Identification of the Sea Urchin Egg Myosin Binding Protein Gene

by

Laura R. Shea, B.S.

Submitted in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in the

Biology

Program

SCHOOL OF GRADUATE STUDIES

YOUNGSTOWN STATE UNIVERSITY

AUGUST 1999

Identification of the Sea Urchin Egg Myosin Binding Protein Gene

by

Laura R. Shea, B.S.

I hereby release this thesis to the public. I understand this thesis will be housed at the Circulation Desk of the University library and will be available for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature: Laura R. Shea 8/13/99
Student Date

Approvals: Gary R. Walker 8/13/99
Thesis Advisor Date

8/13/99
Committee Member Date

8/13/99
Committee Member Date

8/16/99
Dean of Graduate Studies Date

Abstract

A novel myosin binding protein designated 53K (p53EMBP) has been identified in unfertilized sea urchin eggs. (Yabkowitz and Burgess, 1987). Antibody against p53EMBP was used to select recombinant cDNAs from a sea urchin oocyte library. An approximately 1,000 base pair sequence was obtained, a very small fragment of a much larger gene. Based on Southern blot analysis it was found that this gene codes for a very large mRNA. This indicated that the gene was much larger than the approximately 1,000 base pair sequence that had been obtained thus far. In addition, some data suggest that p53EMBP is a fragment of a much larger protein (Yabkowitz and Burgess, 1987). The goal of this study is to obtain the entire EMBP gene from sea urchin genomic DNA. Sea urchin genomic DNA was isolated from sperm of sea urchin species Stongylocentrotus purpuratus. Following the isolation of genomic DNA, a restriction digest was done to create appropriate ends in the DNA obtained. DNA agarose gel electrophoresis was performed and the proper size fragments were obtained by electroelution of the gel. Vector DNA (λ DNA) also prepared by a restriction digest was ligated to the insert genomic DNA. Next, the ligated DNA were packaged into phage particles *in vitro* and then were used to infect a E.Coli culture. The bacteria containing the genomic library were grown on bacterial lawns in the hopes that plaques would be formed. In continuing work on this project p53EMBP cDNA probes will be used to screen the library. This will enabled us to obtain the gene for the sea urchin myosin binding protein designated p53EMBP.

Acknowledgments

I would first like to thank with utmost appreciation my wonderful thesis advisor Dr. Gary Walker. Without all of Dr. Walker's time, patience, understanding, and help, the work included here in this thesis would not have been possible. I would like to thank Dr. David K. Asch who has really been like a co-advisor on this project for all of his time, advise, and help on this project. I would like to thank Dr. Paul C. Peterson for taking time out of his busy schedule as a department chair to serve on my thesis committee. I would like to thank Ms. Renee' A. Pitts for all of the valuable information she gained for this project by searching protein databases on the World Wide Web. I would like to thank my fellow graduate students in the Walker lab, Mr. Nader Atway, Ms. Staci Raab, and Mr. Thomas Watkins for all of their support and advise. My parents, William T. Shea, and Inez P. Haddock, and my sister, Kathleen A. Shea I would like to thank for all of their love and support. For always supporting my academic pursuits I would like to thank my uncle, Maurice J. Shea. And lastly, I would like to thank Mr. Bryan E. Mosora for all of his unconditional love and support.

Table of Contents

I. Abstract	iii
Acknowledgments	iv
List of Figures	vii
List of Tables	ix
II. Introduction	1
The Sea Urchin Egg Myosin Binding Protein	16
Strongylocentrotus Myosin Binding Protein	17
III. Materials and Methods	22
Experimental Design	22
Construction of the Sea Urchin Genomic Library	25
DNA Isolation	25
Restriction Digest	28
Density Gradient Centrifugation	30
Gel Purification	34
Ligation, Packaging, and Plating	35
IV. Results	40
V. Discussion	53

Future Work 60

VI. Literature Cited 63

VII. Appendix A 64

List of Figures

Figure 1.	The muscle sarcomere	1
Figure 2.	Myosin I and myosin II	4
Figure 3.	Actin and myosin in non-muscle cells	6
Figure 4.	Similarities in myosin binding proteins	13
Figure 5.	Block diagram of MyBP-C and MyBP-H	15
Figure 6.	DNA Sequence of p53EMBP	18
Figure 7.	Comparison of p53EMBP and VAP-1	20
Figure 8.	Overview of genomic library construction	23
Figure 9.	Selection of the gene for p53EMBP	25
Figure 10.	Serial dilution of packaging reaction	39
Figure 11.	Purified sea urchin sperm DNA	41
Figure 12.	Attempted DNA isolation	43
Figure 13.	DNA isolation containing impurities	44
Figure 14.	Restriction digest with diluted enzyme Mbo I in a 1:10 concentration.	45
Figure 15.	Restriction digest with diluted enzyme Mbo I in a 1:50 concentration	46

Figure 16.	Glycerol Gradient Gel	48
Figure 17.	Gel purification gel	50
Figure 18.	DNA ligation gel	51
Figure 19	Bam HI digested DNA	52
Figure 20.	Test ligation plate with plaques	57
Figure 21.	DNA gel showing inserts	60

List of Tables

Table 1 Selected DNA using UV absorption 42

II. Introduction

The work included in this study shows how the gene for the sea urchin egg myosin binding protein (p53EMBP)is to be cloned from genomic DNA. In 1987 a novel sea urchin myosin binding protein was discovered in the laboratory of David Burgess (Yabkowitz and Burgess, 1987). This protein was given the name 53kDa protein. Furthermore, 53kDa protein was shown to effect the solubility of myosin at low ionic strengths. It was found by Northern blot analysis that this protein first designated as 53kDa protein codes for a very large mRNA (unpublished data). This indicated that this protein was much larger than the approximately 1,000 base pairs that had been isolated.

In order to understand the background for this study it is necessary to first include the structure and function of myosin in the muscle sarcomere and the contractile ring of non-muscle cells.

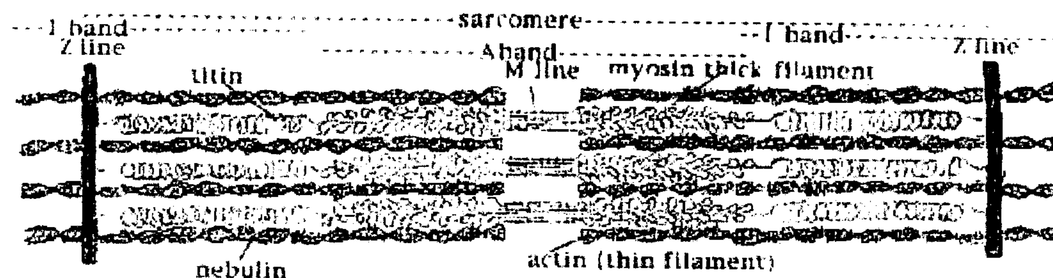


Figure 1. The muscle sarcomere
This figure is redrawn from Alberts fig.16-89 and fig.16-95 pg. 851 and 855.

Myosin is a motor protein that moves along actin filaments and is plentiful in skeletal muscle. In skeletal muscle, it forms a major part of the contractile apparatus (Alberts et al., 1994). Muscle contraction is produced by the interaction of myosin filaments with actin filaments. This occurs by the sliding of actin filaments against myosin filaments. In the presence of ATP, the myosin head regions (which are found at the end of the myosin filaments) become attached to actin filaments. Next, a conformational change occurs in the myosin head region. This conformational change allows the myosin filament to encounter the actin filament and then become unattached. Accessory proteins aid the actin and myosin in accomplishing this task. These proteins hold the actin and the myosin in a parallel overlapping manner. In many non-muscle cells and in smooth muscle cells, the contraction is performed with actin and myosin in much the same way as it is above with skeletal and cardiac muscle. However, in smooth muscle and in non-muscle cells the contractile units are smaller. Furthermore, Ca^{2+} regulated phosphorylation of a myosin light chain control the activity and assembly in non-muscle cells (Alberts et al., 1994).

Because there are more than one type of myosin molecule, it is necessary to distinguish between them. Myosin molecules comprise all of the

actin motor proteins that have been identified so far. When myosin molecules bind to actin filaments they hydrolyze ATP into ADP and P_i. This is how myosin was first identified in skeletal muscle. The myosin found in muscle is called myosin II. The myosin II molecule has two heads and a rod like tail. The head region contains the motor activity as well as the ATPase activity. Myosin II has two heavy chains each of which becomes a pair of light chains at the head region. The function of the myosin II molecule is to move actin filaments past each other. The function of the tail region of the myosin II molecule is to permit polymerization of the molecules into bipolar filaments. Myosin II is also found in the cell cortex as well as the contractile ring of cell division. Furthermore, myosin II is thought to be responsible for the tension in stress fibers and the tension in which to keep the cell surface firm (Alberts et al., 1994). It should also be mentioned that Ca^{2+} dependent phosphorylation of the myosin II molecule increases its interaction with actin as well as enables its assembly into short bipolar filaments (Alberts et al., 1994).

Also found in non-muscle cells is a type of myosin known as myosin I. This type of myosin is smaller than myosin II. It is thought that myosin I is the predecessor from which myosin II originated. Myosin I like myosin II also has a motor head region as well as the ability to hydrolyze ATP (Alberts

et al., 1994). The structure of these two molecules can better be seen in figure two.

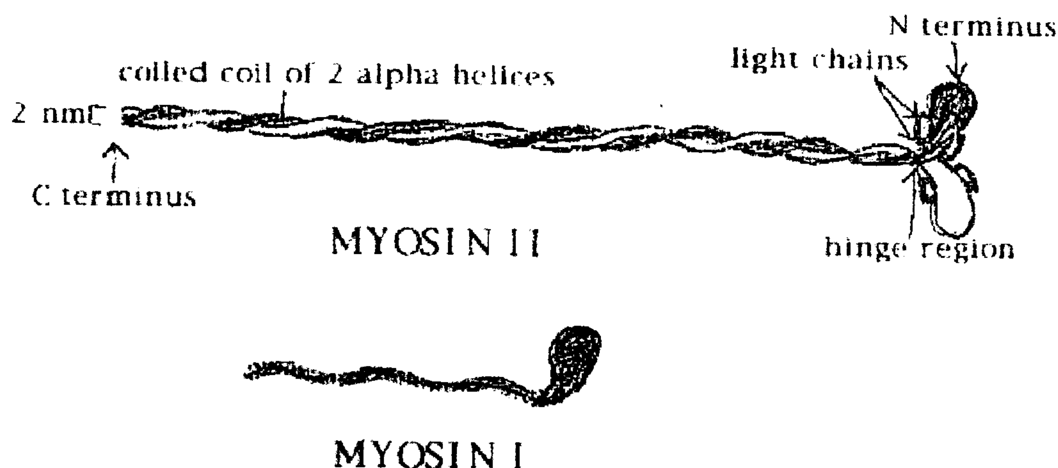


Figure 2. Myosin I and myosin II. This fig. is redrawn from Alberts et al., fig. 16-69 and fig. 16-70. (This figure is not drawn to scale)

It was first thought that the muscle thick filament was bare. However, in the 1970's a entire class of proteins were found to associate with the myosin. This was the beginning of the discovery of myosin-binding proteins (Gautel, 1996). At first these proteins were thought to be contaminants of myosin preparations. Later it was found that these proteins make up almost four percent of the mass of the myofibril (Seiler et al., 1996). Most of the myosin binding proteins identified have been in skeletal and cardiac muscles. The first myosin binding proteins found were C-protein, H-protein, and X-protein. C-protein and H-protein are now known as MyBP-C and MyBP-H respectively. X-protein is now thought to be the slow isoform of C-protein

and not considered a separate myosin binding protein (Gautel, 1996). In conjunction with the thick filaments of muscle sarcomeres, the myosin binding proteins: titin, nebulin, myomesin, 86-kDa protein (MyBP-H), and MyBP-C have been identified (Yabkowitz and Burgess, 1987).

