2010-06-01

RNA Recognition by the Caenorhabditis elegans Embryonic Determinants MEX-5 and MEX-3: A Dissertation

John M. Pagano Jr.
University of Massachusetts Medical School

RNA RECOGNITION BY THE *CAENORHABDITIS ELEGANS* EMBRYONIC DETERMINANTS MEX-5 AND MEX-3

A Dissertation Presented

By

JOHN MICHAEL PAGANO, JR.

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

JUNE 1, 2010

BIOCHEMISTRY AND MOLECULAR PHARMACOLOGY

RNA RECOGNITION BY THE *CAENORHABDITIS ELEGANS* EMBRYONIC DETERMINANTS MEX-5 AND MEX-3

A Dissertation Presented
By

JOHN MICHAEL PAGANO, JR.

The signatures of the Dissertation Defense Committee signifies completion and approval as to style and content of the Dissertation

Sean Ryder, Ph.D., Thesis Advisor

Allan Jacobson, Ph.D., Member of Committee

Marian Walhout, Ph.D., Member of Committee

Scot Wolfe, Ph.D., Member of Committee

Geraldine Seydoux, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

Kendall Knight, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Department of Biochemistry and Molecular Pharmacology

June 1, 2010

DEDICATION

This thesis dissertation is dedicated to the memory of my friend Lisa M. McCoig who passed away on Saturday, February 6, 2010 at the age of 26 years old.

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor Sean Ryder. I really admire the degree of excellence and professionalism that you hold as a research scientist. I've learned how to think critically and ask the right questions to challenging scientific problems. Additionally, you've taught me the importance of communicating my work. I feel like I have been stretched far beyond what I thought I was capable of when I started working in your lab. I'm grateful that I was able to work in your lab and feel privileged to receive the quality of training that you have given me.

I would also like to thank all the members of the Ryder lab. You have been a great support through the duration of my time here. Thank you for reading the numerous drafts of abstracts, manuscripts, posters and proposals. Also for sitting through all of my practice talks over and over again. In particular, I want to thank Brian Farley. Thanks so much for helping me throughout the course of graduate school. I've enjoyed all the time we've had, especially traveling all over the country to various meetings. Also, thank you for all the help you have given me in writing this dissertation. I appreciate all of your critical comments and suggestions.

I would also like to thank the members of my dissertation committee, which include Scot Wolfe, Allan Jacobson, Ken Knight, Marian Walhout, and Geraldine Seydoux. Thank you all for sacrificing your time and investing in my graduate career here at the University of Massachusetts Medical School.

ABSTRACT

Post-transcriptional regulation of gene expression is a mechanism that governs developmental and cellular events in metazoans. In early embryogenesis, transcriptionally quiescent cells depend upon maternally supplied factors such as RNA binding proteins and RNA that control key decisions. Morphogen gradients form and in turn pattern the early embryo generating different cell types and spatial order. In the nematode *Caenorhabditis elegans*, the early embryo relies upon several RNA binding proteins that control mRNA stability, translation efficiency, and/or mRNA localization of cell fate determinants essential for proper development.

MEX-5 and MEX-3 are two conserved RNA-binding proteins required to pattern the anterior/posterior axis and early embryo. Mutation of either gene results in a maternal effect lethal phenotype with proliferating posterior muscle into the anterior blastomeres (Muscle EXcess). Several cell-fate determinants are aberrantly expressed in *mex-5* and *mex-3* embryos. Both proteins are thought to interact with *cis*-regulatory elements present in 3'-UTRs of target RNAs controlling their metabolism. However, previous studies failed to demonstrate that these proteins regulate maternal transcripts directly.

This dissertation presents a thorough assessment of the RNA binding properties of MEX-5 and MEX-3. Quantitative biochemical approaches were used to determine the RNA binding specificity of both proteins. MEX-5 has a relaxed specificity, binding with high affinity to linear RNA containing a tract of six

or more uridines within an eight-nucleotide window. This is very different from its mammalian homologs Tristetraprolin (TTP) and ERF-2. I was able to identify two amino acids present within the MEX-5 RNA binding domain that are required for the differential RNA recognition observed between MEX-5 and TTP. MEX-3 on the other hand is a specific RNA binding protein, recognizing a bipartite element with flexible spacing between two four-nucleotide half-sites. I demonstrate that this element is required for MEX-3 dependent regulation *in vivo*. Previous studies only identify a small number of candidate regulatory targets of MEX-5 and MEX-3. The defined sequence specificity of both proteins is used to predict new putative targets that may be regulated by either protein. Collectively, this study examines the RNA binding properties of MEX-5 and MEX-3 to clarify their role as post-transcriptional regulators in nematode development.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

Portions of this dissertation have appeared in separate publications or as part of publications

Chapter III and one paragraph of chapter V were adapted from:

Pagano, J. M., Farley, B. M., Essien, K. I., and Ryder, S. P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. Proc Natl Acad Sci U S A *106*, 20252-20257.

Chapter IV was adapted from:

Pagano, J. M., Farley, B. M., McCoig, L. M., and Ryder, S. P. (2007). Molecular Basis of RNA Recognition by the Embryonic Polarity Determinant MEX-5. J Biol Chem *282*, 8883-8894.

Brian M. Farley performed most of the bioinformatics in Chapter III and IV and made Figure 3.5 and Figure 5.1

The *glp-1* 3'-UTR reporter strain used in Figure 4.11 was generated by Brian M. Farley

CHAPTER I

Introduction to MEX-5 and MEX-3; post-transcriptional control in nematode development

Embryogenesis is one of the most tightly regulated processes in the life of an animal. From the time of fertilization, a large number of decisions are made that cause a single pluripotent cell to differentiate into a living organism. One remarkable feature across almost all organisms is that these developmental events occur in cells at a time when there is little or no transcription. For example, the Zebrafish embryo does not begin zygotic transcription until the 512-cell stage when the midblastula transition occurs (Kane and Kimmel, 1993). In most species, zygotic transcription does not occur until after several rounds of division. At this point, embryos have already made a number of critical decisions such as establishing the embryonic axis, the timing of cell divisions and cell-fate specification. Therefore, the early embryo depends upon post-transcriptional regulatory mechanisms to control the variations in gene expression needed for these events to occur (Farley and Ryder, 2008; Kuersten and Goodwin, 2003). Prior to the onset of transcription, the maternal load is essential for the fertilized zygote to successfully undergo differentiation. Maternally supplied cytoplasmic factors including RNA binding proteins and mRNAs are required to decode the genetic information supplied to the early embryo. However, very little is known as to how an mRNA is recognized by specific RNA binding proteins as well as the network of RNA targets that are identified by such proteins.

One organism that has been used to study post-transcriptional control of gene expression in early development is the nematode *Caenorhabditis elegans*. In 1974 Sydney Brenner introduced *C. elegans* as an organism to study various

processes including germline development and embryogenesis (Brenner, 1974). This free-living soil nematode is a convenient animal to study these processes. *C. elegans* have a rapid period of embryogenesis of approximately 16 hours. Moreover, there are a fixed number of somatic cells and the entire lineage of each cell has been mapped (Sulston and Horvitz, 1977; Sulston et al., 1983). This makes it possible for the detailed study of gene function and understanding fundamental processes such as cell-fate specification at single cell resolution. In addition, worms are transparent, allowing all of the developmental stages to be viewed from a microscope; a variety of molecular techniques may be employed to visualize factors that are involved in early development.

Anatomy of the adult C. elegans germline

C. elegans are predominantly hermaphroditic organisms, meaning that they have both male and female reproductive organs. Several tightly regulated developmental steps are required for the worm to grow into a functional sexual reproductive organism. A wild-type adult hermaphrodite consists of two gonads where gametes are produced (Figure 1.1). Within each gonad, the distal end consists of a population of mitotically dividing stem cells. As these cells progress away from the mitotic region, they transition into the meiotic cell cycle. When the mitotic germ cells transition into meiosis, their cell membrane breaks down forming a syncytium where the nuclei migrate to the gonad wall. In this region, the same cytoplasm is shared by all nuclei. This is called the pachytene stage

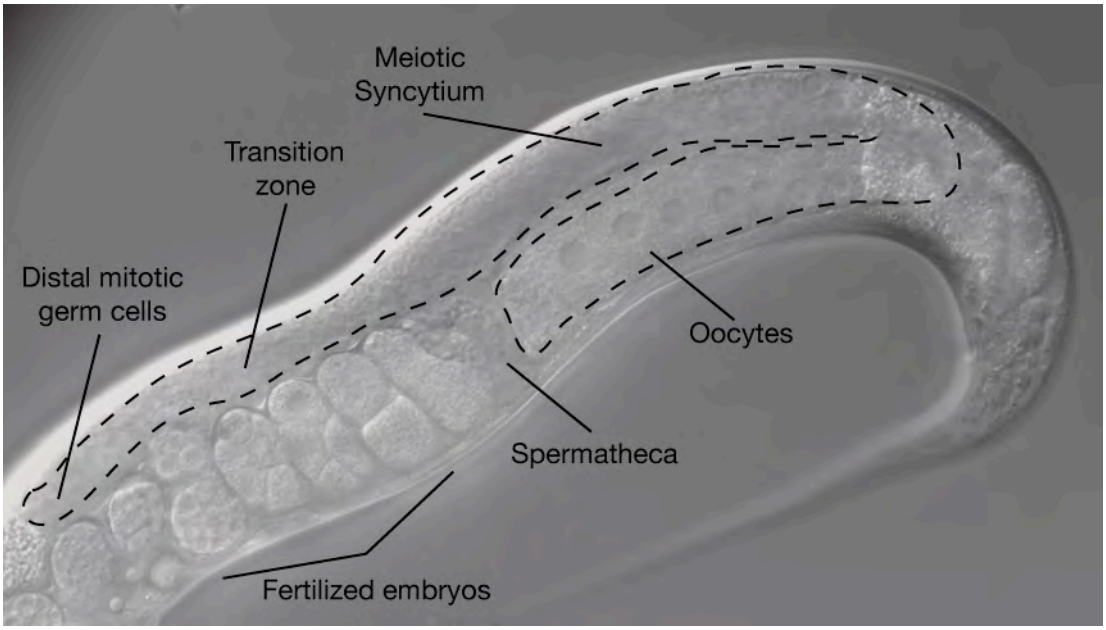

Meiotic
Syncytium
Transition
zone
Distal mitotic
germ cells
Oocytes
Spermatheca
Fertilized embryos

Figure 1.1

Figure 1.1: Anatomy of a *C. elegans* germline (dotted line). The distal end consists of a population of mitotically dividing stem cells. These transition into meiosis forming a syncytium where nuclei migrate to the walls of the gonad. At the bend of the gonad arm, nuclei recellularize and form ooctes. The most proximal oocyte is fertilized as it passes through the spermatheca. Fertilized embryos are retained in the adult uterus for several rounds of division before being layed.

where nuclei are arrested in meiosis I. Here, the cytoplasmic content of the oocyte is synthesized. Nuclei then recellularize as they move further down the gonad arm forming immature oocytes that are arrested in diakinesis. The most proximal oocyte then undergoes maturation and ovulates into the spermatheca where it is fertilized by spermatocytes.

Embryonic development in C. elegans

As a result of fertilization, the *C. elegans* embryo rapidly undergoes several polarized events that are critical for establishment of the anterior/posterior axis and segregating a number of determinants that are necessary for cell-fate specification (Pellettieri and Seydoux, 2002). The initial cue of axis polarity is dependent upon the sperm, which triggers downstream events causing changes with the internal cytoplasmic flow and cytoskeletal rearrangements (Munro et al., 2004). Furthermore, a network of maternal factors begin establishment of the anterior/posterior axis by coordinating the polarized distribution of several cytoplasmic factors, including a number of putative RNA binding proteins (Bowerman et al., 1997; Pellettieri and Seydoux, 2002). Together, these events set the stage for a series of cellular divisions that define the fate of a number of all blastomeres within the early embryo.

The anterior/posterior axis of the worm is determined as soon as the sperm enters the mature oocyte. Prior research has discovered that the position of sperm entry designates the posterior end of the embryo (Goldstein and Hird,

1996). Goldstein and Herd demonstrated that by manipulating the entry point of the sperm, it was possible to change the anterior/posterior axis of the zygote, but otherwise the embryo underwent normal development. Later studies showed that the spatial cue for polarity is actually derived from nucleation of microtubules from the sperm-derived centrosomes (or sperm asters) (Sadler and Shakes, 2000; Wallenfang and Seydoux, 2000). The actin cytoskeleton has also been shown to play a role in initiating cell polarity with non-muscle myosin II heavy chain (NMY-2). By injecting *nmy-2* antisense RNA into *C. elegans* ovaries, Guo and Kemphues observed embryonic partitioning defects and mislocalization of downstream polarity factors (Guo and Kemphues, 1996). Another study further established that the one-cell embryo contains a dynamic and contractile actomysin network that becomes destabilized at the sperm entry point (Munro et al., 2004). As a result, a flow of cortical NMY-2 and F-actin is initiated toward what becomes the anterior pole.

After the initial cue from the sperm entry, a network of maternal factors termed PAR proteins (abnormal embryonic PARtitioning of cytoplasm) further establishes the anterior/posterior axis of the zygote. The *par* genes encode PDZ domain proteins (PAR-3 and PAR-6) and kinases (PKC-3, PAR-1, and PAR-2) (Boyd et al., 1996; Etemad-Moghadam et al., 1995; Guo and Kemphues, 1995; Hung and Kemphues, 1999; Tabuse et al., 1998). As a result of the cortical flow established by the sperm entry, PAR-3, PAR-6, and PKC-3 are localized to the anterior pole, whereas PAR-1 and PAR-2 are localized to the posterior. The

localization of the PAR proteins is interconnected in such a way that PAR-3, PAR-6, PKC-3 antagonizes PAR-2. PAR-1 localization is dependent upon all of the other PAR proteins. These asymmetries create distinct cortical domains that are critical in maintaining anterior/posterior polarization (Cuenca et al., 2003; Nance, 2005). Two other proteins, PAR-4 and PAR-5 are distributed uniformly in the embryo, yet mutation of either gene prevents cell polarity (Morton et al., 2002; Watts et al., 2000). Consequently, the partitioning established from the PAR proteins coordinates downstream events that are pivotal for cell-fate specification. This cascade of events causes the asymmetric distribution of several cytoplasmic factors, including a number of putative RNA binding proteins that are essential for specifying various cell types (Figure 1.2).

Cell-fate Specification

As the proximal oocyte becomes fertilized, the early embryo undergoes several asymmetric divisions, giving rise to a number of founder cells –each specifying a different tissue type (Figure 1.3) (Rose and Kemphues, 1998). The establishment of the anterior/posterior axis causes the zygote to divide into two unequal daughter cells. The larger anterior blastomere is AB and the smaller posterior blastomere is P_1. AB divides to give ABa and ABp, where differing fates including establishing right/left asymmetries result from interactions with neighboring cells. The AB founder cell ultimately specifies the anterior pharynx, hypodermis and neurons. P_1 on the other hand goes through another

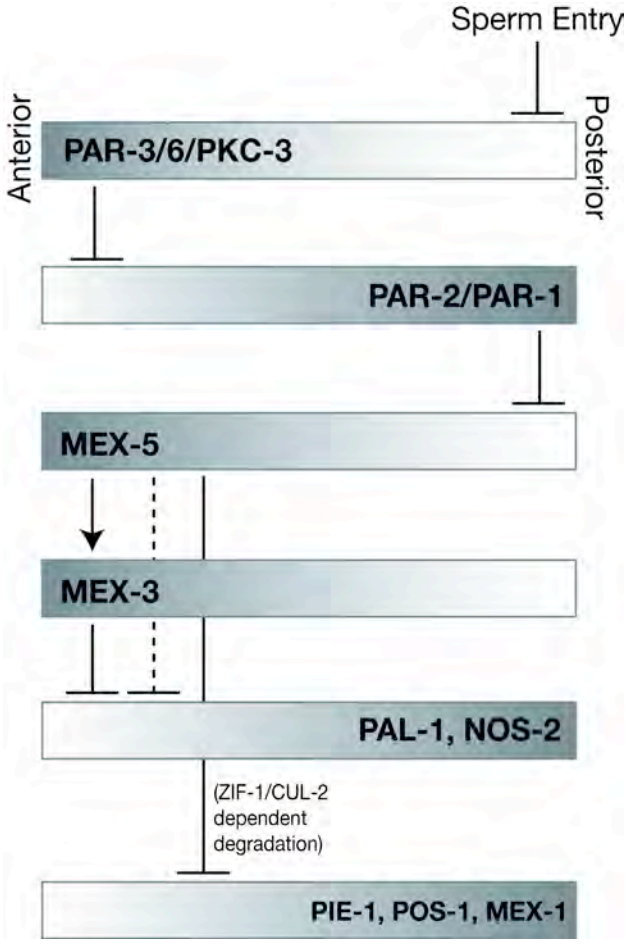

Sperm Entry
Anterior
Posterior
PAR-3/6/PKC-3
PAR-2/PAR-1
MEX-5
MEX-3
PAL-1, NOS-2
(ZIF-1/CUL-2
dependent
degradation)
PIE-1, POS-1, MEX-1

Figure 1.2

Figure 1.2: Epistatic model describing the distribution of maternal factors. Sperm entry causes the asymmetric distribution of PAR proteins. The PAR network causes MEX-5 and and MEX-3 to be restricted to the anterior blastomere(s) of the early embryo. MEX-5 and MEX-3 are required for the polarized accumulation of several cell-fate determinants to the posterior blastomere(s). This schematic summarizes polarized events that occur between the one to four-cell stage embryo in *C. elegans*. The shaded region of each rectangle represents higher levels of protein.

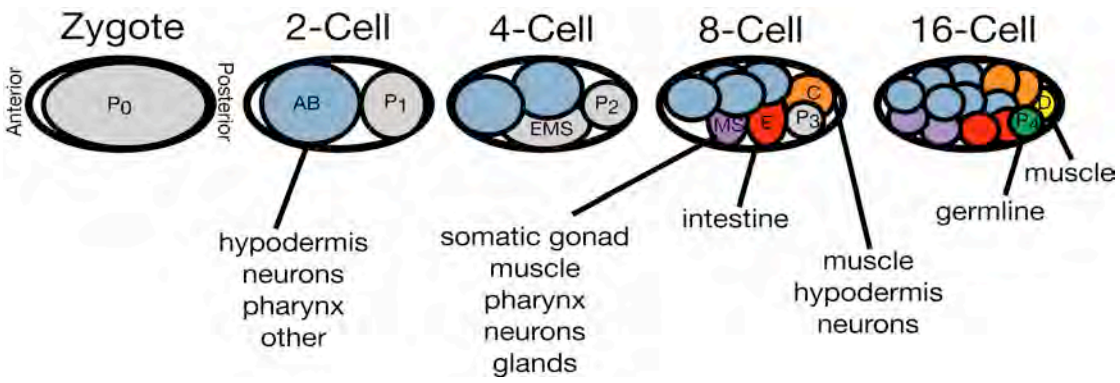

Zygote
2-Cell
4-Cell
8-Cell
16-Cell
Anterior
Posterior
P0
AB
P1
EMS
P2
C
MS
E
P3
D
P4
muscle
germline
hypodermis
neurons
pharynx
other
somatic gonad
muscle
pharynx
neurons
glands
intestine
muscle
hypodermis
neurons

Figure 1.3

Figure 1.3: Early *C. elegans* embryogenesis. The fertilized zygote begins with several asymmetric cellular divisions. This gives rise to several founder cells (AB, MS, E, C, D, and P_4) where each cell has a defined type of tissue that it will specify. The various kinds of tissue that are produced by each founder cell are labeled.

asymmetric division yielding EMS and P_2. The daughters of EMS form the founder cells E, which specifies intestine, and MS, specifying body wall muscle and pharynx. P_2 then goes through another asymmetric division resulting in P_3 and the founder cell C, specifying muscle tissue and hypodermis. The last founder cells are established as P_3 divides resulting with the daughter cells D and P_4, where D specifies muscle and P_4 being the precursor to the germline lineage. In total, early patterning of the *C. elegans* embryo results in six critical founder cells that determine the entire tissue composition of the worm.

Founder cells are specified by maternally supplied cell-fate determinants. These include a number of transcription factors such as SKN-1, PAL-1, and PIE-1. The gene *skn-1* encodes a bZIP class transcription factor that is required to specify EMS fates (Bowerman et al., 1992). PIE-1 on the other hand is CCCH finger protein required to specify the fate of the P_2 blastomere and its descendents (Mello et al., 1992). One of its main roles in embryogenesis is to repress transcription in the germline lineage. The gene *pal-1* encodes a caudal like homeodomain protein required to specify the C and D blastomeres (Hunter and Kenyon, 1996). It activates an embryonic muscle gene regulatory network in these cell lineages (Lei et al., 2009). These examples are just a few of several cell-fate determinants that are critical for the specification of various founder cells in the early embryo. Their spatiotemporal regulation is essential for correct patterning early on in development.

AB Lineage

Initially, ABa and ABp are equivalent and eventually take on very different cell-fate pathways. ABa produces cells that form the pharynx, while ABp does not generate pharyngeal cells, but rather neuronal and excretory cells (Sulston et al., 1983). Remarkably, both blastomeres can replace each other and take the other's place. Priess and Thomson performed an experiment where ABa and ABp were interchanged early upon division; the resulting embryo developed into a newly hatched larva with organs that appeared completely normal (Priess and Thomson, 1987). At the four-cell stage of the embryo, the AB daughter ABp receives a signal from the blastomere P_2. This signal causes ABp to be different from ABa and contributes to dosal/ventral polarity. Indeed, if ABp is not in contact with P_2, ABp will develop in a similar manner to ABa.

The Notch receptor GLP-1 is a major determinant required to specify the AB lineage. The gene *glp-1*, which plays a critical role in germline proliferation, is also essential for cell polarity between ABa and ABp. Temperature sensitive *glp-1* mutant worms, when shifted to nonpermissive temperatures, result in embryos where the ABp blastomere is abnormal and differentiates similar to ABa (Hutter and Schnabel, 1994). Furthermore, work done by Mello and Priess identified a second maternal gene, *apx-1*, that works with *glp-1* to specify the ABp blastomere (Mello et al., 1994). In a screen for maternal effect lethal mutations with excess production of pharyngeal tissue, mutations in the *apx-1* gene were identified that produced twice as much pharynx as in wild-type. This mutant has

a phenotype very similar to the *glp-1* mutant. The *apx-1* gene encodes a Delta ligand that is localized to the boundary of P_2 and ABp (Mickey et al., 1996). GLP-1 is present on the surface of both AB cells, but only ABp is in position to receive the APX-1 signal (Evans et al., 1994).

There are several other factors that are essential for specifying the anterior founder cell AB and preserving its integrity. These include genes that function as translational regulators, controlling the spatiotemporal expression of maternally supplied factors. For example, *glp-1* mRNA is present throughout the entire embryo upon fertilization and early embryogenesis. Yet, the protein isn't expressed until the two-cell stage and only in the anterior blastomeres (Evans et al., 1994). Alternatively, the translational repression of other posterior determinants including *skn-1*, *pal-1, pie-1*, and many others depend on factors that restrict their expression from the anterior end. The RNA binding proteins MEX-5 and MEX-3 are two anterior determinants that are essential for preventing AB from taking on a posterior-like fate; they are predicted post-transcriptional regulators of maternally supplied mRNAs in AB.

Post-transcriptional regulation of maternal RNAs in C. elegans development

As alluded to earlier, post-transcriptional regulation of maternal RNAs is one of the primary mechanisms to control gene expression in developmental processes –this includes both germline development and embryogenesis. At a time when transcription is largely quiescent, variation in gene expression is

achieved by post-transcriptional regulatory mechanisms. Most mRNAs that are required for early embryogenesis are transcribed in the germline from germ cells during mitosis or early meiosis and silenced until later developmental stages. Oftentimes misregulation of these factors is destructive to the worm resulting in sterility or embryonic lethality.

It is known that *cis*-regulatory elements present in 3'-UTRs of maternal transcripts are necessary for the timing and location of translation. A recent study has demonstrated that the 3'-UTRs of maternal mRNA are the primary regulators of gene expression in the germline (Merritt et al., 2008). RNA binding proteins are predicted to function as *trans*-acting factors that interact specifically with functional elements present in 3'-UTRs to promote mRNA turnover, RNA localization, or regulate its translatability. A large number of RNA binding proteins have been identified that are required to regulate mRNAs spatially and temporally (Hunter, 2005). Understanding how these proteins recognize their RNA targets is essential to identify *bona fide* regulatory targets and delineate regulatory networks responsible to specify cell-fate. As such, we will discuss the two major anterior determinants MEX-5 and MEX-3 in greater detail.

Introduction to the anterior determinant MEX-5

The anterior determinant MEX-5 is required to establish somatic/germline asymmetry in the early embryo. The *mex-5* gene was first discovered by Schubert, *et al.* in a genetic screen for maternal-effect lethal mutants with a

muscle excess phenotype (*mex* phenotype) where the embryo develops an abnormally large amount of muscle tissue (Schubert et al., 2000). This phenotype is very similar to mutants where the transcription factor SKN-1 is misexpressed in the anterior blastomeres. In a *mex-5;skn-1* double mutant embryos no longer express excess muscle tissue, and experiments showed that SKN-1 protein is misexpressed in the anterior of a *mex-5* mutant. This demonstrates that MEX-5 is required for the spatiotemporal expression of SKN-1 preventing its expression in the anterior and that MEX-5 plays a crucial role in specification of the anterior blastomere. It is not clear whether MEX-5 regulates *skn-1* directly.

In the same study, the Priess lab also identified the homolog *mex-6* based on sequence identity (Schubert et al., 2000). MEX-6 is approximately 70% identical and 85% similar to MEX-5 in amino acid sequence. However, RNAi experiments as well as *mex-6* mutants reveal that this gene is non-essential. Knockdown of *mex-6* levels or *mex-6* alleles results in viable progeny that grow into normal adults. Yet, *mex-5;mex-6* double mutants result in a highly penetrant terminal phenotype with embryos that lack germ cells and anterior muscles. Based on sequence similarity its thought that these proteins have partially overlapping biochemical functions.

MEX-5 is a CCCH-type tandem zinc finger protein

MEX-5 contains two predicted CCCH type zinc finger domains in its amino acid sequence (Schubert et al., 2000). This motif is a conserved RNA binding domain that is present in several well-known RNA binding proteins. One of the best-characterized CCCH TZF proteins is the mammalian homolog tristetraprolin (TTP) (Lai and Blackshear, 2001; Lai et al., 1999). TTP is required to regulate levels of tumor necrosis factor α (TNF-α). TTP is induced by TNF-α and other factors, functioning in a negative feedback loop regulating the concentration of TNF-α. Regulation of TNF-α is at the mRNA level as TTP binds to an AU-rich element (ARE) in the 3'-UTR promoting its deadenylation and degradation (Lai et al., 2000). Biochemical studies have identified the sequence requirements for high affinity binding by the zinc finger domain of TTP. Using fluorescence polarization measurements, members of the Wilson lab performed mutagenesis studies on the minimal ARE substrate and identified nucleotides that are necessary for TTP binding (Brewer et al., 2004). TTP binds with highest affinity (K_d <20nM) to RNA substrates containing AUUUA or AUUUUA flanked by uridylate residues.

An NMR structure of the TIS11d/ERF2 TZF domain, a mammalian homolog of TTP, bound to sequence UUAUUUAUU has been solved (Hudson et al., 2004). Specificity of RNA recognition is achieved by hydrogen bonds between Watson-Crick edges of the bases and amide and carbonyl functional groups of the ERF2 backbone. Additionally, this protein/RNA complex is

stabilized through stacking interactions of aromatic side chains and RNA bases. Further detail of this complex is described in Chapter III. As MEX-5 is a close family member of TTP and ERF2 (~50% identical), it is possible that it also recognizes RNA in a similar manner. However, previous studies show no evidence that this is the case.

PAR proteins function upstream of MEX-5

Upon fertilization, MEX-5 protein is restricted predominantly to the anterior blastomeres until the four-cell stage (see Chapter III for a detailed description). The localization of MEX-5 depends highly upon the PAR proteins, which includes the anterior determinant PAR-3 and the posterior determinant PAR-1 (Figure 1.2) (Schubert et al., 2000). Mutation of either of these genes results in equal distribution of MEX-5 in all blastomeres of the early embryo. A similar expression pattern of MEX-5 is observed when two other PAR proteins, PAR-4 or PAR-5 are depleted from embryos (Morton et al., 2002; Tenlen et al., 2008). On the contrary, the distribution of PAR proteins appears normal at the one-cell stage in a *mex-5;mex-6* double mutant embryo (Schubert et al., 2000). This implies that the PAR proteins are upstream of MEX-5. A recent study has shed some light on how PAR proteins function to control MEX-5 localization. It appears that phosphorylation of MEX-5 by PAR-1 and PAR4 are necessary for its anterior localization (Tenlen et al., 2008). Priess and coworkers suggest that phosphorylation of MEX-5 may promote its association with actomyosin, perhaps

restricting MEX-5 mobility. As an aside, MEX-5 is also suggested to play a role in promoting the expansion of the posterior domain of the one-cell zygote (Cuenca et al., 2003). PAR-1 and MEX-5 are thought to function in a feedback loop, decreasing levels of MEX-5 in the posterior, which limits the expansion of the posterior domain.

MEX-5 is required for the asymmetric distribution of cell fate determinants

MEX-5/6 (refers to experiments performed with or describing both MEX-5 and MEX-6) is required for the correct expression pattern of several cell-fate determinants (Schubert et al., 2000). Along with SKN-1, a number of germline proteins MEX-1, PIE-1 and POS-1, as well as the posterior determinant PAL-1 depend upon MEX-5/6 for their correct localization pattern (Figure 1.2). In addition, two proteins that are normally expressed in the anterior, GLP-1 and MEX-3, have decreased levels in *mex-5;mex-6* double mutants, suggesting that MEX-5 is required to promote their protein expression. Moreover, work by others has shown that MEX-5 may regulate the RNA abundance of *mex-1*, *pos-1*, and a third determinant *nos-2* (D'Agostino et al., 2006; Tenlen et al., 2006). Additionally, RNA-rich granules called P granules that specify the germline lineage are delocalized in *mex-5;mex-6* mutants. In wild-type embryos, P granules are polarized to the posterior germline lineage, yet in the absence of MEX-5/6 they are found in all blastomeres.

MEX-5 regulation of nos-2 mRNA

It is possible that MEX-5 is regulating some of the determinants mentioned above at the post-transcriptional level. Studies from the *C. elegans* Nanos homolog *nos-2* support this idea (Subramaniam and Seydoux, 1999). NOS-2 is required for the incorporation of primordial germ cells into the gonad. It also is necessary to maintain germ cell viability during larval development. The translation of *nos-2* mRNA is restricted primarily to the germline lineage being turned on in the precursor P4 at the 28-cell stage. Work by D'Agostino *et al.* has shown that the protein expression pattern of NOS-2 relies solely upon the 3'-UTR of *nos-2* (D'Agostino et al., 2006). A *nos-2* 3'-UTR reporter was shown to recapitulate the endogenous expression pattern of NOS-2. This reporter uses GFP fused to histone H2B, which drives the GFP signal to the nucleus. The GFP signal is expressed ectopically in all cells of the early embryo when levels of *mex-5/6* are reduced using RNAi. Moreover, it was shown that *nos-2* RNA is stabilized when knocking down *mex-5/6* RNA levels, suggesting that MEX-5/6 is required for the degradation of *nos-2* RNA. Based on these experiments, MEX-5/6 is predicted to be a post-transcriptional regulator of *nos-2* (Figure 1.2).

MEX-5 activation of ZIF-1

MEX-5 may not only function through post-transcriptional regulatory mechanisms. The somatic degradation of three TZF P-lineage determinants (PIE-1, POS-1 and MEX-1), along with MEX-5 was previously shown to require

one of the CCCH finger domains within each protein (Reese et al., 2000). The first zinc finger of PIE-1, POS-1 and MEX-1, and the second finger of MEX-5 are necessary to target their turnover in the somatic blastomeres, suggesting that these proteins may be targeted for degradation by factors that interact with the zinc finger domains. In an attempt to identify *trans*-acting factors that were responsible for CCCH finger dependent degradation, the Seydoux lab conducted a yeast-two hybrid screen for proteins that interact with the zinc fingers of PIE-1 (DeRenzo et al., 2003). In this study, they identified a new protein that they named ZIF-1 (zinc finger interacting factor 1) –a SOCS-box protein that interacts with and is thought to recruit the E3 ubiquitin ligase subunit elongin C. It is suggested that ZIF-1 acts as a substrate for E3 ubiquitin ligase subunits promoting the ubiquitination and degradation of target CCCH finger proteins. ZIF-1 was shown to interact with all zinc finger proteins mentioned above including MEX-5. Moreover, *zif-1* is required for their somatic degradation. These studies demonstrated that high levels of MEX-5 early in the soma are required to activate ZIF-1 dependent degradation of PIE-1, MEX-1, POS-1, and MEX-5 itself (Figure 1.2). Based on these experiments we can conclude that the localization of some cell fate determinants is at least in part due to a post-translational mechanism. Yet, it is not clear how MEX-5 activates *zif-1*. Understanding the RNA binding properties of MEX-5 may clarify this discrepancy.

