Targeted gene replacement to mimic familial hypertrophic cardiomyopathy in Caenorhabditis elegans

Craig Alan Almeida

University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

Recommended Citation

https://scholars.unh.edu/dissertation/1883

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
Targeted gene replacement to mimic familial hypertrophic cardiomyopathy in Caenorhabditis elegans

Abstract
The identification in invertebrates of genes homologous to human disease genes provides the possibility to broaden the spectrum of model organisms to include experimental invertebrates. The goal of this study is to test the possibility of using Caenorhabditis elegans as an animal model for the inherited human heart disorder familial hypertrophic cardiomyopathy. In 10-30% of the affected families that have been studied, this disease is caused by mutations in the $\beta$-cardiac myosin heavy chain (MHC) gene. The most common of these mutations results in the replacement of an evolutionarily conserved arginine residue (R403) with glutamine (R403Q). While this information has been useful for diagnostic purposes, it has revealed little about the molecular basis of FHC. How do these mutations affect cardiac muscle structure and contractility and result in cardiac hypertrophy? Using a novel gene replacement approach, I introduced the corresponding mutation (R404Q) into the unc-54 gene, which encodes the predominant myosin heavy chain in the body-wall muscle of C. elegans. The phenotype of the unc-54(R404Q) mutant is essentially wild-type. Detailed characterization of the R404Q mutant has yielded information which highlights differences between nematode and vertebrate striated muscle which may be key factors in the pathogenesis of the disease.

Keywords
Biology, Molecular, Biology, Genetics, Chemistry, Biochemistry

This dissertation is available at University of New Hampshire Scholars' Repository: https://scholars.unh.edu/dissertation/1883
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
TARGETED GENE REPLACEMENT TO MIMIC FAMILIAL HYPERTROPHIC CARDIOMYOPATHY IN Caenorhabditis elegans

by

Craig A. Almeida
B.S., Bridgewater State College, 1991

DISCUSSION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Biochemistry
May, 1996
This dissertation has been examined and approved.

Dissertation Director, John J. Collins
Associate Professor of Biochemistry and
Molecular Biology and Graduate Program in Genetics

Anita S. Klein
Associate Professor of Biochemistry and
Molecular Biology and Chair of Graduate Program in Genetics

Andrew P. Laudano
Associate Professor of Biochemistry and Molecular Biology

Samual C. Smith
Professor of Animal and Nutritional Sciences

Thomas M. Davis
Associate Professor of Plant Biology and Graduate Program in Genetics

5/1/96
Date
DEDICATION

I would like to dedicate this work to my parents, Joseph and Carol Ann Almeida.

Mom and Dad,
The words don't exist to possibly express the pride I feel at being your son. You have been supportive of everything I have done and I am secure in knowing that you will always have faith in me and what I do. Thank you for being there every time I have turned and needed you. I love you both so very much.

Craig
ACKNOWLEDGEMENTS

I have learned a great deal over the last five years and much of that is the result of the many discussions with John. I am very appreciative of his time, patience, guidance and encouragement. However..."To know John is to wait for John." Enough said! I want to thank all of those who have passed through the lab during my tenure at UNH. There are those who I would specifically like to recognize...Mindy, for being there all along the way, being a sparring partner and her advice and witiness; Christi for her companionship, support, lunchtime walks to the bookstore and e-mail messages; Michelle for her "sitcom life" and being a fellow commiserator during the first two years of graduate school; Jeremy for his insight and support; Holly for her support, laughter and persistent chipping away at my shell; Kevin for his friendship, sense of humor, and help and support on many occasions - you are one in a million; Queta for her taking charge and her enjoyment of M & Ms; Jen for being the eye in the center of the storm; Joanna for being...well, Joanna, Joanna, Joanna; Stacey for turning me on to tequila; Carrie for her laugh and help in lab; Ken for his help with the polarized light microscopy.

There are others along the way to whom I would like to express my thanks. Andreas for being an original hooahaa; Nicole for tuning me into Stonehill; Michael for his help and suggestions; Anita for her concern and support, Andy for his "Alrights!"; and Sam and Tom for their guidance.

I would like to give special thanks and recognition to Sandy Whelan. She saw in me what I couldn't see in myself. Without her support, encouragement and kind words I wouldn't have come as far as I have nor have accomplished as much. The words "Thank You" are so inadequate for the gratitude I wish to express for the support that she gave me at the beginning of the long road I have traveled the last seven years.
TABLE OF CONTENTS

DEDICATION ........................................................................ iii
ACKNOWLEDGEMENTS ........................................................ iv
LIST OF TABLES ................................................................ vi
LIST OF FIGURES ................................................................ vii
ABSTRACT ........................................................................ viii

SECTION .............................................................................. PAGE

INTRODUCTION .................................................................... 1
MATERIALS AND METHODS .............................................. 23
RESULTS ............................................................................ 38
DISCUSSION ........................................................................ 65
LITERATURE CITED ............................................................. 78
LIST OF TABLES

TABLE 1: Primers used in the representative PCR reactions and the corresponding amplified segments of the *unc-54* gene......41

TABLE 2: Post embryonic development times and brood size determinations for N2 and TW390...............................60
LIST OF FIGURES

FIGURE 1: The amino acid sequence of myosin heavy chains............4
FIGURE 2: Sarcomere structure.............................................6
FIGURE 3: Electron micrographs of an insect flight muscle viewed in
cross section.................................................................7
FIGURE 4: The sliding-filament model of muscle contraction...........8
FIGURE 5: The mechanism of skeletal muscle contraction...............9
FIGURE 6: The myosin molecule.............................................11
FIGURE 7: Dissection of myosin by proteases...........................12
FIGURE 8: Developmental stages of *Caenorhabditis elegans*.........15
FIGURE 9: Vertebrate cross-striated and *C. elegans* obliquely striated
muscle structures.........................................................19
FIGURE 10: *C. elegans* body-wall muscle ultrastructure............21
FIGURE 11: Construction of the unc-54(R404Q) mutation using PCR...40
FIGURE 12: Locations of primers used to generate the
    \textit{pUNK-54(R404Q)} clone \textit{in vitro}............................41
FIGURE 13: PCR products from the 1st round of \textit{in vitro} mutation....42
FIGURE 14: PCR product from the 2nd round of \textit{in vitro} mutation...43
FIGURE 15: Loss of \textit{PmlI} restriction site............................44
FIGURE 16: R404Q mutation results in the loss of a \textit{PmlI} site........46
FIGURE 17: Sequence of the R404 region of \textit{pUNK-54(R404Q)}.......47
FIGURE 18: Position of \textit{BamH}I restriction fragment..................51
FIGURE 19: Locations of R404 and \textit{r360::TcI} insertion sites in
    \textit{unc-54}..............................................................53

vii
FIGURE 20: Loss of PmlI site in transgenics............................55
FIGURE 21: Sequence analysis of the R404 region of unc-54 from
N2 and R3...............................................................56
FIGURE 22: Genomic Southern blot analysis of unc-54::Tcl and
revertants...............................................................58
FIGURE 23: Polarized light microscopy of C. elegans body-wall
muscles...............................................................63
ABSTRACT

TARGETED GENE REPLACEMENT TO MIMIC FAMILIAL HYPERTROPHIC CARDIOMYOPATHY IN CAENORHABDIS ELEGANS

by

Craig A. Almeida
University of New Hampshire, May 1996

The identification in invertebrates of genes homologous to human disease genes provides the possibility to broaden the spectrum of model organisms to include experimental invertebrates. The goal of this study is to test the possibility of using Caenorhabditis elegans as an animal model for the inherited human heart disorder familial hypertrophic cardiomyopathy. In 10-30% of the affected families that have been studied, this disease is caused by mutations in the β-cardiac myosin heavy chain (MHC) gene. The most common of these mutations results in the replacement of an evolutionarily conserved arginine residue (R403) with glutamine (R403Q). While this information has been useful for diagnostic purposes, it has revealed little about the molecular basis of FHC. How do these mutations affect cardiac muscle structure and contractility and result in cardiac hypertrophy? Using a novel gene replacement approach, I introduced the corresponding mutation (R404Q) into the unc-54 gene, which encodes the predominant myosin heavy chain in the body-wall muscle of C. elegans. The phenotype of the unc-54(R404Q) mutant is essentially wild-type. Detailed characterization of the R404Q mutant has yielded information which highlights differences between nematode and vertebrate striated muscle which may be key factors in the pathogenesis of the disease.
INTRODUCTION

One of the major challenges in biomedical research is the understanding of the molecular mechanism underlying human diseases. For various inherited disorders there are a knowledge of the genetic basis. Information regarding an affected gene can provide insight into the molecular basis of the disorder. However, the ability to exploit this information to investigate the molecular details is hampered by the inability to perform experiments directly on human subjects. Often, this obstacle can be circumvented by the use of an "animal model." Frequently, the problem is addressed in another organism similar to humans - mouse, rat, chimpanzee, etc. From the rapid expansion in nucleic acid sequence information from a wide range of organisms it is now known that all organisms share much more at the molecular genetic level than previously imagined. This suggests that it may be possible for certain situations to expand the animal model approach to other organisms, including the popular model organisms such as the fruit fly \textit{(Drosophila melanogaster)}, nematode \textit{(Caenorhabditis elegans)} and zebrafish \textit{(Brachydanio rerio)}.

This dissertation presents efforts to exploit powerful approaches available in \textit{C. elegans} to investigate the molecular basis of an inherited human heart disorder, namely familial hypertrophic cardiomyopathy. This approach is justified because the disease is the result of mutations in the $\beta$-cardiac myosin heavy chain (MHC). Myosin is the basic component of muscle in all organisms. In addition to muscle structure and function being highly conserved, the role of MHC is also well conserved in evolution. The two long-term responses that the human heart undergoes in response to physiological or pathological stress are hypertrophy and dilation. Hypertrophy results from enlargement of existing, terminally differentiated, cardiocytes which ceased proliferating shortly after birth. Elucidation of the molecular basis of hypertrophy would have wide-spread implications. It
would improve diagnosis, allowing for the identification of high risk patients. In addition, it would lead to new and improved treatment protocols. Finally, it would improve our fundamental understanding of hypertrophy in response to inherited diseases.

An example of a disease state that demonstrates hypertrophy of the myocardium is familial hypertrophic cardiomyopathy (FHC). The disease is heterogeneous not only in its clinical symptoms but also in the morphological expression in the heart. The degree and regional distribution of hypertrophy within a single family is highly variable. The reasons for this are unknown at this point. FHC is the most common inherited heart disorder in the human population. It is estimated that the prevalence of the disease within the population is 1 to 10 per 10,000 (Codd, et al., 1989; Bjarnason, et al., 1982). In fact, it is the most common cause of death in the young. In particular, nearly 50% of sudden death in young competitive athletes (i.e. less than 35 years of age) can be attributed to FHC (Marian, 1995; Maron et al., 1986; Maron et al., 1982; McKenna et al., 1981; Maron et al., 1978). The sudden death associated with FHC occurs frequently during or just following exercise, and is often times unexpected because it is the first manifestation of the disorder in asymptomatic, apparently healthy young individuals (Maron et al., 1982; Maron et al., 1986; Maron et al., 1978; McKenna et al., 1981; Burke et al., 1991). Symptoms include: syncope, angina and heart failure (Dausse et al., 1993). Common features of FHC pathology include left ventricular hypertrophy, disorganized myocytes and muscle fiber disarray (Maron and Farrans, 1978; Roberts and Ferrans, 1982; Maron et al., 1984).

FHC is inherited in an autosomal dominant pattern (Fananapazir and Epstein, 1994; Anan et al., 1994). Molecular genetic studies over the past several years have revealed linkage of FHC with regions on human chromosomes 1, 11, 14, 15, 16 and 18 (Marian and Roberts, 1994). The precise locus responsible for FHC has been identified for three of these linkage groups: the β-cardiac myosin heavy chain gene on chromosome 14, α-tropomyosin gene on chromosome 15 and cardiac troponin-T gene on chromosome 1.
(Marian and Roberts, 1994; Thierfelder, 1994). In addition, preliminary studies have identified the genes encoding the α-cardiac myosin heavy chain and cardiac myosin binding protein-C proteins as candidate genes for chromosome 14 and 11-associated FHC, respectively (van den Berg, et al., 1995; Gautel, et al., 1995).

The most commonly affected of these loci is the β-cardiac myosin heavy chain (MYH7) gene on chromosome 14; 10-30% of hypertrophic cardiomyopathy kindreds possess mutations within this gene (Hengstenberg and Schwartz, 1994; McKenna et al., 1995; Rayment, et al., (PNAS) 1995). At least twenty-nine different missense mutations and one deletion in this gene have been reported that cosegregate with FHC (Keating, 1994; Dausse, et al., 1993; Hengstenberg, et al., 1993; Watkins, et al., 1992). All the missense mutations are within exons coding for the globular head or head-rod junction of the β-cardiac myosin heavy chain and affect evolutionarily conserved amino acids (Nishi, et al., 1995). The first and most frequently reported mutation replaces an arginine (R403) with a glutamine (R403Q). Affected members of certain families have also been shown to possess substitutions of tryptophan and leucine for this arginine (Dausse, et al., 1993).

Myosin primary structure is highly conserved through evolution. The R403 residue is invariant among 15 sequenced MHCs from organisms as different as humans and amoeba representing an evolutionary span of over 600 million years (Geisterfer-Lowrance, 1990). In addition, R403 resides in a domain containing several other highly conserved residues, as shown in Figure 1.

