

A DC-STAMP domain within *C. elegans* sperm protein SPE-42 is required for
fertilization

A THESIS

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF THE
UNIVERSITY OF MINNESOTA

BY

Janel Pung

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Tim Kroft, Ph.D.

August 2013

Acknowledgements

I would like to offer my sincere gratitude to my advisor Dr. Tim Kroft for his unfailing support, guidance, funding, and expertise during the course of my research; and to my committee members Dr. Jennifer Liang and Dr. Jon Rumbley for their intellectual contributions, suggestions, and encouragements. I would like to recognize Mr. Luke Wilson, a laboratory technician, for his willingness and efforts placed in support of the work presented in this project as well as undergraduate students Ms. Abigail Richie and Mr. Bryce Meizcowski. I would like to express my appreciation for the opportunity given by the University of Minnesota-Duluth Integrated Biosciences Graduate Program to work in a challenging and rewarding academic program in addition to providing financial backing through graduate teaching. The National Science Foundation (NSF) via Dr. Tim Kroft funded most of the work presented in this thesis. My final thanks goes to my family, friends, and fellow IBS students for all their encouragement.

Abstract

New life requires a sperm and an egg to progress through stages of recognition, binding and fusion once they meet at the site of fertilization. Relatively little is known regarding the common molecular mechanisms of fertilization that unite all metazoans as well as the divergences that are necessary for species specificity. The nematode *C. elegans* has been instrumental in the discovery of several sperm and egg genes that are required for fertilization. Spermatogenesis defective (*spe*) gene *spe-42* functions at the time gametes meet. *spe-42* mutant sperm look and behave like wild type sperm, but fail to fertilize oocytes. The *spe-42* family is present in all organisms that use sperm and eggs and thus may be evolutionarily as old as the sperm and egg system itself. SPE-42 is predicted to be a six-pass sperm plasma membrane protein and contains three essential domains, a large extracellular domain, a RING finger domain, and a dendritic cell-specific transmembrane protein (DC-STAMP) domain. The original DC-STAMP protein is required for cell-cell fusion events unrelated to fertilization such as fusion of preosteoclasts into osteoclasts. The presence of DC-STAMP domains in these otherwise unrelated protein families suggests that this domain is involved in the mediation of membrane fusion. Amino acids within the SPE-42 DC-STAMP domain that showed conservation among the many DC-STAMP domains analyzed were mutated to explore the effect on protein activity. One amino acid was shown to be absolutely essential for function, 2 amino acids nearly erased function and 4 showed mild effects on function. One triple amino acid substitution had no effect on protein function. We also showed that the homologous *C. briggsae* SPE-42 DC-STAMP domain is functional within *C. elegans* SPE-42 DC-STAMP domain. These results support our hypothesis that the DC-STAMP domain is required for SPE-42 function and suggest it is also critical for membrane fusion events mediated by the canonical DC-STAMP protein.

Table of Contents

Acknowledgements	i
Abstract	ii
Table of Contents	iii
List of Figures	iv
List of Tables	v
<i>Chapter One: Caenorhabditis elegans, fertilization and the SPE-42 protein</i>	1
Implications of nematode fertility research	1
Caenorhabditis elegans: a simple model of the molecular mechanisms underlying fertilization	6
spe-42 and let-479: spe-9 class genes	15
A RING finger is located in the SPE-42 C-terminal domain	18
SPE-42 extracellular hydrophilic domain	20
The SPE-42 protein family	25
<i>Chapter Two: Sperm protein, SPE-42, and the DC-STAMP domain</i>	27
Introduction	27
Materials and Methods	28
Sequence alignments	28
Worm strains and handling	29
Plasmid DNA constructs	30
Transgenic C. elegans lines	33
Transgenic rescue experiments	34
Results	35
Amino acid substitutions	37
Effects of amino acid substitutions on fertility	39
DC-STAMP domains are compatible only between closely related proteins	43
Discussion	46
Functional contribution of 8 DC-STAMP conserved amino acids in the SPE-42 DC- STAMP domain	46
DC-STAMP domains are functionally interchangeable only within closely related proteins	50
Potential function of the DC-STAMP domain	51
<i>Chapter Three: Future Directions</i>	52
Bibliography	60
Appendix	67

List of Figures

Figure 1.	<i>C. elegans</i> anatomy.....	8
Figure 2.	Spermatogenesis pathway	10
Figure 3.	<i>spe-9</i> class phenotype.....	13
Figure 4.	Predicted SPE-42 structure with notable domains.....	16
Figure 5.	Projected functions of DC-STAMP.....	23
Figure 6.	SPE-42 DC-STAMP domain-containing homologous sequence alignment.....	36
Figure 7.	<i>spe-42</i> extrachromosomal array	40
Figure 8.	Diagram of amino acids substitutions in SPE-42 DC-STAMP	41
Figure 9.	<i>let-479</i> and <i>spe-42</i> DC-STAMP domain swap plasmid constructs. ...	45
Figure 10.	Amino acid substitution overlap extension PCR strategy.....	67
Figure 11.	DC-STAMP swap overlap PCR strategy.....	68
Figure 12.	Transgenic crossing scheme	68

List of Tables

Table 1. SPE-42 DC-STAMP homology.	37
Table 2. Transgenic rescue with substitution mutations in DC-STAMP	42
Table 3. Transgenic rescue by SPE-42 DC-STAMP chimeric constructs.....	44
Table 4. SPE-42 predicted membrane topology.	69
Table 5. Oligos used to make mutations in wild type SPE-42 plasmid pTK15....	70

Chapter One: *Caenorhabditis elegans*, fertilization and the SPE-42 protein

Implications of nematode fertility research

Successful reproduction involves the uniting of a haploid sperm and egg to form a viable diploid zygote and this process is critical for the continuation of all sexually reproducing metazoans. Both gametes require precise differentiation to play their specific roles during fertilization. The study of the components and molecular mechanisms of fertilization will bring us a step closer to understanding the processes involved, including proteins involved in gamete recognition, binding, and fusion. Fertility studies have been conducted with several organisms, but *Caenorhabditis elegans* is particularly suited for the study of sperm contributions to fertility, the primary focus of this work.

Traditional model systems for fertility studies include sea urchins (*Arbacia punctulata*, *Strongylocentrotus franciscanus/purpuratus*), mice (*Mus musculus*), fruit flies (*Drosophila melanogaster*), green algae (*Chlamydomonas reinhardtii*), and more recently, the nematode *C. elegans*. Researchers of these organisms have made many advances in the study of sperm contributions to successful reproduction via gamete activation and cell-cell interactions. Sea urchin has been a useful organism for observing the biochemistry and physiology of fertilization for a century and many discoveries of species-specific sperm-egg binding and cell signaling originated with this organism. Sea urchin sperm mobilize towards the egg stimulated by hormones such as *S. purpuratus* speract

and *A. punctulata* resact emitted from the sea urchin jelly-like egg coat (SUZUKI 1995; WARD *et al.* 1985). Egg jelly protein receptors, REJs, on the sperm bind to a polymer in the egg jelly bringing about the acrosome reaction (AR) and formation of the acrosomal process (AP) (MOY *et al.* 1996; SEGALL and LENNARZ 1979; VACQUIER and MOY 1997). The acrosome, a cap-like structure located in the sperm head, contains enzymes needed to penetrate the oocyte coat. The contents are released through disintegration of the acrosomal membrane. The extended AP forms from a store of actin within the apex of the sperm. After the AR, the sperm penetrate the egg jelly and reach the vitelline envelope leading to the attachment of the sperm adhesive protein ligand bindin which attaches to the egg bindin receptor 1 (EBR1) on the egg surface during fertilization. (NEILL and VACQUIER 2004; VACQUIER and MOY 1977).

The mouse has been the primary model organism for the study of fertilization in eutherian mammals. Mouse sperm bind to the zona pellucida (egg coat), which consists of three glycoproteins (ZP1, ZP2, and ZP3). First identified in mice, the sperm receptor glycoprotein ZP3, has a single human homolog and was formerly thought to be the sole initiator of the AR, but recent evidence suggests that sperm acrosome react while passing through the cumulus cell layer or while traveling through the zona on their way to the vitelline envelope (BLEIL and WASSARMAN 1983; GAHLAY *et al.* 2010; WASSARMAN 1990). ZP3 is now thought to participate in an essential interaction with ZP2 for binding of

sperm to the zona (GAHLAY *et al.* 2010). ZP2 and ZP3 are closely associated in the zona and are required for sperm binding. One hypothesis is that ZP3 and ZP2 form a scaffold that sperm bind. ZP2 knockout mice produce few (but not zero) ovulated eggs, but no 2-cell embryos (RANKIN *et al.* 1999). ZP1 knockout mice produce smaller litters along with disorganized zona pellucidae, but this phenotype is much less severe than ZP2 or ZP3 phenotypes, indicating ZP1 is not required directly for sperm binding or fertilization (RANKIN *et al.* 2001). Several potential ligand or receptor proteins in sperm and egg have been identified, but only the Immunoglobulin (Ig) superfamily member, Izumo, in sperm and CD9 (cluster of differentiation) in eggs have been shown to disrupt fertilization at the point where sperm and egg plasma membranes meet (INOUE *et al.* 2005; KAJI *et al.* 2000; LE NAOUR *et al.* 2000). Researchers of these organisms have contributed heavily to the general mechanisms of fertilization. However the communication that occurs between the gametes at the molecular level and the details of the fusion event are still mostly unidentified. *C. elegans* is particularly ideal for sperm research for the following reasons: it is inexpensive, forward genetic screens yield significantly more sperm mutants than similar screens in other organisms, the sperm-egg fusion process is simplified, and sperm mutants are easier to maintain.

C. elegans presents an efficient venue for defining the essential components leading up to and directly involved with fertilization. Many *C.*

C. elegans proteins are conserved in all sequenced species that use sperm and egg. Therefore, analysis of the molecular investigations of sperm-egg interactions that occur during *C. elegans* fertilization are pertinent to most multicellular organisms. Of all known human and rodent genes, approximately 40% have homologous genes to *C. elegans* genes (SHAYE and GREENWALD 2011; WHEELAN *et al.* 1999). The knowledge gained may one day aid in reestablishing the faltering biodiversity of animal life on earth or help in the management of nematodes or other parasites harmful to animal and plant species (THOMAS *et al.* 2004). Because of the intricacy of fertility and the functional diversity of proteins regulating the mechanisms of fertility, the simplicity of *C. elegans*' fertilization compared to other models allows us to focus on fertilization without interference from complex preliminary steps such as penetration of the egg coat. It simplifies the process for study and generates easier models for application to more elaborate organisms. *C. elegans* provides the opportunity to investigate the molecular aspects of fertility and apply the knowledge gained to human fertility and, at the same time, open doors to new ways of combating parasites. This simple organism with upwards of 20,000 genes (Claverie 2001) has opened doors to the study of many fields including ageing, neuromuscular biology, neurology, disease, developmental biology and fertility (BRENNER 1974; WOOD 1988). Several genes within *C. elegans* have been identified as essential for spermatogenesis or fertilization, including the

spermatogenesis defective (*spe*) gene *spe-42*. Homologous fertility genes of humans and *C. elegans*, such as *spe-42*, provide a way to infer methods of treatment for human infertility, specifically among males.

In the United States, approximately 7.4% of married couples are infertile with 50% of the cases due to male infertility (CHANDRA *et al.* 2005; HOFHERR *et al.* 2011). Because few of the molecular steps of fertilization have been identified, many cases of male infertility remain unexplained. As causes are determined, couples can make informed decisions concerning their chance of conception by means of the various infertility treatments. Each gametic contribution to the success of fertilization is complex and therefore must be considered individually. Focusing on male-factor infertility will help determine any possible male genetic mutations contributing to the overall infertility of the couple. Assays of the interactions prior to fertilization of sperm and egg in an appropriate environment would be useful. Thousands of dollars per couple in fertility treatment can potentially be saved if the cause of infertility is the result of an irreversible or untreatable genetic defect of the sperm. This could cut down on the nationwide infertility treatment cost to patients of nearly \$1.5 billion per year, much of which is not covered by standard medical insurance (COLLURA 2006).

Nematoda is one of the most numerous and diverse phyla of multicellular animals on earth. Of the more than 26,000 named nematodes, over 15,000 are parasitic (HUGOT *et al.* 2001). *C. elegans* is a non-pathogenic model organism

that can be used to effectively study nematode fertility, which in turn, can be used to control parasitic organisms. Although some methods controlling agricultural parasites are available, new strategies of control and prevention should be continually considered because old methods have failed as a result of harmful effects due to toxicity. Diseases associated with parasites cost billions of dollars yearly to both humans and crops. The study of *C. elegans* fertility may lead to possible targets for eradication of infectious parasites at the gametic level.

Caenorhabditis elegans: a simple model of the molecular mechanisms underlying fertilization

Caenorhabditis elegans, a 1 mm long free-living nematode, is used extensively as a model organism because it is genetically tractable, has a 3-day generation time and has a fully sequenced genome consisting of nearly 100 million base pairs (Wormbase.org WS226 release notes). Its transparent body allows the internal processes of *C. elegans* fertilization to be observed in live worms. The rapid *C. elegans* life cycle allows quick completion of complicated genetic crosses that would take years in other organisms, and the strains can be saved long-term by freezing young larvae at -80° C. It also has a simple and easily manipulated reproductive biology that makes it particularly suited for investigations into the molecular mechanisms of sperm-oocyte recognition, binding and fusion within the reproductive tract (GREENSTEIN 2005; SINGSON

2001; YAMAMOTO *et al.* 2006). One especially advantageous trait of *C. elegans* is that homozygous mutant males are more easily generated and maintained in populations than in other organisms because of the hermaphroditic form.

The main mode of *C. elegans* reproduction is through self-fertilization of the hermaphroditic form because hermaphrodites constitute 99.9% of the natural population (CHASNOV and CHOW 2002). However, cross progeny are also produced by the insemination of hermaphrodites with male-derived sperm that out-compete hermaphrodite-derived sperm resulting in preference for outcross progeny (LAMUNYON and WARD 1995; LAMUNYON and WARD 1998; WARD and CARREL 1979). The males have probably been retained in small numbers as a means to introduce genetic variation in the population. Production of males in a population of hermaphrodites is due to spontaneous loss of a sex chromosome during meiosis. Hermaphrodites have a total of 12 chromosomes with sex being determined by the presence of 2 (XX) sex chromosomes. The males have 11 chromosomes, one of which is its sex chromosome (XO). It is likely that male sperm out-compete hermaphrodite sperm as an evolutionary result of the rarity of males (MURRAY *et al.* 2011).

The hermaphrodite goes through definitive and recognizable life stages that make it possible to observe the time of oocyte fertilization. During the hermaphrodite life cycle, the sperm are made at the L4 stage of development and stored for later use. After production and storage of sperm, the

hermaphrodite ceases sperm production and switches to oocyte formation. Therefore, sperm limit each hermaphrodite's final progeny count. The hermaphrodite contains a bi-lobed gonad system that allows for fertilization from two routes (Figure 1).

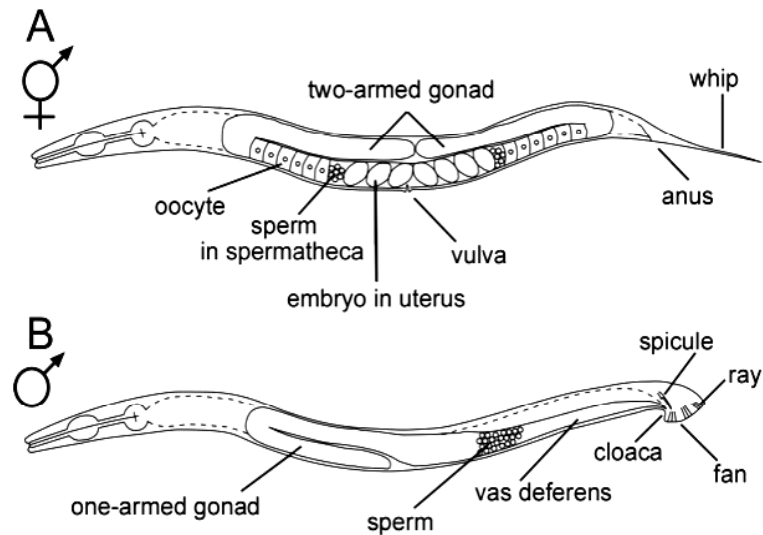


Figure 1. *C. elegans* anatomy. The hermaphrodite has a bi-lobed gonad that allows for fertilization from two directions. Males utilized their rays to locate the vulva prior to copulation. Sperm are transferred into the hermaphrodite through the spicules (modified from a figure by Tom Ratliff).