However, myosin also plays a large role in non-muscle cells (Alberts et al., 1994). Myosin binding proteins also exist in non-muscle cells. This is because myosin undergoes changes in distribution and polymerization in cells and it is believed that the myosin binding proteins regulate the "super molecular organization" of myosin in the cytoplasm (Yabkowitz and Burgess, 1987). In the cytoplasm of eukaryotic cells, actin is known to be the most abundant protein. It comprises five to ten percent of the total protein in eukaryotic cells. In addition, myosin makes up about one-fourth of the amount of actin (Voet, 1995). Studies have shown that in some eukaryotic cells actin and myosin II come together for very short periods, accomplish a task, and then disassemble. This has been shown to occur in cell division. In this case, actin and myosin II filaments form a structure called a contractile ring. This contractile ring aids in cytokinesis by appearing during M phase of cell division. The contractile ring pulls on the plasma membrane and helps to contract the middle of the cells thereby forming two daughter cells (Alberts

et al., 1994). After cell division is complete the myosin II molecules separate (Alberts et al.,1994).

Stress fibers in fibroblasts are another example of the existence of actin and myosin in non-muscle cells. Here they are smaller and not as organized as in muscle cells. Stress fibers attach to the plasma membrane at focal contacts and at intermediate filaments that surround the cell nucleus. The stress fibers permit the cells to apply tension to collagen molecules around them (Alberts et al., 1994).

In epithelial cells, actin and myosin also play a role. Actin filaments stretch across the cytoplasm from cell to cell junction. Tension is formed across a "multicellular sheet". Also, in epithelial cells actin and myosin are found in adhesion belts. In adhesion belts they fold epithelial cell sheets in the developing embryo (Alberts et al., 1994). See figure 3 below.

Dividing Cell **Epithelial Cell** **Fibroblast Cell**

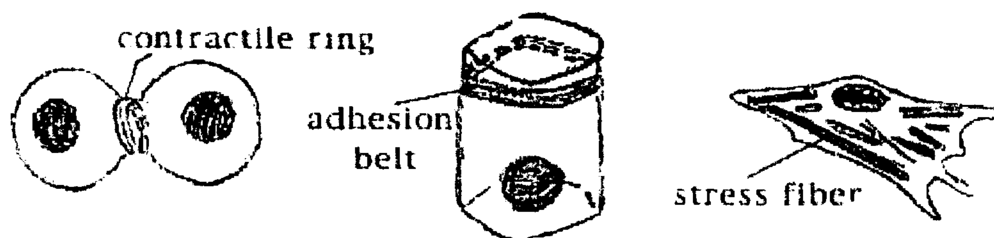


Figure 3. Actin and myosin in non-muscle cells. This figure is redrawn from Alberts et al., fig. 16-72.

General Description of Myosin Binding Proteins

Titin

Titin (which was first known as Connectin) is the largest protein known to date with its molecular weight of 3000 kiloDaltons. Titin is the third most abundant myofibrillar protein. It should be noted that it is third only to actin and myosin (Wang, 1996). In skeletal muscle titin molecules extend from the thick filaments M line to the Z line (see fig.1)(Alberts et al., 1994). This means that titin is over 1μm in length (Higgins et al., 1994). It is believed that titin molecules keep the myosin thick filaments centered (Alberts et al., 1994) and provides the resting tension in the muscle sarcomere (Ayme-Southgate et al., 1991).

Due to the extremely large molecular weight of titin, it took a longer time to discover it. It wasn't until researchers reduced the acrylamide concentration of protein gels from between 6% and 15% to only 2% that titin was discovered. The titin molecule has a rod with a beaded substructure. This is due to the Ig and Fn III domains of the titin molecule (Labeit et al., 1997). It has been shown that between six and twelve titin molecules associate with each muscle thick filament. At the point where titin molecules attach to the thick filament, it is believed to be the place of filament assembly (Eilertsen et al., 1994).

Titin is also believed to exist in non-muscle cells. It is thought to play a role in the organization of the cytoskeletal myosin II filaments (Keller, 1995). It was first discovered in non-muscle cells as a protein called T-protein. T-protein had the same molecular weight, molecular morphology, and immunocrossreactivity as titin. Therefore, it is believed to be the non-muscle cell isoform of titin (Eilertsen et al., 1994).

The brush border of isolated intestinal epithelial cells is an excellent place to characterize cytoskeletal components and this is where the non-muscle cell isoform of titin was discovered. Here, the non-muscle cell isoform of titin is found in a region called the terminal web region. Moreover, it is here at the terminal web region that titin could play a role in the association of myosin II with the cytoskeleton (Eilertsen et al., 1994). In studies by Eilertsen et al., it was found that only one non-muscle cell titin molecule associates with one bipolar filament. This is in contrast to the already noted six to twelve titin molecules that associate per muscle thick filament (Eilertsen et al., 1994).

It is also worthy to note that Eilertsen et al. found that titin-myosin interactions are isoform specific. He noted that non-muscle cell titin would not bind to muscle myosin and muscle titin would not bind to non-muscle cell myosin. This supports the theory that non-muscle cell titin is a factor in

cytoskeletal organization of myosin II bipolar filaments (Eilertsen et al.; 1994).

Titin is not only found in vertebrate muscles and the cytoskeleton but also mini-titins are found in invertebrate muscle. These mini-titins include the C. elegans protein, twitchin, and the Drosophila protein, projectin. Myosin II filaments are found to associate with all of the known titins. However, the role that each of the titins play in muscle structure and function may differ (Keller, 1995).

Twitchin

As already mentioned one of the mini-titins is known as twitchin. Both titin and twitchin contain large numbers of two-conserved amino acid motifs that occur in "regular arrays" (Ayme-Southgate et al., 1991). One of the motifs is similar to the fibronectin type III domains and the other is similar to immunoglobin C2 domains (Ayme-Southgate et al., 1991). The sequences of these two proteins show marked similarities. Not to mention the fact that they both have a protein kinase domain near the C-terminus. Even though there has been remarkable similarity between titin and twitchin sequences, their roles in function appear to be different (Higgins et al., 1994). Animals that lack the gene for twitchin have a constant twitching of the body wall

muscles. This shows that twitchin may be involved in the contraction/relaxation cycle of muscle (Ayme-Southgate et al., 1991).

New members of the Ig superfamily

Because of its Ig domains, twitchin was the first intracellular protein that was considered a member of the Ig superfamily. Now, due to the reoccurring Ig and Fn III domains a whole "family" of intracellular and muscle proteins have added a new branch to the Ig superfamily (Benian et al., 1996). This "family" can be broken down into groups based on their location in the sarcomere and their function. The first of these groups is the already mentioned titin family and consists of twitchin, titin/connectin, and projectin (Benian et al., 1996). The second group includes telokin of smooth muscle, and MLCKs (myosin light chain kinases) of smooth muscle and non-muscle cells. It should be noted that the phosphorylation by MLCKs is necessary in order to initiate contraction in smooth muscle and in non-muscle cells. Moreover, telokin has only one Ig domain. Furthermore, the third group in the new branch of the Ig superfamily consists of C-protein and 86-kDa protein (MyBP-H)(Benian et al., 1996). The proteins of this group can be found at the A-band of vertebrate striated muscle. Specifically, to the inner third of the A-band at transverse stripes of 43 ηm spacing. A fourth group is a single

protein called kettin (Benian et al., 1996). Kettin is not associated with myosin or the thick filament. In addition, the last group includes M-protein, skelemin, and myomesin. All the proteins of the last group are located at the M-line in vertebrate striated muscle (Benian et al., 1996).

Nebulin

Another well-known giant myosin binding protein is nebulin. Nebulin has a molecular weight of 600-900 kDa. There are 150-200 repeating structural domains in a nebulin molecule. Each one of these domains interacts with actin molecules in the thin filament. Nonetheless, nebulin is known to associate with the thin filament only in vertebrate skeletal muscle (Keller, 1995). The nebulin molecule extends from one end of the actin thin filament to the other and during muscle development regulates the assembly of actin and length of the actin thin filament (Alberts et al., 1994). Nebulin is not known to exist in non-muscle cells. Nevertheless, because of the regulation of actin filament length there may be a possibility that nebulin could exist in non-muscle cells. A protein called N-protein was discovered in non-muscle cells that had a comparable molecular weight to that of nebulin. However, unlike the cellular isoform of titin it was later found that it had properties that were different from nebulin (Keller, 1995).

Projectin

Another of the mini-titins is a protein similar in size to twitchin. It is the Drosphila protein known as projectin. It has been identified in the connecting filaments of insect flight muscle (Ayme-Southgate et al., 1991). Partial sequences of projectin shows the same repeat patterns as sequences of twitchin. Therefore, it is thought that projectin is the insect homologue of twitchin. The partial sequences that show the same repeat patterns can be precisely aligned with sequences of twitchin (see fig.4) (Higgins et al., 1994). When analyzed in honeybees, a single molecule of projectin extends from Z line to the thick filament (Benian et al., 1996). In Drosophila, in the A band, the myosin binding protein projectin is located. Moreover, in the Drosophila I band a smaller projectin molecule is located. So, these two isoforms for this protein are both derived from one gene (Benian et al., 1996) by differential RNA splicing.

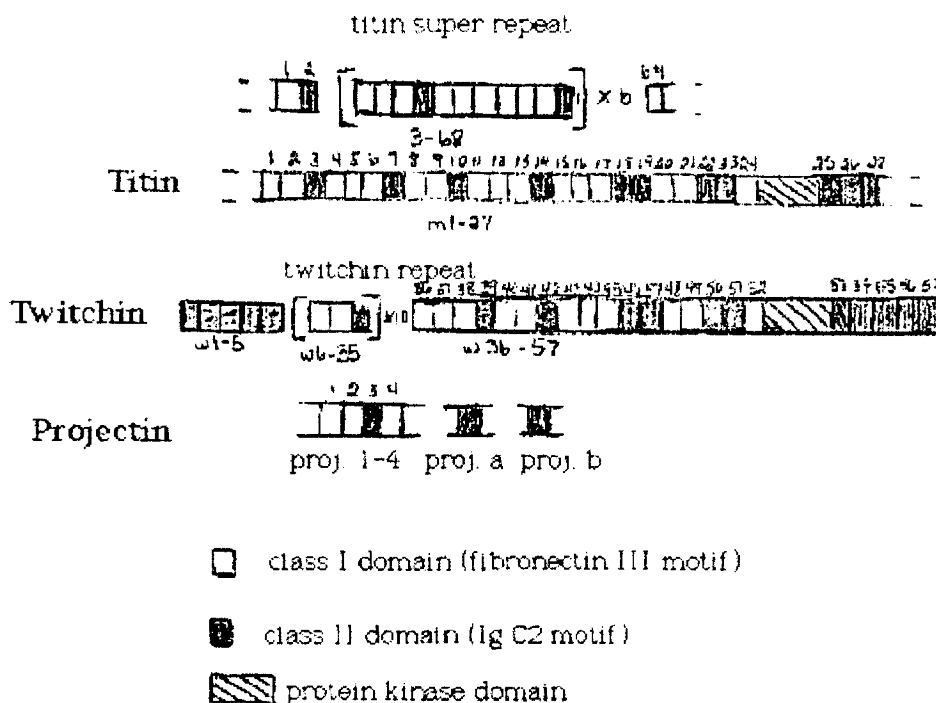


Figure 4. Similarities in myosin-binding proteins. This figure is redrawn from Higgins et al. fig. 1. Layout of Fn II, Ig domains, and kinase domains. The first titin sequence is a series of eleven domain super repeats from the center of the protein. The second lies towards the C-terminus. The titin domains from the super repeat are numbered one through sixty-nine. The class I and class II in the second titin sequence are numbered m1 to m27.

Myomesin

Myomesin anchors titin in the region of the M-band (van der Ven and Furst, 1997). Myomesin has a molecular weight of 185-kDa (Grove et al., 1985). It is thought to play a role in maintenance and assembly of the myofibril (Eppenberger et al., 1981) and in myosin thick filament organization and alignment (Yabkowitz and Burgess, 1987).

Myosin-binding protein-M

Myosin-binding protein M anchors titin in the region of the M-band (van der Ven and Furst, 1997). MyBP-M has a molecular weight of 165-kDa (Grove et al., 1985).

Myosin-binding protein-C

MyBP-C is localized to the A-band of striated muscle. The region in the A-band that it is localized is called the C zone. It is called the C zone due to the fact that MyBP-C is localized there (Vaughan et al., 1993). The C zone is the cross-bridge bearing zone. MyBP-C and MyBP-H have remarkable similarities including structural similarities. These two proteins are highly charged. Ultraviolet spectra and amino-acid composition are very similar in MyBP-C and MyBP-H. Both MyBP-C and MyBP-H are rich in proline residues (Vaughan et al., 1993). However, these two proteins show different distribution in the A-bands.

