal partners that function in these other cell types.

Sequence analysis shows that the similarity between the copine domain of CPNA-1 and human copine family members is ~30% and our preliminary data indicates that currently identified human copine family members have alternatively spliced isoforms that contain a copine domain without C2 domains. Thus, it is possible that there is a true human homologue of CPNA-1 and that it functions in muscle similarly to CPNA-1 in nematode muscle. For example, a putative human CPNA-1 may interact with human UNC-89 (obscurin).

4.2 A LIM-9/SCPL-1 Complex Interacts with the C-terminal Portion of UNC-89

Previously, the Benian lab reported that each of the protein kinase domains of UNC-89 interact with SCPL-1, a CTD-type protein phosphatase (Qadota et al. 2008a). My studies revealed that SCPL-1 also interacts with LIM-9 (FHL), a protein that the Benian lab first discovered as an interactor of UNC-97 (PINCH) and with UNC-96, components of an M-line costamere in nematode muscle. My studies further demonstrated that LIM-9 interacts with UNC-89 through its first kinase domain (PK1) and a portion of unique sequence lying between the two kinase domains. I proposed at least 2 models, not necessarily mutually exclusive, for the function of a LIM-9/SCPL-1/UNC-89 complex: (1) two anti-parallel UNC-89 molecules are linked at their C-termini through interactions with LIM-9 and SCPL-1; (2) a single UNC-89 polypeptide forms a loop with its possibly

flexible interkinase reigon that is stabilized through interactions with LIM-9 and SCPL-1. Evidence that the UNC-89/SCPL-1 and UNC-89/LIM-9 interactions occur in vivo include: (1) UNC-89 and SCPL-1 (Qadota et al. 2008a), and UNC-89 and LIM-9 (chapter 2; Xiong et al. 2009) co-localize at M-lines using antibody staining. Single molecule AFM studies suggest that the interkinase region of UNC-89 is highly elastic (T. Garcia, D. Greene, G. Benian and A. Oberhauser, unpub. data). (2) In body wall muscle cells in which SCPL-1 is transiently over-expressed, there are accumulations of SCPL-1, and near these areas of the myofilament lattice, the M-line organization of UNC-89, as detected by anti-interkinase antibodies is broken, diffuse or missing (chapter 2; Xiong et al. 2009). This is consistent with my model that SCPL-1 acts as a bridge between neighboring UNC-89 polypeptides at their C-termini.

Given my models, I would have expected that loss of function of either *lim-9* or *scpl-1*, or both, would result in a phenotype similar to *unc-89* mutant alleles that lack the kinase domain containing isoforms (alleles *st79* and *tm752*; Ferrara et al. 2005). I have found (data not shown) that *st79* and *tm752* show a characteristic defect in myosin organization, that is distinct from the pattern of myosin disorganization seen when all the giant isoforms of UNC-89 are missing (e.g. in mutant allele *su75*). However, I have not found an obvious defect in myosin organization, or organization of UNC-89 in either single knockout alleles of *lim-9* or *scpl-1*, or in *lim-9(gk106); scpl-1(RNAi)*, or *lim-9; scpl-1* double RNAi. Because the phenotype of *lim-9; scpl-1* is not like *st79* or *tm752*, I hypothesize that yet an additional protein or proteins link the kinase domain containing portions of UNC-89. Screening of the 2-hybrid library with the interkinase region of UNC-89 revealed no new interactors. Perhaps the 2-hybrid approach for this region of

UNC-89 has reached its limit. An alternative approach that might be pursued in the future would be to overexpress HA tagged-kinase and/or interkinase regions in transgenic worms, and use anti-HA IP to pull out interacting proteins and identify them by mass spectrometry. Another approach would be to subject *lim-9; scpl-1* double mutants to RNAi for the ~3500 muscle-expressed genes (Meissner et al. 2009), and screen for animals that show a myosin disorganization similar to *tm752*.