The oocytes are pushed through one of the two spermathecae, where the sperm are contained, and fertilized eggs are passed into a common uterus before being laid onto the growth plate. After copulation, the male-derived sperm will travel through the uterus to either spermatheca. Male-derived sperm typically increase the progeny from about 300 to up to 1000 or more. This is because the males continually provide more sperm through repeated mating (HODGKIN 1988).

If a healthy hermaphrodite self-fertilizes, almost all of their sperm fertilize oocytes. Multiple sperm are consistently pushed out of the spermatheca by oocytes headed to the uterus and are drawn back to the spermatheca for another chance by a polyunsaturated fatty acid-derived signal released from oocytes (KUBAGAWA *et al.* 2006). Major sperm protein (MSP) acts as an actin substitute in the pseudopod of *C. elegans* sperm enabling sperm to travel to the spermatheca (SEPSENWOL *et al.* 1989; THERIOT 1996). Oocyte meiotic maturation and ovulation is signaled by a MSP gradient progressing from the spermatheca into the nearest gonad caused by vesicles carrying MSP that bud from the sperm pseudopod. This gradient also stimulates contraction of gonad sheath cells leading to ovulation (CHENG *et al.* 2008; HARRIS *et al.* 2006; KOSINSKI *et al.* 2005; MILLER *et al.* 2001; MILLER *et al.* 2003).

Spe phenotype & coadjuvant proteins

The *C. elegans* spermatogenesis pathway is relatively simple compared to its mammalian counterpart, facilitating the isolation of spermatogenesis defective (*spe*) mutants at each step of the pathway. Germ cells develop gradually as they proceed from the distal to proximal end of the gonad. The cells remain connected to a common central canal called the rachis until primary spermatocytes bud off followed by differentiation and meiosis (Figure 2).

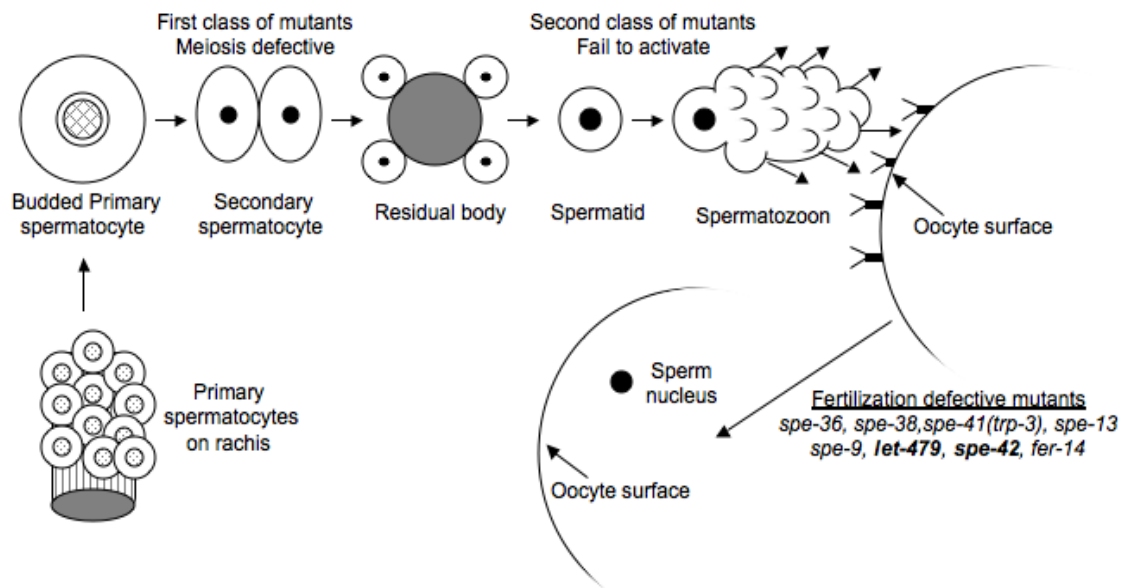


Figure 2. Spermatogenesis pathway. Germ cells develop as they proceed along the gonad. The cells remain connected to the rachis until primary spermatocytes bud off followed by differentiation and meiosis. *SPE-42*, like all *spe-9* class proteins, acts at the site where sperm and egg meet. The other two spermatogenesis defective mutant classes either impede morphogenesis into round spermatids or fail to activate and develop a pseudopod (modified from [L'HERNAULT 2006]).

The eight sperm-specific fertilization genes listed in this work are among the more than 60 genes that have been identified using a straightforward genetic screen for worms with mutations that affect spermatogenesis (L'HERNAULT 1997). Hermaphrodites with a *spe* mutation produce sterile sperm and lay only oocytes. However, their infertility can be rescued by wild type male sperm provided by a male. This shows that the *spe* mutation affects only sperm. The amoeboid spermatozoa of *C. elegans* lack the typical flagellated appearance of most mammalian spermatozoa, but the duties common to all spermatozoa such as migration to the fertilization site and species-specific recognition, binding and

fusion are all present (Figure 2). Unlike *C. elegans* oocytes, mouse and sea urchin eggs have a thick extracellular coat which sperm must bind to and pierce to reach the oocyte plasma membrane (BEMBENEK *et al.* 2007). The absence of this thick coat enables *C. elegans* to be used to identify later-acting mutants in the spermatogenesis pathway. This is a successful method of eliminating mutants that never reach the oocyte plasma membrane due to defects in sperm interactions with a thick outer egg coat (HE *et al.* 2003; STEIN *et al.* 2004). Sperm mutants can be divided into three general classes (Figure 2). The first class of mutants arrest development during meiosis and the organism makes few to no round spermatids (WARD and CARREL 1979). The second class of mutants affects sperm activation. The numbers of round spermatids produced are the same as wild type, but the spermatids do not activate to form pseudopods, and are incapable of fertilizing oocytes. The third class of sperm mutants, called *spe-9* class genes, affect the final stages of fertilization disrupting sperm signaling, recognition, oocyte binding or sperm-egg fusion.

spe-9 class genes

spe-9 was the first fertilization-defective gene to be cloned and, like the other *spe-9* class mutants, produces spermatozoa that are wild type in appearance and competitive behavior, but fail to fertilize oocytes despite direct contact with the oocyte plasma membrane (SINGSON *et al.* 1998). The original

spe-9 phenotypic traits are used as the comparative standard for all similar *spe-9* class genes. The *spe-9* class genes are determined by a series of morphological and behavioral tests. First, *spe-9* class mutants are sterile, but retain normal ovulation, male mating behavior and sperm competition ability (KROFT *et al.* 2005; SINGSON *et al.* 1999). When hermaphrodites are inseminated with male-derived sperm, the resultant progeny are largely outcross due to the larger, faster constitution of the male-derived sperm (LAMUNYON and WARD 1995; LAMUNYON and WARD 1998; WARD and CARREL 1979). The ability of the sperm to successfully fertilize oocytes appears not to be a factor in its competitive ability to take the place of the smaller hermaphrodite-derived sperm in the spermathecae (LAMUNYON and WARD 1998; WARD and CARREL 1979). For *spe-9* class genes, this is seen by a large decrease in progeny counts. Secondly, mutant sperm are not noticeably different from wild type sperm morphologically. Third, *spe-9* class mutant male sperm are able to travel to a hermaphrodite's spermatheca (KROFT *et al.* 2005; L'HERNAULT and SINGSON 2000). They are identified in genetic crosses by visual recognition of an unfertilized oocyte versus a fertilized egg (Figure 3). The *C. elegans* gene of focus, *spe-42*, is recognized as a *spe-9* class gene as are seven additional genes including *spe-9*, *spe-13*, *spe-36*, *spe-38*, *spe-41(trp-3)*, *let-479*, and *fer-14*.

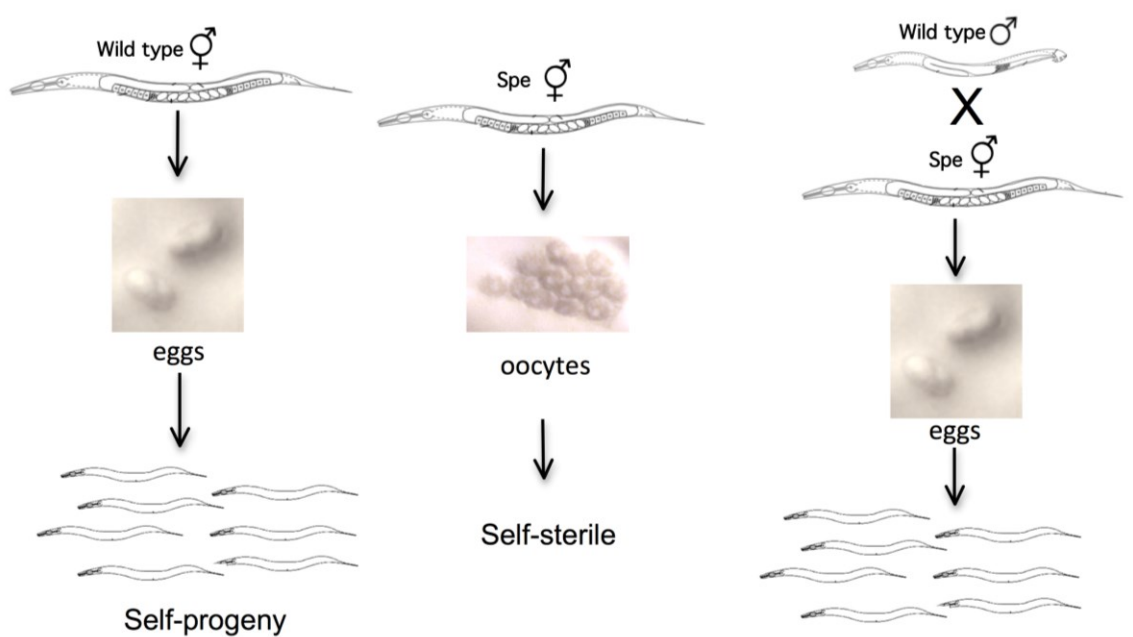


Figure 3. *spe-9* class phenotype. All spermatogenesis defective (*spe*) mutants result in unfertilized oocytes that often clump together. Wild type eggs are translucent with a clear outer shell while oocytes appear brown. These spermatogenesis defective mutants can be maintained by outcrossing with wild type males and picking heterozygotes.

All cloned *spe-9* class genes encode predicted transmembrane proteins and SPE-9 is a single pass transmembrane protein. It contains 10 epidermal growth factor (EGF) repeats essential for function and likely forms a ligand receptor pair with an oocyte protein. After binding receptors, EGFs promote cell growth, propagation, and differentiation (HERBST 2004). SPE-9 has been shown to localize to the sperm plasma membrane through the use of an antibody. It is probable that SPE-9 is a participant in egg adhesion and/or signaling steps of oocyte fertilization (ZANNONI *et al.* 2003). Two candidate oocyte receptors for SPE-9 were EGG-1 and EGG-2. However, *in vitro* and S2 culture tests indicate

the two do not have sufficient ability to bind alone and may require additional binding partners (SINGSON *et al.* 2008). Point mutations in the EGF-like repeats or deletion of the transmembrane domain of SPE-9 resulted in temperature sensitivity, reduced fertility or sterility suggesting that these regions of the protein are essential for its function. SPE-9 is thought to participate in signaling since very little of the protein is needed for successful fertilization (SINGSON *et al.* 2008).

In spermatids, proteins SPE-38 and SPE-41(TRP-3) both localize to the spermatozoa plasma membrane after being carried there by vesicles called membranous organelles (MOs) that are exclusive to nematode sperm (WARD *et al.* 1983). During spermatid activation, MOs fuse with the plasma membrane and release their contents onto the spermatozoon plasma membrane. SPE-38, like SPE-9, localizes to the pseudopod and is structurally comparable to mouse protein CD9 (cluster of differentiation) (CHATTERJEE *et al.* 2005). Mouse gametes do not undergo fusion when CD9, an oocyte cell surface molecule, is missing (KAJI *et al.* 2000). SPE-41, unlike SPE-38, remains spread over the entire plasma membrane following MO fusion (XU and STERNBERG 2003). Worms with a mutation in *spe-41*, also known as *trp-3* because it encodes a transient receptor potential canonical (TRPC) calcium channel, are not completely sterile. This shows that SPE-41/TRP-3 is not essential for fertilization but its absence likely disrupts the calcium influx needed to achieve sperm-oocyte fusion

(JUNGNICKEL *et al.* 2001; JUNGNICKEL *et al.* 2003). Another possibility is that an essential interaction between SPE-38 and SPE-41 is being interrupted since SPE-38 function is needed for translocation of SPE-41/TRP-3 to and from fused MOs to the plasma membrane (GUNASEKARAN *et al.* 2012).

The genes *spe-13* and *spe-36* are not yet cloned, but interestingly all known SPE-13 mutant alleles are temperature sensitive, which suggests that the available alleles do not represent the null phenotype or that the process in which SPE-13 functions is inherently temperature sensitive (SINGSON *et al.* 2008; WARD and MIWA 1978). Little is known about *spe-36*. FER-14 is predicted to be a novel single pass transmembrane protein like SPE-9 and may also be involved in a ligand receptor pair interaction between sperm and eggs based on predicted membrane topology (our unpublished data).

spe-42 and let-479: spe-9 class genes

The final two known *spe-9* class genes are *spe-42* and *let-479*. The homologous SPE-42 and LET-479 proteins (Table 1 and Figure 4) are believed to be involved in sperm-egg recognition, binding, or fusion. *spe-42*, located on chromosome V, is expressed only in the male germline as shown from northern blot data (KROFT *et al.* 2005). *spe-42* was added to the list of *spe-9* class proteins after the following criterion were met: (1) the sperm appear morphologically identical to wild type sperm; (2) male-derived sperm are able to

travel to the site of fertilization in hermaphrodites; (3) mutants are sterile despite normal ovulation, male mating, and sperm competition. (KROFT *et al.* 2005).

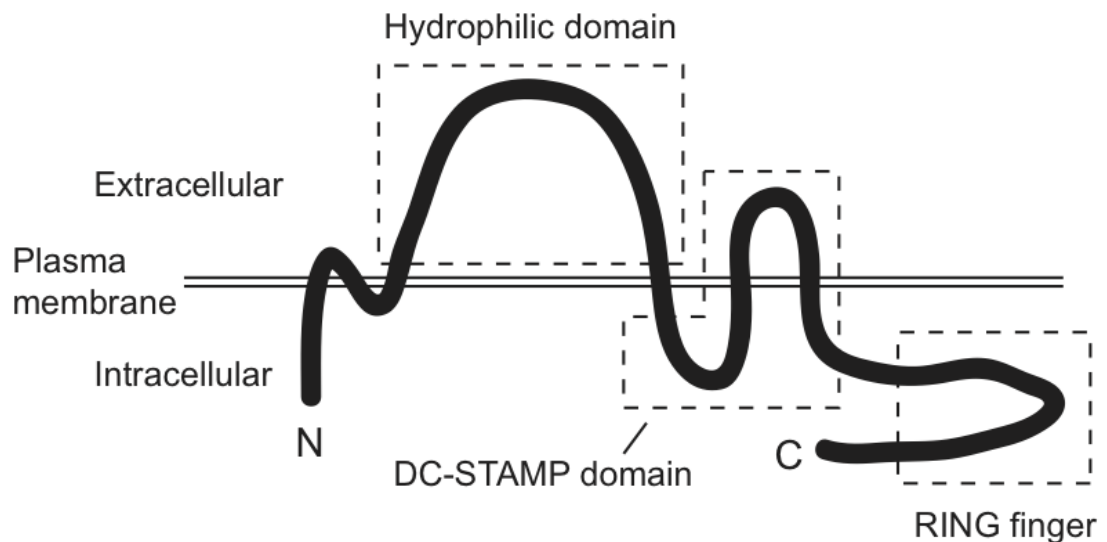


Figure 4. Predicted SPE-42 structure with notable domains. SPE-42 is a predicted 6-pass transmembrane DC-STAMP domain-containing protein with 2 helical passes within the DC-STAMP domain. The RING-finger domain is contained completely within the cytosol while the large hydrophilic extracellular domain is entirely outside the cell. The figure is stretched linearly, but the transmembrane protein helices may interact with each other, influencing protein structure.