Introduction to the cell fate determinant MEX-3

The KH (hnRNP K homology) domain protein MEX-3 is a second determinant that plays a vital role in specifying the AB blastomere in *C. elegans* early embryos. The gene *mex-3* was first identified in a screen for maternal-effect lethal mutations that caused proliferation of muscle in the anterior embryo (Draper et al., 1996). The anterior blastomeres of *mex-3* embryos take on a fate that is reminiscent of the posterior founder cell lineage C with excess muscle tissue and hypodermis. In addition, these embryos have additional germline precursor cells. As opposed to two germ cells, *mex-3* embryos have three to six cells that resemble germ cells.

Several years later, it was discovered that MEX-3 along with a second RNA binding protein, GLD-1, are essential in maintaining totipotency in the germline (Ciosk et al., 2006). The gene *gld-1* is required for meiotic cell cycle progression during oogenesis and worms that lack GLD-1 result in ectopic proliferation of mitotically dividing stem cells (Francis et al., 1995). It was then discovered that worms lacking both MEX-3 and GLD-1 develop transdifferentiated germline tumors that develop various somatic tissues such as neuron and muscle cells (Ciosk et al., 2006). In addition, MEX-3 was also shown to play a redundant role with another RNA binding protein, PUF-8, in promoting the mitotic proliferation of germline stem cells (Ariz et al., 2009). These findings demonstrate that MEX-3 has multiple roles, which include cell fate specification of the early embryo and preserving the integrity of the germline.

MEX-3 is a dual KH domain protein

MEX-3 contains two highly conserved KH RNA binding domains. Of the ten *mex-3* alleles identified, three contain a missense mutation within one of the KH domains (Draper et al., 1996). This motif is found in several known RNA and ssDNA binding proteins (Valverde et al., 2008). Though nucleic acid recognition by KH domains is versatile, they typically consist of a binding pocket that accommodates four unpaired bases mainly through Van der Waals forces and hydrophobic interactions. Yet, prior studies have not defined the requirements for specific RNA recognition by MEX-3.

C. elegans MEX-3 is orthologous to four human MEX-3 proteins (hMEX-3A, -3B, -3C, -3D/TINO) (Buchet-Poyau et al., 2007; Donnini et al., 2004). These proteins are discussed in greater detail in Chapter IV. Qualitative binding experiments suggest that these proteins also interact with RNA implying that hMEX-3 proteins are involved in post-transcriptional regulation (Courchet et al., 2008). However, the precise sequence specificity has not been determined for any of the human homologs.

MEX-3 expression is dependent upon MEX-5/6

In early embryos, MEX-3 is localized predominantly to the anterior blastomeres until the four-cell stage (see Chapter IV for a complete description of the MEX-3 expression pattern) (Draper et al., 1996). Expression of MEX-3 protein in embryos is dependent upon MEX-5/6 activity (Figure 1.2) (Schubert et

al., 2000). Embryos that lack MEX-5/6 have low levels of MEX-3 that is distributed uniformly as opposed to asymmetrically. Work from the Hunter lab suggests that MEX-5/6 may function to activate MEX-3 in the anterior through protein/protein interactions (Huang et al., 2002). Using a yeast-two hybrid screen for MEX-3 interacting proteins, a number of RNA binding proteins were identified, among which MEX-6 was discovered. They demonstrated that MEX-6 is necessary to promote MEX-3 activity in the anterior blastomeres ABa and ABp and describe how the absence of MEX-5/6 in the posterior may subject MEX-3 to inactivation and consequently turned over by other factors. However, the precise mechanism of this regulation remains unclear.

MEX-3 regulation of pal-1 and nos-2

MEX-3 is thought to function as a post-transcriptional regulator of maternal mRNAs to specify AB as well as germline development. Hunter and Kenyon revealed the cause of excess muscle in *mex-3* embryos as they studied the transcription factor *pal-1* (Hunter and Kenyon, 1996). PAL-1 is required to specify the posterior blastomere C for muscle tissue differentiation. They showed that this factor is misexpressed into the anterior blastomeres in a *mex-3* mutant background. They also found that injection of *pal-1* antisense RNA maternally in *mex-3* mutant worms dramatically reduced the amount of body-wall muscle cells observed in *mex-3* embryos. Based on these experiments, they

conclude that *pal-1* function is required for AB-derived muscle tissue in *mex-3* mutant embryos.

In conjunction with these studies, Hunter and coworkers showed that both MEX-3 and GLD-1 function to repress *pal-1* translation in the germline (Mootz et al., 2004). They demonstrated that GLD-1 mediates a developmental switch controlling *pal-1* levels in the distal germline. GLD-1 represses *pal-1* in the distal end of the germline and as GLD-1 levels diminish around the bend of the gonad arm, *pal-1* repression is taken over by MEX-3 as it accumulates in the proximal germline. As mentioned previously, a *gld-1;mex-3* double mutation gives a transdiffentiated germline phenotype, which includes somatic muscle tissue. This is partly due to misregulation of *pal-1* as in a *gld-1;mex-3;pal-1(RNAi)* worm, the number of muscle cells is reduced, partly rescuing the *gld-1;mex-3* phenotype (Ciosk et al., 2006).

A second putative target of MEX-3 is *nos-2*. Similar to *mex-5* mutant embryos, worms that are deficient in MEX-3 misexpress NOS-2 both spatially and temporally (Jadhav et al., 2008). As mentioned earlier, NOS-2 is not expressed until the germline precursor cell P4. However, in *mex-3* mutant worms, the nanos homolog is present in all blastomeres beginning at the one-cell stage. MEX-3 is believed to regulate both *nos-2* and *pal-1* mRNA translation in a 3'-UTR dependent manner (Figure 1.2); evidence that supports this is described in detail in Chapter IV.

It is clear that both MEX-5 and MEX-3 play critical roles in patterning the early embryo, specifically the anterior founder cell AB. Moreover, MEX-3 has been shown to function with the translational regulator GLD-1 to preserve the pluripotent state of the germline. Both MEX-5 and MEX-3 have highly conserved RNA binding domains and are thought to regulate a number of maternal mRNAs through post-transcriptional regulatory mechanisms (See Figure 1.2 for a summary of gene interactions). However, there has been little evidence demonstrating direct regulation of target mRNAs by MEX-5 and MEX-3. Furthermore, the requirements for RNA recognition by these two proteins have not been previously described.

This dissertation tests the hypothesis that MEX-5 and MEX-3 are specific RNA binding proteins that promote anterior development through targeted regulation of maternal transcripts. I provide a thorough dissection of the RNA binding properties of MEX-5 and MEX-3 using quantitative biochemical approaches. In Chapter II, I also present a brief survey of the predominant experimental strategy used to monitor protein/RNA interactions in the proceeding chapters. The remaining two chapters (Chapter III and Chapter IV) provide data that describe the basis of RNA recognition of MEX-5 and MEX-3.

CHAPTER II

Quantitative methods to monitor protein-nucleic acid interactions using fluorescent probes

Abstract

Sequence-specific recognition of nucleic acids by proteins is required for nearly every aspect of gene expression. Quantitative binding experiments are a useful tool to measure the ability of a protein to distinguish between multiple sequences. Here, we describe the use of fluorophore-labeled oligonucleotide probes to quantitatively monitor protein/nucleic acid interactions. We highlight two complementary experimental methods, fluorescence polarization and fluorescence electrophoretic mobility shift assays, that enable quantitative measurement of binding affinity. We also present two strategies for post synthetic end-labeling of DNA or RNA oligonucleotides with fluorescent dyes. The approaches discussed here are efficient and sensitive, providing a safe and accessible alternative to the more commonly used radioisotopic methods.

Introduction

A myriad of cellular processes require sequence specific recognition of a nucleic acid sequence by a protein. For example, transcription factors bind to specific DNA elements to enhance or repress transcription at a defined locus (Barrasa et al., 2007; Carrera and Treisman, 2008; Deplancke et al., 2006; Noyes et al., 2008). Likewise, RNA-binding proteins coordinate translation, mRNA localization and stability, pre-mRNA splicing through association with defined sequences in target transcripts (Iwasaki et al., 2009; Johnstone and Lasko, 2001; Jurica and Moore, 2003; Nilsen, 2002; Singh and Valcarcel, 2005; Varnum et al., 1991). As such, it is critically important to understand the basic mechanisms by which DNA and RNA binding proteins identify their appropriate target sequences.

There are many ways a protein can recognize a specific DNA or RNA sequence. Most DNA binding proteins rely on variations in the pattern of hydrogen bond acceptors and donors in the major groove to achieve sequence specific recognition (Seeman et al., 1976). Others identify specific sequences through their relative flexibility that induces distortions in the DNA duplex (Kim et al., 1993; McClarin et al., 1986). RNA-binding proteins can recognize specific sequences through readout of the hydrogen bonding patterns on the Watson-Crick and Hoogsteen faces of the bases, or through shape specific recognition of RNA secondary and tertiary structure (Hall, 2005; Howe et al., 1994; Hudson et al., 2004; Valegard et al., 1994).

To identify the determinants of specificity, it is often necessary to measure quantitative parameters of binding, such as relative binding affinity and stoichiometry (Ryder et al., 2004; Ryder and Williamson, 2004). There are a number of methods that can be used to monitor protein/nucleic acid complexes (Royer and Scarlata, 2008; Ryder et al., 2008; Wong and Lohman, 1993). One of the more prominent techniques is the electrophoretic mobility shift assay (EMSA) (Ryder et al., 2008). EMSA is a powerful technique that allows visualization of the interaction between a protein and a labeled DNA or RNA probe. Equilibration reactions that include a fixed concentration of a ^{32}P-radiolabeled oligonucleotide probe and varying amounts of protein are run on a native polyacrylamide gel to separate bound from free nucleic acid. Gels are dried and then exposed to film or a phosphorimager plate in order to determine the fraction of bound probe as a function of protein concentration. This data can be fit to determine the apparent equilibrium dissociation constant.

The use of radioactive material has the advantage of making the assay extremely sensitive. Very small amounts of ^{32}P can be detected using common equipment. Another advantage is that the labeled and unlabeled probes are chemically similar, reducing the chance that the binding reaction is perturbed due to indirect effects caused by the label. However, radiolabled probes also have several disadvantages, including safety, environmental, and regulatory challenges. Moreover, radioactive labels are costly, and due to the short half-life

of ^{32}P, probes have to be frequently re-labeled and radioisotope stocks re-ordered, costing time and money.

Practical alternatives to radioisotopes exist. The use of fluorescent probes has become a favorable alternative in many biochemical assays (Royer and Scarlata, 2008). With modern instrumentation, detection of fluorescent probes now rivals that of ^{32}P-labeled probes. In this survey, we present two methods to generate end-labeled fluorescent probes using commercially available fluorescent dyes and simple aqueous phase chemical reactions that can be performed in any laboratory (Czworkowski et al., 1991; Reines and Cantor, 1974). In addition, we present two complementary approaches to measure the affinity of a protein for labeled DNA and RNA probes that rely on different physical properties of the complex—fluorescence polarization (FP) and fluorescence EMSA (F-EMSA) (Hellman and Fried, 2007; Pagano et al., 2007; Ryder et al., 2008). Both assays can be applied to the same equilibration reactions because the FP measurements do not destroy the sample. Furthermore, the F-EMSA method conveniently enables the gels to be analyzed immediately after electrophoresis, avoiding the need to dry and expose the gel and reducing the time requirement of the assay.

We have used the approach described herein to study four RNA-binding proteins required for germline development and/or embryogenesis in the nematode *C. elegans* (Farley et al., 2008; Pagano et al., 2009; Pagano et al., 2007) and two proteins that regulate oligodendrocyte differentiation in

vertebrates. Others have used it to study the association of bacteriophage PP7 coat protein and vertebrate ZBP variants to their cognate RNA sequences (Chao et al., 2008; Chao et al., 2010). The data show that the method is generally applicable to a wide variety of protein-RNA complexes.

Labeling Strategies

5’-end labeling of DNA and RNA oligonucleotides

Chemically synthesized single-stranded DNA and RNA oligonucleotides can be efficiently post-synthetically labeled at the 5’-end using a two-step semi-enzymatic synthesis strategy described by Czworkowski and coworkers (Figure 2.1A) (Czworkowski et al., 1991). In the first step, T4 polynucleotide kinase (PNK) is used to transfer the γ-phosphorothioate group from 5'-*O*-(3-thio) adenosine triphosphate (ATPγS) to the 5’-hydroxyl of the oligonucleotide probe. Next, following an ethanol precipitation step, 5-iodoacetamidofluorescein (5-IAF) is reacted with the phosphorothioate group on the probe in HEPES pH 7.4-buffered water to create a covalent adduct between the oligonucleotide and the fluorophore. Unreacted dye is removed by ethanol precipitation followed by a size exclusion spin column. Analytical denaturing polyacrylamide gel electrophoresis is used to assess purity. The reaction yield is measured comparing the absorbance at 491 nm (fluorescein) to that at 260 nm (nucleotides and fluorescein, Figure 2.1B). We typically observe 80-95% labeling efficiency.

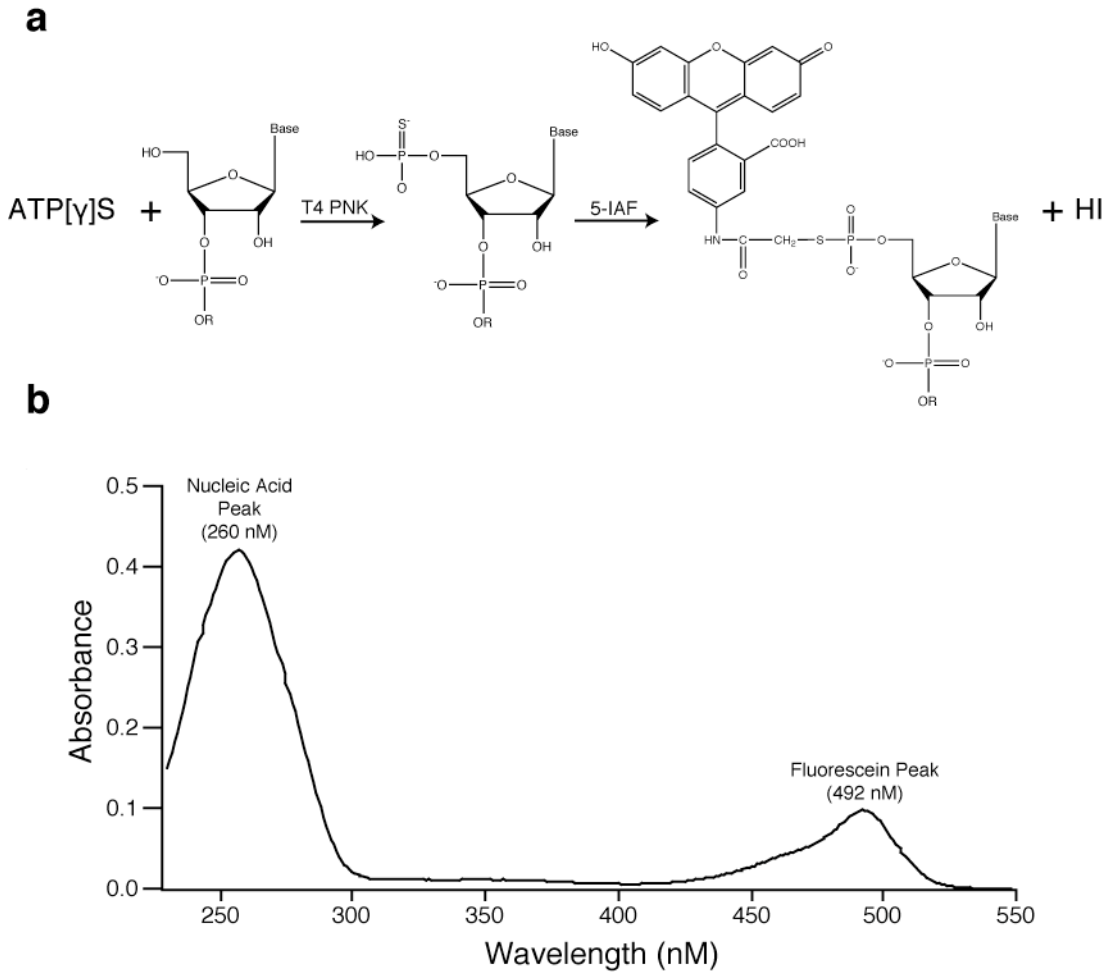

a
ATP[γ]S
+
T4 PNK
5-IAF
+ HI
Base
HO
OH
OR
COOH
b
Absorbance
Wavelength (nM)
Nucleic Acid Peak (260 nM)
Fluorescein Peak (492 nM)
0.0
0.1
0.2
0.3
0.4
0.5
250
300
350
400
450
500
550

Figure 2.1

Figure 2.1: 5’-end labeling of DNA and RNA oligonucleotides. **A.** The 5’-end labeling strategy involves reacting unphosphorylated DNA or RNA with T4 PNK and ATP[γ]S to yield a thiol reactive group on the 5’ end. 5-iodoacetamidofluorescein (5-IAF) is then reacted with the thiol group. **B.** An example of a UV spectrum of a DNA oligonucleotide that was labeled with 5-IAF on the 5’-end. The DNA absorbs at 260 nM and the fluorescein label absorbs at 492 nM.

The reaction is simple, efficient, and does not require pre-existing chemical modification of the DNA or RNA oligonucleotide. DNA probes labeled in this way can be used as primers in primer extension, PCR, or RT-PCR assays; the label does not interfere with subsequent enzymatic applications. The reaction is compatible with other iodoacetamide fluorescent dyes, including 7-diethylamino-3-((4'-(iodoacetyl)amino)phenyl)-4-methylcoumarin (DCIA) and tetramethylrhodamine-5-iodoacetamide, enabling a wide variety of downstream applications in several colors. Moreover, the approach is not limited to chemically synthesized DNA or RNA oligonucleotides. In principal, any substrate of PNK can be labeled in this manner.

3'-end labeling of RNA

In 1974, Reines and Cantor described a simple synthetic method to conjugate fluorescent dyes to RNA molecules at the 3'-end (Reines and Cantor, 1974). This approach takes advantage of the vicinal hydroxyls present only at the 3'-end of ribonucleotides to achieve specific labeling without the need to protect other sites in the RNA. In the first step of the reaction, the RNA is incubated with sodium periodate ($NaIO_4$) in mildly acidic sodium acetate buffer (pH 5.1) in order to oxidize the vicinal hydroxyls to aldehyes (Figure 2.2A). Following an ethanol precipitation step, the oxidized RNA is then reacted with fluorescein-5-thiosemicarbazide (FTSC) in order to generate a covalent thiosemicarbazone linkage with the RNA. This can then be reduced with sodium cyanoborohydride

to limit chemical reversibility, but this is not usually necessary. The labeled RNA is resolved from unreacted dye by ethanol precipitation followed by a size exclusion spin column. The products are analyzed as described above in the 5'-end labeling procedure (Figure 2.2B).

As above, the reaction is straightforward and does not require prior modification of the RNA oligonucleotide. Any molecule with vicinal hydroxyls can be labeled this way, including ribonucleotides and longer RNA molecules produced by *in vitro* transcription. This is an important caveat; if *in vitro* transcribed RNA is used as the substrate, care must be taken to purify the RNA away from other nucleotides present in the transcription reaction. Several aldehyde reactive dyes are available. The chemistry is the same for any thiosemicarbazide or hydrazide, including commercially available Alexa dyes (Invitrogen) in a wide variety of colors, and the useful affinity tag biotinamidocaproyl hydrazide (BACH). As with the prior strategy, labeling efficiency is usually greater than 80%.

There are, of course, other suitable labeling strategies, including PCR or *in vitro* transcription with fluorescent nucleotide analogs in order to body-label the nucleic acid; or incorporation of amino, azido, or thiol groups during chemical synthesis of DNA or RNA oligonucleotides that can react with a wide variety of fluorescent dyes. These strategies work well and can be used to make labeled material for the quantitative applications outlined below. However, we prefer the methods presented above as any oligonucleotide can be labeled without the

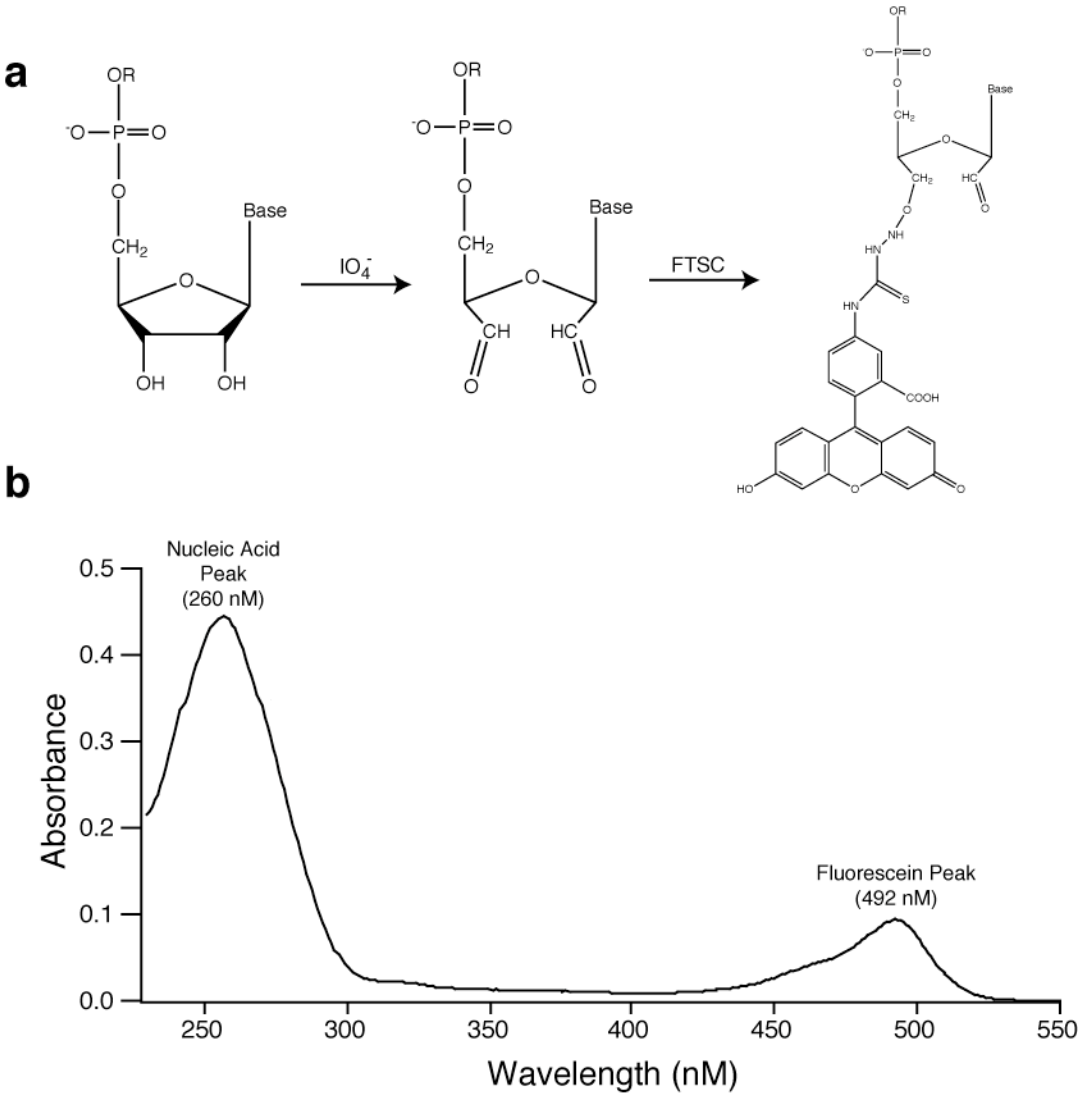

a
OR
Base
IO_4^-
FTSC
b
Nucleic Acid
Peak
(260 nM)
Fluorescein Peak
(492 nM)
Absorbance
0.0
0.1
0.2
0.3
0.4
0.5
250
300
350
400
450
500
550
Wavelength (nM)

Figure 2.2

Figure 2.2: 3’-end labeling of RNA oligonucleotides. **A.** The 3’-end labeling strategy involves reacting vicinal hydroxyls of an RNA molecule through sodium periodate cleavage followed by an addition reaction with Fluorescein 5-thiosemicarbazide (FTSC). **B.** An example of a UV spectrum of an RNA oligonucleotide that was labeled with FTSC on the 3’-end. The RNA absorbs at 260 nM and the fluorescein label absorbs at 492 nM.

need to incorporate reactive moieties during synthesis, and certain applications require an end label rather then random incorporation of a body label. Thus, the methods presented here remove a barrier to the use of fluorescent dyes by making post-synthetic labeling as convenient as current strategies used to radiolabel nucleic acids.

Quantitative fluorescence methods to monitor protein-nucleic acid interactions

FP and F-EMSA are two effective techniques that can be used to determine the binding affinity and specificity of a DNA or RNA-binding protein (Figure 2.3). Both assays rely on different physical properties of the complex; FP measures the extent that a complex tumbles within the lifetime of the fluorophore, while EMSA depends upon the ability of a protein to alter the migration rate of a labeled probe within an electric field in a gel. Both techniques require recombinant protein that has been purified to >95% homogeneity and a fluorescently labeled DNA or RNA probe.

Fluorescence polarization assays

Fluorescence polarization is a useful technique to study the thermodynamic and kinetic properties of a protein nucleic acid interaction. FP takes advantage of the change in tumbling properties of a fluorescent ligand upon binding to a larger macromolecule; in our case a fluorescent DNA or RNA probe and a nucleic acid

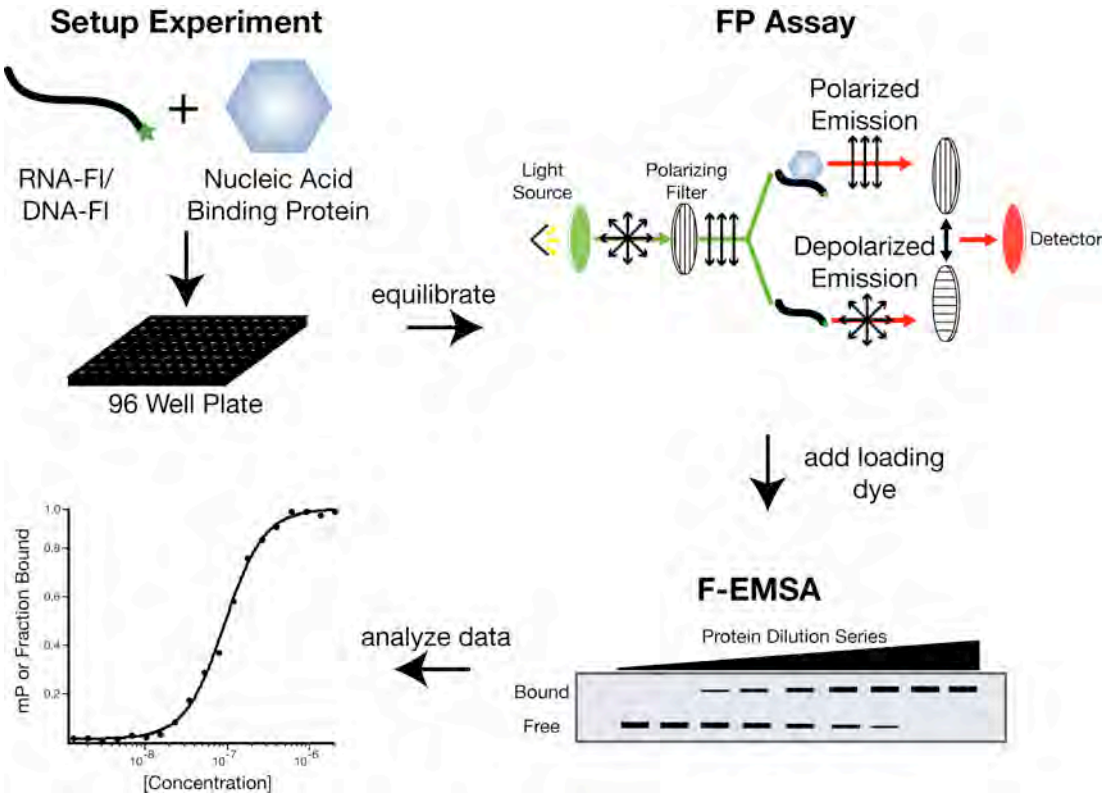

Setup Experiment
RNA-Fl/
DNA-Fl
Nucleic Acid
Binding Protein
96 Well Plate
equilibrate
FP Assay
Light
Source
Polarizing
Filter
Polarized
Emission
Depolarized
Emission
Detector
add loading
dye
F-EMSA
Protein Dilution Series
Bound
Free
analyze data
mP or Fraction Bound
[Concentration]

Figure 2.3

Figure 2.3: An overall schematic of the presented binding assays. The experiment is typically setup in a 96-well plate (black) format. A dilution series of the nucleic acid binding protein is equilibrated with trace amounts of fluorescently labeled oligonucleotides. After the experiment has been given time to reach equilibrium, the fluorescence polarization (FP) of each sample is measured. In the FP illustration shown, a sample is excited with polarized light. Emitted light is then measured using filters both parallel and perpendicular to the plain of excitation. Samples can then be run on a native gel (F-EMSA) directly after measuring FP. Measurements of either polarization (mP) or the fraction of RNA bound is then plotted as a function of protein concentration. The experimental design described provides two complementary approaches that depends upon different physical properties of the complex and enables quantitative measurements of the interaction.

binding protein. Polarized light is used to excite the fluorophore, and emitted light is measured in planes both parallel and perpendicular to the plane of excitation. A labeled probe that is free in solution tumbles more rapidly because of its comparatively low molecular weight, leading to depolarization of the emitted light. However, tumbling is reduced when the nucleic acid is bound to protein, causing an increase in polarization. The maximal extent of the increase is dependent on the size of the protein, the size of the labeled probe, and the lifetime of the fluorophore.

To measure the apparent affinity of the protein nucleic acid complex, a series of equilibration reactions are set up with varying concentration of protein and fixed trace amounts of labeled probe. After equilibration, polarization is determined using using a fluorometer or a fluorescence plate reader equipped with polarizers. The extent of polarization is influenced by the relative concentration of each fluorescent species in the equilabration reaction. As such, polarization is directly proportional to the fraction of bound probe at each protein concentration. It is also important to note that if higher order species are formed, or if there is unreacted dye in the equilibration reaction, these will contribute to the apparent polarization value as well. The effective polarization, expressed in units of millipolarization (*mP*), is related to the fluorescence intensity (*I*) in the parallel (*para*) and perpindicular (*perp*) planes by equation (i). Most plate readers capable of FP measurements output this value automatically.

(i) $$mp = 1000 \times \left[\frac{Ipara - Iperp}{Ipara + Iperp} \right]$$

The data are then plotted as a function of protein concentration and fit to the Hill equation (ii) in order to determine the apparent equilibrium dissociation constant (K_d) and the apparent cooperativity (n) between a macromolecule (P_t) and its ligand (Figure 2.4). The equation presented here uses the maximum (m) and base (b) signals as normalization factors that represent the polarization values at the upper and lower asymptote of the titration (Hill, 1910).

(ii) $$\phi = b + (m - b)\left[\frac{1}{1 + (K_d / [P_t])^n} \right]$$

Other parameters that can be measured by this approach include the stoichiometry of the complex, determined by repeating the equilibration reactions at elevated probe concentration and fitting mP as a function of molar ratio to a quadratic equation derived by Rambo and coworkers (Rambo and Doudna, 2004). FP is especially amenable to kinetic measurements, as the change in polarization can be measured in real time in many instruments.

The primary limitation of FP is that it requires a relatively small labeled ligand, usually ten kilodaltons or smaller when using fluorescein dyes. The lifetime of the fluorophore defines the size limit of the ligand. If the labeled probe

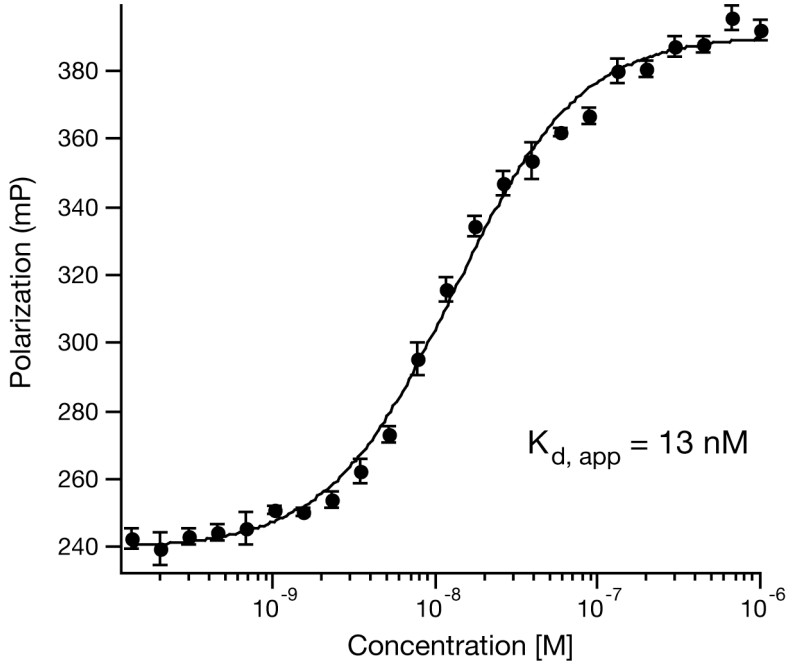

Polarization (mP)
380
360
340
320
300
280
260
240
10^{-9}
10^{-8}
10^{-7}
10^{-6}
Concentration [M]
$K_{d, app}$ = 13 nM

Figure 2.4

Figure 2.4: A sample fluorescence polarization assay. Raw fluorescence polarization data of the RNA binding protein MEX-3 interacting with one of its target RNA sequences is shown (Pagano et al., 2009). The data is given in units of millipolarization (mP) and the equilibrium dissociation constant ($K_{d,app}$) is given for the complex. Shown is a single replicate experiment.

cannot efficiently depolarize the emitted light within the lifetime of the fluorophore, it is not possible to measure the change in depolarization caused by binding to a protein. Other factors that contribute to the size limit include the shape and flexibility of the ligand. We typically use this approach for single stranded RNA probes that are thirty nucleotides or less. Smaller probes may be necessary for double stranded DNA or RNA.