The same FHC phenotype results from mutations in α-tropomyosin, cardiac troponin T and β-cardiac MHC. These are integral components of the sarcomere, the basic contractile unit of muscle. Therefore, FHC is believed to be a disease of the sarcomere (Thierfelder et al., 1994). The two major constituents of the contractile apparatus are the proteins myosin and actin. The sarcomere is formed from the interdigitation of myosin thick filaments and actin thin filaments and muscle contraction results from the interaction.
<table>
<thead>
<tr>
<th>Species</th>
<th>Amino Acid Sequence</th>
<th>R403 Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardiac β (human)</td>
<td>EADKAVLMEANSADLLKLQH</td>
<td>R403</td>
</tr>
<tr>
<td>Perinatal (human)</td>
<td>V---A---QS---A---Y</td>
<td>R403</td>
</tr>
<tr>
<td>Embryonic (human)</td>
<td>V---T---S---A---F</td>
<td>R403</td>
</tr>
<tr>
<td>Cardiac β (rat)</td>
<td></td>
<td>R403</td>
</tr>
<tr>
<td>Cardiac α (rat)</td>
<td></td>
<td>R403</td>
</tr>
<tr>
<td>Embryonic (rat)</td>
<td>V---T---S---A---F</td>
<td>R403</td>
</tr>
<tr>
<td>Adult (rabbit)</td>
<td>V---A---QS---A---Y</td>
<td>R403</td>
</tr>
<tr>
<td>Embryonic (chicken)</td>
<td>V---A---A---A---Y</td>
<td>R403</td>
</tr>
<tr>
<td>Smooth (chicken)</td>
<td>A-Q-VCH---I-VI-FIRSILT</td>
<td>R403</td>
</tr>
<tr>
<td>Non-muscle (chicken)</td>
<td>A-Q-VSH-L-I-VI-FTR-ILT</td>
<td>R403</td>
</tr>
<tr>
<td>C. elegans unc54</td>
<td>E-ASNMY-ICCEEP---A-TR</td>
<td>R403</td>
</tr>
<tr>
<td>C. elegans myla</td>
<td>EPACMCY-ILVQF---A-TR</td>
<td>R403</td>
</tr>
<tr>
<td>C. elegans my2a</td>
<td>D-EA-KCF-IL-EEF---A-TR</td>
<td>R403</td>
</tr>
<tr>
<td>C. elegans my3a</td>
<td>D-LNA-AML-ICAF---A-TR</td>
<td>R403</td>
</tr>
<tr>
<td>Drosophila</td>
<td>GGVSTK-F-CPT-E-Y-N-LK</td>
<td>R403</td>
</tr>
<tr>
<td>Dictyostelium</td>
<td>ALPASTVF-V-PSE-E-ME</td>
<td>R403</td>
</tr>
<tr>
<td>Acanthamoeba</td>
<td>LNTA-E-L-VSA-G-KA-LS</td>
<td>R403</td>
</tr>
<tr>
<td>Yeast</td>
<td>S-IY---N---VDEK-PQHTLR</td>
<td>R403</td>
</tr>
</tbody>
</table>

**FIGURE 1.** The amino acid sequence of myosin heavy chains from a variety of species and the amino acid sequence of exon 13 in FHC patients with a point mutation in this exon. The R403 position of the human β-cardiac MHC (R404 in unc-54) is marked with an *.* Identical residues are represented by dashes and the location of invariant residues are indicated by boxes. The amino acid sequence is represented by the single letter code (R=arginine; Q=glutamine). [Adapted from Geisterfer-Lowrance et al., (1990) Cell 62: 999-1006.]
between these filaments. Electron micrographs of animal skeletal and cardiac muscle show alternating light and dark bands and, therefore, they are often referred to as being striated (Figure 2). A cross-sectional view of vertebrate striated muscle shows that thick and thin filaments are arranged in a regular hexagonal array with each thick filament surrounded by six thin filaments while each of the thin filaments are within close proximity to only three thick filaments (Figure 3).

In 1954 two independent research groups, Andrew Huxley and Ralph Niedergerke, and Hugh Huxley and Jean Hanson put forward the sliding filament model of muscle contraction (Huxley and Niedergerke, 1954; Huxley and Hanson, 1954). The fundamental feature of this model is that the thick and thin filaments slide past one another in the contraction/relaxation cycle of muscle. Initial evidence for this model came from microscopic examination of muscle, which demonstrated that the length of the sarcomere decreased due to a shortening in the length of the I band and H zone but not the A band upon contraction (Figure 4). Force generation during muscle contraction is mediated by crossbridges between thick and thin filaments and is powered by the hydrolysis of ATP. How chemical energy released upon ATP hydrolysis is transduced to force generation needed to power the movement of the actin filaments by myosin is still not clearly understood. The cyclic mechanism of vertebrate skeletal muscle contraction which incorporates both ATP hydrolysis and sliding filament models is illustrated in Figure 5. The binding of ATP to a domain of the globular head of myosin produces a conformational change which results in the dissociation of myosin from actin. The ATP molecule is rapidly hydrolyzed to ADP and P$_i$. Concomitant with hydrolysis further conformational changes may occur in the myosin head which favor its binding to actin. The myosin head then becomes activated and P$_i$ is released, followed by a large conformational change that produces what is referred to as the powerstroke. Actin thin filaments move approximately 10 nm relative to myosin during a single powerstroke. It is also during the powerstroke
FIGURE 2. Sarcomere structure. Low magnification electron micrograph of a longitudinal section through a skeletal muscle cell of a rabbit, showing portions of two adjacent myofibrils and the definition of a sarcomere. Below is a schematic diagram of a single sarcomere, showing the origin of the dark and light bands seen in the electron micrographs. [Source: Garrett, R.H. and Grisham, C.M. Biochemistry (1995).]
**Figure 3.** Electron micrographs of an insect flight muscle viewed in cross-section. The myosin filaments, the large black spots, are arranged in a regular hexagonal lattice, with the actin filaments, small black spots, placed regularly between them. [Source: Garrett, R.H. and Grisham, C.M. *Biochemistry* (1995).]
FIGURE 4. The sliding-filament model of muscle contraction. Two sarcomeres, schematically shown extended and contracted.
 FIGURE 5. The proposed mechanism of skeletal muscle contraction. The stages in the interaction of a single myosin globular head with an actin thin filament are shown. Three free energy of ATP hydrolysis drives a conformation change in the myosin head, resulting in net movement of the myosin head along the actin filament. [Source: Matthews, C.K. and van Holde, K.E. *Biochemistry* (1990).]
that ADP is released which leads to the rapid binding of an ATP molecule, allowing for the cycle to be repeated. Solution kinetic studies and fiber experiments have demonstrated that the transduction of chemical energy to mechanical force (displacement of actin) occurs during the release of $P_i$ and ADP rather than the ATP hydrolysis step.

Biochemical (Cooke, 1986) and genetic (Wood, 1988) analysis has provided a great deal of information about MHC structure-function relationships. The myosin heavy chain monomer contains a globular head and helical tail. The functional unit of myosin is a hexamer comprised of two MHC monomers, whose helical tails are entwined to form a coiled-coil tail, and four myosin light chains (MLC) - two bound to each of the globular heads (Cooke, 1986) (Figure 6). In addition to MLC binding, the MHC head region has been shown to possess functional domains involved in ATP-binding and hydrolysis and actin binding (Karn, et al., 1983; Dibb, et al., 1989; Mitchell, et al., 1989). The MHC helical tails are responsible for self-association of myosin hexamers in the formation of the thick filament backbone. The myosin rods pack in a bipolar fashion within thick filaments. This results in a "bare zone" in the central portion of the filament and is the location for M line proteins which cross-link the thick filaments providing the ordered array within the myofilaments. The remainder of the filament has exposed globular heads which form crossbridges with the actin component of the thin filaments.

Limited proteolytic digestion of myosin with trypsin produces two fragments designated heavy meromyosin (HMM) and light meromyosin (LMM) (Figure 7). The HMM fragment (150 kD) consists of a rod-like tail with two globular heads while the LMM fragment (70 kD) is a 180 Å α-helical rod-like tail. HMM fragments are capable of binding MLC and actin, and possesses ATPase activity, LMM does not. Further proteolytic cleavage of HMM with papain produces two fragments designated S1 and S2. The S1 fragments are individual 95 kD globular heads, which retain the binding and hydrolysis properties of HMM. The remainder of the HMM fragments, 55 kD rods, are the S2
FIGURE 6. The myosin molecule. The model depicts the six polypeptide chains. The two globular heads are connected by the intertwining of two α-helices in the rodlike tail of the molecule. The globular headpieces each carry two noncovalently bound myosin light chains. [Source: Garrett, R.H. and Grisham, C.M. Biochemistry (1995).]
**FIGURE 7.** Dissection of myosin by proteases. Trypsin cleavage cuts in the tail to produce light meromyosin (LMM) and heavy meromyosin (HMM). Treatment of heavy meromyosin with the protease papain digests part of the stalk structure, allowing separation of the two headpieces, with their bound light chains, as S1 fragments, and releases the remainder of the tail as an S2 fragment. [Source: Garrett, R.H. and Grisham, C.M. *Biochemistry* (1995).]
fragments. Further tryptic digestion of S1 yields three principal fragments, often referred to in the description of the chemical and physical properties of myosin. The N-terminal 25 kD fragment possesses the ATP binding pocket, while regions from both the central 50 kD and C-terminal 20 kD fragments contribute to the actin binding domain.

The actin (i.e. thin) filaments to which myosin binds resembles two strands of beads twisted around one another to form a two-stranded helix. The individual "beads" are G-actin monomers with a molecular weight of 42 kD. Each monomer is a two-domain molecule and can bind one ADP or ATP. Due to the asymmetrical nature of the domains the molecule has a polarity and monomers polymerize in a head to tail fashion to form the fiber. Therefore, the actin filament has a defined directionality, like thick filaments.

Evolutionary conservation and a causative association with FHC suggest that the R403 containing domain plays an important role in MHC function. However, this role remains unknown. Some insight into this question was provided by the recent determination of the three-dimensional structure of chicken skeletal myosin S1 by Rayment and colleagues (1993a). The human β-cardiac MHC and chicken skeletal MHC have 79% sequence identity (Rayment, et al., 1995). Using the molecular structures of actin and myosin a model of the rigor complex was constructed (Rayment, et al., 1993). R405, the residue in chicken MHC corresponding to R403 in the human β-cardiac MHC, was located within a loop at the edge of the actin binding pocket. This loop is formed by residues R405 - K415. The loop extends away from the myosin molecule and toward actin and forms a close contact with residues P332 - E334 of actin. The conservation of myosin sequences from diverse species suggests that their three-dimensional structures would be very similar. Therefore, R405/R403 may be functionally important in the actin-myosin complex through a direct interaction with actin or may contribute to the stability of the loop (Rayment, et al., 1993b; Rayment, et al., 1995). Such a role for R403 in MHC function would have direct implications for the molecular basis of FHC. For example, the R403Q mutation may result
in an altered and thereby less efficient interaction between myosin and actin, thereby leading to hypertrophy in response to increased exertion of the myocardium. Determination of the three-dimensional structure of a protein has been used to gain insight in the elucidation of the molecular basis of disease states such as ornithine transcarbamylase deficiency, X-linked agammaglobulinemia, neurodegenerative prion diseases, and Alzheimer's disease (Tuchman et al., 1995; Vihinen et al., 1995; Huang et al., 1994; Struthers et al., 1991).

This dissertation describes efforts to investigate the function of the role of the R403-containing domain in MHC function and how it is affected by the R403Q mutation. An "animal model" was generated by the introduction of the corresponding mutation into the MHC of the muscle of an experimentally tractable organism with subsequent analysis of its effects. I have addressed this problem in the free-living soil nematode Caenorhabditis elegans, a favorite model organism for molecular genetic analysis of a wide range of problems in biology. Figure 8 depicts the four larval and adult developmental stages of the C. elegans hermaphrodite. Over the past thirty years C. elegans has emerged as a useful organism for such studies because of its small size (50-80 microns in diameter and 1 millimeter in length), ease of culture and manipulation in the laboratory. In addition, its mode of reproduction lends itself to genetic analysis. There are two sexes, hermaphrodites which are self-fertilizing and males which arise spontaneously at a low but detectable frequency (0.1%) and capable of cross-fertilizing with hermaphrodites. Also beneficial is a short life cycle of approximately 3.5 days and an average lifespan of 20 days (Kenyon, et al., 1993). A self-fertilizing hermaphrodite can produce a brood size of about 300, whereas one that has been cross-fertilized is capable of producing upwards of 800 progeny (Epstein and Shakes, 1995). Self-fertilized XX hermaphrodites yield hermaphrodite progeny exclusively with the exception of a rare XO male (<1/500) arising as the result of meiotic nondisjunction of the X chromosome (Hodgkin et al., 1979). The large brood size
aids in the recovery of mutants following exposure to ionizing radiation or chemical mutagenesis (Epstein and Shakes, 1995). Its simplistic anatomy (959 somatic nuclei in the hermaphrodite and 1031 in the male) has allowed for the characterization of every cell and cell contact for in the wild-type animal. Microscopic analysis of \textit{C. elegans} is facilitated because it is transparent. Lastly, \textit{C. elegans} has a small genome (8 X 10\(^7\) bp per haploid genome) (Brenner, 1974; Wood, 1988; Epstein and Shakes, 1995). Many of these features are important for my studies.

Other features of particular importance such as mapping, linkage analysis, complementation and epistasis are routine genetic procedures in \textit{C. elegans}. The entire \textit{C. elegans} genome has been cloned into yeast artificial chromosomes (YACs) and in an ordered, overlapping set of cosmids (Coulson \textit{et al.}, 1986; Coulson, \textit{et al.}, 1988). The cloning of \textit{C. elegans} genes has been facilitated by the alignment of the physical map with the very extensive genetic map. There is an expected completion date of 1998 for the determination of the nucleotide sequence of the entire \textit{C. elegans} genome.

Additionally, \textit{C. elegans} is ideal for the studies I have undertaken for three reasons. First, much is known about \textit{C. elegans} muscle cells. Due to the highly conserved nature of muscle proteins and structure the study of muscle in other animals has been aided by the information that has come from its analysis in \textit{C. elegans} (Waterston, 1988).