Depending on the muscle which it is isolated MyBP-C has different numbers of phosphorylation sites. In the isoform found in cardiac muscle there are four phosphorylation sites. Furthermore, these sites are kept under control with help from cAMP and calmodulin regulated kinases. Whereas, in skeletal muscle there is only one phosphorylation site, this could indicate a regulatory function for the phosphorylation sites in cardiac muscle (Weisberg

and Winegrad, 1996). Fast and slow isoforms of MyBP-C have been identified. No isoforms of MyBP-C have been identified in non-muscle or in smooth muscle (Seiler et al., 1996). MyBP-C could play a role in myosin thick filament organization and alignment (Yabkowitz and Burgess, 1987).

Myosin-binding protein-H

Myosin-binding protein-H is the protein formally known as 86 kDa protein. MyBP-H is also localized to the C zone of the A-band of striated muscle. Both fast and slow isoforms have been identified in MyBP-H (Seiler et al., 1996). Not unlike MyBP-C and myomesin, MyBP-H is thought to play a role in myosin thick filament alignment and organization (Yabkowitz and Burgess, 1987).

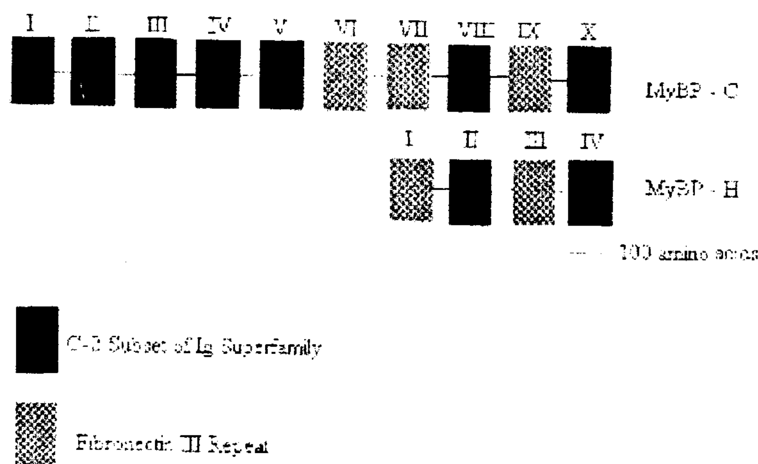


Figure 5. Block diagram representing motif structure of MyBP-H and MyBP-C. This fig. is redrawn from Fig. 5 in Vaughan et al. Note the conserved motif structure of C-terminal of each molecule.

The sea urchin egg myosin binding protein

The 53 kDa myosin binding protein found in sea urchin eggs (Yabkowitz and Burgess, 1987) is different from the skeletal myosin binding proteins in a number of ways. The 53 kDa protein (p53EMBP) binds to myosin in a nucleotide dependent manner. Furthermore, it effects the solubility of myosin at low ionic strengths (Yabkowitz and Burgess, 1987; Walker et al., 1991). This is believed to be because of the area in which the p53EMBP binds to the myosin molecule. The head-rod junction of the myosin molecule is important in myosin structure and function. Hence, any proteins that bind at this region are likely to effect myosin's function (Walker et al., 1991). It was found that the p53EMBP binds to myosin at the head-rod junction region (Walker et al., 1991). Therefore, showing participation in myosin regulation because this is an important area in the regulation of myosin (Walker et al., 1991). This is consistent with p53EMBP effects on the solubility of myosin.

Antibody against p53EMBP was used to select recombinant cDNAs from a sea urchin oocyte cDNA library. The major recombinant 53 kDa protein (p56rEMBP) studied so far contains a number of glycine rich domains (unpublished data). Furthermore, the 53 kDa cDNAs contain a

Casein kinase phosphorylation site (Ck2) and ATP binding motifs (unpublished data). The p56rEMBP also shows protein kinase activity similar to that demonstrated for p53EMBP (unpublished data).

It is believed that the 53kDa obtained thus far is only a fragment of a very high molecular weight protein. Only a fragmentary sequence of this gene has been obtained thus far. This 1,000 base pair sequence is only an extremely small piece of a much larger gene. Northern blot analysis indicates that this gene codes for a very large mRNA (unpublished data). This indicates that the gene is much larger than the approximately 1,000 base pair sequence that has already been obtained. In addition, some data suggest that the 53kDa protein is a fragment of a much larger protein (Yabkowitz and Burgess, 1987). This study is interested in obtaining and preliminary characterization of the entire EMBP gene.

SPMBP: Strongylocentrotus Myosin Binding Protein

Amino acid sequence analysis was done on the world wide web using the DNA sequence obtained for a portion of the myosin binding protein (figure 6). This allowed for the comparison of the EMBP sequence to other known proteins. As shown in figure 7 this sequence is both rich in the amino acid glycine and has repeated sequences (unpublished data).

gaattccgat attgagagac ctgaccttga tgtcagtggg gatgcagacc

ttccatcagg aggagttggc ctggatgttg gaggagggat cggaggcgga

ctcggtggag gactagacat tgatgccaat ggtcctgatg ttgacatcaa

ggggccaaaa gttggaggtg acatctcagg .cccagacct tgatgtgagt

ggacccgatc tggatatcga tgtagatgga aagaaaaagg gaaaaggtgg

atttggattt ggaatgaaaatgcccaaa.. ....ttcggat

ttggaggcca tggcaaaggt gatattgacg tagatgcaga cgttgatatt

gagagacctg accttgatgt cagtggggat gcagaccttc catcaggagg

agttggcctg gatatcggtg gtggagctgg aggtaatatt ggaatt.....

Figure 6. The cDNA sequence from a p53EMBP clone. This was used as the starting point for searching protein database SWISS-PROT

Following the search of the SWISS-PROT database punitive homologs of p53EMBP were identified. A vesicle associated protein (VAP-1) was one of the homologs identified. The VAP-1 protein was cloned from the sea urchin species Strongylocentrotus Purpuratus the same species that p53EMBP was been identified in. The characteristics of the protein known as

VAP-1 include: tissue specificity in the egg cortex, contains four RNA recognition motifs (RNP), the subcellular location is initially a peripheral membrane protein that is associated with the microsomal membrane fractions and that it may be targeted to the nucleus later in development, and it may function as a multidomain RNA-binding protein and play a role in nuclear RNA processing and in early development (Barton et al., 1992).

Furthermore, a protein database was used to obtain the protein sequence of the DNA sequence shown in figure 6. Shown in figure 7 part A is the most likely amino acid sequence of SPMBP. This is further compared to the amino acid sequence for VAP-1 shown in part B of figure 7.

Part A.

NSDIERPDLDVSGDADLPSGGVGLDVGGGIGGGLGGGLDIDANGPDVDIKGP

KVGGDISGPDLDVSGPDLDIDVDGKKKGKGGFGFGMetKMetPKFGFGGHGK

GDIDVDADVDIERPDLDVSGDADLPSGGVGLDIGGGAGGNIGI

Part B.

TDIGGGLDVG GGLRGGLDID AKGPDVDIKG PKVGGDISGP DLDVSGPDLD

IDGGGKKGKG GFGFGLKMPK FGFGGHGKGD IDVDADVDIE

RPDLDVSGDA DLPSGGVGLD VGGGIGGGLG GGLDIDANGP DVDIKGPKVG

GDISGPDLDV SGPDLDIDVD GKKKGKGGFG FGMKMPKFGF

GGHGKGDIDV DADVDIERPD LDVSGDADLP SGGVGLDVGG GIGGGLGGGL

DIDANGPDVD IKGPKVGGDI SGPDLDVSGP DLDIDVDGKK KGKGGFGFGL

KIPKFMDPTF GFGGHGKGDI DVDADGGVVI PEGDIKVKTG KPDIGGDVDL

PSGGVDLDVG GGIGGGLGGG LDIDAKGPDV DIKGPRVGGD ISGPDLDVSG

PDLDIDGDGK KKGKGGFGFG LKMPKFGFGG HGKGDIDVDA

DVDIERPDLN VSG

Figure 7. Comparison of EMBP and VAP-1. Comparison of the amino acid sequence thought to be the most likely representation of EMBP (Shown in Part A) and the known sequence for the protein known as VAP-1 (Shown in Part B). The sequences that are the same in each of the protein sequences are underlined.

Specific Aim of this study

Once the entire gene is obtained for the EMBP then it can be confirmed that p53EMBP is homologous to the protein known as VAP-1. However, this is unable to be known for certain until the entire gene for p53EMBP is obtained. This study has two primary aims. The first one is to

construct a genomic DNA library from sea urchin sperm DNA. The second one is to identify and isolate the entire gene for p53EMBP.

III. Materials and Methods

Experimental Design

The first specific aim of this project was to construct a genomic DNA library from sea urchin sperm DNA (this can be seen in figure 8). This was to be done by first isolating genomic DNA from sea urchin sperm. A restriction digest was then performed on the genomic DNA using a enzyme in order to cut the DNA into certain size pieces. The isolated genomic DNA fragments that are cut by the restriction digest needs to be incorporated into a vector. In this project a bacteriophage lambda vector was used. The bacteriophage lambda vector also has to be cut with the same restriction enzyme as the isolated DNA in order to enable the sticky ends to be ligated together. This creates left and right arms of the lambda vector . The central portion of the lambda vector is not needed and can be discarded. The arms and restriction digested DNA are then ligated together. Once the ends are ligated together the ligated DNA can then be packaged into phage heads. This creates a library of genomic DNA within the phages. The phage heads are then used to infect a E.Coli culture. The E.Coli is then plated out on agar plates. The E.Coli will then grow a

confluent bacterial lawn on the agar. If the ligated DNA has been incorporated into the E.Coli, plaques will be obtained. The presence of a phage plaque on a bacterial lawn indicates a recombinant phage bearing a insert. Once plaques are obtained the library can be screened using a DNA probe of the gene of interest. In this case the gene of interest is p53EMBP (this can be better seen in figure nine).

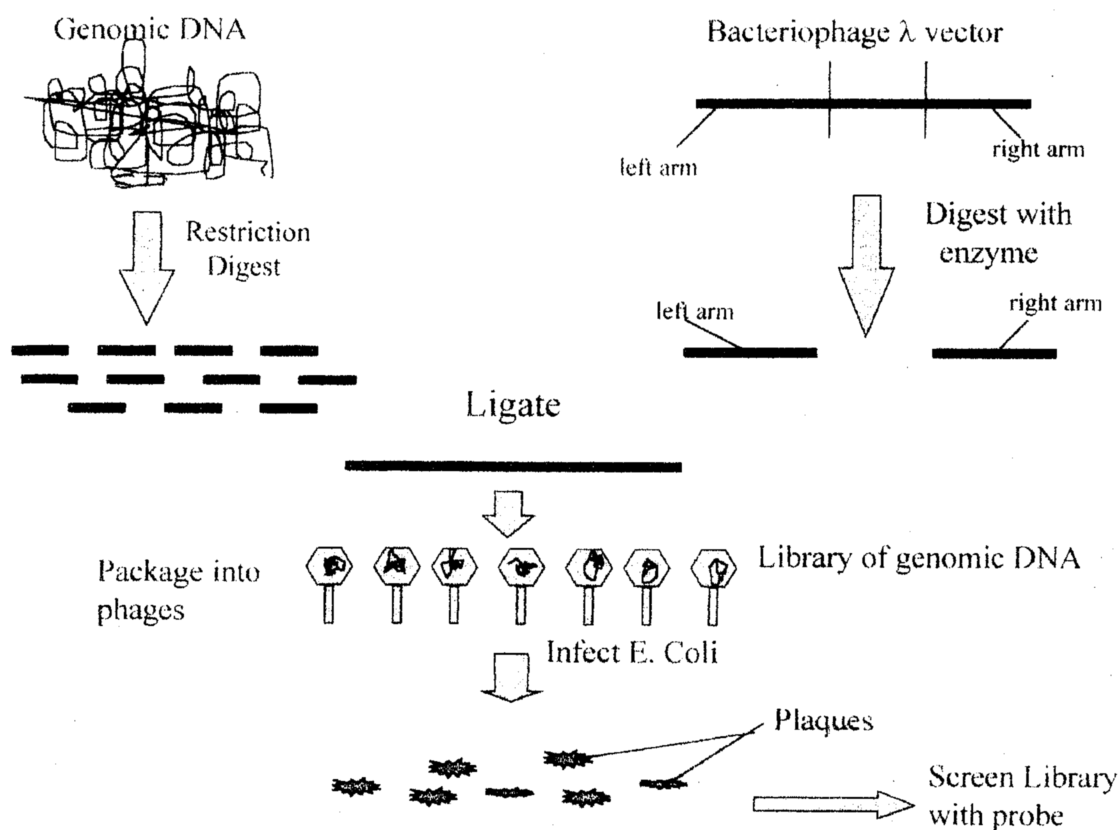


Figure 8. Overview of genomic library construction

Once plaques have been obtained, the absorption of the phage DNA to nitrocellulose will occur by placing the nitrocellulose on the surface

of the plates containing the plaques. The phage protein is then dissolved with NaOH leaving the recombinant DNA denatured so that it can stick to the filter. The filter is then incubated with a ^{32}P labeled probe. In this case the probe is from two EMBP cDNAs. They will be allowed to hybridize to complimentary sequences on the nitrocellulose and the position of the clone having the DNA we are looking for can be revealed by autoradiography. Now the desired clone can be selected and manufactured in a bacterial host .