While much is yet to be discovered for the structure of SPE-42 and the recently cloned, LET-479, both proteins have valid predicted membrane topologies (Figure 4). SPE-42 is predicted to be a six-pass transmembrane protein based on the results from 11 computer algorithms (Appendix: Table 4), completed by Dr. Jon Rumbley (BERNSEL *et al.* 2008; CSERZO *et al.* 2004;

GANAPATHIRAJU *et al.* 2008; JONES 2007; JURETIC *et al.* 2002; KALL *et al.* 2004; KALL *et al.* 2005; SHEN and CHOU 2008; TUSNADY and SIMON 2001; VIKLUND and ELOFSSON 2004; VIKLUND and ELOFSSON 2008). Most of these algorithm programs were developed specifically for prediction of transmembrane proteins, membrane topology, and helical components. A few programs like OCTOPUS have shown success in transmembrane protein prediction and incorporate additional data such as secondary structures to determine topology. The program OCTOPUS uses the amino acid sequence and incorporates data from homologous alignment results and hydrophobicity values to assess possible residue representations. The capabilities and preferences of each residue are considered and various locations inside, outside or spanning the membrane of the cell are evaluated based on the amino acid environment. Finally, a procedure including all the gathered information is used to calculate a final topology. The 11 programs used for the SPE-42 topology prediction employed comparable methods to OCTOPUS and a consensus of the data produced the final prediction. The membrane topology prediction shows the protein areas that span the membrane and which domains are likely to be inside or outside the cell.

SPE-42 is predicted to have a minimum of three distinct functional domains: a large extracellular loop, a DC-STAMP (dendritic cell-specific transmembrane protein) domain and a C-terminal cytoplasmic domain containing a RING finger (Figure 4). The N- and C-termini are predicted to be within the

cytosol. Initial experimentation suggests that certain cysteines in both the C-terminal domain and the extracellular loop are essential for the function of SPE-42. This work has significant implications for understanding sperm-egg interactions in other species because SPE-42 homologs are present in all metazoan species with sequenced genomes, including humans.

A homology structure of the RING-finger domain was recently completed (WILSON *et al.* 2011), and the LET-479 RING is likely similar in structure to the SPE-42 RING finger. The predicted membrane topology of SPE-42 and LET-479 are identical with respect to the relative location of (1) transmembrane domains, (2) the extracellular hydrophilic domain, (3) the DC-STAMP domain, and (4) the RING domain.

A RING finger is located in the SPE-42 C-terminal domain

The SPE-42 C-terminal cytoplasmic domain contains eight conserved cysteines that are predicted to form a RING finger (Figure 4). Mutation of any of these cysteines results in up to a >90% reduction in fertility compared to wild type controls, suggesting the RING finger is functional (WILSON *et al.* 2011). *C. elegans* shares common ancestry with another, reproductively isolated, nematode, *C. briggsae* (STEIN *et al.* 2003). The *C. briggsae* genome encodes a homolog to *C. elegans* SPE-42 that shows strikingly low divergence from the *C. elegans* protein (85% identity and 93% similarity) (KROFT *et al.* 2005). Despite

the high conservation of amino acid sequence between the two proteins, when transgenes that express the *C. briggsae spe-42* gene were crossed into *spe-42* mutants, the level of rescue was very poor suggesting the two proteins are not wholly interchangeable. The most significant divergence in the protein sequences is in the final 29 amino acids of both homologs, located C-terminal to the RING finger (WILSON *et al.* 2011). To test the importance of this sequence for protein function, constructs were made in which the final 29 amino acids of the *C. elegans* protein were substituted for the respective *C. briggsae* sequence and vice versa. These transgenes were crossed into *C. elegans* with a *spe-42* mutation. Replacement of the *C. briggsae* final 29 amino acids with *C. elegans* sequence did not significantly improve the protein's ability to rescue the *spe-42* sterility defect. Likewise, replacement of the *C. elegans* final 29 amino acids with *C. briggsae* sequence did not significantly reduce rescuing ability of the transgene (WILSON *et al.* 2011). These data suggested that the last 29-30 amino acids of these proteins are either not required for SPE-42 to function or that the conserved amino acids located in this area are adequate for SPE-42 to function. To test the first hypothesis, a construct was made in which the final 29 amino acids of SPE-42 were deleted. The resulting transgenic worms displayed wild type numbers of progeny following a cross into a *spe-42* mutant background, demonstrating that these amino acids are expendable (our unpublished data).

All RING domain proteins that have an experimentally determined function are E3 ubiquitin ligases. E3 ligases tag specific proteins for recycling or degradation. E3 ligases are the final proteins in a pathway involving E1 ubiquitin-activating enzymes (E1) and E2 ubiquitin-conjugating enzymes (E2). E1 activates the cascade and E2 interacts with a specific E3 ligase to pass along the ubiquitin to a target protein (DESHAIES and JOAZEIRO 2009; HUIBREGTSE *et al.* 1995; YIN *et al.* 2009). Although SPE-42 is dissimilar to other RING finger proteins, the fact that it contains this domain leaves open the possibility that it functions as an E3 ligase (WILSON *et al.* 2011).

SPE-42 extracellular hydrophilic domain

Our topology prediction shows a large hydrophilic extracellular domain located between the third and fourth transmembrane helices of SPE-42 (Figure 4). Six conserved cysteines, with 9-13 amino acids between them, have been identified within the large extracellular domain. Although this domain contains no predicted protein motifs, all 6 of the cysteines are required for protein function (our unpublished data). This domain is predicted to be outside the cell and available for interaction with egg proteins on the oocyte surface. Changing any of these cysteines (C) to alanine (A) drastically reduced function of SPE-42 in transgenic rescue experiments, suggesting they are required for protein function.

SPE-42 DC-STAMP domain

The DC-STAMP domain was named for the DC-STAMP (dendritic cell-specific transmembrane protein) protein in which it was initially described. The SPE-42 DC-STAMP domain includes the fifth and sixth transmembrane helices. Because these helices span the plasma membrane, portions of the DC-STAMP domain are cytoplasmic and other portions are extracellular (Figure 4). Although the DC-STAMP protein was discovered in dendritic cells of the immune system, it has been shown to be essential for bone homeostasis. Bone tissue contains specialized cells called osteoclasts (OCs) and osteoblasts. Osteoclasts break down bone matrix in a carefully balanced relationship with osteoblasts, which build up bone matrix to maintain a healthy bone density. Defects in osteoclastogenesis lead to osteopetrosis, the build-up of bone, and defects in osteoblastogenesis lead to bone loss or osteoporosis. Osteoporosis is common among women after menopause as a result of lower estrogen levels. OCs are inhibited by estrogen and, as estrogen stores lower with age, OCs are less inhibited, and bones become less dense (OURSLER 1994; OURSLER *et al.* 1993). DC-STAMP acts when preosteoclasts (POCs) fuse to become mature multinucleate osteoclasts (Figure 5). The absence of DC-STAMP protein results in complete loss of POC fusion in knockout (KO) mice (YAGI *et al.* 2005). DC-STAMP KO mice suffer osteopetrosis because the mononuclear POCs do a poor job of resorbing bone compared to multinuclear OCs (YAGI *et al.* 2005).

Macrophages and preosteoclasts originate from the same cell lineage. DC-STAMP is also necessary for macrophages to fuse into foreign body giant cells (FBGCs). A FBGC forms as a result of the presence of a foreign body that is too large for a macrophage to engulf (Figure 5). DC-STAMP is believed to be a membrane-bound ligand involved directly with fusion or it may only act in the signaling pathway to a yet unknown fusion molecule (MENSAH *et al.* 2010).

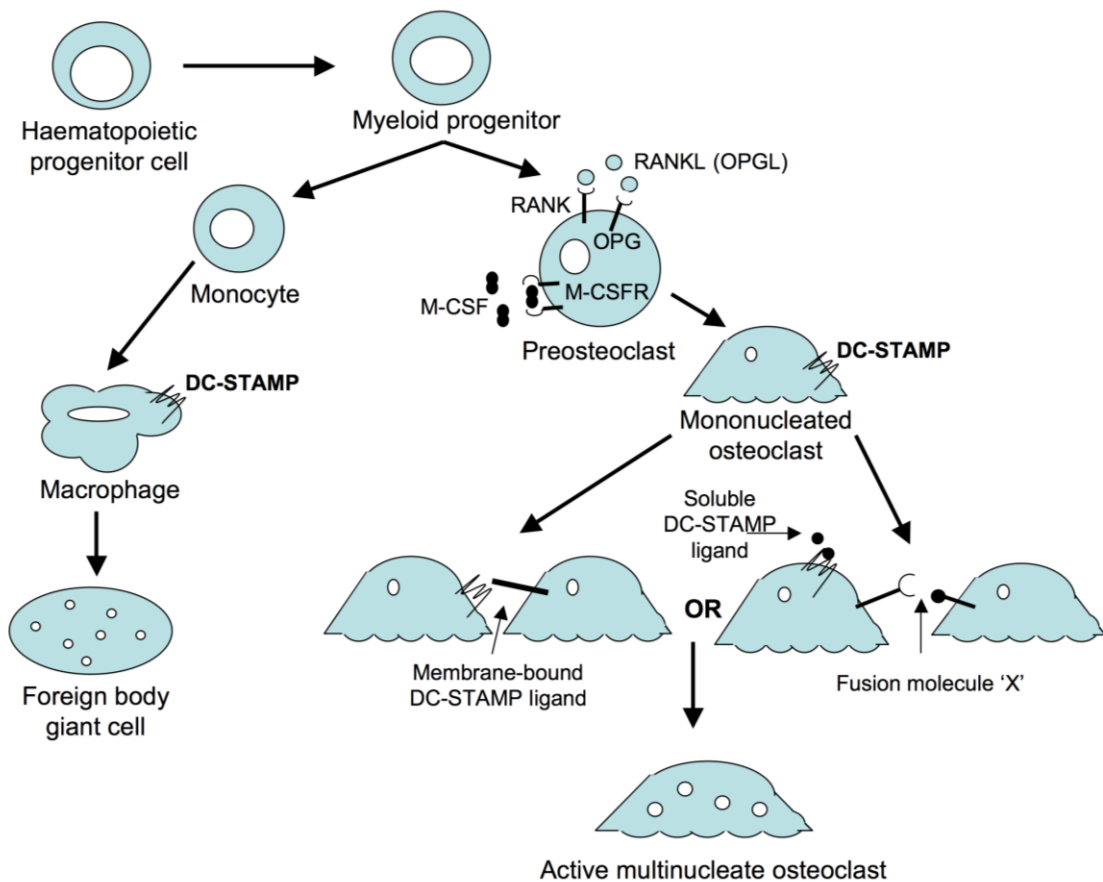


Figure 5. Projected functions of DC-STAMP. DC-STAMP is necessary for macrophages and osteoclast precursors (POCs) to fuse into foreign body giant cells and multinuclear osteoclasts, respectively. Myeloid progenitor cells differentiate into POCs in the presence of M-CSF and RANKL. M-CSF binds to its receptor, M-CSFR while RANKL binds to either RANK or OPG to maintain osteoclast levels. The resulting mononucleated osteoclasts fuse into active multinucleate osteoclasts through the active regulation of DC-STAMP. DC-STAMP is thought to act as a receptor to either a soluble ligand or a ligand bound in adjacent osteoclasts. This either directly initiates fusion or promotes the actual fusion 'X' molecule(s) to activate (MENSAH *et al.* 2010). (Figure modified from [VIGNERY 2005]).

Osteoclast function is regulated by tumor necrosis factor (TNF) and receptor (TNFR)-like proteins: osteoprotegerin (OPG), receptor activator of nuclear factor (NF)- κ B (RANK), and RANK ligand (RANKL) in addition to

secreted cytokine macrophage-colony stimulating factor (M-CSF). M-CSF binds to macrophage-colony stimulating factor receptor (M-CSFR) to influence the hematopoietic cells to differentiate into appropriate cell types (Figure 5). TNF-alpha is involved in regulation of immune cells and therefore it is not surprising that RANKL (a TNF-like ligand) induces the formation of OCs since OCs originate in bone marrow (LAM *et al.* 2000). RANKL is also known as osteoclast differentiation factor (ODF), osteoprotegerin ligand (OPGL) and [TNF]-related activation-induced cytokine (TRANCE). RANKL's many names indicate its importance in generating POCs. Osteoprotegerin (OPG) is a diversionary TNFR receptor on POCs that competes with RANK for RANKL. The purpose of the diversion is to aid in building up bone density by decreasing the production of OCs. If RANKL is able to bind to its receptor, RANK, on osteoclasts, OCs are activated. These molecules are the main players in regulation of osteoclast action (SUDA *et al.* 1999). RANK is also involved in communication between dendritic cells and T cells due to its presence on the surface of dendritic cells (THEILL *et al.* 2002).

Macrophages expressing the DC-STAMP protein fuse into FBGCs to defend against foreign material in the body. OCs and FBGCs both engulf other particles and arise from hematopoietic cells despite their different tasks in the body. As seen in OCs, the absence of DC-STAMP protein in KO mice prevents macrophage fusion into multinucleate FBGCs. DC-STAMP is highly expressed

in OCs, but not highly expressed in macrophages, seemingly contradicting the notion that DC-STAMP is the essential fusion factor. However, foreign body giant cell formation by macrophage fusion is negatively affected by the loss of DC-STAMP (YAGI *et al.* 2005). Recently it has been shown that loss of DC-STAMP also inhibits formation of Langhans giant cells, which are multinucleate cells formed by the fusion of macrophages (MURPHY *et al.* 2007; SAKAI 2012). Langhans giant cells appear in granulomatous conditions such as tuberculosis (SAKAI 2012).

C. elegans SPE-42 shares a homologous domain with the original DC-STAMP protein described above. SPE-42 also acts at a time and place where a plasma membrane fusion event takes place, the fusion between sperm and egg to form a zygote. The presence of a DC-STAMP domain in both proteins and the analogous time and place of protein action suggests that the DC-STAMP region may be directly involved with the fusion event at the molecular level. DC-STAMP is believed to be a surface receptor required for OCP fusion and therefore SPE-42 may act as a surface receptor for sperm egg fusion.

The SPE-42 protein family

A BLAST search uncovered SPE-42 homologs in a diverse group of metazoans, from mosquitoes to humans. Each organism has two SPE-42 homologs (discussed in detail in chapter 2) (ALTSCHUL *et al.* 1990; KROFT *et al.*

2005), and *C. elegans* itself has a homolog, *let-479*, present in its theoretical genome (our unpublished data). In contrast, plant species with fully sequenced genomes have no obvious homologs to SPE-42, LET-479 or other proteins with a DC-STAMP domain. These data show that SPE-42 and its homologs are only present in fertilization events involving metazoan sperm, which suggests the possibility that these proteins evolved along with the process of sexual reproduction. One of the SPE-42 homologs is a protein in *Drosophila* called Snky (sneaky). This protein is more similar to LET-479, and the second (unidentified) homologous protein in *Drosophila* is believed to be the functional homolog to SPE-42. In *Drosophila* fertilization, the egg engulfs sperm with the sperm plasma membrane still intact. Only after the sperm is within the egg, does sperm plasma membrane breakdown (PMBD) occur. Snky is an acrosomal membrane protein that facilitates PMBD, but leads to male sterility when mutated because the sperm plasma membrane fails to break down (WILSON *et al.* 2006). As mentioned earlier, several conserved cysteines in *C. elegans* extracellular domain that were mutated led to a non-functional protein. When the second cysteine of Snky's large extracellular domain is mutated, PMBD fails. The role that Snky plays in fertilization is a strong statement that all the homologous proteins in other species are involved in this process (WILSON *et al.* 2006).

Chapter Two: Sperm protein, SPE-42, and the DC-STAMP domain

Introduction

Sperm and egg combine during fertilization in a series of steps involving recognition, binding, and fusion to form an embryo. *C. elegans* has emerged as a useful model organism for the discovery of proteins involved in the molecular interactions that take place during fertilization. The *C. elegans* self-fertile reproductive biology is an effective system for identification of genes necessary for fertilization, potentially elucidating the molecular interactions that occur between gametes. Hermaphrodites produce both sperm and eggs, yet the small percentage of males in the population can introduce new genetic information via copulation. The ability of hermaphrodites to self-fertilize contributes to the easy identification of sperm mutations. Hermaphrodites with defective sperm lay only oocytes, but wild type sperm from males can rescue their sterility. This demonstrates that the *spe* mutations affect only sperm (L'HERNAULT 1997).

While *C. elegans* amoeboid spermatozoa have a different appearance than most mammalian sperm, they still must perform the same actions for fertilization to succeed. These shared actions include traveling to the site of fertilization, recognizing the oocyte in a manner, and species-specific fusion into a zygote.