It should also be mentioned that more recently, muscle phenotypes have been described for both *lim-9* and *scpl-1* loss of function mutants. As part of their large-scale RNAi screen for muscle-expressed genes that result in disorganization of GFP tagged myosin A in adults, Meissner et al., identified *lim-9*. Using their multi-generational RNAi method (different from the method I used), they reported a mild disorganization of GFP::myosin A, at a low frequency. In addition, in Qadota et al. (2008a), the Benian lab reported that loss of function of *scpl-1* resulted in a mild egg laying muscle phenotype (Egl), but no obvious defect in sarcomere organization or function. However, using a more quantitative assay of motility developed by Hang Lu's lab at Georgia Tech., an undergraduate in our lab, John Nahabedian, has demonstrated that loss of function of *scpl-1* resulted in increased ability of the worm to maximally bend. Using the same assay, *unc-89* displays reduced ability to bend maximally. Thus, in the future, we could use the Meissner RNAi method to knockdown *lim-9* in an *scpl-1* mutant background and look for a sarcomere disorganization similar to *tm752*, and/or explore defects in maximal bending (including looking for epistasis with *unc-89* mutants).

SCPL-1 is a member of the CTD phosphatase family, members of which were previously implicated in transcriptional regulation. However, SCPL-1 is not localized to

the nucleus. Until our lab's report on SCPL-1 (Qadota et al. 2008a), the CTD phosphatases were known to be involved in the regulation of transcription either through dephosphorylation of the C-terminal domain (CTD) of the largest subunit of RNA polymerase II (as is the case for FCPs, which have both a phosphatase and a BRCT domain), or in dephosphorylating Smad transcription factors (in the case of SCPs, "small CTD phosphatases", that have only the phosphatase domain, like SCPL-1). Our lab's results implicate an absolutely new function for this class of protein phosphatases: in sarcomeres, they are somehow involved in giant kinase function or signaling. Given the fact that humans also have SCP proteins (SCP1-3), and the level of homology between these proteins and *C. elegans* SCPL-1 within the phosphatase domains is high (63-66% identity), it is possible that a human SCP may also interact with human obscurin. The lab has expressed the phosphatase domain of SCPL-1b as an MBP fusion protein and shown that it has phosphatase activity in vitro towards a model substrate, and its enzymatic properties are similar to previously characterized FCP and SCPs (in terms of pH~5.0 optimum and sensitivity to characteristic inhibitors)(Qadota et al. 2008a). And yet this phosphatase domain alone is also sufficient for interaction with UNC-89 PK1 and PK2 regions. Therefore, it would be interesting to determine if these two functions of the SCPL-1 phosphatase domain can be separated. For example, perhaps mutants could be obtained in which the phosphatase activity is destroyed and the binding activity to UNC-89 remains, and vice versa.

5. References

Ackermann, M. A., L. Y. Hu, et al. (2009). "Obscurin interacts with a novel isoform of MyBP-C slow at the periphery of the sarcomeric M-band and regulates thick filament assembly." Mol Biol Cell **20**(12): 2963-2978.

Alberts, B., Johnson, A., Lewis, et al. (2002). "Cell biochemistry and biosynthesis". In Molecular Biology of The Cell. New York, Garland Science.

Altun, Z.F. and Hall, D.H. 2009. "Muscle system, introduction". In WormAtlas. doi:10.3908/wormatlas.1.6 (http:// www.wormatlas.org).

Bagnato, P., V. Barone, et al. (2003). "Binding of an ankyrin-1 isoform to obscurin suggests a molecular link between the sarcoplasmic reticulum and myofibrils in striated muscles." J Cell Biol **160**(2): 245-253.

Bang, M. L., T. Centner, et al. (2001). "The complete gene sequence of titin, expression of an unusual approximately 700-kDa titin isoform, and its interaction with obscurin identify a novel Z-line to I-band linking system." Circ Res **89**(11): 1065-1072.

Barral, J. M., C. C. Bauer, et al. (1998). "Unc-45 mutations in *Caenorhabditis elegans* implicate a CRO1/She4p-like domain in myosin assembly." J Cell Biol **143**(5): 1215-1225.

Barral, J. M., A. H. Hutagalung, et al. (2002). "Role of the myosin assembly protein UNC-45 as a molecular chaperone for myosin." Science **295**(5555): 669-671.

Benian, G. M., A. Ayme-Southgate, et al. (1999). "The genetics and molecular biology of the titin/connectin-like proteins of invertebrates." Rev Physiol Biochem Pharmacol **138**: 235-268.