Second, molecular genetic analysis of muscle is well developed in this organism (Waterston, 1988). Over two dozen \textit{C. elegans} genes have been identified that encode muscle components (Wood, 1988). One of these, \textit{unc-54}, encodes the predominant myosin heavy chain of the body-wall muscle of this organism. This gene is well characterized at both genetic and molecular levels. Hundreds of loss-of-function mutations in \textit{unc-54} have been isolated and confer a very distinctive and recognizable phenotype: mutant animals are paralyzed and egg-laying defective (Karn, \textit{et al.}, 1983). The gene has been cloned and its genomic and cDNA sequences determined; the first for a MHC gene (or
protein) from any organism (Karn, et al., 1983). The unc-54-encoded MHC exhibits extensive sequence similarity with MHCs from other organisms and the R403 containing domain is highly conserved (Geisterfer-Lowrance, et al., 1990). In the UNC-54 protein, the corresponding arginine residue is R404. In addition, the β-cardiac and unc-54-encoded MHC proteins are the predominant of two isoforms expressed within their respective tissues. Adult human myocardium possesses both α- and β-cardiac MHC, but in different proportions within the atrial and ventricular walls. The β MHC isoform is the principal myosin component of the ventricles of the heart, whereas the α MHC isoform is predominant in the atria. The unc-54 gene product is the predominant of two myosin isoforms expressed in the body-wall muscles of C. elegans.

Third, methods for germ-line transformation and targeted gene replacement are available in this organism. Transformation of C. elegans is accomplished by microinjection of transforming DNA into the gonad of hermaphrodite worms. Injected DNA commonly forms tandem arrays which can be maintained extrachromosomally or integrated into the chromosomes via γ-irradiation (B. Herman, pers. comm.). A strategy for targeted alteration of the genome has been demonstrated using transgene-instructed repair of double stranded breaks in DNA induced by transposon activity. The method was first developed using P elements in Drosophila (Gloor et al., 1991) and subsequently in C. elegans using Tc1 elements (Plasterk and Groenen, 1992). The technique exploits certain properties of these transposable elements. While those studies demonstrated the feasibility of this approach in flies and worms, respectively, my use of this technique to investigate the R404-containing domain function represents its first successful use to answer a biological question. The strategies of the approach taken in this study as well as those for use of the technique in general will be discussed.

This dissertation describes the use of the soil nematode Caenorhabditis elegans to investigate the function of the R403 residue and the domain in which it resides.
While the structure of vertebrate and *C. elegans* sarcomeres are analogous, there are three principal differences: (Waterston, 1988). (1) The nature of the striation pattern. In both vertebrate and *C. elegans* muscle the thick and thin filaments run along the longitudinal axis of the muscle cell. However, their arrangement along that axis differs. Vertebrate myofilaments are parallel to the long axis while *C. elegans* myofilaments are arranged obliquely (Figure 9). The organization of obliquely striated fibers is described as pennate. The oblique striation is the result of adjacent filaments around the circumference of the cell being offset from one another along the longitudinal axis. While the filaments are parallel with the long axis, the stagger between them results in the birefringent striations being at a 6° angle to the filament axis. Pennate-fiber muscle is more powerful than parallel muscle of the same weight because more fibers are able to work in parallel. The possible implication of pennate organization on the work described in this dissertation will be discussed later.

(2) Size and composition of thick and thin filaments. The thick filaments of *C. elegans* are larger than those of vertebrates, with an average length of 9.7 μm and a diameter at the center of 33.4 nm which tapers to about 14 nm distally. This compares to vertebrate thick filaments with a length of 1.6 μm and a diameter of 12-14 nm. The thin filaments of *C. elegans* and vertebrates share the same 6 fold difference as seen for the thick filaments, (6μm vs. 1 μm, respectively) whereas their diameters (8 nm) are comparable. The two also differ in the relative arrangement of the thick and thin filaments. Each of the *C. elegans* thick filaments are surrounded by 12 thin filaments (Waterston, *et al.*, 1980), whereas in vertebrates six thin filaments surround each thick filament (Amos, 1985). Similar to other invertebrates and in contrast to vertebrates, *C. elegans* thick filaments have a paramyosin core.

(3) Attachment of the muscle cells. Both vertebrate and nematode muscle cells contain attachment plaques at their ends and it is to these structures that the (+) end of the longitudinally arranged thin filaments attach. In vertebrate muscle there are attachment
Figure 9. Vertebrate cross-striated and *C. elegans* obliquely striated muscle structures. (A) Vertebrate striated muscle. A vertebrate muscle cell is multinucleate and 10-50 microns in diameter and can be many millimeters long. Z discs anchor thin filaments which interdigitate with thick filaments. A bands contain thick filaments which are about 1.5 microns long. The cross striations are orthogonal to the filament axes. (B) *C. elegans* obliquely striated muscle. A *C. elegans* adult nematode contains 95 mononucleate muscle body-wall muscle cells arranged in four strips running along the body axis, two dorsally and two ventrally. The analogs of vertebrate Z discs are dense bodies to which body-wall thin filaments are attached at one end. Thick filaments are approximately 10 microns long. While the filaments lie parallel to the longitudinal axis of the animal the adjacent units are staggered. The result is striations that are at an oblique angle of about 6° to the long axis of the filaments. [Source: Anderson, P. (1989) Molecular genetics of nematode muscle. *Annu. Rev. Genet.* 23: 507-25.]
plaques at the ends of skeletal muscle where it joins the tendon and specialized junctions (i.e. intercalated disks) between cardiac muscle cells. While it is unquestionable that there is tension transmitted between *C. elegans* muscle cells by way of attachment plaques, the majority appears to be transferred directly, via lateral attachments, to the cuticle.

The adult *C. elegans* hermaphrodite has 135 muscle cells, 95 of which comprise the body-wall musculature (Waterston, 1988; Anderson, 1989). These cells are rhomboid-shaped, mononucleate, multisarcomeric and located in four longitudinal strips along the body axis (Figure 10). The left and right subdorsal quadrants, each containing 24 muscle cells, oppose the left and right subventral quadrants which contain 23 and 24 muscle cells, respectively. The arrangement of overlapping pairs of muscle cells is more pronounced in the anterior portion of the animal than the posterior. A basement membrane separates the myofilament lattice, located in a 1-2 μm deep region, from the hypodermis that underlies the exterior cuticle. M lines and dense bodies, the component of Z lines, anchor the thin and thick filaments, respectively to the plasma membrane of the muscle cell by way of a series of lateral attachments. Small projections from these attachment sites extend into the basement membrane of the hypodermal syncytium which possess an extensive array of filaments that make contact with the cuticle. These lateral attachments when brought closer together during contraction of the sarcomere shorten the animal. The coordination of the process by the nervous system can produce bends in the animal leading to the progression of a sinusoidal wave along the body to provide backward or forward movement.

I have exploited many of the features that *C. elegans* has to offer to investigate the molecular basis of FHC caused by the most common mutation R403Q in human β-cardiac MHC. To do so I introduced the corresponding mutation (R404Q) into the *unc-54* gene using a novel gene replacement approach and have characterized its effects on a range of features using genetic, molecular, ultrastructural and behavioral analysis. The result is that in all cases there is no detectable difference between the FHC mutant and wild type.

20
Figure 10. *C. elegans* body-wall muscle ultrastructure. Polarized light micrograph of a young adult showing right and left dorsal muscle quadrants, which are separated by a dark strip (arrowhead). The curved arrow indicates the margins of elongate, spindle-shaped cells which are arranged in overlapping pairs in the anterior portion of the animal, with progressively less overlap and less pairing toward the posterior end. Bar represents 50 μm. [Source: Wood, W.B. ed. *The nematode Caenorhabditis elegans*. (1988).]
Possible explanations for this and implications for MHC and muscle function, and the molecular basis of FHC will be discussed.
MATERIALS AND METHODS

NEMATODE GROWTH CONDITIONS:

Growth and manipulation of *Caenorhabditis elegans* strains was as described (Brenner, 1974). The strain OP50 of *Escherichia coli* was supplied as the food source on solid nematode growth media (NGM: 3 g NaCl, 17 g agar, 2.5 g peptone, per liter, and 1 mM CaCl₂, 1 mM MgSO₄, 1 mM KH₂PO₄ pH 6.0, and 5 μg/mL cholesterol). DNA isolated for use in Southern blot analysis was obtained from worms grown in liquid culture at 20°C containing M9 buffer (3 g KH₂PO₄, 6 g Na₂HPO₄, 50 mg NaCl, per liter and 1 mM MgSO₄), S-medium (10 μg/mL cholesterol, 10 mM potassium citrate pH 6.0, 3 mM CaCl₂, 3 mM MgSO₄, 0.5 mM EDTA, 0.25 mM FeSO₄, 0.1 mM MnCl₂, 0.1 mM ZnCl₂, 0.1 mM CuSO₄) and 40X *E. coli* strain P90C (Miller et al., 1977) as the food source.

*Caenorhabditis elegans* STRAINS:

Strains used for injection were:

<table>
<thead>
<tr>
<th>STRAIN</th>
<th>GENOTYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 var. Bristol</td>
<td>wild-type</td>
</tr>
<tr>
<td>TR250</td>
<td><em>unc-54(r856r857)</em></td>
</tr>
<tr>
<td>TW333</td>
<td><em>mut-2</em>r459)<em>unc-54(r360::Tc1)</em></td>
</tr>
</tbody>
</table>

Strains used for genetic analysis:

<table>
<thead>
<tr>
<th>STRAIN</th>
<th>GENOTYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW186</td>
<td><em>mut-2(r459)</em></td>
</tr>
<tr>
<td>TR656</td>
<td><em>unc-54(r360::Tc1)</em></td>
</tr>
<tr>
<td>TW390</td>
<td><em>unc-54(cj289Gr)</em></td>
</tr>
<tr>
<td>BC200</td>
<td><em>unc-22(s12)</em></td>
</tr>
</tbody>
</table>
CA1  
unc-54(cj289Gr); unc-22(s12)

TW192  
unc-105(n490); mut-2(r459) unc-54(cj177)

CA2  
unc-54(cj289Gr); unc-105(n490)

The originally isolated gap replaced hermaphrodite was used to establish a clonal population:

<table>
<thead>
<tr>
<th>LINE</th>
<th>GENOTYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>R3</td>
<td>mut-2(r459) unc-54(cj289Gr)</td>
</tr>
</tbody>
</table>

For genetic manipulation of R3 and TW390, males were recovered following incubation of L4 larva at 30°C for 6 hours.

**PLASMIDS:**

The following plasmids were used in this study:

<table>
<thead>
<tr>
<th>PLASMID</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR#124</td>
<td>genomic clone of the wild-type <em>unc-54</em> gene</td>
</tr>
<tr>
<td>TW#26</td>
<td><em>Bam</em> HI restriction fragment of TR#124 cloned into <em>Bam</em> HI site of pBR322 (Sutcliffe, 1979; Watson, 1988)</td>
</tr>
<tr>
<td>TW#27</td>
<td>identical to TW#26 except for the base substitutions GT -&gt; AA (nucleotides 3278, 32-79 [Karn et al., 1983])</td>
</tr>
<tr>
<td>TW#28</td>
<td>identical to TR#124 except for the base substitutions GT -&gt; AA (nucleotides 3278, 32-79 [Karn et al., 1983])</td>
</tr>
<tr>
<td>pBR322</td>
<td>cloning vector (Sutcliffe, 1979; Watson, 1988)</td>
</tr>
<tr>
<td>pRF4</td>
<td>genomic clone of the <em>su1006</em> allele of the <em>rol-6</em> gene</td>
</tr>
</tbody>
</table>

**PLASMID CONSTRUCTIONS:**

TR#124 represents the *Pst I - Xba I* restriction fragment of the wild-type *unc-54* gene [nucleotides 285-9405 (Karn, 1983)] inserted into *pEMBL8* (Dente et al., 1983) cut...
with Pst I and Eco RI. An Xba I/Eco RI adapter was used to modify the pEMBL8 EcoRI site to an Xba I site. The construction of TW#26 was initiated by cutting TR#124 with Bam HI (bp 1571-4401) and ligating the 2.8 kb fragment at the Bam HI site of pBR322 (Sutcliffe, 1979; Watson, 1988). The Bam HI fragment possessed the R404 site and greater than 1 kb of flanking DNA on each side. A 144 bp Kpn I (bp 3203)-Nco I (bp 3347) fragment, containing the R404 site, from TW#26 was replaced with a Kpn I-Nco I mutagenized PCR product possessing the desired base substitutions to codon 404 necessary for glutamine to be substituted for arginine. This plasmid is called TW#27. TW#28 was constructed by removing the Bam HI fragment, possessing the corresponding FHC mutation, from TW#27 and ligating it into the Bam HI site full length unc-54 clone possessing the R404Q mutation. pRF4 represents a 4 kD C. elegans genomic DNA Eco RI fragment containing the su1006 allele of the rol-6 gene (16) in the Bluescribe vector (Kramer et al., 1990). Plasmids used for injection were prepared by alkali lysis and purified by precipitation with polyethylene glycol as described (Sambrook, et al., 1989). Injection concentration of plasmids was 100 μg/mL.

**PLASMID ISOLATIONS:**

Overnight cultures of 1-3 mL of TYE (8 g NaCl, 10 g bactotryptone, 5 g yeast extract, per liter) and ampicillin (50 μg/mL) were inoculated with bacteria containing the plasmid of interest and grown at 37°C. Plasmids were then isolated using one of two methods.

**Magic™ Minipreps DNA Purification kit:**

Plasmids were isolated using the Magic™ Minipreps DNA Purification kit (Promega, Madison, WI) as described by the manufacturer.

**STET preparation:**
The bacteria were pelleted by centrifugation at 14,000 g for 20 secs and then resuspended in 210 µL STET (0.1 M NaCl, 10 mM Tris-Cl, 1 mM EDTA pH 8, 5% Triton X-100). Tubes were vortexed after adding 15 µL of freshly prepared lysozyme (10 mg/mL), then placed in boiling water for 45 secs. followed by centrifugation at 4,000 g for 10 mins at room temperature. The pellet was removed and discarded and the nucleic acids precipitated by adding 500 µL of 100% ethanol at -20°C and 22 µL 3 M NaOAc followed by inverting the tube several times and then placing at -20°C for at least 30 mins. The nucleic acids were pelleted by centrifugation at 14,000 g for 10 mins. at 4°C. The supernatant was decanted and 200 µL 70% ethanol at 4°C was used to rinse the pellet followed by centrifugation at 14,000 g for 10 mins. at 4°C. The supernatant was removed and the pellet was air dried for 30 mins and then resuspended in 50 µL TE (10 mM Tris-Cl, 1 mM EDTA pH 8). 1 µL RNase A (5 mg/mL stock) was added and then incubated at 37°C for 30 mins. The volume of the sample was increased to 250 µL with TE to which 250µL of chlorophorm:phenol (1:1) was added. The tubes were vortexed for 1 min and then centrifuged at 14,000 g for 5 mins. at room temperature. The aqueous phase was transferred to a new tube and the plasmid was precipitated using ethanol as described above. The DNA pellet was resuspended in 30 µL TE.