In this work, the focus is on the SPE-42 DC-STAMP domain, which is common among many proteins and important for cell fusion, although the

molecular mechanism of action is not known. We used a mutagenesis approach to determine essential portions of the DC-STAMP domain for SPE-42 function based on its conservation between several homologous proteins. We aimed to identify specific amino acids within the DC-STAMP domain that are required for protein function. We identified candidates for mutation based on conservation among DC-STAMP proteins and the specific properties of each amino acid. Lastly, we asked whether DC-STAMP domains from the different proteins could substitute for each other. Each altered protein was then expressed in a self-sterile *spe-42* mutant strain and the ability of each transgene to rescue SPE-42 function was determined. This allowed us to measure the importance of individual amino acids or larger scale substitutions for SPE-42 function based on the number of progeny compared with wildtype transgenic lines.

Materials and Methods

Sequence alignments

The protein sequences of SPE-42 and LET-479 were from cDNAs ((KROFT *et al.* 2005);our unpublished data). SPE-42 and LET-479 were BLAST (Basic Local Alignment Search Tool) searched for similar sequences (ALTSCHUL *et al.* 1990; KROFT *et al.* 2005). Nine different common organisms with proteins that displayed homology to SPE-42 and contained DC-STAMP regions were chosen for comparison. The sequences were aligned using PROMALS (PROfile Multiple

Alignment with Local Structure; (PEI and GRISHIN 2007) and the portion of each sequence that aligned with the DC-STAMP domain of SPE-42 was reentered into PROMALS for a direct comparison of the DC-STAMP regions. The amino acids within the DC-STAMP regions were analyzed for high conservation among species. Dr. Jon Rumbley directly compared SPE-42 and LET-479 to note differences. These data as well as the possible structural effects on the protein were used to compile a list of possible essential amino acids for mutation.

Worm strains and handling

Worm culture and genetic manipulations were performed according to standard methods, and wild type was Bristol N2 (BRENNER 1974). All mutations are in the N2 genetic background. Growth temperature was 20° C unless described otherwise. The following mutant alleles, markers and genetic balancers were used: *him-8(e1489)* IV (HODGKIN *et al.* 1979); *nT1[qIs51]* (IV;V) (FERGUSON and HORVITZ 1985); *mIs10* V (K. LIU and A. FIRE, personal communication), *spe-42(tm2421)* V (S. MITANI, personal communication); *spe-42(tn1231)* V (KROFT *et al.* 2005)).

Plasmid DNA constructs

A *spe-42* wild type plasmid construct, pTK15, was subjected to site-directed mutagenesis to make changes that lead to amino acid substitutions in the SPE-42 DC-STAMP domain. The template, pTK15, includes 1409 bp of 5' promoter sequence and 265 bp of 3' UTR sequence in addition to the *spe-42* coding region and contains no genes other than *spe-42* (WILSON et al. 2011). In separate experiments, seven individual amino acids and one group of three were substituted with a preferred *C. elegans* codon to produce the new amino acid or group. Using *C. elegans* preferred codons ensures strong protein expression (SHARP and BRADNAM 1997; STENICO *et al.* 1994).

Quickchange site-directed mutagenesis (Agilent Technologies) was used according to the manufacturer's instructions to create 6 of the 7 single amino acids substitutions except Phusion polymerase (Finnzymes) was used instead of Pfu Turbo DNA polymerase. The oligonucleotide primers shown in Appendix, Table 5 and template pTK15 were utilized according to manufacturer's protocol. Primer annealing conditions were set based on T_m values of each oligonucleotide pair. The Quickchange method was not successful in generating the desired amino acid substitution for the final single amino acid substitution R615A. As a result, this amino acid substitution and the three amino acid mutation, FFP 624-626 AAG were generated using an overlap PCR strategy (Appendix: Figure 10). Both substitutions involved three PCR reactions. For amino acid substitution

R615A: PCR I, and PCR II used pTK15 (*C. elegans spe-42* genomic DNA) as template. Primers TK508 / TK257 (Appendix: Figure 10: A, C) were used for PCR I, and primers TK509 / TK243 (Appendix: Figure 10: B, D) were used for PCR II. After gel purification, these products were used as template for PCR III with primers TK257 and TK243 (Appendix: Figure 10: A, D). For amino acid substitutions FFP 624-626 AAG: The same overlap extension PCR procedure with the steps just described was run using primers TK510 / TK257 (Appendix: Figure 10: A, C) in PCR I and TK511 / TK243 (Appendix: Figure 10: B, D) in PCR II. The product was sequenced to confirm the presence of the FFP 624-626 AAG mutation. The two gel purified constructs were cloned separately into the XhoI and XbaI sites in pTK15, replacing the wild type sequence. Ligated DNA and plasmids from the site-directed mutagenesis were transformed using either heat-shock at 42°C or electroporation into *E coli* cells followed by growth on LB-ampicillin plates at 37°C). Individual colonies were grown up in liquid cultures, DNA was isolated (Wizard Plus SV Minipreps, Promega), and sequencing was performed to confirm the presence of the desired mutations.

To splice the *C. elegans let-479* DC-STAMP region into *C. elegans spe-42* and vice versa, another overlap PCR strategy was used. Genomic DNA was used in this case to minimize the loss of any regulatory regions prior to splicing. First, a PCR reaction with SPE-42 genomic DNA (1 µl of 1:50 dilution of pTK15 plasmid miniprep DNA) and primers TK257 and TK626 (Appendix: Figure 11: A,

C) was run to amplify one attachment site for the insertion. Next, a PCR reaction with LET-479, pTK113 genomic DNA (1 µl of 1:20 dilution of pTK13 plasmid miniprep DNA), and primers TK625 and TK628 (Appendix: Figure 11: B, E) was run to amplify the LET-479 DC-STAMP region. Finally, a PCR reaction with SPE-42 genomic DNA (1 µl to 1:50 dilution of pTK15 DNA) and primers TK627 and TK243 (Appendix: Figure 11: D, F) was run to amplify the second attachment site for the insertion. The three products were run on a 1% agarose gel and purified (Wizard SV Gel and PCR Clean-up System, Promega). These products (26ng, 87ng, and 25 ng respectively) were used as template in a fourth PCR reaction with primers TK257 and TK243; these primers were added after the first 5 cycles to ensure the desired product (Appendix: Figure 11: A, F). The gel-purified product from the fourth PCR reaction and pTK15 were digested with restriction enzymes AgeI and NdeI and gel purified again. Ligation of the LET-479 DC-STAMP insert and pTK15 was performed with T4 DNA ligase and transformed into *E. coli*. The resultant plasmid, pTK109, contains a 1056 bp genomic DNA fragment of *C. elegans let-479* that encodes the DC-STAMP domain flanked by the 5' and 3' portions of *spe-42* genomic DNA.

To splice SPE-42 DC-STAMP region into LET-479, the same overlap PCR strategy was used. First, a PCR reaction with LET-479 genomic DNA (1 µl of 1:50 dilution of pTK113 miniprep DNA) and primers TK257 and TK657 (Appendix: Figure 11: A, C) was run to amplify one attachment site for the

insertion. Next, a PCR reaction with SPE-42, pTK15 genomic DNA (1 µl of 1:20 dilution of pTK15 plasmid miniprep DNA) and primers TK656 and TK659 (Appendix: Figure 11: B, E) was run to amplify the SPE-42 DC-STAMP region. Finally, a PCR reaction with LET-479 genomic DNA (1 µl to 1:50 dilution of pTK113 DNA) and primers TK658 and TK243 (Appendix: Figure 11: D, F) was run to amplify the second attachment site for the insertion. The three products were run on a 1% agarose gel and purified (Wizard SV Gel and PCR Clean-up System, Promega). The products of each of these were used as template in a fourth PCR reaction with primers TK257 and TK243; these primers were added after the first 5 cycles to ensure the desired product (Appendix: Figure 11: A, F). The gel-purified product from the fourth PCR reaction and pTK113 were digested with restriction enzymes NruI and NsiI and gel purified. Ligation of the SPE-42 DC-STAMP insert and pTK113 was performed with T4 DNA ligase and transformed into *E. coli*. The resultant plasmid, pTK119, contains a 789 bp genomic DNA fragment of *C. elegans spe-42* that encodes the DC-STAMP domain flanked by the 5' and 3' portions of *let-479* genomic DNA (Appendix: Figure 10).

Transgenic C. elegans lines

Extrachromosomal transgenic arrays were created by co-injecting 100 ng/µl of marker plasmid pPD118.20 (*myo-3::GFP*, 1997 Fire Lab Vector Kit) with

each *spe-42* DNA construct (10 ng/μl) into gravid Bristol N2 (BRENNER 1974) wild type hermaphrodites using standard methods (MELLO and FIRE 1995; MELLO *et al.* 1991). The pPD118.20 plasmid drives GFP expression in body wall muscle under the control of the *myo-3* promoter (MELLO and FIRE 1995). Transgenic lines were followed for three generations to ensure stability of transgene inheritance. At least three independently derived lines were obtained for each transgenic construct to control for protein expression variability typical in extrachromosomal arrays. Only lines in which all worms appeared healthy were used for transgenic rescue experiments.

Transgenic rescue experiments

Males from strain SL968 *him-8(e1989) IV; mIs10 V* (integrated *myo-2::GFP*, which expresses GFP in pharynx muscle) were mated with hermaphrodites from each transgenic line to be analyzed. Mates consisted of 12 L4 or adult males and 3 L4 stage hermaphrodites (Appendix: Figure 12). Males bearing both pharynx (*myo-2::GFP*) and body wall (*myo-3::GFP*) GFP were crossed into *spe-42 (tm2421 or tn1231)* hermaphrodites, which are both completely sterile. Hermaphrodites from this cross with both pharynx and body wall GFP (*spe-42/mIs10 V; xyEx* transgene) were allowed to self-fertilize at 25°C. Progeny with only body wall GFP (*spe-42 (tm2421 or tn1231) V* homozygotes plus the *spe-42* transgene) were picked onto individual plates at 25°C and

progeny were counted to determine the effect of the mutation on fertility. Each hermaphrodite was transferred to a new plate daily and maintained at 25°C. Counts were continued until the first day each hermaphrodite produced no live progeny.

Results

SPE-42 and LET-479 protein sequence were run through BLAST (ALTSCHUL *et al.* 1990), PROMALS (PEI and GRISHIN 2007) and GeneDoc (NICHOLAS and NICHOLAS 1997) to determine the presence of proteins or protein regions of sequenced metazoans with high homology (Figure 6 and Table 1). The highly conserved amino acids among homologous SPE-42 proteins indicate fundamental amino acids within many metazoans including humans. The sequence alignment was used to select candidates for mutagenesis studies and future targets of study. *C. elegans* LET-479 and the rest of the mammalian proteins shown are postulated homologous proteins to *C. elegans* SPE-42 with similar sequence in the DC-STAMP domain. The level of homology within the DC-STAMP domain of the similar proteins is indicated by the similarity shown among the proteins compared. We searched the proteomes of plant and animal species without a *bona fide* sperm and egg mode of fertilization for SPE-42 and LET-479 homologs but found none. There was no significant homology to SPE-

42 or LET-479 proteins in the algae, *Chlamydomonas reinhardtii* or the protozoan, *Dictyostelium discoideum* among others (data not shown).

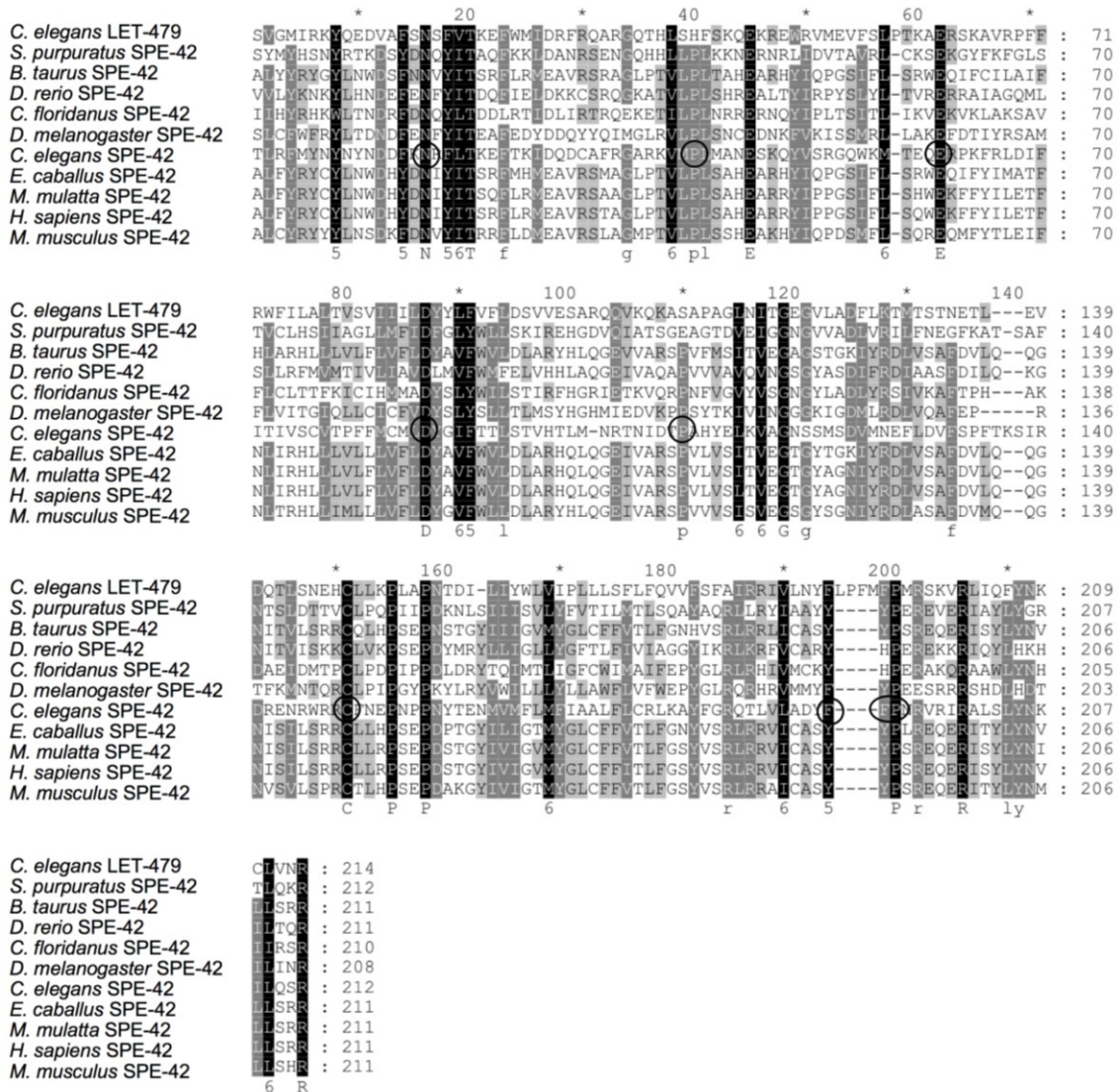


Figure 6. SPE-42 DC-STAMP domain-containing homologous sequence alignment. The highly conserved amino acids (black / grey) indicate possible essential amino acids within many metazoans including humans. This sequence alignment was used to distinguish candidates for mutagenesis studies and future targets of study. Circled amino acids are those substituted in this study. The first protein listed is *C. elegans* LET-479 and the rest of the mammalian proteins shown are postulated homologous proteins to *C. elegans* SPE-42 with similar sequence in the DC-STAMP domain. In order shown: *S. purpuratus*, *B. tuarus*, *D. rerio*, *C. flordanus*, *D. melanogaster*, *C. elegans*, *E. caballus*, *M. mulatta*, *H. sapien*, and *M. musculus*.