Benian, G. M., J. E. Kiff, et al. (1989). "Sequence of an unusually large protein implicated in regulation of myosin activity in *C. elegans*." Nature **342**(6245): 45-50.

Benian, G. M., S. W. L'Hernault, et al. (1993). "Additional sequence complexity in the muscle gene, unc-22, and its encoded protein, twitchin, of *Caenorhabditis elegans*." Genetics **134**(4): 1097-1104.

Benian, G. M., T. L. Tinley, et al. (1996). "The *Caenorhabditis elegans* gene unc-89, required fpr muscle M-line assembly, encodes a giant modular protein composed of Ig and signal transduction domains." J Cell Biol **132**(5): 835-848.

Bonnemann, C. G. and N. G. Laing (2004). "Myopathies resulting from mutations in sarcomeric proteins." Curr Opin Neurol **17**(5): 529-537.

Bowman, A. L., D. H. Catino, et al. (2008). "The rho-guanine nucleotide exchange factor domain of obscurin regulates assembly of titin at the Z-disk through interactions with Ran binding protein 9." Mol Biol Cell **19**(9): 3782-3792.

Bowman, A. L., A. Kontrogianni-Konstantopoulos, et al. (2007). "Different obscurin isoforms localize to distinct sites at sarcomeres." FEBS Lett **581**(8): 1549-1554.

Brenner, S. (1974). "The genetics of *Caenorhabditis elegans*." Genetics **77**: 71–94.

Butler, T. M., S. U. Mooers, et al. (2010). "The N-terminal region of twitchin binds thick and thin contractile filaments: redundant mechanisms of catch force maintenance." J Biol Chem **285**(52): 40654-40665.

Callis, T. E., B. C. Jensen, et al. (2010). "Evolving molecular diagnostics for familial cardiomyopathies: at the heart of it all." Expert Rev Mol Diagn **10**(3): 329-351.

Church, D. L. and E. J. Lambie (2003). "The promotion of gonadal cell divisions by the *Caenorhabditis elegans* TRPM cation channel GON-2 is antagonized by GEM-4 copine." Genetics **165**(2): 563-574.

Consortium., C. e. S. (1998). "Genome sequence of the nematode *C. elegans*: a platform for investigating biology." Science **282**(5396): 2012-2018.

Cooper, G. M. (2000). "The Cytoskeleton and Cell Movement. " In The Cell: A Molecular Approach, ed. G. M. Cooper. Sunderland (MA), Sinauer Associates.

Cowling, B. S., M. J. McGrath, et al. (2008). "Identification of FHL1 as a regulator of skeletal muscle mass: implications for human myopathy." J Cell Biol **183**(6): 1033-1048.

Creutz, C. E., J. L. Tomsig, et al. (1998). "The copines, a novel class of C2 domain-containing, calcium-dependent, phospholipid-binding proteins conserved from Paramecium to humans." J Biol Chem **273**(3): 1393-1402.

Damer, C. K., M. Bayeva, et al. (2007). "Copine A is required for cytokinesis, contractile vacuole function, and development in Dictyostelium." Eukaryot Cell **6**(3): 430-442.

Ferrara, T. M., Flaherty, D. B. & Benian, G. M. (2005). "Titin/connectin-related proteins in *C. elegans*: a review and new findings." J. Muscle Res. Cell Motil. **26**(6-8): 435-447.

Flaherty, D. B., K. M. Gernert, et al. (2002). "Titins in *C.elegans* with unusual features: coiled-coil domains, novel regulation of kinase activity and two new possible elastic regions." J Mol Biol **323**(3): 533-549.

Forbes, J. G., D. B. Flaherty, et al. (2010). "Extensive and modular intrinsically disordered segments in *C. elegans* TTN-1 and implications in filament binding, elasticity and oblique striation." J Mol Biol **398**(5): 672-689.

Ford-Speelman, D. L., J. A. Roche, et al. (2009). "The rho-guanine nucleotide exchange factor domain of obscurin activates rhoA signaling in skeletal muscle." Mol Biol Cell **20**(17): 3905-3917.

Francis, R. and R. H. Waterston (1991). "Muscle cell attachment in *Caenorhabditis elegans*." J Cell Biol **114**(3): 465-479.

