

University of Massachusetts Medical School

eScholarship@UMMS

GSBS Dissertations and Theses

Graduate School of Biomedical Sciences

2012-08-09

Sequence and Target Specificity of the *C. elegans* Cell Fate Specification Factor POS-1: A Dissertation

Brian M. Farley

University of Massachusetts Medical School

Let us know how access to this document benefits you.

Follow this and additional works at: https://escholarship.umassmed.edu/gsbs_diss



Part of the Amino Acids, Peptides, and Proteins Commons, Animal Experimentation and Research Commons, Biochemistry, Biophysics, and Structural Biology Commons, Cell and Developmental Biology Commons, Cells Commons, Embryonic Structures Commons, and the Nucleic Acids, Nucleotides, and Nucleosides Commons

Repository Citation

Farley BM. (2012). Sequence and Target Specificity of the *C. elegans* Cell Fate Specification Factor POS-1: A Dissertation. GSBS Dissertations and Theses. <https://doi.org/10.13028/2k07-xr46>. Retrieved from https://escholarship.umassmed.edu/gsbs_diss/629

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.

SEQUENCE AND TARGET SPECIFICITY OF THE *C. ELEGANS* CELL FATE
SPECIFICATION FACTOR POS-1

A Dissertation Presented

By

BRIAN MATTHEW FARLEY

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

AUGUST 9, 2012

BIOCHEMISTRY AND MOLECULAR PHARMACOLOGY

SEQUENCE AND TARGET SPECIFICITY OF THE *C. ELEGANS* CELL FATE
SPECIFICATION FACTOR POS-1

A Dissertation Presented
By

Brian Matthew Farley

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

Sean P. Ryder, Ph.D., Thesis Advisor

Craig C. Mello, Ph.D., Member of Committee

Melissa J. Moore, Ph.D., Member of Committee

Scot A. Wolfe, Ph.D., Member of Committee

Traci M. T. Hall, 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

Phillip D. Zamore, 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

Biochemistry and Molecular Pharmacology

August 9th, 2012

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my thesis advisor, Sean Ryder. Without his training and guidance, I would not be the researcher I am today.

I would also like to thank the members of my laboratory for their assistance throughout my doctoral work.

ABSTRACT

In most metazoans, early embryogenesis is controlled by the translational regulation of maternally supplied mRNA. Sequence-specific RNA-binding proteins play an important role in regulating early embryogenesis, yet their specificities and regulatory targets are largely unknown. To understand how these RNA-binding proteins select their targets, my research focused on the *C. elegans* CCCH-type tandem zinc finger protein POS-1. Embryos lacking maternally supplied POS-1 die prior to gastrulation, and exhibit defects in the specification of pharyngeal, intestinal, and germline precursor cells. To identify the regulatory targets that contribute to the POS-1 mutant phenotype, we set out to determine the sequence specificity of POS-1 *in vitro*, and then use this information to identify regulatory targets *in vivo*.

Using a candidate-based search, we identified a twelve-nucleotide fragment of the *mex-3* 3' untranslated region (3' UTR) to which POS-1 binds with high affinity. Using quantitative fluorescent electrophoretic mobility shift assays, I determined the affinity of the RNA-binding domain of POS-1 for a panel of single nucleotide mutations of this sequence, and then defined a consensus binding element based on this dataset. POS-1 recognizes the degenerate element $UAU_{2-3}RDN_{1-3}G$, where R is any purine (adenosine or guanine), and D is any base except cytosine. A bioinformatics analysis revealed the presence of this element in approximately 40% of *C. elegans* 3' UTRs, suggesting that POS-1 is capable of binding to and perhaps regulating many transcripts *in vivo*. POS-1 binding sites alone are not sufficient to pattern the expression of a reporter, suggesting that other factors may contribute to POS-1 specificity.

To address the mechanism of POS-1-mediated translational regulation, I investigated the translational regulation of the *C. elegans* Notch homolog *glp-1*. Previous work demonstrated that *glp-1* translation is repressed in the early embryo in a POS-1-dependent fashion, though it was not clear if this regulation was direct. The *glp-1* 3' UTR contains two POS-1 binding sites within five nucleotides of each other, and these sites are within a thirty-nucleotide region of the 3' UTR required for proper spatiotemporal translation of *glp-1*. The POS-1 sites overlap with a negative regulatory element that is recognized by GLD-1, and a positive regulatory element recognized by an unknown factor. Both POS-1 and GLD-1 bind to an RNA containing these sites *in vitro*, and POS-1 competes with GLD-1 for binding. Both proteins are required for translational repression of a *glp-1* 3' UTR reporter in embryos. Furthermore, only one of the two POS-1 binding sites is required for repression, and the required site is wholly contained within a previously characterized positive regulatory element. Based on this, we propose that POS-1 does not regulate its targets by recruiting regulatory machinery, but instead by competing with factors that do. Thus, sites of POS-1 regulation are highly context dependent, which may contribute to POS-1 specificity.

TABLE OF CONTENTS

| | |
|---|------|
| TITLE PAGE | i |
| SIGNATURE PAGE | ii |
| ACKNOWLEDGEMENTS | iii |
| ABSTRACT | iv |
| TABLE OF CONTENTS | vi |
| LIST OF FIGURES | viii |
| LIST OF TABLES | ix |
| PREFACE | x |
| | |
| CHAPTER I: Sequence specific RNA-binding proteins in <i>C. elegans</i> development | 1 |
| <i>C. elegans</i> hermaphrodite germline physiology and embryogenesis | 8 |
| Function and regulation of <i>glp-1</i> | 20 |
| Sequence specific RNA-binding proteins in <i>C. elegans</i> development | 31 |
| | |
| CHAPTER II: RNA target specificity of the embryonic cell fate determinant POS-1 | 41 |
| ABSTRACT | 42 |
| INTRODUCTION | 43 |
| RESULTS | 48 |
| POS-1 binds weakly to TTP and MEX-5 binding sites | 48 |
| POS-1 binds to the 3' UTR of multiple genes required for embryogenesis | 51 |
| POS-1 forms an equimolar complex with the <i>mex-3</i> fragment | 57 |
| POS-1 recognizes a twelve nucleotide fragment within the <i>mex-3</i> UTR | 58 |
| Determination of the POS-1 consensus sequence | 61 |
| The PRE is necessary but not sufficient to pattern the expression of a reporter | 65 |
| Prevalence of the PRE in <i>C. elegans</i> 3' UTRs | 68 |
| DISCUSSION | 72 |
| SUPPLEMENTAL DISCUSSION | 75 |
| MATERIALS AND METHODS | 77 |
| | |
| CHAPTER III: POS-1 represses translation of <i>glp-1</i> by antagonizing multiple RNA-binding proteins | 84 |
| ABSTRACT | 85 |
| INTRODUCTION | 86 |
| RESULTS | 94 |
| Identification of a second PRE in the <i>glp-1</i> 3' UTR | 94 |
| POS-1 and GLD-1 recognize the <i>glp-1</i> 3' UTR in a sequence-specific manner | 95 |
| The PREs are equivalent and independent | 99 |
| POS-1 antagonizes GLD-1 binding to the <i>glp-1</i> 3' UTR | 100 |
| Mutations in the <i>glp-1</i> 3' UTR are specific for either POS-1 or GLD-1 | 103 |
| Both POS-1 and GLD-1 binding are required for repression of <i>glp-1</i> in embryos | 108 |
| The 3' PRE mutations disrupt the GDE in the germline | 113 |
| SPN-4 does not directly activate <i>glp-1</i> translation in the germline | 114 |
| DISCUSSION | 120 |

| | |
|---|-----|
| MATERIALS AND METHODS | 123 |
| ACKNOWLEDGEMENTS | 128 |
| CHAPTER IV: The ramifications of promiscuous RNA binding specificity | 129 |
| POS-1 specificity | 131 |
| Translational regulation of <i>glp-1</i> | 135 |
| Genome-wide identification of functional regulatory elements | 140 |
| BIBLIOGRAPHY | 143 |

LIST OF FIGURES

| | |
|--|-----|
| Figure 1.1: The <i>C. elegans</i> gonad and early embryo | 9 |
| Figure 1.2: The cell lineage of the early <i>C. elegans</i> embryo | 14 |
| Figure 1.3: RNA-binding proteins involved in <i>C. elegans</i> gametogenesis and Embryogenesis | 17 |
| Figure 1.4: Translational repression of <i>glp-1</i> | 23 |
| Figure 1.5: Regulatory elements in the <i>glp-1</i> 3' UTR | 26 |
| Figure 2.1: POS-1 is required for the specification of multiple cell fates in the early <i>C. elegans</i> embryo | 44 |
| Figure 2.2: POS-1 is a specific RNA-binding protein with a different specificity than Either TTP or MEX-5 | 49 |
| Figure 2.3: POS-1 binds to the <i>mex-3</i> UTR fragment with apparent 1:1 stoichiometry | 55 |
| Figure 2.4: POS-1 specifically recognizes a twelve nucleotide fragment of the <i>mex-3</i> 3' UTR | 59 |
| Figure 2.5: POS-1 recognizes the sequence UA(U ₂₋₃)RD(N ₁₋₃)G | 63 |
| Figure 2.6: The PRE is necessary but not sufficient for patterned regulation of a reporter | 67 |
| Figure 2.7: The PRE is abundant in <i>C. elegans</i> 3' UTRs | 71 |
| Figure 3.1: POS-1 and GLD-1 expression is anti-correlated with GLP-1 expression | 90 |
| Figure 3.2: POS-1 and GLD-1 binding sites lie within regulatory elements of the <i>glp-1</i> 3' UTR | 93 |
| Figure 3.3: POS-1 and GLD-1 recognize the <i>glp-1</i> 3' UTR in a sequence-specific manner | 98 |
| Figure 3.4: POS-1 antagonizes GLD-1 binding to the <i>glp-1</i> fragment | 102 |
| Figure 3.5: POS-1 RBD binds to the <i>glp-1</i> fragment with a similar affinity to that of POS-1 1-206 | 105 |
| Figure 3.6: POS-1 or GLD-1 specific mutations in the <i>glp-1</i> 3' UTR | 107 |
| Figure 3.7: POS-1 and GLD-1 binding are independently required to repress a <i>glp-1</i> 3' UTR reporter in embryos | 110 |
| Figure 3.8: Mutation of the 3' PRE decreases <i>glp-1</i> reporter expression independently of POS-1 | 116 |
| Figure 3.9: SPN-4 activates <i>glp-1</i> translation indirectly | 119 |
| Figure 4.1: A cluster of binding sites is present in the <i>glp-1</i> 3' UTR | 139 |

LIST OF TABLES

| | |
|---|-----|
| Table 2.1: Electrophoretic mobility shift assay with <i>glp-1</i> 3' UTR fragments | 52 |
| Table 2.2: Electrophoretic mobility shift assay with maternal mRNA fragments | 54 |
| Table 3.1: Dissociation constants of POS-1 and GLD-1 for variants of the <i>glp-1</i> fragment | 97 |
| Table 3.2: Transgenic worm strains used in this study | 112 |

PREFACE

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

Chapter II is adapted from:

Farley BM, Pagano JM, Ryder SP. 2008. RNA target specificity of the embryonic cell fate determinant POS-1. *RNA* **14**: 2685-2697.

Chapter III is adapted from the submitted manuscript:

Farley BM and Ryder SP. POS-1 represses *glp-1* translation by antagonizing multiple RNA-binding proteins. Submitted.

John M. Pagano generated the transgenic reporter strain used in Figure 2.6.

Work from the following publications was not included in this dissertation:

Pagano JM, Farley BM, McCoig LM, Ryder SP. 2007. Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. *J Biol Chem* 282(12).

Farley BM, Ryder SP. 2008. Regulation of maternal mRNAs in early development. *Crit Rev Biochem Mol Biol* 43(2).

Zearfoss NR, Farley BM, Ryder SP. 2008. Post-transcriptional regulation of myelin formation. *Biochim Biophys Acta* 1779(8).

Pagano JM, Farley BM, Essien KI, Ryder SP. 2009. RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. *Proc Natl Acad Sci* 106(48).

Zearfoss NR, Clingman CC, Farley BM, McCoig LM, Ryder SP. 2011. Quaking regulates *Hnrnpa1* expression through its 3' UTR in oligodendrocyte precursor cells. *PLoS Genet* 7(1).

Wright JE, Gaidatzis D, Senften M, Farley BM, Westhof E, Ryder SP, Ciosk R. 2011. A quantitative RNA code for mRNA target selection by the germline fate determinant GLD-1. *EMBO J* 30(3).

Kalchhauser I, Farley BM, Pauli S, Ryder SP, Ciosk R. 2011. FBF represses the Cip/Kip cell-cycle inhibitor CKI-2 to promote self-renewal of germline stem cells in *C. elegans*. *EMBO J* 30(18).

Chapter I

Sequence specific RNA-binding proteins in

C. elegans development

Prior to fertilization, oocytes remain arrested in meiosis, sometimes for a period of many years. Once fertilization takes place, the oocyte becomes active and a cascade of rapid events occurs. The maternal pronucleus completes meiosis and fuses with the paternal pronucleus, which triggers a series of cell divisions that dramatically increase the number of cells present in the embryo without increasing its volume. While these divisions are taking place, the rough body plan of the developing organism is established. Body axes are specified, the germline and soma are differentiated from one another, and the identities of the future endoderm, mesoderm, and ectoderm are established. What is remarkable about this process is that it occurs largely without the benefit of zygotic transcription. In sea urchins (*S. purpuratus*), nematodes (*C. elegans*), fruit flies (*D. melanogaster*), zebrafish (*D. rerio*), and african clawed frogs (*X. laevis*), zygotic transcription is completely inhibited for at least the first 1-2 cycles of division, and embryos can reach the hundred cell stage even when transcription is inhibited (reviewed in Tadros and Lipshitz 2009). Transcription is reduced or inhibited in the final stages of mouse oocyte development (Moore and Lintern-Moore 1978), and it reinitiates at roughly the first embryonic division (Moore 1975). Embryogenesis occurs at a slower pace in mice versus other metazoans, so transcription is inhibited for twenty hours or more after fertilization. Thus, the earliest events in the development of most multicellular organisms is driven by translational regulation of maternally supplied mRNAs (reviewed in Farley and Ryder 2008).

The transcription of mRNAs encoding proteins required to carry out the developmental programs of early embryogenesis occurs during oogenesis, but the

activities of these proteins are not required until after fertilization. Thus, transcripts encoding such proteins must be translationally repressed within the developing oocyte, as well as stored for later use in the embryo. This phenomenon is particularly noticeable in the specification of germ cell fates in flies, worms, frogs, and zebrafish. During oogenesis, each of these species generates granular structures containing mRNAs and RNA-binding proteins called germ plasm. After fertilization, germ plasm localizes to a specific region of the embryo, and eventually segregates to only a few cells by a variety of mechanisms. The subset of cells that inherit germ plasm are fated to produce the entire germ line in the adult organism. This illustrates the role that the localization of maternally supplied factors plays in the specification of cell fates in the early embryo.

Outside of germ line specification, many maternally supplied transcripts exhibit localized expression patterns during embryogenesis. For example, during *Drosophila* embryogenesis, mRNAs transcribed from approximately 20% of fly genes exhibit restricted expression patterns. Lecuyer and colleagues examined the localization patterns of approximately 3400 transcripts (about 25% of the genome) in the syncytial *Drosophila melanogaster* embryo by high resolution fluorescent *in situ* hybridization. Of these, approximately 2300 were expressed in the embryo, and of the expressed transcripts, 71% (approximately 1600) exhibit a restricted localization pattern (Lécuyer et al. 2007). The patterns observed can be grouped into at least thirty-five distinct localization categories, with slight differences between members of each category. Extrapolating the proportion of localized transcripts observed in this study to the entire genome, it is expected that nearly 7000 transcripts have a distinct expression pattern in the embryo (Lécuyer et al.

2007). Both the diversity in localization patterns observed, and the number of localized transcripts point to a broad network of transcript-specific localization that is active in the *Drosophila* embryo, and likely other species as well.

Throughout the entire process of oogenesis and early embryogenesis, specific mRNAs must be selected for participation in the appropriate regulatory pathways. Only a subset of the mRNAs actively transcribed within developing germ cells are localized to cytoplasmic granules or are otherwise translationally inhibited. After fertilization, the timing and localization of translation of maternally supplied mRNAs is precisely controlled, and a diverse ensemble of translational outcomes is essential. How are mRNAs targeted to the appropriate regulatory machinery within developing oocytes and early embryos?

mRNAs are sorted into the appropriate regulatory pathways by virtue of elements contained within their primary sequence. Short motifs or structural elements that appear only in a subset of mRNAs are sufficient for segregating those mRNAs into different regulatory pathways. These elements are often, but not always, found upstream of the start codon and downstream of the stop codon in regions named the 5' and 3' untranslated regions (UTRs). Modular RNA-binding proteins recognize these elements through one or more RNA-binding domains. These proteins also typically contain one or more catalytic or regulatory domains. Thus, the RNA-binding protein serves to connect specific transcripts with the appropriate regulatory machinery (reviewed in Lunde et al. 2007).

Two distinct strategies for recognizing specific RNA sequences have evolved. One strategy involves RNA-binding proteins that associate with a small RNA, such as a

microRNA (miRNA) or a small interfering RNA (siRNA) (reviewed in Ghildiyal and Zamore 2009). These short RNAs base pair with the target sequence, and thus direct the RNA-binding protein to the target. As these proteins rely on base pairing between their associated small RNA and the target sequence, the sequence specificities of small-RNA associated proteins are at least partially understood (reviewed in Pasquinelli 2012). The other strategy employed for recognizing specific sequence elements involves RNA-binding proteins that make direct contact with the sequence element. There are many different types of RNA-binding domains, each with a different sequence specificity (reviewed in Glisovic et al. 2008). As these proteins do not derive their specificity from an associated small RNA, it is much more challenging to predict the specificity of any member of this class of proteins. Thus, our understanding of the sequence specificity of RNA-binding domains that directly contact their target RNAs is far from complete.

RNA-binding proteins are common in the genomes of eukaryotes. Bioinformatic predictions based on homology to previously identified RNA-binding domains estimate that somewhere between 2% and 3% of all proteins encoded by the genomes of *D. melanogaster* and *C. elegans* are capable of recognizing RNA (Lasko 2000; Lee and Schedl 2006). As these estimates only consider genes with sequences that are homologous to known RNA-binding domains, it is possible that it is an underestimate. A recent study performed by the Hentze lab identified the ensemble of proteins that interact with polyadenylated mRNA within HeLa cells (Castello et al. 2012). Many of the proteins identified have never been experimentally shown to interact with RNA, and 315 proteins that have no RNA-related functional annotation reproducibly associate with

poly(A) mRNA (Castello et al. 2012). Thus, there is a large network of transcripts that are regulated post-transcriptionally, and a large number of potential specificity factors that regulate those targets. One of the main challenges in the field of post-transcriptional regulation is mapping the specificity factors to their regulatory targets. This is especially challenging for sequence specific RNA-binding proteins, as they usually recognize a specificity element that is between 4 and 12 nucleotides long (Keene 2007) and is likely to be quite common in the transcriptome.