SYNTHETIC OLIGONUCLEOTIDES:

The following are the oligonucleotides used for site-directed mutagenesis by PCR and DNA sequencing and were purchased from Operon, Alameda, CA:

<table>
<thead>
<tr>
<th>PRIMER</th>
<th>SEQUENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>JC24</td>
<td>5'-GTTCAAGCAACGCCACG-3'</td>
</tr>
<tr>
<td>JC25</td>
<td>5'GGTTCCGACCTTGAACCTTGCTGGCTTGGTC-3'</td>
</tr>
<tr>
<td>JC26</td>
<td>5'-GCCAGTGGACACACACGGA-3'</td>
</tr>
<tr>
<td>JC27</td>
<td>5'-GACCAAGCCACACTAAGTGCAAGTCCGGAAAC-3'</td>
</tr>
</tbody>
</table>

26
NEMATODE DNA ISOLATIONS:

DNA was isolated from nematodes by one of two ways depending on its intended use.

PCR DNA Isolation:

Worms were grown on NGM plates and washed off with 500 μL M9 buffer. The worms were pelleted by centrifugation at 14,000 g for 15 secs at room temperature. The supernatant was removed and the worms were resuspended in 500 μL of sterile still water and again pelleted. Following the removal of the supernatant the worms were resuspended in 200 μL of worm lysis buffer (50 mM KCl, 2.5 mM MgCl2, 10 mM Tris pH 8.2, 0.45% NP40, 0.45% Tween 20, 0.01% gelatin) to which 2 μL proteinase K (10 mg/mL stock) (USB, Cleveland, OH) was added and then placed in at -80°C (dry ice/ethanol bath) for 10 mins. The tubes were then incubated at 65°C for 1 hr. followed by 95°C for 15 mins. to inactivate the proteinase K. The isolated DNA was stored at -20°C and 1 μL was used at template per PCR reaction.

Southern blot DNA isolation:

Worms were grown in liquid culture until near exhaustion of the food supply. They were then transferred to a 15 mL Falcon tube and pelleted by centrifugation at 3,000 r.p.m. in an IEC clinical centrifuge for 10 secs. and the supernatant was removed by aspiration. Bacteria were washed from the worms by resuspending the worm pellet in 10 mL of milli-Q water followed by centrifugation and removal of the supernatant as previously described. A total of three washes was done. The worms were then captured by floating them on a bed of 35% sucrose as follows: The worm pellet obtained following the third wash was resuspended in 7 mL of milli-Q water at 4°C to which 7 mL of a 70% sucrose solution was added. The tube was inverted sharply several times to obtain thorough mixing of the sucrose. The tube was then centrifuged at 3,000 r.p.m. for 3 mins. The worms were withdrawn from the top of the tube and transferred to a new 15 mL
Falcon tube. The worms were washed three times as described above with the exception that DNA disruption buffer (0.2 M NaCl, 0.1 M Tris, 0.05 M EDTA, 0.5% SDS, pH8.5) was used in place of milli-Q water. Following the final wash and spin, all but 500 µL of DNA disruption buffer was removed and the DNA was stored at -20°C.

The DNA was purified as follows: The 500 µL of DNA in disruption buffer was transferred to a 600 µL microfuge tube to which 10 µL of proteinase K (10 mg/mL stock) was added. The tube was mixed and incubated at 65°C for 30 mins. at which time an additional 5 µL of proteinase K was added and the reaction was incubated at 65°C for another 20 mins. The tube was then cooled to room temperature and extracted with 515 µL of chloroform:phenol (1:1) with vigorous shaking and inversion of the tube for 4 mins. The aqueous and organic layers were separated from one another by centrifugation at 14,000 g for 5 mins. at room temperature. The aqueous phase was transferred to a new microfuge tube and the extraction with chloroform:phenol (1:1) was repeated twice more followed by a single extraction with chloroform only. Nucleic acids were precipitated by adding 1.2 mL of 100% ethanol at -20°C and 50 µL 3 M NaOAc followed by inverting the tube several times and then placing at -20°C for at least 30 mins. The nucleic acids were pelleted by centrifugation at 14,000 g for 10 mins. at 4°C. The supernatant was decanted and 200 µL 70% ethanol at 4°C was used to rinse the pellet followed by centrifugation at 14,000 g for 10 mins. at 4°C. The supernatant was removed and the pellet was air dried for 30 mins and then resuspended in 200 µL TE (10 mM Tris-Cl, 1 mM EDTA pH 8). 2 µL RNase A (5 mg/mL stock) was added and then incubated at 37°C for 30 mins. at which time an additional 2 µL RNase A was added and the tube was incubated for another 30 mins. at 37°C. The tube was then extracted twice with phenol:chloroform (1:1) followed by a chloroform:isoamyl alcohol (24:1) extraction as described above. The DNA was precipitated as described above using 500 µL of ethanol and 25 µL of 3 M sodium acetate. The pellet was resuspended in 75 µL of TE. 2 µL of the sample was electrophoresed on a
0.8% SeaKem LE agarose (FMC, Rockland, ME), 1X TBE (45 mM Tris pH 8, 45 mM Borate pH 8 and 5 mM EDTA pH 8) gel to check for the integrity and concentration of the DNA and the presence of RNA.

**POLYMERASE CHAIN REACTIONS:**

The PCR reactions were made up as follows:

<table>
<thead>
<tr>
<th>Volume</th>
<th>Component</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>38.1 µL</td>
<td>UV-irradiated sterile still water</td>
<td></td>
</tr>
<tr>
<td>0.4 µL</td>
<td>deoxyribonucleotide mix [25 mM/nucleotide]</td>
<td>Pharmacia, Piscataway, NJ</td>
</tr>
<tr>
<td>5 µL</td>
<td>Taq DNA poly 10X Buffer B [Mg²⁺ free]</td>
<td>Promega, Madison, WI</td>
</tr>
<tr>
<td>3 µL</td>
<td>MgCl₂ [25 mM]</td>
<td>Promega, Piscataway, NJ</td>
</tr>
<tr>
<td>0.5 µL</td>
<td>Taq DNA polymerase (Storage Buffer B)</td>
<td>Promega, Piscataway, NJ</td>
</tr>
<tr>
<td>1 µL</td>
<td>DNA primer #1 [100 pmol]</td>
<td>OPERON, Alameda, CA</td>
</tr>
<tr>
<td>1 µL</td>
<td>DNA primer #1 [100 pmol]</td>
<td></td>
</tr>
<tr>
<td>1 µL</td>
<td>template</td>
<td></td>
</tr>
<tr>
<td>50 µL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

25 µL of mineral oil was used as an overlay for the reactions. The reactions were placed in one of two thermal cyclers, a Perkin Elmer Cetus DNA Thermal Cycler (Perkin Elmer Cetus, Norwalk, CT) or Coy Tempcycler (Coy Laboratory Products, Ann Arbor, MI), and the following profile parameters were used: an initial 4 min. denaturation of the template at 94°C followed by twenty-nine cycles of 94°C for 1 min., 55°C for 1 min., 72°C for 2 mins. with a final 10 min. extension at 72°C. Post-PCR reactions were stored at -20°C.
SITE-DIRECTED MUTAGENESIS BY PCR:

The polymerase chain reaction (PCR)-based strategy to change the arginine codon (CGT) to a glutamine codon (CAA) was performed as described (Higuchi, et al., 1988). TR#124 was used as template in both first round reactions. In one of the two reactions the flanking primer JC24 along with the mutagenic primer JC25 were used, whereas in the other reaction the flanking and mutagenic primers were JC26 and JC27. Excess primers were removed by excising the PCR products from a 2% NuSieve GTG agarose (FMC, Rockland, ME), 1X TAE (4.8g Tris, 1.1 mL glacial acetic acid, 2 mL 0.5 M EDTA pH 8) gels following electrophoresis. The excised band was heated to 94°C for 10 mins. and 1μL was used as template in the second round of PCR. The final PCR product was sequenced directly using the CircumVent™ Thermal Cycle Sequencing kit with Vent (exo-) DNA polymerase as described by the manufacturer (NEB, Beverly, MA) to confirm the presence of the desired base changes. It was then digested with Kpn I and Nco I so as to achieve directional cloning in the preparation of TW#27.

CYCLE SEQUENCING:

DNA samples were sequenced using the CircumVent™ Thermal Cycle Sequencing kit with Vent (exo-) DNA polymerase (New England Biolabs, Beverly, MA) following the directions of the manufacturer as to reaction preparation and thermal cycler profiles. Samples were placed in boiling water for 2 mins. and then immediately on ice prior to being electrophoresed on a 6% urea-polyacrylamide gel [12 mL 40% acrylamide (19:1 acrylamide:bis), 15 mL 5X TBE, 36 g urea, 20 mL milli-Q water, 450 μL 10% ammonium persulfate, 20 μL TEMED]. Following electrophoresis the urea was removed from the gel by soaking it for 15 mins. in a mixture of 5% methanol, 5% glacial acetic acid and 90% water and then vacuum dried. Kodak medical X-ray film (Eastman Kodak Co., Rochester, NY) was exposed to the gel for autoradiography.
DNA EXTRCTIONS FROM AGAROSE GELS:

PCR amplification products were separated by electrophoresis on 2% NuSieve GTG agarose (FMC, Rockland, ME), 1X TAE (4.8g Tris, 1.1 mL glacial acetic acid, 2 mL 0.5 M EDTA pH 8) gels. The desired band(s) were excised from the gel and their weight(s) determined. DNA extraction was performed using USBioclean MP™ kit (USB, Cleveland, OH) or β-agarase I (New England Biolabs, Beverly, MA) as described by the manufacturer.

GERM-LINE TRANSFORMATION:

Transformation by syncytial injection was performed essentially as described (Fire, 1986; Mello et al., 1991). Needles were pulled on a Kopf model 710C Vertical Pipette Puller (David Kopf Instruments). Eppendorf Microloaders (Eppendorf, Pittsburgh, PA) were used to load the needles used for injections. Needle tips were broken against Fisherbrand disposable micro-pipets (Fisher, Madison, WI) placed on a coverslip at approximately a 22° angle to the needle. The injections were done using a Zeiss IM microscope equipped with differential interference contrast optics. Injections were performed at 400X magnification using a Zeiss water objective. The needle was held and positioned using a Zeiss micromanipulator.

ISOLATION OF GERM-LINE TRANSFORMANTS:

Injected worms were placed singly on NGM plates seeded with E. coli OP50. If the parent was wild-type the progeny were screened for the Rol marker phenotype, whereas improved movement was screened for if the parent worm was paralyzed. Scored animals were picked singly to a fresh plate to establish a line.
ISOLATION OF GENE-REPLACED ANIMALS:

Injected worms were placed singly on NGM plates seeded with *E. coli* OP50. F₁ hermaphrodites with improved movement were isolated singly and their progeny (F₂) were screened for nonparalyzed animals. Single nonparalyzed F₂ animals were put on plates to establish clonal populations.

CONSTRUCTION OF STABLE TRANSFORMANTS BY γ-IRRADIATION:

Approximately 20 young adult transformed animals with the Rol phenotype were placed on a plate and exposed to γ radiation (20,000 rads/hour) from Cs¹³⁷ for 11.5 minutes for a total dosage of 3,800 rads. F₁ Rol hermaphrodites were picked singly to a new plate and their progeny (F₂) were screened for Rol animals. Single Rol F₂ animals were transferred to new plates to confirm 100% segregation of the Rol phenotype in the F₃ generation.

DETERMINATION OF LIFESPAN:

Eggs were placed on NGM plates seeded with *E. coli* OP50. Upon hatching, the time was noted and the L₁ animal was transferred to a freshly seeded plate. During the adult stage the animal was transferred to a freshly seeded plate every two days when it was producing progeny and once a week when it was no longer fertile. Death of an animal was scored when there was no movement even upon prodding and there was no pharyngeal pumping.

DETERMINATION OF BROOD SIZE:

Single L₄ animals were placed on plates seeded with *E. coli* OP50. The animal was transferred to a freshly seeded plate every day and the number of progeny were scored on the plate from which it was transferred.
STATISTICAL ANALYSIS:

The means of the brood sizes and lifespans of N2 and TW390 were compared using Minitab.

SOUTHERN BLOT ANALYSIS:

Genomic and plasmid DNA samples were digested with the restriction endonuclease \textit{Bam} HI and separated by electrophoresis on a 0.8% SeaKem LE agarose (FMC, Rockland, ME), 1X TBE gel. The gel was depurinated by two 15 min. soakings in 0.25 M HCl, denatured by soaking in 1.5 M NaCl, 0.5 M NaOH for 30 mins. and neutralized with a 30 min. soak in 1.5 M NaCl, 1 M Tris-HCl pH 7.5. The gel was then soaked in 20X SSC (3 M NaCl, 0.3 M sodium citrate) overnight. The transfer of DNA to a nitrocellulose membrane (NitroPlus, MSI, Westboro, MA) was by capillary action using 20X SSC for at least an 8 hr. period of time. The membrane was allowed to air dry and then the DNA was fixed to the membrane using UV transillumination for 2 mins. 6X SSC was used to rewet the membrane and then it was placed in prehybridization buffer [6X SSC, 5X Denhardt's reagent (USB, Cleveland, OH), 100 \textmu g/mL denatured herring sperm DNA and 50\% deionized formamide (USB, Cleveland, OH) at 42°C for 1.5 hours. pBR322 and TW#27 were used as probes following labeling with alpha-\textsuperscript{32}P-dATP using the Random Primed Labeling kit (USB, Cleveland, OH) and the procedure provided by the manufacturer. Column chromatography (Sambrook \textit{et al.}, 1989) was used to separated nonincorporated nucleotides from the probe. The membrane was incubated with the probe in hybridization buffer (identical to prehybridization buffer) at 42°C overnight. 3X, 1X and 0.3X SSC was used to wash the membrane for 10, 30 and 30 mins., respectively, at 65°C. Kodak medical X-ray film (Eastman Kodak Co., Rochester, NY) was exposed to the membrane at -80°C for autoradiography.
RESTRICTION ENDONUCLEASE DIGESTS:

Restriction endonucleases and buffers were purchased from New England Biolabs (Beverly, MA) and digests were performed as described by the manufacturer. The total volume of each digest was 20 μL.