Fukuzawa, A., S. Idowu, et al. (2005). "Complete human gene structure of obscurin: implications for isoform generation by differential splicing." J. Muscle Res. Cell Motil. **26**(6-8): 427-434.

Fukuzawa, A., S. Idowu, et al. (2005). "Complete human gene structure of obscurin: implications for isoform generation by differential splicing." J Muscle Res Cell Motil **26**(6-8): 427-434.

Fukuzawa, A., S. Lange, et al. (2008). "Interactions with titin and myomesin target obscurin and obscurin-like 1 to the M-band: implications for hereditary myopathies." J Cell Sci **121**(Pt 11): 1841-1851.

Funabara, D., C. Hamamoto, et al. (2007). "Unphosphorylated twitchin forms a complex with actin and myosin that may contribute to tension maintenance in catch." J Exp Biol **210**(Pt 24): 4399-4410.

Funabara, D., S. Watabe, et al. (2003). "Twitchin from molluscan catch muscle: primary structure and relationship between site-specific phosphorylation and mechanical function." J Biol Chem **278**(31): 29308-29316.

Furukawa, T., Y. Ono, et al. (2001). "Specific interaction of the potassium channel beta-subunit minK with the sarcomeric protein T-cap suggests a T-tubule-myofibril linking system." J Mol Biol **313**(4): 775-784.

Gottschalk, A., R. B. Almedom, et al. (2005). "Identification and characterization of novel nicotinic receptor-associated proteins in *Caenorhabditis elegans*." EMBO J **24**(14): 2566-2578.

Granzier, H. L. and S. Labeit (2004). "The giant protein titin: a major player in myocardial mechanics, signaling, and disease." Circ Res **94**(3): 284-295.

Grater, F., J. Shen, et al. (2005). "Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations." Biophys J **88**(2): 790-804.

Greene, D. N., T. Garcia, et al. (2008). "Single-molecule force spectroscopy reveals a stepwise unfolding of Caenorhabditis elegans giant protein kinase domains." Biophys J **95**(3): 1360-1370.

Hannak, E., K. Oegema, et al. (2002). "The kinetically dominant assembly pathway for centrosomal asters in *Caenorhabditis elegans* is gamma-tubulin dependent." J Cell Biol **157**(4): 591-602.

Heinrich, C., C. Keller, et al. (2010). "Copine-III interacts with ErbB2 and promotes tumor cell migration." Oncogene **29**(11): 1598-1610.

Hu, S. H., J. Y. Lei, et al. (1994). "Crystallization and preliminary X-ray analysis of the auto-inhibited twitchin kinase." J Mol Biol **236**(4): 1259-1261.

Hua, J., P. Grisafi, et al. (2001). "Plant growth homeostasis is controlled by the Arabidopsis BON1 and BAP1 genes." Genes Dev **15**(17): 2263-2272.

Huxley, H. and J. Hanson (1954). "Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation." Nature **173**(4412): 973-976.

Jambunathan, N., J. M. Siani, et al. (2001). "A humidity-sensitive Arabidopsis copine mutant exhibits precocious cell death and increased disease resistance." Plant Cell **13**(10): 2225-2240.

Kobe, B., J. Heierhorst, et al. (1996). "Giant protein kinases: domain interactions and structural basis of autoregulation." EMBO J **15**(24): 6810-6821.

Kontrogianni-Konstantopoulos, A., M. A. Ackermann, et al. (2009). "Muscle giants: molecular scaffolds in sarcomerogenesis." Physiol Rev **89**(4): 1217-1267.

Kontrogianni-Konstantopoulos, A., D. H. Catino, et al. (2004). "Obscurin regulates the organization of myosin into A bands." Am J Physiol Cell Physiol **287**(1): C209-217.

Kontrogianni-Konstantopoulos, A., D. H. Catino, et al. (2006). "Obscurin modulates the assembly and organization of sarcomeres and the sarcoplasmic reticulum." FASEB J **20**(12): 2102-2111.

Kontrogianni-Konstantopoulos, A., E. M. Jones, et al. (2003). "Obscurin is a ligand for small ankyrin 1 in skeletal muscle." Mol Biol Cell **14**(3): 1138-1148.