Multiple techniques have been developed to assess which sequences an RNA-binding protein recognizes *in vivo*. RIP-Chip (Keene et al. 2006), CLIP (Licatalosi et al. 2008), and PAR-CLIP (Hafner et al. 2010) all rely on specific immunoprecipitation of RNA-binding protein-mRNA complexes from tissues. Each of these techniques allows for a genome-scale identification of bound mRNAs in a model-unbiased manner. The antibodies can be raised against the endogenous protein, or an epitope-tagged version can be expressed as a transgene in the organism of interest. RIP-Chip is typically performed without chemical or ultraviolet crosslinking (Keene et al. 2006), which allows RNA-binding proteins to repartition in extract and potentially associate with mRNAs they never bind to in a biologically relevant context (Riley et al. 2012). To prevent repartitioning, two immunoprecipitation techniques that crosslink RBPs prior to tissue homogenization have been developed. Crosslinking and immunoprecipitation (CLIP) uses short wavelength UV radiation to crosslink RBPs to the mRNAs they interact with *in vivo* (Licatalosi et al. 2008). This permits the use of harsh pulldown and washing conditions, which allows for a significant reduction of non-specifically bound material

relative to RIP-Chip. However, crosslinking with short wavelength UV radiation is not very efficient, and CLIP suffers from low recovery efficiencies (Hafner et al. 2010). To address this issue, the Tuschl lab developed a technique called Photoactivatable-Ribonucleoside Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP). This method incorporates a photoactivatable crosslinker (usually 4-thiouracil) into the mRNA of the tissue to be used in the pulldown. This modification crosslinks efficiently using long wavelength UV radiation. In addition, after the crosslinks have been reversed, a predictable base conversion occurs, which permits the identification of locations within the recovered mRNA that the RBP was interacting with directly (Hafner et al. 2010). Modified ribonucleosides can be introduced into living organisms, such as *C. elegans*, which allows for crosslinking from living tissue (Jungkamp et al. 2011). All three of these techniques produce a wealth of information regarding what an RNA-binding protein binds to *in vivo*. However, binding of an RNA-binding protein to an mRNA may not always result in a regulatory event, so the set of all mRNAs associated with a given RBP may include some mRNAs that are not regulated by that RBP.

To study the problem of RNA-binding protein functional target specificity, a combination of *in vitro* and *in vivo* experimentation is essential. The gametogenesis and embryogenesis of the roundworm *Caenorhabditis elegans* provides an attractive model system in which to address these questions. The vast majority of animals are self-fertile hermaphrodites (reviewed in Riddle et al. 1997), with a well-defined switch from spermatocyte to oocyte production (Riddle et al. 1997). Gametogenesis occurs sequentially within the germline, allowing for all steps of gametogenesis to be observed

within one animal (Riddle et al. 1997). The generation time of *C. elegans* is short (at 20 °C, an entire life cycle is complete in approximately 3 days) (Riddle et al. 1997), permitting for rapid proliferation of worms and rapid progress through gametogenesis and development. Animals are transparent, permitting developmental events to be readily observed in intact animals. The expression of genes within the animal can be robustly knocked down via RNA interference (RNAi), allowing for the biological function of genes of interest to be conveniently studied (Fire et al. 1998). Transgenes can be introduced into the worm by a variety of methods (Stinchcomb et al. 1985; Mello et al. 1991; Kelly et al. 1997; Praitis et al. 2001; Frøkjaer-Jensen et al. 2008), which provides an opportunity to investigate post-transcriptional regulatory events via reporters. As described in more detail below, RNA-binding proteins play a central role in many of the events of development, ranging from the initial entry of mitotically dividing germ cell precursors into meiosis (reviewed in Kimble and Crittenden 2007) to cell fate specification events in the developing embryo (reviewed in Maduro 2010).

***C. elegans* hermaphrodite germline physiology and embryogenesis**

Each hermaphrodite animal contains two complete gonad arms that each begin with a small population of mitotically dividing germline stem cells and terminate at a shared uterus in the middle of the animal (Figure 1.1). The distal end contains the mitotically dividing cells (Hirsh et al. 1976), while the proximal end joins with the uterus.

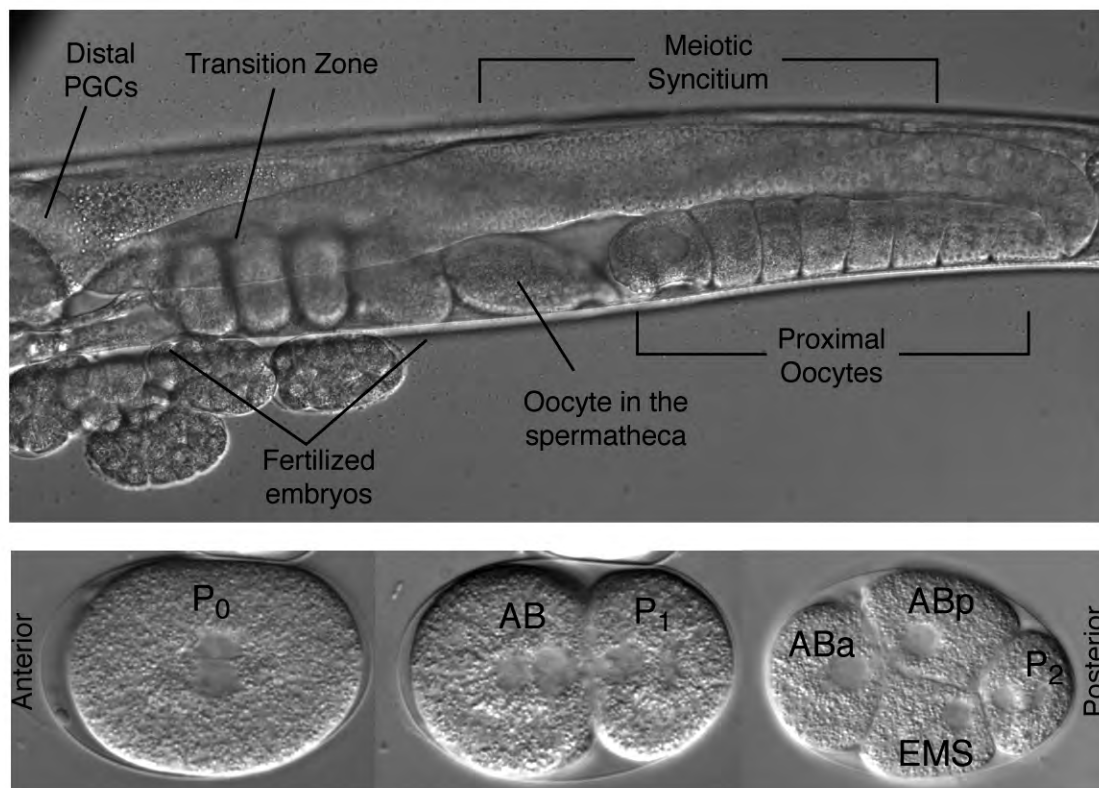


Figure 1.1

Figure 1.1. The *C. elegans* gonad and early embryo. Differential interference contrast (DIC) micrographs of the *C. elegans* gonad and early embryo with significant developmental events labeled. Labels on embryonic cells denote the identity of the cell.

The gonad arm is almost as long as the worm (Hirsh et al. 1976), and it folds over on itself to fit within the body of the worm. As germ cells develop into gametes, they travel towards the proximal end of the germline (Hirsh et al. 1976). A somatic cell with multiple long processes that run along the basement membrane of the gonad caps each gonad arm (Kimble and White 1981). This cell is called the distal tip cell which expresses the membrane-bound Delta ligand LAG-2 (Henderson et al. 1994) that instructs the cells it contacts to remain in mitosis. As germ cells migrate away from the distal tip cell, they transition into meiosis, migrate to the periphery of the gonad arm, and lose a portion of their membranes (Hirsh et al. 1976). This generates a cylindrical syncytium (Hirsh et al. 1976), where hundreds of nuclei all share the same cytoplasm. It is within the syncytial region of the gonad that synthesis of mRNAs and proteins required for early embryogenesis takes place; concordant with this, transcriptional activity increases significantly in this region (Gibert et al. 1984). Large ribonucleoprotein granules called P-granules arise in the syncytial region as well (Strome and Wood 1982; Jungkamp et al. 2011). These granules contain many RNA-binding proteins and mRNAs that are required for early embryogenesis, and are first located around the periphery of the germ cell nuclei (Licatalosi et al. 2008; Schisa et al. 2001; Bezares-Calderón et al. 2010). The cytoplasm of the syncytial region also contains RNP granules similar to P-bodies that contain a different complement of mRNAs and proteins relative to P granules (Hafner et al. 2010; Noble et al. 2008; Jungkamp et al. 2011). The nuclei in the syncytial

region arrest in prophase of meiosis I, and do not progress further in meiosis until just prior to fertilization (Henderson et al. 1994; McCarter et al. 1999; McNally and McNally 2005).

As the germ nuclei round the bend of the gonad arm and enter the proximal region, transcription is silenced and a small fraction of the nuclei recellularize and become oocytes. The germ cell nuclei that do not become oocytes are degraded via apoptosis (Gumienny et al. 1999). Cytoplasmic streaming from the syncytium deposits factors required for embryogenesis into the developing oocytes (Wolke et al. 2007; Nadarajan et al. 2009). The immature oocytes approach the spermatheca, an organ that contains the spermatocytes generated during the final larval stage of the worm (Ward et al. 1981). The spermatheca secretes a protein called Major Sperm Protein (MSP), which causes the resumption of meiosis (Miller et al. 2001). As the oocyte enters the spermatheca, meiosis I and II are completed, and the oocyte is fertilized (Ward and Carrel 1979).

Fertilization triggers a cascade of events that specify the anterior-posterior axis of the developing embryo. The sperm entry point determines the posterior pole of the embryo by disrupting an actin-myosin network, which causes the cortex of the fertilized embryo to be pulled towards the anterior (Munro et al. 2004). This allows for a number of proteins to associate with the posterior cortex, thus polarizing the embryo along the anterior-posterior axis (Etemad-Moghadam et al. 1995; Boyd et al. 1996). During polarization, the male and female pronuclei meet in the center of the embryo, fuse and the

single nucleus migrates towards the posterior of the embryo along with the P granules (Strome and Wood 1982; 1983). The first mitotic division takes place along the anterior-posterior axis approximately 45 minutes after fertilization, resulting in two unequally sized daughter cells. The larger cell is located in the anterior, and is the first of six founder cells that will be generated over the course of early embryogenesis (Sulston and Horvitz 1977; Kimble and Hirsh 1979; Sulston et al. 1983). Each of these cells is fated to produce only a limited subset of the tissue types present in the adult worm, and their fates are specified shortly after they are born (Sulston et al. 1983). The anterior cell is called AB and will produce pharyngeal and hypodermal tissues (Figure 1.2; (Sulston et al. 1983). The smaller posterior cell is termed P₁, and contains all of the P-granules deposited in the one cell embryo (Sulston et al. 1983). This cell will divide asymmetrically three more times, each time producing another founder cell and a P-lineage cell that inherits the P-granules. The first asymmetric division of P₁ generates the proto-founder cell EMS, the second gives rise to C, and the third produces D (Sulston et al. 1983). EMS divides again to make the founder cells E and MS (Sulston et al. 1983), which are fated to produce intestine and mesoderm, respectively (Sulston et al. 1983). Both C and D make muscle (Sulston et al. 1983). The last P-lineage cell, P₄ divides one final time during the 100-cell stage of the embryo, producing Z2 and Z3 (Sulston et al. 1983), which populate the entire germline during larval development.

Many RNA-binding proteins play regulatory roles during each of the major events during gametogenesis and early embryogenesis. During the switch from mitosis to

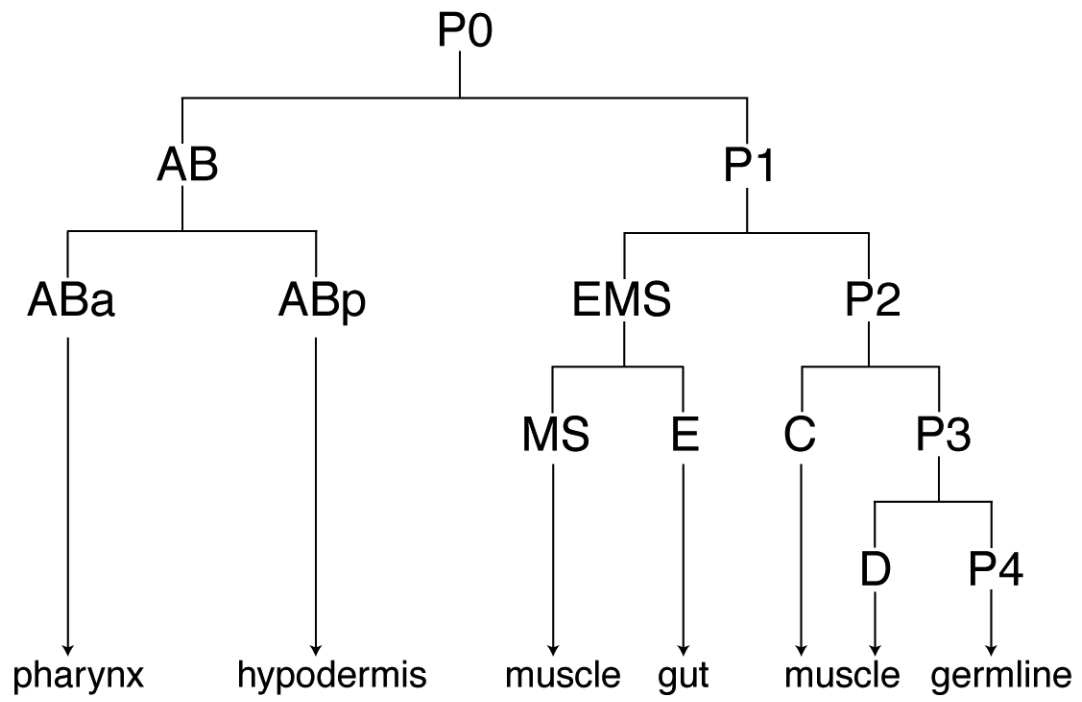


Figure 1.2

Figure 1.2. The cell lineage of the early *C. elegans* embryo. Each founder cell is labeled with the tissue type it is fated to produce.

meiosis in the distal end of the germline, the PUF (Pumilio and FBF) proteins FBF-1 and FBF-2 (reviewed in Tadros and Lipshitz 2009; Crittenden et al. 2002), the Nanos homolog NOS-2 (reviewed in Farley and Ryder 2008; Subramaniam and Seydoux 1999) and the CPEB-related protein FOG-1 promote mitosis (Lécuyer et al. 2007; Barton and Kimble 1990), while the atypical cytoplasmic poly(A) polymerase GLD-2 in complex with the KH-domain containing GLD-3 (Lécuyer et al. 2007; Wang et al. 2002), and the STAR-domain protein GLD-1 promote meiosis (Keene 2007; Francis et al. 1995a) (Figure 1.3). In addition to GLD-1, oogenesis requires the CPEB homolog CPB-3 and its interacting partner DAZ-1 (Licatalosi et al. 2008; Hasegawa et al. 2006; Hafner et al. 2010; Jungkamp et al. 2011), the PUF proteins PUF-5, PUF-6, and PUF-7 (Licatalosi et al. 2008; Lublin and Evans 2007), and the KH-domain containing protein MEX-3 (Hafner et al. 2010; Pagano et al. 2009; Jungkamp et al. 2011). Apoptosis of germ cell nuclei is regulated by the RNA-binding proteins CAR-1, CGH-1 and GLA-3 (Henderson et al. 1994; Boag et al. 2005), and oocyte maturation requires the CCCH-type tandem zinc finger proteins OMA-1 and OMA-2 (Detwiler et al. 2001; Shimada et al. 2006).

Once fertilization occurs, a cascade of RNA-binding protein localization and translation occurs. The CCCH-type tandem zinc finger proteins MEX-5 and MEX-6 (Schubert et al. 2000), as well as the RNA Recognition Motif (RRM) containing protein SPN-4 (Gomes et al. 2001) and the KH-domain containing protein MEX-3 promote anterior fates in the early embryo (Draper et al. 1996), while the CCCH-type tandem zinc finger proteins POS-1 and PIE-1 (Mello et al. 1996; Tabara et al. 1999), and the STAR-

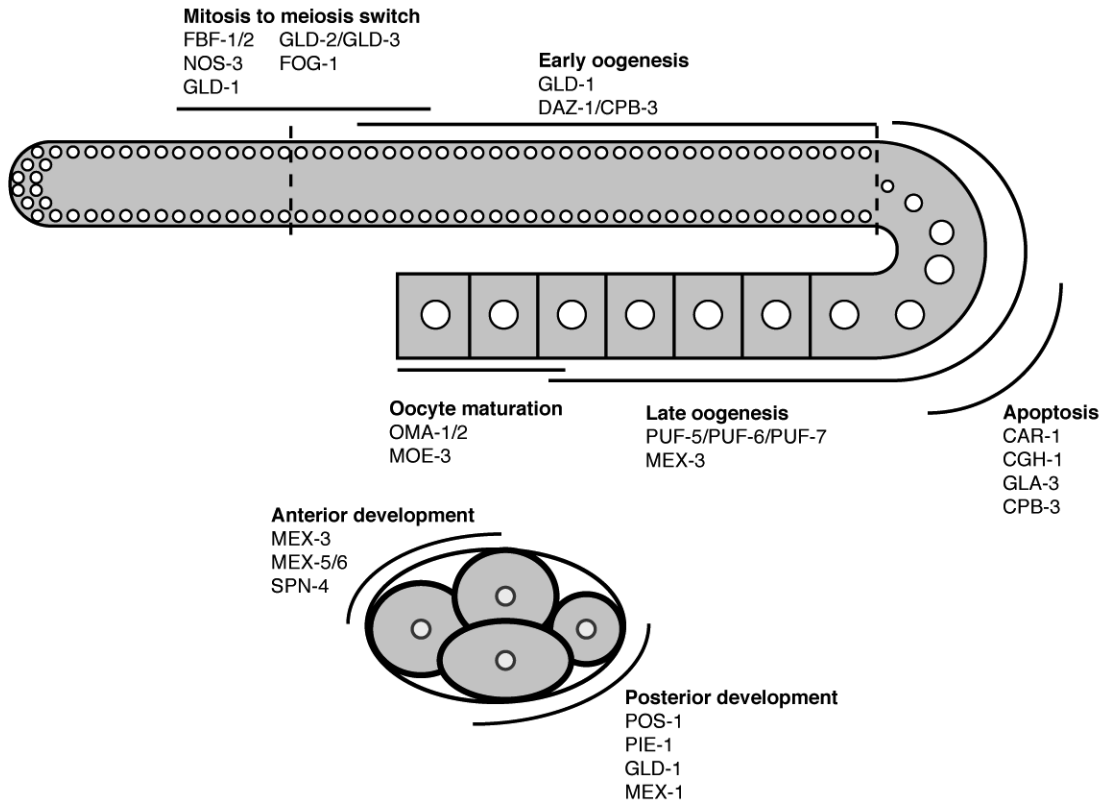


Figure 1.3

Figure 1.3. RNA-binding proteins involved in *C. elegans* gametogenesis and embryogenesis. The RNA-binding proteins that participate in the listed developmental events are located in the approximate region of the gonad or embryo in which they are expressed.

domain containing protein GLD-1 promote posterior fates (Jones et al. 1996) (Figure 1.3). Very little is known about the suite of targets that each of these proteins interacts with, and the mechanism of target regulation remains largely unknown for most of the RNA-binding proteins required for *C. elegans* development.