Table 1. SPE-42 DC-STAMP homology

<u>Organism</u>	<u>Similarity</u>	<u>Identity</u>
<i>Drosophila melanogaster</i> (fruit fly)	47%	21%
<i>Strongylocentrotus purpuratus</i> (sea urchin)	47%	19%
<i>Homo sapiens</i> (human)	46%	23%
<i>Macaca mulatta</i> (rhesus monkey)	46%	23%
<i>Mus musculus</i> (mouse)	45%	21%
<i>Danio Rerio</i> (zebrafish)	44%	21%
<i>Bos taurus</i> (cow)	44%	23%
<i>Equus caballus</i> (horse)	43%	21%
<i>Camponotus floridanus</i> (ant)	42%	22%
<i>Caenorhabditis elegans let-479</i> (nematode)	41%	21%

Amino acid substitutions

Amino acids were chosen for substitution on the basis of conservation between SPE-42 homologs and predicted importance for protein function or structure (Figure 6). Three of the conserved amino acids chosen for mutation carry an electrostatic charge: aspartic acid (D518), glutamic acid (E493), and arginine (R615). These were chosen because the charge suggests each of these may be necessary in either an active site or in a protein-protein electrostatic interaction. The one noticeably conserved cysteine (C581) in the DC-STAMP domain was chosen because cysteines are often crucial for disulfide bond interactions or possibly for metal binding. While there are no cysteines nearby within the DC-STAMP region that could form a disulfide bond, the many cysteines within the extracellular loop are possible bonding partners. Another

possibility is a cysteine-containing ligand or partner in a complex that may bind the region. The unique location of the asparagine (N448) and the fact that it is nearly 100% conserved across the species made it an interesting target for mutation. The side-chain of asparagine can form hydrogen bonds with the peptide backbone and therefore it is often located at the beginning or end of alpha helices, which may be the case here. For these reasons, this asparagine (N448) was chosen to mutate, as its location indicated a possible conformational role, which might explain its high conservation. Toward the C-terminal end of the DC-STAMP region, there is a three amino acid long conserved polypeptide consisting of two phenylalanines (F) and a proline (P). The location of these three amino acids, curiously similar among many species (Figure 6), was a possibility for a hydrophobic pocket or contact surface necessary for protein-protein bonding. Two additional single conserved prolines (P471 and P540) were substituted with glycines (G) because they likely play an important role in protein structure. Proline is rigid as a result of its cyclic side chain and may only assume a limited number of angles and conformations whereas glycine is the most versatile of the amino acids and allows for many configurations. Proline is rarely present in alpha helices and beta sheets because it prefers to fold or bend its side chain, which is capable of only one hydrogen (accepting) bond. Proline is frequently found at the start of a helix or along the edges of beta sheets.

Effects of amino acid substitutions on fertility

A homozygous *spe-42* mutant hermaphrodite lays unfertilized oocytes (Figure 7). However, wild type SPE-42 protein expressed in a transgenic hermaphrodite is capable of rescuing function in a *spe-42* mutant genetic background hermaphrodite. Therefore, if the substitution of an amino acid(s) disrupt(s) the function of *spe-42* in the extrachromosomal array, it is assumed that the protein fails at some point in its operations. The number of offspring produced in comparison to wild type determines the fundamental nature of the amino acid to the protein. If the transgene fails to rescue, the hermaphrodite will lay unfertilized oocytes and the substituted amino acid is considered essential (Figure 7).

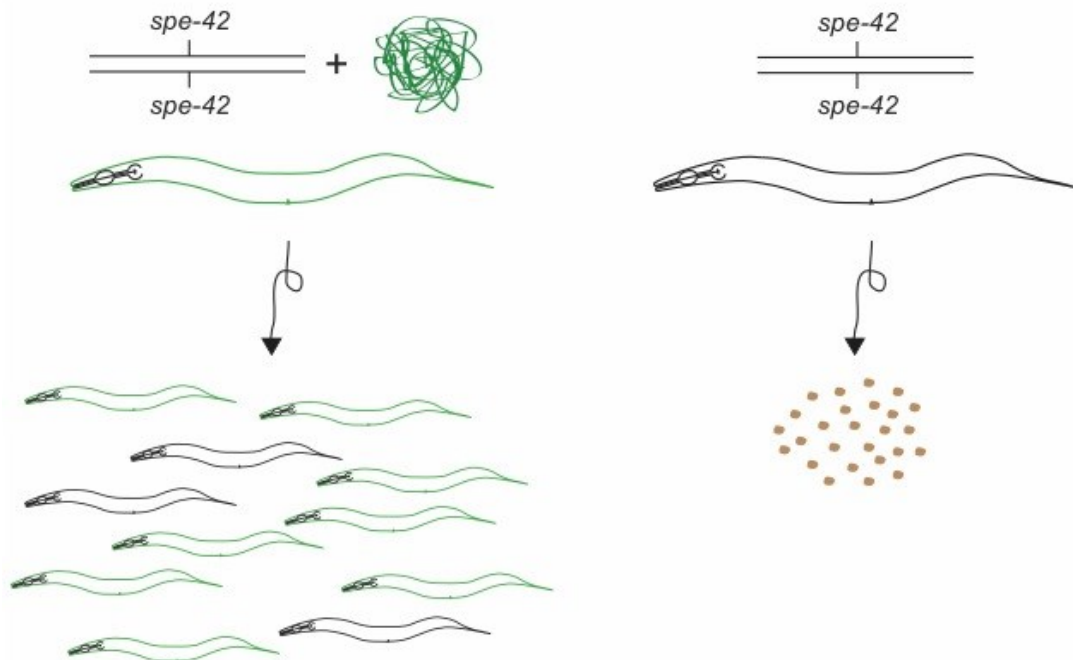


Figure 7. *spe-42* extrachromosomal array. Each nematode was injected with a plasmid containing GFP and a second plasmid containing SPE-42 with the mutation of interest. The two plasmids come together to form an extra ‘chromosome’ made up of several hundred repeats of the mutated GFP tagged gene. *spe-42* mutants lay unfertilized oocytes; they also lay unfertilized oocytes if a transgene fails to rescue.

The substitution mutations of asparagine (N448) and cysteine (C581) to alanine (A) showed a small reduction in fertility (~62% rescue; Table 2, Figure 8) compared to wild type transgenes, whereas the substitution of proline (P472) to glycine (G) had a drastic effect on fertility (0% rescue; Table 2, Figure 8). Glutamic acid (E493) and aspartic acid (D518) replaced by alanine (A) displayed significant loss in fertility, ~18% and ~8% rescue respectively (Table 2, Figure 8). The mutation of arginine (R615) to alanine (A) had an inconclusive effect on fertility (20%-100% rescue). The triple amino acid substitution including

phenylalanine (F624-F625) and proline (P626) to alanine (A) and glycine (G) had very little affect on fertility (>94% rescue).

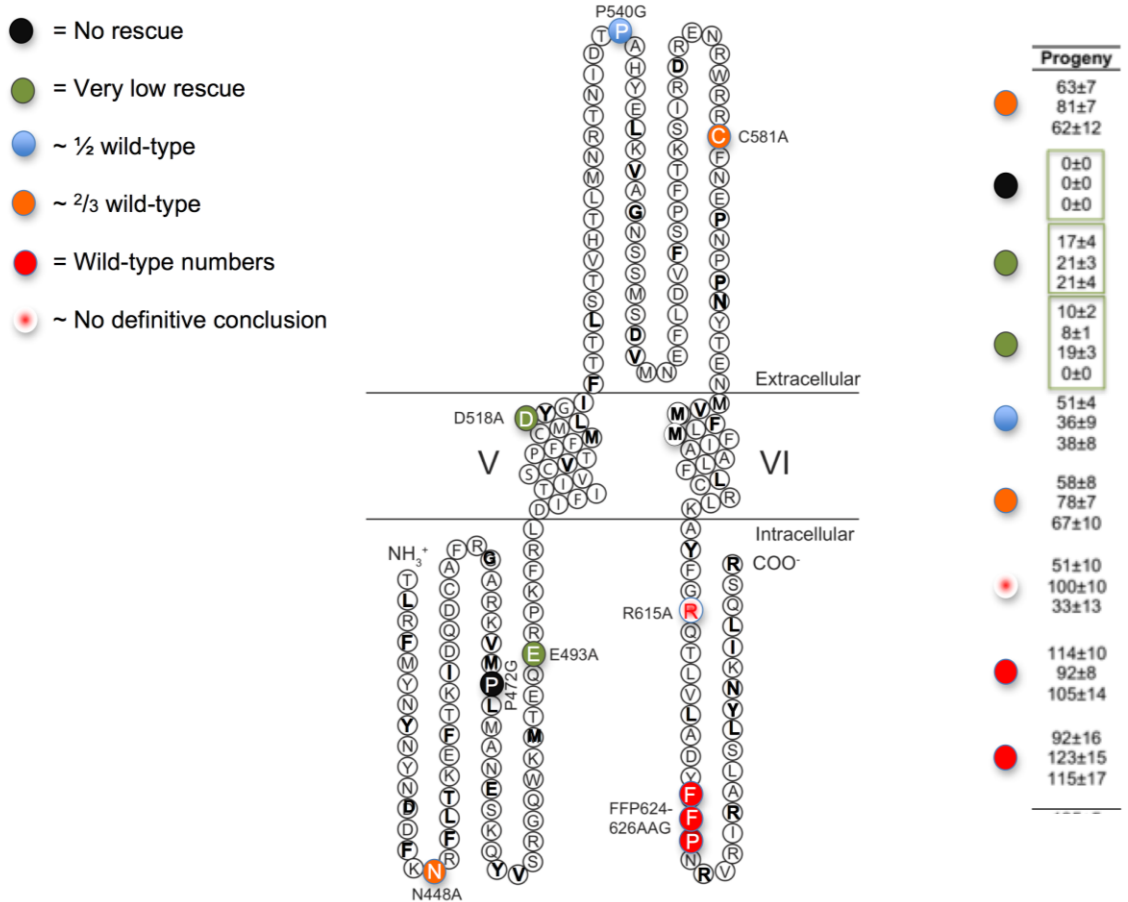


Figure 8. Diagram of amino acids substitutions in SPE-42 DC-STAMP. The substitutions were chosen based on conservation across metazoan species and possible effects on protein structure or function. Each color corresponds to a particular amino acid substitution in the SPE-42 protein and its effect on reproduction as shown by the progeny counts on the right. Highly conserved amino acids that were not mutated are in bold.

Table 2. Transgenic rescue with substitution mutations in DC-STAMP

Mutation ^a	Transgene ^b	Progeny ^c	N
N448A	xyEx247	63±7	11
	xyEx259	81±7	11
	xyEx237	62±12	9
P472G	xyEx166	0±0	12
	xyEx167	0±0	12
	xyEx170	0±0	13
E493A	xyEx195	17±4	12
	xyEx198	21±3	13
	xyEx265	21±4	11
D518A	xyEx196	10±2	14
	xyEx200	8±1	12
	xyEx236	19±3	12
	xyEx172	0±0	12
P540G	xyEx206	51±4	13
	xyEx202	36±9	12
	xyEx203	38±8	12
C581A	xyEx193	58±8	13
	xyEx199	78±7	12
	xyEx266	67±10	10
R615A	xyEx276	51±10	10
	xyEx277	100±10	10
	xyEx278	33±13	11
FFP624-626AAG	xyEx232	114±10	12
	xyEx251	92±8	10
	xyEx267	105±14	9
wild type	ebEx498	92±16	11
	xyEx175	123±15	10
	xyEx177	115±17	11
-	N2 (wild type) ^d	185±5	10
	<i>spe-42(tm1231)</i> ^d	0	24
	<i>spe-42(tm2421)</i> ^d	0	15

^a Live progeny counted for 3 independently-derived transgenes for each substitution.

^b Transgenes are in a null *spe-42(tm1231 or tm2421)* genetic background.

^c Counts conducted until the first day each worm produced no live progeny. Counts ± SEM.

^d These worms do not carry a transgene.

DC-STAMP domains are compatible only between closely related proteins

The insertion of the SPE-42 DC-STAMP domain into the LET-479 DC-STAMP region and vice versa (Table 3; Figure 9) resulted in a nonfunctional chimeric protein in both cases (no rescue). Both homozygous *spe-42* mutant worms containing transgenes with the LET-479 DC-STAMP domain in the SPE-42 protein and homozygous *let-479* mutant worms containing transgenes with the SPE-42 DC-STAMP domain in the LET-479 protein were sterile.

Table 3. Transgenic rescue by SPE-42 DC-STAMP chimeric constructs.

Transgene ^a	Transgene source ^b	DC-STAMP source	Progeny ^c	n
<i>xyEx254</i>	<i>spe-42</i>	<i>C.elegans let-479</i>	<1	9
<i>xyEx256</i>	<i>spe-42</i>	<i>C.elegans let-479</i>	0	10
<i>xyEx257</i>	<i>spe-42</i>	<i>C.elegans let-479</i>	0	10
<i>xyEx300</i>	<i>spe-42</i>	<i>C. briggsae spe-42</i>	128±13	10
<i>xyEx301</i>	<i>spe-42</i>	<i>C. briggsae spe-42</i>	147±19	9
<i>xyEx308</i>	<i>spe-42</i>	<i>C. briggsae spe-42</i>	118±17	9
<i>xyEx295</i>	<i>spe-42</i>	mouse <i>spe-42</i>	0	9
<i>xyEx296</i>	<i>spe-42</i>	mouse <i>spe-42</i>	0	10
<i>xyEx298</i>	<i>spe-42</i>	mouse <i>spe-42</i>	<1	9
<i>xyEx303</i>	<i>spe-42</i>	mouse DC-STAMP	0	10
<i>xyEx304</i>	<i>spe-42</i>	mouse DC-STAMP	0	10
<i>xyEx245</i>	<i>let-479</i>	<i>C. elegans spe-42</i>	<1	8
<i>ebEx498</i>	<i>spe-42</i>	<i>C. elegans spe-42</i>	92±16	11
<i>xyEx175</i>	<i>spe-42</i>	<i>C. elegans spe-42</i>	114±17	10
<i>xyEx177</i>	<i>spe-42</i>	<i>C. elegans spe-42</i>	112±18	11
N2 (wild type) ^d	-	-	184±5	10
<i>spe-42(tn1231)</i> ^d	-	-	0	24
<i>let-479(tm4655)</i> ^d	-	-	0	13

^a Transgenes are in a *spe-42(tn1231 or tm2421)* genetic background.

^b *C. elegans* genes were used as the base to build all transgenic constructs.

^c Counts conducted as described in Table 1. Counts ± SEM.

^d These worms do not carry a transgene.

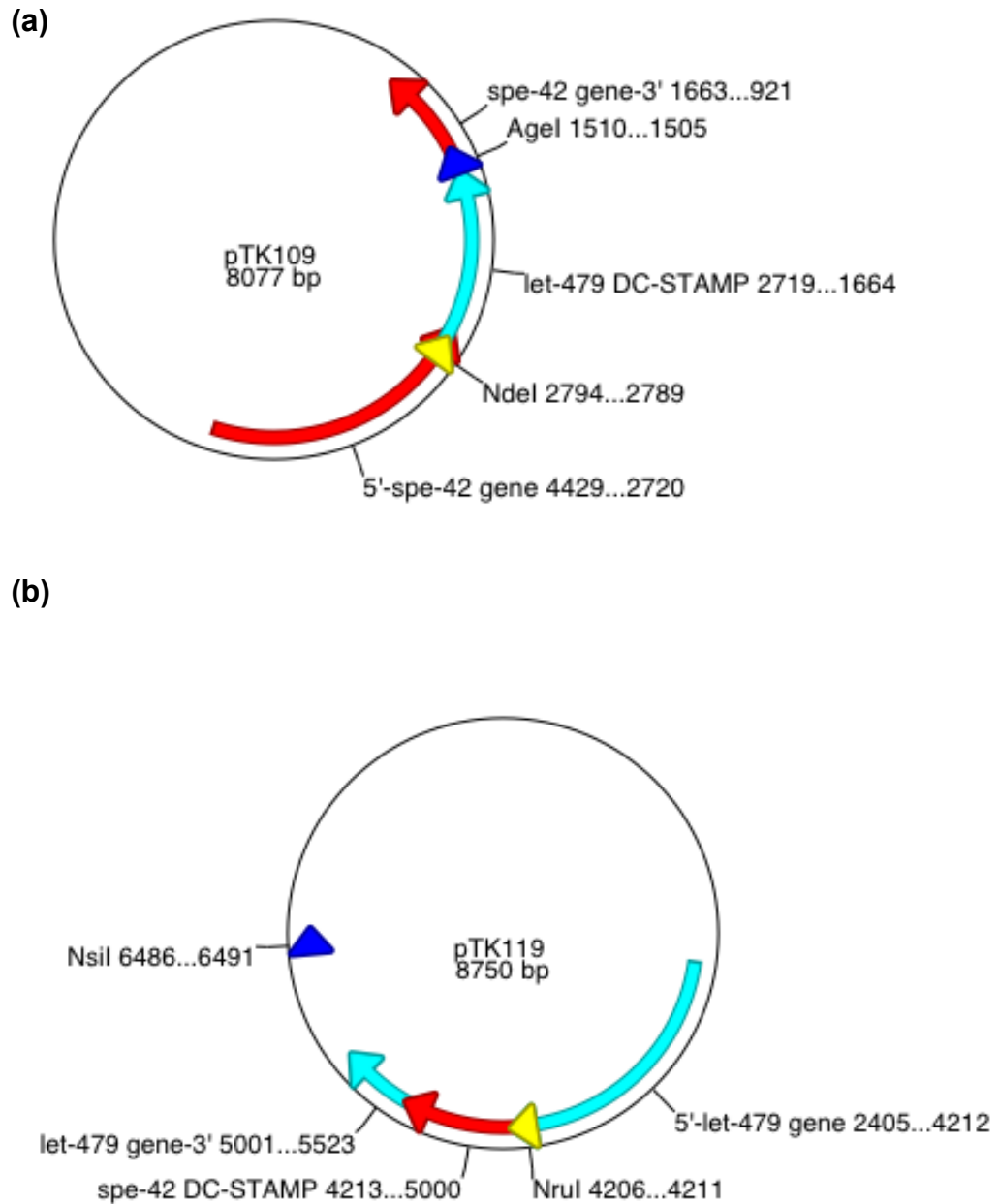


Figure 9. *let-479* and *spe-42* DC-STAMP domain swap plasmid constructs. Restriction enzymes *Agel* and *NdeI* were used to insert LET-479 DC-STAMP domain into SPE-42 plasmid, pTK109 (a). Restriction enzymes *NsiI* and *NruI* were used to insert SPE-42 DC-STAMP domain into LET-479 plasmid, pTK119 (b).