Electrophoretic mobility shift assays

A different but complementary approach to measure the affinity of an interaction between a nucleic acid and a protein is EMSA. This method relies on the ability of a protein to influence the migration of a labeled nucleic acid through a native polyacrylamide or agarose gel (Hellman and Fried, 2007; Ryder et al., 2008). The migration depends upon the length, and as a result the overall charge, as well as the shape of the nucleic acid. If association of a protein perturbs these parameters, it can change the rate of migration, enabling separation of bound and free probe. As with FP, a series of equilibration reactions are set up such that protein concentration varies while a fixed, trace amount of labeled probe remains the same in each reaction. After equilibration, the reactions are loaded onto the gel and subjected to electrophoresis in order to separate bound from free RNA. Then, the fraction of bound RNA is determined, and then fit as a function of protein concentration to the Hill equation as above. The primary advantage of this approach is that it enables visualization of both

free and bound probe. Thus, if multiple proteins bind to the DNA or RNA sequence, this usually manifests as multiple shifted species. Moreover, if the probe has degraded, this can be assessed as well. The disadvantage of this approach is that it does not provide a true equilibrium measurement. Complexes with fast off rates will dissociate in the time it takes to load the sample, perturbing the apparent affinity.

Traditionally, EMSA is performed with radioactively labeled oligos in vertical gel equipment (Hellman and Fried, 2007; Ryder et al., 2008). We have observed that fluorescent probes provide sufficient sensitivity so that they can be used in experiments that require that the labeled oligo be present in low nanomolar amounts. To achieve sufficient sensitivity, we adapted the EMSA approach to horizontal submarine polyacrylamide gels, which enables loading of more sample per gel compared to standard vertical gel equipment. Moreover, the horizontal format enables multiplexing of experiments as more than one comb can be used to create wells in the gel matrix. Immediately following electrophoresis, the gel is imaged using a fluorescent capable phosphoimager equipped with the appropriate laser to excite the fluorophore (Fuji FLA-5000 or similar, Figure 2.5). The assay conveniently uses the same experimental setup used in FP, so only a single series of binding reactions is required for both assays. Since the FP is nondestructive, samples may be electrophoresed immediately following the polarization measurement. Samples are run on a

native acrylamide gel separating free nucleotides from those that are bound to protein.

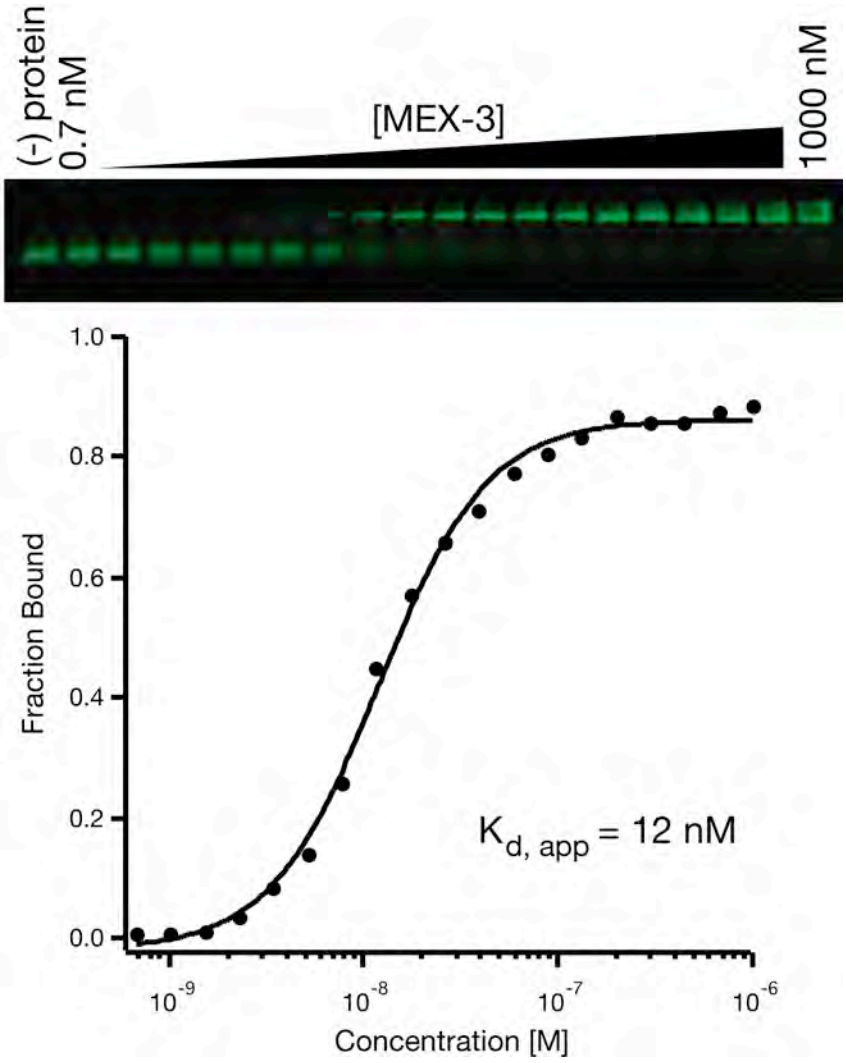

(-) protein
0.7 nM
[MEX-3]
1000 nM
1.0
0.8
0.6
0.4
0.2
0.0
Fraction Bound
$K_{d, app}$ = 12 nM
10^{-9}
10^{-8}
10^{-7}
10^{-6}
Concentration [M]

Figure 2.5

Figure 2.5: A sample fluorescent gel shift assay. The RNA binding protein MEX-3 interacting with one of its target RNAs is shown (Pagano et al., 2009). Fluorescently labeled RNAs are shown in the gel as green bands. A plot of the fraction RNA bound at varying MEX-3 concentrations is given. The equilibrium dissociation constant ($K_{d,app}$) is given for the complex.

Concluding Remarks

In this survey, we focused primarily on quantitative approaches to monitor protein-nucleic acid interactions with the use of fluorescently labeled probes. There are a number of ways to label an oligonucleotide using chemical and enzymatic strategies. Labeled probes can be used in both FP and EMSA to provide a dual readout of the same experimental setup. This approach is fast, safe, and affordable providing an effective strategy to perform quantitative biochemical experiments to dissect protein/nucleic acid interactions *in vitro*. These approaches are adaptable to many different strategies demonstrating the versatility and usefulness of fluorescence methodologies.

CHAPTER III

Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5

Abstract

Embryonic development requires maternal proteins and RNA. In *Caenorhabditis elegans*, a gradient of CCCH tandem zinc finger (TZF) proteins coordinates axis polarization and germline differentiation. These proteins govern expression from maternal mRNAs by an unknown mechanism. Here we show that the TZF protein MEX-5, a primary anterior determinant, is an RNA-binding protein that recognizes linear RNA sequences with high affinity but low specificity. The minimal binding site is a tract of six or more uridines within a nine to thirteen nucleotide window. This sequence is remarkably abundant in the 3'-untranslated region of *C. elegans* transcripts, demonstrating that MEX-5 alone cannot specify mRNA target selection. In contrast, human TZF homologs tristetraprolin (TTP) and ERF-2 bind with high specificity to UUAUUUAUU elements. We show that mutation of a single amino acid in each MEX-5 zinc finger confers TTP-like specificity to this protein. We propose that divergence of this discriminator residue modulates the RNA-binding specificity in this protein class. This residue is variable in nematode TZF proteins, but is invariant in other metazoans. Therefore, the divergence of TZF proteins and their critical role in early development is likely a nematode-specific adaptation.

Introduction

Embryogenesis is the process by which a fertilized oocyte transforms into a multicellular organism. Although the zygote contains all of the information required for development, zygotic DNA alone is not sufficient to drive patterning. Somatic cell nuclear transfer experiments, like those used to clone Dolly the sheep, demonstrate that maternal factors present in the oocyte cytoplasm are needed for the initiation of development (Wilmut et al., 1997). These maternal factors are proteins and quiescent mRNAs (Seydoux and Braun, 2006); they coordinate early development prior to the onset of zygotic transcription.

In the nematode worm *Caenorhabiditis elegans*, polarization of the body axes occurs after fertilization and requires several highly conserved maternal factors termed PAR proteins (Etemad-Moghadam et al., 1995; Goldstein and Hird, 1996; Guo and Kemphues, 1995; Kemphues et al., 1988; Kirby et al., 1990; Levitan et al., 1994; Tabuse et al., 1998; Wallenfang and Seydoux, 2000). Prior to fertilization, these proteins are uniformly distributed in the cytoplasm. Once the sperm penetrates the oocyte, they localize to opposing cortical domains in a process that requires microtubules derived from the asters of the sperm pronucleus. The PAR network coordinates asymmetric translation of several cell signaling proteins (Mello et al., 1994; Thorpe et al., 1997) (*glp-1, apx-1, mom-2,* and *mom-5*) and transcription factors (Maduro et al., 2005) (*skn-1, pal-1,* and *pop-1*) encoded by maternal mRNAs present throughout the 1-cell embryo. The

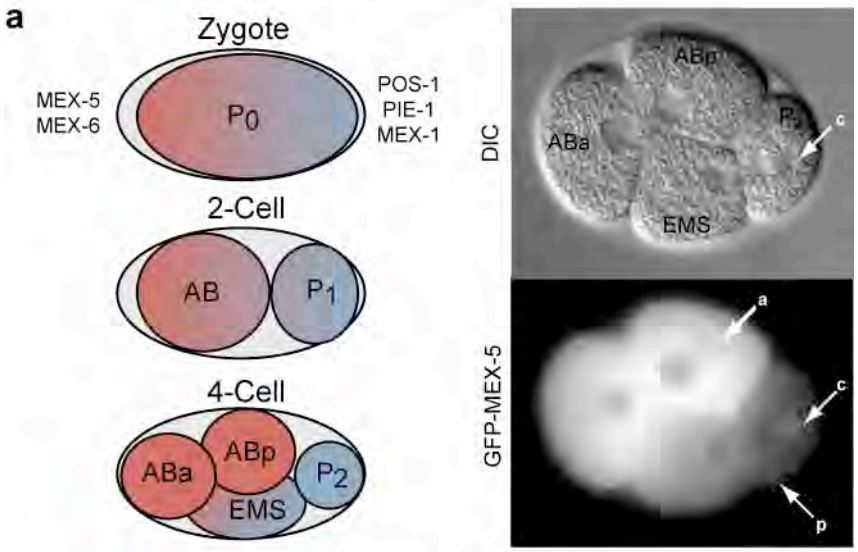
PAR proteins are thought to locally deactivate maternal RNA-binding proteins thereby modulating the stability or translation efficiency of maternal mRNAs.

Consistent with this hypothesis, posterior localization of PAR-1 promotes anterior localization of two putative RNA-binding proteins, MEX-5 and MEX-6 (Cuenca et al., 2003; Schubert et al., 2000). Though these proteins are 70% identical, their contributions to development only partially overlap. Disruption of *mex-5* causes embryonic death with a terminal phenotype that includes proliferation of muscle (MEX = Muscle Excess). In contrast, deletion of *mex-6* does not affect viability. Two major roles in development have been attributed to MEX-5. First, it controls segregation of the germline from the soma by activating *zif-1*, which promotes anterior turnover of three germline-specific maternal proteins (POS-1, PIE-1, and MEX-1) (DeRenzo et al., 2003). The overall result of this pathway is a gradient of MEX-5/6 from anterior to posterior and an opposing gradient of POS-1, PIE-1, and MEX-1 (Figure 3.1a). Second, MEX-5 plays a relatively uncharacterized role in maintaining PAR polarity via a feedback loop with PAR-1 (Cuenca et al., 2003). It is not yet clear if the two roles are linked at the molecular level. Additionally, there may be other roles for MEX-5 that have not yet been described. For example, residual posterior MEX-5 accumulates on the posterior centrosome and in P-granules, RNA-rich bodies that segregate with and determine the germline lineage (Cuenca et al., 2003; Schubert et al., 2000). The functional ramifications of this localization are not known.

MEX-5, MEX-6, POS-1, PIE-1, and MEX-1 are all CCCH-type tandem zinc finger proteins (Figure 3.1B, hereafter TZF). This class is typified by tristetraprolin (TTP), a mammalian protein involved in regulating inflammation response by destabilizing TNF-α transcripts (Carballo et al., 1998; Lai et al., 1999; Taylor et al., 1996; Varnum et al., 1991). The expression pattern of several key maternal transcripts is perturbed in TZF mutants leading to the hypothesis that they directly regulate maternal mRNA stability or translation efficiency (D'Agostino et al., 2006; Ogura et al., 2003; Reese et al., 2000; Schubert et al., 2000; Tabara et al., 1999). If so, then the network of maternal RNA regulation in the embryo may be governed by differences in the RNA-binding specificity of each protein. Consistent with this model, TTP is an exquisitely specific RNA-binding protein; it recognizes nonameric UUAUUUAUU sequences present in the 3'-untranslated region (UTR) of its targets (Brewer et al., 2004; Carballo et al., 1998; Lai et al., 1999; Lai et al., 2006). An NMR structure of the mammalian TTP homolog ERF-2 (also known as Tis11D) demonstrates that each zinc finger individually recognizes a UAUU repeat (Hudson et al., 2004). In contrast, an interaction between any of the TZF proteins from *C. elegans* and RNA has not been demonstrated, and as such their mRNA target specificity has not been explored.

MEX-5 and MEX-6 diverge from TTP in a few notable ways (Figure 3.1B): (i) nine amino acids rather than eight separate the first two cysteines in each zinc finger, (ii) the spacing between fingers is lengthened, and (iii) several highly

conserved amino acids that contribute to RNA-binding in mammalian TZF proteins are not conserved in MEX-5 and MEX-6. These differences could impact the ability of MEX-5 and MEX-6 to bind to RNA. Moreover, MEX-5 has been shown to interact with ZIF-1 protein in a yeast two-hybrid assay (DeRenzo et al., 2003), suggesting that it may not regulate this factor at the RNA level. We set out to describe the RNA-binding properties of MEX-5 in order to probe its role in patterning the anterior/posterior axis.


a
Zygote
MEX-5
MEX-6
P0
POS-1
PIE-1
MEX-1
2-Cell
AB
P1
4-Cell
ABa
ABp
P2
EMS
DIC
GFP-MEX-5
ABp
ABa
P2
EMS
a
c
p

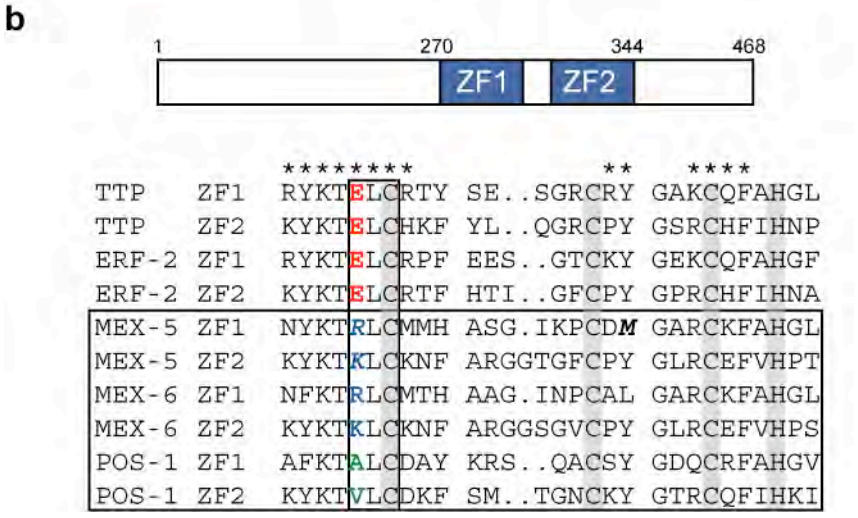

b
1
270
344
468
ZF1
ZF2
TTP ZF1 RYKTELCRTY SE..SGRCRY GAKCQFAHGL
TTP ZF2 KYKTELCHKF YL..QGRCPY GSRCHFIHNP
ERF-2 ZF1 RYKTELCRPF EES..GTCKY GEKCQFAHGF
ERF-2 ZF2 KYKTELCRTF HTI..GFCPY GPRCHFIHNA
MEX-5 ZF1 NYKTRLCMMH ASG.IKPCDM GARCKFAHGL
MEX-5 ZF2 KYKTKLCKNF ARGGTGFCPY GLRCEFVHPT
MEX-6 ZF1 NFKTRLCMTH AAG.INPCAL GARCKFAHGL
MEX-6 ZF2 KYKTKLCKNF ARGGSGVCPY GLRCEFVHPS
POS-1 ZF1 AFKTALCDAY KRS..QACSY GDQCRFAHGV
POS-1 ZF2 KYKTVLCDKF SM..TGNCKY GTRCQFIHKI

Figure 3.1

Figure 3.1: *C. elegans* tandem zinc finger proteins. **A.** Reciprocal gradients of the TZF proteins MEX-5, MEX-6, POS-1, PIE-1, and MEX-1 from the 1–4 cell stages of development. MEX-5 and MEX-6 are represented in red, POS-1, PIE-1, and MEX-1 are represented in blue. DIC and fluorescence images of GFP-MEX-5 in a live 4-cell embryo are shown. The pattern was observed in live embryos by Seydoux and co-workers with GFP reporters (Cuenca et al., 2003) and by Priess and colleagues using immunofluorescence in fixed embryos (Schubert et al., 2000). The protein accumulates predominantly in the cytoplasm of the anterior blastomeres (a, ABa and ABp), but is also in P-granules (p) and on the centrosome of P_2 (c). The cells are labeled in the DIC image. **B.** Domain structure of MEX-5. The location of the two CCCH zinc fingers (ZF1 and ZF2) is shown in blue. The numbers represent the primary amino acid sequence. Alignment of *C. elegans* zinc finger domains MEX-5, MEX-6, and POS-1 with mammalian TTP and ERF-2. The *C. elegans* proteins are boxed. Gray bars denote the three cysteines and one histidine that coordinate the zinc ion. The asterisks represent points of contact with RNA observed in the NMR structure of ERF-2 (Hudson et al., 2004). The adenosine recognition pocket is boxed. The discriminator position is colored red for acidic side chains, blue for basic side chains, and green for hydrophobic side chains. Amino acids in italics were mutated in the present work.

RESULTS

MEX-5 and MEX-6 bind to ARE repeat elements

To clarify the role of MEX-5 and MEX-6 in development, we set out to test whether these proteins, like TTP, bind with high affinity and specificity to RNA. A recombinant fragment of each protein comprising the TZF domain was purified from bacteria and tested for the ability to interact with an established TTP binding sequence – the AU-rich element (ARE) of TNF-α mRNA (Brewer et al., 2004). Two approaches were employed to measure affinity to this RNA: electrophoretic mobility shift (EMSA) and fluorescence polarization (FP) assays. Both methods reveal that MEX-5 and MEX-6 bind to TNF-ARE RNA. The EMSA experiments show that multiple binding sites are present in this sequence. The apparent dissociation constant ($K_{d,\ app}$) of MEX-5 for TNF-ARE RNA is 17 ± 1 nM by EMSA and 14 ± 4 nM by FP (Figure 3.2A, B, Table 1). Similarly, MEX-6 binds to this RNA with an affinity of 4 ± 3 nM by EMSA and 12 ± 3 nM by FP (Figure 3.2A). Both proteins are capable of binding to RNA with high affinity. Furthermore, because the EMSA and FP results are nearly equivalent, it is clear that both assays can effectively monitor RNA binding by these proteins.

TNF-ARE RNA contains several UUAUUUAUU repeat sequences and therefore can bind multiple molecules of TTP (Brewer et al., 2004). To determine the affinity of MEX-5 for a shorter RNA variant containing only one TTP site, we repeated the binding analyses with ARE13 RNA (AUUUAUUUAUUUA). The apparent dissociation constant for this sequence is 55 ± 15 nM by EMSA and 97

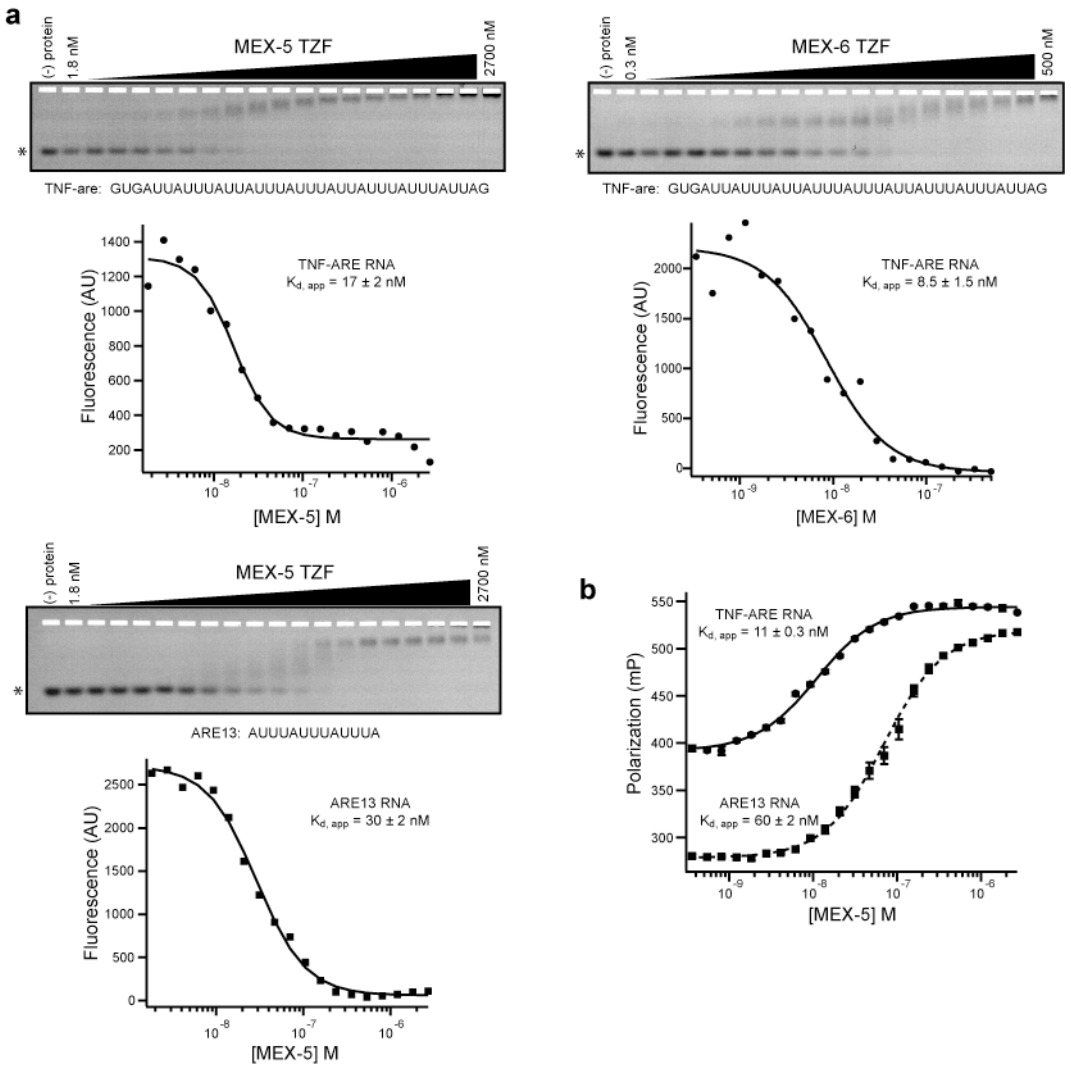

a
(-) protein
1.8 nM
MEX-5 TZF
2700 nM
TNF-are: GUGAUUAUUUAUUAUUUAUUUAUUAUUUAUUUAUUAG
Fluorescence (AU)
TNF-ARE RNA
$K_{d, app}$ = 17 ± 2 nM
[MEX-5] M
(-) protein
0.3 nM
MEX-6 TZF
500 nM
TNF-are: GUGAUUAUUUAUUAUUUAUUUAUUAUUUAUUUAUUAG
TNF-ARE RNA
$K_{d, app}$ = 8.5 ± 1.5 nM
[MEX-6] M
(-) protein
1.8 nM
MEX-5 TZF
2700 nM
ARE13: AUUUAUUUAUUUA
ARE13 RNA
$K_{d, app}$ = 30 ± 2 nM
[MEX-5] M
b
Polarization (mP)
TNF-ARE RNA
$K_{d, app}$ = 11 ± 0.3 nM
ARE13 RNA
$K_{d, app}$ = 60 ± 2 nM
[MEX-5] M

Figure 3.2

Figure 3.2: MEX-5 is an RNA-binding protein. **A.** MEX-5 binds to the ARE element of TNF-a mRNA 3'-UTR and a shorter variant termed ARE13 by electrophoretic mobility shift. The interaction of MEX-6 with TNF-ARE RNA is shown for comparison. The unshifted RNA is denoted by an asterisk for each gel. The sequences of the RNA are shown. A fit of free RNA fluorescence intensity as a function of MEX-5 concentration is presented below each gel. The $K_{d, app}$ and fit error is given for the specific gel above. **B.** Fluorescence polarization analysis of MEX-5 binding to TNF-a ARE and ARE13 RNA. Raw polarization values are presented. A fit of the data is shown, and the $K_{d, app}$ and error are as in panel A.

± 4 nM by FP (Figure 3.2). In contrast, TTP binds to ARE13 RNA with 5-10 fold tighter affinity (Brewer et al., 2004). Together, the results show that MEX-5 binds to UUAUUUAUU repeat sequences, but with an overall affinity that is weaker than TTP.

MEX-5 binds to regulatory elements in glp-1and nos-2 3'-UTRs

Prior work demonstrates that several factors are aberrantly expressed in *mex-5* mutants (Schubert et al., 2000). Ectopic expression of five proteins (SKN-1, PIE-1, MEX-1, POS-1, and PAL-1) and reduced levels of two others (GLP-1 and MEX-3) result from *mex-5* mutation. Furthermore, MEX-5 is required to activate *zif-1*, which in turn targets TZF proteins for degradation (DeRenzo et al., 2003). Lastly, recent studies reveal that MEX-5 is required for anterior degradation of *mex-1, nos-2*, and *pos-1* transcripts (D'Agostino et al., 2006; Tenlen et al., 2006). Although the results are consistent with MEX-5 regulating a network of maternal genes, this regulation has not been shown to be a direct result of MEX-5 binding to target mRNAs.

Extended UAUU sequence repeats similar to the TNF-ARE are not present in the 3'-UTR of any candidate MEX-5 regulatory target. However, several functional regulatory elements have been identified in the 3'-UTR of *glp-1* and *nos-2* mRNAs (Figure 3.3A, B). Two translational control elements are present in the 3'-UTR of *glp-1* mRNA, a spatial control region (SCR) and a temporal control region (TCR) (Evans et al., 1994). Furthermore, five elements

(subA-E) found in the 3'-UTR of *nos-2* mRNA coordinate translational silencing, mRNA localization, and 3'-end formation (D'Agostino et al., 2006). To determine if MEX-5 can bind these functional elements, we performed EMSA and FP experiments with overlapping thirty nucleotide fragments of the TCR and all five elements from the *nos-2* 3'-UTR (Figure 3.3, Table 3.1). Surprisingly, MEX-5 binds with high affinity to all of the TCR fragments and three of the elements from *nos-2* mRNA (TCR1-4, subA, subC, subE, $K_{d, app}$ ~ 25 – 100 nM). MEX-5 binds moderately to *nos-2* subD ($K_{d, app}$ = 200 ± 20 nM) and poorly to *nos-2* subB ($K_{d, app}$ 400 ± 50 nM by FP, > 1 µM by EMSA).

Because a single shifted species is observed with TCR2 RNA, we decided to investigate the stoichiometry of the complex by repeating the FP experiments with elevated TCR2 RNA concentration (Figure 3.3C, D). The apparent stoichiometry is approximately one to one (equivalence point N is 0.9 ± 1 by a quadratic fit), demonstrating that there is only one binding site in this RNA and that the recombinant protein is nearly 100% active (Figure 3.3D). Consistent with these results, analytical gel filtration chromatography reveals that the TZF domain is predominantly monomeric at concentrations well above the apparent dissociation constant for this RNA sequence (17 µM, $K_{d, app}$ = 31 ± 9 nM, Figure 3.3D).

Inspection of the RNA fragments reveals that all of the interacting sequences contain a tract of six to eight uridines within an eight-nucleotide window. This feature is absent in *nos-2* subB, suggesting that MEX-5 requires

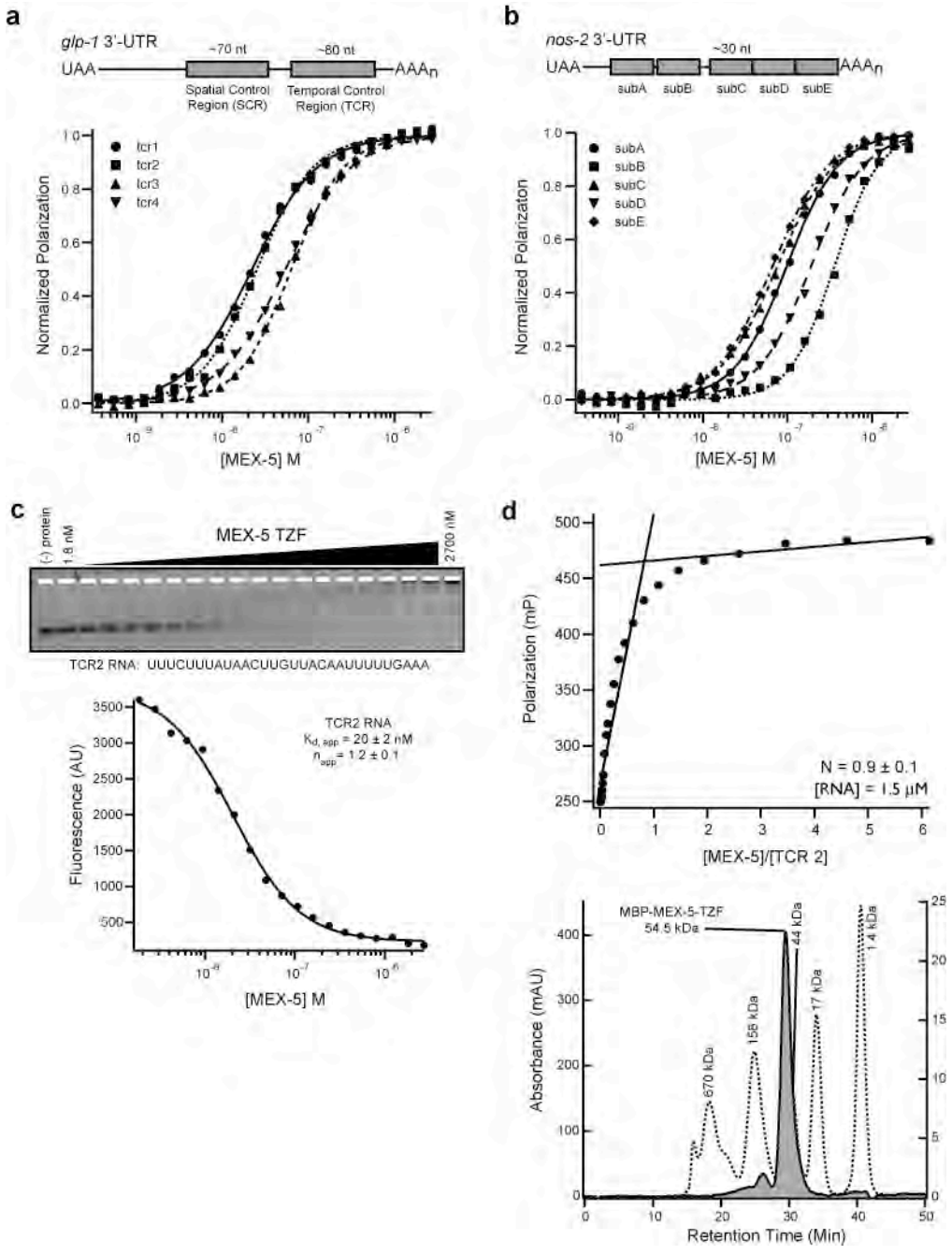

a
glp-1 3'-UTR
~70 nt
~80 nt
UAA
AAAn
Spatial Control Region (SCR)
Temporal Control Region (TCR)
tcr1
tcr2
tcr3
tcr4
Normalized Polarization
[MEX-5] M
b
nos-2 3'-UTR
~30 nt
subA
subB
subC
subD
subE
c
(-) protein
1.8 nM
MEX-5 TZF
2700 nM
TCR2 RNA: UUUCUUUAUAACUUGUUACAAUUUUUGAAA
TCR2 RNA
Kd, app = 20 ± 2 nM
napp = 1.2 ± 0.1
Fluorescence (AU)
d
Polarization (mP)
N = 0.9 ± 0.1
[RNA] = 1.5 µM
[MEX-5]/[TCR 2]
MBP-MEX-5-TZF
54.5 kDa
670 kDa
155 kDa
44 kDa
17 kDa
1.4 kDa
Absorbance (mAU)
Retention Time (Min)

Figure 3.3

Figure 3.3: MEX-5 can interact with regulatory elements in the 3'-UTR of *glp-1* and *nos-2* transcripts. **A.** Regulatory elements in *glp-1* transcripts. Plots of normalized polarization as a function of MEX-5 concentration are shown. Uncorrected mP values ranged between 263 and 530. **B.** Regulatory elements in the 3'-UTR of *nos-2* mRNA. Curves are as in panel A. The mP values ranged between 274 and 530 **C.** EMSA of MEX-5 binding to TCR2 RNA. The fit shown below is as per figure 2. The $K_{d,\ app}$ and n_{app} are given for the experiment shown. Errors represent the uncertainty of the fit. **D.** Stoichiometric binding of MEX-5 to TCR2 RNA. The total RNA concentration is shown. The lines represent linear fits to the pre- and post- saturation data. The stoichiometry is approximated from the intersection of the two lines, and from a fit of the data to the quadratic equation. The equivalence point N is given from the quadratic fit. The error represents the uncertainty of the fit. **D.** MEX-5 is predominantly monomeric at a concentration well above the $K_{d,\ app}$ for TCR2 RNA (17 µM). The calculated molecular weight of MBP-MEX-5 (gray) and five standards (dashed line) are shown.

this element in order to bind. To test this model, a mutant version of the second TCR fragment where all of the uridines are replaced by adenosine was prepared. As expected, this sequence does not bind to MEX-5 (Table 3.1). The results show that MEX-5 does not require UAUU repeats in order to bind to RNA with high affinity, and suggest that MEX-5 can bind to sequences harboring an extended uridine tract.