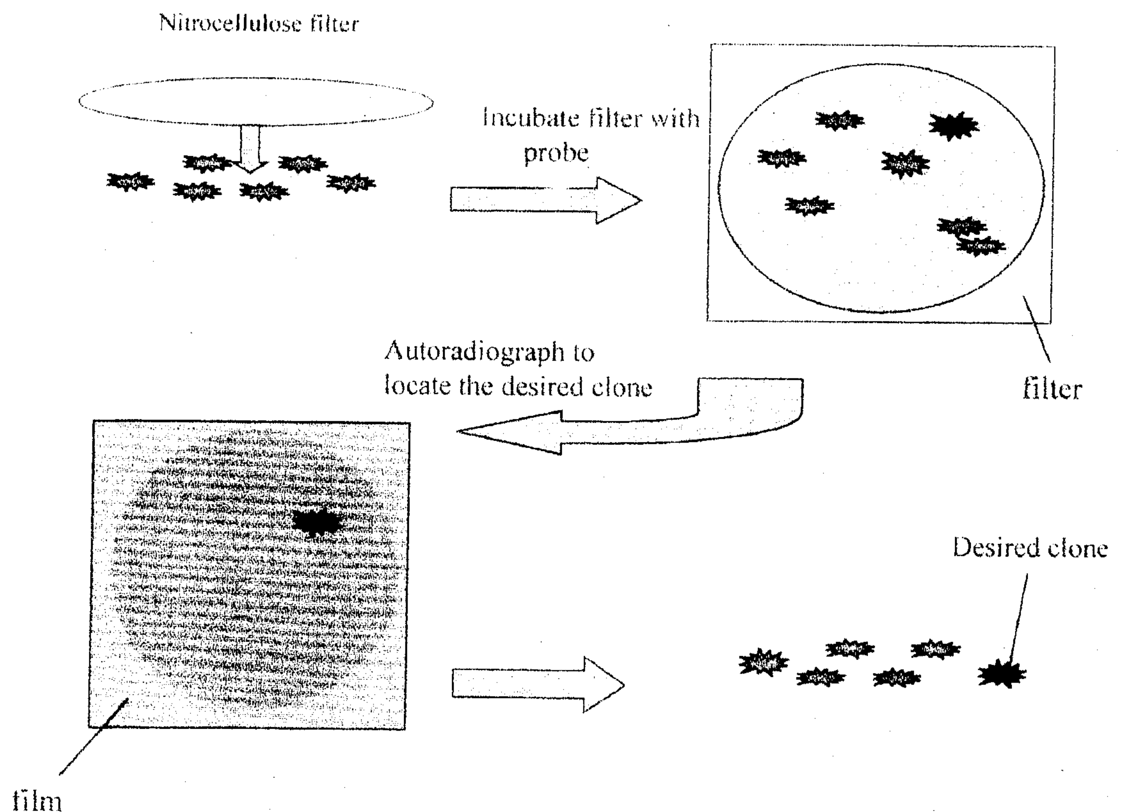


Figure 9. Selection of the gene for p53EMBP.

Construction of the sea urchin genomic library

DNA Isolation

In order to begin the creation of the sea urchin genomic library, the first step was to isolate the DNA. The source of the sea urchin genomic DNA was sperm from the sea urchin species Strongylocentrotus purpuratus. This procedure has been performed according to the protocol 9.16 from Molecular Cloning, *Analysis and Cloning of Eukaryotic Genomic DNA* (Maniatis et al., 1989). The sea urchin sperm used was collected and frozen. The sperm was taken out of the freezer and thawed by placing the conical tube containing the sperm in a beaker of warm water. The amount of sperm to be used for the isolation was weighed to calculate its weight in grams. The amount of sperm used in the isolation was approximately 3 grams. The sperm was placed into a Erlenmeyer flask and 40 milliliters of extraction buffer was added. Next, the sperm and extraction buffer (see Appendix A) solution was pippetted up and down using an electric pippetter. This was to break up the mass of sperm which was clumped together. The sperm and extraction buffer solution was then placed into a 37 °C incubator for a period of one hour. After removing from the incubator the proteinase K was added to a

concentration of 10mg/ml. Therefore, 1 ml of proteinase K was added. The Erlenmeyer flask containing the solution was then placed into a incubator/shaker for three hours at 50°C (Maniatis et al., 1989). Proteinase K was then added and placed into the incubator/shaker at 50°C for three hours (Maniatis et al., 1989). The mixture as agitated every few minutes to resuspend. Following the three hour incubation the solution was placed into two 50 ml conical tubes. An equal volume of phenol was added to each of the solutions in the incubator (Maniatis et al., 1989). The conical tubes were then turned end over end for ten minutes to mix the two phases and then centrifuged for thirty minutes at 3600g. The aqueous phase was extracted off of each of the conical tubes and placed into a new conical tube. The phenol extraction was repeated to a total of three times. Then 0.27 volumes of 7.5M ammonium acetate were added (5.4 mls to each conical tube because each contained approximately 20 mls). The solutions in the two conical tubes were then added back together into a Erlenmeyer flask. In addition, two volumes of 100% ethanol at room temperature were added and swirled (Maniatis et al., 1989). The DNA precipitated immediately. The amount of the isolated DNA filled the flask and was then removed by a glass hook which was made by melting a glass

pipette (Maniatis et al., 1989). The DNA was placed into a new conical tube and 70% ethanol was added to 50mls. The DNA was centrifuged for eight minutes at 3600g. The ethanol was decanted and more 70% ethanol was added to a volume of 50mls. This was centrifuged at 3600g for eight minutes. After the ethanol was decanted the pellet was allowed to dry (but not completion) by leaving the tube open (Maniatis et al., 1989). Five mls of TE (pH 8.0)were added to the conical tube and the DNA was allowed to dissolve. The DNA in TE was allowed to rock on the moving platform for twenty-four hours. An aliquot of isolated DNA was then run out on both a 0.3% agarose gel (see Appendix A) and a 1% agarose gel (see Appendix A) assess the size and state of the DNA (Maniatis et al., 1989). Due to the high concentration of the isolated DNA, one µl of DNA was placed into ten µl of TAE. And then one µl of that was loaded into three separate wells of each of the two gels (This can better be seen in figure 11 of the results section).

In separate DNA isolations another method was also followed in order to attempt to obtain DNA with the greatest purity and largest size. Following the third phenol extraction, the aqueous phase was allowed to dialyze against four liters of TE (pH 8.0) (see Appendix A). This was

done by taking the aqueous phase and pippetting it into a dialysis bag which was then clamped at both ends. Enough room was left in the dialysis bag to allow the solution to increase by 1.5 - 2.0 fold. The dialysis bag was placed into one liter of TE in a beaker containing a stir bar. The solution was allowed to dialyze for approximately twelve hours while stirring. This procedure was followed until the four liters of TE had been used to dialyze the solution containing the DNA (Maniatis et al., 1989).

Restriction Digest

The next step in the preparation of the DNA library was to digest the DNA using restriction enzymes. Two different methods were used to digest the sea urchin genomic DNA. At first a partial digest of the sea urchin genomic DNA was completed. Lastly, a complete digest of the sea urchin genomic DNA was accomplished.

Partial restriction digest of the isolated genomic DNA

Partial digestion of sea urchin DNA using the restriction enzyme Mbo I was done (Mbo I cuts at GATC) (Maniatis et al., 1982). This procedure was at first performed using only a very small amount of DNA. Once the optimum partial digest conditions were established, the digestion was necessarily scaled up. The enzyme was diluted for the small scale partial digest using a dilution buffer (see Appendix A). At first the partial digest

was performed diluting the enzyme Mbo I in a 1:10 concentration. This was done by taking one μl of the enzyme and diluting it in nine μls of dilution buffer (see Appendix A).This proved not to be dilute enough. Therefore, enzyme Mbo I was then diluted in a 1:50 concentration by taking 49μl of the dilution buffer and adding 1μl of the concentrated enzyme Mbo I. Next an eppendorf® tube of the following was made: 1 μl of the DNA isolated earlier, 1 μl of the 1:50 dilution of enzyme Mbo I, 2 μl of 10x Rxn buffer (obtained from Promega), and 16 μl ddH2O.

The eppendorf® tube containing the above was then placed in the incubator at 37°C for 5 minutes. After precisely 5 minutes the eppendorf® tube was placed into a hot water bath set at 65°C. Next, four μl of stop buffer was added to the eppendorf® tubes and the DNA digested by Mbo I was run out on a 0.3% agarose gel. After the optimal time point and dilution of the enzyme was determined the conditions were scaled up and a large scale digestion of the DNA was completed.

The digest was then performed again by scaling up the smaller scale digestion. However, at first when these conditions were employed and the DNA was viewed on a long DNA gel after being fractionated on a glycerol gradient (see below) the amount of DNA used was too little to view. So,

when scaling up some alterations were made. The amount of dH2O used was replaced by DNA. The conditions used were as follows: 340 µl of the isolated DNA, 20 µl of the 1:50 concentration of enzyme Mbo I, and 40 µl of Rxn buffer C. The above ingredients were added to a eppendorf® tube and immediately placed into an incubator at 37°C. After exactly 5 minutes the eppendorf® tube containing the reaction was placed into a -80°C freezer. This step was necessary to stop the reaction.

Complete restriction digest of the isolated genomic DNA

A complete digest of the isolated DNA was also attempted. A small scale restriction digest was done using the enzyme Bam H1. Once optimal conditions were determined then the reaction was scaled up. Added together were: 50 µl isolated genomic DNA, 75 µl dH2O, 15 µl 10x buffer, and 2 µl enzyme Bam H1. Once added together the digest was allowed three hours for completion by incubation at 37°C.

Density Gradient Centrifugation

This was done by first making 500 mls of NET buffer according to the Maniatis protocol for sucrose gradients(2.85)(See Appendix A). Then a 40% glycerol solution and a 10% glycerol solution were made with the NET buffer (see Appendix A). The denser glycerol (40%) solution into a centrifuge tube

that is placed straight in a test tube rack. Carefully, the lighter (10%) glycerol solution was slowly pippetted on top of the denser glycerol solution. The interface is visible at this point. The test tube was then closed with a rubber stopper. Next, the test tube rack was slowly laid down on its side and the glycerol solutions were allowed to diffuse with one another for three hours. When the three hours were up the test tube rack was very slowly straightened back up (Abe and Davis, 1986). The restriction digested DNA was then added to the top of the gradient. A blank gradient was then weighed and balanced with the gradient containing the restriction digested genomic DNA. The gradients were then loaded into a swinging bucket rotor and centrifuged for 18 hours at 26K.

Sucrose gradients were also made. A 60% sucrose solution and a 15% sucrose solution were made (see Appendix A). The gradients were then made the same as above with the glycerol gradients. The restriction digested DNA was added to the top of the gradient. A blank gradient was then weighed and balanced with the gradient containing the restriction digested genomic DNA. The gradients were then loaded into a swinging bucket rotor and centrifuged for 18 hours at 26K.

The next step was to fractionate the gradients. Forty eppendorf® tubes were numbered from 1-40 for each of the gradients. The eppendorf® tubes

were then placed in order in a long row so that they would be handy for the fractionation. The gradient to be fractionated was then clamped onto a stand and a rubber stopper was placed onto the top of the test tube. An 18 gauge needle was placed into the bottom of the test tube and approximately 0.5ml fractions were collected in the numbered eppendorf® tubes.