LIGATION:

The DNAs used were PCR fragments, cloned genomic unc-54 fragments and the cloning vector pBR322 (Sutcliffe, 1979; Watson, 1988). The DNAs were digested with specific restriction endonucleases. 1 μL of calf intestinal phosphatase (New England Biolabs, Beverly, MA) was added to the digested vector, and the reaction was incubated at 37°C for 1 hr. The digested vector and the restriction fragments were separated on a 2% NuSieve™ GTG agarose (FMC, Rockland, ME), 1X TAE gel. The desired bands were excised, extracted, ethanol precipitated and resuspended in 20 μL. 2 μL of the resuspended DNA was electrophoresed on a 0.8% SeaKem LE agarose (FMC, Rockland, ME) gel to determine relative concentrations. The concentration of insert was 5-10 times that of vector and the ligation reaction consisted additionally of the following: 2.5 μL 10X T4 DNA ligase buffer, 1 μL T4 DNA ligase (Promega, Madison, WI) and sterile glass distilled water to bring the volume up to 25 μL. The reaction tubes were incubated between 1 hr. and overnight at 15°C.

TRANSFORMATIONS:

Competent E. coli, strain DH5α, were thawed on ice. The following amounts of DNA and cells were combined: 1 μL of a 1:10 dilution of the ligation reaction was added to 50 μL of cells for the experimental samples; for the positive control 5 μL of uncut pUC19 (0.5 ng) was added to 10 μL of cells; and no DNA was added to 10 μL of cells for the negative control. The DNA and cells were mixed gently but thoroughly and left on ice
for 30 mins. The DNA/cell mixtures were heatshocked at 37°C for 30 secs. and then placed on ice for greater than 2 mins. 1 mL TYE (8 g NaCl, 10 g bactotryptone, 5 g yeast extract, per liter) was added to each of the cell/DNA mixtures which were shaken for 1 hr. at 37°C. 200 µL of the experimental and negative control inoculums and 50 µL of the positive control inoculum was spread on TYE/AMP (50 µg/mL) plates which were then incubated overnight at 37°C. Single colonies were used to inoculate overnight cultures from which the plasmid was isolated.

**GENETIC ANALYSIS:**

**Construction of TW333 [genotype mut-2(r459)unc-54(r360::Tc1)]:**

To activate excision of Tc1 from *unc-54(r360::Tc1)*, TR656 hermaphrodites were crossed with TW186 males to introduce the *mut-2* mutator background. F1 heterozygous, nonparalyzed hermaphrodites were isolated singly and their progeny (F2) were screened for paralyzed animals. Single paralyzed F2 animals were put on plates to establish clonal populations. These lines were then screened for reversion to *unc-54*+. Twenty-five to thirty independent populations, for each line, were initiated simultaneously with 10 hermaphrodites per 60 mm Petri plate. The populations grew until either a non paralyzed revertant was identified or the food supply was exhausted. The plates were then counted and the reversion frequency calculated as the ratio of the total number of revertants per the total number of worms screened per line. TW333 represents the line with the highest reversion frequency, 4 X 10⁻⁴.

**Construction of TW390 [genotype unc-54(cj289Gr)]:**

The wild-type Bristol background was substituted for the *mut-2* mutator background in R3 by backcrossing males with TR250 hermaphrodites, using the *unc-54(r856r857)* mutation as an *unc-54* marker in *trans* to follow the *cj289Gr* locus. Several F1 nonparalyzed hermaphrodites were plated singly and their progeny (F2) were

35
screened for nonparalyzed animals which were plated singly to confirm homozygosity of the unc-54(cj289Gr) allele and establish a clonal population. The backcrossing series was continued for a total of 6 crosses, yielding TW390.

Construction of CA1 [genotype unc-54(cj289Gr):unc-22(s12)] :

TW390 males were crossed with BC200 hermaphrodites and F1 nonparalyzed, nontwitcher heterozygotes were picked singly and their progeny (F2) screened for twitchers. Isolated twitchers were picked singly to establish clonal populations. The frequency of twitchers was calculated as the ratio of the total number of twitchers per the total number of F2 progeny. The presence or absence of the unc-54(cj289Gr) allele in these lines was determined by PCR amplifying and sequencing the R404 region of unc-54. CA1 represents a strain homozygous for unc-54(cj289Gr) and unc-22(s12).

Construction of CA2 [genotype unc-54(cj289Gr):unc-105(n490)] :

TW192 hermaphrodites (paralyzed) were backcrossed with wild-type Bristol N2 males. Several F1 hermaphrodites were plated singly and their progeny screened for hypercontracted animals. The hypercontracted phenotype is due to the presence of the homozygous unc-105(n490) allele and absence of the suppressor unc-54(cj177) allele. Several hypercontracted hermaphrodites were plated singly to confirm loss of the unc-54(cj177) allele and to establish a clonal population. TW390 males were crossed with hypercontracted hermaphrodites and nonhypercontracted F1 hermaphrodites were plated singly. The F2 generation was screened for hypercontracted animals which were picked singly to establish clonal lines. The frequency of hypercontracted worms was calculated as the ratio of the total number of hypercontracted animals per the total number of F2 progeny. The R404 region of unc-54 was PCR amplified and sequenced to confirm the presence or absence of the unc-54(cj289Gr) allele in these lines.
**POST-EMBRYONIC DEVELOPMENT AND BROOD SIZE:**

Eggs were picked singly and time zero was recorded at the time of hatching. The end point of post-embryonic development was recorded when the first egg was laid. Brood size was determined by counting the number of progeny that hatched from the eggs laid by a single worm. Statistical analysis was performed using Minitab.

**MICROSCOPIC ANALYSIS:**

Polarized light microscopy was performed by Ken Norman (University of British Columbia at Vancouver) on the body-wall muscle of young adult males as described (Waterston *et al.*, 1980) using a Zeiss WL microscope equipped with strain free objectives and condensor and a 50-W Hg lamp.
RESULTS

The initial goal of this study was to generate a strain of *Caenorhabditis elegans* containing a mutation corresponding to one that causes an inherited heart disorder in humans. The mutation changes a single amino acid in the myosin heavy chain (MHC) protein, replacing an arginine residue (R) with a glutamine (Q). When this mutation is present in the human β-cardiac MHC gene (R403Q) it results in the lethal heart condition familial hypertrophic cardiomyopathy. This work describes the introduction of the corresponding mutation (R403Q) into the unc-54 gene, which encodes the major MHC in the body-wall muscle of *C. elegans*. To accomplish this a mutant unc-54 clone containing the R404Q mutation was generated *in vitro* and then introduced into the germ line of *C. elegans* to obtain unc-54(R404Q) transgenic lines. This was done with the goal of using the resulting mutant to investigate the role of this conserved arginine residue, and the affect of the arginine-to-glutamine mutation, on MHC function, and muscle structure and contraction. The rationale for this approach is that MHC structure and function are highly conserved in evolution and this arginine residue is invariant in MHC proteins from a wide spectrum of organisms. Hence, the role of this arginine residue is likely general to MHC proteins in all these organisms. Therefore, it should be possible to study these problems in an experimentally tractable organism like *C. elegans*. This animal offers many important advantages for such a study including an array of powerful methods and extensive knowledge of its muscle at the genetic and cellular levels.

A particular goal of this work was to exploit the R404Q mutant I generated, using genetic suppression analysis, to identify factors that interact with this MHC domain. *C. elegans* is ideal for this approach. However, in order to use it, the mutation introduced must confer a typical unc-54 mutant phenotype. This phenotype is very distinctive and
recognizable; unc-54 mutants are paralyzed and egg-laying defective. Unfortunately, this did not turn out to be the case, as described below.

**Construction of unc-54(R404Q) in vitro.**

To generate a mutant unc-54 gene encoding the R404Q mutation it was necessary to change the R404 arginine codon (CGU) to a glutamine codon (CAA). This was accomplished using a PCR-based in vitro mutagenesis strategy as illustrated in Figure 11 (Higuchi et al., 1988). Two overlapping "mutagenic oligos" were designed and used with two flanking primers (see Table 1 and Figure 12) to produce overlapping PCR products containing the CGU -> CAA double base substitutions (Figure 13). In the next step, these overlapping mutant fragments were fused via "recombinant PCR" (Figure 14) to generate a 227 bp fragment which represents the region of unc-54 containing the R404Q mutation. A restriction fragment from within this PCR product containing the R404Q mutation was cloned into the full-length unc-54 gene contained in the plasmid pUNK-54 (see Methods and Materials). Throughout this dissertation, the clone and genomic copy of the wild-type unc-54 gene will be referred to as pUNK-54 and unc-54, respectively and the clone and genomic copy of the unc-54 gene containing the R404Q mutation as pUNK-54(R404Q) and unc-54(R404Q), respectively.

To confirm the presence of the R404Q mutation (CGU -> CAA) within pUNK-54(R404Q) two tests were used. First, a restriction enzyme cleavage site change resulting from this mutation was exploited. The presence of the FHC mutation results in the loss of a PmII site that is present in the wild-type sequence (Figure 15). Therefore, digestion with PmII is useful as a diagnostic test. The wild-type and mutant unc-54 clones were digested with BamHI and PmII. (BamHI generates a 2.8 kb restriction fragment containing the R404 region.) As predicted, PmII cut within this BamHI fragment from the wild-type clone pUNK-54, but failed to cut the same fragment from the R404Q mutant clone,
FIGURE 11. Construction of the \textit{unc-54(R404Q)} mutation using PCR. Primers JC24 and JC26 flank the R404 site. Primers JC25 and JC27, which overlap and include the R404 site, contain the appropriate nucleotide substitutions (indicated by \_\_\_\_) to introduce the R404Q mutation.
Table 1. Primers used in the representative PCR reactions and the corresponding amplified segments of the *unc-54* gene.

<table>
<thead>
<tr>
<th>PCR Reaction</th>
<th>Primer</th>
<th>Primer Sequence</th>
<th>Corresponding nucleotide location in the <em>unc-54</em> gene</th>
<th>Length of amplified product</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>JC24</td>
<td>5'-GTTCACAGCCGCCACG-3'</td>
<td>3159-3294</td>
<td>136 bp</td>
</tr>
<tr>
<td></td>
<td>JC25</td>
<td>5'-GGTTCGACCTTGACTTGGCTTGGTC-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>JC27</td>
<td>5'-GACCAAGCACA AGTCAAGGGTTGAAC-3'</td>
<td>3267-3385</td>
<td>119 bp</td>
</tr>
<tr>
<td></td>
<td>JC26</td>
<td>5'-GCCAAGTTGACACACACGGA-3'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>JC24</td>
<td>5'-GTTCAGCAGCGCCACG-3'</td>
<td>3159-3385</td>
<td>227 bp</td>
</tr>
<tr>
<td></td>
<td>JC26</td>
<td>5'-GCCAAGTTGACACACACGGA-3'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 12. Locations of primers used to generate the *pUNK-54(R404Q)* clone *in vitro*. JC24 and JC26 are primers which flank the R404 site. JC25 and JC27 are mutagenic primers containing the appropriate double base substitutions to change the arginine codon (CGT) to one for glutamine (CAA). ———, exons; ——, introns; —, 3' untranslated region.
FIGURE 13. PCR products from the first round of *in vitro* mutagenesis. The right lane shows the two products from the two separate reactions of the first round of symmetric PCR. Amplification with JC24 and JC25 yields a 136 bp product while a 119 bp product is obtained with JC26 and JC27. The left lane contains a 136 bp marker.
Figure 14. PCR product from the second round of *in vitro* mutagenesis. The left lane shows the products of the second round of symmetric PCR. A 227 bp product is generated using the PCR products from the first round reactions as template along with the flanking primers, JC24 and JC26. The right lane shows the products generated in the first round of symmetric PCR. The center lane contains a molecular weight marker.
Figure 15. Loss of PmlII restriction site. The double base substitution necessary to change the arginine codon (CGU) to a glutamine codon (CAA) results in the loss of a PmlII restriction site that is present in the wild-type unc-54 sequence.
pUNK-54(R404Q) (Figure 16). As a more definitive test, the R404 region of pUNK-54(R404Q) was amplified by PCR and the product was sequenced in both directions, confirming the presence of the CGT -> CAA double substitution (Figure 17). The PmlII site was used to monitor the state of the unc-54(R404Q) allele during all subsequent genetic manipulations described in this dissertation. When such manipulations were completed, the presence of the double base substitution was confirmed by PCR amplification and nucleotide sequence determination.

**Generation of unc-54(R404Q) mutants via germ-line transformation.**

The pUNK-54(R404Q) clone generated in vitro was introduced into the C. elegans germ line to determine its effect on movement and egg-laying ability. Transformation of C. elegans is accomplished by microinjection of transforming DNA into the gonad of the adult hermaphrodite. The DNA that is injected commonly forms large extrachromosomal arrays composed of tandem copies of the injected DNA (Fire, 1986; Mellow et al., 1991). Transgenic animals were identified via coinjection of a "marker gene," a plasmid containing a clone of a dominant mutant allele of the rol-6 gene (i.e. pRF4). Expression of this allele confers a distinct visible "roller" phenotype, making it easy to identify transformed animals.