MEX-5 binds to polyuridine

TTP displays an 80-fold preference for UAUU repeat sequences over polyuridine (Brewer et al., 2004). In contrast, our data show that MEX-5 binds with high affinity to uridine-rich RNAs lacking canonical ARE motifs. To test whether uridine nucleotides are sufficient to promote MEX-5 binding, EMSA and FP experiments were performed with a thirty-nucleotide polyuridine sequence. MEX-5 binds to this RNA with an apparent K_d of 29 ± 6 nM by EMSA and 23 ± 2 nM by FP (Figure 3.4A). The data show that MEX-5 can bind to polyuridine with affinity similar to that of TNF-ARE. This demonstrates that MEX-5 binds with less specificity than TTP.

To further probe MEX-5 specificity, EMSA and FP experiments were performed with fifteen-nucleotide polyuridine, polyadenosine, polycytidine, and polyguanosine sequences. As before, MEX-5 binds to polyuridine with slightly weaker affinity than a similar length AUUUA repeat RNA (ARE13), but does not bind to polycytidine, polyadenosine, or polyguanosine (Figure 3.4A, Table 3.1).

Analysis of MEX-5 TZF binding to RNA

RNA ID	Sequence	Gene/3′-UTR position	EMSA $K_{d,app}$	FP $K_{d,app}$
			n_M	
TNF-ARE	GUGAUUAUUUAUUAUUUAUUUAUUAUUUAUUUAUUUAG	TNF-α, 459–470	17 ± 1	14 ± 4
ARE13	AUUUAUUUAUUUA	NA[a]	55 ± 15	97 ± 4
TCR1	UUGUUUUUUCUCCUUUUCUUUAUAACUUGU	*glp-1, 242–272*	29 ± 3	25 ± 7
TCR2	UUUCUUUAUAACUUGUUACAAUUUUUGAAA	*glp-1*, 256–285	37 ± 18	31 ± 9
TCR3	UUACAAUUUUUGAAAUUCCCUUUUUUGACA	*glp-1*, 270–300	66 ± 22	52 ± 12
TCR4	UUCCCUUUUUUGACAGGCUUUUAUUACACUGUAA	*glp-1*, 286–319	46 ± 15	62 ± 11
SubA	CAAUACUUUUUUAUAUCGGGUCCAACCCGUUUA	*nos-2*, 15–47	67 ± 12	94 ± 2
SubB	UACAAGCUUUCACAAACAGAUAGUUUAU	*nos-2*, 53–80	>1000	400 ± 50
SubC	CCCGUUCAUAGCCUUUAUUGAUUCCAAAUUU	*nos-2*, 88–118	68 ± 11	78 ± 13
SubD	CCCAUCUCACACUUUUCUACGGUAU	*nos-2*, 119–143	400 ± 60	200 ± 20
SubE	ACCAUUUACUUUUUCUGCUAAUAAUCAAUUAUUAAUA	*nos-2*, 144–180	62 ± 12	70 ± 12
TCR2 U-A	AAACAAAAACAAGAAACAAAAAAAGAAA	NA	>1000	>1000
U30	UUUUUUUUUUUUUUUUUUUUUUUUUUUUUU	NA	29 ± 6	23 ± 2
U15	UUUUUUUUUUUUUUU	NA	150 ± 4	140 ± 4
C15	CCCCCCCCCCCCCCC	NA	>1000	>1000
A15	GGGGGGGGGGGGGGG	NA	>1000	>1000
G15	AAAAAAAAAAAAAAA	NA	>1000	>1000

[a] NA, not applicable.

Table 3.1

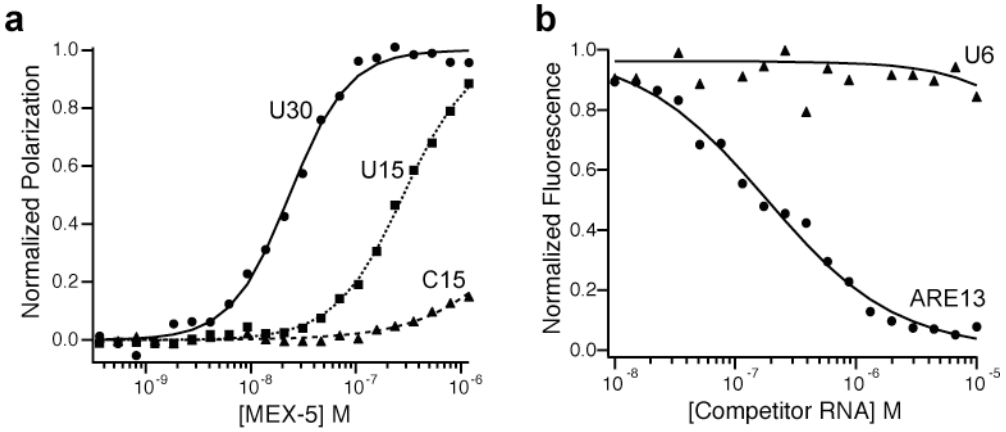

a
Normalized Polarization
1.0
0.8
0.6
0.4
0.2
0.0
U30
U15
C15
10-9
10-8
10-7
10-6
[MEX-5] M
b
Normalized Fluorescence
1.0
0.8
0.6
0.4
0.2
0.0
U6
ARE13
10-8
10-7
10-6
10-5
[Competitor RNA] M

Figure 3.4

Figure 3.4: MEX-5 binds to polyuridine sequences. **A.** Interaction of MEX-5 with polyuridine 30, polyuridine 15, and polycytidine 15 sequences by fluorescence polarization. The curves are as in figure 3. **B.** Competition experiments of unlabeled ARE13 RNA and U6 RNA into the complex of MEX-5 with labeled ARE13 RNA.

Together, the results show that uridine tracts are both necessary and sufficient to promote MEX-5 binding.

In order to identify the shortest RNA fragment that can bind to MEX-5, competition EMSA and FP experiments were performed with even shorter polyuridine sequences (U6 and U9). Competition experiments were favored with shorter RNA sequences to prevent the possibility that steric hindrance by the 3'-fluorescein label would influence the apparent affinity. Disruption of the complex between MEX-5 and labeled ARE13 RNA was used to probe the affinity of unlabeled competitor sequences. The apparent dissociation constant of the competitor ($K_{c, app}$) was determined by a fit of the data to the Lin and Riggs equation (Lin and Riggs, 1972; Weeks and Crothers, 1992). By self-competition, $K_{c, app}$ for ARE13 is equivalent within error to the $K_{d, app}$ determined by direct titration (Figure 3.4B). In contrast, U6 and U9 are unable to compete for MEX-5 binding ($K_{c, app} > 1$ μM). Together, the results demonstrate that the minimal MEX-5 binding site is greater than nine nucleotides in length.

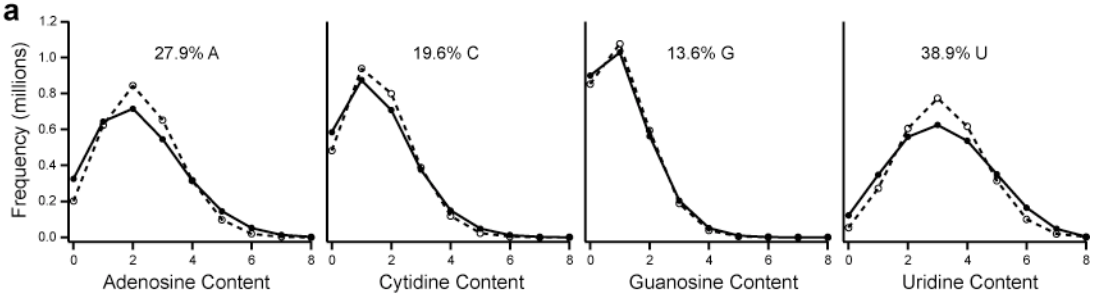
Uridine tracts are remarkably abundant in C. elegans 3'-UTR sequences

It is possible that MEX-5 specifically regulates only those maternal genes that contain uridine tracts in their 3'-UTRs. To identify potential MEX-5 targets, we searched every annotated 3'-UTR in Wormbase release WS165 for octamer sequences containing at least six uridines using the PATSCAN algorithm (Dsouza et al., 1997). Amazingly, 91% of 3'-UTRs harbor at least one potential

MEX-5 binding site. This demonstrates that *C. elegans* 3'-UTRs are remarkably rich in uridine tracts, and implies that MEX-5 alone cannot specify mRNA targets for regulation.

To further explore this hypothesis, we repeated the PATSCAN search with every possible octamer sequence (4^8 sequences) in the *C. elegans* 3'-UTR database. There are 11,938 unique 3'-UTRs annotated in Wormbase. These UTRs are rich in uridine and adenosine but are relatively poor in guanosine and cytidine. To determine if tracts of uridine are overrepresented, we determined the theoretical distribution of octamer units weighted by the relative base content in 3'-UTRs using a binomial distribution. If a given base has a propensity to segregate into tracts of high and low base content, the frequency of octamers containing six to eight occurrences of the base will be greater than expected from the random distribution. Similarly, the frequency of octamers containing zero to three base occurrences will also be greater than the expected amount. Therefore, a plot of octamer frequency versus the number of base occurrences will appear more broad and shallow than the theoretical random distribution (Figure 3.5A). In *C. elegans* 3'-UTRs, the observed distribution of uridine and to a lesser extent adenosine reveals a propensity to segregate into high and low base content tracts, while the observed distributions of guanosine and cytidine very closely match the random distribution (Figure 3.5A). This is consistent with a selective pressure that favors runs of uridine and adenosine. When the

a



b

Octamer ID	Freq.	Octamer ID	Freq.	Octamer ID	Freq.
UUUUUUUU	3686	AAUUUUUU	1966	UGUUUUUU	1717
UUUUUUUC	2904	UAUUUUUU	1960	UUGUUUUU	1696
AUUUUUUU	2869	GUUUUUUU	1932	UUUGUUUU	1628
UUUUUUCU	2521	UUUUAUUU	1919	UUUUGUUU	1605
UUUUUUUG	2346	CUUUUUUU	1915	AUUUAUUU	1579
UUUUUUGU	2330	UCUUUUUU	1897	UCAUUUUU	1471
UUUAUUUU	2142	UUUUCUUU	1893	UAAUUUUU	1469
UUAUUUUU	2106	UUCUUUUU	1788	UUUUUAAA	1450
UUUUUCUU	2091	UUUUUAUU	1778	UUUUUUAA	1448
UUUUUGUU	2065	UUUCUUUU	1718	UUUUUUUA	1448

Figure 3.5

Figure 3.5: Analysis of the distribution of base content in octamer windows of all annotated *C. elegans* 3'-UTRs. **A.** A plot of the relative frequency of octamer sequences with zero through eight occurrences of a given base (dark line) compared the expected frequency based on a random distribution weighted by the nucleotide distribution (dashed line). The fraction of each base present in *C. elegans* 3'-UTR space is shown above each chart. **B.** The identity of the top 30 most frequent octamer elements present in *C. elegans* 3'-UTRs and the frequency of occurrence in Wormbase release WS165 is shown.

distribution of observed octamer sequences is plotted as a function of relative frequency, we find that the most frequently observed octamer is U8, and 29 of the top 30 octamers contain at least six uridines (Figure 3.5B). This highlights the preponderance of uridine tracts in the 3'-UTRs of *C. elegans* genes, and demonstrates that MEX-5 does not bind with sufficient specificity to select mRNAs for regulation on its own.

MEX-5 and PIE-1 localization do not depend on MEX-3

If MEX-5 regulates specific maternal transcripts, it must do so as part of a complex with a more specific RNA-binding protein. If so, a likely candidate is MEX-3. MEX-3 is a KH domain RNA-binding protein that is also required for anterior development (Draper et al., 1996). Mutation of *mex-3* results in the same terminal phenotype as *mex-5*, and the protein displays a similar distribution throughout the 1-4 cell stages of development. Moreover, MEX-6, which is 70% identical to MEX-5, was identified in a yeast two-hybrid screen for MEX-3 interacting proteins (Huang et al., 2002). Finally, the localization pattern of MEX-3 protein is perturbed in *mex-5* mutants (Schubert et al., 2000).

Because the localization pattern of MEX-3 depends on MEX-5 (Schubert et al., 2000), it is possible that the two proteins localize in a complex. If so, then anterior MEX-5 accumulation might depend upon the presence of MEX-3. In addition, because MEX-5 activation of *zif-1* drives posterior accumulation of PIE-1 (DeRenzo et al., 2003), localization of this protein should be altered in the

absence of MEX-3 if it serves as a *zif-1* co-activator. To test these hypotheses, we knocked down MEX-3 levels using RNAi in worms expressing GFP-MEX-5 or GFP-PIE-1 by feeding them *Escherichia coli* expressing double stranded RNA targeted against *mex-3* transcripts. The cellular distribution of GFP-MEX-5 and GFP-PIE-1 in embryos was determined by wide field fluorescence microscopy. No difference is observed in the expression pattern of GFP-MEX-5 or GFP-PIE-1 in *mex-3* RNAi embryos (n = 135, n = 127, respectively; only embryos in the 1-8 cell stage were counted) compared to control embryos (n = 101, n = 111, Figure 3.6). More than 90% of the embryos laid onto *mex-3* RNAi plates failed to hatch (n > 500), indicating that RNAi was disrupting *mex-3* function. We conclude that MEX-3 is not required for MEX-5 accumulation in the anterior or for activation of *zif-1*. The results argue against a functional role for a MEX-3/MEX-5 complex. However, we cannot rule out the possibility that they co-regulate other maternal transcripts.

A single residue in each finger defines MEX-5 RNA-binding specificity

The NMR structure of the mammalian TTP homolog ERF-2 reveals that each zinc finger recognizes adjacent UAUU repeats (Hudson et al., 2004). In order to understand the molecular basis for differential MEX-5 RNA-binding specificity, we prepared a homology model of MEX-5 based on the ERF-2 structure using SWISS-MODEL (Schwede et al., 2003) (Figure 3.7A). The model reveals several amino acid differences in the RNA-binding interface that may

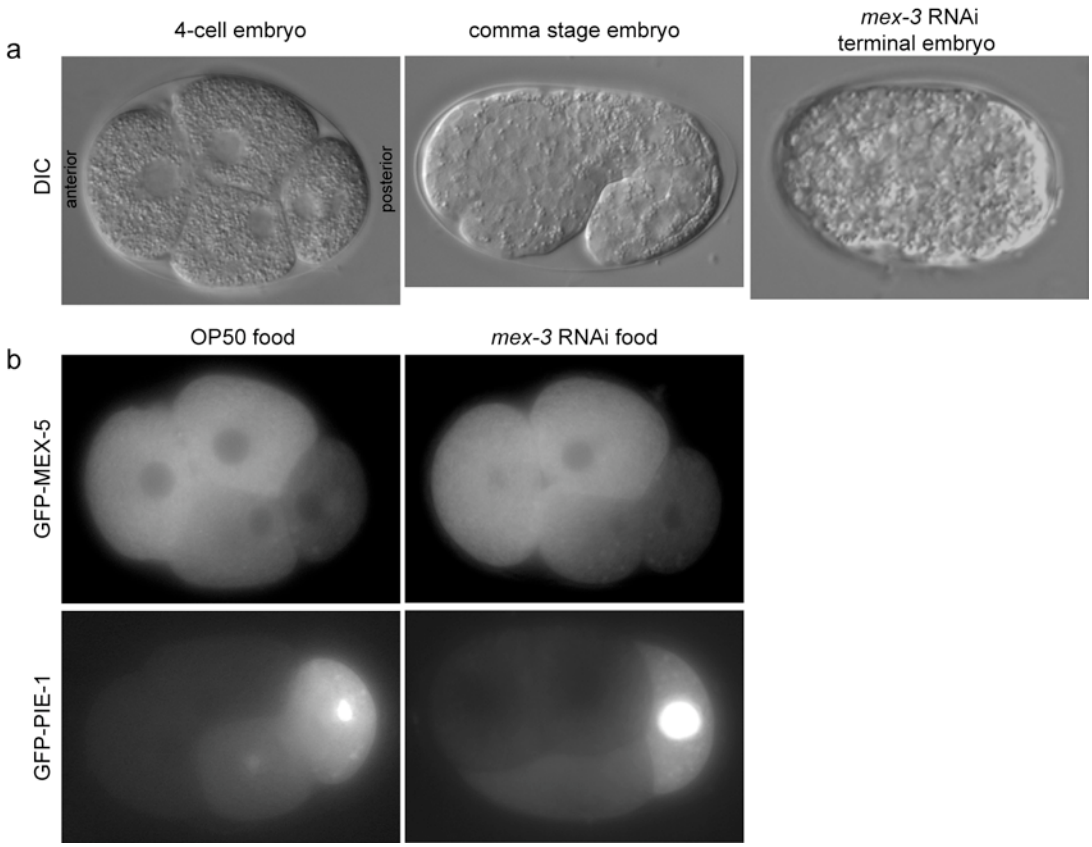

a
4-cell embryo
comma stage embryo
mex-3 RNAi
terminal embryo
DIC
anterior
posterior
b
OP50 food
mex-3 RNAi food
GFP-MEX-5
GFP-PIE-1

Figure 3.6

Figure 3.6: MEX-5-GFP and PIE-1-GFP expression in *mex-3* RNAi embryos. **A.** DIC images of a 4-cell and a comma stage embryo compared to the terminal *mex-3* RNAi phenotype. The anterior and posterior pole are labeled in the 4-cell embryo. **B.** Comparison of the 4-cell stage localization pattern of GFP-MEX-5 or GFP-PIE-1 in untreated or *mex-3* RNAi embryos.

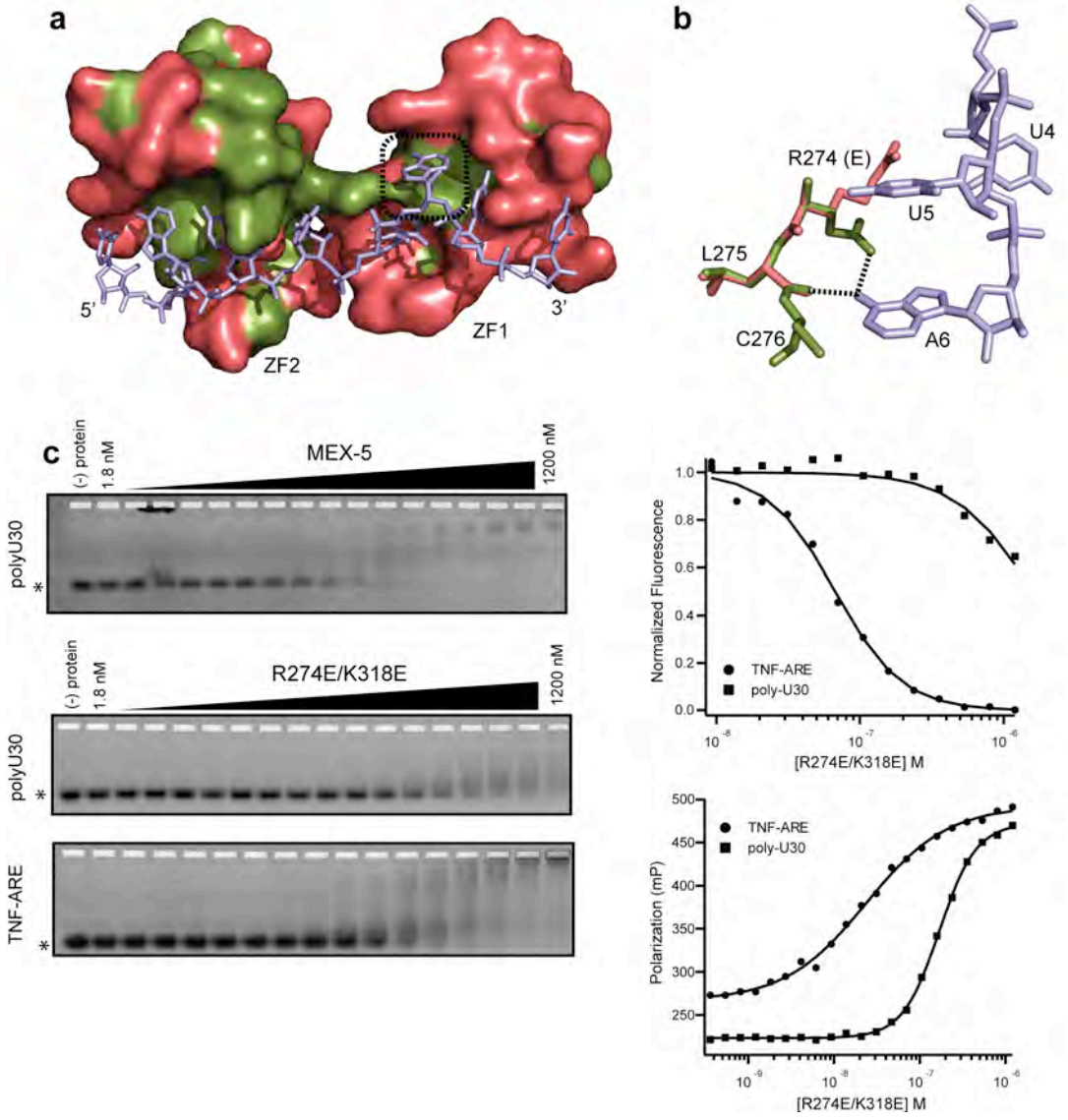

a
5'
3'
ZF1
ZF2
b
U4
R274 (E)
U5
L275
C276
A6
c
(-) protein
1.8 nM
MEX-5
1200 nM
polyU30
*
(-) protein
1.8 nM
R274E/K318E
1200 nM
polyU30
*
TNF-ARE
*
Normalized Fluorescence
TNF-ARE
poly-U30
[R274E/K318E] M
Polarization (mP)
TNF-ARE
poly-U30
[R274E/K318E] M

Figure 3.7

Figure 3.7: A discriminator amino acid in each finger defines RNA-binding specificity. **A.** A homology model of MEX-5 bound to RNA based upon the NMR structure of ERF-2 (Hudson et al., 2004). Amino acids that are conserved between the two are colored green. Amino acids that are different are in red. The RNA (5'-UUAUUUAUU-3') is denoted in blue. The box denotes the adenosine recognition pocket. **B.** Recognition of adenosine by MEX-5 (red) and ERF-2 (green). The numbering corresponds to the sequence in MEX-5. E denotes the position of the discriminator residue, which is an arginine in MEX-5 zinc finger 1 and a glutamate in ERF-2 zinc finger 1. **C.** Electrophoretic mobility shift of wild-type MEX-5 with polyuridine 30 RNA and the R274E/K318E mutant with polyuridine 30 and TNF-ARE RNA. Unshifted RNA is denoted by an asterisk. A fit of the normalized free fluorescence intensity as a function of R274E/K318E concentration is shown for TNF-ARE and polyU30 RNA. Fluorescence polarization data are shown for R274E/K318E mutant with TNF-ARE and polyU30 RNA.

contribute to the difference in specificity. Most notably, three adjacent residues (E, L, C) in each finger of ERF-2 combine to form an adenosine recognition pocket. In MEX-5, the glutamate residue is replaced with arginine (R274) in the first finger and lysine (K318) in the second (Figure 3.1C). In the NMR structure, the glutamate side chain forms a hydrogen bond with the exocyclic amine of the adenosine base. In the homology model, the basic residues rotate away from the adenosine and form backbone contacts with adjacent nucleotides (Figure 3.7B). The model predicts that loss of base-specific hydrogen bonds and formation of additional backbone contacts contributes to the relaxed specificity of MEX-5. If so, mutation of the basic residues to glutamate might confer TTP-like specificity to this protein.

To test this hypothesis, we prepared mutant variants of MEX-5 where either or both of these basic residues were replaced with the glutamate residue present in TTP (R274E, K318E, and R274E/K318E). A substitution at an unrelated position in the RNA interface (M288Y) was prepared as a control. First, we examined the ability of the mutants to interact with the TNF-ARE sequence. The R274E and the K318E mutations do not affect the interaction, while the double mutation binds with four-fold reduced affinity to this sequence (Table 3.2). A small loss in affinity is expected due to increased electrostatic repulsion from the glutamate residues. The M288Y mutation has no effect on TNF-ARE affinity.

In contrast, the effect of glutamate mutations on binding to the thirty-nucleotide polyuridine RNA is significantly more drastic. The individual glutamate mutations reduce binding by two to four fold, while the double mutation significantly reduces binding (>15 fold by EMSA, 7.5 fold by FP, Figure 3.7C, Table 3.2). As before, the M288Y mutation has no effect. Because all variants retain the ability to bind to TNF-ARE RNA, and because all are soluble and monomeric as determined by analytical size exclusion chromatography, we conclude that reduction in polyuridine binding results from a change in binding specificity rather than from protein misfolding effects. The data demonstrate that a single glutamate in each finger confers the ability to discriminate between UAUU repeat RNA and polyuridine, thus defining the molecular basis for the difference in RNA discrimination between MEX-5 and TTP/ERF-2.

Comparison of mutant MEX-5 RNA-binding specificity

	Polyuridine 30 RNA		***TNF-a ARE RNA***		
Protein	**$K_{d, app}$ (nM)**	**Relative Affinity**	**$K_{d, app}$ (nM)**	**Relative Affinity**	**Fold Discrimination**
MEX-5	29 ± 6	1	18 ± 1	1	1.6
M288Y	28 ± 0.5	1	12 ± 7	0.7	2.3
R274E	28 ± 4	1	16 ± 2	0.9	1.8
K318E	110 ± 30	3.9	28 ± 1.5	1.6	4.0
R274E/K318E	>1000	>35	70 ± 7	4	>15
MEX-5 (FP)	23 ± 2	1	14 ± 4	1	1.5
R274E/K318E (FP)	170 ± 7	7.5	23 ± 3	1.6	7.5

Table 3.2

Discussion

Accurate post-transcriptional regulation of gene expression requires the ability to distinguish specific transcripts from the total pool of cellular RNA. Regulatory factors target mRNA through sequence-specific or structure-specific interactions (Auweter et al., 2006). In the case of microRNAs, the mechanistic basis for mRNA recognition is easy to conceptualize; complimentary base pairing of a seven or eight nucleotide seed drives mRNA discrimination (Brennecke et al., 2005). In contrast, the sequence code recognized by eukaryotic RNA-binding proteins usually includes a greater degree of degeneracy.

In *C. elegans*, early embryogenesis requires a class of divergent CCCH-type tandem zinc finger proteins. They segregate to opposite poles of the 1-cell embryo in response to PAR protein activity (Cuenca et al., 2003; DeRenzo et al., 2003; Mello et al., 1996; Schubert et al., 2000). Because they are homologous to mammalian TTP, most are thought to regulate maternal gene expression at the RNA level. Our data show that two of the *C. elegans* TZF proteins, MEX-5 and MEX-6, can in fact bind to RNA with high affinity but with relaxed specificity compared to TTP. The results are consistent with the hypothesis that they directly regulate maternal transcripts, but also suggest that accessory factors are required for mRNA targeting.

Implications for RNA recognition by other CCCH-type TZF proteins

The difference in specificity between MEX-5 and TTP is governed by a single amino acid substitution within a highly conserved region of each finger. Our data indicate that the amino acid identity of this discriminating position defines a code that predicts the specificity of other TZF proteins. We predict that glutamate residues encode selectivity for UUAUUUAUU RNA, while basic residues lead to promiscuous binding to uridine-rich RNA sequences including, but not limited to, UAUU-repeat sequences, as seen in TNF-ARE RNA.

To determine the variability of the discriminator residue in TZF proteins, we performed a BLAST search of the *C. elegans* genome using the TZF domain of MEX-5 as a query (Figure 3.8). Sixteen hits result from this analysis. Variability is observed at the discriminator position, but the other two residues that comprise the adenosine recognition pocket are conserved. Of the sixteen, only MEX-5 and MEX-6 contain basic amino acids in both fingers at the discriminator position. One protein, F38B7.1, has a glutamate at both positions similar to TTP and ERF-2. We predict that this protein binds specifically to UAUU repeat RNAs. Interestingly, five proteins (POS-1, Y116A8C.19, Y116A8C.20, Y57G11C.25, and F38C2.5) have small hydrophobic residues in both discriminator positions, an alanine in finger one and a valine in finger two.

Gene Identity	ZF1 A-pocket	ZF2 A-pocket	
TTP	**E**LC	**E**LC	Acidic
ERF-2	**E**LC	**E**LC	
F38B7.1	**E**LC	**E**AC	
MEX-5	**R**LC	**K**LC	Basic
MEX-6	**R**LC	**K**LC	
POS-1	**A**LC	**V**LC	Small hydrophobic
Y116A8C.19	**A**LC	**V**LC	
Y116A8C.20	**A**LC	**V**LC	
Y57G11C.25	**A**LC	**V**LC	
F38C2.5	**A**LC	**V**LC	
F38C2.7	**S**LC	**V**LC	Mixed
C35D6.4	**S**LC	**V**LC	
Y116A8C.17	**S**LC	**V**LC	
OMA-1	**V**IC	**K**LC	
OMA-2	**V**IC	**K**LC	
MOE-3	**V**IC	**K**LC	
MEX-1	**A**LC	**Q**LC	
PIE-1	**R**LC	**Q**IC	

Figure 3.8

Figure 3.8: Identity of the discriminator amino acid (bold) in the adenosine recognition pocket of each finger for all *C. elegans* TZF proteins (boxed) is presented. The acidic, basic, hydrophobic, and mixed TZF protein classes are labeled.

Experiments with recombinant POS-1 protein indicate that it binds to RNA with specificity that is different from both MEX-5 and TTP (Farley et al., 2008). Therefore, small hydrophobic residues at the discriminator position may define a third specificity class for TZF proteins.

The remaining TZF proteins have a combination of basic, small hydrophobic, or acidic/polar amino acids at the discriminator position. Three proteins required for oocyte maturation (OMA-1, OMA-2, and MOE-3) have a valine in the first finger and a lysine in the second. We predict these proteins will display hybrid specificity between POS-1 and MEX-5. In contrast, PIE-1 contains an arginine in finger one and a glutamine in finger two. Like glutamate, glutamine can accept a hydrogen bond from adenosine, but it can also donate a hydrogen bond to the O6 carbonyl of a guanosine base as well. F38C2.7, C35D6.4, and MEX-1 have the combination of a small hydrophobic residue in one finger and a polar amino acid (serine or asparagine) that could theoretically accept and/or donate a base-specific hydrogen bond in the RNA complex. The variations of the discriminator residue imply differences in RNA-binding specificity by the divergent *C. elegans* TZF proteins. Further experiments will determine the possible RNA interactions allowed by each discriminator amino acid within this class of proteins.

Remarkably, the variations in the TZF protein discriminator residue found in *C. elegans* are absent in higher eukaryotes. Blast analysis reveals that every vertebrate CCCH-TZF protein in GenBank has a glutamate residue at both

discriminator positions. In contrast, homologs with variable discriminators are present in *Caenorhabditis briggsae* and *Caenorhabditis remanei*, and in more divergent Nemata including parasitic species. Since the identity of the discriminator correlates with its RNA-binding specificity and thus its molecular function, the critical role of divergent TZF proteins during early development must be a special adaptation of this phylum.

The role of the PAR proteins in establishing cell polarity is conserved from worms to flies to mammals (Pellettieri and Seydoux, 2002). Asymmetric expression of signaling proteins and transcription factors from maternal mRNAs is also highly conserved (Seydoux and Braun, 2006). Yet the cassette of RNA-binding proteins that connect these two layers is clearly highly divergent (Johnstone and Lasko, 2001). Defining the basis of RNA-binding protein function in early development will provide a framework by which mechanistic differences in the regulation of maternal mRNAs contribute to variability of metazoan body plan.

Materials and Methods

Protein expression constructs

Fragments of *mex-5* and *mex-6* containing the TZF domain (amino acids 236-350 and 250-400, respectively) were amplified from ORFeome clones (Open Biosystems) and sub-cloned into the vector pMal-c (NEB). Mutations of pMal-MEX-5(236-350) were prepared by site-directed mutagenesis using quickchange (Stratagene).