For the fractions collected from the glycerol gradient, every other fraction collected was run out on a 0.3% long agarose gel (see Appendix A) overnight. Three µls of stop buffer was added to the fractions to be run on the gel. The voltage on the gel was run at 45 volts for approximately 20 minutes until the dye ran out of the wells and then the voltage was turned down to 16 volts for 16.5 hours. The gel was dyed with ethidium bromide and viewed on a large lightbox.

For the fractions collected from the sucrose gradient, every other fraction was also run out on a 0.3% long agarose gel however, with the sucrose gradient the fractions to be run on the gel had to be precipitated first. Therefore, 300µl of each of the fractionated samples of the sucrose gradient had two volumes of 100% ethanol added to them in order to precipitate the DNA. The fractions were then microfuged at 7000g. The supernatant was then discarded. Thirty µls of TE were placed onto each of the pellets and

placed on the shaker for ten minutes. Three µls of stop buffer was added to each of the samples. Lastly, twenty µls of the fractionated, precipitated DNA was loaded onto the 0.3% long agarose gel and electrophoresed overnight for 16.5 hours. The gel was run at 45 volts for approximately 20 minutes until the dye front was out of the wells and then the voltage was turned down to 16 volts. The next day the gel was stained with ethidium bromide and viewed on a large lightbox.

Gel Purification

Once the restriction digest is completed the next step is to make sure that the fragments to be ligated are of a certain size range. Once the partial restriction digest was performed, a long 0.3% agarose gel was made. The comb which is used to make wells in a agarose gel were taped together in this case in order to create one large trough. One hundred µls of the partially digested genomic DNA and five µls of Stop buffer were loaded into the trough. The gel was then electrophoresed overnight for sixteen and a half hours. Starting with forty-five volts until the dye was out of the trough and then turned down to sixteen volts for the remainder of the time. The next day the gel was stained with ethidium bromide and viewed on a ultraviolet lightbox. Using a razor blade a portion of the gel containing the high

molecular weight DNA was cut out. The gel containing the band of high molecular weight DNA was then placed into a dialysis bag and the ends were clamped. The dialysis bag was then placed into a electrophoresis chamber containing 0.5X TAE and electrophoresed at 60 volts for approximately an hour. This step allows the DNA to be removed from the gel and into the surrounding fluid contained in the dialysis bag. In the next step a column was then primed with three mls of high salt buffer by pushing the buffer through the column with a syringe. Then column was then subjected to a second wash with three mls of low salt buffer. The DNA which was collected from the dialysis bag was then pushed through the column with even pressure. The DNA was then eluted from the column with four hundred µls of high salt buffer and into a eppendorf® tube. The DNA was then extracted with four hundred µls of phenol, microfuged at 14,000 for five minutes, and the top layer was collected into a fresh eppendorf® tube. The DNA is next extracted with 400 µl of chloroform, microfuged at 14,000 for five minutes, and the top layer was then collected into a fresh eppendorf® tube. Next, five hundred µls of isopropanol was added and spun for fifteen minutes, the top layer is drawn off. The top layer was then washed one time with 80% EtOH, spun for five minutes at 14,000, the liquid was then removed from the pellet and

dried. The DNA was then resuspended in twenty µls of TE (Tris-EDTA).

Ligation, Packaging, and Plating

DNA Ligation

The ligation was then done by adding together the genomic DNA fragments, the pre-digested lambda vector arms, 10x ligation buffer, and ligase. The following concentrations were used: 2 µls lambda vector arms, 6 µls genomic DNA fragments, 2 µls 10x ligation buffer, 1 µls ligase, and 10 µls sterile H2O.

Once added together the ligation mixture was then run overnight in the PCR machine on the ligation setting at 14°C. The next day the packaging reaction was performed.

Cultures of E.Coli were grown in order to get ready for plating. An overnight culture of E.Coli strain KW392 in 2 mls of LB broth (see Appendix A) was grown. Added to a sterile tube containing LB was a toothpick which was used to scrape the top of the frozen glycerol stock of E.Coli strain KW392. The tube was then placed in a incubator/shaker overnight at 37°C. From the overnight culture the next day a two hour culture of E.Coli strain KW392 was grown. Five hundred µls of the overnight culture of KW392 was placed into fifty mls of LB and placed into

the incubator/shaker at 37°C for a couple of hours. Once the culture was grown up the cells were placed into a centrifuge tube and pelleted by centrifugation. The cells were then resuspended in 25 mls of $MgSO_4$ to prepare them for plating.

Packaging

Next, the ligation reaction is packaged by adding together the ligation mixture and the packaging extracts. The vector DNA (λ phage) was purchased prepared by enzyme digestion and removal of unnecessary intermediate fragments was already completed. A Stratagene gigapak® packaging extract tube was removed from the -80°C freezer and placed on ice. The packaging extract tube is held between fingers until it just begins to thaw. Ligation of sea urchin Mbo I digestion DNA to Mbo I digested λ DNA was then carried out. After ligation, the phage DNA were packaged into phage particles *in vitro*. The packaged phage, then, were used to infect E.Coli culture KW392 (Maniatis et al., 1982). In doing so, immediately eleven μls of the thawed ligation mixture was added to the tube with the packaging extracts. It was then placed on ice. Carefully the tube was stirred with a pippette tip being careful not to introduce bubbles. The tube was then microfuged briefly in order to ensure that all the contents were at the bottom

of the tube. The tube was then incubated at room temperature for exactly one hour and forty-five minutes. After the incubation period was over, 500 mls of SM buffer was added to the tube. Next, a serial dilution using SM buffer was done to prepare for plating (this can be better seen in figure ten below). This allows for better distinction of plaques. The phages and cells are then incubated at 37°C for fifteen minutes. Following the incubation three mls of soft agar was added to each tube. The contents of each of the four tubes are then dumped onto four agar plates consisting of NZCYM media. The plates were then incubated overnight at 37°C. The bacteria containing this λ library will then be grown on bacterial lawns. Plaques should form overnight. The plaques are then harvested. This last step results in the amplification of the library (Maniatis et al., 1982).

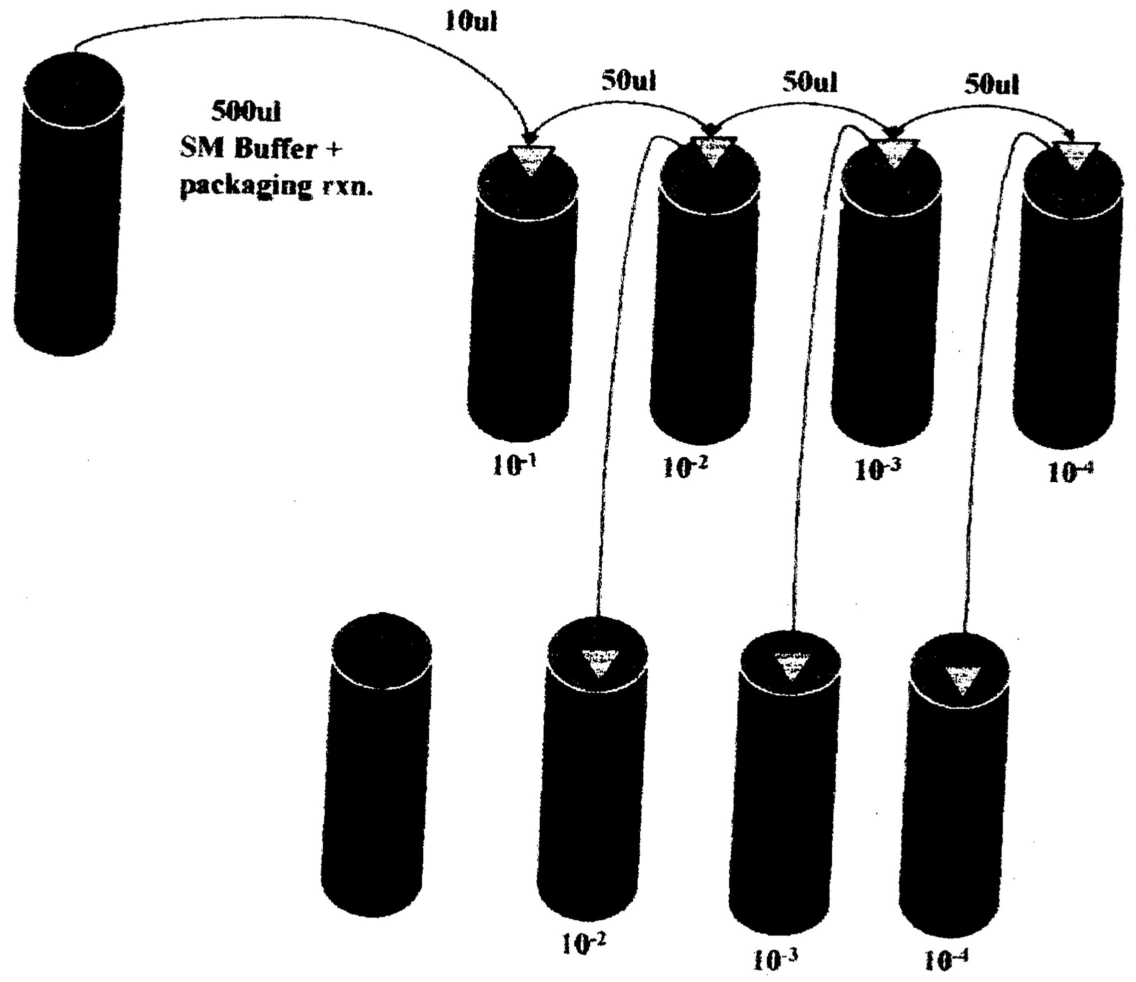


Figure 10. Serial dilutions of the packaging reactions

IV. Results

The gene coding for EMBP is expected to be large based on Northern blot analysis. The mRNA is larger than the 28s RNA which corresponds to 4,718 nucleotides (unpublished data).

The isolation of the DNA from sea urchin sperm yielded high molecular weight DNA. This procedure had been repeated several times with varying results. The DNA showing the highest level of purity and largest size was selected to continue on to the next step in genomic library creation which was the restriction digest. In the DNA from the selected isolation, the DNA banded at about 23kbp on 0.3% agarose gels (Figure 9). It indicated a more or less uniform size, with smaller amounts of intermediate fragments. The small size indicated by a smeared appearance. High molecular weight DNA is what is desired in order to exclude randomly broken DNA that will not give good restriction digested ends. The selected DNA has been quantified using UV absorption. Purity was assessed by a A_{260}:A_{280} ratio, and this was found to be 1.92. This indicates highly purified DNA. The amount of DNA purified was 73.5μg/ml total of 73μg (table 1).

A_{260} ηm = 1.4570
A_{280} ηm = 0.759
concentration = A_{260} ηm = 1 = 50μg/ml
so, 1.4570 x 50μg/ml
73.5 μg/ml

Table 1. Selected DNA using UV absorption

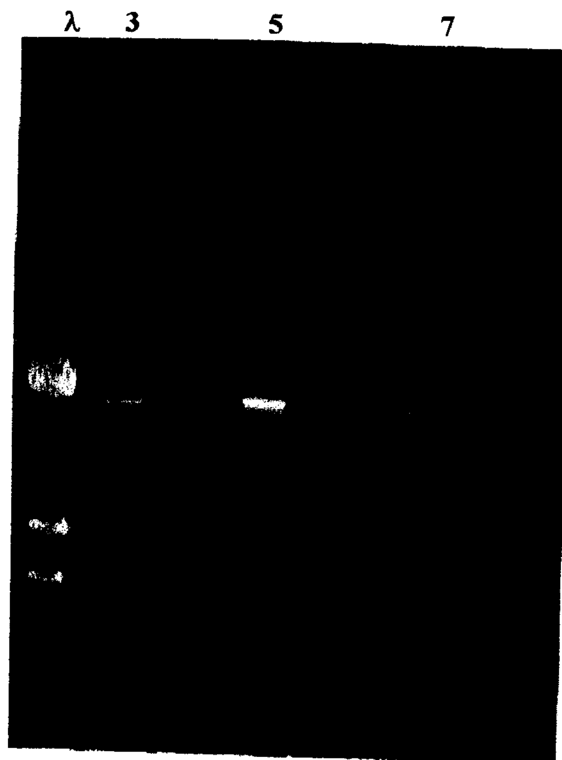


Figure 11. Shown here is isolated sea urchin sperm DNA with the highest degree of purity and largest size. Lane 1 is the Hind III digested lambda DNA ladder used as a size marker, Lane 2 is blank, Lane 3 isolated sperm DNA, 15μl load, Lane 4 is blank, Lane 5 isolated sperm DNA, 24μl load. Lane 6 is blank. Lane 7 isolated sperm DNA, 8μl load. The isolated DNA shown here was chosen to create the genomic library.