In humans the R403Q mutation is dominant. The first set of transformation experiments tested whether unc-54(R404Q) confers a dominant mutant phenotype in C. elegans. To address this, pUNK-54(R404Q) was coinjected with pRF4 into the wild-type C. elegans strain Bristol N2. As a control a group of worms were injected solely with pRF4 for comparison of the two Rol phenotypes. pUNK-54(R404Q)/pRF4 transgenic animals exhibited a phenotype different from that of the uninjected animals and transgenic worms expressing the mutant rol-6 marker gene alone. pUNK-54(R404Q)/pRF4 animals moved slower and had smaller brood sizes, suggesting that R404Q might exert a dominant
FIGURE 16. R404Q mutation results in the loss of a PmlI site. pUNK-54 and pUNK-54(R404Q) were digested with BamHI and PmlI and then separated on a 1% agarose gel. The presence of the PmlI site will digest the 2.8 kb BamHI fragment yielding three restriction fragments of sizes 1604, 1124 and 102 bp (left lane). The introduction of the R404Q mutation will reduce the number of restriction fragments to two due to the loss of the PmlI site. The right lane shows the sizes of the restriction fragments, 1604 and 1226 bp.
FIGURE 17. Sequence of \textit{pUNK-54}(R404Q) and \textit{pUNK-54}. Sequence on the left, from \textit{pUNK-54} shows the arginine codon CGT at the R404 site, whereas the sequence of \textit{pUNK-54}(R404Q) on the right shows a G $\rightarrow$ A transition and a T $\rightarrow$ A transversion resulting in an Arg $\rightarrow$ Gln substitution at amino acid residue 404.
effect in *C. elegans*. However, it was clear that this effect was not a typical *unc-54* mutant phenotype. In the final analysis, these results were inconclusive because of problems inherent to extrachromosomal arrays. These arrays typically contain hundreds of copies of the injected sequence. High copy expression can result in a range of phenotypic effects. For example, high copy expression of the wild-type *unc-54* gene can generate a phenotype similar to an *unc-54* mutant (Fire and Waterston, 1989). These features make it difficult to interpret the phenotype of the transformants.

A second set of injections looked at the question from the opposite angle. Rather than ask whether expression of *pUNK-54(R404Q)* results in a mutant phenotype it was asked if it could rescue an *unc-54* mutant. Rescue to wild-type would demonstrate that *unc-54(R404Q)* encodes a functional myosin heavy chain that does not confer a mutant phenotype in *C. elegans*. To test this, *pUNK-54(R404Q)* was injected into a strain that exhibits a strong *unc-54* mutant phenotype, *unc-54(r293)*. The *r293* mutation is a small deletion; spontaneous reversion of this allele is undetectable. Eight revertant lines were isolated. However, an additional problem with extrachromosomal arrays was encountered. Extrachromosomal arrays are unstable during mitosis and meiosis which gives rise to mosaic expression and unstable transgenic lines. To eliminate this problem, arrays were integrated into genomic DNA. This is accomplished using a method developed by R.K. Herman at the University of Minnesota. Transgenic animals were exposed to γ-radiation (see Materials and Methods) resulting in double-stranded breaks in genomic DNA. In the repair of these breaks the ends of the arrays can be ligated to the newly created ends of genomic DNA resulting in the integration of the array within a chromosome.

For each of the eight revertant lines a stable *unc-54(r293);Is[unc-54(R404Q)]* integrated line was established. These lines exhibited near wild-type movement and egg-laying ability in comparison to the transgenic *pUNK-54* rollers. This is consistent with the interpretation that the R404Q mutation confers a weak mutant phenotype in *C. elegans.*

48
However, this evidence comes from transgenic animals in which \textit{unc-54(R404Q)} is contained in high copy number on tandem arrays of DNA, integrated at random locations in the genome. To eliminate uncertainties due to overexpression from high copy arrays, and chromosomal position effects on expression from integrated arrays, I sought to generate animals homozygous for an \textit{unc-54(R404Q)} mutation at the normal \textit{unc-54} locus.

\textbf{Generation of homozygous \textit{unc-54(R404Q)} mutant via targeted gene replacement.}

To determine unequivocally the phenotype caused by \textit{unc-54(R404Q)}, I introduced this mutation into the \textit{unc-54} gene on chromosome I, using a transposon-based gene replacement strategy. The mutant obtained by this approach provided clear and unambiguous information on the phenotype conferred by \textit{unc-54(R404Q)}. It also provided a stable mutant line for characterizing this phenotype in greater detail. The method employed was developed in \textit{Drosophila} using P elements (Gloor et al., 1991). Subsequently, the feasibility of this method was demonstrated in \textit{C. elegans}, using the transposable element Tc1 (Plasterk and Groenan, 1992). The technique exploits certain properties of transposable elements in the genomes of these organisms. Upon excision, these elements generate double-stranded gaps in the chromosome at the empty site. These gaps are repaired by a template-dependent gap repair process. Usually the sister chromatid or homologous chromosome serves as template. However, homologous sequences introduced by germ-line transformation can serve as "ectopic" templates (Gloor et al., 1991). This provides a way to introduce specific changes into a gene by incorporating them into a cloned copy of the gene in \textit{vitro}, then supplying that altered sequence as a template for double stranded gap repair in \textit{vivo}. Results in both \textit{Drosophila} and \textit{C. elegans} demonstrate that sequence changes up to 2 kb from the excision site can be introduced into
the chromosomal copy of a gene by this method. In summary, this method requires two components:

1. a target allele - a transposon insertion in or near the gene into which the desired mutation will be introduced. The inserted element must be a) active for excision, generating the required double stranded gap, and b) within 1-2 kb of the site to be altered.

2. an ectopic template containing the desired mutation (and the ability to introduce this template into the germ-line of animals containing the target allele).

The target allele used in the gene replacement experiment to introduce the R404Q mutation into the unc-54 gene was r360. This allele contains the transposon Tc1 inserted in unc-54 at a position 438 bp from the R404 site (Eide and Anderson, 1988) (Figure 18). As stated above, to be useful as a target allele the element must be active for excision. Germ-line reversion of unc-54(r360) is undetectable in the genetic background in which it was isolated (<2.3 X 10^-6), suggesting that the inserted element might not be excising. Tc1 activity is known to depend upon genetic background. For example, in C. elegans variety Bergerac, Tc1 transposition is frequent and a significant source of spontaneous mutations (Moerman and Waterston, 1984; Eide and Anderson, 1985; Moerman et al., 1986). In other varieties, such as Bristol and DH424, its activity is undetectable. Excision of Tc1 results in reversion to a wild-type phenotype only when the element is active in the germ line (Moerman et al., 1986; Moerman and Waterston, 1984). Collins et al., (1987) isolated and characterized a mutator designated mut-2(r459) that increases transposition and excision of Tc1, as well as other transposons, in the germ-line of C. elegans. Tc1 activity in the germ-line of variety Bergerac is increased 50-fold in the presence of mut-2(r459). Therefore, if the lack of reversion of r360 in Bergerac reflects a low a frequency of Tc1
Figure 18. Position of *Bam*HI restriction fragment. The 2.8 kb *Bam*HI restriction fragment contained in plasmid subclone TW#27 is shown below the genomic structure of the *unc-54* gene.
excision from *unc-54*, introduction of *mut-2(r459)* may increase the frequency of excision of this Tc1 element to a level that results in detectable reversion.

To test whether *mut-2(r459)* activates excision of the *r360* Tc1 element a *mut-2(r459); unc-54(r360::Tc1)* double mutant was constructed. Homozygous *r459* hermaphrodites were crossed with *r360/+* males. Nonparalyzed hermaphrodites were isolated singly and their progeny (F2) were screened for paralyzed animals (*i.e. r360/r360*). Clonal populations were established from single paralyzed F2 animals. These lines were then screened for reversion to *unc-54*+. Several of the lines reverted. The line with the highest frequency of reversion, 4X10^-4, has been given the strain designation TW333. An indicator of the presence of *mut-2(r459)* in this strain is that it exhibits the Him (*high incidence of males*) phenotype of *mut-2* mutator mutants. This demonstrates that Tc1 is excising from its site of insertion in *unc-54* in the presence of the *mut-2* mutator. Therefore, TW333 possesses a suitable target allele for gene replacement at the R404Q site: a Tc1 element 438 bp from the R404 site that is active for excision.

The **ectopic template** used to achieve targeted gene replacement was a 2.8 kb *BamHI* restriction fragment from *pUNK-54(R404Q)* containing a portion of *unc-54(R404Q)* including the R404Q site and greater than 1 kb of flanking DNA on each side (Figure 19). The *BamHI* fragment was inserted into the plasmid vector pBR322 (Sutcliffe, 1979; Watson, 1988). The resulting clone is designated TW#27. Use of a segment of *unc-54* rather than the entire gene eliminates concerns about expression from the injected DNA.

The gene replacement approach to isolate homozygous *unc-54(R404Q)* mutants was carried out as follows. Briefly, TW#27 was introduced into the gonad of TW333 [*unc-54(r323::Tc1); mut-2(r459)*] hermaphrodites by microinjection. TW333 exhibits a typical *unc-54* mutant phenotype, paralyzed and egg-laying defective. Forty-five injected worms were plated singly and their progeny screened for animals exhibiting improved
Figure 19. Locations of R404 and r360::Tc1 insertion sites in unc-54. (A) r360::Tc1 insertion site in unc-54 shown relative to the protein-coding regions of the gene. The limits of unc-54 shown here are the AUG translational initiation codon and the AAUAAA polyadenylation signal (Karn et al., 1983). ---, exons; ----, introns; ----, 3' untranslated region. (B) DNA sequence of R404Q site and Tc1 insertional junctions shown below the corresponding wild-type DNA sequences. Sequences corresponding to the Tc1 element are indicated (<...Tc1...>). The nucleotide positions in unc-54 are described by Karn et al., (1983). Adapted from Eide and Anderson, 1988.
movement and egg-laying ability. Based on the rescue experiments described above, this is the phenotype expected for an animal in which the R404Q mutation was copied into the \textit{unc-54} gene from the ectopic template provided. It could not be determined how many of the 45 injected worms produced transgenic progeny containing arrays of TW#27. This is because the marker gene typically used to identify transgenic animals generates a "roller" phenotype which is not observable in the paralyzed TW333 background. Therefore, the injected animals were plated singly and the progeny were screened directly for non-paralyzed revertants.

Nine independent revertants arose in the F2 through F4 generations from the 45 populations screened. Each was picked to a separate plate and allowed to self-fertilize to establish homozygous revertant lines. All nine revertants exhibited wild-type or near wild-type (see below) movement and egg-laying ability. Each was analyzed to determine if the chromosomal \textit{unc-54} locus possesses the R404Q mutation. It is also possible that reversion resulted from the normal process of Tc1 excision followed by interruption of the normal homologue-dependent gap repair process. Such events would be expected to leave the R404 site unchanged and leave a wild-type sequence or a small "footprint" at the Tc1 excision site. The R404 region of \textit{unc-54} from each of the nine revertants was amplified by PCR and digested with \textit{PmlII}. \textit{PmlII} failed to digest the PCR product from one of the nine lines identifying this line as a candidate for an \textit{unc-54(R404Q)} mutant line (Figure 20). This PCR product was sequenced in both directions and confirmed the presence of the desired CGT->CAA double substitution (Figure 21). The revertant line will be referred to as R3. It is assumed that the remaining eight revertants arose from excision followed by repair not involving the ectopic template. These lines were not characterized further.

The lines of evidence described above (reversion, \textit{PmlII} digestion pattern, and nucleotide sequence) are all consistent with gene replacement having occurred in R3. However, it remained possible that PCR amplification of copies of the ectopic template
FIGURE 20. Loss of PmlI site in transgenics. The R404 region of the unc-54 gene from the nine revertants, obtained following injection of the 2.8 kb BamHI ectopic template, was amplified by PCR using the flanking primers JC24 and JC26. The 227 bp PCR products were digested with PmlI as a diagnostic test for the alteration of the R404 site. The digests were size separated by gel electrophoresis. Wild-type sequence yields three fragments of size 108, 102 and 17 bp, whereas sequence containing the R404Q mutation upon PmlI digestion produces restriction fragments of 210 and 17 bp. The odd numbered lanes contain undigested PCR fragments. The even numbered lanes contain PmlI digested PCR fragments. All of the even numbered lanes with the exception of lane 2 shows the production of 108 and 102 bp fragments. The template DNA used to generate the PCR product in Lanes one and two was derived from line R3.
remaining in the nucleus generated the product analyzed. To address this concern, the presence of TW#27 sequences was tested for using genomic Southern blot analysis. Genomic DNA from wildtype strain Bristol N2, TW333 [unc-54(r360::Tc1)mut-2(r459)], and R3, as well as the plasmid TW#27 (to serve as an internal control) were digested with BamHI. TW#27 was used as a probe to determine the presence of pBR322 sequence (i.e. backbone of TW#27) that may have persisted in the revertant R3. Figure 22 shows the results of this analysis: no pBR322 sequences were detected in revertant line R3. Hybridization to a 4.4 kb band representing the pBR322 backbone of TW#27 was seen only in the control lanes containing BamHI digested TW#27 diluted 1,000- and 10,000-fold, comparable to genomic concentrations. However, hybridization to the 2.8 kb BamHI restriction fragment is seen in each of the lanes. Since equal amounts of DNA were loaded in each lane (data not shown) and each of the genomic DNA lanes show the same degree of hybridization it can be concluded that R3 does not have additional copy numbers of unc-54 in its genome. Therefore, arrays containing TW#27 are no longer present in R3, confirming that the PmlII digestion and nucleotide sequence did indeed reflect incorporation of unc-54(R404Q) at the unc-54 locus. This demonstrates that the injected ectopic template formed an array which was used to repair the double stranded gap created from the excision of the Tc1 element from the r360 site, followed by the expected loss of the array due to its inherent instability. The hybridization pattern also indicates the loss of the Tc1 element (1.6 kb) in the R3 BamHI restriction fragment due to the presence of a 2.8 kb band, as seen in N2, rather than 4.4 kb, as seen in TW333. Collectively, these results establish that R3 is homozygous for unc-54(R404Q) at the unc-54 chromosome locus.

R404Q does not result in a characteristic unc-54 mutant phenotype.