In general, these RNA-binding proteins are thought to regulate their specific regulatory targets through binding to specific regulatory sequences within their 3' UTRs. Indeed, post-transcriptional regulation is the primary method of gene regulation within the germline and early embryo. In a study performed by the Seydoux lab, the expression patterns of fluorescent transgenic reporters bearing a gene's promoter and a 3' UTR that does not confer patterned regulation were compared with the expression patterns of reporters that have a pan-germline promoter and that gene's specific 3' UTR (Merritt et al. 2008). For all genes tested whose encoded proteins are expressed in the germline, in oocytes, or in the early embryo, the 3' UTR reporter construct mostly recapitulated the endogenous protein's expression pattern. On the other hand, reporters driven by the gene's promoter typically exhibited unpatterned expression throughout the germline and embryo, indicating that post-transcriptional regulation through the 3' UTR drives gene expression during oogenesis and embryogenesis. The 3' UTR reporters for sperm-expressed genes did not faithfully recreate the corresponding endogenous protein's expression pattern, suggesting that gene expression in spermatocytes is regulated by alternative mechanisms (Merritt et al. 2008). While a general model has been proposed

for how RNA-binding proteins recognize their specific targets, the details for most individual proteins are still unknown.

Given the large number of RNA-binding proteins expressed in the germline, as well as their distinct expression patterns, it is likely that an mRNA transcribed in the syncytial region destined for embryonic translation is part of a dynamic mRNP complex whose composition changes as a function of position within the germline. To understand the complete makeup of such an mRNP complex would require a detailed understanding of each of the RNA-binding proteins involved, their specific binding sites within the mRNA, and any co-factors required to elicit regulatory responses. Our knowledge of RNA-binding protein mediated regulation is not yet sufficient to understand the complete makeup and dynamics of such an mRNP complex, but a comprehensive picture is beginning to emerge for a handful of regulatory targets in the *C. elegans* germline. One such target is the Notch homolog *glp-1*, which is required for multiple signaling events in the germline and early embryo.

Function and regulation of *glp-1*

GLP-1 is a member of the Notch family of transmembrane receptors, which are type I transmembrane proteins. The extracellular portion of the protein contains multiple epidermal growth factor (EGF) motifs, followed by three repetitions of the LIN-12/Notch repeat sequence (Yochem and Greenwald 1989). A membrane-spanning spacer sequence that contains two conserved cysteines lies immediately after the LNR sequences. When GLP-1 associates with one of its ligands, this spacer sequence is cleaved at the conserved

cysteines, which liberates the intracellular domain of GLP-1 (Crittenden et al. 1994). This domain contains repeated CDC-10/Ankyrin motifs, as well as a PEST-destabilization domain (Yochem and Greenwald 1989). After cleavage, the intracellular domain of GLP-1 enters the nucleus and associates with LAG-1. The GLP-1/LAG-1 complex is a transcriptional activator (Neves et al. 2007) that promotes expression of genes downstream of the GLP-1 receptor.

GLP-1 protein is required for two classes of inductive events during gametogenesis and embryogenesis. It is expressed on the surface of mitotically dividing cells in the distal arm of the gonad (Crittenden et al. 1994), where it receives a signal from the processes of the distal tip cell (Kimble and White 1981). This signal is the membrane-associated ligand LAG-2 (Henderson et al. 1994), which instructs GLP-1 expressing cells to remain in mitosis, and thus controls proliferation of germline precursor cells. The other class of inductive events takes place in the early embryo. *glp-1* mRNA is maternally supplied, and GLP-1 is initially expressed in two cell embryos on the surface of the anterior blastomere, AB (Evans et al. 1994). After division of AB, GLP-1 continues to be expressed on the surface of both daughter cells, ABa and ABp (Evans et al. 1994). The embryonic ligand for GLP-1 is APX-1, which is expressed on the surface of P₂. As P₂ only makes contact with ABp, GLP-1 signaling polarizes the descendants of AB (Mello et al. 1994; Mango et al. 1994; Evans et al. 1994). In ABp, this inductive signaling mechanism causes the expression of the *ref-1* family of transcription factors, which prevents pharyngeal fate specification and promotes

hypodermal fates (Neves et al. 2007). GLP-1 participates in three more inductive signaling events in the embryo which contribute to pharyngeal development, left-right body axis specification, and excretory cell fate specification (Priess 2005).

glp-1 is essential for both of these processes. Loss of function mutations of *glp-1* were identified in two separate genetic screens; one screened specifically for sterile phenotypes, while the other screened for maternal effect embryonic lethal mutations. Hermaphrodite worms that have no zygotic *glp-1* contribution are sterile, and produce gonad arms that contain approximately 10 germ cells (Austin and Kimble 1987), indicating that *glp-1* is required for mitotic proliferation within the germ line. Embryos lacking maternally supplied *glp-1* fail to produce pharyngeal tissue, and die during embryogenesis (Priess et al. 1987). A dominant gain-of-function mutation (*oz112*) of *glp-1* was identified in a subsequent screen (Berry et al. 1997). This mutation causes constitutively activated GLP-1 signaling, and consequently produces a tumorous germ line full of cells that do not exit from mitosis (Berry et al. 1997). Thus, GLP-1 signaling is both necessary and sufficient for mitotic proliferation of the germline, and tight regulation of GLP-1 activity is essential for viable gametogenesis.

GLP-1 protein is expressed at the distal end of the gonad, as well as the anterior of the four-cell stage embryo, but *glp-1* mRNA is expressed throughout the entire germ line as well as all cells of the early embryo (Evans et al. 1994) (Figure 1.4). Thus, *glp-1* mRNA is translationally repressed in four distinct developmental environments: the syncytial region of the gonad, recellularizing oocytes, mature oocytes, and the early embryo. Translational repression is mediated through the 369 nucleotide long 3' UTR, as

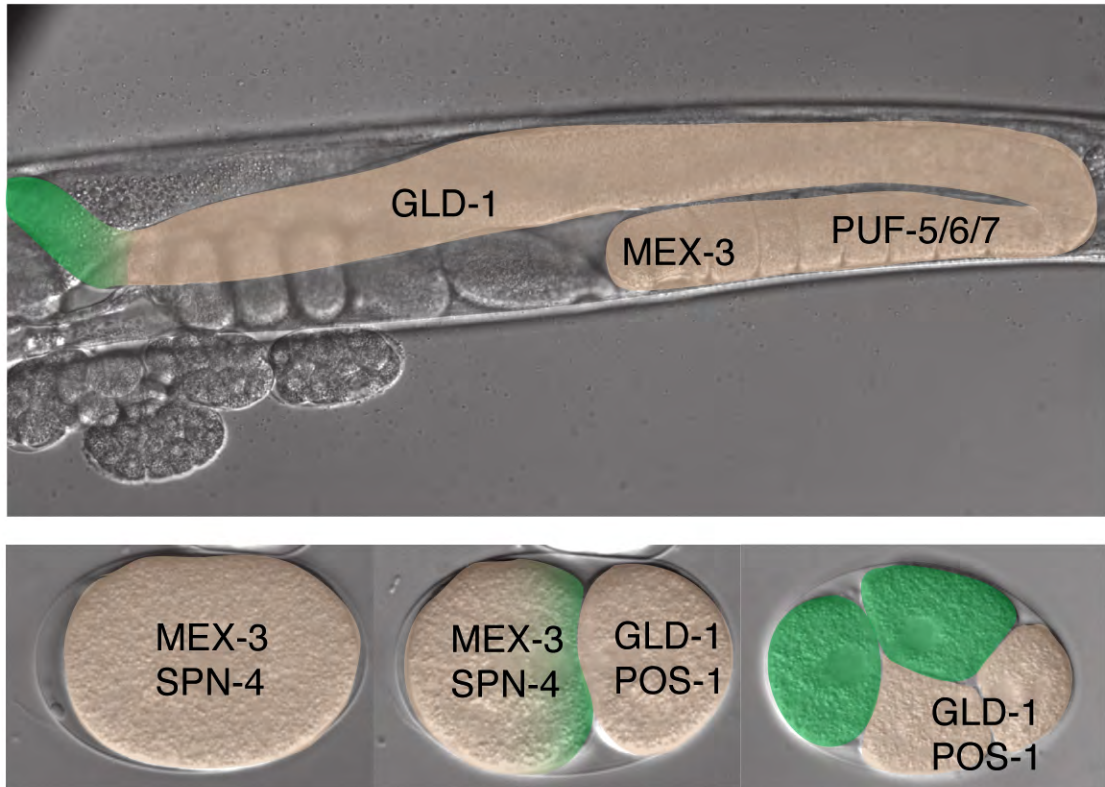


Figure 1.4

Figure 1.4. Translational repression of *glp-1*. The expression pattern of *glp-1* mRNA is highlighted in tan, while the expression of GLP-1 protein is shown in green. RNA-binding proteins involved in regulating *glp-1* are labeled in the approximate regions in which they participate.

demonstrated through the use of *in vitro* transcribed capped, polyadenylated reporter mRNAs encoding *E. coli* beta-galactosidase (Evans et al. 1994). When a version of this mRNA that does not contain the *glp-1* 3' UTR is injected into the syncytial region of adult hermaphrodites, beta-galactosidase expression is observed throughout the syncytial region, in oocytes, and in all cells of embryos (Evans et al. 1994). When this reporter is fused to the *glp-1* 3' UTR and similarly injected, beta-galactosidase expression is restricted to the anterior cells of embryos, which recapitulates the expression pattern of endogenous GLP-1 protein (Evans et al. 1994). Thus, the *glp-1* 3' UTR is sufficient for translational repression in the hermaphrodite germ line.

The *glp-1* 3' UTR contains two regulatory elements that confer spatio-temporal control of *glp-1* translation: the Temporal Control Region (TCR) and the Spatial Control Region (SCR; Figure 1.5). Removal of the TCR in the context of an injected mRNA reporter results in reporter expression throughout the embryo, but otherwise normal expression in the germ line, while removal of the SCR results in both ectopic embryonic expression and expression throughout the germline (Evans et al. 1994). A 34-nucleotide portion of the SCR is sufficient to confer patterned regulation on an unpatterned 3' UTR, suggesting that a subset of the SCR is sufficient to give rise to the appropriate pattern of translation of *glp-1* (Marin and Evans 2003). This fragment of the SCR can be further subdivided into two regulatory sub-elements: the *Glp-1* Repression Element (GRE) and the *Glp-1* De-repression Element (GDE) (Marin and Evans 2003). Mutations within the GRE result in expression of a reporter throughout the gonad and all cells of the early embryo, while mutations within the GDE result in no reporter expression at all within the

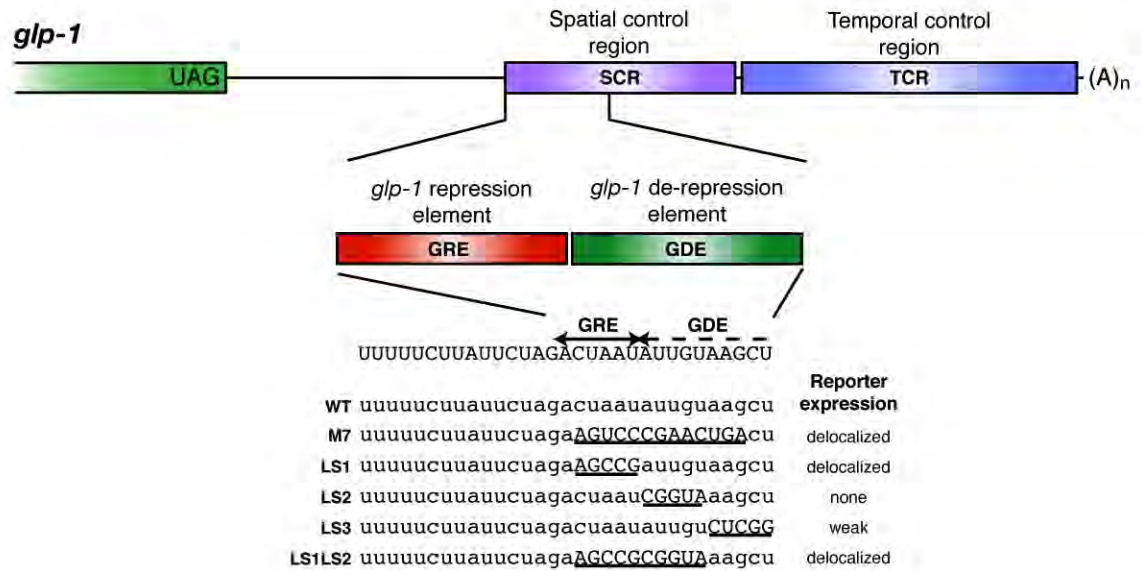


Figure 1.5

Figure 1.5. Regulatory elements in the *gfp-1* 3' UTR. Top, schematic of regulatory elements identified in Evans, *et al* 1994 and Marin and Evans, 2003. Bottom, fine mutagenesis performed in the context of a reporter reveals two subelements within the SCR. The mutations listed result in the reporter expression pattern described to the right.

germline or embryo (Marin and Evans 2003). The GDE mutations do not appreciably reduce the level of reporter mRNA present in the germline or embryo, suggesting that the GDE is required solely for activation of translation (Marin and Evans 2003).

Seven different sequence-specific RNA-binding proteins contribute to the expression pattern of endogenous GLP-1 in the germline and embryo (Figure 1.4). Once the mitotically dividing germ cell precursors enter meiosis, translation of *glp-1* is repressed by GLD-1, which continues to repress translation throughout the syncytial region of the gonad (Francis et al. 1995b; 1995a; Marin and Evans 2003). Null mutants of *gld-1* exhibit a germline tumor similar to the *glp-1* gain-of-function mutation, supporting the hypothesis that GLD-1 is a translational repressor of *glp-1* (Francis et al. 1995a). Furthermore, the GRE contains a GLD-1 binding site, suggesting that this translational repression is direct. As the germ nuclei destined to become oocytes begin to recellularize, the expression level of GLD-1 is reduced (Jones et al. 1996), and the PUF proteins PUF-5, PUF-6, and PUF-7 are responsible for translationally repressing *glp-1* (Lublin and Evans 2007). Knockdown of all three genes simultaneously results in expression of GLP-1 in oocytes, and these oocytes fail to mature (Lublin and Evans 2007). PUF-5, PUF-6, and PUF-7 are no longer expressed in maturing oocytes, and it is thought that MEX-3 represses translation of *glp-1* in maturing oocytes (Pagano et al. 2009). Knockdown of *mex-3* in a strain harboring a fluorescent reporter bearing the *glp-1* 3' UTR results in ectopic reporter expression in all cells of the early embryo (Pagano et al. 2009), suggesting that MEX-3 is required for *glp-1* translational repression.

After fertilization, five proteins are required to regulate the translation of *glp-1* mRNA. MEX-5, MEX-6, and SPN-4 are all expressed in the anterior of the one-cell embryo (Schubert et al. 2000; Gomes et al. 2001; Ogura et al. 2003). Knockdown of MEX-5 and MEX-6 results in no apparent expression of GLP-1 protein (Schubert et al. 2000), while mutations in SPN-4 result in a similar expression pattern (Ogura et al. 2003). POS-1 and GLD-1 are required to translationally repress *glp-1* in the posterior of the embryo; POS-1 is expressed in the posterior of the one-cell embryo shortly after fertilization (Tabara et al. 1999), while GLD-1 is expressed in the posterior starting from the four-cell stage (Jones et al. 1996). Knockdown of either POS-1 or GLD-1 results in expression of GLP-1 throughout the embryo (Ogura et al. 2003; Marin and Evans 2003). Thus, the regulation of *glp-1* translation requires a carefully coordinated suite of RNA-binding proteins that each contribute to the translation of *glp-1* at precise developmental phases of *C. elegans* oogenesis and embryogenesis. The composition of the *glp-1* mRNP is likely to be dynamic as it makes its way through the gonad and embryo, as RNA-binding proteins and perhaps other regulatory factors required to translationally repress or activate *glp-1* mRNA have limited domains of expression throughout the gonad. No single group of factors has yet been identified that translationally represses *glp-1* mRNA throughout the entire process of oogenesis and embryogenesis, and given the factors already identified which play a role in the regulation of *glp-1*, this is not likely to be the case.

To investigate the translational regulation of *glp-1*, or any other transcript, throughout development, a detailed understanding of which factors are binding to which

regions of the mRNA during which stages of development is required. For *glp-1*, we have a limited amount of insight as to the binding sites for each of the factors known to participate in translational regulation. The specificity of GLD-1 has been determined (Ryder et al. 2004), and a GLD-1 binding site is located within the GRE (Marin and Evans 2003). Mutations that overlap with the GLD-1 binding site result in increased reporter translation (Marin and Evans 2003). Thus, it is likely that GLD-1 directly regulates *glp-1* translation through this site. By yeast three-hybrid assay, POS-1 and SPN-4 have been shown to associate with the SCR and TCR of the *glp-1* 3' UTR, respectively (Ogura et al. 2003). This narrows down the potential region that each protein recognizes to approximately 100 nucleotides, but it provides no information about the specific binding sites either protein recognize, the affinities of either protein for the *glp-1* 3' UTR, or the sequence specificities of either protein. There are two matches to the MEX-3 consensus sequence in the *glp-1* 3' UTR (Pagano et al. 2009). One is located within the SCR, and the other lies upstream in a region of the 3' UTR that is dispensable for patterned regulation (Marin and Evans 2003; Pagano et al. 2009). The ability of MEX-3 to bind to either of these sites has not been determined, nor is it known if either plays a role *in vivo*. There is no information regarding PUF-5, PUF-6 or PUF-7 directly binding to the *glp-1* 3' UTR, so it is not known if regulation by these factors is direct. Thus, our understanding of how *glp-1* is regulated through time is incomplete, as is our understanding of most mRNAs throughout the germline. With a better understanding of how sequence-specific RNA binding proteins select their targets, we could begin to map the mRNPs that are required for translational regulation throughout the germline.

Sequence specific RNA-binding proteins in *C. elegans* development

As discussed previously, many proteins containing RNA-binding domains are required to perform a variety of regulatory roles in oogenesis and embryogenesis. However, the sequence specificities of only a few are known, and the complete set of regulatory targets for any RBP is still unknown. To highlight the diversity of RNA-binding protein mediated translational regulation in *C. elegans*, I will discuss the RNA-binding proteins FBF, GLD-1, MEX-3, and POS-1 as examples.

FBF

FBF (*fem-3* binding factor) is the collective term for two nearly identical proteins encoded by *fbf-1* and *fbf-2*. FBF is required to make the switch from spermatocyte production to oocyte production in larval worms (Zhang et al. 1997; Kraemer et al. 1999), and it is also necessary to maintain a population of mitotically dividing germ cell precursors in the distal arm of the gonad (Crittenden et al. 2002). Single mutations of either gene result in germ line development that is essentially wild-type, while double mutations result in an adult germline filled with spermatocytes and lacking a pool of mitotic cells (Zhang et al. 1997). At least three regulatory targets of FBF have been identified: *fem-3*, which encodes a novel protein that promotes spermatogenesis and represses oogenesis (Zhang et al. 1997), *gld-1*, which encodes a RNA-binding protein that inhibits mitosis (Crittenden et al. 2002), and *gld-3*, which encodes a specificity factor for

the cytoplasmic poly(A) polymerase GLD-2 (Eckmann et al. 2004). Together, GLD-2 and GLD-3 are positive regulators of transcripts that promote meiosis, such as *gld-1* (Suh et al. 2006). FBF has been shown to directly associate with each of these transcripts both *in vitro* and *in vivo*, and a single nucleotide mutation that disrupts the FBF binding site in the *fem-3* 3' UTR results in a germline that produces exclusively spermatocytes (Crittenden et al. 2002), supporting the hypothesis that FBF generally acts as a direct repressor of translation.