The procedure of isolating the DNA was performed several times under varying conditions. In the end the DNA shown in figure eleven was chosen to continue on with the restriction digest. However, isolation of sea urchin DNA was attempted from a different source other than sea urchin sperm. This source was sea urchin eggs. This was done following the same protocol from Maniatis et al. as used with the sea urchin sperm. After the isolation was done the purity was assessed by U.V. absorption and the A_{260} to A_{280} ratio was found to be : 2.012/1.936 $\cong$ 1.039. This ratio suggests impurities in the DNA prep. The ratio expected should be greater than 1.75. A 0.3% agarose gel was run and the results showed a large amount of RNA (figure 12). The highest molecular weight of the RNA was at the 2.0kbp marker and increasingly became lower molecular weight.

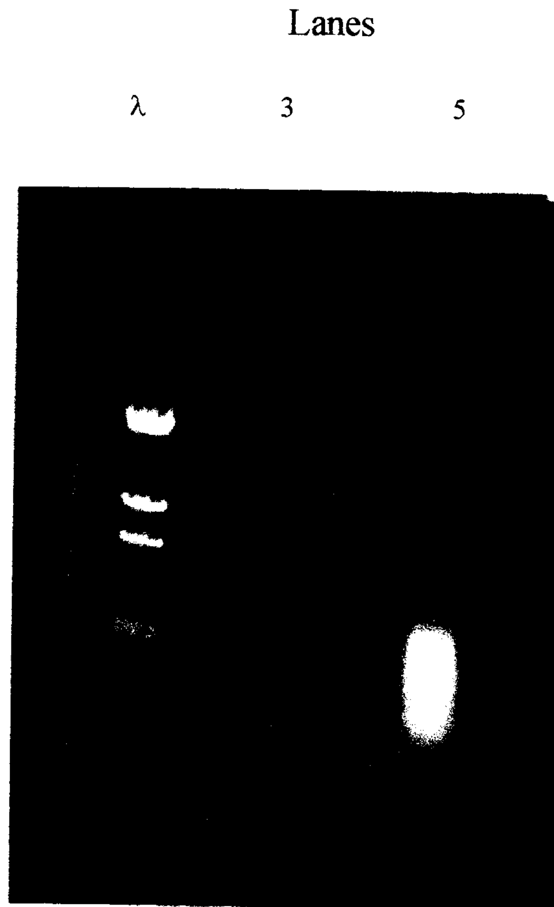


Figure 12. Attempted DNA isolation. In this attempted DNA isolation large amounts of RNA were obtained, this is indicated slightly in lane 3 and by the large illuminated appearance in lane 5.

In a different sea urchin sperm DNA isolation, results turned out fairly well and it was thought that this would be the DNA used to create the sea urchin genomic library. However, it was later found out that this DNA contained a high level of impurities. The DNA shown in figure 11 was then used to attempt library construction. This DNA in this isolation attempt can

be seen in figure 13 below. The sperm used in this isolation were freshly collected.

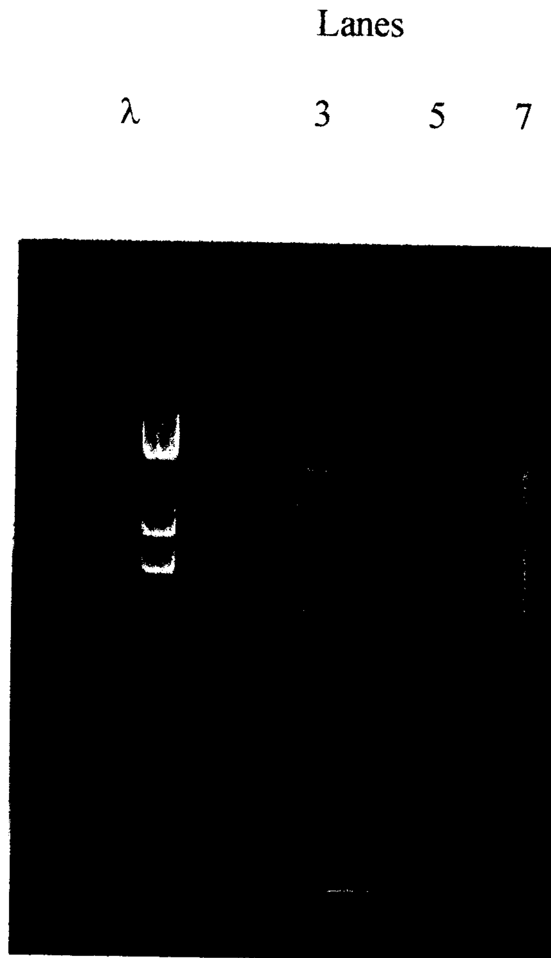


Figure 13. DNA isolation containing impurities.
At first it was thought that the DNA from this isolation would be the one used to create the genomic library. It was later found out that this DNA contained a high level of impurities. Lanes 3, 5, and 7 are different amounts of DNA from the same isolation.

Isolated genomic DNA was subjected to limited restriction enzyme digestion using Mbo I. In the gel shown below in figure 14, the restriction digest was performed using different time points of diluted enzyme (Mbo I), in a 1:10 concentration (see Appendix A). The goal was to obtain fragments

of approximately 20kbp. In the four time points indicated here the fifteen minute and the twenty minute time points were entirely too long. Entirely too much digestion of the genomic DNA occurred. Therefore, the digest was redone using diluted enzyme in a 1:50 concentration.

Figure 14. Restriction digest with diluted enzyme Mbo I in a 1:10 concentration.
Lane 1 shows EcoRI digested lambda DNA used as a standard, Lane 2 is blank, Lane 3 has a five minute time point, Lane 4 is blank, Lane 5 has a ten minute time point, Lane 6 is blank, Lane 7 shows a fifteen minute time point, and Lane 8 is showing a twenty minute time point.

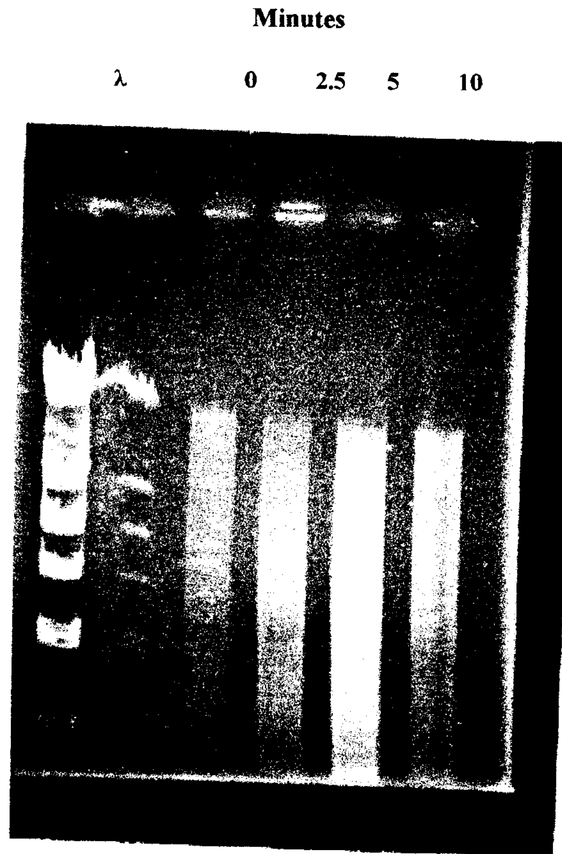


Figure 15. Restriction digest with enzyme Mbo I diluted 1:50. Lane 1 shows the EcoRI digested lambda DNA ladder, Lane 2 is blank (however, some of the ladder spilled over), Lane 3 is a zero time point of undigested DNA, Lane 4 is a 2.5 minute time point, Lane 5 is a 5 minute time point, Lane 6 shows a ten minute time point.

Nevertheless, the restriction digested genomic DNA with the enzyme in a 1:10 dilution failed to produce the high molecular weight DNA needed. As shown in figure 15 three more time points were run. This time the concentration of the enzyme Mbo I was in a 1:50 dilution. The fragments at the 2.5 minute time point appear barely to be cut at all when compared to the

zero time point. The five minute time point was selected as the optimal time point and the digest was necessarily scaled up.

In order to isolate restriction DNA fragments, the digests were subjected to density gradient centrifugation. Glycerol or sucrose gradients were made, centrifuged, and fractionated. Every other fraction was electrophoresed overnight on a 0.3% agarose large format gel. Density gradient centrifugation was repeated various times, all with unsuccessful results. The gradients did not separate the DNA according to size as had been planned (see figure 16 below). Both glycerol and sucrose gradients were tried (results of sucrose gradient not shown) as an effort to produce uniform gradients. A gradient maker was also tested in establishing gradients. However, due to mechanical problems with the gradient maker this method was abandoned.

Fractions

λ 2 3 5 6 8 9 11 12 14

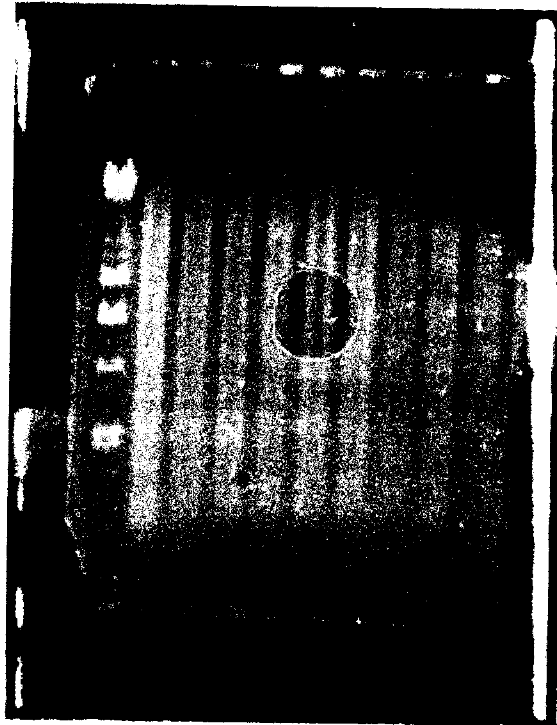


Figure 16. Glycerol Gradient Gel. This gel shows fractions that were electrophoresed on a 0.3% agarose gel. As shown the lanes do not slope downward as one would expect a properly separated gradient to look. Lane 1 shows the Hind III Digested DNA ladder. Lane 2 shows fraction #2. Lane 3 shows fraction #3. Lane 4 shows fraction #5. Lane 5 shows fraction #6. Lane 6 shows fraction #8. Lane 7 shows fraction #9. Lane 8 shows fraction #11. Lane 9 shows fraction #12. Lane 10 shows fraction #14. Lane 2 signifies the bottom of the gradient and Lane 14 signifies the top.

In an effort to determine the problem in the density gradient separation, test gradients were run using lambda DNA. This would allow optimalization of conditions without using the isolated genomic DNA. This method showed that gradients were unable to be established. If a gradient was established, when analyzed by gel

electrophoresis and the lanes ran in order, the lanes should slope in a downward fashion.

Due to the inability to establish a gradient with density gradient DNA separation on sucrose or glycerol gradients, a different method to obtain the optimal size DNA fragments was done. This method involves gel purification (gel purification was only necessary for the genomic DNA that was digested under partial digest conditions).

The results of the gel purification turned out to be very good on the first try. The excess small fragments were gotten rid of upon purification from the 0.3% large gel and the approximately 20 kbp fragments necessary for inserting into the lambda vector were obtained (this can be seen in figure 17).