Kruger, M. and W. A. Linke (2009). "Titin-based mechanical signalling in normal and failing myocardium." J Mol Cell Cardiol **46**(4): 490-498.

Labeit, S. and B. Kolmerer (1995). "Titins: giant proteins in charge of muscle ultrastructure and elasticity." Science **270**(5234): 293-296.

Lange, S., D. Auerbach, et al. (2002). "Subcellular targeting of metabolic enzymes to titin in heart muscle may be mediated by DRAL/FHL-2." J Cell Sci **115**(Pt 24): 4925-4936.

Lange, S., E. Ehler, et al. (2006). "From A to Z and back? Multicompartment proteins in the sarcomere." Trends Cell Biol **16**(1): 11-18.

Lange, S., K. Ouyang, et al. (2009). "Obscurin determines the architecture of the longitudinal sarcoplasmic reticulum." J Cell Sci **122**(Pt 15): 2640-2650.

Lange, S., F. Xiang, et al. (2005). "The kinase domain of titin controls muscle gene expression and protein turnover." Science **308**(5728): 1599-1603.

Lei, J., X. Tang, et al. (1994). "Protein kinase domain of twitchin has protein kinase activity and an autoinhibitory region." J Biol Chem **269**(33): 21078-21085.

Li, Y., M. Gou, et al. (2010). "Requirement of calcium binding, myristoylation, and protein-protein interaction for the Copine BON1 function in Arabidopsis." J Biol Chem **285**(39): 29884-29891.

Lin, X., H. Qadota, et al. (2003). "*C. elegans* PAT-6/actopaxin plays a critical role in the assembly of integrin adhesion complexes in vivo." Curr Biol **13**(11): 922-932.

Linke, W. A. (2008). "Sense and stretchability: the role of titin and titin-associated proteins in myocardial stress-sensing and mechanical dysfunction." Cardiovasc Res **77**(4): 637-648.

Liu, J. and B. Rost (2003). "NORSp: Predictions of long regions without regular secondary structure." Nucleic Acids Res **31**(13): 3833-3835.

Liu, J., H. Tan, et al. (2002). "Loopy proteins appear conserved in evolution." J Mol Biol **322**(1): 53-64.

Luther, P., Squire, J.(1978). "Three-dimensional structure of the vertebrate muscle M-region." J. Molecular Biology **125** (5): 314-324.

Ma, K., J. G. Forbes, et al. (2006). "Titin as a giant scaffold for integrating stress and Src homology domain 3-mediated signaling pathways: the clustering of novel overlap ligand motifs in the elastic PEVK segment." J Biol Chem **281**(37): 27539-27556.

Mackinnon, A. C., H. Qadota, et al. (2002). "*C. elegans* PAT-4/ILK functions as an adaptor protein within integrin adhesion complexes." Curr Biol **12**(10): 787-797.

Mayans, O., P. F. van der Ven, et al. (1998). "Structural basis for activation of the titin kinase domain during myofibrillogenesis." Nature **395**(6705): 863-869.

McLachlan, A. D. and J. Karn (1982). "Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle." Nature **299**(5880): 226-231.

Meissner, B., A. Warner, et al. (2009). "An integrated strategy to study muscle development and myofilament structure in *Caenorhabditis elegans*." PLoS Genet **5**(6): e1000537.

Mercer, K. B., D. B. Flaherty, et al. (2003). "*Caenorhabditis elegans* UNC-98, a C2H2 Zn finger protein, is a novel partner of UNC-97/PINCH in muscle adhesion complexes." Mol Biol Cell **14**(6): 2492-2507.

Mercer, K. B., R. K. Miller, et al. (2006). "*Caenorhabditis elegans* UNC-96 is a new component of M-lines that interacts with UNC-98 and paramyosin and is required in adult muscle for assembly and/or maintenance of thick filaments." Mol Biol Cell **17**(9): 3832-3847.

Miller, D. M., 3rd, I. Ortiz, et al. (1983). "Differential localization of two myosins within nematode thick filaments." Cell **34**(2): 477-490.

Miller, R. K., H. Qadota, et al. (2006). "UNC-98 links an integrin-associated complex to thick filaments in *Caenorhabditis elegans* muscle." J Cell Biol **175**(6): 853-859.