Three separate chimeric transgenic lines in which *C. elegans* SPE-42 DC-STAMP domain is replaced with mouse SPE-42 protein's DC-STAMP domain were created. A chimeric with two transgenic lines with the mouse DC-STAMP protein replacing the *C. elegans* SPE-42 DC-STAMP region were also made. The five chimeras were unable to rescue function of the native DC-STAMP region of the SPE-42 protein (no rescue; Table 3). However, the insertion of *C. briggsae* SPE-42 DC-STAMP domain into the *C. elegans* SPE-42 DC-STAMP region resulted in wild type levels of progeny (Table 3).

Discussion

Functional contribution of 8 DC-STAMP conserved amino acids in the SPE-42 DC-STAMP domain

Each of the amino acids substitutions were analyzed and interpreted based on the mutants' ability to rescue progeny compared to wildtype transgenic lines. First of all, the amino acid substitution of the cysteine (C581) to alanine (A) resulted in a small loss in progeny compared to wildtype transgenic lines, but not as much as some of the other mutations (Table 2). Therefore, it is not likely that the cysteine C581 participates in a disulfide bond with any cysteines from nearby or adjacent regions because of the high rescue. All of the conserved cysteines in the extracellular loop were found to be essential for protein function and appear to participate in disulfide bonds or other interactions (WILSON *et al.* 2011).

Therefore it was worthwhile to test the conserved cysteine (C518) present in the extracellular portion of the DC-STAMP domain. It is yet possible that the cysteine is able to contribute to stabilization of an alpha helix, which may account for the slight decrease in protein function. The five nearby arginines (Figure 8) might provide enough protein structural stability through multiple hydrogen bonds to minimize structural effects. The cysteine potentially contributes to securing the attachment of a SPE-42 ligand and therefore its absence causes a decrease in protein function. Despite the reason, the rescue is not perfect, resulting in the decreased numbers from wild type.

The amino acid substitution of the asparagine (N448) to alanine (A) also resulted in noticeably fewer progeny (Table 2). The fact that the progeny counts were lower than wild type supports the idea that the asparagine forms a hydrogen bond with the peptide backbone and possibly contributes to one of the ends of an alpha helix and this cannot occur with the substituted alanine. However, there are three asparagines within the next 10 amino acids closer to the plasma membrane that possibly provide enough stability through hydrogen bonding with the backbone to maintain a useful conformation (Figure 8).

Amino acid substitution of proline (P540) to glycine (G) in the extracellular matrix showed a reduction in progeny of more than 50%. This demonstrates that the conserved proline is useful for structural integrity in this location, possibly serving a role in a tight turn where the versatility of glycine was able to mimic the

conformation resulting in some rescue. The second proline substitution (P472) resulted in complete loss of protein function indicating that it plays an irreplaceable conformational or binding role in which the substituted glycine could not rescue. The location of this proline sandwiched among three highly conserved amino acids with propensities for forming alpha helices along with a few additional amino acids (M, A, L, K) on both sides with high propensities for alpha helices seems to indicate a kink or turn in a short coil (PACE and SCHOLTZ 1998). Perhaps the kink caused by proline maintains a structure that signals an action inside the cell after fusion that substitution with glycine cannot provide.

Substitution of both glutamic acid (E493) and aspartic acid (D518) with alanine (A) resulted in a drastic reduction in rescued progeny compared with wildtype numbers. These data support the hypothesis that they participate in an active site or in protein-protein interaction. Since D518 is predicted to be located either within the plasma membrane or near the surface, it could easily be used to securely bind an egg plasma membrane protein or other extracellular ligand. D518 may be necessary for good interaction but not critical, which would account for the small amount of rescue observed. E493 is 100% conserved across the species chosen for analysis. It is likely to participate in an important intercellular protein interaction simply because (1) substitution with alanine reduces protein function to 18% of the wild type level, (2) it has an electrostatic charge and (3)

none of the surrounding amino acids show significant conservation among the species compared.

Lastly, the triple amino acid substitution involving phenylalanine and proline (FFP624-626) appeared to have no effect on protein functionality. The final substitutions yielded progeny numbers equivalent to wild type indicating that these amino acids were not crucial and therefore unimportant for maintaining protein structure. We hypothesized that these nonpolar amino acids form a hydrophobic pocket used for binding due to the bulky side chains and since they were conserved in many species. The substitution to simpler nonpolar amino acids would eliminate the possibility for a pocket used for binding, however this had no measurable effect. These results do not support the hypothesis that it is a hydrophobic pocket used for binding. An interesting additional experiment would be to substitute one polar amino acid to see if that change affects function. It is possible that a substrate interacts here, but all that it needs is a hydrophobic surface. A deletion of one of the three may also help determine if they are involved in a protein turn.

All of the transgenic constructs were present in extrachromosomal arrays composed completely of simple repeats of the injected experimental DNA and marker DNA. The likelihood that the protein expression observed from the constructs is accurately localized in the cell for all constructs is high because wild-type SPE-42 within an extrachromosomal array was able to provide enough

functional protein to rescue function. In our lab, transgenic SPE-42 proteins were epitope-tagged to observe and confirm the site of expression, but the small amount of protein present was not detectable. Regardless of the ability to detect the transgenic protein in this way, the amount of protein expressed from the transgenic arrays was enough to rescue function to near wild type levels and was therefore viable and able to travel to the cellular site of action. However, the possibility remains that the non-functional proteins observed in these experiments were not reaching the active site or were being degraded prior to reaching the active site (WILSON *et al.* 2011).

DC-STAMP domains are functionally interchangeable only within closely related proteins

Although *C. elegans* SPE-42 and LET-479 share significant homology at the amino acid level, the DC-STAMP domains are not functionally interchangeable. It is clear that the evolutionary divergence between the two *C. elegans* proteins is significant enough to result in failure of the DC-STAMP domains from one of the proteins to function in the other. However, the possibilities remain that (1) they interact at some point in the cell, (2) one signals the other's action, or (3) they simply perform similar functions. This result is reasonable if the DC-STAMP domain does in fact act as a selective receptor before or during gamete fusion, signaling steps further along in the pathway.

The chimeras of two homologs to *C. elegans* SPE-42, including *M. musculus* SPE-42 and *C. briggsae* SPE-42, point to the conserved areas of the SPE-42 DC-STAMP domain that are essential to function in *C. elegans*. Experimentally, the DC-STAMP domains of *C. elegans* SPE-42 and its *M. musculus* SPE-42 homolog were not able to function interchangeably indicating that the DC-STAMP domains of *C. elegans* and *M. musculus* are not similar enough to result in even moderate rescue of progeny. Contrariwise, the swap of the DC-STAMP domains of *C. elegans* SPE-42 and *C. briggsae* SPE-42 led to wild type numbers of progeny suggesting the fundamental importance of their similar amino acids. These results may imply that (1) nematodes require additional conserved amino acids for successful fertilization, (2) mice SPE-42 DC-STAMP contains the essential amino acids for *C. elegans*, but the dissimilar regions actually hinder fertilization.

Potential function of the DC-STAMP domain

SPE-42 may interact with its binding partner(s) to (1) determine sperm fitness before fusion with the egg, to (2) signal the egg that distribution of the sperm material is complete, or to (3) signal after fusion to initiate degradation of membrane proteins no longer needed that may lead to polyspermy (WILSON *et al.* 2011). In *Chlamydomonas*, it has been shown that 2 proteins, FUS1 and HAP2, essential for fusion, are promptly signaled for degradation in response to

the fusion event (LIU *et al.* 2010). In a similar way, the RING-finger portion of SPE-42 may be cleaved and released into the cytoplasm like the mouse RING finger protein 13 (RNF13) in response to the fusion event. RNF13, an E3 ubiquitin ligase, is similar to SPE-42 because it is a transmembrane protein and contains a C-terminal RING-finger domain located in the cytosol (BOCOCK *et al.* 2009). When the RING-finger domain of the RNF13 protein is cleaved from the plasma membrane, this segment is able to take part in ubiquitination. This may indicate that after the SPE-42 protein completes its regulatory function for the cell and cleaves the RING-finger domain, the protein is then tagged for degradation. DC-STAMP is currently thought to act as a receptor and is known to affect fusion events, therefore it is possible that the ligand(s) of the DC-STAMP portion, signal cleavage and release of the RING-finger into the cytoplasm after initiating fusion of sperm and egg.

Chapter Three: Future Directions

The study of fertility among model organisms over the past century has led to abundant explanations of the intricate steps of fertilization including most recently the nematode, *C. elegans*. Despite these advances, the specific molecules that function in this process, their interactions and how they coordinate the fusion event continue to be a mystery. The simple reproductive biology of *C. elegans* combined with the power of forward and reverse genetics, are beginning to allow the identification of the proteins and molecules involved in fertilization

and other cell fusion events. Even though elaborate anatomical descriptions of fertilization are available, the discovery of the molecule or pathways underlying this event is still an exciting venue.

The efforts of this work primarily focused on the structural and functional picture of the enigmatic operations of the SPE-42 protein that occur just prior to or during the fertilization event. Because solved structures of DC-STAMP proteins or similar DC-STAMP domain-containing proteins are not yet available, the forging of the structural and functional significance is slow and tedious. After mutating hypothetical essential amino acids to discern a few of the most important amino acids for protein action, we concluded that some of the amino acids were essential. This information will be used to aid in prediction of SPE-42 structure through continued use of protein structure analysis computer programs and further transgenic rescue assays. As stated earlier, the DC-STAMP region was predicted to be a transmembrane domain by primary amino acid analysis through the use of several protein algorithms.

The discovery of the respective functions of the *spe-42*, *let-479* and DC-STAMP proteins will be greatly aided by solving each protein structure. For this reason, investigating and testing predicted structures are indispensable to ultimately determining the function of SPE-42 and LET-479. In this work, we were able to establish a few of the amino acids necessary for SPE-42 function. This information will assist in evaluating blocks of conservation among similar

functioning proteins and sifting out the structural motifs present in the proteins as well as the overall structural families in which the protein belongs. Additionally, our data will facilitate analysis of structural components such as active sites, folding patterns and overall protein architecture

We have now determined a few of the essential, moderately essential and non-essential amino acids in the SPE-42 DC-STAMP region. The three amino acid substitutions that yielded little to no progeny suggest the protein is non-functional without them. The possibility that the protein is not traveling to the site of action is not likely because of the ability to rescue seen with the other amino acid substitutions. Single or triple amino acid substitutions are also less likely to affect protein localization than gross changes. This case can be further tested by substitution of these amino acids to amino acids with similar properties. This experiment should show progeny yields only slightly lower than wild type transgenic levels if the amino acid is essential to function and reinforce the likelihood that the protein is reaching its active site. There is still the possibility that the protein is not reaching the active site, but the results are still interesting since a sole amino acid substitution conceivably causes either misfolding and/or degradation of the entire protein.

The four amino acid substitutions that yielded a smaller, but noticeable decrease in progeny (33.3%, 60.1%, 36.2%, 42.2%) show that the protein is able to shape itself adequately slightly more than half the time. This may indicate

these amino acids function as part of a group so that loss of a single member is not catastrophic. There may also be nearby amino acid(s) that can compensate for their loss, allowing the protein to regain structure and function. The amino acids may also be near a binding site so that their loss imposes enough instability to disrupt perfect binding without compromising function completely.

In order to provide a workable picture of proteins containing a DC-STAMP protein, the importance of discovering the functional amino acids and their role cannot be underestimated. Protein function often cannot be determined without deciphering the 3D structure of the protein. Most proteins need to be able to undergo conformational changes to complete their various roles in cell pathways. Many factors contribute to protein structure that may render the protein active or inactive. Some of the contributing factors based on the amino acid sequence are hydrogen bonding, Van der Waals forces, hydrophobic packaging, and ionic interactions. The methods for determining function of proteins all include several limitations that need to be considered. Experimental constraints include (1) tedious initial investigations of unknown proteins or protein regions, (2) proteins hard to tag in their natural environment, (3) proteins that change conformations, or (4) proteins that serve multiple functions. The functional significance of the DC-STAMP domain is yet to be discovered, but considering its status as a functional domain and the diversity of proteins within which it resides, it will

prospectively have far reaching effects into bone, arthritis, and fertility research because of its obvious role in fusion events.

Now that it has been shown that *C. briggsae* SPE-42 DC-STAMP region can be used in place of *C. elegans* SPE-42 DC-STAMP, a more thorough comparison of amino acids can be done to determine which are essential or conserved. A comparison of the two proteins should also provide clues that will aid in determination of a more accurate picture of its 3D shape.

A further investigation replacing LET-479 DC-STAMP with DC-STAMP domains from *M. musculus* and *C. briggsae* may shed more light on the possibility of an interaction between SPE-42 and LET-479 proteins. One objective would be to place *M. musculus* SPE-42 DC-STAMP and *M. musculus* LET-479 DC-STAMP into a self-sterile SPE-42/LET-479 background to test the possibility that these proteins interact. If the mutant proteins rescued function, this would indicate that they form a complex that requires a specific conformation in order for the protein to function. It would also show that DC-STAMP regions are interacting with one another. Another quick test to check for a possible interaction between SPE-42 and LET-479 would be to overexpress both proteins individually and then together in a S2 cell culture to see if overexpression leads to fusion of the cells. This would show that there is the possibility of fusion taking place in the presence of both proteins.

To determine the likelihood of SPE-42 and LET-479 being present only in species utilizing sperm and eggs and to look at the divergence from DC-STAMP and DC-STAMP-like proteins, a broader search of all sequenced plants and plant-like species should be conducted. These data would be used to construct a phylogenetic relationship between SPE-42 and DC-STAMP protein families and drawn upon to find relative estimates of the time these proteins diverged.

Once an antibody for *spe-42* is developed, several experiments will be possible. First, a wild type SPE-42 protein can be detected and its active location observed. Once the site of action is determined, transgenic constructs with amino acids substitutions can be observed to make sure fertilization fails because the protein is inactive as a result of the amino acid substitution and not because the protein does not reach the active site. For cases when the protein is partially active, the protein may not reach the active site because a certain conformation is less likely to occur.

There is also a possibility of identifying an extracellular ligand. If the active location of SPE-42 is known, an inventory of ligands present at the time may be tested as a match for SPE-42. SPE-42 might also act as an E3 ubiquitin ligase based on the presence of the RING-finger domain. We plan to screen known *C. elegans* E2 ubiquitin conjugating enzymes for interaction with SPE-42 and LET-479. Next, we will conduct a yeast two hybrid screen to search for binding partners for *C. elegans* SPE-42 and LET-479.

Through collaboration with an x-ray crystallographer, utilization of the hanging drop technique could be attempted to acquire a protein crystal and determine the structure of SPE-42 via its diffraction pattern. In addition, the primary structure could be entered into PONDR (Predictor of Naturally Disordered Regions) to determine the areas of disorder (ROMERO *et al.* 2002). Disordered regions contribute to functions such as protein binding and conformational changes that result in specificity or affinity of protein binding, and management of protein life span (DUNKER *et al.* 2002).