FBF is a founding member of the PUF (Pumilio and FBF) family of RNA-binding proteins. PUF proteins are typified by an RNA-binding domain that contains eight repeats of a motif comprised of three alpha helices (Zamore et al. 1997; 1999; reviewed in Wickens et al. 2002). These repeats are packed together in a curved structure with each repeat having a similar orientation. This generates an RNA-binding surface on the concave face of the RNA-binding domain. Most PUF proteins recognize an eight nucleotide sequence that begins with the trinucleotide UGU and ends with the dinucleotide UA. In canonical PUF domains, each one of the PUF repeats makes direct contact with one of the eight bases in the recognition element (Wang et al. 2001). However, the optimal binding sequence for FBF contains nine nucleotides (UGURNNAUA, where R is a purine and N is any base) (Opperman et al. 2005; Bernstein et al. 2005). A series of crystal structures of the PUF domain of FBF in association with a variety of target sequences reveals the structural basis of flexible RNA recognition by FBF (Wang et al. 2009). Each of the bases contained within the UGU and

AUA trinucleotides is involved in stacking interactions with amino acids of FBF, and the permitted identities of these nucleotides are thus significantly restricted. Bases at positions 4, 5, and 6, on the other hand, do not form stacking interactions with amino acids of the RNA-binding domain, permitting their identities to be more flexible. The bases at positions 5 and 6 are flipped away from the RNA-binding domain and stacked directly with base 4, which allows FBF to accommodate an additional nucleotide relative to other PUF proteins (Wang et al. 2009; Koh et al. 2011). Depending on the identities of the bases at positions 4 through 6, these bases can adopt a variety of conformations. Thus, flexibility within an RNA recognized by an RNA-binding protein can allow that RNA-binding protein to recognize a variety of sequences.

As a result of FBF's degenerate specificity *in vitro*, it is capable of associating with many transcripts *in vivo*. A genome-wide analysis of FBF-associated transcripts identified 1350 transcripts that are reproducibly immunoprecipitated with FBF from adult, hermaphrodite *C. elegans* extracts (Kershner and Kimble 2010). 67% of these transcripts contain at least one consensus FBE, compared with approximately 30% of all transcripts in *C. elegans*, demonstrating that this strategy enriches for FBF-associated transcripts (Kershner and Kimble 2010). However, even among the most enriched transcripts, only 85% contained FBEs, suggesting that FBF may be capable of recognizing sequences outside of its defined consensus.

FBF is proposed to have two distinct methods of regulating the transcripts with which it associates: recruitment of the CCF-1/Pop2p deadenylase, and recruitment of the atypical cytoplasmic poly(A) polymerase GLD-2. FBF is a negative regulator of *gld-1*

translation in the germline of adult hermaphrodites, and translational repression is achieved through a high-affinity FBF binding site within the *gld-1* 3' UTR. Using an *in vitro* system based on the *gld-1* 3' UTR bearing a short poly(A) tail, Suh and colleagues demonstrated that deadenylation can take place in the presence of purified recombinant FBF and CCF-1, and that this activity is dependent on the presence of a high affinity FBF binding site (Suh et al. 2009). FBF associates directly with CCF-1 by yeast two-hybrid assay, suggesting that FBF directly recruits CCF-1 to 3' UTRs to which it binds, and thus represses translation by shortening the poly(A) tail (Suh et al. 2009). Outside of the germline, FBF is capable of activating translation in neurons by recruiting GLD-2. One target of FBF is *egl-4*, which encodes a kinase required to mediate odor adaptation in *C. elegans*. The *egl-4* 3' UTR contains an FBF binding site, and mutants that disrupt this binding site have decreased levels of EGL-4 protein expression, but no change in the expression level of *egl-4* mRNA (Kaye et al. 2009). These mutants also fail to adapt to noxious odors. Decreased EGL-4 expression is also observed in a *gld-2* mutant, suggesting that cytoplasmic polyadenylation is required for EGL-4 translation (Kaye et al. 2009). Given that FBF associates directly with GLD-2 by yeast two-hybrid assay (Suh et al. 2009), it is possible that FBF is directly recruiting GLD-2 to the *egl-4* 3' UTR and thus stimulating its translation by extending its poly(A) tail. The basis of the apparently opposite regulatory activities of FBF is not well understood, but it is likely to depend on the cytoplasmic content of the cell where regulation is taking place, the sequence context of the FBF binding site, or some combination of both.

GLD-1

GLD-1 is essential for oogenesis, hermaphrodite spermatogenesis and embryogenesis. The germline progenitor cells of *gld-1(null)* mutant worms enter meiosis but exit at the pachytene stage and return to mitosis; thus, the gonad is transformed into a mitotically dividing germline tumor and no viable oocytes or embryos are generated (Francis et al. 1995a). Less severe mutant alleles of *gld-1* result in defects in spermatogenesis and oogenesis without the formation of a germline tumor (Francis et al. 1995b). Adult hermaphrodite worms treated with *gld-1(RNAi)* produce a limited number of fertilized embryos prior to germline tumor formation, and these embryos arrest prior to gastrulation (Marin and Evans 2003). GLD-1 is expressed in the syncytial region of the gonad, as well as the posterior of embryos beginning at the four-cell stage (Jones et al. 1996). GLD-1 has been shown to repress translation of numerous mRNAs including *glp-1*, *tra-2*, and *rme-2* (Marin and Evans 2003; Lee and Schedl 2001). *glp-1* encodes a Notch receptor required for mitosis as well as inductive signaling events in the early embryo. *Tra-2* encodes a membrane protein of unknown function that is required to initiate the switch to oogenesis in late larval worms (Doniach 1986; Kuwabara et al. 1992), while *rme-2* encodes a yolk receptor that is translated in recellularizing oocytes and is required for yolk uptake (Grant and Hirsh 1999). GLD-1 is thus capable of repressing the translation of factors required for germline development events that occur both prior to and after the expression of GLD-1.

Within the syncytial region of the germline, as well as in the posterior cells of early embryos, GLD-1 localizes to cytoplasmic granules. General cytoplasmic staining of GLD-1 is also observed in both cases (Jones et al. 1996; Noble et al. 2008). In both embryos and the germline, these granules are distinct from P-granules, as GLD-1 immunofluorescence does not colocalize with that of the P-granule component PGL-1 (Noble et al. 2008). Extensive analysis of the composition of the embryonic granules has not been performed, but the germline granules contain the RNA-binding proteins CGH-1 and CAR-1, and do not contain the 5' decapping enzyme DCAP-2 (Noble et al. 2008). CAR-1 and CGH-1, but not PGL-1, also immunoprecipitate with GLD-1 from worm extracts, further supporting GLD-1 being a component of these granules (Noble et al. 2008). Given the absence of machinery required for turnover of mRNAs, these germline granules may be required for the storage of translationally inactive mRNA destined for use in developing oocytes or post-fertilization. Consistent with this hypothesis, cytoplasmic germline granules also contain mRNAs for genes required later in oogenesis or embryogenesis such as *rme-2*, *pos-1*, and *glp-1* (Noble et al. 2008; Scheckel et al. 2012). An injected *in vitro* transcribed reporter containing the *glp-1* 3' UTR localizes to these granules. Reporter constructs lacking the *glp-1* SCR fail to localize in the syncytial region and are translated, suggesting that the SCR mediates repression in the distal arm of the gonad, and also that granular localization is necessary for translational repression of *glp-1* (Noble et al. 2008).

GLD-1 contains a STAR RNA-binding domain. STAR domains comprise a maxi-KH domain flanked by two conserved domains termed Qua1 and Qua2 (Ryder et al. 2004). GLD-1 is an obligate dimer, and dimerization is mediated through the Qua1 domain (Ryder et al. 2004). Truncations of the protein that lack this domain not only fail to dimerize, but also bind RNA with a ten fold weaker affinity, indicating that dimerization is essential for high-affinity RNA-binding (Ryder et al. 2004). A recent crystal structure of the GLD-1 Qua1 domain shows that each protomer of the dimer is oriented perpendicularly, which would suggest that the two RNA-binding surfaces of the GLD-1 dimer are distant from one another, and a substantial amount of RNA looping is required for binding (Beuck et al. 2010). Consistent with this model, GLD-1 has a footprint of approximately 20 nucleotides on a high affinity fragment of the *tra-2* 3' UTR (Ryder et al. 2004). GLD-1 also requires two distinct elements for the highest affinity binding: a core hexameric sequence of UACU(A/C)A, and an upstream UA dinucleotide (Ryder et al. 2004). Thus, when bound to RNA, GLD-1 is likely to occupy much more sequence than just its consensus binding site.

In addition to the storage granules described above, GLD-1 also associates with the F-box containing protein FOG-2 (Clifford et al. 2000). F-box domains typically recruit ubiquitination machinery, which usually results in targeting of the ubiquitinated protein for degradation by the proteasome. The association between GLD-1 and FOG-2 is required for spermatogenesis, and as each protein is essential, it is unlikely that GLD-1 is being degraded as a result of FOG-2 association (Clifford et al. 2000). The FOG-

2/GLD-1 complex may target other proteins required for translation that are associated with the same transcript, and thus repress translation by targeting translational machinery for degradation (Clifford et al. 2000). Because GLD-1 dimerizes (Ryder et al. 2004) and recruits multi-subunit regulatory complexes (Clifford et al. 2000), GLD-1 likely has a large footprint on the RNAs with which it associates.

MEX-3

MEX-3 contains a dual-KH domain, and is expressed at both the distal end of the germline as well as the anterior of early embryos (Draper et al. 1996). In the absence of MEX-3, AB (the anterior blastomere of the two cell embryo) produces muscle instead of hypodermal and pharyngeal tissues (Draper et al. 1996). This defect is thought to be the result of mis-regulation of the transcription factor PAL-1, which specifies muscle and whose activity is normally restricted to the posterior C and D blastomeres (Draper et al. 1996). In embryos lacking MEX-3, PAL-1 is expressed in the anterior of the embryo, which suggests that MEX-3 plays a role in translationally repressing *pal-1* in the anterior (Draper et al. 1996). MEX-3 also translationally represses the nanos homolog NOS-2 in embryos (Subramaniam and Seydoux 1999), which may indicate that MEX-3 regulates many transcripts.

Single mutations of *mex-3* do not result in any apparent germline defects, but two double mutations have been discovered that result in substantially transformed germlines. Mutation of *mex-3* simultaneously with *puf-8*, a PUF family RNA-binding protein, results

in a ten-fold reduction of the number of cells within the germline (Ariz et al. 2009). Mutation of either gene alone results in only a two-fold decrease in the number of mitotically proliferating cells, demonstrating redundant control of mitosis by these proteins (Ariz et al. 2009). If *glp-1* is also mutated at the same time, these cells can exit mitosis and differentiate into both spermatocytes and oocytes, suggesting that MEX-3 and PUF-8 redundantly promote mitotic proliferation only (Ariz et al. 2009). A double mutation of *mex-3* and *gld-1* results in a germline containing cells that have transdifferentiated into a variety of somatic types, including neurons and muscle (Ciosk et al. 2006). Thus, it is possible that MEX-3 regulates multiple transcripts required for preserving germ-cell fate in concert with other RNA-binding proteins.

The specificity of the MEX-3 RNA-binding domain has been determined, and it recognizes the highly degenerate sequence DKAGN₀₋₈UHUA, where D is A, G, or U, K is G or U, N is any base, and H is A, C, or U (Pagano et al. 2009). This binding site is present in the 3' UTRs of approximately 30% of *C. elegans* genes, including the known regulatory targets *pal-1* and *nos-2* (Pagano et al. 2009). Given the high number of candidate MEX-3 binding sites, as well as the apparently redundant nature of MEX-3 regulation in the germline, MEX-3 may require other, unknown factors for specific regulation of its targets. In addition, the mechanism of MEX-3 mediated translational repression is also unknown.

POS-1

POS-1 is one of 17 CCCH-type tandem zinc finger RNA-binding proteins (Tabara et al. 1999), in *C. elegans* (reviewed in Kaymak et al. 2010). POS-1 is expressed in the posterior of embryos starting after fertilization, and is restricted to the posterior blastomeres until the four cell stage. After this point, POS-1 expression is only detectable in the P-lineage, where it remains until at least the hundred-cell stage. Throughout early embryogenesis, POS-1 expression is predominantly cytoplasmic (Tabara et al. 1999). Embryos lacking POS-1 arrest prior to gastrulation with defects in three tissue types: AB-derived pharynx, E-derived intestine, and P-lineage-derived germ cell precursors (Tabara et al. 1999). The first divisions of the P-lineage are also abnormal. In wild-type embryos, the pace of division of the P-lineage is slow relative to the other blastomeres, but in POS-1 mutant embryos, the P-lineage cell divides at a similar rate to the somatic blastomeres (Tabara et al. 1999). The pleiotropic nature of the *pos-1* mutant phenotype suggests it regulates multiple critical targets during embryogenesis. To understand the biological role that POS-1-mediated translational regulation plays in the early embryo, we set out to determine its *in vitro* specificity, and then use that information to predict and identify critical regulatory targets. Chapter II describes the *in vitro* biochemical characterization of the RNA-binding domain of POS-1, including a comprehensive determination of its sequence specificity. Chapter III describes the POS-1-mediated regulation of *glp-1* in concert with the RNA-binding protein GLD-1.

Chapter II

RNA target specificity of embryonic cell

fate determinant POS-1

ABSTRACT

Specification of *Caenorhabditis elegans* body axes and cell fates occurs prior to the activation of zygotic transcription. Several CCCH-type tandem zinc finger (TZF) proteins coordinate local activation of quiescent maternal mRNAs after fertilization, leading to asymmetric expression of factors required for patterning. The primary determinant of posterior fate is the TZF protein POS-1. Mutants of *pos-1* are maternal effect lethal with a terminal phenotype that includes excess pharyngeal tissue and no endoderm or germline. Here, we delineate the consensus POS-1 recognition element (PRE) required for specific recognition of its target mRNAs. The PRE is necessary but not sufficient to pattern the expression of a reporter. The PRE is distinct from sequences recognized by related proteins from both mammals and nematodes, demonstrating that variants of this protein family can recognize divergent RNA sequences. The PRE is found within the 3' untranslated region of 227 maternal transcripts required for early development, including genes involved in endoderm and germline specification. The results enable prediction of novel targets that explain the pleiotropy of the *pos-1* phenotype.

INTRODUCTION

During the first few rounds of cell division in the development of a metazoan, body axes are formed, cell fates are specified, and a rough body plan is established. In many species, these events occur without the benefit of zygotic transcription, relying instead on the asymmetric localization and translation of specific maternally-supplied mRNAs (reviewed in Farley and Ryder 2008). A network of *cis*-acting regulatory elements and cognate *trans*-acting specificity factors, including RNA-binding proteins and microRNAs, is required to ensure that mRNAs are appropriately regulated. Frequently, the *cis*-acting regulatory elements are found in the 3' untranslated region (3' UTR) of mRNAs, as this region is not actively translated and is therefore readily accessible to the *trans*-acting regulatory factors (reviewed in Kuersten and Goodwin 2003).

Early embryogenesis in the nematode *Caenorhabditis elegans* requires the regulation of an extensive network of maternally supplied mRNAs, as the onset of zygotic transcription is delayed. Prior to the initiation of zygotic transcription, the fates of all of the founder cells that produce the tissues and organs present in the adult worm are established (Sulston et al. 1983) (Figure 2.1). This process begins soon after fertilization, when the zygote divides asymmetrically across the anterior-posterior axis. The larger anterior daughter is the first founder cell, while the smaller posterior daughter is the progenitor of the germline. The germline progenitor repeats this pattern of asymmetric cell division three more times, eventually giving rise to six total founder cells

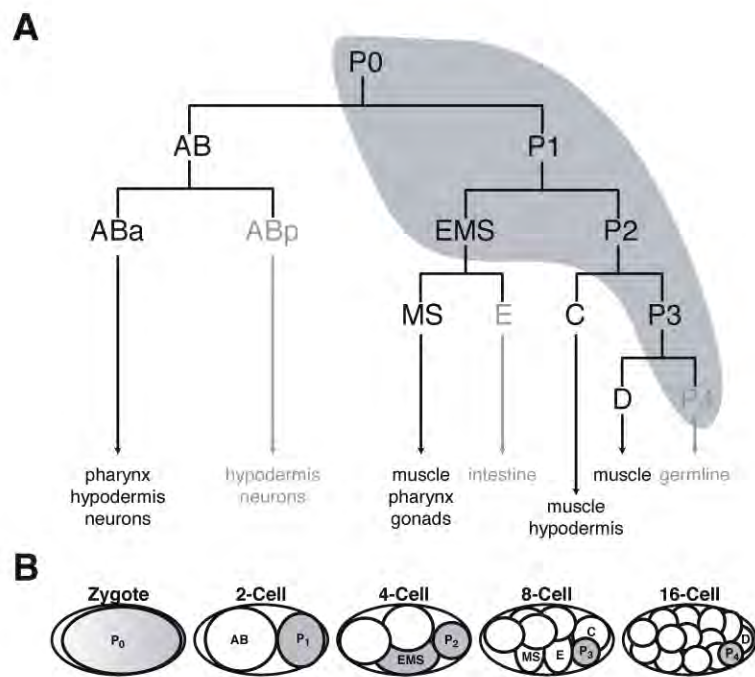


Figure 2.1

Figure 2.1. POS-1 is required for the specification of multiple cell fates in the early *C. elegans* embryo. A. The early *C. elegans* cell lineage is shown, with each founder cell labeled with the tissue types it produces. The expression of POS-1 protein is shown in grey, and each founder cell specified incorrectly in *pos-1* mutant embryos is grayed out. B. Schematic of POS-1 expression in early embryos. POS-1 expression is shown in grey, and each founder cell is labeled.

that together are the progenitors of all the tissue types present in the adult worm (Sulston et al. 1983). Most cells of the zygote do not begin transcription until the four cell stage, and transcription does not begin in the germline lineage until just prior to gastrulation. (Mello et al. 1996; Seydoux et al. 1996; Seydoux and Dunn 1997). Thus, post-transcriptional regulation of maternal transcripts by maternal RNA-binding proteins is the primary mechanism that drives cell fate specification in the early embryo.

POS-1 is a critical *trans*-acting factor required for *C. elegans* early embryogenesis (Tabara et al. 1999). POS-1 accumulates in the posterior of the fertilized zygote and is inherited asymmetrically at each division (Figure 2.1). Embryos lacking POS-1 fail to hatch; the terminally differentiated embryos lack intestine and germ cell precursors, and have excess pharyngeal tissue. Each of these tissue types is derived from a different founder cell, indicating that POS-1 is required for multiple cell fate specification events.

Three genes have been identified whose expression is perturbed in *pos-1* mutants: *glp-1*, *apx-1* (Ogura et al. 2003), and *nos-2* (D'Agostino et al. 2006). *nos-2* encodes a protein similar to *Drosophila* Nanos that is required for germ cell development and migration during gastrulation (Subramaniam and Seydoux 1999). *glp-1* encodes a cell surface receptor homologous to *Drosophila* Notch, and *apx-1* encodes a ligand homologous to *Drosophila* Delta that is recognized by GLP-1 (Fehon et al. 1990; reviewed in Artavanis-Tsakonas et al. 1999). APX-1 and GLP-1 are required to pattern anterior development; their interaction at the two-cell stage polarizes the anterior blastomere as it divides, causing its posterior daughter (ABp) to adopt a hypodermal fate (Mello et al. 1994). In *pos-1* mutants, GLP-1 is aberrantly expressed in all cells of the

early embryo, while APX-1 expression is undetectable (Ogura et al. 2003). This prevents polarization of the two-cell stage anterior blastomere, causing the formation of excess pharyngeal tissue at the expense of hypodermis.