Moerman, D. G., G. M. Benian, et al. (1988). "Identification and intracellular localization of the unc-22 gene product of *Caenorhabditis elegans*." Genes Dev **2**(1): 93-105.

Moerman, D. G. and A. Fire (1997). "Muscle Structure, Function and Development". In *C. elegans* II, New York, Cold Spring Harbor Laboratory Press.

Moerman, D. G. and B. D. Williams (2006). "Sarcomere assembly in *C. elegans* muscle." In WormBook , ed. The *C. elegans* Research Community (http://www.wormbook.org).

Mues, A., P. F. van der Ven, et al. (1998). "Two immunoglobulin-like domains of the Z-disc portion of titin interact in a conformation-dependent way with telethonin." FEBS Lett **428**(1-2): 111-114.

Nicholas, G., M. Thomas, et al. (2002). "Titin-cap associates with, and regulates secretion of, Myostatin." J Cell Physiol **193**(1): 120-131.

Nonet, M. L., K. Grundahl, et al. (1993). "Synaptic function is impaired but not eliminated in *C. elegans* mutants lacking synaptotagmin." Cell **73**(7): 1291-1305.

Norman, K. R., S. Cordes, et al. (2007). "UNC-97/PINCH is involved in the assembly of integrin cell adhesion complexes in *Caenorhabditis elegans* body wall muscle." Dev Biol **309**(1): 45-55.

Ono, K., R. Yu, et al. (2006). "*Caenorhabditis elegans* kettin, a large immunoglobulin-like repeat protein, binds to filamentous actin and provides mechanical stability to the contractile apparatuses in body wall muscle." Mol Biol Cell **17**(6): 2722-2734.

Phillips, G. N., Jr., J. P. Fillers, et al. (1986). "Tropomyosin crystal structure and muscle regulation." J Mol Biol **192**(1): 111-131.

Probst, W. C., E. C. Cropper, et al. (1994). "cAMP-dependent phosphorylation of Aplysia twitchin may mediate modulation of muscle contractions by neuropeptide cotransmitters." Proc Natl Acad Sci U S A **91**(18): 8487-8491.

Puchner, E. M., A. Alexandrovich, et al. (2008). "Mechanoenzymatics of titin kinase." Proc Natl Acad Sci U S A **105**(36): 13385-13390.

Qadota, H. and G. M. Benian (2010). "Molecular structure of sarcomere-to-membrane attachment at M-Lines in *C. elegans* muscle." J Biomed Biotechnol **2010**: 864749.

Qadota, H., A. Blangy, et al. (2008b). "The DH-PH region of the giant protein UNC-89 activates RHO-1 GTPase in *Caenorhabditis elegans* body wall muscle." J Mol Biol **383**(4): 747-752.

Qadota, H., L. A. McGaha, et al. (2008a). "A novel protein phosphatase is a binding partner for the protein kinase domains of UNC-89 (Obscurin) in *Caenorhabditis elegans*." Mol Biol Cell **19**(6): 2424-2432.

Qadota, H., K. B. Mercer, et al. (2007). "Two LIM domain proteins and UNC-96 link UNC-97/pinch to myosin thick filaments in *Caenorhabditis elegans* muscle." Mol. Biol. Cell **18**(11): 4317-4326.

Quinzii, C. M., T. H. Vu, et al. (2008). "X-linked dominant scapuloperoneal myopathy is due to a mutation in the gene encoding four-and-a-half-LIM protein 1." Am J Hum Genet **82**(1): 208-213.

Rogalski, T. M., G. P. Mullen, et al. (2000). "The UNC-112 gene in *Caenorhabditis elegans* encodes a novel component of cell-matrix adhesion structures required for integrin localization in the muscle cell membrane." J Cell Biol **150**(1): 253-264.

Samson, T., N. Smyth, et al. (2004). "The LIM-only proteins FHL2 and FHL3 interact with alpha- and beta-subunits of the muscle alpha7beta1 integrin receptor." J Biol Chem **279**(27): 28641-28652.

Schallus, T., K. Feher, et al. (2009). "Structure and dynamics of the human muscle LIM protein." FEBS Lett **583**(6): 1017-1022.

Schessl, J., Y. Zou, et al. (2008). "Proteomic identification of FHL1 as the protein mutated in human reducing body myopathy." J Clin Invest **118**(3): 904-912.