Purification of recombinant proteins

TZF domains from MEX-5, MEX-6, and mutants thereof were expressed and purified from *E. coli* JM109 cells as C-terminal fusions to maltose binding protein. Liquid cultures grown at 37 deg. C were induced in mid log phase with 0.1 mM IPTG. Zinc acetate was added to a final concentration of 100 µM at the time of induction. Harvested cells were resuspended in lysis buffer (50 mM Tris pH 8.8, 200 mM NaCl, 2 mM DTT, EDTA free protease inhibitor tablet (Roche), 100 µM $Zn(OAc)_2$) and lysed by sonication. Soluble protein was purified over an amylose column (NEB). Fractions containing the fusion protein were pooled and dialyzed into Q buffer (50 mM Tris, pH 8.8, 20 mM NaCl, 2 mM DTT, and 100 µM $Zn(OAc)_2$) and then further purified over a Hi-trap Q HP column (GE Healthcare). Final purification was achieved by combining fractions containing the protein and dialyzing them into S buffer (50 mM MOPS, pH 6.0, 20 mM NaCl, 2 mM DTT, 100 µM $Zn(OAc)_2$) before running them over a Hi-trap S HP column (GE

Healthcare). Care was taken to minimize the amount of time the protein was exposed to pH 6.0 buffer. Pure fractions as determined by coomassie-stained SDS-PAGE were combined and dialyzed into storage buffer (20 mM Tris, pH 8.0, 20 mM NaCl, 100 μM $Zn(OAc)_2$, 2 mM DTT). After dialysis, the protein concentration was determined using Beer's law by measuring absorbance at 280 nm and a calculated extinction coefficient determined using the ProtParam server (Gasteiger et al., 2003). The protein was concentrated to approximately 50 μM before storage at 4 deg. C.

Analytical Gel Filtration Chromatography

A Superdex 200 10/300 GL column (10mm x 300 mm, GE Healthcare) was used to determine the apparent molecular weight of MBP-MEX-5 (236-350) and its mutants. The column was equilibrated on an AKTA FPLC with two column volumes of filtration buffer (50 mM Tris, pH 8.0, 300 mM NaCl) at a flow rate of 0.5 ml min^{-1} prior to loading the protein sample. Approximately 50 μl of sample (17 μM) was loaded on to the column and eluted with 1.5 column volumes of filtration buffer. Retention time was determined in relation to standards (Bio-Rad).

RNA labeling protocol

All of the RNA sequences used in this work were prepared by chemical synthesis and deprotected/lyophylized as the manufacturer directed (Dharmacon

or Integrated DNA technologies). Lyophilized samples were resuspended in 300 µl of TE buffer, pH 8.0, and the concentration was measured by determining the absorbance at 260 nm using a calculated extinction coefficient based on the nucleotide content.

Fluorescein 5-thiosemicarbazide (FTSC, Invitrogen) was used to 3'-end label each RNA via the method of Reines and Cantor (Reines and Cantor, 1974). A typical 50 µl reaction consisted of 0.5 nanomoles RNA, 100 mM NaOAc, pH 5.1, and 5 nmoles $NaIO_4$. After a ninety minute incubation at room temperature, the sample was ethanol precipitated with 1 µl RNase free glycogen (Invitrogen 20 µg/µl), 5 M NaCl (1/20 the volume), and 2 volumes of 100% ice-cold ethanol. The resulting pellet was resuspended in 50 µl of 100 mM NaOAc, pH 5.1 containing 1 mM FTSC. This reaction was incubated overnight at 4°C and unreacted label was removed using a Roche G-25 spin column. The labeling efficiency was determined by calculating the ratio of fluorescein absorbance at 490 nm to RNA-fluorescein absorbance at 260 nm. Typical efficiencies were 60–80%.

Electrophoretic mobility shift assays

Electrophoretic mobility shift assays were used to measure the binding activity of recombinant MEX-5 and MEX-6 to fluorescein-labeled RNA oligonucleotides. Typical reactions consisted of 2–4 nM labeled RNA equilibrated with varying concentrations of protein in equilibration buffer for 3

hours. Equilibration buffer is 0.01% IGEPAL CA630 (a mild detergent used to prevent adhesion or protein and RNA to tubes, microplates, and gel wells), 0.01 mg/ml tRNA (a polyanionic non-specific binding inhibitor), 10 mM Tris, pH 8.0, 100 uM $Zn(OAc)_2$, and 100 mM NaCl. The RNA was heated to 60 deg. C and allowed to cool to room temperature before use. Immediately prior to loading, one-fifth volume of 30% v/v glycerol, 0.01% w/v bromocresol green was added to each reaction as a dye marker. A 40 µl sample of each reaction (100 µl total) was loaded onto a 1% agarose gel (EMD Biosciences, some lot to lot variability was observed) in 1X TB buffer. The gels were run for 40 minutes at 120 volts then immediately scanned using a fluor-imager (Fujifilm FLA-5000) with a blue laser at 473 nm. The fluorescence intensity of unbound RNA was determined as a function of protein concentration using ImageGuage software. The data were fit to a sigmoidal dose response function (equation 1) in order to determine the half maximal saturation point ($K_{d,\,app}$):

1) $$\phi = b + (m - b)\frac{1}{1 + \left(K_{d,app} / P\right)^{n}}$$

where m is the maximal signal, b is the minimal signal, P is the protein concentration, and n is the apparent Hill coefficient. It is important to note that $K_{d,\,app}$ is not equivalent to the thermodynamic equilibrium dissociation constant for RNA sequences that contain multiple overlapping binding sites. In all cases, the

reported value is the average of at least three experiments and the reported error is the standard deviation.

Competition assays were performed as above except a constant concentration of sub-saturating recombinant MEX-5 was used in the equilibration while varying concentrations of unlabeled competitor RNA were added to the reaction. The apparent dissociation constant of the competitor RNA was determined by a fit of the data to a quadratic solution of the Lin and Riggs equation (Lin and Riggs, 1972; Weeks and Crothers, 1992) as described (Ryder et al., 2004).

Fluorescence Polarization assays

Equilibration reactions (100 µL volume) were set up using the same conditions as the electrophoretic mobility shift experiments above in 96-well black plates (Greiner). The apparent fluorescence polarization was determined using a Victor 3 plate reader (Perkin Elmer) equipped with fluorescein sensitive filters and polarizers. A total of five reads were measured for each experiment and the average and standard deviation of the millipolarization value (mP) were calculated for each protein concentration. The data were fit to equation 1 in order to extract the apparent dissociation constant. The reported value is the average of at least three experiments and the error is the standard deviation.

Stoichiometric binding experiments were performed as above except the reactions were supplemented with unlabeled RNA to a final concentration of 1.5

μM. The elevated concentration of RNA enables determination of the apparent stoichiometry by measuring the equivalence saturation point. This value was estimated by plotting polarization as a function of molar equivalents of protein to RNA and performing linear fits to pre- and post-saturation data. The equivalence point was determined by the intersection point of the two lines, and separately by a fit of the data to the quadratic equation as described (Rambo and Doudna, 2004).

UTR sequence analysis

C. elegans 3'-UTRs were retrieved from Wormbase release WS165. To determine the frequency of each possible octamer sequence in 3'-UTR-space, a Ruby script was written to enumerate each possible octamer and send it to the pattern searching tool PATSCAN (Dsouza et al., 1997). The PATSCAN output files were analyzed using standard UNIX text processing tools. The total number of occurrences of each class of A, C, G, or U-containing octamers was determined by summing up the number of occurrences of each octamer in that class. The theoretical distribution of each class was determined by a binomial distribution weighted by the fractional proportion of each base in all *C. elegans* 3'-UTRs.

RNAi

C. elegans strains expressing GFP-MEX-5 (JH1448) or GFP-PIE-1 (JH1327) were obtained from the *Caenorhabidits* Genetics Center and cultured by propagating animals with the roller phenotype. Embryos were harvested from young adult hermaphrodites grown on OP50 food by bleach treatment and then deposited on NGM plates seeded with OP50 or NGM-IPTG plates seeded with *mex-3* RNAi food (generously provided by Dr. Craig Mello) and cultured as described (Brenner, 1974). Embryos were collected from gravid adult hermaphrodites by dissecting the worms in M9 on a 4% agarose pad with a fine gauge needle. DIC and GFP images were collected with live specimens using a Zeiss Axioskop microscope with 40X or 100X objectives.

CHAPTER IV

RNA recognition by the embryonic cell-fate determinant and germline totipotency factor MEX-3

Abstract

Totipotent stem cells have the potential to differentiate into every cell type. Renewal of totipotent stem cells in the germline, and cellular differentiation during early embryogenesis, rely upon post-transcriptional regulatory mechanisms. The *Caenorhabditis elegans* RNA binding protein MEX-3 plays a key role in both processes. MEX-3 is a maternally-supplied factor that controls the RNA metabolism of transcripts encoding critical cell fate determinants. However, the nucleotide sequence specificity and requirements of MEX-3 mRNA recognition remain unclear. Only a few candidate regulatory targets have been identified, and the full extent of the network of MEX-3 targets is not known. Here, we define the consensus sequence required for MEX-3 RNA recognition, and demonstrate that this element is required for MEX-3 dependent regulation of gene expression in live worms. Based on this work, we identify several candidate MEX-3 targets that help explain its dual role in regulating germline stem cell totipotency and embryonic cell fate specification.

Introduction

Totipotent stem cells have the capacity to differentiate into an entire organism. Multicellular organisms maintain a population of totipotent stem cells in the germline in order to assure reproductive potential (Seydoux and Braun, 2006). These stem cells give rise to the gametes that eventually produce the next generation. The gametes that arise from these cells encode all of the information necessary to pattern the development of a new organism. This information is carried in the form of DNA, epigenetic marks, and cytoplasmic components. All of this information must be established and maintained in totipotent germline stem cells.

After fertilization, totipotency is lost in most cellular lineages as tissues and organs begin to differentiate. In many organisms, cell-fate specification occurs early during embryogenesis at a time when the organism's genome is transcriptionally quiescent (Farley and Ryder, 2008; Seydoux and Braun, 2006). Thus, post-transcriptional regulation of maternally and paternally-supplied gene products provides the basis for the loss of totitpotency and cell fate specification.

In the nematode *Caenorhabditis elegans*, the conserved dual KH (hnRNP K homology) domain protein MEX-3 is required for both maintaining totipotency in the germline and cell fate specification in the early embryo (Figure 4.1A) (Ciosk et al., 2006; Draper et al., 1996). Worms with a homozygous null *mex-3* mutation have a fully penetrant maternal-effect lethal phenotype, resulting in embryos that fail to undergo body morphogenesis and produce excess muscle

and hypodermal cells (Draper et al., 1996). Furthermore, these embryos produce between three and six cells that resemble germline progenitor cells, as opposed to two in a wild-type embryo. In addition, worms that are mutant for *mex-3* and a second KH domain protein, *gld-1,* are sterile with a germline tumor that contains numerous trans-differentiated cells forming a cell mass similar to a germline teratoma (Ciosk et al., 2006).

In the germline, MEX-3 protein is expressed in distal germ cells and maturing ooctyes (Draper et al., 1996). Upon fertilization, cytoplasmic MEX-3 is present throughout the entire embryo and then becomes restricted predominantly to the anterior founder cell (AB) at the end of the two-cell stage (Figure 4.1B-D). After the four-cell stage, MEX-3 begins to disappear from the embryo. A small amount of MEX-3 remains in the posterior germline lineage, where it localizes to RNA rich bodies including germ granules in early embryos and CAR-1/CGH-1 granules in late oocytes (Draper et al., 1996; Gallo et al., 2008; Jud et al., 2007; Jud et al., 2008).

A family of related human proteins (hMex-3A–D and TINO) are differentially recruited to RNA granules involved in post-transcriptional regulatory mechanisms (Buchet-Poyau et al., 2007; Donnini et al., 2004). In addition to a region of MEX-3 homology, these factors contain a carboxy-terminal RING finger domain and numerous phosphorylation sites. Qualitative experiments show that hMex-3A, -3B, and -3C bind directly to RNA homopolymers while hMex-3D/TINO binds to a sequence that contains AU-rich elements (Buchet-Poyau et al., 2007;

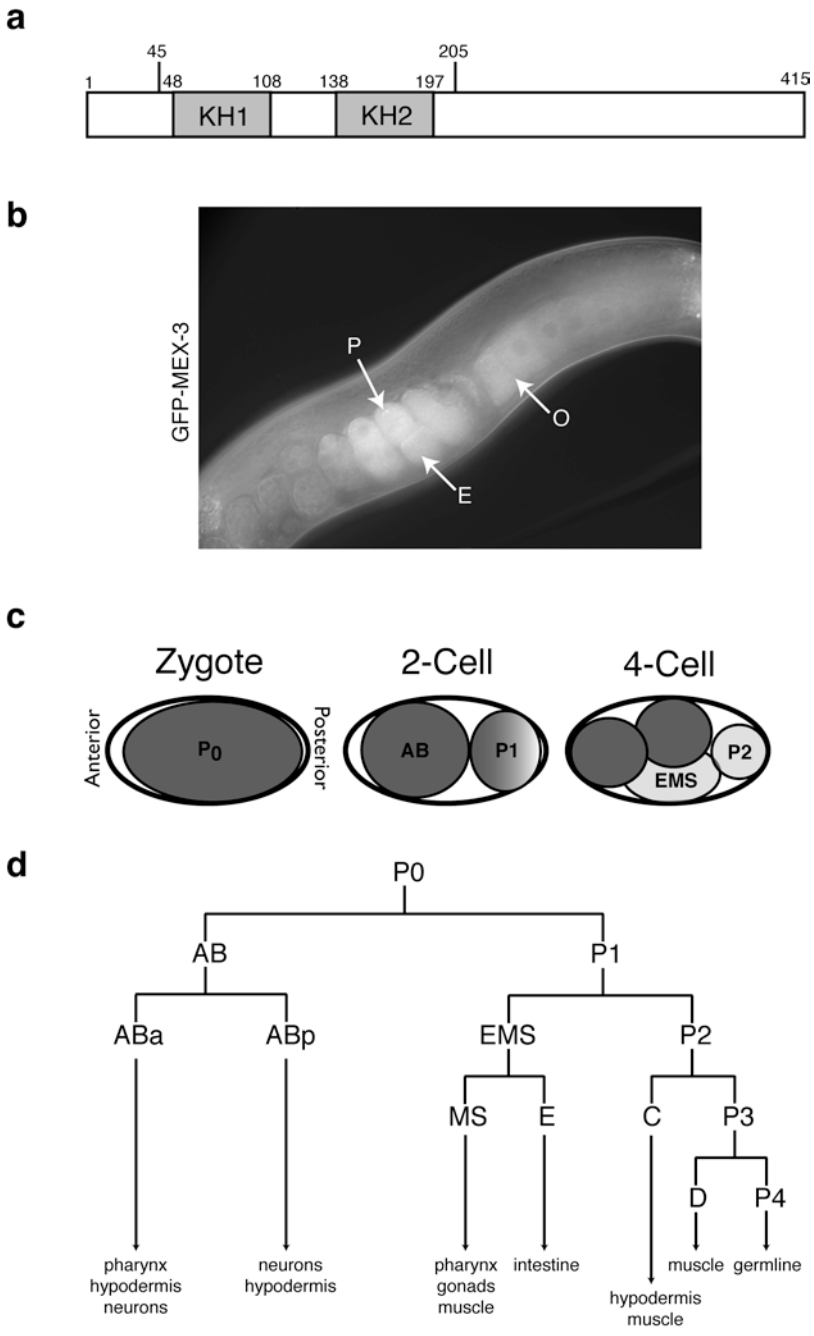

a
45
205
1
48
108
138
197
415
KH1
KH2
b
GFP-MEX-3
P
O
E
c
Zygote
2-Cell
4-Cell
Anterior
Posterior
P0
AB
P1
P2
EMS
d
P0
AB
P1
ABa
ABp
EMS
P2
MS
E
C
P3
D
P4
pharynx
hypodermis
neurons
neurons
hypodermis
pharynx
gonads
muscle
intestine
hypodermis
muscle
muscle
germline

Figure 4.1

Figure 4.1: MEX-3 in *C. elegans* development. **A.** Domain structure of the RNA binding protein MEX-3. The two KH domains (KH1 and KH2) are shaded dark grey. The numbers represent the primary amino acid sequence. **B.** A transgenic worm expressing GFP::MEX-3 (GFP and Nomarski overlay). This strain recapitulates most of the expression pattern of endogenous MEX-3 (Draper et al., 1996), being translated in oocytes [O], early embryos [E], and localized to p-granules [P]. However, expression is not observed in the distal end of the germline, and polarized expression from the 2-4 cell stage is less pronounced than endogenous MEX-3. **C.** A model representing the MEX-3 expression pattern in the 1-4 cell embryo. At the end of the 2-cell stage MEX-3 is restricted predominantly to the anterior blastomere (AB), shown in dark grey. The blastomeres AB, EMS, and P are labeled. **D.** Cellular lineage of early embryogenesis.

Donnini et al., 2004). The precise sequence specificity and affinity of the MEX-3 family has not been defined.

Because MEX-3 contains conserved RNA-binding domains, it likely exerts a role in development at the post-transcriptional level. Consistent with this idea, several lines of evidence indicate that MEX-3 regulates the spatial and temporal translation of two key targets, *pal-1* and *nos-2*. PAL-1 is a Caudal-like homeodomain protein required to specify the posterior blastomere C (Hunter and Kenyon, 1996). Its protein expression pattern is anti-correlated with that of MEX-3 in early embryos, as expected if MEX-3 negatively regulates *pal-1* translation. PAL-1 expression is not observed until the four-cell stage where it accumulates in the nuclei of the posterior blastomeres. *mex-3* mutant embryos show ectopic expression of PAL-1, and the anterior blastomeres take on a C-like fate resulting in excess muscle in the anterior (Draper et al., 1996). The gene encoding the Nanos homolog NOS-2 is also dependent upon MEX-3 for its protein expression pattern (Jadhav et al., 2008). This protein is required for the proper development of primordial germ cells and their incorporation into the somatic gonad. In the early embryo, NOS-2 is not observed until the sixteen-cell stage in the germline precursor P_4. However, in *mex-3* mutant embryos, NOS-2 is expressed ectopically throughout the entire embryo (Jadhav et al., 2008).

Translational reporter experiments suggest that MEX-3 is regulating its targets in a 3'-UTR dependent manner (Hunter and Kenyon, 1996; Jadhav et al., 2008). An RNA reporter containing the 3'-UTR of *pal-1* fused to *lacZ* is

translated in early embryos in a pattern that matches endogenous PAL-1 (Hunter and Kenyon, 1996). In the absence of MEX-3, ectopic expression of the reporter is observed in oocytes and anterior blastomeres. Furthermore, recent work demonstrates that MEX-3 represses *nos-2* mRNA translation through its 3'-UTR (Jadhav et al., 2008). Using a GFP::H2B translational reporter transgene, several *cis*-regulatory elements within the 3'-UTR required for the spatial and temporal control of *nos-2* mRNA translation (subA-E) were identified (D'Agostino et al., 2006). Both *in vitro* and *in vivo* experiments suggest that MEX-3 represses *nos-2* translation by interacting with a repeat sequence present within the regulatory elements subB and subC (Jadhav et al., 2008), yet the precise specificity determinants remain unknown.

Ectopic expression of NOS-2 in the *mex-3* mutant likely explains the presence of extra germline precursor cells in terminal mutant embryos. Likewise, ectopic expression of PAL-1 likely explains the presence of excess muscle. However, the *mex-3* mutant phenotype, and its role in maintaining germline totipotency, suggests that MEX-3 regulates additional mRNA targets. In order to determine the nucleotide binding specificity of MEX-3 and identify new candidate regulatory targets, we set out to define the determinants of MEX-3 binding and map the *cis*-acting elements within *nos-2* and *pal-1* required for MEX-3 regulation.

Results

Identification of a high affinity MEX-3 RNA aptamer

To better characterize how MEX-3 recognizes its mRNA targets we performed an affinity elution-based *in vitro* selection experiment (SELEX) (Figure 4.2A) (Ellington and Szostak, 1990; Loughlin et al., 2009; Tuerk and Gold, 1990). With a starting pool of ssRNA comprised of 30-nucleotides of randomized sequence, RNA target sequences were selected by immobilizing the RNA binding domain of MEX-3 fused to the C-terminus of maltose binding protein (MBP) on amylose resin. To monitor the progress of the selection, an electrophoretic mobility shift assay (EMSA) was performed where fluorescently labeled RNA samples from either pool 0, pool 4, or pool 7 were equilibrated with varying concentrations of MEX-3 and resolved on a native polyacrylamide slab gel (Figure 4.2B). After seven rounds of selection and enrichment, a pool of RNA that binds with high affinity to MEX-3 was identified.

To analyze the sequences present within pool 7, the cDNA was cloned and DNA from individual transformants was sequenced. Out of 69 recovered sequences, 56 segregate into two major classes (Figure 4.3A). The remaining sequences display no obvious similarity (Figure 4.2C). The first group is comprised of 39 highly-related sequences with only 0-3 variable nucleotides between them. The other main group consists of 17 sequences that are highly purine-rich. To test whether MEX-3 binds with high affinity to either of these sequence classes, EMSA experiments were performed with the most abundant

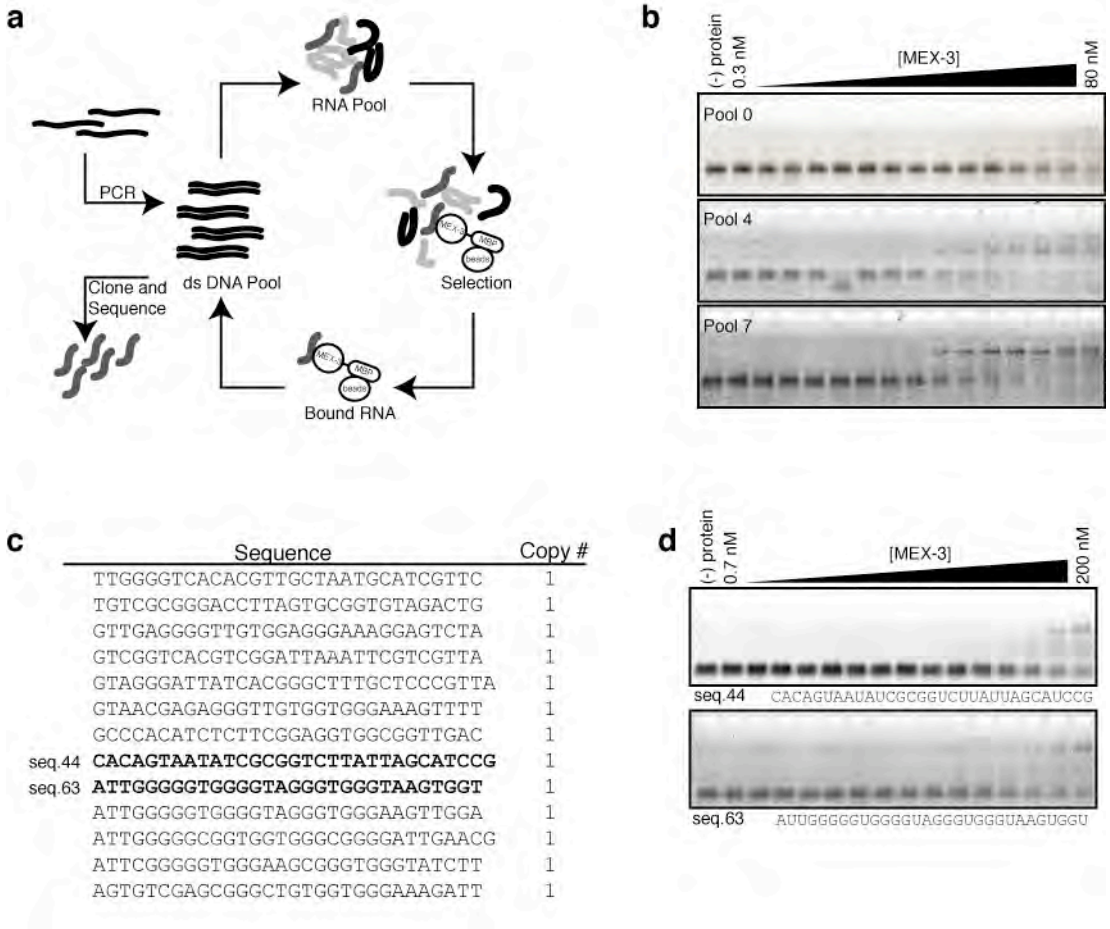

a
RNA Pool
PCR
ds DNA Pool
Clone and Sequence
Selection
Bound RNA
b
(-) protein
0.3 nM
[MEX-3]
80 nM
Pool 0
Pool 4
Pool 7
c
Sequence
Copy #
TTGGGGTCACACGTTGCTAATGCATCGTTC 1
TGTCGCGGGACCTTAGTGCGGTGTAGACTG 1
GTTGAGGGGTTGTGGAGGGAAAGGAGTCTA 1
GTCGGTCACGTCGGATTAAATTCGTCGTTA 1
GTAGGGATTATCACGGGCTTTGCTCCCGTTA 1
GTAACGAGAGGGTTGTGGTGGGAAAGTTTT 1
GCCCACATCTCTTCGGAGGTGGCGGTTGAC 1
seq.44 CACAGTAATATCGCGGTCTTATTAGCATCCG 1
seq.63 ATTGGGGGTGGGGTAGGGTGGGTAAGTGGT 1
ATTGGGGGTGGGGTAGGGTGGGAAGTTGGA 1
ATTGGGGGCGGTGGTGGGCGGGGATTGAACG 1
ATTCGGGGGTGGGAAGCGGGTGGGTATCTT 1
AGTGTCGAGCGGGCTGTGGTGGGAAAGATT 1
d
(-) protein
0.7 nM
[MEX-3]
200 nM
seq.44 CACAGUAAUAUCGCGGUCUUAUUAGCAUCCG
seq.63 AUUGGGGGUGGGGUAGGGUGGGUAAGUGGU

Figure 4.2

Figure 4.2: *In vitro* selection of MEX-3 RNA aptamers. **A.** A schematic of the affinity elution-based in vitro selection experiment. **B.** An electrophoretic mobility shift assay (EMSA) of MEX-3 bound to fluorescently labeled Pool0, Pool4, or Pool7. **C.** Thirteen sequences from the selection that are unrelated to the two main classes of sequences. The sequences that are bold were used in EMSA experiments. **D.** An EMSA experiment with fluorescently labeled seq.44 and seq.63.

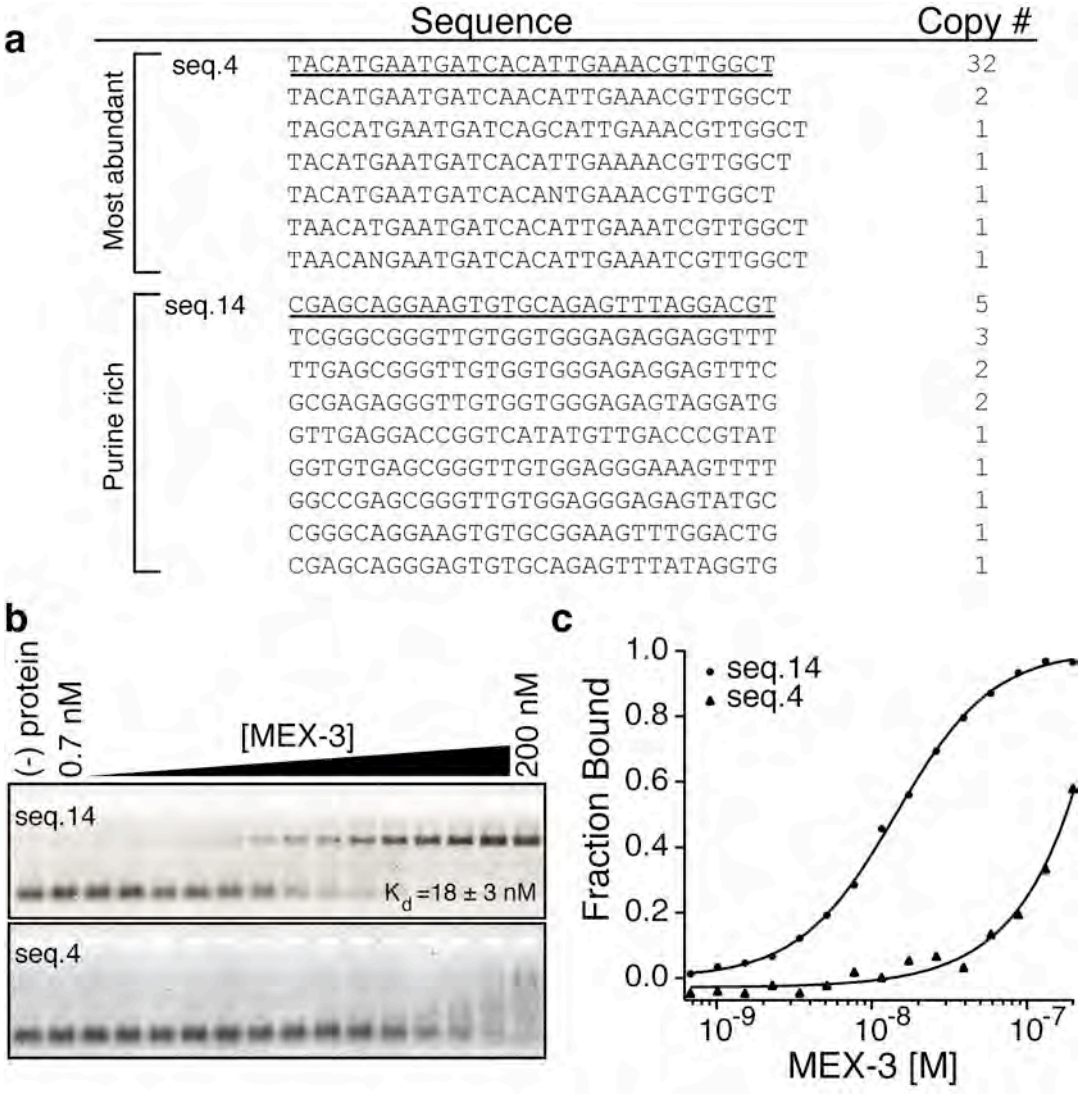

a
Sequence
Copy #
Most abundant
seq.4
TACATGAATGATCACATTGAAACGTTGGCT 32
TACATGAATGATCAACATTGAAACGTTGGCT 2
TAGCATGAATGATCAGCATTGAAACGTTGGCT 1
TACATGAATGATCACATTGAAAACGTTGGCT 1
TACATGAATGATCACANTGAAACGTTGGCT 1
TAACATGAATGATCACATTGAAATCGTTGGCT 1
TAACANGAATGATCACATTGAAATCGTTGGCT 1
Purine rich
seq.14
CGAGCAGGAAGTGTGCAGAGTTTAGGACGT 5
TCGGGCGGGTTGTGGTGGGAGAGGAGGTTT 3
TTGAGCGGGTTGTGGTGGGAGAGGAGTTTC 2
GCGAGAGGGTTGTGGTGGGAGAGTAGGATG 2
GTTGAGGACCGGTCATATGTTGACCCGTAT 1
GGTGTGAGCGGGTTGTGGAGGGAAAGTTTT 1
GGCCGAGCGGGTTGTGGAGGGAGAGTATGC 1
CGGGCAGGAAGTGTGCGGAAGTTTGGACTG 1
CGAGCAGGGAGTGTGCAGAGTTTATAGGTG 1
b
(-) protein
0.7 nM
[MEX-3]
200 nM
seq.14
K_d=18 ± 3 nM
seq.4
c
Fraction Bound
1.0
0.8
0.6
0.4
0.2
0.0
seq.14
seq.4
10^{-9}
10^{-8}
10^{-7}
MEX-3 [M]

Figure 4.3

Figure 4.3: Analysis of MEX-3 RNA aptamers. **A.** The two main sequence classes recovered after seven rounds of selection are shown. The first class is the most abundant with 39 sequences and the second class is purine-rich. Underlined are the examples tested for binding activity. The copy number of each clone is listed next to its sequence. **B.** An EMSA of MEX-3 bound to *seq.14* and *seq.4* RNA. **C.** A graph of MEX-3 versus fraction bound of *seq.14* and *seq.4* RNA is plotted and fit with the Hill equation. The calculated apparent equilibrium dissociation constant ($K_{d,app}$) for *seq.14* is reported as the average ± one standard deviation of at least three replicates.

RNA sequence from each group. Quantitative analysis reveals that MEX-3 binds to the purine-rich example (*seq.14*) with high affinity ($K_{d,app}$= 18 ± 3 nM; Figure 4.1B and C). In contrast, MEX-3 binds with lower affinity to the RNA example (*seq.4*) from the most abundant group ($K_{d,app}$= 160 ± 6 nM). It is not clear why *seq.4* and its variants dominate the *in vitro* selection yield. To test if MEX-3 binds with high affinity to any of the unrelated RNAs, EMSA experiments were performed with two representative sequences. Neither sequence binds to MEX-3 with high affinity (Figure 4.2D). Because *seq.14* binds to MEX-3 with 9-fold higher affinity than *seq.4*, we chose to investigate *seq.14* in more detail.