The phenotype of the original R3 isolate was that of a thick, slow moving, coiled worm, with low fecundity. This phenotype might have resulted from the R404Q mutation.
FIGURE 22. Genomic Southern blot analysis of *unc-54::TcI* and revertants. R3 and 5A are spontaneous revertants of TW333 following transformation. TW333 strain is *unc-54(r360::TcI);mut-2(r459)*. Wild-type strain is *unc-54+*. TW327 contains 2.8 kb *unc-54* BamHI fragment cloned into pBR322. The DNA was digested with BamHI and separated on a 1.0% agarose gel. The radiolabeled probe was TW#27. Sizes in kilobases, are indicated to the right of the blot.
However, it is also typical of the mut-2 mutant background. mut-2 mutants tend to be "sick" due to increased transposon activity and the resulting accumulation of deleterious mutations. To distinguish between these possibilities, R3 was backcrossed six times by wildtype strain Bristol N2 using a marked unc-54 chromosome to follow the FHC locus. The resulting strain, TW390 [unc-54(cj289Gr)], is homozygous for unc-54(R404Q) and is essentially wild-type for both movement and egg-laying ability. By these criteria, unc-54(R404Q) does not confer a mutant phenotype in C. elegans. I characterized this strain in several ways to determine if it exhibited more subtle defects that might result from the unc-54(R404Q) mutation.

Characterization of post-embryonic development time, brood size and life span.

The C. elegans life-cycle consists of four larval stages, designated L1 through L4, and an adult stage. Each is punctuated by a molt. An L1 hermaphrodite possesses 81 somatic muscle cells; that number increases to 135 in the adult. The unc-54 gene product is expressed in all but the 20 pharyngeal muscles. A consequence of such a large number of muscle cells may be that an alteration of their structure or function could have an consequence on development. The effect of the R404Q mutation on post-embryonic development time, the time period between hatching of the L1 larvae and laying of the first egg by an adult, was investigated. Freshly laid eggs were placed on a plate and monitored to record the times of their hatchings. Newly hatched L1 larvae were transferred singly to new plates and monitored to record the time when the first egg was laid. Table 2, row 2, shows that the post-embryonic development time for TW390 is the same as that of wildtype strain N2. Similarly, there is no statistical difference between their brood sizes (Table 2, row 3). A self-fertilized TW390 hermaphrodite produces about 315 progeny compared with 326 from an N2 hermaphrodite. However, there is a marked statistical difference
Table 2. Post embryonic development times, brood size and Life Span determinations for wild-type N2 strain and TW390.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Post-Embryonic Development Time (p=0.55)</th>
<th>Brood Size (p=0.051)</th>
<th>Life Span (p&lt;0.0001)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2</td>
<td>60.50 hours ± 0.29 (n=26)</td>
<td>326.0 ± 5.74 (n=21)</td>
<td>13.4 days ± 0.343 (n=20)</td>
</tr>
<tr>
<td>TW390</td>
<td>60.20 hours ± 0.39 (n=22)</td>
<td>315.0 ± 6.73 (n=22)</td>
<td>21.8 days ± 1.03 (n=17)</td>
</tr>
</tbody>
</table>
(p < 0.0001) between the life spans of R404Q mutants and N2. When grown at 20°C the mean life span for N2 is 13.4 days versus 21.8 days for TW390. The basis of this extended life span was not understood. It should be noted that C. elegans life spans reported in the literature vary from 14 to 22 days (Wood et al., 1988; Kenyon et al., 1993; Johnson and Wood, 1982). To determine the effect of the genetic background on the lifespan, TW390 males were crossed with N2 hermaphrodites and the progeny of F1 hermaphrodites were screened by PCR amplification and PmlI digestion to obtain siblings homozygous for unc-54(cj289Gr) or unc-54+. The life spans of progeny of F2 homozygotes were measured and determined to be 15.5 days on average. These data would seem to indicate that the extended lifespan for the TW390 strain previously recorded was due to a secondary mutation.

**Interaction of unc-54(R404Q) with other muscle mutants.**

A goal of this project was to use "genetic suppression analysis" to identify genes encoding proteins that interact with the R404Q-containing domain of the unc-54 gene product. This approach relied on the R404Q mutation conferring a visible mutant phenotype, which it did not. I conducted further studies in an effort to identify a more subtle phenotype conferred by cj289Gr that I might be able to exploit for suppression analysis.

The ability to suppress unc-22(s12) was investigated. Certain unc-54 mutants with near normal mobility are dominant suppressors of a specific allele of the unc-22 gene, unc-22(s12). unc-22 encodes another muscle protein termed "twitchin," so named for the characteristic body twitch exhibited by unc-22 mutants. Suppression of unc-22(s12) provides a sensitive measure of contractile defects in these weak unc-54 mutants. If cj289Gr suppressed unc-22(s12), the nontwitching double mutant might provide a screen for suppression analysis. The method would entail mutagenizing the non-twitcher double

61
mutants followed by screening for twitchers. The reisolation of twitchers would be due to the suppression of the R404Q mutation by a defect introduced into an interacting protein. To test for an interaction between unc-54(R404Q) and unc-22(s12), I constructed the unc-54(cj289Gr);unc-22(s12) double mutant. No suppression of twitching was observed. This precludes the use of this system for suppression analysis. Further, it suggests that the R404Q mutation does not even result in a slight contractile defect in the body-wall muscle.

The ability of cj289Gr to suppress unc-105(n490) was also investigated. The n490 mutation results in severe hypercontraction of the body-wall muscles resulting in a hypercontracted worm with a body length that is approximately one-fourth that of wild-type. The hypercontracted phenotype is suppressed by mutations in several muscle component-encoding genes, including unc-54. If unc-54(cj289Gr) suppresses unc-105(n490), it would provide a sensitive measure of contractile defects in these weak unc-54 mutants, and the nonhypercontracted double mutant would provide an attractive screen for suppression analysis (i.e. screen for suppressors of the R404Q mutation by reisolating hypercontracted worms). To test for an interaction between unc-105(n490) and unc-54(R404Q) I constructed the unc-54(cj289Gr);unc-105(n490) double mutant. The lack of suppression of the hypercontracted phenotype precluded the use of this system for suppression analysis and provided added evidence that unc-54(cj289Gr) does not result in a contractile defect.

**Polarized light microscopy of R404Q mutants.**

Polarized light microscopy is useful in the analysis of muscle structure in living animals. This provides a means to analyze mutations which affect this structure from those affecting movement. Comparison between wild-type and unc-54(R404Q) body-wall muscle shows no difference in ultrastructure (Figure 23). The bright A bands, which correspond to the regions containing thick filaments, dark I bands which represent areas of
FIGURE 23. Polarized light microscopy of *C. elegans* body-wall muscles. The birefringent (light) bands are A bands containing thick as well as thin filaments. The non-birefringent (dark) bands are the I bands containing thin filaments only. The refractile bodies are the dense or Z bodies to which the thin filaments are attached. (Micrographs courtesy of Ken Norman, University of British Columbia at Vancouver.)
thin filaments only, and the dense bodies therein are comparable in both wild-type and mutant muscle.
DISCUSSION

Familial hypertrophic cardiomyopathy is a phenotypically heterogeneous disease linked with regions on human chromosomes 1, 11, 14, 15, 16 and 18 (Marian and Roberts, 1994). It has been six years since the elucidation of the first locus involved in this disorder, MYH7, the β-cardiac myosin heavy chain (MHC) gene, and the identification of a missense mutation (R403Q) which cosegregates with inheritance of the disease in certain kindred (Geisterfer-Lowrance et al., 1990). However, how this mutation (and other mutations subsequently identified within this and other "FHC genes") results in FHC remains unclear. The goal of the studies described in this dissertation was to elucidate the molecular basis of FHC caused by the R403Q mutation. To do this the soil nematode Caenorhabditis elegans was used as an animal model for this human heart disorder. The plan involved (1) introducing the corresponding mutation into a C. elegans MHC gene and (2) using the C. elegans mutant to investigate the molecular basis of this disorder, and the function of a conserved arginine residue in myosin heavy chain function. To accomplish the first goal a novel gene replacement technique was used that exploits certain features of the C. elegans transposon Tc1. Using this approach the mutation corresponding to R403Q was introduced into the C. elegans MHC gene unc-54. This successful alteration of the wild-type unc-54 gene at the normal chromosomal locus represents the major accomplishment of this dissertation. However, the "FHC mutant" generated by this approach did not exhibit any apparent mutant phenotype. This precluded further genetic analysis such as a screen for suppressors of the FHC mutant. More detailed characterization of the R404Q mutant has yielded information which highlights differences between nematode and vertebrate striated muscle which may be key factors in the pathogenesis of the disease.

65
Reverse Genetics in *C. elegans*.

Scientists have resorted to animal models to investigate problems related to human disease for many decades. Many factors converge to make this approach even more important now and in years to come. First, there is a greater understanding of common genes, pathways and molecules among various organisms. Second, from the number of genome projects currently underway numerous homologues to human disease loci have been found and the number will increase. Study of the function of these genes and the effect of certain mutations on their function can provide valuable insight into the molecular basis of many disorders. Certain invertebrates have emerged as model organisms for study of basic genetic, molecular and biochemical processes shared by all eukaryotes. Discovery of human disease gene homologues in these model organisms opens the door to their use to understand the molecular basis of many human disorders. Current examples are the study in *C. elegans* of amyloid precursor proteins involved in Alzheimer's disease (Daigle and Li, 1993), and in *D. melanogaster*, the study of copper/zinc superoxide dismutase, the enzyme involved in Lou Gehrig's disease (Philips et al., 1995). Collectively, the above events suggest the use of *C. elegans* and *Drosophila* as models where appropriate (i.e. same gene, same pathway, same anatomy/physiology). It is important to develop a tool for introducing desired mutations into genomes of experimental animals to create animal models. The method requires the ability to target specific changes to the nucleotide sequence to mimic diseased human genetic loci (e.g. Alzheimer's and Lou Gehrig's diseases and cystic fibrosis). The construction of animal models of specific human diseases is important for a number of reasons. Primarily, animal models of human genetic diseases allow for experimentation to investigate the molecular basis of the disorder which is most often not feasible with human subjects. Models also allow for the study of the development and progression of the disease. Slowing or hampering progression is often the direct goal of treatment and this requires knowledge of the biochemistry occurring during successive
phases of the disorder. Experimental organisms with short generation times provide an opportunity to study the time course of a disorder within a reasonable time frame. The testing of hypotheses regarding the pathogenesis of a disorder often times requires obtaining diseased and non-diseased tissue for analysis and this may not be practical or feasible from an afflicted individual. Organisms serving as models are a readily accessible source of material which can often be obtained in abundance. Animal models also provide the ability to determine the effect of genotype on phenotype. The genetic background in which a specific gene defect resides may have an effect on the severity of the disorder. Animal models can be manipulated to investigate the effect of a specific mutation in a variety of backgrounds.

The current ability to obtain transgenic animals with relative ease enables one to introduce mutations which correspond to those occurring in the human population, as well as those which would lead to partial or complete inactivation of the gene. Therefore, reverse genetics can be of use in the study of the effect of specific changes to a gene in vivo to obtain greater knowledge of a gene's function. An example of this would be the knockout of the c-src gene in mice to directly test its role. The expected result was a dramatic effect on phenotype if not lethality due to its proposed role in the cell cycle. Surprisingly, the mice exhibited only a mild bone defect due to aberrant osteoclast function (Soriano et al., 1991). The lack of a drastic affect from the c-src mutation seems to indicate that its role is not indispensable in the cell as had been thought. In addition, what can also be learned by reverse genetics is the interplay between structure and function. The precise mimicry of a human disorder in an animal may or may not reveal a similar phenotype. No matter what the outcome, information regarding similarities and differences between the species can be gained. In particular, the insight could be in the developmental and biochemical differences of the organisms. Novel therapeutic approaches to human diseases

67
may be developed as a result of the understanding of species specific variations. Lastly, animal models enable the testing of various treatment strategies prior to human trials.

Traditionally, human disorders have been mimicked and studied using mammalian models. Experimental organisms of choice have been the mouse, chimpanzee, rat and hamster. However, the reasons given above suggest in many cases it would be appropriate to use *C. elegans* as a model. One attractive feature of *C. elegans* is its small genome. The haploid genome size is \(1 \times 10^8\) bp, eight times that of *Saccharomyces*, about one-half that of *Drosophila*, and almost one-thirtieth that of humans. *C. elegans* genetic and physical maps are detailed and extensive (Epstein and Shakes, 1995). The cloning and sequencing of the *C. elegans* genome has and will continue to identify many genes of interest. Some of those genes will have human homologues which have been identified as loci for various disorders. Reverse genetics approaches that would make it possible to introduce specific alterations into such genes would provide a very powerful tool for addressing questions of function. In the mouse targeted alteration is possible through homologous recombination. The lack of an ability to introduce desired alterations into the *C. elegans* genome by a homologous recombination strategy necessitates the development of alternative approaches. Traditionally, mutational analysis in *C. elegans* has been limited to the reintroduction, by microinjection, of mutated clones into the gonad of the worm. The injected sequences commonly forms large, multicopy arrays. There are problems inherent to expression from an extrachromosomal array. One such problem is the instability of the arrays during mitosis and meiosis which gives rise to mosaic expression and unstable transgenic lines. Further complicating matters, high copy expression of a wild-type gene can result in a mutant phenotype (Fire and Waterston, 1989). However, the introduction of an exact change at a given locus can be accomplished using the Tc1-mediated gene replacement approach described in the Materials and Methods and Results sections. For the work of this dissertation, the introduction of the desired base substitutions into the *unc-54* gene on
chromosome I to change the arginine codon (CGT) to a glutamine codon (CAA) took advantage of the mechanism for the repair of a double stranded gap following excision of the transposable element Tc1 from a nearby insertion site. The repair is template dependent and most often uses the sister chromosome or chromatid. However, an ectopic template supplied through germ-line transformation can be used. This technique for the targeted alteration of the *C. elegans* genome has the potential for use in the generation of gene fusions, insertions, deletions and replacements. The goal of these forms of alteration are the elucidation of structure, function and location of gene expression and or activity.