Scholl, F. A., P. McLoughlin, et al. (2000). "DRAL is a p53-responsive gene whose four and a half LIM domain protein product induces apoptosis." J Cell Biol **151**(3): 495-506.

Siegman, M. J., D. Funabara, et al. (1998). "Phosphorylation of a twitchin-related protein controls catch and calcium sensitivity of force production in invertebrate smooth muscle." Proc Natl Acad Sci U S A **95**(9): 5383-5388.

Simmer, F., M. Tijsterman, et al. (2002). "Loss of the putative RNA-directed RNA polymerase RRF-3 makes *C. elegans* hypersensitive to RNAi." Curr Biol **12**(15): 1317-1319.

Small, T. M., K. M. Gernert, et al. (2004). "Three new isoforms of *Caenorhabditis elegans* UNC-89 containing MLCK-like protein kinase domains." J Mol Biol **342**(1): 91-108.

Southgate, R. and A. Ayme-Southgate (2001). "Alternative splicing of an amino-terminal PEVK-like region generates multiple isoforms of *Drosophila* projectin." J Mol Biol **313**(5): 1035-1043.

Timmons, L., D. L. Court, et al. (2001). "Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*." Gene **263**(1-2): 103-112.

Tomsig, J. L. and C. E. Creutz (2002). "Copines: a ubiquitous family of Ca(2+)-dependent phospholipid-binding proteins." Cell Mol Life Sci **59**(9): 1467-1477.

Tskhovrebova, L. T., J. (2003). "Titin: properties and family relationships." Nat. Rev. Mol. Cell Biol. **4**(9): 679-689.

Udd, B. (2008). "Third filament diseases." Adv Exp Med Biol **642**: 99-115.

Vikstrom, K. L., S. H. Seiler, et al. (1997). "The vertebrate myosin heavy chain: genetics and assembly properties." Cell Struct Funct **22**(1): 123-129.

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

Waterston, R. H., Ed. (1988). "Muscle" in The Nematode *Caenorhabditis elegans*. New York, Cold Spring Harbor Laboratory Press.

Waterston, R. H., J. N. Thomson, et al. (1980). "Mutants with altered muscle structure of *Caenorhabditis elegans*." Dev Biol **77**(2): 271-302.

Williams, B. D. and R. H. Waterston (1994). "Genes critical for muscle development and function in *Caenorhabditis elegans* identified through lethal mutations." J Cell Biol **124**(4): 475-490.

Windpassinger, C., B. Schoser, et al. (2008). "An X-linked myopathy with postural muscle atrophy and generalized hypertrophy, termed XMPMA, is caused by mutations in FHL1." Am J Hum Genet **82**(1): 88-99.

Wixler, V., D. Geerts, et al. (2000). "The LIM-only protein DRAL/FHL2 binds to the cytoplasmic domain of several alpha and beta integrin chains and is recruited to adhesion complexes." J Biol Chem **275**(43): 33669-33678.

Wootton, J. C. and S. Federhen (1996). "Analysis of compositionally biased regions in sequence databases." Methods Enzymol **266**: 554-571.

Wootton, J. C. and S. Federhen (1996). "Analysis of compositionally biased regions in sequence databases." Methods Enzymol. **266**: 554-571.

Xiong, G., H. Qadota, et al. (2009). "A LIM-9 (FHL)/SCPL-1 (SCP) complex interacts with the C-terminal protein kinase regions of UNC-89 (obscurin) in *Caenorhabditis elegans* muscle." J Mol Biol **386**(4): 976-988.

Yang, H., Y. Li, et al. (2006). "The C2 domain protein BAP1 negatively regulates defense responses in Arabidopsis." Plant J **48**(2): 238-248.

Yochem, J., T. Gu, et al. (1998). "A new marker for mosaic analysis in *Caenorhabditis elegans* indicates a fusion between hyp6 and hyp7, two major components of the hypodermis." Genetics **149**(3): 1323-1334.

Young, P., E. Ehler, et al. (2001). "Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly." J Cell Biol **154**(1): 123-136.

Zarkower, D. (2006). "Somatic sex determination." In WormBook, ed. The *C. elegans* Research Community (http://www.wormbook.org).