Ideally this research will contribute to the growing body of information concerning proteins and molecules related to fertilization and other fundamental fusion events in the body. These steps gradually lead to substantial increases in the general knowledge of this highly important field. In addition to the human benefits within multiple fields of medicine, endangered species, including those in captivity could profit from increased knowledge of the process of fertility. The more we understand the biological processes of model organisms, the more we can deduce solutions to problems with those processes. Outcomes of experiments conducted with *C. elegans* will compliment studies done in other model organisms while shortening the experimental periods and expense. *C. elegans* is a vital part of the pool of organisms used for fertilization studies because there is diminished complexity leading up to the fertilization event. This prevents causal assumptions of infertility that are actually related to problems

taking place earlier in the fertilization pathway. Finally, fertilization studies have an impact on many other systems, which is apparent by the diversity of the DC-STAMP region of SPE-42. These studies will lead to greater knowledge of interactions between cells as well as those inside the cell. The biological pathways and molecules operating inside and between cells within model organisms are worthwhile to investigate and lead to greater understanding of how the organisms function. The fusion of sperm and oocyte is similar to the fusion of other cells within the body as indicated by the fusion of preosteoclasts among many others. This is a fortunate side effect of all studies conducted in the name of fertilization leading to immeasurable gains in multiple fields of research.

Bibliography

- ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS and D. J. LIPMAN, 1990 Basic local alignment search tool. *J Mol Biol* **215**: 403-410.
- BEMBENEK, J. N., C. T. RICHIE, J. M. SQUIRRELL, J. M. CAMPBELL, K. W. ELICEIRI *et al.*, 2007 Cortical granule exocytosis in *C. elegans* is regulated by cell cycle components including separase. *Development* **134**: 3837-3848.
- BERNSEL, A., H. VIKLUND, J. FALK, E. LINDAHL, G. VON HEIJNE *et al.*, 2008 Prediction of membrane-protein topology from first principles. *Proc Natl Acad Sci U S A* **105**: 7177-7181.
- BLEIL, J. D., and P. M. WASSARMAN, 1983 Sperm-egg interactions in the mouse: sequence of events and induction of the acrosome reaction by a zona pellucida glycoprotein. *Dev Biol* **95**: 317-324.
- BOCOCK, J. P., S. CARMICLE, S. CHHOTANI, M. R. RUFFOLO, H. CHU *et al.*, 2009 The PA-TM-RING protein RING finger protein 13 is an endosomal integral membrane E3 ubiquitin ligase whose RING finger domain is released to the cytoplasm by proteolysis. *FEBS J* **276**: 1860-1877.
- BRENNER, S., 1974 The genetics of *Caenorhabditis elegans*. *Genetics* **77**: 71-94.
- CHANDRA, A., G. M. MARTINEZ, W. D. MOSHER, J. C. ABMA and J. JONES, 2005 Fertility, family planning, and reproductive health of U.S. women: Data from the 2002 National Survey of Family Growth. National Center for Health Statistics., pp. in *Vital Health Stat.*
- CHASNOV, J. R., and K. L. CHOW, 2002 Why Are There Males in the Hermaphroditic Species *Caenorhabditis elegans*? *Genetics* **160**: 983-994.
- CHATTERJEE, I., A. RICHMOND, E. PUTIRI, D. C. SHAKES and A. SINGSON, 2005 The *Caenorhabditis elegans spe-38* gene encodes a novel four-pass integral membrane protein required for sperm function at fertilization. *Development* **132**: 2795-2808.
- CHENG, H., J. A. GOVINDAN and D. GREENSTEIN, 2008 Regulated trafficking of the MSP/Eph receptor during oocyte meiotic maturation in *C. elegans*. *Curr Biol* **18**: 705-714.
- CLAVERIE, J., 2001 Gene Number: What if There are Only 30,000 Human Genes? *Science* **291**: 1255-1257.
- COLLURA, B., 2006 The Costs of Infertility Treatment, pp.
- CSERZO, M., F. EISENHABER, B. EISENHABER and I. SIMON, 2004 TM or not TM: transmembrane protein prediction with low false positive rate using DAS-TMfilter. *Bioinformatics* **20**: 136-137.
- DESHAIES, R. J., and C. A. JOAZEIRO, 2009 RING domain E3 ubiquitin ligases. *Annu Rev Biochem* **78**: 399-434.
- DUNKER, A. K., C. J. BROWN, J. D. LAWSON, L. M. IAKOUCHEVA and Z. OBRADOVIC, 2002 Intrinsic Disorder and Protein Function. *Biochemistry* **41**: 6573-6582.

- FERGUSON, E. L., and H. R. HORVITZ, 1985 Identification and characterization of 22 genes that affect the vulval cell lineages of the nematode *Caenorhabditis elegans*. *Genetics* **110**: 17-72.
- GAHLAY, G., L. GAUTHIER, B. BAIBAKOV, O. EPIFANO and J. DEAN, 2010 Gamete Recognition in Mice Depends on the Cleavage Status of an Egg's Zona Pellucida Protein. *Science* **329**: 216-219.
- GANAPATHIRAJU, M., N. BALAKRISHNAN, R. REDDY and J. KLEIN-SEETHARAMAN, 2008 Transmembrane helix prediction using amino acid property features and latent semantic analysis. *BMC Bioinformatics* **9 Suppl 1**: S4.
- GREENSTEIN, D., 2005 Control of oocyte meiotic maturation and fertilization, pp. Wormbook, doi/10.1895/wormbook.1891.1853.1891, <http://www.wormbook.org> in *Wormbook*, edited by T. C. E. R. COMMUNITY.
- GUNASEKARAN, S., I. CHATTERJEE, S. RAHIMI, M. K. DRUZHININA, K. LIJUN *et al.*, 2012 The sperm surface localization of the TRP-3/SPE-41 Ca(2+)-permeable channel depends on SPE-38 function in *Caenorhabditis elegans*. *Developmental Biology* **365**: 376-383.
- HARRIS, J. E., J. A. GOVINDAN, I. YAMAMOTO, J. SCHWARTZ, I. KAVERINA *et al.*, 2006 Major sperm protein signaling promotes oocyte microtubule reorganization prior to fertilization in *Caenorhabditis elegans*. *Dev Biol* **299**: 105-121.
- HE, Z. Y., C. BRAKEBUSCH, R. FASSLER, J. A. KREIDBERG, P. PRIMAKOFF *et al.*, 2003 None of the integrins known to be present on the mouse egg or to be ADAM receptors are essential for sperm-egg binding and fusion. *Dev Biol* **254**: 226-237.
- HECKMAN, K. L., and L. R. PEASE, 2007 Gene splicing and mutagenesis by PCR-driven overlap extension. *Nature Protocols* **2**: 924-932.
- HERBST, R. S., 2004 Review of epidermal growth factor receptor biology. *International Journal of Radiation Oncology*Biological*Physics* **59**: 21-26.
- HODGKIN, J., 1988 Sexual dimorphism and sex determination, pp. 243-279 in *The Nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- HODGKIN, J., H. R. HORVITZ and S. BRENNER, 1979 Nondisjunction mutants of the nematode *Caenorhabditis elegans*. *Genetics* **91**: 67-94.
- HOFHERR, S. E., A. E. WIKTOR, B. R. KIPP, D. B. DAWSON and D. L. VAN DYKE, 2011 Clinical diagnostic testing for the cytogenetic and molecular causes of male infertility: the Mayo Clinic experience. *J Assist Reprod Genet* **28**: 1091-1098.
- HUGOT, J. P., P. BAUJARD and S. MORAND, 2001 Biodiversity in helminths and nematodes as a field of study: an overview. *Nematology* **3**: 199-208.
- HUIBREGTSE, J. M., M. SCHEFFNER, S. BEAUDENON and P. M. HOWLEY, 1995 A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc Natl Acad Sci U S A* **92**: 2563-2567.

- INOUE, N., M. IKAWA, A. ISOTANI and M. OKABE, 2005 The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs. *Nature* **434**: 234-238.
- JONES, D. T., 2007 Improving the accuracy of transmembrane protein topology prediction using evolutionary information. *Bioinformatics* **23**: 538-544.
- JUNGNICKEL, M. K., H. MARRERO, L. BIRNBAUMER, J. R. LEMOS and H. M. FLORMAN, 2001 Trp2 regulates entry of Ca²⁺ into mouse sperm triggered by egg ZP3. *Nature Cell Biology* **3**: 499-502.
- JUNGNICKEL, M. K., K. A. SUTTON and H. M. FLORMAN, 2003 In the beginning: lessons from fertilization in mice and worms. *Cell* **114**: 401-404.
- JURETIC, D., L. ZORANIC and D. ZUCIC, 2002 Basic charge clusters and predictions of membrane protein topology. *J Chem Inf Comput Sci* **42**: 620-632.
- KAJI, K., S. ODA, T. SHIKANO, T. OHNUKI, Y. UEMATSU *et al.*, 2000 The gamete fusion process is defective in eggs of CD9-deficient mice. *Nat Genet* **24**: 279-282.
- KALL, L., A. KROGH and E. L. SONNHAMMER, 2004 A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* **338**: 1027-1036.
- KALL, L., A. KROGH and E. L. SONNHAMMER, 2005 An HMM posterior decoder for sequence feature prediction that includes homology information. *Bioinformatics* **21 Suppl 1**: i251-257.
- KOSINSKI, M., K. McDONALD, J. SCHWARTZ, I. YAMAMOTO and D. GREENSTEIN, 2005 *C. elegans* sperm bud vesicles to deliver a meiotic maturation signal to distant oocytes. *Development* **132**: 3357-3369.
- KROFT, T. L., E. J. GLEASON and S. W. L'HERNAULT, 2005 The spe-42 gene is required for sperm-egg interactions during *C. elegans* fertilization and encodes a sperm-specific transmembrane protein. *Dev Biol* **286**: 169-181.
- KUBAGAWA, H. M., J. L. WATTS, C. CORRIGAN, J. W. EDMONDS, E. SZTUL *et al.*, 2006 Oocyte signals derived from polyunsaturated fatty acids control sperm recruitment in vivo. *Nat Cell Biol* **8**: 1143-1148.
- L'HERNAULT, S. W., 1997 Spermatogenesis, pp. 271-294 in *Caenorhabditis elegans II*, edited by D. L. RIDDLE, BLUMENTHAL, T., MEYER, B. J. AND PRIESS, J. R. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- L'HERNAULT, S. W., 2006 Spermatogenesis, pp. Wormbook, doi/10.1895.1891, <http://www.wormbook.org> in *Wormbook*, edited by T. C. E. R. COMMUNITY.
- L'HERNAULT, S. W., and A. W. SINGSON, 2000 Developmental Genetics of Spermatogenesis in the Nematode *Caenorhabditis elegans*, pp. 109-119 in *The Testis: From Stem Cell to Sperm Function*, edited by E. GOLDBERG. Serono Symposia USA.
- LAM, J., S. TAKESHITA, J. E. BARKER, O. KANAGAWA, F. P. ROSS *et al.*, 2000 TNF- α induces osteoclastogenesis by direct stimulation of macrophages exposed to permissive levels of RANK ligand. *Journal of Clinical Investigation* **106**: 1481-1488.

- LAMUNYON, C. W., and S. WARD, 1995 Sperm precedence in a hermaphroditic nematode (*Caenorhabditis elegans*) is due to competitive superiority of male sperm. *Experientia* **51**: 817-823.
- LAMUNYON, C. W., and S. WARD, 1998 Larger sperm outcompete smaller sperm in the nematode *Caenorhabditis elegans*. *Proc R Soc Lond B Biol Sci* **265**: 1997-2002.
- LE NAOUR, F., E. RUBINSTEIN, C. JASMIN, M. PRENANT and C. BOUCHEIX, 2000 Severely reduced female fertility in CD9-deficient mice. *Science* **287**: 319-321.
- LIU, Y., M. J. MISAMORE and W. J. SNELL, 2010 Membrane fusion triggers rapid degradation of two gamete-specific, fusion-essential proteins in a membrane block to polygamy in *Chlamydomonas*. *Development* **137**: 1473-1481.
- MELLO, C., and A. FIRE, 1995 DNA transformation. *Methods Cell Biol* **48**: 451-482.
- MELLO, C. C., J. M. KRAMER, D. STINCHCOMB and V. AMBROS, 1991 Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *Embo J* **10**: 3959-3970.
- MENSAH, K. A., C. T. RITCHLIN and E. M. SCHWARZ, 2010 RANKL induces heterogeneous DC-STAMP(lo) and DC-STAMP(hi) osteoclast precursors of which the DC-STAMP(lo) precursors are the master fusogens. *J Cell Physiol* **223**: 76-83.
- MILLER, M. A., V. Q. NGUYEN, M. H. LEE, M. KOSINSKI, T. SCHEDL *et al.*, 2001 A sperm cytoskeletal protein that signals oocyte meiotic maturation and ovulation. *Science* **291**: 2144-2147.
- MILLER, M. A., P. J. RUEST, M. KOSINSKI, S. K. HANKS and D. GREENSTEIN, 2003 An Eph receptor sperm-sensing control mechanism for oocyte meiotic maturation in *Caenorhabditis elegans*. *Genes Dev* **17**: 187-200.
- MOY, G. W., L. M. MENDOZA, J. R. SCHULZ, W. J. SWANSON, C. G. GLABE *et al.*, 1996 The sea urchin sperm receptor for egg jelly is a modular protein with extensive homology to the human polycystic kidney disease protein, PKD1. *J Cell Biol* **133**: 809-817.
- MURPHY, K., P. TRAVERS and M. WALPORT, 2007 *Janeway's Immunobiology*. Garland Science.
- MURRAY, R. L., J. L. KOZLOWSKA and A. D. CUTTER, 2011 Heritable determinants of male fertilization success in the nematode *Caenorhabditis elegans*. *BMC Evolutionary Biology* **11**.
- NEILL, A. T., and V. D. VACQUIER, 2004 Ligands and receptors mediating signal transduction in sea urchin spermatozoa. *Reproduction* **127**: 141-149.
- NICHOLAS, K. B., and H. B. J. NICHOLAS, 1997 GeneDoc: a tool for editing and annotating multiple sequence alignments **4**.

- OURSLER, M. J., 1994 Osteoclast synthesis and secretion and activation of latent transforming growth factor beta. *Journal of Bone and Mineral Research* **9**: 443-452.
- OURSLER, M. J., B. L. RIGGS and T. C. SPELSBERG, 1993 Glucocorticoid-induced activation of latent transforming growth factor-beta by normal human osteoblast-like cells. *Endocrinology* **133**: 2187-2196.
- PACE, C. N., and J. M. SCHOLTZ, 1998 A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal* **75**: 422-427.
- PEI, J., and N. V. GRISHIN, 2007 PROMALS: towards accurate multiple sequence alignments of distantly related proteins. *Bioinformatics* **23**: 802-808.
- RANKIN, T., P. TALBOT, E. LEE and J. DEAN, 1999 Abnormal zonae pellucidae in mice lacking ZP1 result in early embryonic loss. *Developmental* **126**: 3847-3855.
- RANKIN, T. L., M. O'BRIEN, E. LEE, K. WIGGLESWORTH, J. EPPIG *et al.*, 2001 Defective zonae pellucidae in Zp2-null mice disrupt folliculogenesis, fertility and development. *Development* **128**: 1119-1126.
- ROMERO, P., A. K. DUNKER, X. LI and Z. OBRADOVIC, 2002 PONDRA, pp. Molecular Kinetics, Inc.
- SAKAI, H., 2012 The CD40-CD40L axis and IFN- γ play critical roles in Langhans giant cell formation. *International Immunology* **24**: 5-15.
- SEGALL, G. K., and W. J. LENNARZ, 1979 Chemical characterization of the component of the jelly coat from sea urchin eggs responsible for induction of the acrosome reaction. *Dev Biol* **71**: 33-48.
- SEPSENWOL, S., H. RIS and T. M. ROBERTS, 1989 A unique cytoskeleton associated with crawling in the amoeboid sperm of the nematode, *Ascaris suum*. *The Journal of Cell Biology* **108**: 55-66.
- SHARP, P. M., and K. R. BRADNAM, 1997 Codon Usage in *C. elegans*, pp. 1053-1057 in *C. elegans II*, edited by D. L. RIDDLE, T. BLUMENTHAL, B. J. MEYER and J. R. PRIESS. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- SHAYE, D. D., and I. GREENWALD, 2011 OrthoList: a compendium of *C. elegans* genes with human orthologs. *PLoS ONE* **6**: 20085.
- SHEN, H., and J. J. CHOU, 2008 MemBrain: improving the accuracy of predicting transmembrane helices. *PLoS ONE* **3**: e2399.
- SINGSON, A., 2001 Every sperm is sacred: fertilization in *Caenorhabditis elegans*. *Dev Biol* **230**: 101-109.
- SINGSON, A., J. S. HANG and J. M. PARRY, 2008 Genes required for the common miracle of fertilization in *Caenorhabditis elegans*. *Int J Dev Biol* **52**: 647-656.
- SINGSON, A., K. L. HILL and S. W. L'HERNAULT, 1999 Sperm competition in the absence of fertilization in *Caenorhabditis elegans*. *Genetics* **152**: 201-208.