The mRNA encoding *glp-1* is present in every cell of the embryo until the eight-cell stage, but GLP-1 protein is expressed only in the anterior of the embryo (Ogura et al. 2003). The *glp-1* 3' UTR is both necessary and sufficient to direct this expression pattern, as microinjected mRNA encoding a lacZ reporter under control of the *glp-1* 3' UTR is expressed in a pattern identical to endogenous *glp-1* (Evans et al. 1994). This suggests that *glp-1* mRNA is translationally repressed through its 3' UTR in the posterior of the embryo. Two seventy-nucleotide elements within the 3' UTR, termed the spatial control region (SCR) and the temporal control region (TCR), are required for both spatial and temporal patterning. Within the SCR are two regulatory sub-elements required for translational repression (*glp-1* repression element, or GRE) and translational activation (*glp-1* de-repression element, or GDE). By yeast-three hybrid, POS-1 associates with both the SCR and TCR, suggesting that it represses translation of *glp-1* through direct association (Ogura et al. 2003). However, given the complex pleiotropic phenotype of *pos-1* mutant embryos (Tabara et al. 1999), it is unlikely that derepression of *glp-1* translation drives all of the observed patterning defects.

POS-1 is one of several nematode CCCH-type tandem zinc finger proteins (hereafter TZF) required for oocyte maturation and early development (Mello et al. 1996; Schubert et al. 2000; Shimada et al. 2006; Shirayama et al. 2006). These proteins are related to tristetraprolin (TTP), a mammalian factor that promotes the deadenylation and

subsequent rapid turnover of tumor necrosis factor alpha (TNF α) mRNA by direct association with its 3' UTR (Carballo et al. 1998; Lai et al. 1999; 2005). TTP binds to the sequence UUAUUUAUU, present in multiple copies within the AU-rich element (ARE) of the TNF α 3' UTR, with high affinity and specificity (Worthington et al. 2002; Blackshear et al. 2003; Brewer et al. 2004). In contrast, the nematode TZF protein MEX-5, required for anterior development, binds with high affinity but relaxed specificity to uridine-rich sequences (Schubert et al. 2000; Pagano et al. 2007). Little is known about the binding specificity of the other members of the nematode TZF protein family, including POS-1. To probe the basis for specific mRNA recognition by POS-1, and to facilitate prediction of novel POS-1 regulatory targets, we set out to delineate the RNA-binding specificity of this protein.

RESULTS

POS-1 binds weakly to TTP and MEX-5 binding sites

We first asked whether POS-1 binds to RNA with the same specificity as TTP or MEX-5. Recombinant POS-1 TZF domain (amino acids 80–180) was expressed as a C-terminal fusion to maltose binding protein (MBP) in *Escherichia coli* and purified to near homogeneity. The recombinant protein was used in quantitative electrophoretic mobility shift assays (EMSA) with fluorescein end-labeled RNAs encoding high affinity binding sequences recognized by TTP or MEX-5: ARE13 and poly(U)-13, respectively (Brewer et al. 2004; Pagano et al. 2007) (Figure 2.2). Varying concentrations of POS-1 were equilibrated with trace labeled RNA, and the bound RNA was resolved from the free

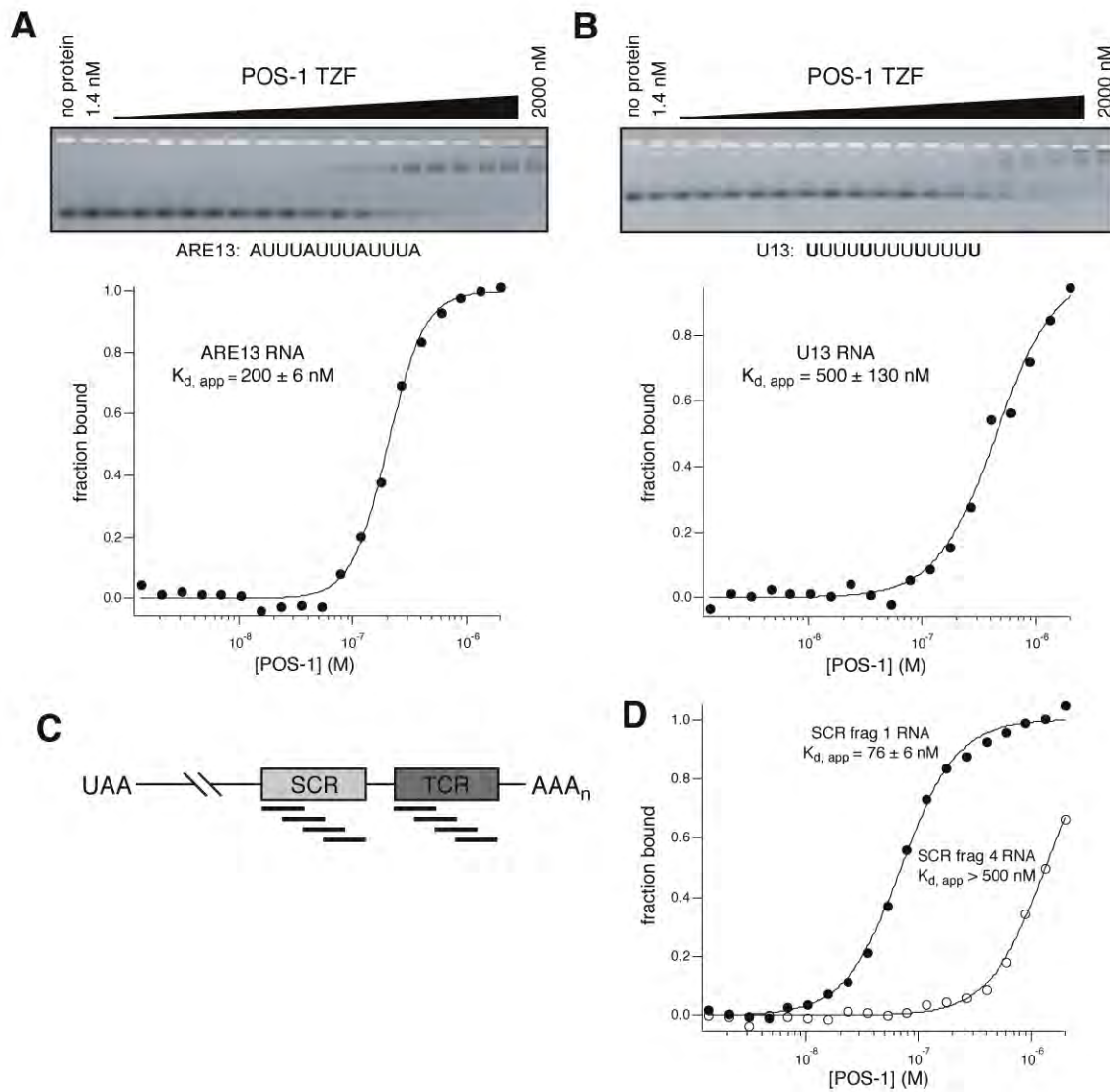


Figure 2.2

Figure 2.2. POS-1 is a specific RNA-binding protein with a different specificity than either TTP or MEX-5. A. POS-1 binding to ARE13 was measured by electrophoretic mobility shift assay. A representative gel is shown, with a plot of POS-1 concentration versus fraction bound below. The reported $K_{d, app}$ is the average \pm one standard deviation of three independent replicates. B. POS-1 binding to U13 was measured as in A. C. Schematic of tiled fragments of the *glp-1* SCR and TCR. D. POS-1 binding to fragments one and four of the *glp-1* SCR was measured as in A.

RNA by gel electrophoresis. The fraction of bound RNA at each concentration of POS-1 was determined using a FUJI FLA-5000 fluorimager and the apparent equilibrium dissociation constant ($K_{d, app}$) was determined by a fit of the data to the Hill equation. POS-1 binds to ARE13 with modest affinity ($K_{d, app} = 200 \pm 6$ nM), approximately 50 times weaker than the previously published affinity of TTP for ARE13 (Brewer et al. 2004). Likewise, POS-1 binds to poly(U)-13 RNA ($K_{d, app} = 500 \pm 130$ nM), but the affinity is reduced compared to MEX-5's affinity for the same sequence (Pagano et al. 2007). Together, the results demonstrate that though POS-1 binds directly to RNA recognized by TTP and MEX-5, it does so with significantly reduced affinity.

POS-1 binds to the 3' UTR of multiple genes required for embryogenesis

We hypothesized that POS-1 binds with higher affinity to a sequence determinant that is different from TTP or MEX-5. In an attempt to identify a high affinity POS-1 interacting sequence, we constructed a library of tiled fragments of the *glp-1* SCR and TCR. These elements were previously demonstrated to associate with POS-1 by yeast three hybrid analysis (Ogura et al. 2003). Each fragment is approximately thirty nucleotides long and overlaps the previous fragment by fifteen nucleotides (Figure 2.2). The association of POS-1 with these fragments was assayed by EMSA (Figure 2.2, Table 2.1). POS-1 binds to three of the four SCR fragments, and to all four TCR fragments. The affinity of POS-1 is highest for the fragments near the 5' end of the SCR ($K_{d, app} = 76 \pm 6$), which contain the GRE and GDE. The data indicate that one or more high affinity POS-1 binding sites are present in both the SCR and the TCR. However, a comparison of

Table 2.1: Electrophoretic mobility shift assay with *glp-1* 3' UTR fragments

| Name | Sequence ^a | K _{d, app} (nM) | n |
|------------|---|--------------------------|---|
| SCR frag 1 | UUAUUCUAGACUAAU AUUGUAAG CUAUAAG | 76 ± 6 | 3 |
| SCR frag 2 | AUUGUAAG CUAUAAGUUGUAGAAUAAUUAU | 97 ± 6 | 3 |
| SCR frag 3 | UUGUAGAAUAAUUAUUGAUCCAAAUCAGAU | 120 ± 20 | 3 |
| SCR frag 4 | UGAUCCAAAUCAGAUUAAGAGUAUAA | 500 ± 100 | 3 |
| TCR frag 1 | UUGUUUUUCUCCUUUUUUUAUAACUUGU | 150 ± 20 | 3 |
| TCR frag 2 | UUUCUUUAUAACUUGUUACAAUUUUUGAAA | 120 ± 20 | 3 |
| TCR frag 3 | UUACAAUUUUUGAAAUCCCUUUUUUGACA | 310 ± 30 | 3 |
| TCR frag 4 | UUCCCUUUUUUGACAGGCUUUUUAUUACACUGUAA | 189 ± 7 | 3 |

^a Boldface denotes PRE

Each reported K_{d, app} is the mean ± one standard deviation of the number of replicates

listed to the right.

the sequences of each fragment does not easily reveal the determinants of the high affinity interactions.

In a second approach to identify sequences that POS-1 binds with high affinity, we constructed a library of sequences from maternal mRNA 3' UTRs that contain at least two UAUU elements (TTP half sites) with no more than two intervening nucleotides. We rationalized that POS-1 may bind to similar determinants as TTP, but with altered spacing between those determinants due to an increase in the number of amino acids that link the individual zinc fingers (17 in POS-1 compared to 12 in TTP). The library was biased to include only UTR fragments from genes that are post-transcriptionally regulated during early development (Bowerman et al. 1993; Mello et al. 1994; Draper et al. 1996; Hunter and Kenyon 1996; Gomes et al. 2001). The affinity of POS-1 for each sequence in the library was assayed by EMSA (Table 2.2). In total, six UTR fragments from five genes (*apx-1*, *mex-3*, *pal-1*, *skn-1*, and *spn-4*) were tested. POS-1 binds to two of these sequences with high affinity. These include a short fragment of the *pal-1* 3'-UTR ($K_{d,app} = 84 \pm 3$ nM), and a longer fragment of the *mex-3* 3'-UTR that contains multiple UAUU sequences (*mex-3* fragment:

AACUAUUAUUAUUUGUUAUUCAUAUUUU, $K_{d,app} = 47 \pm 7$ nM) (Figure 2.3).

MEX-3 is a KH domain RNA-binding protein required for specifying the fates of anterior blastomeres during early embryogenesis (Draper et al., 1996). MEX-3 acts by inhibiting the developmental program that specifies body wall muscle. In wild-type embryos, this tissue type is produced exclusively by one of the posterior founder cells, while in *mex-3* mutant embryos, it is produced ectopically by descendants of the two-cell

Table 2.2: Electrophoretic mobility shift assay with maternal mRNA fragments

| Name | Sequence ^{a,b} | K _{d, app} (nM) | n |
|---------------------|--|--------------------------|---|
| <i>apx-1</i> frag | GUUU <u>AUUUU</u> U <u>AUU</u> AU | 120 ± 10 | 3 |
| <i>pal-1</i> frag 1 | AUUU <u>AUU</u> A <u>AUUUU</u> | 167 ± 6 | 3 |
| <i>pal-1</i> frag 2 | CUU UAUUUAUUGU | 84 ± 3 | 3 |
| <i>skn-1</i> frag | AGUU <u>AUUUC</u> <u>AUU</u> AU | 130 ± 30 | 3 |
| <i>spn-4</i> frag | ACGU <u>AUU</u> G <u>AUUUU</u> | 250 ± 20 | 3 |
| <i>mex-3</i> frag | AACU <u>AUU</u> UAUU AUUUG UU <u>AUU</u> CA <u>AUUUU</u> | 47 ± 7 | 3 |
| <i>poly(C)-15</i> | CCCCCCCCCCCCCCC | >1000 | 1 |

a Underline denotes TTP half site

b Boldface denotes PRE

Each reported K_{d, app} is the mean ± one standard deviation of the number of replicates listed to the right.

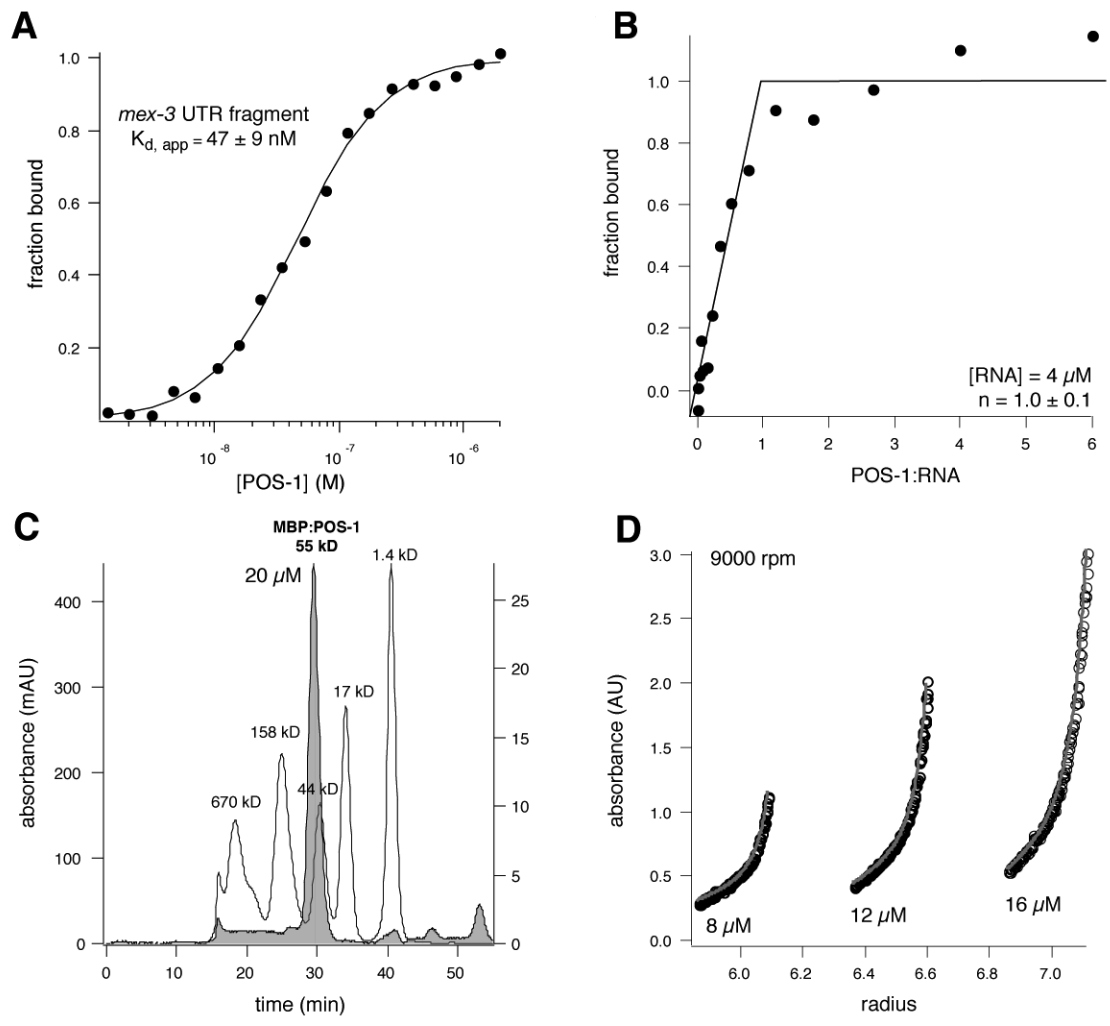


Figure 2.3

Figure 2.3. POS-1 binds to the *mex-3* UTR fragment with apparent 1:1 stoichiometry.

A. POS-1 binding to the *mex-3* UTR fragment was measured as in figure 1A. B.

Stoichiometric binding assay between POS-1 and *mex-3* RNA. The total RNA concentration for the experiment is indicated. N was determined from a quadratic fit (Rambo & Doudna, 2004). C. Gel filtration chromatogram of recombinant POS-1. Gel filtration standards are displayed in white. Molecular weights of each standard peak are indicated. POS-1 elution profile is displayed in grey, and the protein concentration and apparent molecular weight are indicated. D. Equilibrium analytical ultracentrifugation traces of recombinant POS-1. The rate of rotation, as well as the protein concentration for each trace is indicated. Grey lines represent the fit to a monomer-Nmer equilibrium.

stage anterior daughter (Draper et al., 1996). This suggests that MEX-3 functions during the same stage of embryogenesis as POS-1, and also that MEX-3 activity must be restricted to the anterior of the embryo. Consistent with this, both MEX-3 protein and *mex-3* mRNA are asymmetrically distributed in the anterior of the embryo at the two and four cell stages in a pattern that is anti-correlated with POS-1 (Draper et al., 1996). POS-1 may regulate the expression of *mex-3* by repressing its translation, promoting the turnover of *mex-3* mRNA, or both. To further explore the potential regulatory relationship between POS-1 and *mex-3*, we decided to further characterize the interaction between *mex-3* mRNA and POS-1.

POS-1 forms an equimolar complex with the mex-3 fragment

POS-1 binds to the *mex-3* 3' UTR fragment with the highest affinity of all tested sequences. This could be due to the presence of multiple binding sites, or it might be due to the presence of a single site that binds to POS-1 with higher affinity. To distinguish between these possibilities, the stoichiometry of the complex between POS-1 and this RNA fragment was determined by EMSA (Figure 2.3). Varying concentrations of POS-1 were equilibrated with a fixed, elevated concentration of unlabeled *mex-3* RNA supplemented with a trace amount of fluorescently labeled *mex-3* RNA. After equilibration, the bound RNA was resolved from the free RNA by gel electrophoresis and the fraction of bound RNA was determined by fluorimetry. The data were fit to a quadratic model of saturable binding (Rambo and Doudna 2004). The apparent stoichiometry of the POS-1:*mex-3* complex is one to one (saturation point = 1.0 ± 0.1),

consistent with the hypothesis that the *mex-3* RNA contains a single, high affinity binding site.