In this study 45 worms were injected and plated singly. From the screened progeny 9 revertants were analyzed for the presence of the R404Q mutation within the unc-54 gene. One of those 9 lines possessed the appropriate base changes within the R404 codon to code for glutamine. This study was not able to calculate a frequency for transgene instructed repair due to the inability to determine the number of transgenic progeny from the injected animals. The marker phenotype is undetectable in the unc-54(r360::Tc1) background because the animals are paralyzed. The Rol phenotype can be scored only following a reversion event resulting in the animal's ability to move. However, it should be noted that the desired result was obtained with what would be considered a low number of injected animals. The reported frequency of approximately 10^{-5} [i.e. 1 in 100,000 transgenic animals reverted and was demonstrated to have undergone gene replacement (Plasterk and Groenan, 1992)] is high enough to make it a feasible approach for genome alteration.

There are several hurdles that must be overcome in using this technique. First, its use relies upon obtaining a Tc1 insertion within approximately 2 kb of the desired alteration site. The unc-54(r360::Tc1) site used in this work was 438 bp downstream of the R404 site. The PCR based sib-selection method developed by Rushforth and Anderson (1993) can be an aid in the isolation of a usable element if one is not already available. This
method is a PCR-based approach to isolate transposon insertions without relying on the detection of a phenotype. Therefore, silent insertions can be isolated as well as those resulting in a mutant phenotype. Silent insertions are ideal for the introduction of a knock-out or loss of function mutation. Those insertions resulting in a mutant phenotype are suitable for the introduction of subtle/altered or gain of function mutations. A mutant library of Tc1 alleles established using this protocol provides an available pool of potential target alleles for use in gene replacement (Epstein and Shakes, 1995). Another requirement of the gene replacement technique is that the Tc1 element must be actively excising from its site of insertion. The mut-2 mutator may have to be introduced to either activate or increase the frequency of excision of the element. Germ-line reversion is undetectable for unc-54(r360::Tc1) in the genetic background in which it was isolated (<2.3 X 10^{-6}), suggesting that the inserted element might not be excising (Eide and Anderson, 1988). A possibility for the element's inactivity may be that it is crippled. This may be due to a sequence alteration within the element or adjacent to the insertion site acquired during its transposition or insertion into its current resident site, thereby disabling it and preventing its further movement. Another possibility may be that element excision is at a frequency too low for detection. The introduction of the mut-2 mutator background enhanced r360 reversion to wild type (4 X 10^{-4}) demonstrating that Tc1 is excising from its site of insertion in unc-54. Also to be considered, if possible, is the expected phenotype arising from the desired alteration. This can be useful in the final step of isolating an animal with a targeted alteration. The expected phenotype can be discerned by injection of a mutated intact clone into wild-type and mutant animals and screening for a visible mutant phenotype or rescue to wild type, respectively. If from this or other experiments it is determined that expression of the mutant gene will have little to no effect then it would be best to inject a loss-of-function Tc1 allele and screen for rescue. Conversely, if the expected phenotype is other than wild-type then a silent Tc1 allele should be used and the progeny of the injected
animals screened for mutants. Again, the sib selection PCR strategy would be valuable in the isolation of a silent Tc1 insertion.

Also, several features about the generation of an ectopic template must be considered. This work used a template consisting of only a portion of the \textit{unc-54} gene to eliminate concerns about expression from the injected DNA. The initial work by Gloor \textit{et al.} (1991) using P elements in \textit{Drosophila}, as well as the follow up work by Plasterk and Groenen (1992) using Tc1 in \textit{C. elegans}, demonstrate that the excision site is widened by exonuclease activity resulting in conversion tracts up to 1.4 and 2 kilobases, respectively. However, this may represent an upper limit. A much more workable distance may be much less. Base changes were introduced within 200 bases of the Tc1 site within the \textit{unc-22} gene 76\% of the time a transgene was used, and the efforts described in these studies represent the successful replacement of two bases 440 bases from the Tc1 insertion site.

Therefore, a suitable template is one which extends one to two kilobases on either side of the Tc1 site. This template is coinjected with a plasmid containing a gene that possesses a mutation that when expressed will provide an easily detected marker phenotype. The marker phenotype can be relied upon for the presence of the ectopic template since the extrachromosomal arrays formed following injection of the DNA consists of tandem arrays of the plasmids in a ratio approximate to the ratio injected.

The most difficult phase of this method is the detection of an animal with a targeted alteration. It may be safe to assume that most often the element employed for the method will result in a mutant phenotype and the desired alteration will also result in a mutant phenotype. The degree of similarity between the two phenotypes may play a role in the isolation of a sequence altered animal. The PCR based sib-selection may be the best means to detect a potentially gene replaced animal. In addition to the absence of a PCR product using one primer of a pair flanking the insertion site and the other within the inverted repeat of the element the detection of a PCR product using primers flanking the insertion site
could be used for further verification. It must be kept in mind that the second PCR reaction will produce a product from the extrachromosomal array as well as the chromosomal locus. The presence of the desired alteration could then be confirmed depending upon its nature. An insertion or deletion if large enough may be detectable based on a difference in PCR product size. Alternatively, base changes may introduce or destroy a restriction site. However, the task may be as simple as screening the transgenic animals for a visible phenotype that had been previously determined (e.g. from the prior injection experiments) or is expected (e.g. loss of a silent Tc1 insertion and gain of a missense, insertion, deletion or out of frame mutation in a gene with a known mutant phenotype). Potential animals can then be analyzed by PCR as described above. Once a sequence-altered hermaphrodite is obtained it is placed on a plate and allowed to propagate and progeny are selected that have lost the extrachromosomal array as evidenced by the loss of the marker phenotype. This can then be verified by Southern analysis. The region altered can then be PCR amplified (this time without concern of amplification from the extrachromosomal array) followed by direct sequencing of the product.

The final concern is the effect of the genetic background of the animal. The usefulness of the technique relies upon the excision of the transposable element and in particular within the germ-line of the animal. It may be necessary to introduce the mut-2 mutator to activate the element. The consequence is that mut-2 mutants tend to be "sick" due to increased transposable element activity and the resultant accumulation of deleterious mutations. Therefore, prior to analysis of the alterations the mutant should be backcrossed by a wild type strain a number of times to cross out the mut-2 background effects.

The unc-54(R404Q) mutant strain was created using the novel gene replacement technique in order to investigate the molecular basis of FHC caused by the R403Q mutation in human β-cardiac MHC. Characterization of the unc-54 FHC mutant strain revealed no apparent defects when compared to wild-type, based on several criteria. First, animals
homozygous for the R404Q mutation are indistinguishable from wild-type in their movement, egg-laying ability, broodsize, developmental timing and ability of males to mate. Second, the FHC mutation failed to act like other unc-54 mutations in that it did not relieve the unc-105(n490) hypercontracted phenotype or suppress the unc-22(sI2) twitcher phenotype. Third, polarized light microscopy failed to detect abnormalities in the overall organization of the myofilament lattice in the body-wall muscle of R404Q mutants. One difference in phenotype is that the R404Q mutant's life span is one-and-a-half times that of wild type. The apparent cause of this is not the R404Q mutation but a secondary mutation. Siblings, either homozygous for the wild-type unc-54 locus or homozygous for the R404Q mutation, that were isolated following the cross of TW390 males with N2 hermaphrodites have lifespans comparable to that of wild-type N2 strain.

**Possibilities for the lack of an effect of the R404Q mutation.**

Studies that have been conducted on the FHC mutations in North American families reveal that the R403Q mutation is associated with early onset of symptoms, high penetrance and a high incidence of premature sudden cardiac death (Durand et al., 1995). How can this pronounced effect on cardiac muscle function be reconciled with the apparent lack of effect of the corresponding R404Q mutation on MHC function in *C. elegans* body-wall muscle? One possibility is that the role of the R403 residue (and surrounding domain) in MHC function is important in human heart muscle but not nematode body-wall muscle. This would seem unlikely due to the fact that the arginine is invariant over a wide phylogeny. This would suggest its having a fundamental role in MHC function. A second is that R404Q may have a minor affect on MHC function, the consequence of which is detectable only after a certain period of time. In essence, a twenty day old worm may not be comparable to a twenty year old heart.
Perhaps only significant disruptions in MHC function result in the typical *unc-54* mutant phenotype while only minor disruptions can cause FHC. MHC defects strong enough to result in paralyzed worms would be too severe to allow survival long enough to be represented as FHC alleles in humans. Conversely, MHC defects mild enough to be represented as FHC alleles would be too weak to have phenotypic consequences in *C. elegans*. This could be tested by introducing mutations into *unc-54* that correspond to any of a number of the more than 30 mutations within the β-cardiac MHC that cosegregate with FHC. The limiting factor would be the availability of a Tc1 insertion in close proximity to the site destined for alteration. The lack of a preexisting Tc1 element in the desired region can be overcome by the use of sib-selection to obtain an insertion of interest. Since the R404Q mutation mimics one of the strongest FHC mutations and it lacks a distinctive phenotype it is predicted that all other corresponding mutations would be silent in the worm.

Another possibility is that this apparent discrepancy may reflect intrinsic differences in myofilaments of *C. elegans* versus vertebrates. The fibers in body-wall muscle of *C. elegans* run obliquely to the long axis of the muscle cell ("penate" organization) while the fibers in cardiac muscle run parallel to the long axis ("parallel" organization). Pennate-fiber muscle is more powerful than parallel muscle of the same weight because more fibers are able to work in parallel. Consequently, *C. elegans* muscle may have a built in mechanism to compensate for a decrease in force production in the muscle of R404Q mutants. At the outset of this work the evolutionary conservation and association with FHC indicate that R403 and the surrounding amino acids are critical for β-cardiac MHC function and MHC function in general. However, the function of this domain remained unknown. In the course of this study insight has been gained by recent studies. The recent elucidation of the crystal structure of the myosin head and actin-myosin complex have revealed that this residue is within a loop that is positioned to interact with several residues of actin (Rayment
et al., 1993a; Rayment et al., 1993b). Also an in vitro motility assay that allows for the monitoring of the translocation of fluorescently labeled actin filaments by myosin bound to a coverslip has demonstrated that human β-cardiac MHC containing the R403 mutation translocated rabbit skeletal actin filaments at a rate approximately five times slower than that of wild-type MHC. The crystallographic data and the in vitro motility studies both suggest that the R403-containing domain has a role in the interaction of myosin with actin. The role of this residue in interactions with actin could be tested using in vitro motility assays to measure the rate of translocation of actin filaments by wild-type and unc-54(R404Q)-encoded MHCs.

The R404Q mutation may result in an altered and thereby less efficient interaction between myosin and actin, as the above crystallographic and in vitro motility results in addition to the hypertrophy of cardiac muscle would seem to indicate. Elizabeth Bucher's lab at the University of Pennsylvania is developing a method for the determination of force generation by a single muscle cell. This system would allow for the detection of a difference in force produced upon contraction between wild-type and FHC mutant muscle.

The above evidence may also lead one to the possibility that the lack of a phenotype is due to a species difference in the actin thin filaments. However, this is unlikely due to the extreme conservation in sequence between the human and C. elegans actins. Among the 377 residues of actin only 19 are variant and based on the crystallographic data, none of those residues are in the region on the actin surface that interacts with myosin.

Another possibility has to do with the hypothesis that the increase in muscle mass is the result of a stimulus such as pressure or volume overload in the myocardium. In other words, the hypertrophy is secondary and arises due to the attempt to compensate for the impaired contractile function of the muscle brought about by the primary genetic defect. It may be that C. elegans body-wall muscle does not undergo stress similar to that experienced by human heart tissue. As a consequence the body-wall muscle does not
become hypertrophied. A final possibility may be that the mutation is not in the proper genetic background conducive to the formation of the disease. Familial hypertrophic cardiomyopathy, previously thought to be due to a single gene defect in affected individuals, may in fact require the expression of other genes that have a role in the ultimate phenotype.

**Potential Investigations.**

There are a number of avenues available to this study for future investigation. "Enhancer analysis" may identify genes encoding proteins that interact with the R404 domain of the *unc-54* gene product. This genetic approach entails mutagenizing *unc-54(R404Q)* mutants (no visible defects) and screening their progeny for worms with impaired movement. Also mutagenized would be the *unc-54(R404Q); unc-105(n490)* double mutant with their progeny being screened for non hypercontracted worms. This will identify mutations that confer such a phenotype in combination with R404Q, suggesting an interaction between the two encoded proteins. This approach might identify actin genes (see below) or genes encoding other muscle components.

The two-hybrid system (Chien *et al.*, 1991) is a method that detects proteins that are capable of interacting with a known protein. In addition it provides the immediate availability of the cDNA clone of the gene encoding that protein. One "hybrid" (recombinant DNA molecule) contains DNA encoding the protein domain of interest fused to DNA encoding the DNA binding domain of the Lex A protein. The second "hybrid" consists of a cDNA library fused to DNA encoding the transactivation domain of the Gal-4 protein. A reporter plasmid possessing the Lex A binding sequence upstream of the *lac-Z* gene is used to detect interactions between the two "hybrid" proteins. If an interaction occurs between the protein domain of interest and a protein encoded by cDNA in the library the Gal-4 activation domain will activate *lac-Z* expression.
Wild-type and *unc-54(R404Q)* fragments could be fused to the Lex A binding site while the second hybrid would possess cDNA fragments from *C. elegans*, human and murine libraries. Clones which demonstrate an ability to interact with the *unc-54* wild-type and mutant fragments could then be analyzed to determine their identity. This would entail sequencing those clones and searching the available sequence databases to identify homology with known proteins. In addition to this study yielding information regarding the role of the R404 domain regarding its interacting with other muscle proteins, it would provide direct information concerning the identity of such interacting proteins.

This work set out to generate an animal model for FHC by the introduction of the R404Q mutation into the predominant MHC of the body-wall muscle of *C. elegans*. The justification for this approach is based on the conserved nature of MHC and muscle structure and function. While no effect of the R404Q mutation was detected details of a method for targeted alteration of the genome were elucidated. The gene replacement technique is a valuable tool in the study of a gene's function. Its use will increase with the progression of the genome project to investigate the function of genes for which nothing is known except its sequence. Therefore, caution must be taken when investigating the possibility of the use of an experimental organism as a model for a human disease due to the presence of subtle or intrinsic differences between the organism and the human.
LITERATURE CITED


78


80