- SINGSON, A., K. B. MERCER and S. W. L'HERNAULT, 1998 The *C. elegans spe-9* gene encodes a sperm transmembrane protein that contains EGF-like repeats and is required for fertilization. *Cell* **93**: 71-79.
- STEIN, K. K., P. PRIMAKOFF and D. MYLES, 2004 Sperm-egg fusion: events at the plasma membrane. *J Cell Sci* **117**: 6269-6274.
- STEIN, L. D., Z. BAO, D. BLASIAR, T. BLUMENTHAL, M. R. BRENT *et al.*, 2003 The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol* **1**: E45.
- STENICO, M., A. T. LLOYD and P. M. SHARP, 1994 Codon usage in *Caenorhabditis elegans*: delineation of translational selection and mutational biases. *Nucleic Acids Research* **22**: 2437-2446.
- SUDA, T., N. TAKAHASHI, N. UDAGAWA, E. JIMI, M. T. GILLESPIE *et al.*, 1999 Modulation of osteoclast differentiation and function by the new members of the tumor necrosis factor receptor and ligand families. *Endocrine Review* **20**: 345-357.
- SUZUKI, N., 1995 Structure, function and biosynthesis of sperm-activating peptides and fucose sulfate glycoconjugate in the extracellular coat of sea urchin eggs. *Zoolog Sci* **12**: 13-27.
- THEILL, L. E., W. J. BOYLE and J. M. PENNINGER, 2002 RANK-L and RANK: T cells, bone loss, and mammalian evolution. *Annual Review of Immunology* **20**: 795-823.
- THERIOT, J. A., 1996 Worm sperm and advances in cell locomotion. *Cell* **84**: 1-4.
- THOMAS, C. D., A. CAMERON, R. E. GREEN, M. BAKKENES, L. J. BEAUMONT *et al.*, 2004 Extinction risk from climate change. *Nature* **427**: 87-180.
- TUSNADY, G. E., and I. SIMON, 2001 The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**: 849-850.
- VACQUIER, V. D., and G. W. MOY, 1977 Isolation of bindin: the protein responsible for adhesion of sperm to sea urchin eggs. *Proc Natl Acad Sci U S A* **74**: 2456-2460.
- VACQUIER, V. D., and G. W. MOY, 1997 The fucose sulfate polymer of egg jelly binds to sperm REJ and is the inducer of the sea urchin sperm acrosome reaction. *Dev Biol* **192**: 125-135.
- VIGNERY, A., 2005 Macrophage fusion: the making of osteoclasts and giant cells. *J Exp Med* **202**: 337-340.
- VIKLUND, H., and A. ELOFSSON, 2004 Best alpha-helical transmembrane protein topology predictions are achieved using hidden Markov models and evolutionary information. *Protein Sci* **13**: 1908-1917.
- VIKLUND, H., and A. ELOFSSON, 2008 OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar. *Bioinformatics* **24**: 1662-1668.

- WARD, G. E., C. J. BROKAW, D. L. GARBERS and V. D. VACQUIER, 1985 Chemotaxis of *Arbacia punctulata* spermatozoa to resact, a peptide from the egg jelly layer. *J Cell Biol* **101**: 2324-2329.
- WARD, S., and J. S. CARREL, 1979 Fertilization and sperm competition in the nematode *Caenorhabditis elegans*. *Dev Biol* **73**: 304-321.
- WARD, S., E. HOGAN and G. A. NELSON, 1983 The initiation of spermiogenesis in the nematode *Caenorhabditis elegans*. *Dev Biol* **98**: 70-79.
- WARD, S., and J. MIWA, 1978 Characterization of temperature-sensitive, fertilization-defective mutants of the nematode *Caenorhabditis elegans*. *Genetics* **88**: 285-303.
- WASSARMAN, P. M., 1990 Profile of a mammalian sperm receptor. *Development* **108**: 1-17.
- WHEELAN, S. J., M. S. BOGUSKI, L. DURET and W. MAKALOWSKI, 1999 Human and nematode orthologs--lessons from the analysis of 1800 human genes and the proteome of *Caenorhabditis elegans*. *Gene* **238**: 163-170.
- WILSON, K. L., K. R. FITCH, B. T. BAFUS and B. T. WAKIMOTO, 2006 Sperm plasma membrane breakdown during *Drosophila* fertilization requires sneaky, an acrosomal membrane protein. *Development* **133**: 4871-4879.
- WILSON, L. D., J. M. SACKETT, B. D. MIECZKOWSKI, A. L. RICHIE, K. THOEMKE *et al.*, 2011 Fertilization in *C. elegans* requires an intact C-terminal RING finger in sperm protein SPE-42. *BMC Developmental Biology* **11**.
- WOOD, W. B., 1988 Introduction to *C. elegans* biology, pp. 1-16 in *The nematode Caenorhabditis elegans*, edited by W. B. WOOD. Cold Spring Harbor Press, Cold Spring Harbor, NY.
- XU, X. Z., and P. W. STERNBERG, 2003 A *C. elegans* sperm TRP protein required for sperm-egg interactions during fertilization. *Cell* **114**: 285-297.
- YAGI, M., T. MIYAMOTO, Y. SAWATANI, K. IWAMOTO, N. HOSOGANE *et al.*, 2005 DC-STAMP is essential for cell-cell fusion in osteoclasts and foreign body giant cells. *J Exp Med* **202**: 345-351.
- YAMAMOTO, I., M. E. KOSINSKI and D. GREENSTEIN, 2006 Start me up: cell signaling and the journey from oocyte to embryo in *C. elegans*. *Dev Dyn* **235**: 571-585.
- YIN, Q., S. C. LIN, B. LAMOTHE, M. LU, Y. C. LO *et al.*, 2009 E2 interaction and dimerization in the crystal structure of TRAF6. *Nat Struct Mol Biol* **16**: 658-666.
- ZANNONI, S., S. W. L'HERNAULT and A. W. SINGSON, 2003 Dynamic localization of SPE-9 in sperm: a protein required for sperm-oocyte interactions in *Caenorhabditis elegans*. *BMC Dev Biol* **3**: 10.

Appendix

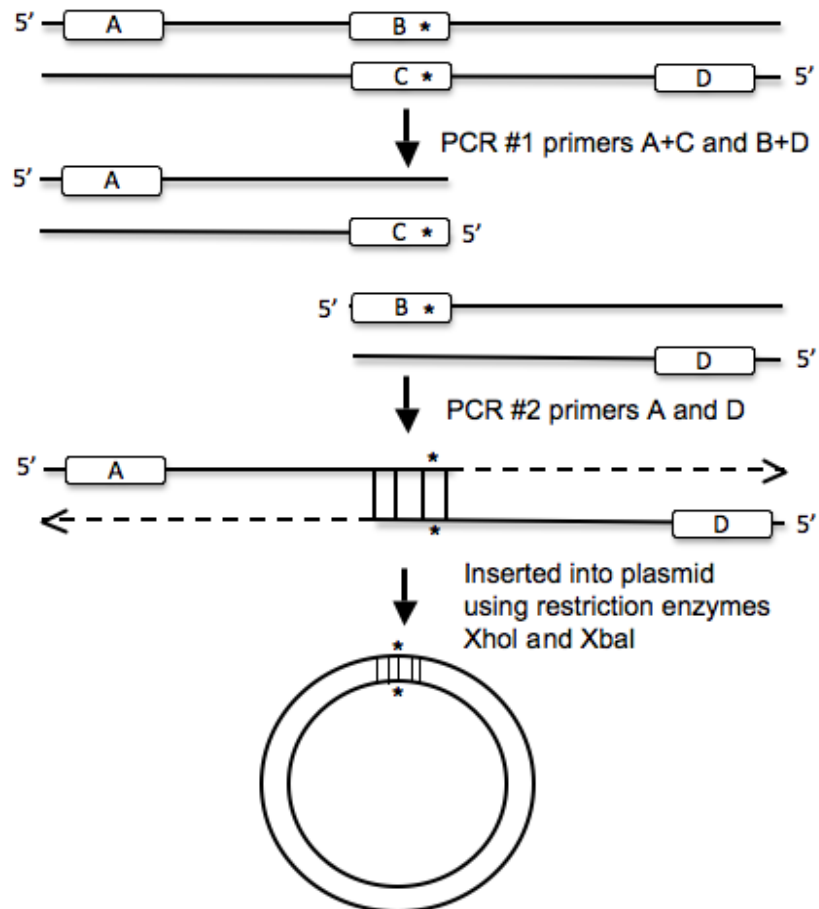


Figure 10. Amino acid substitution overlap extension PCR strategy. This method was used to make amino acid substitutions R615A (pTK71) and FFP 624-626 AAG (pTK72). This strategy involved two PCR reactions using mutagenic primers B* and C* and flanking primers A and D (modified from [HECKMAN and PEASE 2007]).



Figure 11. DC-STAMP swap overlap PCR strategy. This method was used to replace the *C. elegans spe-42* DC-STAMP domain with the same domain from (in separate constructs) the mouse *spe-42* homolog and DC-STAMP gene, the *C. briggsae spe-42* homolog, and the *C. elegans let-479* gene. This strategy involved four PCR reactions. PCR I and PCR III were run with pTK15 (*C. elegans spe-42*) plasmid as template using primers A and C in one reaction and primers D and F in a second reaction, respectively. PCR II was run with primers B and E using fosmid 5dD07 (*let-479* genomic DNA) as template. PCR IV used gel-purified PCR products from PCR reactions I-III as template and was run 5 cycles with no primers followed by 25 cycles with primers A and F. PCR IV and pTK15 were digested with restriction enzymes *AgeI* HF and *NdeI* and ligated.

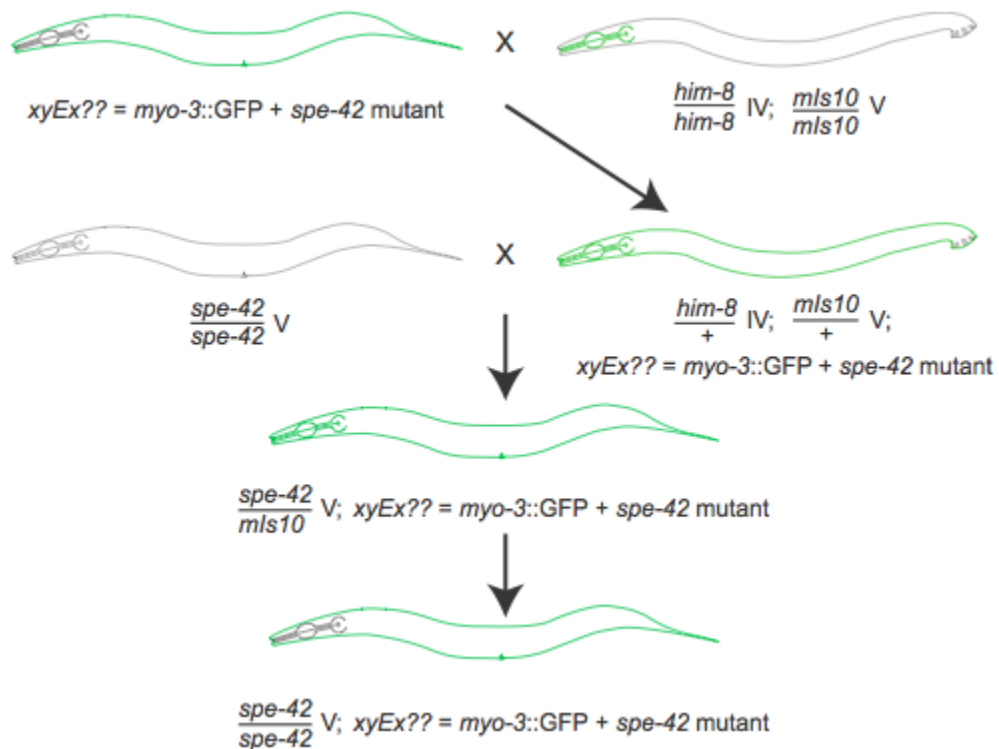


Figure 12. Transgenic crossing scheme. Homozygous *spe-42* (chromosome V) mutant hermaphrodites bearing an extrachromosomal array with the transgene of interest (body wall GFP) were crossed with males bearing a *mls10* marker on chromosome V (pharynx GFP). Male progeny bearing the transgene and *mls10* marker were then crossed with homozygous *spe-42* mutants. Progeny bearing both GFP markers were allowed to self-fertilize and *spe-42* homozygous hermaphrodites with the transgene were picked.

Table 4. SPE-42 predicted membrane topology.

<i>Helices</i>	<i>I</i>	<i>II</i>	<i>III</i>	<i>IIIa</i>	<i>IV</i>		<i>V</i>	<i>VI</i>	<i>VII</i>
Algorithm									
Scampi-msa	84- 104	106- 126	147- 167		414- 434		500- 520	596- 616	
PRODIV	82- 102	107- 127	148- 168		413- 433		501- 521	593- 613	
Octopus	77- 97	112- 132	147- 167	293- 313	413- 433		501- 521	593- 613	
Phobius	83- 102	108- 127	148- 167		405- 431		499- 522	594- 613	
PolyPhobius	79- 99	108- 128	148- 167		406- 432		499- 522	594- 612	
DAS-TMf	79- 99	113	159		403- 433		502- 522	595- 608	
HMMTOP	77- 100	109- 128	149- 167	319- 338	398- 416	502- 522	421- 440	499- 517	
SPLIT	75- 100	110	159	165- 185	399- 417	421- 437	499- 524	593- 609	697- 719
Mem3.0	98	122	146- 165			418- 437	498- 522	582- 606	
Membrain	78- 98	113- 141	145- 160		403- 435		500- 522	593- 608	702- 716
TMPro	83- 99	107- 129	149- 157		412- 432		499- 510	594- 602	704- 712
Consensus	80- 100	108- 128	147- 165		407- 430		500- 521	594- 610	

*Numbers indicate projected locations of helices within SPE-42

Table 5. Oligos used to make mutations in wild type SPE-42 plasmid pTK15.

Plasmid	Codon Change 5'-sense-3'	Oligonucleotides
pTK65 N448A	5'-AAAAACCGT-3' 5'-AAA GCC CGT-3'	TK496: 5'-AATGACGACTTTAA GCC CGTTTTCTCACAAAAG-3' TK497: 3'-TTACTGCTGAAATTT CGG GCAAAAGAGTGTTTTC-5'
pTK66 P472G	5'-ATGCCATTA-3' 5'-ATG GGA TTA-3'	TK498: 5'-GCTCGTAAAGTCATG GGA TTAATGGCAAATGAA-3' TK499: 3'-CGAGCATTTCAGTAC CCT TAATTACCGTTTACTT-5'
pTK67 E493A	5'-CAAGAGCGA-3' 5'-CAA GCT CGA-3'	TK500: 5'-GAAATGACAGAACA GCT CGACCTAAATTCGAT-3' TK501: 3'-CTTTTACTGTCTTGT CGA GCTGGATTTAAAGCTA-5'
pTK68 D518A	5'-CTTGATTAT-3' 5'-CTT GCT TAT-3'	TK502: 5'-CTTTATGTGTATGCTT GCT TATGGTATATTACG-3' TK503: 3'-GAAATACACATACGA CGA ATACCATATAAATGC-5'
pTK69 P540G	5'-ACTCCTGCT-3' 5'-ACT GGA GCT-3'	TK504: 5'-ACAAATATCGATACT GGA GCTCATTATGAATTG-3' TK505: 3'-TGTTTATAGCTATG CCT CGAGTAATACTTAAC-5'
pTK70 C581A	5'-CGATGTTTC-3' 5'-CGA GCT TTC-3'	TK506: 5'-AATCGTTGGAGACG GCT TTCAACGAACCAAATC-3' TK507: 3'-TTAGCAACCTCTGCT CGA AAGTTGCTTGGTTTAG-5'
pTK71 R615A	5'-GGAAGGCAA-3' 5'-GGA GCT CAA-3'	TK508: 5'-GAAGGCATATTTTGG GCT CAAACACTTGATTG-3' TK509: 3'-CTTCCGTATAAAACCT CGA GTTTGTGAACATAAC-5'
pTK72 FFP624- 626AAG	5'-TTCTTTCCG-3' 5'- GCCGCTGGA -3'	TK510: 5'-GTATTGGCAGATTAC GCCGCTGGA AATCGAGTAAGAATTC-3' TK511: 3'-CATAACCGTCTAATG CGGCGACCT TTAGCTCATTCTTAAG-5'

* Changed nucleotides in grey