To ensure that POS-1 binds as a monomer, the oligomerization state of recombinant POS-1 was determined by size exclusion chromatography and by equilibrium sedimentation ultracentrifugation (Figure 2.3). Both methods reveal that POS-1 is predominantly monomeric, even at a concentration near 20 μ M, orders of magnitude higher than the apparent dissociation constant for the *mex-3* RNA. We conclude that monomeric POS-1 binds to the *mex-3* 3' UTR fragment with high affinity as a one to one molar stoichiometric complex.

POS-1 recognizes a twelve nucleotide fragment within the mex-3 UTR

To identify the minimal POS-1 binding site within the *mex-3* 3'-UTR fragment, three overlapping 15-nucleotide sub-fragments of this RNA were synthesized. The affinity of each for POS-1 was determined by EMSA (Figure 2.4). POS-1 binds to the fragment #1 and #2 slightly weaker than the intact sequence (*mex-3* fragment #1 $K_{d, app} = 110 \pm 40$ nM; *mex-3* fragment #2 $K_{d, app} = 89 \pm 5$ nM). In contrast, POS-1 binds to fragment #3 with drastically reduced affinity (*mex-3* fragment #3 $K_{d, app} = 800 \pm 200$ nM). This suggests that a high affinity POS-1 binding site is located in the overlap between the first and second *mex-3* fragments. To test this hypothesis, a twelve nucleotide fragment that corresponds to this overlap (hereafter, *mex-3 min*) was synthesized. This sequence binds with identical affinity, within error, to the original *mex-3* RNA (*mex-3 min* $K_{d, app} = 39 \pm 6$ nM, *mex-3* RNA $K_{d, app} = 47 \pm 9$ nM, Figure 2.4).

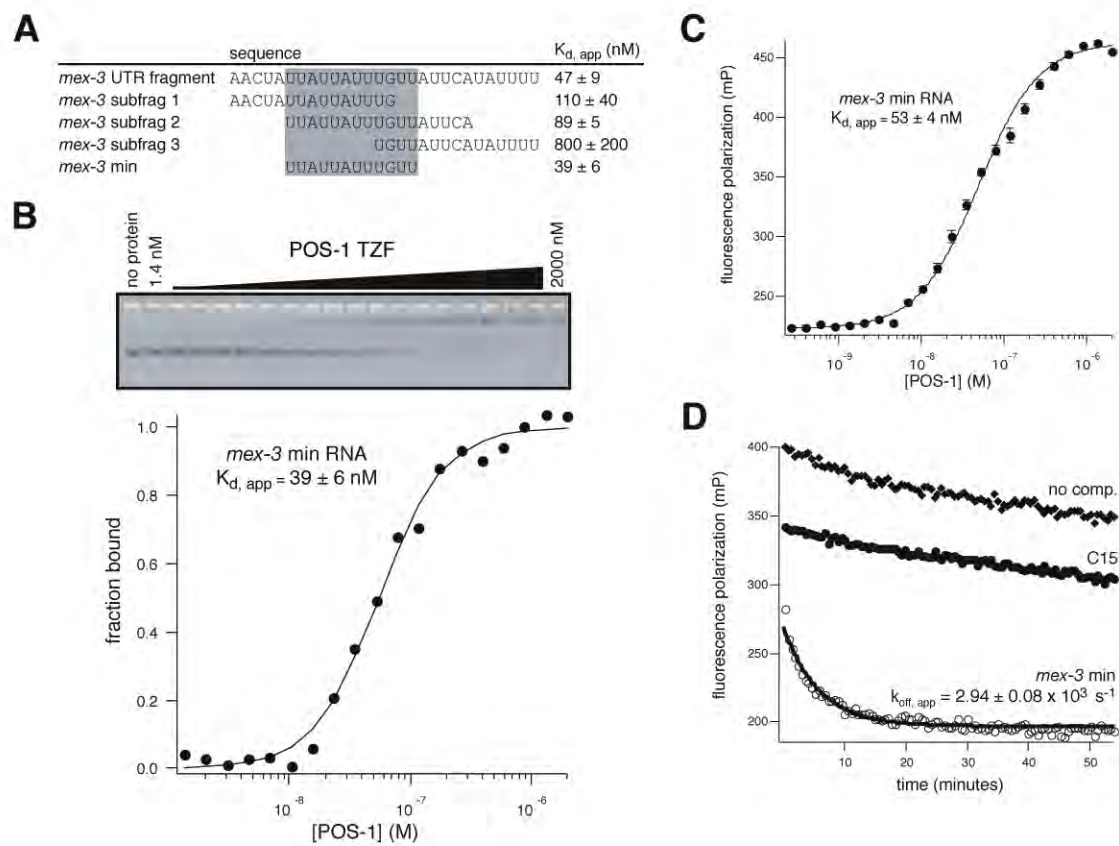


Figure 2.4

Figure 2.4. POS-1 specifically recognizes a twelve nucleotide fragment of the *mex-3* 3' UTR. A. Table of association measurements for POS-1 and *mex-3* fragments. $K_{d,app}$ was measured by electrophoretic mobility shift assay, and the reported values are the average \pm standard deviation for three independent replicates. The *mex-3* min sequence is highlighted in grey. B. Representative electrophoretic mobility shift assay for *mex-3* min. The fit and reported $K_{d,app}$ are as in figure 1A. C. Representative fluorescence polarization assay for *mex-3* min. Each data point is the average \pm standard deviation of five reads of an independent replicate. The data were fit to the Hill equation, and the reported $K_{d,app}$ is the average \pm standard deviation of three independent replicates. D. Representative kinetic fluorescence polarization assays for POS-1 and the indicated RNAs. $k_{off,app}$ was determined by monitoring the change in fluorescence polarization after the addition of a hundred-fold excess of unlabeled competitor RNA, and fitting to a single exponential. Empty circles, unlabeled *mex-3* min competitor. Filled circles, unlabeled C15 competitor. Filled diamonds, no competitor. Solid black line represents the single exponential fit.

To confirm the equilibrium dissociation constant for *mex-3* min, and to develop a convenient assay for kinetic analysis, we assessed the ability of POS-1 to change the polarization of fluorescein end-labeled *mex-3* min RNA in solution. Varying concentrations of protein were equilibrated with limiting labeled *mex-3* min RNA, and the polarization value determined using a fluorescence plate reader. The association of POS-1 with the fluorescently labeled *mex-3* min RNA significantly increases the polarization of the fluorophore, thus providing a parameter to monitor POS-1 binding in real time. The apparent equilibrium dissociation constant was determined by plotting polarization as a function of protein concentration and fitting the data to the Hill equation (Figure 2.4). The $K_{d, app}$ of POS-1 for *mex-3* RNA is 53 ± 4 nM, similar to the $K_{d, app}$ determined by EMSA. Next, the polarization assay was used to determine the dissociation kinetics of the complex. An excess of unlabeled *mex-3* min RNA, poly(C)-15 RNA, or a buffer control was added to a pre-formed complex of POS-1 and labeled *mex-3* min RNA (Figure 2.4). Unlabeled *mex-3* min serves as a trap to prevent reassociation with the labeled RNA over the time course of the experiment. Unlabeled poly(C)-15 RNA does not bind to POS-1 (Table 2), and serves as a non-specific binding control. To control for dissociation due to dilution, an experiment was performed with buffer in place of the unlabeled RNA. The apparent dissociation rate constant ($k_{off, app}$) was determined by fitting the observed change in polarization as a function of time to an exponential decay. The $k_{off, app}$ is $2.94 \pm 0.08 \times 10^{-3} \text{ s}^{-1}$ when unlabeled *mex-3* min is used as a competitor. In contrast, little dissociation is observed when poly(C)-15 or buffer

alone is added. Based on these measurements, and the determination of the $K_{d, app}$ above, the association rate constant can be calculated ($k_{on, calc} = 6.0 \pm 0.4 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$). These results provide a kinetic and thermodynamic framework for detailed analysis of the interaction between POS-1 and the minimal *mex-3* fragment.

Determination of the POS-1 consensus sequence

Next, we set out to determine the sequence determinants within *mex-3* min that contribute to binding. To do this, the change in the standard free energy change ($\Delta\Delta G^\circ$) was measured for every single point mutation of *mex-3* min RNA using EMSA. In total, thirty-six individual mutations were tested, representing every possible single nucleotide substitution at each position of the *mex-3* min sequence (Figure 2.5). Mutation of seven positions causes a significant decrease in affinity (position 2–7, 10: $\Delta\Delta G > 0.5 \text{ kcal/mol}$), while five positions can tolerate any base substitution (position 1, 8, 9, 11, and 12: $\Delta\Delta G < 0.5 \text{ kcal/mol}$). Based on this data, the POS-1 recognition element (PRE) is UAUURDNNG, where R is any purine, D is A, G, or U, and N is any base. Compared to the TTP binding site, the PRE contains an extra purine and displays a greater degree of degeneracy (Brewer et al. 2004). In contrast with the MEX-5 binding sequence, the PRE exhibits strict requirements at several positions, while MEX-5 binds to any uridine-rich sequence (Pagano et al. 2007). The disparity in specificity between POS-1, TTP, and

MEX-5 suggests that the POS-1-RNA interface is significantly different from those of the other proteins.

To investigate the spacing requirements between each of the purines in the PRE, the relative position of each purine nucleotide was varied within a polyuridine background, and the $\Delta\Delta G^\circ$ for each mutant sequence was determined by EMSA (Figure 2.5). POS-1 is tolerant of an additional nucleotide between A3 and A6, and one additional or one fewer nucleotide between A6 and G10. Taking the flexibility of the spacing between purines into account, we expand the PRE to 5'-UA(U₂₋₃)RD(N₁₋₃)G-3'. This consensus is present in all tested sequences that bind to POS-1 with high affinity, including fragments from the *pal-1* and *glp-1* 3' UTRs (*pal-1* fragment #2, $K_{d, app} = 84 \pm 3$ nM and *glp-1* SCR fragment #1, $K_{d, app} = 76 \pm 6$ nM).

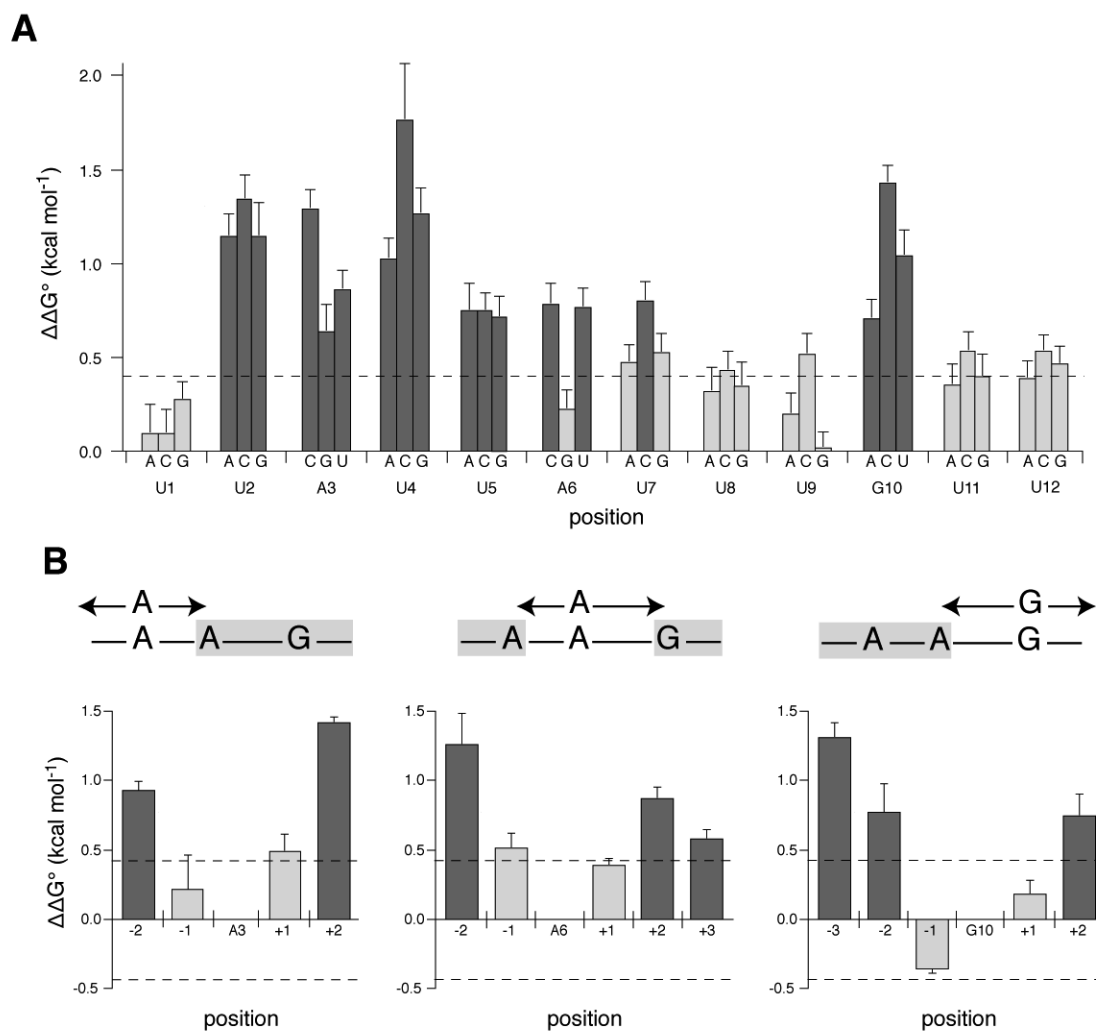


Figure 2.5

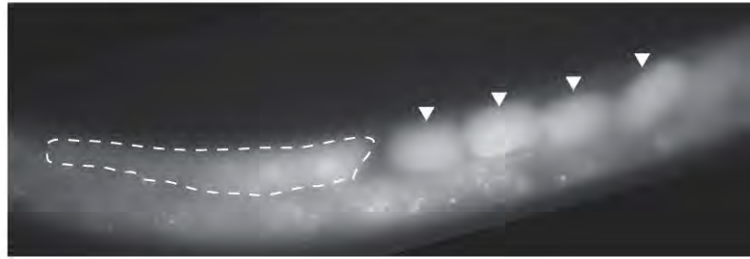
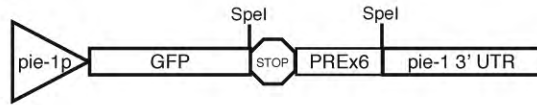
Figure 2.5. POS-1 recognizes the sequence UA(U₂₋₃)RD(N₁₋₃)G. A. Systematic mutagenesis of *mex-3* min. The $K_{d,app}$ for each mutation of *mex-3* min was measured by electrophoretic mobility shift assay and compared to the previously measured $K_{d,app}$ for *mex-3* min to determine the change in standard change of free energy ($\Delta\Delta G^\circ$). Each bar represents the $\Delta\Delta G^\circ$ caused by the base substitution indicated on the x-axis. Substitutions that cause a greater than 0.4 kcal/mol (dotted line, approximately two-fold difference in $K_{d,app}$) change in standard free energy are shown in dark grey. Error bars are the propagation of error derived from the standard deviations of three replicates each of *mex-3* min and the indicated base substitution. B. Spacing mutagenesis of *mex-3* min. The $K_{d,app}$ for a series of double mutations that switch the position of the indicated purine with either an upstream or downstream uracil was measured as in A. Each bar represents the $\Delta\Delta G^\circ$ caused by one double mutation, determined as in A. Double mutations that cause a greater than 0.4 kcal/mol change in standard free energy are shown in dark grey.

The PRE is necessary but not sufficient to pattern the expression of a reporter

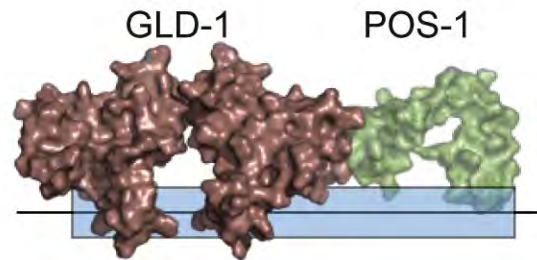
Published work from Marin and Evans identified a 34-nucleotide fragment of the *glp-1* SCR that is sufficient to confer patterned embryonic expression on a reporter gene. Contained within this fragment are two sub-elements, the GRE and GDE, which are required for translational repression and de-repression of *glp-1* mRNA in early embryos, respectively (Marin and Evans 2003). GLD-1, a STAR-domain RNA-binding protein (Lee and Schedl 2001; Ryder et al. 2004) was shown to bind directly to the GRE, while it was unclear what *trans*-acting factor mediated regulation of *glp-1* through the GDE (Marin and Evans 2003). Surprisingly, the PRE is coincident with the GDE. Furthermore, a five-nucleotide substitution within the PRE prevents expression of the reporter protein without changing the expression of the reporter mRNA (Marin and Evans 2003)(Figure 2.6). This indicates that the PRE is necessary for the post-transcriptional regulation of *glp-1* mRNA during early embryogenesis, and suggests that it may play a functional role in the regulation of other genes during the same period.

To test if the PRE is sufficient to confer patterned expression to a reporter, we generated a strain that expresses GFP from the *pie-1* promoter where six tandem copies of the *mex-3* min sequence were inserted into the *pie-1* 3' UTR (Figure 2.6). This UTR does not cause asymmetric expression of reporter transcripts in embryos (Reese et al. 2000), and as such provides a neutral background to test the role of exogenous cis-acting elements. If the PRE is sufficient to cause POS-1 mediated negative regulation, we expect to see a GFP expression pattern that anti-correlates with POS-1. Instead, the transgenic worms express GFP throughout the syncytial germline, in oocytes, and all

A



B



GLD-1 POS-1
 ACUAAUAUUGUAAG
 AAGCCGAUUGUAAG
 ACUAAUCGGUA AAG
 ACUAAUAUUGUAAG

Figure 2.6

Figure 2.6. The PRE is necessary but not sufficient for patterned regulation of a reporter.

A. 6XP_{RE} GFP reporter strain. Above is a schematic of the reporter construct, below is a widefield epifluorescence image of an adult hermaphrodite. The oocyte region is boxed with a dashed line, and embryos are marked with arrows. GFP expression is uniform throughout early embryos. B. Previously published mutational analysis of the *glp-1* SCR revealed two regulatory elements corresponding to predicted POS-1 and GLD-1 binding sites. Above, schematic representation of POS-1 and GLD-1 binding sites within the *glp-1* SCR. Below, mutations previously made within these binding sites and their corresponding effect on the expression of a reporter, from (Marin & Evans, 2003).

cells of the early embryo. This finding suggests that the PRE alone, and thus POS-1 binding alone, is not sufficient to drive UTR dependent regulation, and may require additional factors for post-transcriptional regulation of specific targets. One likely accessory factor is GLD-1. In addition to the close proximity of the PRE and GLD-1 consensus binding sites in the *glp-1* SCR, GLD-1 is co-expressed in the posterior of the early embryo with POS-1 (Marin and Evans 2003), and embryos depleted of GLD-1 express GLP-1 ectopically in all cells of the early embryo (Marin and Evans 2003), in a pattern identical to that found in *pos-1* mutant embryos (Ogura et al. 2003). Mutations that span both the PRE and GLD-1 consensus binding site in the SCR also cause ubiquitous expression of a reporter throughout the embryo (Marin and Evans 2003), strongly suggesting that POS-1 and GLD-1 coordinate the regulation of *glp-1* in early embryos.