Identification of a 12-nucleotide element sufficient for MEX-3 binding

In order to determine the minimal sequence required for binding, truncation analysis was performed where either the 5' or 3'-end of *seq.14* was shortened by three nucleotide fragments (Figure 4.4). The binding affinity of MEX-3 to these sequences was determined by EMSA. As many as fifteen bases from the 5'-end ($K_{d,app}$ < 43 nM) or six bases from the 3'-end ($K_{d,app}$ < 21 nM) can be removed without a dramatic reduction in affinity. The results identify the region that is essential for recognition by MEX-3 (Figure 4.4, shaded region). Based on these results, we designed a 12-nucleotide fragment (*seq.14min*) that encompasses this region and tested its ability to interact with MEX-3. Indeed, MEX-3 binds to this sequence with high affinity ($K_{d,app}$ = 36 ± 5 nM; Figure 4.5).

ID	Sequence	$K_{d, app}$ nM
seq.14	CGAGCAGGAAGUGUGCAGAGUUUAGGACGU	18±3
5′frag1	CGAGCAGGAAGUGUGCAGAGUUUAGGA	16±2
5′frag2	CGAGCAGGAAGUGUGCAGAGUUUA	21±6
5′frag3	CGAGCAGGAAGUGUGCAGAGU	190±40
5′frag4	CGAGCAGGAAGUGUGCAG	NB
5′frag5	CGAGCAGGAAGUGUG	NB
5′frag6	CGAGCAGGAAGU	NB
3′frag1	GCAGGAAGUGUGCAGAGUUUAGGACGU	37±8
3′frag2	GGAAGUGUGCAGAGUUUAGGACGU	14±1
3′frag3	AGUGUGCAGAGUUUAGGACGU	28±3
3′frag4	GUGCAGAGUUUAGGACGU	30±4
3′frag5	CAGAGUUUAGGACGU	43±9
3′frag6	AGUUUAGGACGU	NB

Figure 4.4

Figure 4.4: Truncation analysis of *seq.14* RNA. Binding data for 3' and 5' truncations of *seq.14* RNA. The $K_{d,app}$ for each sequence is given. NB stands for no binding observed. The shaded region represents the nucleotides that are critical for MEX-3 binding.

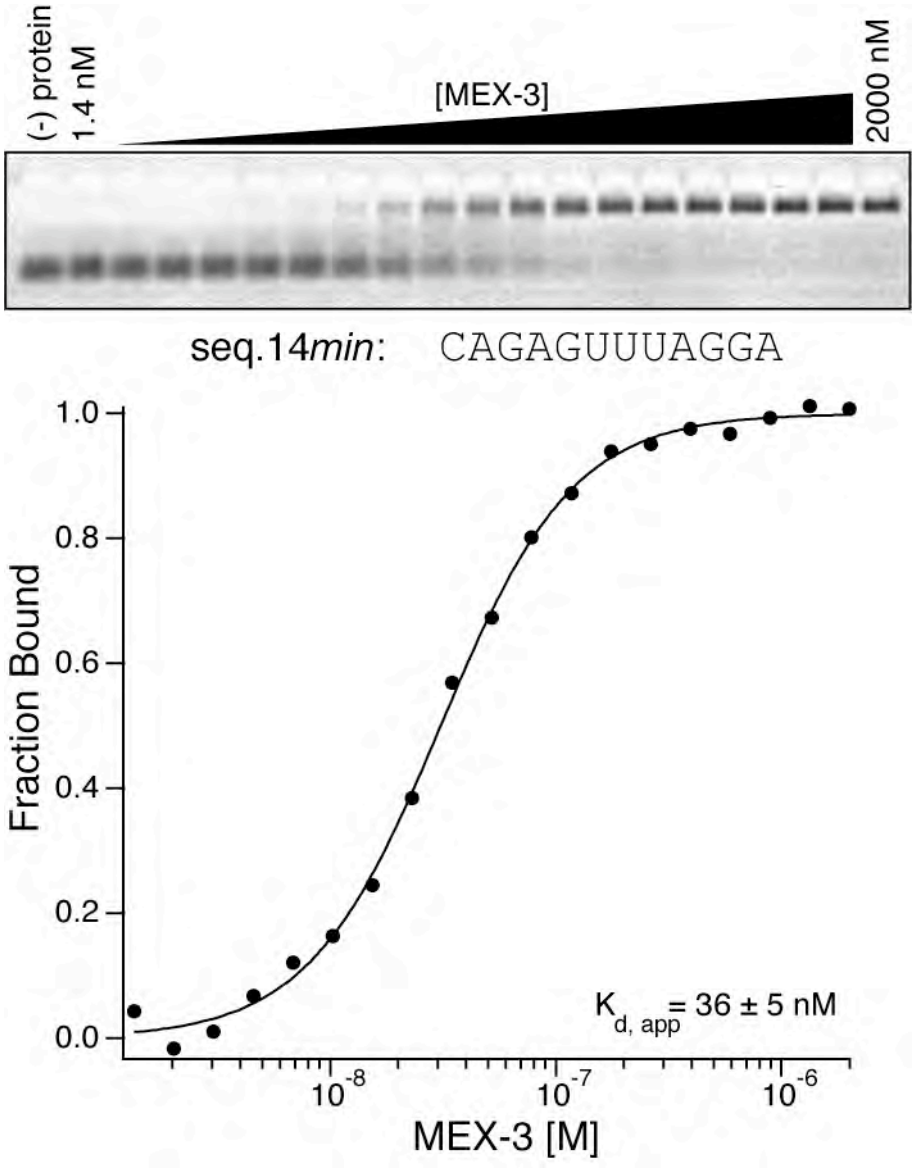

(-) protein
1.4 nM
[MEX-3]
2000 nM
seq.14min: CAGAGUUUAGGA
Fraction Bound
1.0
0.8
0.6
0.4
0.2
0.0
$K_{d, app}$ = 36 ± 5 nM
10^{-8}
10^{-7}
10^{-6}
MEX-3 [M]

Figure 4.5

Figure 4.5: A 12-nucleotide element sufficient for MEX-3 binding. An EMSA experiment showing MEX-3 bound to *seq.14min* RNA with a plot of MEX-3 versus fraction bound of *seq.14min* RNA. The $K_{d,app}$ is reported on the graph.

These experiments reveal the minimal region of the aptamer required for high affinity MEX-3 binding.

Determination of the MEX-3 consensus sequence

To determine the MEX-3 consensus sequence and analyze the thermodynamic contribution of each base, EMSA was employed to measure the change in standard free energy change ($\Delta\Delta G°$) of every single point mutation of *seq.14min* RNA (Figure 4.6A). Mutation of eight positions causes a significant reduction in binding affinity (positions 2-9, $\Delta\Delta G > 0.5$ kcal mol^{-1}), whereas mutation of four positions has little or no effect on binding (positions 1, 10-12, $\Delta\Delta G < 0.5$ kcal mol^{-1}). The cutoff of 0.5 kcal mol^{-1} represents a two-fold reduction in binding affinity at 20 degrees C. Based on this quantitative assay, we define the MEX-3 recognition element (MRE) as (A/G/U)(G/U)AGU(U/A/C)UA. Analysis of this sequence reveals that five positions have a fixed nucleotide specificity, while three positions are partially degenerate allowing one or more base substitution at each position. The first four bases are predominantly purine-rich and the last four bases are AU-rich. MEX-3 contains two KH domains and the typical binding surface of this motif is a four-nucleotide element with variable specificity (Braddock et al., 2002; Gamarnik and Andino, 2000; Lewis et al., 2000; Makeyev and Liebhaber, 2002; Valverde et al., 2008). Therefore, we predict that each KH domain in MEX-3 recognizes a four base half-site within the MRE.

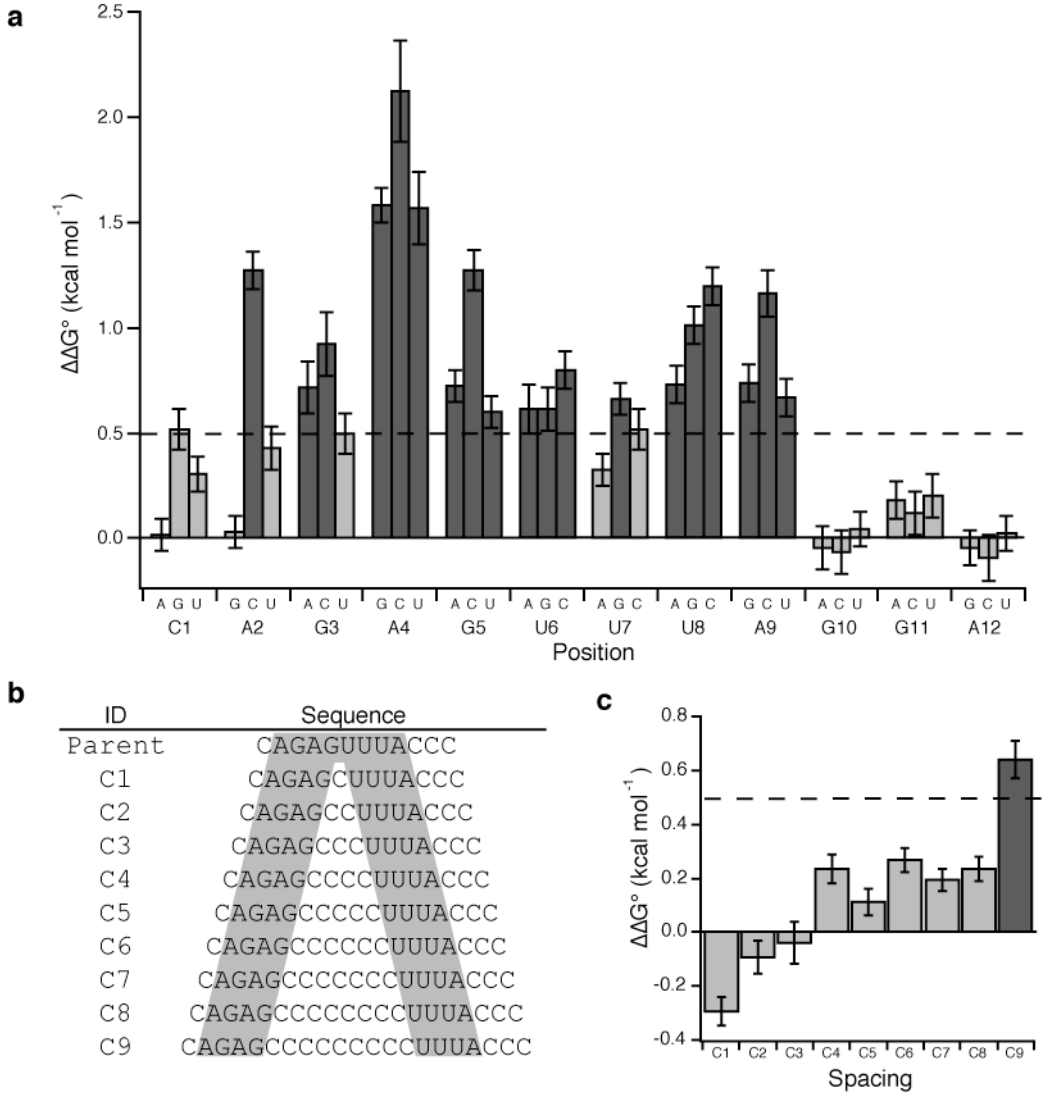

a
2.5
2.0
1.5
1.0
0.5
0.0
ΔΔG° (kcal mol-1)
A G U
C1
G C U
A2
A C U
G3
G C U
A4
A C U
G5
A G C
U6
A G C
U7
A G C
U8
G C U
A9
A C U
G10
A C U
G11
G C U
A12
Position
b
ID
Sequence
Parent CAGAGUUUACCC
C1 CAGAGCUUUACCC
C2 CAGAGCCUUUACCC
C3 CAGAGCCCUUUACCC
C4 CAGAGCCCCUUUACCC
C5 CAGAGCCCCCUUUACCC
C6 CAGAGCCCCCCUUUACCC
C7 CAGAGCCCCCCCUUUACCC
C8 CAGAGCCCCCCCCCUUUACCC
C9 CAGAGCCCCCCCCCCUUUACCC
c
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
ΔΔG° (kcal mol-1)
C1
C2
C3
C4
C5
C6
C7
C8
C9
Spacing

Figure 4.6

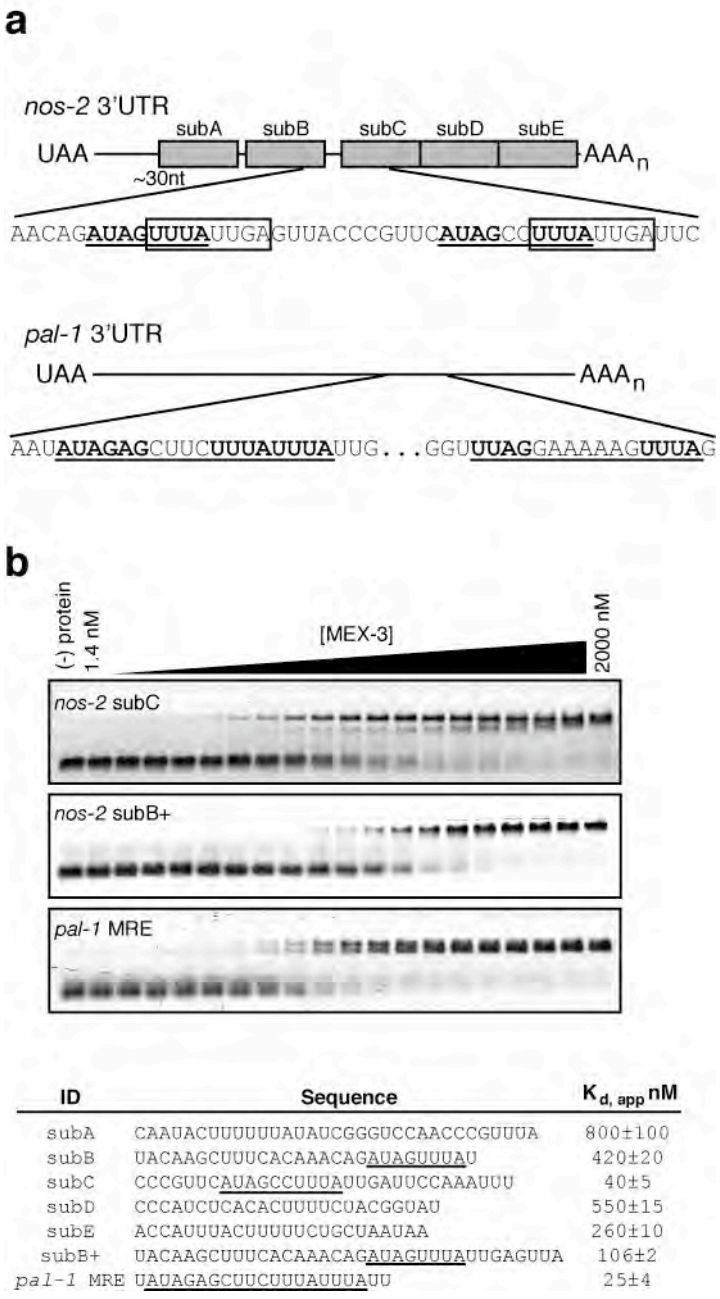
Figure 4.6: Mutagenesis of *seq.14min* RNA. **A.** This sequence was mutated systematically and the $K_{d,app}$ was determined for every single point mutation. This was compared to the $K_{d,app}$ of *seq.14min* to calculate the change in standard free energy change ($\Delta\Delta G°$). The bars represent the $\Delta\Delta G°$ for each point mutation indicated on the *x*-axis. Base substitutions that have greater than a two-fold loss in binding affinity are shown in dark grey (dotted line, $\Delta\Delta G° > 0.5$ kcal mol^{-1}), while mutations that have less than a two-fold loss in binding affinity are in shown in light grey. The error was propagated from the standard deviation of *seq.14min* and the respective point mutation. **B.** The parent sequence containing the eight-nucleotide consensus with a background of cytodine is shown followed by nine spacing mutants (C1-C9). The shaded nucleotides represent each half-site within the MEX-3 consensus sequence. **C.** The $K_{d,app}$ of each spacing mutant was compared to the $K_{d,app}$ of the parent sequence in order to calculate $\Delta\Delta G°$. The spacing mutant that has greater than a two-fold loss in binding affinity is shown in dark grey (dotted line, $\Delta\Delta G° > 0.5$ kcal mol^{-1}).

To test this hypothesis, a series of spacing mutants were made where 0-9 cytidines were inserted between positions G5 and U6 (Figure 4.6B). A background of cytidine was used because this base is not tolerated in the MRE with the exception of one position (U7). These experiments reveal that MEX-3 binds to a bipartite recognition element that can tolerate as many as eight nucleotides in between each specificity determinant (Figure 4.6C). Based on this data, we re-define the MRE is as (A/G/U)(G/U)AGN$_{(0-8)}$U(U/A/C)UA.

MEX-3 binds specifically to MREs present in nos-2 and *pal-1 transcripts*

Next, we asked if the regulatory targets of MEX-3 (*nos-2* and *pal-1*) harbor an MRE within their 3'-UTR. A search of the 3'-UTR of *nos-2* and *pal-1* reveals that both transcripts contain two copies of the MRE (Figure 4.7A). In the case of *nos-2*, the MREs are present within the previously defined regulatory elements of subB and subC (Figure 4.7A). This suggests that the MRE is a functional *cis*-regulatory sequence.

We then asked if MEX-3 binds specifically to the MREs present within the *nos-2* and *pal-1* 3'-UTRs. To address this question, EMSA was performed to determine the binding affinity of MEX-3 to the five known regulatory elements of *nos-2* (subA-E; Figure 4.7B) (D'Agostino et al., 2006) and a 21-nucleotide fragment of the *pal-1* UTR that harbors one MRE. We anticipated that MEX-3 would bind to *nos-2* subB, *nos-2* subC, and the fragment from the *pal-1* UTR, but not to *nos-2* subA, subD, or subE sequences which lack the MRE. As expected



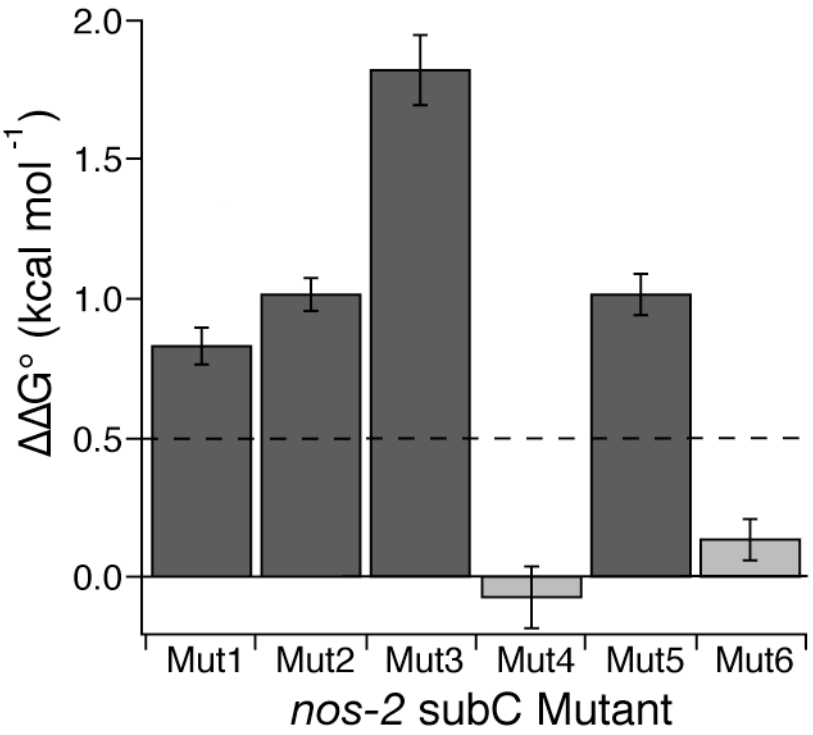
ID	Sequence	$K_{d, app}$ nM
subA	CAAUACUUUUUUAUAUCGGGUCCAACCCGUUUA	800±100
subB	UACAAGCUUUCACAAACAGAUAGUUUAU	420±20
subC	CCCGUUCAUAGCCUUUAUUGAUUCCAAAUUU	40±5
subD	CCCAUCUCACACUUUUCUACGGUAU	550±15
subE	ACCAUUUACUUUUUCUGCUAAUAA	260±10
subB+	UACAAGCUUUCACAAACAGAUAGUUUAUUGAGUUA	106±2
pal-1 MRE	UAUAGAGCUUCUUUAUUUAUU	25±4

Figure 4.7

Figure 4.7: MEX-3 binds specifically to the MRE of *nos-2* and *pal-1*. **A.** Analysis of the *nos-2* 3'-UTR reveals two MRE sites (underlined) within the subB and subC regulatory elements. The half-sites are shown in bold. The conserved repeat element is boxed (D'Agostino et al., 2006). The *pal-1* 3'-UTR is shown with each MRE underlined. **B.** EMSA experiments for *nos-2* subC, *nos-2* subB+, and the *pal-1* MRE are shown. Binding data of MEX-3 with each *nos-2* regulatory element (subA-subE, subB+) and the *pal-1*MRE site. Underlined is the MRE site within each sequence. The $K_{d,app}$ is given for each sequence.

MEX-3 binds with high affinity to *nos-2* subC ($K_{d,app}$= 40 ± 5 nM; Figure 4.7B) and the *pal-1* 3'-UTR fragment ($K_{d,app}$ = 25 ± 4 nM; Figure 4.7B), and with low affinity to *nos-2* subA, subD, and subE ($K_{d,app}$ > 200 nM). Both *nos-2* subC and *pal-1* MRE RNA reveal two shifted species, which is likely due to additional partial binding sites present within the sequences. We were surprised to observe that MEX-3 binds with low affinity to *nos-2* subB ($K_{d,app}$ = 420 ± 20). Because the MRE is near the 3'-end of the sequence, we suspected that the fluorescein label might be disrupting the predicted interaction. A longer variant (subB+) was prepared that includes seven additional nucleotides downstream. Consistent with our hypothesis, MEX-3 binds to this RNA with four-fold higher affinity ($K_{d,app}$ =106 ± 2 nM; Figure 4.7B). The data reveal that MREs are present within both MEX-3 regulatory targets, and that MEX-3 binds to them with high affinity.

To verify that the MRE is essential for the interaction of MEX-3 with its mRNA targets, a series of mutations of the *nos-2* subC element was prepared and the ability of MEX-3 to bind to each was determined (Figure 4.8). Each half-site was mutated such that either AUAG or UUUA was changed to CCCC. In each case, the binding affinity is significantly reduced (mut1, mut2; $\Delta\Delta G > 0.5$ kcal mol^{-1}). The binding affinity was then tested when both half-sites were mutated to CCCC and a dramatic reduction in binding is observed (mut3; $\Delta\Delta G >$ 1.5 kcal mol^{-1}). When the half-site UUUA is mutated to AAAA, a loss in affinity is observed similar to mut1 or mut2 (mut5; $\Delta\Delta G > 0.5$ kcal mol^{-1}). Control mutations outside of the MRE have no significant change in relative binding



ID	Sequence
WT subC	CCCGUUCAUAGCCUUUAUUGAUUCCAAAUUU
Mut1	CCCGUUCAUAGCC**CCCC**UUGAUUCCAAAUUU
Mut2	CCCGUUC**CCCC**CCUUUAUUGAUUCCAAAUUU
Mut3	CCCGUUC**CCCC**CC**CCCC**UUGAUUCCAAAUUU
Mut4	CCCG**AAA**AUAGCCUUUAUUGAUUCCAAAUUU
Mut5	CCCGUUCAUAGCC**AAA**AUUGAUUCCAAAUUU
Mut6	CCCGUUCAUAGCCUUUAUUGA**AAA**CAAAUUU

Figure 4.8

Figure 4.8: Mutagenesis of the *nos-2* subC regulatory element. The bar graph represents the $\Delta\Delta G°$ calculated from the $K_{d,app}$ of each mutant sequence compared to the $K_{d,app}$ of wild-type subC RNA. Mutations that have greater than a two-fold loss in binding affinity are shown in dark grey (dotted line, $\Delta\Delta G° > 0.5$ kcal mol^{-1}). Below the graph are the sequences tested for MEX-3 binding. The MRE half-sites are underlined in the wild-type sequence, and the mutations made are shown in bold (Mut1-6).

affinity (mut4, mut6; $\Delta\Delta G < 0.5$ kcal mol^{-1}). Together, the data show that the MRE is required for the high affinity interaction between MEX-3 and *nos-2* subC.

The expression pattern of a nos-2 3'-UTR reporter depends upon the MRE

To validate the relevance of the MRE consensus *in vivo*, we examined the expression pattern of a *nos-2* 3'-UTR reporter where specific mutations are made to both MRE sites. We employed MosSCI (Frokjaer-Jensen et al., 2008) to generate single copy transgenic strains encoding green fluorescent protein (GFP) fused to histone H2B with the *nos-2* 3'-UTR downstream (Figure 4.9A). The *pie-1* promoter was used in each transgene to drive germline transcription. Based on our biochemical assessment of the interaction between MEX-3 and RNA containing an MRE, the first half-site ATAG was mutated to CCCC in each MRE present within the *nos-2* 3'-UTR (MREmut). We chose to mutate these nucleotides because this half-site has the most dramatic effect on MEX-3 binding (Figure 4.8). The UUUA half site is located within a conserved repeat element that has previously been shown to be required for MEX-3 dependent regulation (D'Agostino et al., 2006; Jadhav et al., 2008).

The expression pattern of the reporter harboring the wild-type *nos-2* 3'-UTR matches the previously reported endogenous NOS-2 expression pattern (Figure 4.9B) (D'Agostino et al., 2006; Subramaniam and Seydoux, 1999). The reporter is absent in early embryos, and is only observed in the germline precursor cells in older embryos. In stark contrast, the transgenic reporter that

contains the *nos-2* MREmut 3'-UTR is observed both in early embryos and in all cells of embryos at approximately the 28-cell stage (Figure 4.9B). Thus, mutation of the MREs disrupts both the spatial and temporal regulation of reporter expression. A similar pattern is observed with the wild-type reporter when *mex-3* mRNA is depleted by RNAi (Figure 4.9B). The results demonstrate that both MEX-3, and the MREs, are required to appropriately pattern the expression of the reporter transgene.

It is interesting to note that reduction of MEX-3 by RNAi appears to have a more dramatic effect on *nos-2* reporter expression than mutating its cis-acting response elements within the *nos-2* 3'-UTR (Figure 4.9B). A possible explanation for this observation is that residual binding of MEX-3 to *nos-2* 3'-UTR mutants leads to partial repression, as only part of the MRE is disrupted in each case. There may also be other specificity determinants from SELEX that we have not distinguished. Another possibility is that MEX-3 might regulate other *trans*-acting factors that feed back to repress *nos-2* translation. Alternatively, other proteins that bind to the *nos-2* 3'-UTR may facilitate MEX-3 binding and this interaction has not been fully disrupted. Our data cannot distinguish between these possibilities.

Candidate MEX-3 regulatory targets based on the MRE

To identify new candidate MEX-3 targets, we used the MRE consensus as a pattern to search annotated *C. elegans* transcripts for potential MEX-3 binding

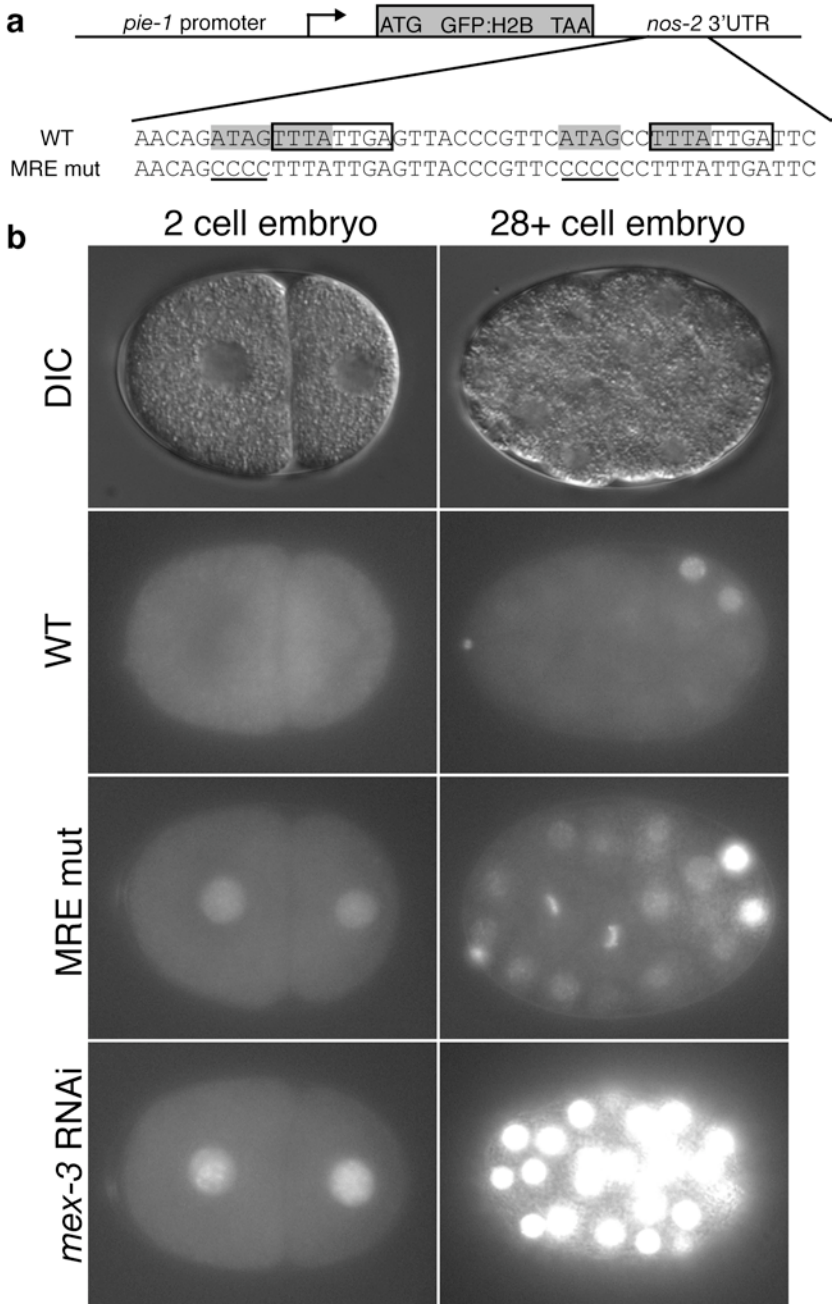

a
pie-1 promoter
ATG GFP:H2B TAA
nos-2 3'UTR
WT
MRE mut
AACAGATAGTTTATTGAGTTACCCGTTCATAGCCTTTATTGATTC
AACAGCCCCTTTATTGAGTTACCCGTTCCCCCCCTTTATTGATTC
b
2 cell embryo
28+ cell embryo
DIC
WT
MRE mut
mex-3 RNAi

Figure 4.9

Figure 4.9: Validation of the MRE within the regulatory target *nos-2*. **A.** A Schematic of the *nos-2* 3'-UTR reporter transgene. A *pie-1* promoter was used to drive germline specific transcription. Below the schematic is the region of the *nos-2* 3'-UTR containing both MRE sites. Each MRE is highlighted in grey. The conserved octamer is surrounded by a box, which was reported previously to be required for translation (D'Agostino et al., 2006). Half of each MRE site was mutated to CCCC (underlined) to make the *nos-2* MREmut reporter construct. **B.** Expression of the GFP::H2B *nos-2* 3'-UTR reporter constructs in early embryos. DIC images of 2-cell stage embryos and later staged (28+cells) are shown. The GFP::H2B reporter expression pattern is compared between the *nos-2* 3'-UTR, *nos-2* MREmut 3'-UTR, and the *nos-2* 3'-UTR reporter in *mex-3* RNAi embryos.

sites. MREs throughout the entire genome were identified using the pattern matching tool PATSCAN (Dsouza et al., 1997). Because 3'-UTRs are the primary determinant of spatial gene expression within the germline, we cross-referenced the predicted MREs with 3'-UTR annotations from release WS190 of the *C. elegans* genome (Merritt et al., 2008). Out of 10,802 genes with an annotated 3'-UTR, 2,834 (26.2%) contain at least one MRE in their 3'-UTR (for dataset refer to Table S1) (Pagano et al., 2009). To identify candidate MEX-3 targets required for embryogenesis, we filtered the results to include only transcripts present in 1-8 cell embryos (Baugh et al., 2003) and that result in an embryonic lethal phenotype when knocked-down by RNAi (Sonnichsen et al., 2005). Based on this analysis, 214 candidate MEX-3 targets were identified (7.5% of all MRE-containing 3'-UTRs, Figure 4.10).