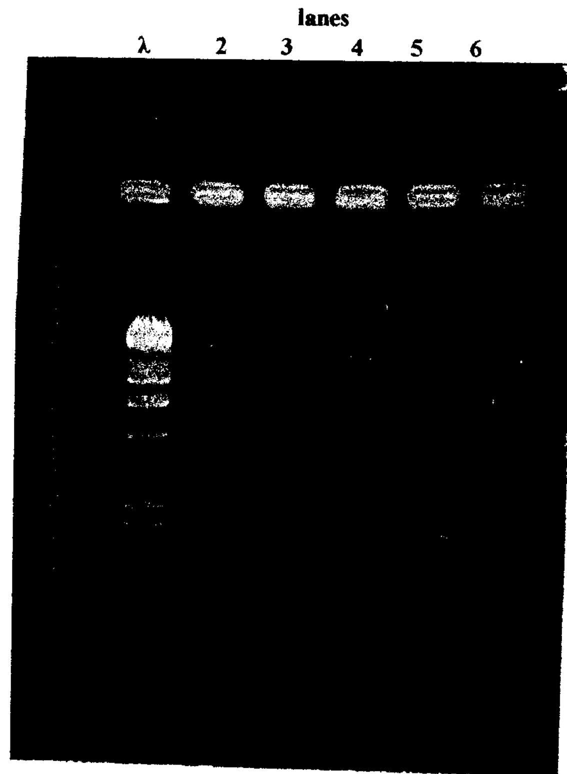


Figure 17. Gel purification gel. Lane 1 is the pre-digested lambda DNA ladder. Lanes 2 and 3 are blank. Lane 4 contains the gel purified DNA of approximately 20 kbp. Lanes 5and 6 are blank. The pre-digested ladder fragments are as follows: (starting from the top) 23kbp, 9.4 kbp, 6.5 kbp, 4.3 kbp, 2.3 kbp, and 2 kbp.

Ligation, packaging, and plating were done. The result was that there were no plaques. This could possibly be due to the fact that ligation may not have occurred. The leftover ligation mixture was run out on a 0.3% agarose gel. As indicated below in figure 16, there are three bands in the lane where the ligation mix was loaded. This may indicate that two of these bands are the lambda DNA arms and the third is the fragments that were to be inserted.

So ligation probably did not take place. The ligation was redone several times with the same results.

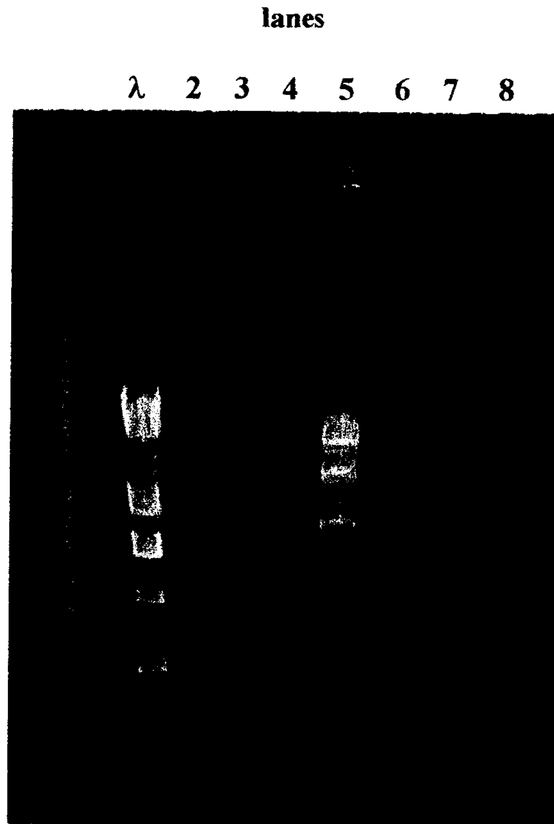


Figure 18. DNA ligation gel. Lane 1 shows the pre digested lambda ladder. Lanes 2, 3, and 4 are blank. Lane 5 shows what was left over from the ligation mixture. Lanes 6, 7, and 8 are blank. Three bands are shown in lane 5 indicating that ligation may not have taken place.

Due to the problems occurring with sea urchin genomic library construction under partial digest conditions, it was decided that the direction taken in creating the library should be changed. Now, a complete restriction digest would be performed. A small scale digest was performed with enzyme Bam HI. Several times the digest was performed and no digestion occurred. This was found to be because the enzyme was old. New enzyme was

purchased and once optimal results were established, the digest was necessarily scaled up. The results can be seen in figure 19 below.

λ 2 3 4 5 6

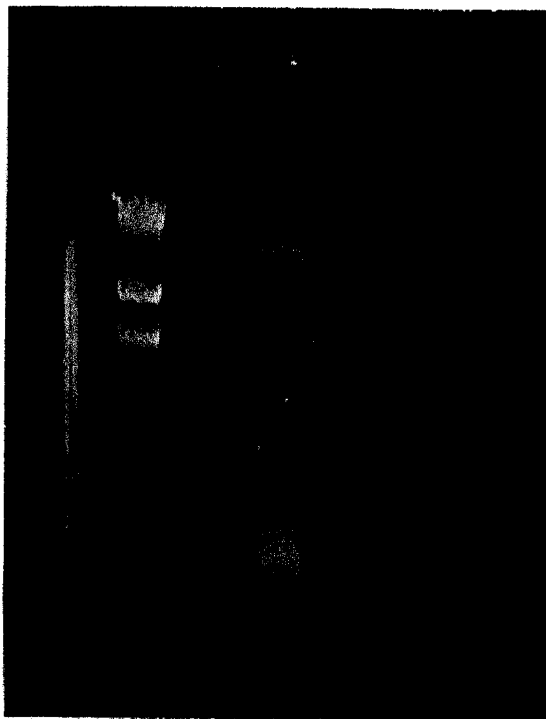


Figure 19. Bam HI digested DNA. This DNA was used for the complete digest of the sea urchin genomic library. Lane 1 shows the lambda DNA digested ladder as a standard. Lane 2 is blank. Lane 3 shows the uncut sea urchin genomic DNA. Lane 4 is blank. Lane 5 shows the complete restriction digested DNA by Bam HI. Lane 6 is blank.

Again, the ligation, packaging, and plating was carried out as before. Once again, plaques were not formed. A test ligation was performed to check the protocol used. The test ligation produced plaques on the agar plates indicating that the procedure used could indeed produce positive results.

V. Discussion

In the initial stage of this project problems were encountered in restriction digestion and the subsequent ligation steps resulting in a failure to construct a DNA library. Thereby hindering accomplishment of the second specific aim, which was to isolate the gene for the sea urchin 53kDa myosin binding protein. A possible conclusion to the difficulties that have occurred while trying to create this sea urchin genomic library is the high quantity of methylated DNA in the sea urchin genome and/or in sperm DNA alone. In previous research it has been found that about forty percent of the sea urchin genome is methylated (Bird et al., 1979) Methylation of this DNA could possibly be the reason that this genomic library was so difficult to create. Furthermore, methylation could have been the reason for many or all of the problems that have arisen in creation of this library.

Although, the DNA isolation was performed various times with various results. In the end the DNA from the very first isolation proved to be of the highest molecular weight and the greatest purity. The isolation was performed several times while changing variables with the hopes of obtaining better results. However, the first isolation was chosen to continue on with rest of the project. Upon isolation of the DNA from sea urchin eggs and gonads it

was found that this tissue contained too much RNA. The best luck in acquiring the large amount of DNA needed was isolated from sea urchin sperm. However, this DNA proved to be contaminated with a large quantity of impurities. Alternate sources of sea urchin DNA should be investigated. However, due to lack of demand a commercial sea urchin genomic library was unable to be located.

The state of sperm DNA could possibly make it difficult to cleave by restriction enzymes (Bird et al., 1979). It is also possible that methylation occurring in the sperm is causing the difficulties and sea urchin DNA from sources other than sperm and eggs should be researched.

The restriction digest had to be performed both as a partial digest and then a complete digest since creating a genomic library with the former proved to be unsuccessful. The ligation performed with the partial digest appeared not to ligate according to the results seen in figure 18. However, this conclusion was based on results from agarose gel electrophoresis which is well-known for uncertainty for results such as these. Although it is hard to confirm from the agarose gel, the results of the gel run from the leftover ligation mixture indicate that no ligation took place. Had recombinant molecules transpired, plaques on the agarose plates would have been indicative of the ligation occurring. It could be that methylation is not

allowing all sequences to be restriction digested. The enzyme that was chosen for the partial restriction digest (Mbo I) is not sensitive to all types of methylation.

Other problems in this project have to do with the inability to establish uniform gradients. When the fractionated gradients were analyzed by gel agarose electrophoresis it was found that the lanes were the same all the way across instead of sloping downward. If the lanes were sloping downward it would indicate that the DNA was in fact separated by molecular weight. In all of the methods chosen to create the gradients this turned out to be the problem. When the gradients were created with both sucrose and glycerol this same dilemma occurred, so the substance used to create the gradients was not the problem. This could have happened because the centrifuge stopped too abruptly or perhaps other problems with the swinging bucket rotor. A gradient maker was also tried in creating these gradients but, due to problems with the tubing this method also had to be abandoned. These gradients were fractionated and weighed before running to see if the gradient was uniform and it proved not to be. If the gradient was uniform weighing out the same quantity of each of the fractions should have showed a gradual difference in weight.

Due to the problems occurring with density gradient centrifugation, gel purification was used to obtain the needed size DNA fragments. This proved to be the method in the project with the least uncertainty. Upon the first try the approximately 20kbp DNA fragments needed were purified. The fragments were then used to attempt the DNA ligation with the λ vector arms that were pre-digested with the same enzyme.

A complete restriction digest was performed when ligation problems occurred in the partial digested sea urchin sperm DNA. This was digested with enzyme Bam HI. Along with the partial digested DNA problems also transpired in the ligation step with sperm DNA that had been digested to completion. Restriction enzymes other than Mbo I and Bam HI that are less sensitive to DNA methylation should be tried as an effort to create recombinant molecules.

The problems occurring in the ligation are not from the procedure chosen. A test ligation was run in order to ensure that the mechanics of the procedure were being done properly. Plaques were formed on the test plates.

This indicates that the problems occurring were indeed from the DNA.

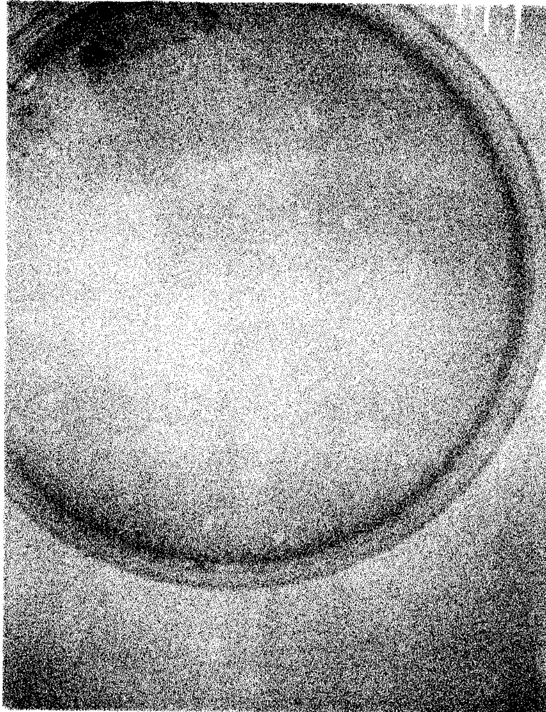
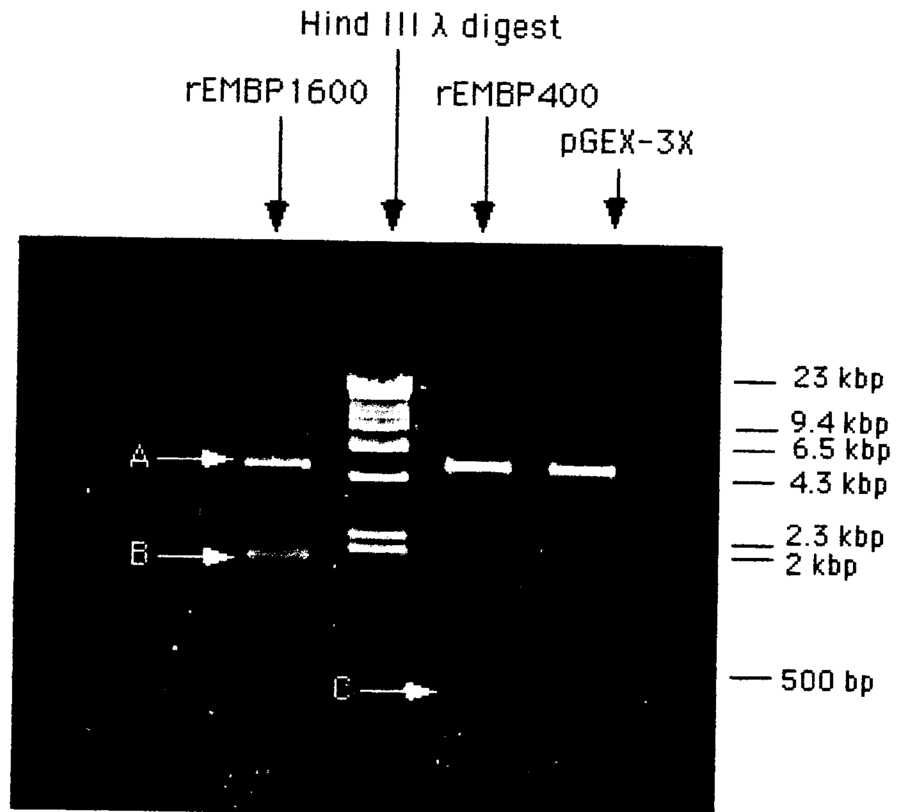


Figure 20. Plaques on agar plate from test ligation. Although hard to see, this plate contained 102 plaques from the test ligation. This plate was the 10^{-2} test plate.