Prevalence of the PRE in C. elegans 3' UTRs

In order to establish a list of candidate POS-1 regulatory targets, we used the pattern matching tool PATSCAN to locate the PRE consensus within all annotated 3' UTRs retrieved from Wormbase release WS180 (Dsouza et al. 1997). Of the 10,201 3' UTRs retrieved, 2,902 (28.4%) contain at least one POS-1 binding site. Because POS-1 is expressed only in early embryos, and *pos-1* mutant embryos have a maternal effect lethal phenotype (Tabara et al. 1999), it is expected that critical POS-1 regulatory targets will be (1) expressed in early embryos and (2) required for early embryogenesis. Of the 2,902 genes that contain a predicted POS-1 binding site, 227 are expressed in one to eight

cell embryos (Baugh et al. 2003) and result in embryos that fail to hatch when silenced by RNAi (Sönnichsen et al. 2005) (Figure 2.7). GLD-1 binding sites are present in 67 of the 227 candidate POS-1 targets, suggesting that interplay between these factors might be a general requirement for POS-1 mediated regulation. We propose that these represent the most likely candidate POS-1 targets relevant to its roles in embryonic patterning.

In addition to *glp-1*, *mex-3*, and *pal-1*, a number of genes required for establishing and maintaining the anterior-posterior axis (*par-1*, *par-3* and *par-5*) (Cuenca et al. 2003) and specifying intestinal fate (*mom-2*, *mom-5*, and *skn-1*) (Rocheleau et al. 1997; Thorpe et al. 1997; Maduro and Pilgrim 1995) contain at least one PRE within their 3' UTR.

Intriguingly, the *pos-1* 3' UTR also has a PRE, suggesting that POS-1 may play a role in regulating its own expression. Neither *apx-1* or *nos-2*, two genes whose expression patterns are perturbed in *pos-1* mutants, contain a predicted POS-1 binding site in their 3' UTRs, suggesting that the role of POS-1 in regulating these genes may be indirect.

To determine if the POS-1 consensus sequence is statistically under-represented, we generated one hundred randomized artificial 3' UTR libraries and determined the frequency of the PRE in each. The artificial libraries were generated using a Markov chain based on the dinucleotide frequencies observed in embryonic 3' UTRs (Figure 2.7).

Dinucleotides were used instead of mononucleotides because a number of dinucleotides occurred more or less frequently than would be expected on the basis of the mononucleotide frequencies alone. The average and standard deviation of the PRE frequency in the 100 artificial libraries establishes the expected number and variance of the PRE associated with random chance. The ratio of the number of actual occurrences

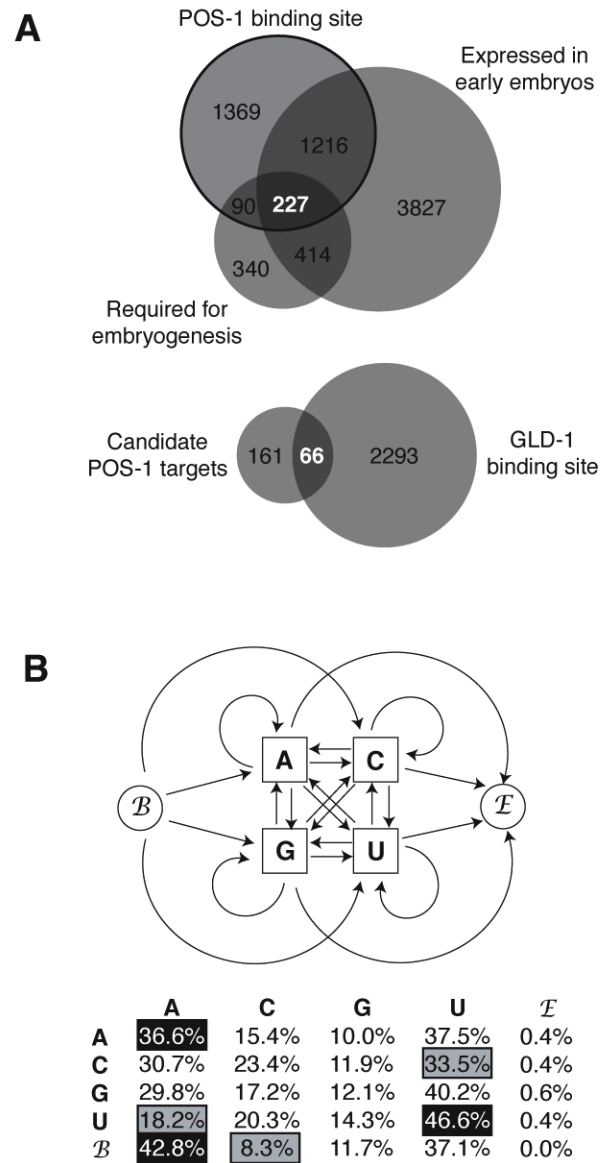


Figure 2.7

Figure 2.7. The PRE is abundant in *C. elegans* 3' UTRs. A. Bioinformatics analysis of genes containing a PRE in their 3' UTR. B. Schematic of the Markov chain used to generate random libraries of 3' UTRs based on the transit probabilities measured from embryonic 3' UTRs. Each shape denotes a state, with the script B and E denoting the begin and end states used to model the beginning and the end of random 3' UTRs, respectively. Arrows represent the transit probabilities from the state the arrow emerges from to the state to which the arrow points. Below are the measured dinucleotide transit probabilities of early embryonic 3' UTRs. Each entry represents the frequency that the state on the left is followed by the state above. Values with a black background are at least five percent above their corresponding mononucleotide frequencies, while values with a grey background are at least five percent below their corresponding mononucleotide frequencies.

to the number of expected occurrences defines the extent to which the PRE is over or under represented.

There are 2,314 occurrences of the PRE in the 3' UTRs of early embryonic transcripts. In contrast, there are 1530 ± 40 PRE occurrences in the set of 100 artificial 3' UTRs. Thus, the PRE is 1.51 ± 0.04 fold over-represented in real embryonic 3'-UTR sequences. This finding is consistent with the hypothesis that POS-1 requires additional specificity factors, such as GLD-1, in order to choose appropriate targets for regulation, or, alternatively, that the network of targets regulated by POS-1 is much larger than previously anticipated.

DISCUSSION

Using *in vitro* biochemical techniques, we have determined that the *C. elegans* TZF protein POS-1 is a sequence specific RNA-binding protein that binds to RNA with novel specificity. POS-1 binds with highest affinity the sequence $UA(U_{2-3})RD(N_{1-3})G$, and with moderate affinity to other sequences containing ARE-like elements. Compared to the related TZF proteins TTP and MEX-5, POS-1 binds to RNA with different specificity, demonstrating that this RNA-binding protein family is capable of binding to diverse sequence determinants. TTP recognizes the sequence UUAUUUAUU (Brewer et al. 2004), while MEX-5 recognizes any sequence with six to eight uridines within an eight nucleotide window (Pagano et al. 2007). The difference of a single amino acid in each zinc finger, required for specific recognition of each of the two adenosines in the TTP recognition element (Hudson et al. 2004), is sufficient to account for the differences

between TTP and MEX-5 specificity (Pagano et al. 2007). Accordingly, the identity of these residues is different in POS-1 compared to either MEX-5 or TTP, suggesting that these amino acids also contribute to the difference in POS-1 specificity. However, POS-1 recognizes three purines compared to two in the TTP binding site, indicating that it must have a third purine recognition site. More work, including structural studies, will be required to understand the molecular basis of differential RNA recognition by POS-1.

The PRE lies within a 34-nucleotide region of the *glp-1* 3' UTR that is required for translational regulation in early embryos (Marin and Evans 2003), and mutations that directly target the PRE abolish expression of a reporter gene, demonstrating that the PRE is a functional *cis*-acting regulatory element required to pattern *glp-1* expression.

However, the PRE is not sufficient to confer POS-1 dependent regulation to an orthogonal 3' UTR, indicating that additional *cis*-acting elements and *trans*-acting factors are required. One likely candidate is GLD-1 (Lee and Schedl 2001; Ryder et al. 2004).

The POS-1 binding site in the *glp-1* 3' UTR is immediately adjacent to a binding site for GLD-1. Mutations of either *pos-1* or *gld-1* result in ectopic GLP-1 expression in all cells of the early embryo, and mutations that disrupt the binding site for either protein lead to aberrant patterning of a reporter (Fig. 6A) (Marin and Evans 2003). It is possible that POS-1 and GLD-1 may function as an RNA-binding complex that binds with enhanced specificity to only a small number of mRNAs. If so, then maternal transcripts that contain binding sites for both proteins are excellent candidate targets that may contribute to aspects of POS-1 mediated development.

The PRE occurs in the 3' UTR of 227 genes expressed in early embryos and required for embryogenesis. These include *mom-2*, *mom-5*, and *skn-1*, all of which are required for the specification of endoderm at the four-cell stage. *mom-2* and *mom-5* encode Wnt and Frizzled homologs, respectively, cell signaling factors that are required to induce endoderm differentiation (Thorpe et al. 1997). *skn-1* encodes a transcription factor that is required to activate the zygotic transcription of genes that promote endoderm fate (Bowerman et al. 1993; Maduro et al. 2001). *pos-1* mutants lack endoderm, suggesting that deregulation of some or all of these mRNAs leads to the failure in endoderm specification. The PRE is also found in the 3' UTR of *par-1*, *par-3*, and *par-5*, genes essential for the establishment of the anterior-posterior axis in the early embryo. POS-1 is not translated until after the anterior-posterior axis has been established, which argues that it cannot play a primary role in establishing this axis. However, the axis must be maintained following establishment, and POS-1 could play a role in this pathway. Maintenance of the anterior-posterior axis requires MEX-5 (Cuenca et al. 2003), which restricts the expression of POS-1 to the posterior of the early embryo. If POS-1 regulates *par* gene expression, then MEX-5 could maintain axis polarization through its molecular function of spatially restricting POS-1.

The data presented here demonstrate that POS-1 binds to RNA with novel specificity compared to homologous TZF domain proteins, indicating that this fold can evolve to recognize diverse RNA sequences. The relative simplicity of the binding consensus suggests a dichotomy between RNA-binding specificity and selection of specific mRNA targets for regulation. The POS-1 recognition element is necessary but

not sufficient to confer patterned expression to a reporter, indicating that additional factors are involved in mRNA target selection. The requirement for recognition by multiple RNA-binding proteins, each with limited sequence specificity, could explain this dichotomy. Our work provides a framework for dissection of the network of maternal transcripts regulated by POS-1 during development, and suggests several interesting candidate targets that can explain the phenotypes observed in *pos-1* mutants.

SUPPLEMENTAL DISCUSSION

The highest affinity sequence identified for POS-1 in this work comes from the *mex-3* 3' UTR to which POS-1 binds with an apparent dissociation constant of 39 nM. If a simple two state model correctly describes the *in vivo* association of POS-1 with its binding sites, POS-1 is uniformly distributed throughout the cytoplasm of the embryo, and POS-1 is present in substantial excess over its binding sites, the intracellular concentration of POS-1 must be at least 400 nM to achieve 90% occupancy of POS-1 binding sites. A *C. elegans* embryo can be approximated by a prolate ellipsoid with a minor axis of 30 μm and a major axis of 60 μm (Hillier et al. 2009), so the approximate volume of an embryo is 750 pL. Thus, to achieve this concentration of POS-1, at least 200 million molecules of POS-1 protein must be present in the one-cell embryo. The total number of mRNA molecules within the embryo is unknown, but estimates for the average number of polyadenylated transcripts within each cell of the 556 cell pre-gastrulation embryo place the number between 12,500 and 25,000 (Hillier et al. 2009). If mRNA turnover does not occur between fertilization and this stage, and all of these

transcripts are maternally supplied, the number of molecules of mRNA present within the one-cell embryo would be between approximately 7 million and 14 million. If *pos-1* mRNA makes up 1% of all mRNAs within the one-cell embryo – making it one of the most abundant mRNAs – there would be a maximum of 140,000 molecules of *pos-1* mRNA present at fertilization. Producing enough POS-1 to reach a concentration of 400 nM would require the translation of 1500 molecules of POS-1 per mRNA molecule within the time it takes to complete one cell division. Based on the assumptions made here, this represents the minimum level of POS-1 translation required to reach this concentration.

This is a very high rate of translation, so some aspect of the POS-1 localization or binding model is likely incorrect. POS-1 may recognize other sequences with an affinity much higher than that observed for the fragment from the *mex-3* 3' UTR, and these sequences may represent the actual regulatory elements through which POS-1 functions. Another possibility is that the RNA-binding domain of POS-1 can have its affinity enhanced by other factors or post-translational modifications, which would reduce the amount of protein required within the embryo. A third possibility is that POS-1 is enriched in specific locations within the embryo, increasing the local concentration of POS-1 in these locations. Another RNA-binding protein with higher affinity or abundance may bind to mRNAs and direct them to these subcellular locations. Thus, POS-1 mediated regulation of specific mRNAs may require additional upstream factors.

In addition to its relatively low affinity for mRNA, the RNA-binding domain of POS-1 also exhibits an *in vitro* calculated association rate constant that is approximately

four orders of magnitude larger than the diffusion rate constant. This implies that binding of the RBD of POS-1 to its mRNA targets could be a slow process. As POS-1 protein is first translated in the one-cell embryo, the slow association of POS-1 with its targets could mean that binding does not occur until the two cell stage or beyond, and that other proteins are required to translationally repress POS-1 targets in the early embryo. Alternatively, the *in vitro* experiments described in this chapter may not fully recapitulate the RNA-binding properties of POS-1. Regardless, the data presented here suggest that there is still much to understand about POS-1 RNA-binding and the role it plays in regulating early embryogenesis.

MATERIALS AND METHODS

Cloning and of purification of POS-1 80-180 pHMTc

The sequence encoding amino acids 80-180 of POS-1 was amplified from the corresponding ORFeome (Open Biosystems) clone via PCR and cloned in frame into the multiple cloning site of pHMTc, a derivative of pMal-c2x (New England Biolabs) that includes an N-terminal 6-his tag and a TEV protease site after MBP (Ryder et al. 2004). This construct was transformed into and expressed in *E. coli* strain BL21 (DE3) Gold (Stratagene). Protein expression was induced with 1 mM IPTG and 100 μ M Zn(OAc)₂. Cells were lysed and purified using an amylose column (New England Biolabs), followed by HiTrap SP and HiTrap Q (GE Healthcare) columns. Purified POS-1 was dialyzed into storage buffer (25 mM Tris-Cl pH 8.0, 25 mM NaCl, 100 μ M Zn(OAc)₂, 2 mM DTT),

concentrated to approximately 100 μ M using a 30,000 MWCO spin concentrator (Vivascience Group), and stored at 4 °C for up to two months.

Preparation of fluorescein-5-thiosemicarbazide labeled RNA

All RNA oligonucleotides used in this study were chemically synthesized by Integrated DNA Technologies. Oligos were fluorescently labeled at the 3' end by periodate oxidation followed by reaction with fluorescein-5-thiosemicarbazide as described (Pagano et al. 2007). Labeled RNA was purified away from unreacted label using a Sephadex G-25 spin column (GE Healthcare). Recovered RNA was assayed by agarose gel electrophoresis and UV-Vis spectrophotometry to determine purity and labeling efficiency, respectively. Labeling efficiencies were typically around 70%.

Fluorescent electrophoretic mobility shift assay

Varying concentrations of purified POS-1 were incubated with 3 nM fluorescently labeled RNA in equilibration buffer for 3 hours at room temperature. Equilibration buffer contained 50 mM Tris pH 8.0, 100 mM NaCl, 0.01 mg/mL tRNA, 0.01% (v/v) IGEPAL, and 100 μ M ZnOAc₂. After three hours, 10 μ L of 0.005% (w/v) bromocresol green in 30% glycerol was added to 100 μ L of each sample and mixed thoroughly. 50 μ L of each mixture was loaded onto a 1% agarose, 1X TB gel and run at room temperature for thirty-five minutes at 120 volts to resolve bound from free RNA. Gels were imaged using a Fuji FLA-5000 laser imager to detect the fluorescently labeled RNA. The fraction of bound RNA was determined by taking the ratio of bound signal to

total signal. This was plotted against the total protein concentration and fit to the Hill equation to determine the apparent dissociation constant:

$$\theta = \frac{([P]_t)^n}{([P]_t)^n + (K_{d,app})^n}$$

where θ is the measured fraction RNA bound, $[P]_t$ is the total protein concentration, $K_{d,app}$ is the apparent dissociation constant, and n is the Hill coefficient. Stoichiometric binding assays were conducted in a similar fashion, except unlabeled RNA was added to a final RNA concentration of 4 μ M, and the data were fit to the following equation (Rambo and Doudna 2004):

$$\theta = \frac{r + K_{d,app} + n - \sqrt{(r + K_{d,app} + n)^2 - 4rn}}{2n}$$

where θ and $K_{d,app}$ are as above, r is the molar ratio of protein to RNA, and n is the molar equivalence point. The $\Delta\Delta G^\circ$ at 20 °C for each of the mutants of *mex-3 min* was determined using the following equation:

$$\Delta\Delta G^\circ = -RT \cdot \ln\left(\frac{K_{d,app} \text{ mex-3 min}}{K_{d,app} \text{ mut}}\right)$$

where R is the gas constant and T is the temperature.

Gel filtration chromatography

The apparent molecular weight of POS-1 80-180 pHMTc relative to standards was determined using a Superdex 200 10/300 GL gel filtration column (GE Healthcare). Approximately 50 μ L of 20 μ M POS-1 was loaded on a column equilibrated with 50 mM

Tris pH 8.0, 300 mM NaCl, and the absorbance at 280 nm was monitored. The apparent molecular weight of POS-1 was determined by comparing the elution volume to the elution volume for molecular weight standards.

Equilibrium analytical ultracentrifugation

Three concentrations of POS-1 (8 μ M, 12 μ M, and 16 μ M) were centrifuged at 9,000, 12,000, and 16,000 rpm, and absorbance across the cell at 280 nm was monitored using a Beckman Optima XL-I analytical ultracentrifuge. Samples were centrifuged until equilibrium had been reached. The resulting traces were fit to the following equation describing a monomer/n-mer transition:

$$A = \sqrt{a^2} \cdot f \cdot e^{\left(\frac{\omega^2}{2RT} \cdot M(1-\bar{v}\rho)(r^2-r_0^2)\right)} + (1-f) \cdot e^{\left(\frac{\omega^2}{2RT} \cdot nM(1-\bar{v}\rho)(r^2-r_0^2)\right)}$$

where A is the measured absorbance, a is a scaling factor, f is the fraction of protein in the monomeric state, ω is the angular velocity, R is the gas constant, T is the temperature, M is the molecular mass of the protein, n is the apparent stoichiometry of the complex, ν is the viscosity of solution, ρ is the density of solution, r is the cuvette radius, and r_0 is a reference distance.

Fluorescence Anisotropy Assay

Varying concentrations of POS-1 and 4 nM fluorescently labeled RNA were equilibrated as described in the electrophoretic mobility shift assay section. Following equilibration, each sample was excited with linearly polarized light, and the parallel and perpendicular fluorescence intensities were measured five times each using a Victor³

1420 Multilabel Counter (Perkin-Elmer) and the apparent polarization determined.

Polarization was plotted against total protein concentration, and fit to the Hill equation as above.