To identify candidate MEX-3 targets that play a role in maintaining germline stem cell totipotency, we filtered our primary search results using microarray data that identify transcripts enriched in the germline (Reinke et al., 2004). Of the germline enriched genes, 527 have at least one MRE (18.6% of all MRE-containing 3'-UTRs, Figure 4.10). Because the role of MEX-3 in the germline is only revealed in the context of a *mex-3 gld-1* double mutant (Ciosk et al., 2006), we rationalized that the most important targets might contain both MEX-3 and GLD-1 binding sites. To identify such transcripts, we searched the 3'-UTRs of germline enriched genes for both the MRE and the STAR-binding element (SBE), recognized by GLD-1 (for dataset refer to Table S1)

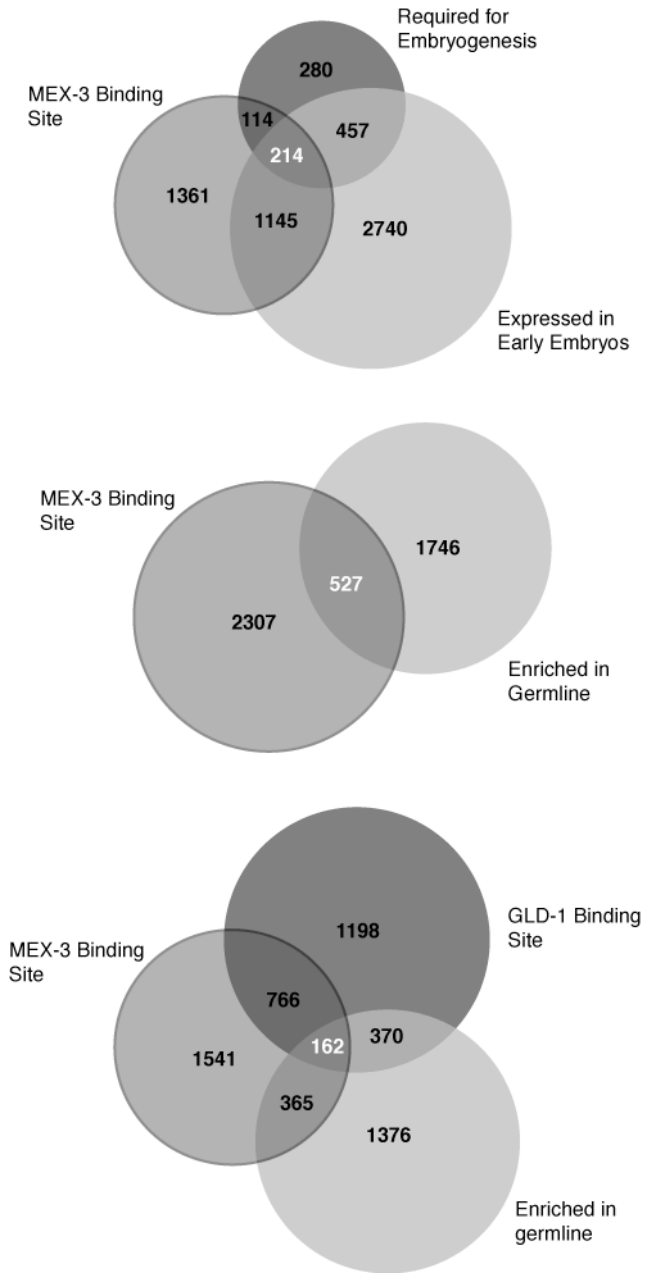

Required for
Embryogenesis
280
MEX-3 Binding
Site
114
457
214
1361
1145
2740
Expressed in
Early Embryos
MEX-3 Binding
Site
1746
527
2307
Enriched in
Germline
1198
GLD-1 Binding
Site
MEX-3 Binding
Site
766
370
162
1541
365
1376
Enriched in
germline

Figure 4.10

Figure 4.10: Bioinformatic analysis of genes that contain an MRE within their 3'-UTR. Genes that harbor a MEX-3 binding site were compared with transcripts that are expressed in early embryos, required for embryogenesis, enriched in the germline, and/or contain a GLD-1 binding site.

(Pagano et al., 2009). Based on this analysis, a total of 162 genes were identified as potential targets of both MEX-3 and GLD-1 (Figure 4.10). Together, the results predict several novel candidate MEX-3 targets that play critical roles both in the embryo and in the germline (Table 4.1).

We wished to examine predicted novel targets in a functional assay. As a first step towards this goal, we determined the MEX-3 dependence of reporter expression in ten transgenic lines harboring germline-specific GFP::H2B::3'-UTR reporters containing at least one MRE. These strains represent a subset of a germline 3'-UTR reporter library developed by Seydoux and colleagues (Merritt et al., 2008). Worms were grown on either OP50 or *mex-3* RNAi food and the expression pattern of GFP was determined (Table 4.2). Two of the ten (*pal-1* and *glp-1* 3'-UTR reporter strains) demonstrated a change in reporter fluorescence upon *mex-3* RNAi.). *pal-1* has previously been demonstrated to be a MEX-3 regulatory target and thus serves as a positive control (Hunter and Kenyon, 1996). The *glp-1* reporter shows an overall increase in GFP fluorescence and ectopic expression in the posterior of early embryos (2-4 cell) when *mex-3* RNA is depleted (Figure 4.11). There are two MEX-3 binding sites in a conserved region of the *glp-1* 3'-UTR. The binding sites surround previously characterized binding sites for the RNA-binding proteins POS-1 and GLD-1 (Farley et al., 2008; Ryder et al., 2004). The location of the binding sites suggests that regulation is direct, and could potentially be influenced by the presence of other RNA-binding proteins. The results identify *glp-1* as a new

candidate MEX-3 regulatory target; yet demonstrate that the presence of an MRE is not the sole determinant of MEX-3 dependent regulation.

Subset of candidate MEX-3 mRNA targets

Gene	Transcript ID	Function	Sites
ced-3	C48D1.2	protease required for apoptosis	4
cpb-1	C40H1.1	CPEB protein	2
cpb-3	B0414.5	CPEB protein	3
cye-1	C37A2.4b	Cyclin E/G1 cell cycle regulator	2
daz-1	F56D1.7	RNA binding protein	1
fog-3	C03C11.2	sex determination	1
gld-1	T23G11.3	RNA binding protein	3
glh-3	B0414.6	germline RNA helicase	2
glh-4	T12F5.3.1	germline RNA helicase	2
glp-1	F02A9.6	NOTCH receptor	2
gna-2	T23G11.2	Glucosamine phosphate N-Acetyl transferase	1
him-3	ZK381.1	sex determination	2
let-711	F57B9.2	mRNA deadenylation	3
mes-3	F54C1.3a	X chromosome silencing	1
mes-6	C09G4.5	X chromosome silencing	1
mex-3	F53G12.5b	RNA binding protein	2
mig-5	T05C12.6b	dishevelled homolog/Wnt signalling	1
mom-4	F52F12.3	Wnt signalling	2
nos-2	ZK1127.1	Nanos homologue	2
nst-1	K01C8.9	larval germline proliferation	1
ooc-3	B0334.11a	axis polarity determinant	1
pal-1	C38D4.6a.2	Caudal homeodomain protein	2
pha-4	F38A6.1a	FoxA transcription factor/pharynx	1
puf-3	Y45F10A.2.1	RNA binding protein	1
puf-6	F18A11.1.1	RNA binding protein	1
puf-7	B0273.2.1	RNA binding protein	1
sox-2	K08A8.2a.1	HMG-box transcription factor	1
spe-41	K01A11.4.2	spermatogenesis	1
spn-4	ZC404.8.1	RNA binding protein	1
tbx-2	F21H11.3.2	T-box transcrption factor/phayrnx	1
tra-2	C15F1.3a.1	sex determination	3
ubc-9	F29B9.6.1	ubiquitin-conjugating enzyme	4

A list of candidate regulatory targets of MEX-3 involved in germline and/or early development in *C. elegans*. The number of MRE sites within the 3'UTR of each gene is listed.

Table 4.1

Results from the *mex-3* RNAi screen

Strain/ ID	Function	OP50 worms imaged	MEX-3 RNAi worms imaged	GFP expression in Seydoux 3'UTR fusions	Observations
gld-1 JH2436	RNA binding protein	13	8	Pachytene	No change in GFP expression between MEX-3 RNAi and OP50 worms
spn-4 JH2311	RNA binding protein	14 (embryos imaged; 2-4 cell stage)	11 (embryos imaged; 2-4 cell stage)	Oocytes	No change in GFP expression between MEX-3 RNAi and OP50 worms
glp-1 *sprSi2*	Notch receptor	40	37	Progenitors	Observed an increase in GFP expression in MEX-3 RNAi oocytes (27/37) and early embryos ($n = 41$) compared to OP50 oocytes (4/40) and embryos ($n = 43$)
pal-1 JH2236	Caudual homeodomain protein	15 (embryos imaged; 2-4 cell stage)	13 (embryos imaged; 2-4 cell stage)	Oocytes	Observed GFP expression in MEX-3 RNAi early embryos; low expression in OP50 early embryos compared to MEX-3 RNAi
daz-1 JH2223	RNA binding protein	13	16	Progenitors	No change in GFP expression between MEX-3 RNAi and OP50 worms
cye-1 JH2261	E/G1 cell cycle regulator	17	11	Mixed	No change in GFP expression between MEX-3 RNAi and OP50 worms
him-3 JH2324	Sex determination	10	9	Pachytene	No change in GFP expression between MEX-3 RNAi and OP50 worms
mes-3 JH2377	X chromosome silencing	11	15	Mixed	No change in GFP expression between MEX-3 RNAi and OP50 worms
spe-41 JH2381	Spermatogenesis	7	10	Sperm	No change in GFP expression between MEX-3 RNAi and OP50 worms
fog-2 JH2207	Feminization of germline	10	12	Ubiquitous	No change in GFP expression between MEX-3 RNAi and OP50 worms

Table 4.2

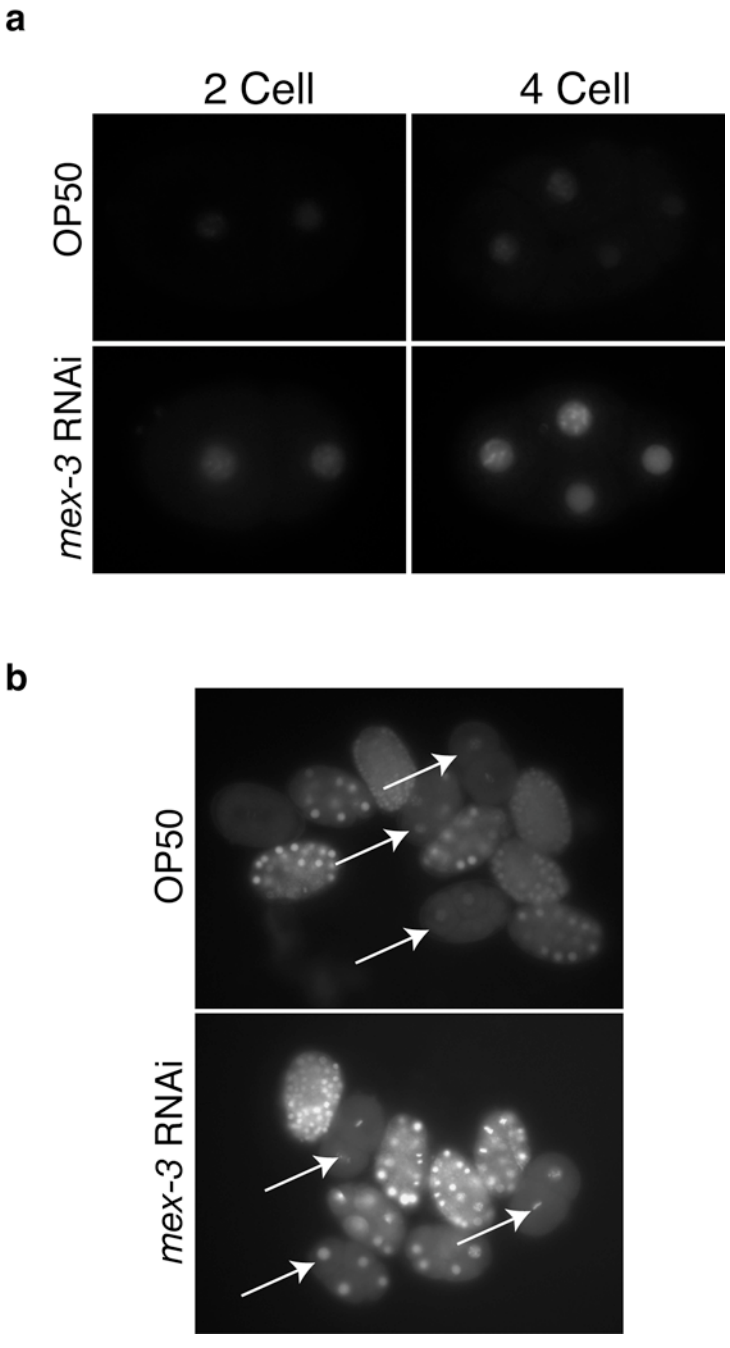


Figure 4.11

Figure 4.11: A glp-1 3'-UTR reporter is dependent upon MEX-3 for its expression pattern. **A.** There is an increase in GFP expression in early embryos when *mex-3* mRNA is knocked down by RNAi. The image compares a 2 and 4 cell embryo when grown on OP50 food or *mex-3* RNAi food. **B.** A group of embryos expressing the *glp-1* 3'-UTR reporter are shown. Arrows are pointing to 2 and 4 cell embryos.

Discussion

Based on *in vitro* selection and biochemical experiments, we have defined the consensus MEX-3 recognition element (MRE: (A/G/U)(G/U)AGN$_{(0-8)}$U(U/A/C)UA). The MRE is a bipartite element that consists of two four-nucleotide motifs separated by 0-8 nucleotides. MEX-3 contains two KH RNA binding domains. We predict that each motif binds specifically to one four-nucleotide half-site. Previous studies have shown that KH domains typically accommodate four nucleotides in their binding pockets, consistent with this hypothesis (Braddock et al., 2002; Gamarnik and Andino, 2000; Lewis et al., 2000; Makeyev and Liebhaber, 2002; Valverde et al., 2008). Alignment of the KH domains between the human hMex-3 proteins and *C. elegans* MEX-3 reveal 79-81% sequence identity within the RNA-binding domain (Buchet-Poyau et al., 2007). Interestingly, part of the MRE is comprised of an AU-rich half-site, which is proposed to be required for hMEX-3D/TINO RNA binding. We predict that the hMEX-3 proteins bind to RNA with similar specificity as their *C. elegans* homolog MEX-3.

The MRE is present in the 3'-UTR of approximately 26.2% of all genes in *C. elegans*. By filtering this list for transcripts expressed in the same tissues at the same time as MEX-3, we have narrowed the candidate targets to 214 transcripts in the embryo and 527 transcripts in the germline. The two previously characterized MEX-3 regulatory targets—*nos-2* and *pal-1*—both contain two MRE sites in their 3'-UTR. Mutating both sites in a *nos-2* reporter leads to

derepression of the transgene, as does reduction of MEX-3 by RNAi. Thus, the MRE is a critical *cis*-acting functional element.

Our search for candidate regulatory targets reveals several genes that play key roles during various stages of development. A number of these genes encode other RNA binding proteins, including *cbp-3*, *puf-3*, *puf-6*, *puf-7*, *spn-4*, *mex-1*, and *gld-1* (Gomes et al., 2001; Jones et al., 1996; Lublin and Evans, 2007; Mello et al., 1992; Ryder et al., 2004; Stumpf et al., 2008). Interestingly, the *mex-3* transcript also has two MREs within its own 3'-UTR. This suggests that MEX-3 may regulate its own mRNA translation. Other candidate mRNA targets encode membrane proteins such as *glp-1* and *ooc-3*. GLP-1 is a Notch homolog essential for mitotic proliferation of germ cells and maintenance of germline stem cells as well as the development of early blastomeres (Hutter and Schnabel, 1994; Hutter and Schnabel, 1995). Our results suggest MEX-3 regulates this transcript, but more work is needed to show that regulation is direct. OOC-3 is a putative transmembrane protein localized to the endoplasmic reticulum and is required for the correct localization of the polarity determinants PAR-2 and PAR-3, two critical proteins that establish asymmetry in the early embryo immediately after fertilization (Basham and Rose, 1999; Pichler et al., 2000).

Of particular interest is how MEX-3 functions to maintain totipotency in the germline. The unusual trans-differentiated germline teratoma phenotype is only observed in worms with mutations in both *mex-3* and *gld-1*, a second KH-domain

RNA binding protein (Ciosk et al., 2006). Moreover, recent evidence indicates that MEX-3 and PUF-8, a Puf-domain RNA binding protein, promote mitotic division of germ cells (Ariz et al., 2009). This suggests that MEX-3 is playing redundant roles with other RNA binding proteins in order to ensure preservation and maintenance of germline stem cells. One potential mechanism by which this may occur is co-regulation of the same mRNA target. A potential target that is recognized by both MEX-3 and GLD-1 is the mRNA encoding the putative transcription factor SOX-2. *C. elegans* SOX-2 is homologous to human SOX2, which is required for embryonic stem cell self-renewal and is one of the factors used to induce pluripotency in induced pluripotency stem cells (iPS) (Fong et al., 2008; Takahashi and Yamanaka, 2006). *C. elegans sox-2* mRNA harbors one MRE and one SBE (GLD-1 binding sequence), consistent with the hypothesis that MEX-3 and GLD-1 play a redundant role in regulating SOX-2 expression.

Here, we reveal the requirements for RNA recognition by MEX-3 and identify *cis*-acting elements in the 3'-UTR of its mRNA targets *nos-2* and *pal-1*. Our results demonstrate that a large number of transcripts contain a MEX-3 recognition element within their 3'-UTR, but caution that the MRE is not the sole determinant of MEX-3-dependent regulation. Our results reveal a number of candidate targets that could potentially explain the diverse roles of MEX-3 in regulating embryonic cell fate specification and maintaining totipotency within the germline. Further work is needed to define which candidate targets are the most important factors contributing to the *mex-3* mutant phenotypes.

Materials and Methods

Protein expression and purification

The sequence encoding amino acids 45-205 of MEX-3 was amplified from the corresponding ORFeome clone (Open Biosystems) and subcloned into pMal-c (New England Biolabs), a protein expression vector that encodes an N-terminal maltose-binding protein (MBP) tag. MBP-MEX-3(45-205) was expressed and purified from BL21 (DE3) Gold (Stratagene) *Escherichia coli*. A liquid culture grown at 37°C was induced at mid-log phase with 1 mM isopropyl 1-thio-β-D-galactopyranoside and grown for 3 hours before harvesting cells. Cells were lysed and purified using an amylose (New England Biolabs) affinity column, followed by HiTrap Q and source 15Q (GE Healthcare) ion exchange columns at 4°C. Purified MEX-3 was dialyzed into storage buffer (25 mM Tris, pH 8.0, 25 mM NaCl, 2 mM DTT) and stored at 4°C.

In Vitro RNA Selection

RNA library design and *in vitro* selection protocols were adapted from published protocols with a few modifications (Hori et al., 2005; Loughlin et al., 2009). The DNA library was amplified from the template 5'-GGGAAGATCTCGACCAGAAG-(N30)-TATGTGCGTCTACATGGATCCTCA with a forward (5' – CGGAATTCTAATACGACTCACTATAGGGAAGATCTCG ACCAGAAG - 3') and reverse (5' – TGAGGATCCATGTAGACGCACATA - 3') primer pair using three cycles of PCR to make the initial double stranded DNA

pool. RNA was transcribed from 10 pmol of starting DNA template and purified as described (Loughlin et al., 2009).

Binding reactions were carried out in 200 µL of selection buffer (25 mM HEPES, pH 7.4, 200 mM NaCL, 75 µg/ml tRNA, 0.01% Igepal CA-630). Between 1-800 nM of purified MBP-MEX-3(45-205) was equilibrated with transcribed RNA in selection buffer for 1 hour and then mixed with amylose resin (New England Biolabs). The amount of protein used was progressively decreased in later rounds to increase stringency. Unbound RNA was separated using a Zeba spin column (Pierce) and washed 2-8 times with 200 µL of selection buffer. MEX-3 bound to RNA was eluted from amylose resin with 10 mM maltose in selection buffer. All reactions were carried out at room temperature. Recovered RNA was phenol/chloroform extracted, ethanol precipitated and resuspended in 10 µL of TE buffer. RNA was then reverse transcribed and amplified with 10-15 rounds of PCR using the SuperScript III One-Step RT-PCR kit with Platinum Taq (Invitrogen). The new DNA pool was *in vitro* transcribed to generate the next RNA pool to be used.

A total of 7 rounds of selection were performed. To increase the selection stringency, rounds 6 and 7 included two washes where the salt concentration was increased to 500 mM NaCl. The DNA from pool 7 was subcloned into pUC18 using EcoRI and BamHI restriction enzymes, and DNA from individual isolates was sequenced.

Electrophoretic Mobility Shift Assay

RNA transcripts or synthetic oligonucleotides (Integrated DNA Technologies) were 3'-end labeled with fluorescein 5-thiosemicarbazide (Invitrogen) following the protocol previously described (Pagano et al., 2007). Electrophoretic mobility shift experiments and data analysis were carried out as previously described with a few modifications (Pagano et al., 2007). Varying concentrations of purified MEX-3 were equilibrated with 3 nM of labeled RNA in equilibration buffer (0.01% IGEPAL, 0.01 mg/ml tRNA, 10 mM Tris, pH 8.0, 100 mM NaCl) for 3 hours. Samples were loaded on a 5% slab polyacrylamide gel in 0.5X TBE buffer. Both the gel and running buffer were pre-chilled to 4°C before running each sample. The gels were run for 60 minutes at 120 volts and immediately scanned using a fluor-imager (Fujifilm FLA-5000) with a blue laser at 473 nm. Due to dissociation in the gel, the 5' Frag3 RNA EMSA experiment was fit to a Langmuir isotherm to determine the apparent equilibrium dissociation constant.

Worm Strain Generation

All transgenic worm strains were made using biolistic transformation (Praitis, 2006; Praitis et al., 2001) or MosSCI (Frokjaer-Jensen et al., 2008). The unspliced coding region and 3'-UTR of *mex-3* was PCR amplified from genomic DNA containing NheI and NsiI restriction sites and subcloned into the plasmid pJH4.52 (a generous gift of Dr. Geraldine Seydoux, Johns Hopkins University)

using the NsiI and SpeI sites to generate the construct pJMP001 (*Ppie-1*::GFP::MEX-3::*mex-3* 3'-UTR). The cloned *mex-3* insert was sequenced confirming 794 bp from the 5' end and 191 bp from the 3' end. The transgenic worm was made using biolistic transformation with *unc-119* rescue (Praitis, 2006; Praitis et al., 2001). An equal mixture of pJMP001 and pDEST-DD03 harboring the *unc-119* gene (a gift of Dr. Marian Walhout, University of Massachusetts Medical School) was co-bombarded to generate the transgenic worm (*sprIs1[Ppie-1::GFP::MEX-3::mex-3 3'-UTR, unc-119(+)]*).

The *nos-2* and *glp-1* 3'-UTR reporter constructs were made using the Gateway system (Walhout et al., 2000). The 3'-UTR of *nos-2* was PCR amplified from genomic DNA containing *attB* sites and recombined into pDONRP2R-P3 with Gateway BP Clonase II Enzyme Mix (Invitrogen) to generate the 3'-UTR entry clone pJMP015. Quickchange was performed on this plasmid to generate the *nos-2* MREmut 3'-UTR entry clone pJMP044. The *glp-1* 3'-UTR was amplified off of pCM5.40 (provided by the third generation Seydoux Lab Vector kit) with *attB* sites and used to generate pBMF3.1. Multisite Gateway reactions were performed using Gateway LR Clonase II Plus Enzyme Mix (Invitrogen) with the plasmids pCG142 or pCM1.111, pCM1.35 (provided by the third generation Seydoux Lab Vector kit), pCFJ150 (a generous gift from Dr. Erik Jorgensen, University of Utah), and either pJMP015 or pJMP044 to generate pJMP046 (*Ppie-1::GFP::H2B::nos-2 3'-UTR*), pJMP049 (*Ppie-1::GFP::H2B::nos-2 MREmut 3'-UTR*), and pBMF4.1 (*Pmex-5::GFP::H2B::glp-1 3'-UTR*). The plasmids

pJMP046, pJMP049, or pBMF4.1 were used along with pCFJ90, pCFJ104, and pJL43.1 and co-injected into the strain EG4322 to generate single copy gene insertions (*sprSi1[Ppie-1::GFP::histone H2B::nos-2 3'-UTR cb-unc-119(+)] II; sprSi2[Ppie-1::GFP::histone H2B::nos-2 MREmut 3'-UTR cb-unc-119(+)] II; sprSi3[Pmex-5::GFP::histone H2B::glp-1 3'-UTR cb-unc-119(+)] II*) following the direct insertion protocol previously described (Frokjaer-Jensen et al., 2008). The integrated worm strain *sprSi3* was confirmed by PCR and strains *sprSi1* and *sprSi2* were confirmed by PCR and sequencing.

RNAi experiments were performed using *mex-3* RNAi food (a generous gift from Dr. Craig Mello). The 3'-UTR reporter strains (JH2436, JH2311, JH2236, JH2223, JH2261, JH2324, JH2377, JH2381, and JH2207) used in the *mex-3* RNAi screen were obtained from the Caenorhabditis Genetics Center (CGC). All strains were grown at 25°C with the exception of *sprIS1*, *sprSI3*, and JH2261, which were grown at room temperature (~20°C). This was done to decrease the stability of GFP::histone H2B reducing background fluorescence that does not recapitulate the expression of endogenous protein. DIC and GFP images were collected with live specimens using a Zeiss Axioskop microscope with 40X or 100X objectives. All GFP::H2B reporter images were taken with the same exposure time and contrasted equally for each strain.

Bioinformatics

MREs were identified in release WS190 of the *C. elegans* genome using the pattern matching tool PATSCAN (Dsouza et al., 1997). Using a custom MySQL database, the predicted MREs were cross-referenced with the 3'-UTR annotations for all transcripts from Wormbase release WS190 to identify genes with 3'-UTR containing candidate MEX-3 binding sites. To identify genes that may be regulated by MEX-3, the results were filtered through genome-wide data sets describing genes expressed during one to eight cell embryos (Baugh et al., 2003), genes required to complete embryogenesis (Sonnichsen et al., 2005), and genes whose expression is increased in the germline relative to an average of all tissues (Reinke et al., 2004).

CHAPTER V

Final summary and concluding remarks

The purpose of this study was to test the hypothesis that MEX-5 and MEX-3 are specific RNA binding proteins that promote anterior development through direct regulation of maternal transcripts in the nematode *C. elegans*. Attention was focused on characterizing the RNA binding specificity as well as the mRNA target specificity of both proteins to further elucidate their role in early development. Both quantitative biochemical approaches and molecular techniques were employed to dissect the binding parameters of MEX-5 and MEX-3.

MEX-5 RNA recognition

I have demonstrated that MEX-5 is an RNA binding protein that recognizes linear RNA containing a tract of six or more uridines within an eight-nucleotide window. This is different from its human homologs TTP and Tis11d, which bind with high specificity to UUAUUUAUU elements (Hudson et al., 2004). To identify residues that may contribute to this difference in RNA recognition between MEX-5 and TTP/ERF-2, a homology model was prepared based on the NMR structure of ERF-2. The NMR structure reveals that a glutamate residue within each CCCH finger makes a base specific hydrogen bond with the adenosine base of UAUU motif (Hudson et al., 2004). MEX-5 on the other hand contains a non-conserved arginine and lysine at analogous positions in the first and second finger, respectively. The homology model predicts that this residue rotates away forming backbone contacts with adjacent nucleotides, a possible

explanation for MEX-5 RNA recognition. Using site directed mutagenesis I was able to confer TTP-like specificity to MEX-5 when both basic residues were mutated to an acidic glutamate, demonstrating a basis for specificity among CCCH TZF proteins. By understading the binding specificity of MEX-5, we show that approximately 91% of annotated 3'-UTRs contains at least one MEX-5 binding site. In addition, 29 out of the top 30 octamers contain at least six uridines.

The role of MEX-5 in early development

As my data has shown, MEX-5 is an RNA binding protein with low specificity, yet the purpose of this function in early development remains unclear. We show that the presence of a MEX-5 binding site is present in 3'-UTR space more frequently than expected from a random distribution. Therefore, contrasting with previous hypotheses, MEX-5 does not bind to RNA with sufficient specificity to drive regulation of a subset of maternal transcripts.

The low RNA binding specificity of MEX-5 has several implications for its function in development. It is possible that it is a general mRNA repressor, silencing all transcripts in the oocyte and the anterior lineage until it is destroyed after the 8-cell stage. This is unlikely because translation of *glp-1* mRNA occurs in the anterior and appears to require MEX-5 (Evans et al., 1994; Schubert et al., 2000). Alternatively, MEX-5 may be an affinity adapter for a more specific RNA-binding protein, regulating expression from just a few maternal transcripts as part

of a cooperative complex. If so, then MEX-5 could behave like *Drosophila melanogaster* Nanos protein, which works in complex with Pumilio to regulate maternal mRNA expression (Asaoka-Taguchi et al., 1999; Curtis et al., 1997; Forbes and Lehmann, 1998). Finally, MEX-5 RNA-binding activity may not contribute to maternal RNA regulation. Instead, RNA-binding activity may be required to target MEX-5 to the posterior centrosome and to P-granules (Figure 3.1A), two subcellular organelles that contain RNA (Alliegro et al., 2006; Cuenca et al., 2003; Pitt et al., 2000; Schubert et al., 2000). If so, then MEX-5 RNA binding activity may play a role in germline maintenance in addition to its role in germline formation.

Further studies should be performed to clarify the purpose of MEX-5 RNA binding activity. One possible approach would be to test if the RNA binding activity is required for embryonic development. This could be addressed by performing rescue experiments using variants of the *mex-5* gene in a *mex-5* mutant background. These constructs could contain the same amino acid changes described in Chapter III that were shown to change the specificity of MEX-5. An additional approach would be to examine the localization of MEX-5 when these same mutations are introduced.

As described in Chapters I and III, MEX-5 is thought to have at least a dual role in the early embryo to specify anterior development. This includes regulation of maternal RNAs post-transcriptionally, and also promoting the somatic degradation of the germline factors PIE-1, MEX-1, POS-1, as well as

MEX-5 itself. This second function is achieved by MEX-5 activating *zif-1*, which causes the ubiquitination and degradation of target CCCH finger proteins (DeRenzo et al., 2003). Although MEX-5 interacts with ZIF-1 through a yeast-2-hybrid assay, it remains unclear whether this activity is responsible to activate *zif-1*. This does not exclude the possibility that MEX-5 activates *zif-1* at the mRNA level, or even indirectly through another factor. My data reveal that MEX-5 has a low RNA binding specificity, so it would be difficult to map possible binding sites within the *zif-1* transcript with high confidence.

MEX-3 RNA recognition

My studies also defined the RNA binding properties of MEX-3. The MEX-3 recognition element (MRE) is defined as (A/G/U)(G/U)AGN$_{(0-8)}$U(U/A/C)UA; a bipartite sequence comprised of two four-nucleotide elements with variable spacing. Both putative mRNA targets of MEX-3, *pal-1* and *nos-2*, contain two MRE sites and I demonstrate that MEX-3 binds specifically to these elements. To test that the MRE is an actual *cis*-regulatory element, I examined the expression pattern of a *nos-2* 3'-UTR reporter when the MRE sites are mutated. Mutation of these sites leads to derepression of the reporter in a similar manner as reducing MEX-3 levels by RNAi. This is evidence that *nos-2* is a direct regulatory target of MEX-3. The MRE is present in the 3'-UTR of approximately 26% of all genes in *C. elegans*. We also find that the MRE is necessary but not sufficient for regulation by MEX-3.

The role of MEX-3 in germline and early development

The RNA binding activity of MEX-3 is clearly responsible for targeted regulation of maternal RNA to promote development of the early embryo. My studies have shown that MEX-3 binds specifically to its putative targets *nos-2* and *pal-1*, yet it remains unclear what other factors are regulated by MEX-3 to specify the AB blastomere. As AB differentiates into pharynx, its possible that MEX-3 is responsible to regulate factors that are necessary for pharyngeal development. Three possible genes are *ubc-9*, *pha-4*, and *tbx-2*, all of which are required to specify ABa-derived pharyngeal muscle (Mango et al., 1994; Roy Chowdhuri et al., 2006). The 3'-UTR of *ubc-9* contains four MRE sites, while *tbx-2* and *pha-4* each have one. Similar to *mex-3* mutant embryos, worms that lack these genes fail to produce pharyngeal tissue from the ABa blastomere. My results also identify *glp-1* as a new putative target, but it is not yet known if this is a direct target. The data suggest that MEX-3 may repress GLP-1 expression in the anterior and posterior blastomeres from the two to four-cell stage, but it is not clear what role MEX-3 plays in regulating GLP-1. It is possible that MEX-3 may prevent too much *glp-1* from being translated in the anterior blastomeres, as it is normally expressed in these cells in wild-type embryos.

As described in Chapters I and IV, MEX-3 also plays a role in germline development and may function redundantly with other factors (GLD-1, PUF-8) to maintain totipotency in the germline. One group of genes that may be regulated by MEX-3 and GLD-1 is the maternal effect sterile (*mes*) genes. The proteins

MES-3 and MES-6, along with a third protein MES-2 form a DNA-methylating Polycomb-like complex that modifies histone H3 and K27 residues thus silencing the X chromosome in germline cells (Bender et al., 2004). Post-transcriptional regulation of the *mes* genes by MEX-3 and GLD-1 may be required to ensure appropriate transcriptional silencing during germline development and in adult germline stem cells. The 3'-UTR of *mes-3* contains both GLD-1 and MEX-3 binding sites, and the 3'-UTR of *mes-6* contains a binding site for MEX-3. GLD-1 is already known to bind specifically to the 3'-UTR of *mes-3* and represses its translation in the meiotic syncitial region of the germline (Xu et al., 2001). Transcriptional deregulation caused by mis-expression of the *mes* genes could drive expression of transcripts that promote cellular differentiation.