If ligation had occurred, perhaps another strain of E.Coli could be tried in order to package the recombinant molecules. Perhaps if recombinant molecule were formed and if they contain methylated portions they may not be being recognized by E.Coli's genome. Trying different strains of E.Coli that might be able to recognize fragments that are methylated should be explored. If indeed the problem with ligation is the sea urchin genome containing large quantities of methylated DNA that are not being recognized other methods of overcoming this will have to be explored.

In the event that recombinant molecules were able to be made and plaques had formed, the second specific aim of this project was to selected for the entire gene for p53EMBP. EMBP cDNA probes (discussed below) were to be used to screen the sea urchin genomic phage DNA library, described above. Once the phage were able to be plated out on bacterial lawns at a density giving distinct plaques, nitrocellulose filter disks would then have been placed over the bacterial lawns. Absorption of phage DNA to nitrocellulose would then occur, and nitrocellulose would then be vacuum dried at 80°C. This fixes the DNA in place. Next, the nitrocellulose would be moistened with a solution containing ^{32}P labeled probe DNA (Maniotis, 1982). The ^{32}P labeled probe DNA would be allowed to hybridize to complimentary sequences present on the nitrocellulose. The nitrocellulose would then be washed extensively under conditions of high stringency, dried, and autoradiographed. The positions of the labeled plaques could then be cross-referenced back to the original cultures. The phage would be recovered from these positive plaques. Each plaque would be subjected to a second round of plaque selection as described above. This would insure clonality.

Once the library is established, it must be probed in hopes of obtaining several clones for p53EMBP. This would be done using two plasmids available as a source of cDNA. These plasmids are pGEX rEMBP 400 and pGEX rEMBP 1600 (pGEX is the plasmid, rEMBP (recombinant egg myosin-binding protein) is the insert, and 400 or 1600 is the number of base pairs) (figure 21). These two sea urchin egg myosin binding protein cDNAs were prepared to be utilized as probes for the screening. The cDNA in these plasmids are of two sizes, 400 bp (rEMBP400) and 1600 bp (rEMBP1600). They would have both be used in order to maximize the number of clones detected and to increase the amount of the gene that would be probed for. The cDNAs would be gel purified and ^{32}P labeled (please see figure 21). These two plasmids code for different portions of p53EMBP. These plasmids are easily purified using the PERFECT prep ™ plasmid DNA kit. The inserted cDNAs can be cut out with EcoR1 (Figure 21) and used as the probe.



1% agarose gel of Eco RI digested
Plasmids. Stained with ethidium bromide.
A, linearized plasmid DNA
B, 1600 bp cDNA coding forEMBP
C, 400 bp cDNA coding for EMBP

Figure 21. This is a DNA gel showing the inserts.

Future work

In order to establish the genomic library alternate sources of DNA other than sea urchin sperm and eggs should be tried. A library could be established from a closely related species such as the star fish in hopes that another species would be similar enough to probe for the gene and the

problems encountered in creation of the sea urchin genomic library could be avoided. If this does not work, perhaps if there are any methods available to demethylate the sea urchin DNA these could be tried in order to establish the sea urchin genomic library. Furthermore, alternate methods of obtaining DNA from the sea urchin eggs in order to avoid the large quantities of RNA could be tried.

Once the problems of creating this genomic library are worked out, and the gene for p53EMBP is obtained further characterization of the gene will be accomplished by analysis of its genetic sequence. This analysis will identify any functional domains that are known. Furthermore, comparative studies can be done between this gene and other known myosin binding protein sequences using recombinant EMBP. Transformation experiments then could be done to understand the role of EMBP in cells. Introduction of altered EMBP genes into sea urchin eggs can be used to address the questions of the developmental role of EMBP. Transformation of other cell lines may suggest the fundamental role EMBP plays in cellular physiology.

Further investigation into the sequence analysis on the world wide web that p53EMBP could indeed be VAP-1, a vesicle transport protein should be attempted. Until the entire gene is isolated for p53EMBP it won't be known for certain whether these two proteins are homologous or not.

Once the gene is isolated characterization of he gene can take place. The restriction mapping of the isolated clones will be useful in characterizing the gene for further manipulations. The specific enzymes to be used should be determined at that time.

Literature Cited

Abe, S and E. Davies. "Quantitative analysis of polysomes using a baseline from uncentrifuged blank gradients." Memoirs of the College of Agriculture, Ehime University 31:(1986)187-199.

Alberts, Bruce, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, and James D. Watson. Molecular Biology of the Cell. New York: Garland Publishing Inc., 1994. 1294 pages.

Ayme-Southgate, A., J. Vigoreaux, G. Benian, and M. Pardue. "*Drosophila* has a twitchin/titin-related gene that appears to encode projectin." Proceedings of the National Academy of Sciences 88 (1991):7973-7977.

Barton, N. R., E.M. Bonder, D. J. Fishkind, R. H.Warren, and M.M. Pratt. "A novel vesicle-associated protein (VAP-1) in sea urchin eggs containing multiple RNA-binding consensus sequences." Journal of Cell Science 103 (1992):797-809.

Benian, Guy M., X. Tang, and T. L. Tinley. "Twitchin and Related Giant Ig Superfamily Members of *C. Elegans* and Other Invertebrates."Advanced Biophysics 33(1996):183-197.

Bird, Adrian P., Mary H. Taggart, and Barbara A. Smith. “Methylated and Unmethylated DNA Compartments in the Sea Urchin Genome.” Cell 17(1979): 889-901.

Eilertsen, Kenneth J., Steven T. Kazmierski, and Thomas C.S. Keller III. “Cellular Titin Localization in Stress Fibers and Interaction with Myosin II Filaments In Vitro.” The Journal of Cell Biology 126 (1994): 1201-1210.

Eppenberger, H.M., J. C. Perriard, U. B. Rosenburg, and E.E.Strehler. “The Mr 165,000 M-protein myomesin: a specific protein of cross-striated muscle cells.” Journal of Cell Biology 89.2 (1981):185-193.

Gautel, Mathias. “The Super-Repeats of Titin/Connectin and their Interactions: Glimpses at Sarcomeric Assembly.” Advanced Biophysics 33 (1996): 27-37.

Grove, B.K., L. Cernyl, J. C. Perriard, and H. M. Eppenburger. “Myomesin and M-protein: expression of two M-band proteins in pectoral muscle and heart during development.” Journal of Cell Biology 101.4 (1985): 1413-1421.

Higgins, Desmond G., S. Labeit, M. Gautel, T. J. Gibson. ”The Evolution of Titin and Related Giant Muscle Proteins.” Journal of Molecular Evolution 38(1994): 395-404.

Keller III, Thomas C.S. "Structure and Function of Titin and Nebulin." Current Opinion in Cell Biology 7(1995): 32-38.

Labeit, Siegfried, Bernhard Kolmerer, and Wolfgang A. Linke. "The Giant Protein Titin Emerging Roles In Physiology and Pathology." Circulation Research 80(1997): 290-294.

Maniatis, T., E.F. Fritsch, and J. Sambrook. Molecular Cloning A Laboratory Mannual. Cold Spring Harbor Laboratory Press (1982). 545 pages.

Maniatis, T., E.F. Fritsch, and J. Sambrook. Molecular Cloning A Laboratory Mannual. 2nd ed. Cold Spring Harbor Laboratory Press (1989).

Seiler, Stephanie H., Donald A. Fischman, and Leslie A. Leinwand. "Modulation of Myosin Filament Organization by C-Protein Family Members." Molecular Biology of the Cell 7(1996): 113-127.

van der Ven, P.F., and Furst, DO. "Assembly of titin, myomesin and M-protein into the sarcomeric M band in differentiating human skeletal muscle cells in vitro." Cell Structure and Function 22.1(1997): 163-171.

Vaughan, Kevin T., F. E. Weber, S. Einheber, and D. A. Fischman. "Molecular Cloning of Chicken Myosin-binding Protein (MyBP) H (86-kDa Protein) Reveals Extensive Homology with MyBP-C (C-protein) with Conserved Immunoglobulin C2 and Fibronectin Type

III Motifs."The Journal of BiologicalChemistry 268.5 (1993): 3670-3676.

Voet, Donald, and Judith G. Voet. Biochemistry. New York: John Wiley and Sons Inc., 1995.1361 pages

Walker, Gary, Rachel Yabkowitz, and David Burgess. "Mapping the Binding Domain of a Myosin II Binding Protein." Biochemistry 30.42(1991): 10206-10210.

Wang, Kuan. "Titin/Connectin and Nebulin: Giant Protein Rulers of Muscle Structure and Function." Advanced Biophysics 33(1996): 123-134.

Weisburg, Andrea, and Saul Winegrad. "Alteration of myosin cross bridges by phosphorylation of myosin-binding protein C in cardiac muscle." Proceedings of the National Academy of Sciences 93.17(1996): 8999-9003.

Yabkowitz, Rachel, and David R. Burgess. "Low Ionic Strength Solubilityof Myosin in Sea Urchin Egg Extracts Is Mediated by a Myosin-binding protein." The Journal of Cell Biology 105 (1987): 927-936.

Appendix A

1.Extraction buffer

10mM Tris-Cl

0.1M EDTA

0.5% SDS

2. TE (for dialysate 4L)

50mM Tris-Cl (pH 8.0)

10mM EDTA (pH 8.0)

3. Dilution buffer

10μl BSA (Bovine Serum Albumin)

10μl Buffer B (made earlier according to Promega protocol)

80μl of dH2O

4. NET buffer (according to the Maniotis protocol for sucrose gradients[2.85])

1M NaCl

20mM TrisCl

5 mM EDTA

brought to pH 8.0.

5. Glycerol solutions

For the 40% solution 40mls of glycerol and 60 mls of NET buffer were added together. For the 10% solution 10mls of glycerol and 90 mls of NET buffer were added together.

6. Sucrose solutions

For the 60% solution approximately 77g of sucrose and 51g of dH_20 were added together. For the 15% sucrose solution approximately 16g of sucrose and 90g of dH_20 were added together.

7. 0.3% agarose gel

0.175g of agarose and 50mls of TAE

8. 1% agarose gel

0.5g of agarose and 50 mls of TAE

9. 0.3% Long agarose gel

0.350g of agarose and 100 milliliters of TAE.

10. SM (Phage storage and dilution): 1L

5.8g NaCl

2.0g $MgSO_4$

50.0ml 1M Tris, pH 7.5

5.0ml 2% Gelatin

11. NZCYM (liquid) : 500ml

5g NZamine (Casein, enzymatic lysate)

2.5g NaCl

0.5g Casamino acids

2.5g Yeast Extract

1g $MgSO_4$

NaOH to pH 7.5

12. NZCYM Plates

Bottom agar: 15g Bacto agar/L

Soft agar: 7g Bacto agar/L

13. LB Broth (1L)

10g Bacto-trypton

5g Bacto-Yeast extract

10g NaCl

to 1000 ml dH_2O

pH to 7.5, autoclave