A mechanistic view of MEX-5 and MEX-3 RNA regulation

Thus far, the mechanism of RNA regulation by MEX-5 and MEX-3 remains unknown. I propose that MEX-5 serves as multifunctional protein in terms of RNA regulation, while MEX-3 functions as a translational repressor. In the case of MEX-5, previous studies show that MEX-5 protein is required to regulate the abundance of *nos-2, mex-1,* and *pos-1* transcript levels, suggesting that MEX-5 may regulate mRNA stability (D'Agostino et al., 2006; Tenlen et al., 2006). On the other hand, MEX-5 is required for GLP-1 protein expression in the anterior blastomeres. The mRNA of *glp-1* is present in all blastomeres, yet the protein is only expressed in the anterior descendents of AB at the four-cell stage (Schubert

et al., 2000). This suggests that MEX-5 may also function as a translational activator. However, no mechanistic studies have been performed to test these hypotheses.

MEX-3 may function as a translational repressor to specify the AB blastomere. In the case of *pal-1*, the mRNA is present throughout the entire embryo at the four-cell stage, yet the protein is expressed only in the posterior blastomeres (Hunter and Kenyon, 1996). This expression pattern of PAL-1 anti-correlates with MEX-3, suggesting MEX-3 translationally represses *pal-1*. Although this proposes a role for MEX-3 regulation, one cannot rule out the possibility that MEX-3 functions to regulate mRNA stability, localization, or other mechanisms. *pal-1* is only one of potentially several mRNA targets of MEX-3, and there are no mechanistic studies of this RNA binding protein thus far.

mRNA target specificity

My findings reveal that MEX-5 and MEX-3 do not recognize RNA targets with very high sequence specificity, but instead are more degenerate. When analyzing every annotated 3'-UTR in *C. elegans*, we observe that over 90% of genes contain a MEX-5 binding site, while 26% contain a MEX-3 binding site. This leads us to the difficult problem of how mRNA target specificity and/or regulation is achieved. One possibility is that these proteins function in multifactor complexes. Specificity may be accomplished through a combination of RNA binding proteins where individually, each has relatively low sequence

specificity, but together are very specific to target mRNAs. Alternatively, the location of a binding site within a given UTR may cause a transcript to be more accessible for regulation by *trans*-acting factors. The determinant *glp-1* is an example that depends upon several RNA binding proteins for its regulation throughout the germline and early embryo (GLD-1, POS-1, MEX-3, MEX-5/6, SPN-4, PUF-5/6/7) (Farley and Ryder, 2008). Analysis of the *glp-1* 3'-UTR reveals a group of previously characterized functional elements present within a highly conserved region of the UTR (Figure 5.1). Perhaps clusters of binding sites found within conserved regions of UTR space sets a precedent for mRNA target specificity. This paradigm would greatly facilitate target identification and help map out regulatory networks for RNA binding proteins.

Another way mRNA target specificity could be achieved is through the presence of multiple binding sites within a transcripts 3'-UTR. As mentioned in Chapter IV, both MEX-3 targets *nos-2* and *pal-1* contain two MREs in their 3'-UTR. *glp-1* is a new putative target and also has two MRE sites within its 3'-UTR. However, MEX-3 is capable of binding to a single MRE with high affinity *in vitro*, and there is no evidence to indicate that a single MRE is not sufficient to support MEX-3 dependent regulation. Future studies will help clarify what role the context of a binding site plays that favors regulation of a transcript. For this to be accomplished, it is critical to continue dissecting the requirements for RNA recognition of individual RNA binding proteins as was done with MEX-5 and MEX-3. In addition to quantitative biochemical approaches, it will be useful to

identify mRNA targets of specific RNA binding proteins using high-throughput approaches *in vivo*. These include techniques such as RNA co-immunoprecipitation with sequencing or RT-PCR to screen candidate mRNA targets. Comparing targets identified *in vivo* with targets predicted *in silico* may reveal regions in sequence space that are preferred for RNA binding.

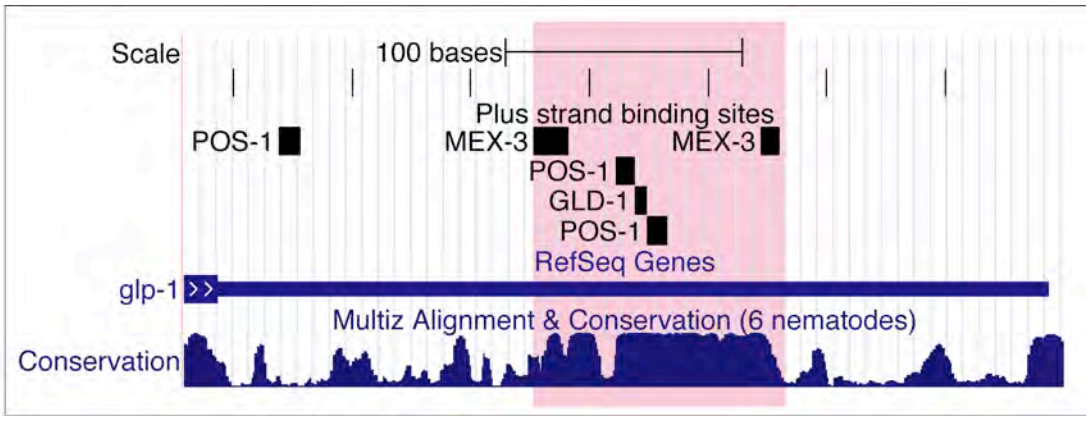

Scale
100 bases
Plus strand binding sites
POS-1
MEX-3
MEX-3
POS-1
GLD-1
POS-1
RefSeq Genes
glp-1
Multiz Alignment & Conservation (6 nematodes)
Conservation

Figure 5.1

Figure 5:1: Presence of functional elements within the *glp-1* 3'-UTR. The binding sites of GLD-1, MEX-3 and POS-1 are located within a highly conserved region of the *glp-1* 3'-UTR. Conservation predictions are derived from *C. briggsae*, *C. remanei*, *C. brenneri*, *C. japonica*, and *Pristionchus pacificus* genome sequence information. The cluster of functional elements is highlighted in pink. This image has been exported from the UCSC genome browser created by the Genome Bioinformatics Group of UC Santa Cruz.

REFERENCES

Alliegro, M. C., Alliegro, M. A., and Palazzo, R. E. (2006). Centrosome-associated RNA in surf clam oocytes. Proc Natl Acad Sci U S A *103*, 9034-9038.

Ariz, M., Mainpal, R., and Subramaniam, K. (2009). C. elegans RNA-binding proteins PUF-8 and MEX-3 function redundantly to promote germline stem cell mitosis. Dev Biol *326*, 295-304.

Asaoka-Taguchi, M., Yamada, M., Nakamura, A., Hanyu, K., and Kobayashi, S. (1999). Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos. Nat Cell Biol *1*, 431-437.

Auweter, S. D., Oberstrass, F. C., and Allain, F. H. (2006). Sequence-specific binding of single-stranded RNA: is there a code for recognition? Nucleic Acids Res *34*, 4943-4959.

Barrasa, M. I., Vaglio, P., Cavasino, F., Jacotot, L., and Walhout, A. J. (2007). EDGEdb: a transcription factor-DNA interaction database for the analysis of C. elegans differential gene expression. BMC Genomics *8*, 21.

Basham, S. E., and Rose, L. S. (1999). Mutations in ooc-5 and ooc-3 disrupt oocyte formation and the reestablishment of asymmetric PAR protein localization in two-cell Caenorhabditis elegans embryos. Dev Biol *215*, 253-263.

Baugh, L. R., Hill, A. A., Slonim, D. K., Brown, E. L., and Hunter, C. P. (2003). Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development *130*, 889-900.

Bender, L. B., Cao, R., Zhang, Y., and Strome, S. (2004). The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans. Curr Biol *14*, 1639-1643.

Bowerman, B., Eaton, B. A., and Priess, J. R. (1992). skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early C. elegans embryo. Cell *68*, 1061-1075.

Bowerman, B., Ingram, M. K., and Hunter, C. P. (1997). The maternal par genes and the segregation of cell fate specification activities in early Caenorhabditis elegans embryos. Development *124*, 3815-3826.

Boyd, L., Guo, S., Levitan, D., Stinchcomb, D. T., and Kemphues, K. J. (1996). PAR-2 is asymmetrically distributed and promotes association of P granules and PAR-1 with the cortex in C. elegans embryos. Development *122*, 3075-3084.

Braddock, D. T., Baber, J. L., Levens, D., and Clore, G. M. (2002). Molecular basis of sequence-specific single-stranded DNA recognition by KH domains: solution structure of a complex between hnRNP K KH3 and single-stranded DNA. Embo J *21*, 3476-3485.

Brennecke, J., Stark, A., Russell, R. B., and Cohen, S. M. (2005). Principles of microRNA-target recognition. PLoS Biol *3*, e85.

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

Brewer, B. Y., Malicka, J., Blackshear, P. J., and Wilson, G. M. (2004). RNA sequence elements required for high affinity binding by the zinc finger domain of tristetraprolin: conformational changes coupled to the bipartite nature of Au-rich MRNA-destabilizing motifs. J Biol Chem *279*, 27870-27877.

Buchet-Poyau, K., Courchet, J., Le Hir, H., Seraphin, B., Scoazec, J. Y., Duret, L., Domon-Dell, C., Freund, J. N., and Billaud, M. (2007). Identification and characterization of human Mex-3 proteins, a novel family of evolutionarily conserved RNA-binding proteins differentially localized to processing bodies. Nucleic Acids Res *35*, 1289-1300.

Carballo, E., Lai, W. S., and Blackshear, P. J. (1998). Feedback inhibition of macrophage tumor necrosis factor-alpha production by tristetraprolin. Science *281*, 1001-1005.

Carrera, I., and Treisman, J. E. (2008). Message in a nucleus: signaling to the transcriptional machinery. Curr Opin Genet Dev *18*, 397-403.

Chao, J. A., Patskovsky, Y., Almo, S. C., and Singer, R. H. (2008). Structural basis for the coevolution of a viral RNA-protein complex. Nat Struct Mol Biol *15*, 103-105.

Chao, J. A., Patskovsky, Y., Patel, V., Levy, M., Almo, S. C., and Singer, R. H. (2010). ZBP1 recognition of beta-actin zipcode induces RNA looping. Genes Dev *24*, 148-158.

Ciosk, R., DePalma, M., and Priess, J. R. (2006). Translational regulators maintain totipotency in the Caenorhabditis elegans germline. Science *311*, 851-853.

Courchet, J., Buchet-Poyau, K., Potemski, A., Bres, A., Jariel-Encontre, I., and Billaud, M. (2008). Interaction with 14-3-3 adaptors regulates the sorting of hMex-3B RNA-binding protein to distinct classes of RNA granules. J Biol Chem *283*, 32131-32142.

Cuenca, A. A., Schetter, A., Aceto, D., Kemphues, K., and Seydoux, G. (2003). Polarization of the C. elegans zygote proceeds via distinct establishment and maintenance phases. Development *130*, 1255-1265.

Curtis, D., Treiber, D. K., Tao, F., Zamore, P. D., Williamson, J. R., and Lehmann, R. (1997). A CCHC metal-binding domain in Nanos is essential for translational regulation. Embo J *16*, 834-843.

Czworkowski, J., Odom, O. W., and Hardesty, B. (1991). Fluorescence study of the topology of messenger RNA bound to the 30S ribosomal subunit of Escherichia coli. Biochemistry *30*, 4821-4830.

D'Agostino, I., Merritt, C., Chen, P. L., Seydoux, G., and Subramaniam, K. (2006). Translational repression restricts expression of the C. elegans Nanos homolog NOS-2 to the embryonic germline. Dev Biol *292*, 244-252.

Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A. M., Grove, C. A., Martinez, N. J., Sequerra, R., Doucette-Stamm, L., Reece-Hoyes, J. S., Hope, I. A.*, et al.* (2006). A gene-centered C. elegans protein-DNA interaction network. Cell *125*, 1193-1205.

DeRenzo, C., Reese, K. J., and Seydoux, G. (2003). Exclusion of germ plasm proteins from somatic lineages by cullin-dependent degradation. Nature *424*, 685-689.

Donnini, M., Lapucci, A., Papucci, L., Witort, E., Jacquier, A., Brewer, G., Nicolin, A., Capaccioli, S., and Schiavone, N. (2004). Identification of TINO: a new evolutionarily conserved BCL-2 AU-rich element RNA-binding protein. J Biol Chem *279*, 20154-20166.

Draper, B. W., Mello, C. C., Bowerman, B., Hardin, J., and Priess, J. R. (1996). MEX-3 is a KH domain protein that regulates blastomere identity in early C. elegans embryos. Cell *87*, 205-216.

Dsouza, M., Larsen, N., and Overbeek, R. (1997). Searching for patterns in genomic data. Trends Genet *13*, 497-498.

Ellington, A. D., and Szostak, J. W. (1990). In vitro selection of RNA molecules that bind specific ligands. Nature *346*, 818-822.

Etemad-Moghadam, B., Guo, S., and Kemphues, K. J. (1995). Asymmetrically distributed PAR-3 protein contributes to cell polarity and spindle alignment in early C. elegans embryos. Cell *83*, 743-752.

Evans, T. C., Crittenden, S. L., Kodoyianni, V., and Kimble, J. (1994). Translational control of maternal glp-1 mRNA establishes an asymmetry in the C. elegans embryo. Cell *77*, 183-194.

Farley, B. M., Pagano, J. M., and Ryder, S. P. (2008). RNA target specificity of the embryonic cell fate determinant POS-1. Rna *14*, 2685-2697.

Farley, B. M., and Ryder, S. P. (2008). Regulation of maternal mRNAs in early development. Crit Rev Biochem Mol Biol *43*, 135-162.

Fong, H., Hohenstein, K. A., and Donovan, P. J. (2008). Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. Stem Cells *26*, 1931-1938.

Forbes, A., and Lehmann, R. (1998). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. Development *125*, 679-690.

Francis, R., Barton, M. K., Kimble, J., and Schedl, T. (1995). gld-1, a tumor suppressor gene required for oocyte development in Caenorhabditis elegans. Genetics *139*, 579-606.

Frokjaer-Jensen, C., Davis, M. W., Hopkins, C. E., Newman, B. J., Thummel, J. M., Olesen, S. P., Grunnet, M., and Jorgensen, E. M. (2008). Single-copy insertion of transgenes in Caenorhabditis elegans. Nat Genet *40*, 1375-1383.

Gallo, C. M., Munro, E., Rasoloson, D., Merritt, C., and Seydoux, G. (2008). Processing bodies and germ granules are distinct RNA granules that interact in C. elegans embryos. Dev Biol *323*, 76-87.

Gamarnik, A. V., and Andino, R. (2000). Interactions of viral protein 3CD and poly(rC) binding protein with the 5' untranslated region of the poliovirus genome. J Virol *74*, 2219-2226.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res *31*, 3784-3788.

Goldstein, B., and Hird, S. N. (1996). Specification of the anteroposterior axis in Caenorhabditis elegans. Development *122*, 1467-1474.

Gomes, J. E., Encalada, S. E., Swan, K. A., Shelton, C. A., Carter, J. C., and Bowerman, B. (2001). The maternal gene spn-4 encodes a predicted RRM protein required for mitotic spindle orientation and cell fate patterning in early C. elegans embryos. Development *128*, 4301-4314.

Guo, S., and Kemphues, K. J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell *81*, 611-620.

Guo, S., and Kemphues, K. J. (1996). A non-muscle myosin required for embryonic polarity in Caenorhabditis elegans. Nature *382*, 455-458.

Hall, T. M. (2005). Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol *15*, 367-373.

Hellman, L. M., and Fried, M. G. (2007). Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. Nat Protoc *2*, 1849-1861.

Hill, A. V. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its oxygen dissociation curve. J Physiol (London) *40*, 4-7.

Hori, T., Taguchi, Y., Uesugi, S., and Kurihara, Y. (2005). The RNA ligands for mouse proline-rich RNA-binding protein (mouse Prrp) contain two consensus sequences in separate loop structure. Nucleic Acids Res *33*, 190-200.

Howe, P. W., Nagai, K., Neuhaus, D., and Varani, G. (1994). NMR studies of U1 snRNA recognition by the N-terminal RNP domain of the human U1A protein. Embo J *13*, 3873-3881.

Huang, N. N., Mootz, D. E., Walhout, A. J., Vidal, M., and Hunter, C. P. (2002). MEX-3 interacting proteins link cell polarity to asymmetric gene expression in Caenorhabditis elegans. Development *129*, 747-759.

Hudson, B. P., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2004). Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. Nat Struct Mol Biol *11*, 257-264.

Hung, T. J., and Kemphues, K. J. (1999). PAR-6 is a conserved PDZ domain-containing protein that colocalizes with PAR-3 in Caenorhabditis elegans embryos. Development *126*, 127-135.

Hunter, C. P., and Kenyon, C. (1996). Spatial and temporal controls target pal-1 blastomere-specification activity to a single blastomere lineage in C. elegans embryos. Cell *87*, 217-226.

Hunter, T. C. E. a. C. P. (2005). Translational Control of Maternal RNAs, In WormBook, ed (The C. elegans Research Community).

Hutter, H., and Schnabel, R. (1994). glp-1 and inductions establishing embryonic axes in C. elegans. Development *120*, 2051-2064.

Hutter, H., and Schnabel, R. (1995). Specification of anterior-posterior differences within the AB lineage in the C. elegans embryo: a polarising induction. Development *121*, 1559-1568.

Iwasaki, S., Kawamata, T., and Tomari, Y. (2009). Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression. Mol Cell *34*, 58-67.

Jadhav, S., Rana, M., and Subramaniam, K. (2008). Multiple maternal proteins coordinate to restrict the translation of C. elegans nanos-2 to primordial germ cells. Development *135*, 1803-1812.

Johnstone, O., and Lasko, P. (2001). Translational regulation and RNA localization in Drosophila oocytes and embryos. Annu Rev Genet *35*, 365-406.

Jones, A. R., Francis, R., and Schedl, T. (1996). GLD-1, a cytoplasmic protein essential for oocyte differentiation, shows stage- and sex-specific expression during Caenorhabditis elegans germline development. Dev Biol *180*, 165-183.

Jud, M., Razelun, J., Bickel, J., Czerwinski, M., and Schisa, J. A. (2007). Conservation of large foci formation in arrested oocytes of Caenorhabditis nematodes. Dev Genes Evol.

Jud, M. C., Czerwinski, M. J., Wood, M. P., Young, R. A., Gallo, C. M., Bickel, J. S., Petty, E. L., Mason, J. M., Little, B. A., Padilla, P. A., and Schisa, J. A. (2008). Large P body-like RNPs form in C. elegans oocytes in response to arrested ovulation, heat shock, osmotic stress, and anoxia and are regulated by the major sperm protein pathway. Dev Biol *318*, 38-51.

Jurica, M. S., and Moore, M. J. (2003). Pre-mRNA splicing: awash in a sea of proteins. Mol Cell *12*, 5-14.

Kane, D. A., and Kimmel, C. B. (1993). The zebrafish midblastula transition. Development *119*, 447-456.

Kemphues, K. J., Priess, J. R., Morton, D. G., and Cheng, N. S. (1988). Identification of genes required for cytoplasmic localization in early C. elegans embryos. Cell *52*, 311-320.

Kim, J. L., Nikolov, D. B., and Burley, S. K. (1993). Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature *365*, 520-527.

Kirby, C., Kusch, M., and Kemphues, K. (1990). Mutations in the par genes of Caenorhabditis elegans affect cytoplasmic reorganization during the first cell cycle. Dev Biol *142*, 203-215.

Kuersten, S., and Goodwin, E. B. (2003). The power of the 3' UTR: translational control and development. Nat Rev Genet *4*, 626-637.

Lai, W. S., and Blackshear, P. J. (2001). Interactions of CCCH zinc finger proteins with mRNA: tristetraprolin-mediated AU-rich element-dependent mRNA degradation can occur in the absence of a poly(A) tail. J Biol Chem *276*, 23144-23154.

Lai, W. S., Carballo, E., Strum, J. R., Kennington, E. A., Phillips, R. S., and Blackshear, P. J. (1999). Evidence that tristetraprolin binds to AU-rich elements and promotes the deadenylation and destabilization of tumor necrosis factor alpha mRNA. Mol Cell Biol *19*, 4311-4323.

Lai, W. S., Carballo, E., Thorn, J. M., Kennington, E. A., and Blackshear, P. J. (2000). Interactions of CCCH zinc finger proteins with mRNA. Binding of tristetraprolin-related zinc finger proteins to Au-rich elements and destabilization of mRNA. J Biol Chem *275*, 17827-17837.

Lai, W. S., Parker, J. S., Grissom, S. F., Stumpo, D. J., and Blackshear, P. J. (2006). Novel mRNA targets for tristetraprolin (TTP) identified by global analysis of stabilized transcripts in TTP-deficient fibroblasts. Mol Cell Biol *26*, 9196-9208.

Lei, H., Liu, J., Fukushige, T., Fire, A., and Krause, M. (2009). Caudal-like PAL-1 directly activates the bodywall muscle module regulator hlh-1 in C. elegans to initiate the embryonic muscle gene regulatory network. Development *136*, 1241-1249.

Levitan, D. J., Boyd, L., Mello, C. C., Kemphues, K. J., and Stinchcomb, D. T. (1994). par-2, a gene required for blastomere asymmetry in Caenorhabditis elegans, encodes zinc-finger and ATP-binding motifs. Proc Natl Acad Sci U S A *91*, 6108-6112.

Lewis, H. A., Musunuru, K., Jensen, K. B., Edo, C., Chen, H., Darnell, R. B., and Burley, S. K. (2000). Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. Cell *100*, 323-332.

Lin, S. Y., and Riggs, A. D. (1972). Lac repressor binding to non-operator DNA: detailed studies and a comparison of eequilibrium and rate competition methods. J Mol Biol *72*, 671-690.

Loughlin, F. E., Mansfield, R. E., Vaz, P. M., McGrath, A. P., Setiyaputra, S., Gamsjaeger, R., Chen, E. S., Morris, B. J., Guss, J. M., and Mackay, J. P. (2009). The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5' splice site-like sequences. Proc Natl Acad Sci U S A *106*, 5581-5586.

Lublin, A. L., and Evans, T. C. (2007). The RNA-binding proteins PUF-5, PUF-6, and PUF-7 reveal multiple systems for maternal mRNA regulation during C. elegans oogenesis. Dev Biol *303*, 635-649.

Maduro, M. F., Hill, R. J., Heid, P. J., Newman-Smith, E. D., Zhu, J., Priess, J. R., and Rothman, J. H. (2005). Genetic redundancy in endoderm specification within the genus Caenorhabditis. Dev Biol *284*, 509-522.

Makeyev, A. V., and Liebhaber, S. A. (2002). The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. Rna *8*, 265-278.

Mango, S. E., Lambie, E. J., and Kimble, J. (1994). The pha-4 gene is required to generate the pharyngeal primordium of Caenorhabditis elegans. Development *120*, 3019-3031.

McClarin, J. A., Frederick, C. A., Wang, B. C., Greene, P., Boyer, H. W., Grable, J., and Rosenberg, J. M. (1986). Structure of the DNA-Eco RI endonuclease recognition complex at 3 A resolution. Science *234*, 1526-1541.

Mello, C. C., Draper, B. W., Krause, M., Weintraub, H., and Priess, J. R. (1992). The pie-1 and mex-1 genes and maternal control of blastomere identity in early C. elegans embryos. Cell *70*, 163-176.

Mello, C. C., Draper, B. W., and Priess, J. R. (1994). The maternal genes apx-1 and glp-1 and establishment of dorsal-ventral polarity in the early C. elegans embryo. Cell *77*, 95-106.

Mello, C. C., Schubert, C., Draper, B., Zhang, W., Lobel, R., and Priess, J. R. (1996). The PIE-1 protein and germline specification in C. elegans embryos. Nature *382*, 710-712.

Merritt, C., Rasoloson, D., Ko, D., and Seydoux, G. (2008). 3' UTRs are the primary regulators of gene expression in the C. elegans germline. Curr Biol *18*, 1476-1482.

Mickey, K. M., Mello, C. C., Montgomery, M. K., Fire, A., and Priess, J. R. (1996). An inductive interaction in 4-cell stage C. elegans embryos involves APX-1 expression in the signalling cell. Development *122*, 1791-1798.

Mootz, D., Ho, D. M., and Hunter, C. P. (2004). The STAR/Maxi-KH domain protein GLD-1 mediates a developmental switch in the translational control of C. elegans PAL-1. Development *131*, 3263-3272.

Morton, D. G., Shakes, D. C., Nugent, S., Dichoso, D., Wang, W., Golden, A., and Kemphues, K. J. (2002). The Caenorhabditis elegans par-5 gene encodes a 14-3-3 protein required for cellular asymmetry in the early embryo. Dev Biol *241*, 47-58.

Munro, E., Nance, J., and Priess, J. R. (2004). Cortical flows powered by asymmetrical contraction transport PAR proteins to establish and maintain anterior-posterior polarity in the early C. elegans embryo. Dev Cell *7*, 413-424.

Nance, J. (2005). PAR proteins and the establishment of cell polarity during C. elegans development. Bioessays *27*, 126-135.

Nilsen, T. W. (2002). The spliceosome: no assembly required? Mol Cell *9*, 8-9.

Noyes, M. B., Christensen, R. G., Wakabayashi, A., Stormo, G. D., Brodsky, M. H., and Wolfe, S. A. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. Cell *133*, 1277-1289.

Ogura, K., Kishimoto, N., Mitani, S., Gengyo-Ando, K., and Kohara, Y. (2003). Translational control of maternal glp-1 mRNA by POS-1 and its interacting protein SPN-4 in Caenorhabditis elegans. Development *130*, 2495-2503.

Pagano, J. M., Farley, B. M., Essien, K. I., and Ryder, S. P. (2009). RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. Proc Natl Acad Sci U S A *106*, 20252-20257.

Pagano, J. M., Farley, B. M., McCoig, L. M., and Ryder, S. P. (2007). Molecular Basis of RNA Recognition by the Embryonic Polarity Determinant MEX-5. J Biol Chem *282*, 8883-8894.

Pellettieri, J., and Seydoux, G. (2002). Anterior-posterior polarity in C. elegans and Drosophila--PARallels and differences. Science *298*, 1946-1950.

Pichler, S., Gonczy, P., Schnabel, H., Pozniakowski, A., Ashford, A., Schnabel, R., and Hyman, A. A. (2000). OOC-3, a novel putative transmembrane protein required for establishment of cortical domains and spindle orientation in the P(1) blastomere of C. elegans embryos. Development *127*, 2063-2073.

Pitt, J. N., Schisa, J. A., and Priess, J. R. (2000). P granules in the germ cells of Caenorhabditis elegans adults are associated with clusters of nuclear pores and contain RNA. Dev Biol *219*, 315-333.

Praitis, V. (2006). Creation of transgenic lines using microparticle bombardment methods. Methods Mol Biol *351*, 93-107.

Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. Genetics *157*, 1217-1226.

Priess, J. R., and Thomson, J. N. (1987). Cellular interactions in early C. elegans embryos. Cell *48*, 241-250.

Rambo, R. P., and Doudna, J. A. (2004). Assembly of an active group II intron-maturase complex by protein dimerization. Biochemistry *43*, 6486-6497.

Reese, K. J., Dunn, M. A., Waddle, J. A., and Seydoux, G. (2000). Asymmetric segregation of PIE-1 in C. elegans is mediated by two complementary mechanisms that act through separate PIE-1 protein domains. Mol Cell *6*, 445-455.

Reines, S. A., and Cantor, C. R. (1974). New fluorescent hydrazide reagents for the oxidized 3'-terminus of RNA. Nucleic Acids Res *1*, 767-786.

Reinke, V., Gil, I. S., Ward, S., and Kazmer, K. (2004). Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development *131*, 311-323.

Rose, L. S., and Kemphues, K. J. (1998). Early patterning of the C. elegans embryo. Annu Rev Genet *32*, 521-545.

Roy Chowdhuri, S., Crum, T., Woollard, A., Aslam, S., and Okkema, P. G. (2006). The T-box factor TBX-2 and the SUMO conjugating enzyme UBC-9 are required for ABa-derived pharyngeal muscle in C. elegans. Dev Biol *295*, 664-677.

Royer, C. A., and Scarlata, S. F. (2008). Fluorescence approaches to quantifying biomolecular interactions. Methods Enzymol *450*, 79-106.

Ryder, S. P., Frater, L. A., Abramovitz, D. L., Goodwin, E. B., and Williamson, J. R. (2004). RNA target specificity of the STAR/GSG domain post-transcriptional regulatory protein GLD-1. Nat Struct Mol Biol *11*, 20-28.

Ryder, S. P., Recht, M. I., and Williamson, J. R. (2008). Quantitative analysis of protein-RNA interactions by gel mobility shift. Methods Mol Biol *488*, 99-115.

Ryder, S. P., and Williamson, J. R. (2004). Specificity of the STAR/GSG domain protein Qk1: implications for the regulation of myelination. Rna *10*, 1449-1458.

Sadler, P. L., and Shakes, D. C. (2000). Anucleate Caenorhabditis elegans sperm can crawl, fertilize oocytes and direct anterior-posterior polarization of the 1-cell embryo. Development *127*, 355-366.

Schubert, C. M., Lin, R., de Vries, C. J., Plasterk, R. H., and Priess, J. R. (2000). MEX-5 and MEX-6 function to establish soma/germline asymmetry in early C. elegans embryos. Mol Cell *5*, 671-682.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003). SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res *31*, 3381-3385.

Seeman, N. C., Rosenberg, J. M., and Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. Proc Natl Acad Sci U S A *73*, 804-808.

Seydoux, G., and Braun, R. E. (2006). Pathway to totipotency: lessons from germ cells. Cell *127*, 891-904.

Singh, R., and Valcarcel, J. (2005). Building specificity with nonspecific RNA-binding proteins. Nat Struct Mol Biol *12*, 645-653.

Sonnichsen, B., Koski, L. B., Walsh, A., Marschall, P., Neumann, B., Brehm, M., Alleaume, A. M., Artelt, J., Bettencourt, P., Cassin, E.*, et al.* (2005). Full-genome RNAi profiling of early embryogenesis in Caenorhabditis elegans. Nature *434*, 462-469.

Stumpf, C. R., Kimble, J., and Wickens, M. (2008). A Caenorhabditis elegans PUF protein family with distinct RNA binding specificity. Rna *14*, 1550-1557.

Subramaniam, K., and Seydoux, G. (1999). nos-1 and nos-2, two genes related to Drosophila nanos, regulate primordial germ cell development and survival in Caenorhabditis elegans. Development *126*, 4861-4871.

Sulston, J. E., and Horvitz, H. R. (1977). Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev Biol *56*, 110-156.

Sulston, J. E., Schierenberg, E., White, J. G., and Thomson, J. N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol *100*, 64-119.

Tabara, H., Hill, R. J., Mello, C. C., Priess, J. R., and Kohara, Y. (1999). pos-1 encodes a cytoplasmic zinc-finger protein essential for germline specification in C. elegans. Development *126*, 1-11.

Tabuse, Y., Izumi, Y., Piano, F., Kemphues, K. J., Miwa, J., and Ohno, S. (1998). Atypical protein kinase C cooperates with PAR-3 to establish embryonic polarity in Caenorhabditis elegans. Development *125*, 3607-3614.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Taylor, G. A., Carballo, E., Lee, D. M., Lai, W. S., Thompson, M. J., Patel, D. D., Schenkman, D. I., Gilkeson, G. S., Broxmeyer, H. E., Haynes, B. F., and

Blackshear, P. J. (1996). A pathogenetic role for TNF alpha in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraprolin (TTP) deficiency. Immunity *4*, 445-454.

Tenlen, J. R., Molk, J. N., London, N., Page, B. D., and Priess, J. R. (2008). MEX-5 asymmetry in one-cell C. elegans embryos requires PAR-4- and PAR-1-dependent phosphorylation. Development *135*, 3665-3675.

Tenlen, J. R., Schisa, J. A., Diede, S. J., and Page, B. D. (2006). Reduced dosage of pos-1 suppresses Mex mutants and reveals complex interactions among CCCH zinc-finger proteins during Caenorhabditis elegans embryogenesis. Genetics *174*, 1933-1945.

Thorpe, C. J., Schlesinger, A., Carter, J. C., and Bowerman, B. (1997). Wnt signaling polarizes an early C. elegans blastomere to distinguish endoderm from mesoderm. Cell *90*, 695-705.

Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science *249*, 505-510.

Valegard, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein-operator complex. Nature *371*, 623-626.

Valverde, R., Edwards, L., and Regan, L. (2008). Structure and function of KH domains. Febs J *275*, 2712-2726.

Varnum, B. C., Ma, Q. F., Chi, T. H., Fletcher, B., and Herschman, H. R. (1991). The TIS11 primary response gene is a member of a gene family that encodes proteins with a highly conserved sequence containing an unusual Cys-His repeat. Mol Cell Biol *11*, 1754-1758.

Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (2000). GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. Methods Enzymol *328*, 575-592.

Wallenfang, M. R., and Seydoux, G. (2000). Polarization of the anterior-posterior axis of C. elegans is a microtubule-directed process. Nature *408*, 89-92.

Watts, J. L., Morton, D. G., Bestman, J., and Kemphues, K. J. (2000). The C. elegans par-4 gene encodes a putative serine-threonine kinase required for establishing embryonic asymmetry. Development *127*, 1467-1475.

Weeks, K. M., and Crothers, D. M. (1992). RNA binding assays for Tat-derived peptides: implications for specificity. Biochemistry *31*, 10281-10287.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Wong, I., and Lohman, T. M. (1993). A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. Proc Natl Acad Sci U S A *90*, 5428-5432.

Xu, L., Paulsen, J., Yoo, Y., Goodwin, E. B., and Strome, S. (2001). Caenorhabditis elegans MES-3 is a target of GLD-1 and functions epigenetically in germline development. Genetics *159*, 1007-1017.