Dissociation rate kinetics

Binding reactions containing 100 nM POS-1 and 4 nM fluorescently labeled *mex-3* min were equilibrated in conditions described above for thirty minutes. Following equilibration, samples were transferred to a 96-well FluoTrak plate (Grenier) containing unlabeled competitor RNA and rapidly mixed. Immediately following addition to competitor RNA, the fluorescence polarization of the sample was measured every twenty seconds as described above. Fluorescence polarization measurements were plotted against time and fit to the following single exponential decay to determine the apparent dissociation rate constant:

$$P = P_0 + A \cdot e^{(-k_{\text{off, app}} \cdot t)}$$

where P is the measured fluorescence polarization, P_0 is the baseline fluorescence polarization, A is the polarization at t=0, $k_{\text{off, app}}$ is the apparent dissociation rate constant, and t is elapsed time.

Bioinformatics

3' UTR sequences were retrieved from release WS180 of Wormbase (www.wormbase.org) using WormMart. To avoid redundancy in the pool of sequences, only the longest 3' UTR for each gene with multiple gene models was used. The pattern

matching tool PATSCAN (Dsouza et al. 1997) was used to identify UTRs that contain POS-1 and/or GLD-1 binding sites. Mononucleotide and dinucleotide frequencies were determined using Perl scripts and standard UNIX text processing tools. Random 3' UTR libraries were constructed via a Markov chain; transit probabilities for each state were determined from the observed dinucleotide distribution and the following equation:

(Newport and Kirschner 1982; Batchelder et al. 1999; Tadros and Lipshitz 2009)

where $p_{i \rightarrow j}$ is the transit probability from nucleotide i to nucleotide j , d_{ij} is the observed frequency of dinucleotide ij , and D_i is the observed frequency of dinucleotides beginning with i .

Error Analysis

For electrophoretic mobility shift and fluorescence polarization assays, the reported value is the average of at least three independent replicates, and the error is \pm one standard deviation. For equilibrium sedimentation experiments, the fit error was determined by plotting the residuals to compare the error distribution of each fitted model.

Reporter Strain Construction

The reporter construct was generated by digesting plasmid pJH 4.52 (a generous gift of Dr. Geraldine Seydoux) with Spe1, and inserting a synthetic DNA duplex containing 3X copies of the *mex-3* min sequence flanked by Spe1 sites. The construct used for all experiments contained two copies of the insert in tandem. The reporter strain

was generated by ballistic transformation using *unc-119* rescue (Praitis et al. 2001). An equal mixture of the reporter plasmid and pDEST-DD03, which harbors the *unc-119* rescuing fragment (a gift of Dr. Marian Walhout), was used to coat the gold particles. Rescued worms were analyzed for GFP expression pattern by fluorescence microscopy with a Zeiss Axioskop microscope.

Chapter III

POS-1 represses translation of *glp-1* by
antagonizing multiple RNA-binding proteins

ABSTRACT

RNA-binding proteins (RBPs) coordinate cell fate specification and differentiation in variety of systems. RNA regulation is critical during oocyte development and early embryogenesis, where RBPs control expression from maternal mRNAs encoding key cell fate determinants. The *Caenorhabditis elegans* Notch homolog *glp-1* coordinates germline progenitor cell proliferation and anterior fate specification in embryos. A network of sequence-specific RBPs is required to pattern GLP-1 translation. Here, we map the cis-regulatory elements that guide *glp-1* regulation by POS-1 and GLD-1. Our results demonstrate that both proteins recognize the *glp-1* 3' untranslated region (UTR) through adjacent, overlapping binding sites, and that POS-1 binding excludes GLD-1 binding. Both factors are required to repress *glp-1* translation in the embryo, suggesting they function in parallel regulatory pathways. Intriguingly, two equivalent POS-1 binding sites are present in the *glp-1* 3' UTR, but only one, which overlaps with a translational de-repression element, is functional *in vivo*. We propose that POS-1 regulates its targets by blocking access of other RBPs to key regulatory sequences, and suggest that antagonism may be a general mechanism that distinguishes functional from non-functional binding sites.

INTRODUCTION

The physiology of a cell is governed by the identity and extent of genes that it expresses. Gene expression is regulated at the transcriptional, post-transcriptional, and post-translational levels. Each aspect is important and necessary to ensure appropriate expression for a given cell type. The relative importance of each varies depending on cell lineage and activity. For example, control of gene expression after transcription is of primary importance during early embryogenesis, when transcription is repressed due to continuous DNA replication (Newport and Kirschner 1982; Batchelder et al. 1999; Tadros and Lipshitz 2009), and in cells where physiology is partially decoupled from the nucleus due to size or morphology, such as neurons (Swanger and Bassell 2011). In such situations, RNA-binding proteins and small RNA-protein complexes function as critical regulatory factors that control mRNA stability, subcellular localization, and translation efficiency to guide cell function.

In order to coordinate gene expression, RNA-regulatory factors must select specific target transcripts from the set of all transcripts present in the cell. A variety of high-throughput methods, including cross-linked immunoprecipitation with deep sequencing (HITS-CLIP, PAR-CLIP) (Licatalosi et al. 2008; Hafner et al. 2010) and RNA immunoprecipitation coupled with array- or deep sequencing-based detection (RIP-CHIP, RIP-SEQ) (Kershner and Kimble 2010; Wright et al. 2011), reveal that some RNA regulatory factors interact with hundreds or thousands of mRNA targets. Though the number is large, it is in many cases less than one might expect based solely upon predictions from corresponding *in vitro* binding studies (Ryder et al. 2004; Bernstein et

al. 2005; Wright et al. 2011). In contrast, functional assays routinely show that only a few RNA targets contribute to the phenotype observed upon loss of the RNA regulatory factor (Lee and Schedl 2001; Hansen et al. 2004; Kalchhauser et al. 2011), suggesting that most binding events do not directly contribute to the major phenotypes associated with loss of the RNA-binding protein. The apparent disparity between *in vitro* binding, *in vivo* binding, and functional data reveals that the basis for specific target selection is not well understood.

Post-transcriptional regulation of mRNA is primarily mediated through cis-regulatory elements present in the untranslated regions (UTRs) of the transcript. These regions are not subject to the evolutionary constraint of the genetic code and thus can primarily serve a regulatory role. A survey of germline expressed genes in the nematode *Caenorhabditis elegans* demonstrates that 3' UTRs are sufficient to drive patterned gene expression in the germline (Merritt et al. 2008). Transgenic worms carrying a fluorescent reporter that only included the 3' UTR of the gene being investigated recapitulated the expression pattern of the endogenous protein in 24 of 30 cases (Merritt et al. 2008). In addition, most transcripts do not contain a unique 5' UTR as 5' end formation in nematodes is mediated primarily by trans-splicing of one of two leader sequences (Allen et al. 2011). Thus, the *C. elegans* germline provides an ideal model system in which the selection of biologically relevant mRNA targets by sequence specific RNA-binding proteins can be examined.

Studies of translationally regulated transcripts in *C. elegans* development reveal that multiple RNA-binding proteins contribute to the regulation of a single mRNA

(Jadhav et al. 2008; Pagano et al. 2009). One example of this is the transcript encoding the *C. elegans* Notch receptor homolog *glp-1*. GLP-1 is a critical regulator of at least two distinct developmental pathways. It is the central regulator of the mitosis to meiosis switch in the distal arm of the gonad (Austin and Kimble 1987), and it is required for specifying endodermal cell fates in the anterior of the four-cell embryo (Mickey et al. 1996). GLP-1 protein is restricted to these locations in the germline and embryo (Figure 3.1); however, *glp-1* mRNA is present throughout the entire gonad and all cells of the early embryo (Evans et al. 1994). *glp-1* translational repression requires at least five different RNA-binding proteins, each repressing translation at different times during development: the STAR-domain protein GLD-1 acts in germ cells entering meiosis (Marin and Evans 2003), the PUF family members PUF-5/6 and PUF-7 act during oogenesis (Lublin and Evans 2007), the KH-domain protein MEX-3 acts after fertilization (Pagano et al. 2009), and both GLD-1 and the CCCH-type tandem zinc finger protein POS-1 are required in the posterior of early embryos (Ogura et al. 2003; Marin and Evans 2003) (Figure 3.1). In addition, the RRM-domain protein SPN-4 is required for translational activation of *glp-1* in the early embryo (Ogura et al. 2003).

Mutational analysis of the *glp-1* 3' UTR has identified a 34-nucleotide region that is sufficient to spatially pattern a reporter (Marin and Evans 2003). This region can be further subdivided into two regulatory elements: the *glp-1* repression element (GRE) and the *glp-1* de-repression element (GDE) (Marin and Evans 2003). Mutations within the GRE result in an expanded reporter expression pattern in the gonad and excess reporter expression in the posterior of early embryos. On the other hand, mutations of the GDE result in either decreased or no reporter expression in either the gonad or embryos

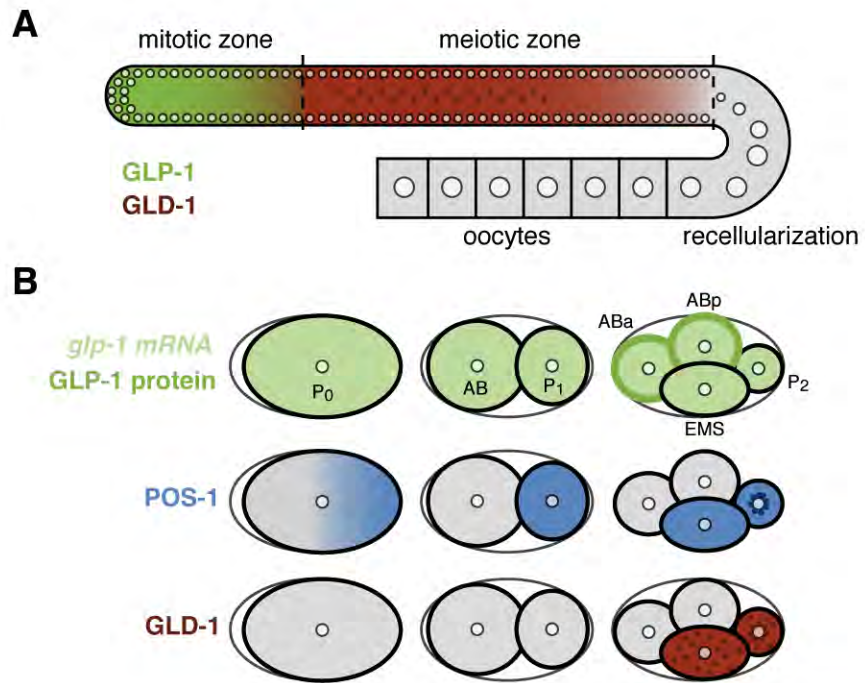


Figure 3.1

Figure 3.1. POS-1 and GLD-1 expression is anti-correlated with GLP-1 expression. A. Schematic of GLD-1 and GLP-1 expression in the germline. GLP-1 expression (green) is restricted to germ cells in mitosis, while GLD-1 (red) is expressed in the meiotic syncytium. GLD-1 is expressed diffusely in the cytoplasm as well as in P-body-like granules. B. Schematic of GLP-1, POS-1 and GLD-1 expression in embryos. Embryos are oriented with the anterior to the left. *glp-1* mRNA (light green) is expressed in all cells of the early embryo, while GLP-1 protein (dark green) is not expressed until the four cell stage, and is restricted to the surface of the two anterior blastomeres (ABa and ABp). POS-1 (blue) is expressed in the posterior cytoplasm of early embryos from the one-cell stage. POS-1 localizes to perinuclear P-granules in the germline P-lineage of embryos. GLD-1 (red) expression begins at the four-cell stage, and it is present in the two posterior blastomeres. GLD-1 is found in the cytoplasm, in perinuclear P-granules in the P-lineage, and in granules distributed throughout both cells of the four-cell stage.

(Marin and Evans 2003). GLD-1 directly associates with the GRE in a sequence specific manner, suggesting that GLD-1 translationally represses *glp-1* through the GRE (Marin and Evans 2003). Given the proximity of the GDE and the GRE, it has been hypothesized that another RNA-binding protein inhibits GLD-1 association with the GRE by binding to the GDE (Marin and Evans 2003) (Figure 3.2). We previously mapped a binding site for POS-1 (PRE; POS-1 recognition element) (Farley et al. 2008) within the GDE that partially overlaps with the GLD-1 binding motif (GBM) (Wright et al. 2011) in the GRE, suggesting that POS-1 could function as the *glp-1* activator through competition with GLD-1 (Figure 3.2). In contrast, mutational studies suggest that POS-1 acts to repress *glp-1* translation, possibly in complex with GLD-1, and a different factor functions as the activator (Marin and Evans 2003). Importantly, none of the mutations that have been made across the GRE and GDE are predicted to exclusively perturb POS-1 binding (Figure 3.2), so it remains unclear what role POS-1 plays in regulating *glp-1*, or if regulation is direct.

POS-1 and GLD-1 exhibit different expression patterns in the early embryo. POS-1 is first observed in the posterior of the one-cell embryo, and is expressed in the posterior blastomeres of the embryo through the four-cell stage. POS-1 is present throughout the cytoplasm of the cells in which it is expressed, as well as perinuclear granules called P-granules in the germline precursor blastomeres (Tabara et al. 1999). In contrast, GLD-1 is first expressed in the posterior blastomeres of the four-cell stage, and it localizes to cytoplasmic granules in the cells in which it is expressed. In addition to the cytoplasmic granules, GLD-1 also localizes to P-granules (Jones et al. 1996) (Figure 3.1).

| | PRE | GBM | PRE | |
|---------------------|---------------------|----------|-------------------------|-------------|
| WT | uuuuucuuauucua | agacuaaa | uuuguaagcu | |
| M7 | uuuuucuuauucua | aga | <u>AGUCCGAACUGA</u> cu | delocalized |
| LS1 | uuuuucuuauucua | aga | <u>AGCCG</u> uuuguaagcu | delocalized |
| LS2 | uuuuucuuauucua | agacuaaa | <u>CGGUA</u> aagcu | none |
| LS3 | uuuuucuuauucua | agacuaaa | uuugu <u>CUCGG</u> | weak |
| LS1LS2 | uuuuucuuauucua | aga | <u>AGCCGCGGUA</u> aagcu | delocalized |
| <hr/> | | | | |
| ΔGBM | uuuuucuuauucua | agac | Caauuuuguaagcu | |
| Δ5'PRE 3'PRE | uuuuuc <u>CC</u> cu | agacuaaa | <u>CC</u> cuguaagcu | |
| Δ5'PRE | uuuuuc <u>CC</u> cu | agacuaaa | uuuguaagcu | |
| Δ3'PRE | uuuuucuuauucua | agacuaaa | <u>CC</u> cuguaagcu | |

Figure 3.2

Figure 3.2. POS-1 and GLD-1 binding sites lie within regulatory elements of the *glp-1* 3' UTR. The *glp-1* repression element (GRE) is adjacent to the *glp-1* de-repression element (GDE, top), and both overlap with predicted POS-1 recognition elements (PRE, blue) and a GLD-1 binding motif (GBM, red). Mutations across this region in the context of a reporter (left, above the line) result in various expression patterns, listed to the right. Below the line, the mutations used in this study are listed.

Thus, POS-1 and GLD-1 may associate with *glp-1* mRNA in different subcellular locations.

To determine the individual roles of both POS-1 and GLD-1 in regulating *glp-1*, we identify separation of function mutations that individually perturb the POS-1 or GLD-1 binding sites *in vitro*, and use them to generate fluorescent reporters to measure their relative contribution to the regulation of *glp-1* expression *in vivo*. The results reveal that POS-1 is indeed a repressor of *glp-1* translation, but it acts independently of GLD-1, and its activity is highly context dependent.

RESULTS

Identification of a second PRE in the glp-1 3' UTR

In Chapter II, we used quantitative *in vitro* binding studies to define the POS-1 consensus recognition element and used the pattern matching tool PatScan (Dsouza et al. 1997) to identify putative PREs in annotated *C. elegans* 3' UTRs based upon this sequence (Farley et al. 2008). This analysis revealed a single PRE in the *glp-1* 3' UTR. However, we noticed that the search pattern does not accurately reflect the thermodynamic measurements in a special case where a single nucleotide deletion compensates for an otherwise deleterious mutation. Specifically, mutation of the adenosine at position six of the PRE to a cytosine reduces binding by 0.6 kcal/mol, while reducing the number of intervening nucleotides between positions six and ten improves binding by the same amount. When we apply the revised pattern to the *glp-1* 3' UTR, we observe a second putative PRE that lies immediately upstream from the GBM, such that

the last nucleotide of the PRE corresponds to the first nucleotide of the GBM. As such, the GBM is flanked by two PREs, each overlapping the GBM by one nucleotide (Figure 3.2).

*POS-1 and GLD-1 recognize the *glp-1* 3' UTR in a sequence-specific manner*

To determine if POS-1 and GLD-1 recognize their predicted binding sites in the *glp-1* 3' UTR, we performed quantitative fluorescent electrophoretic mobility shift (F-EMSA) experiments using purified recombinant POS-1 or GLD-1 RNA binding domains, and a thirty four nucleotide RNA fragment of the *glp-1* 3' UTR that contains both PREs, the GBM, and flanking sequences (Table 3.1, Figure 3.3). The fraction of bound RNA was plotted as a function of total protein concentration and fit to the Hill equation to determine the apparent dissociation constant ($K_{d,app}$) and Hill coefficient (n). By this method, both POS-1 and GLD-1 bind to this fragment with high affinity (POS-1: $K_{d,app} = 19 \pm 2$ nM, $n = 1.3 \pm 0.2$; GLD-1: $K_{d,app} = 70 \pm 10$ nM, $n = 1.0 \pm 0.2$) (Table 3.1, Figure 3.3).

To determine if binding of each protein is dependent on its respective binding sites, we designed RNA oligonucleotides containing mutations in either the PREs or the GBM (Table 3.1, Figure 3.3). Based on previous studies, the mutations were predicted to reduce the affinity of each protein for this sequence by more than 10-fold. Mutation of both PREs results in an 15-fold reduction in affinity for POS-1 ($K_{d,app} = 310 \pm 20$ nM), while mutation of the GBM results in almost complete abrogation of binding ($K_{d,app} > 2000$ nM). Thus, both POS-1 and GLD-1 recognize the *glp-1* 3' UTR in a

Table 3.1. Dissociation constants of POS-1 and GLD-1 for variants of the *glp-1* fragment.

| Identifier | Sequence | POS-1 $K_{d, app}$ (nM) | GLD-1 $K_{d, app}$ (nM) |
|-----------------|----------------------------------|----------------------------|----------------------------|
| <i>glp-1</i> WT | UUUUUCUUAUUCUAGACUAAUAUUGUAAGCU | 19 ± 2 | 70 ± 10 |
| ΔSBE | UUUUUCUUAUUCUAGACCAAUAUUGUAAGCU | 21 ± 1 | > 2000 |
| Δ5' PRE | UUUUUCUCCCUCUAGACUAAUAUUGUAAGCU | 55 ± 4 | 33 ± 7 |
| Δ3' PRE | UUUUUCUUAUUCUAGACUAAACCCUGUAAGCU | 53 ± 1 | 25 ± 3 |
| Δ5' 3' PRE | UUUUUCUCCCUCUAGACUAAACCCUGUAAGCU | 310 ± 20 | 20 ± 3 |

Reported $K_{d, app}$ values are the mean ± one standard deviation of three independent replicates.

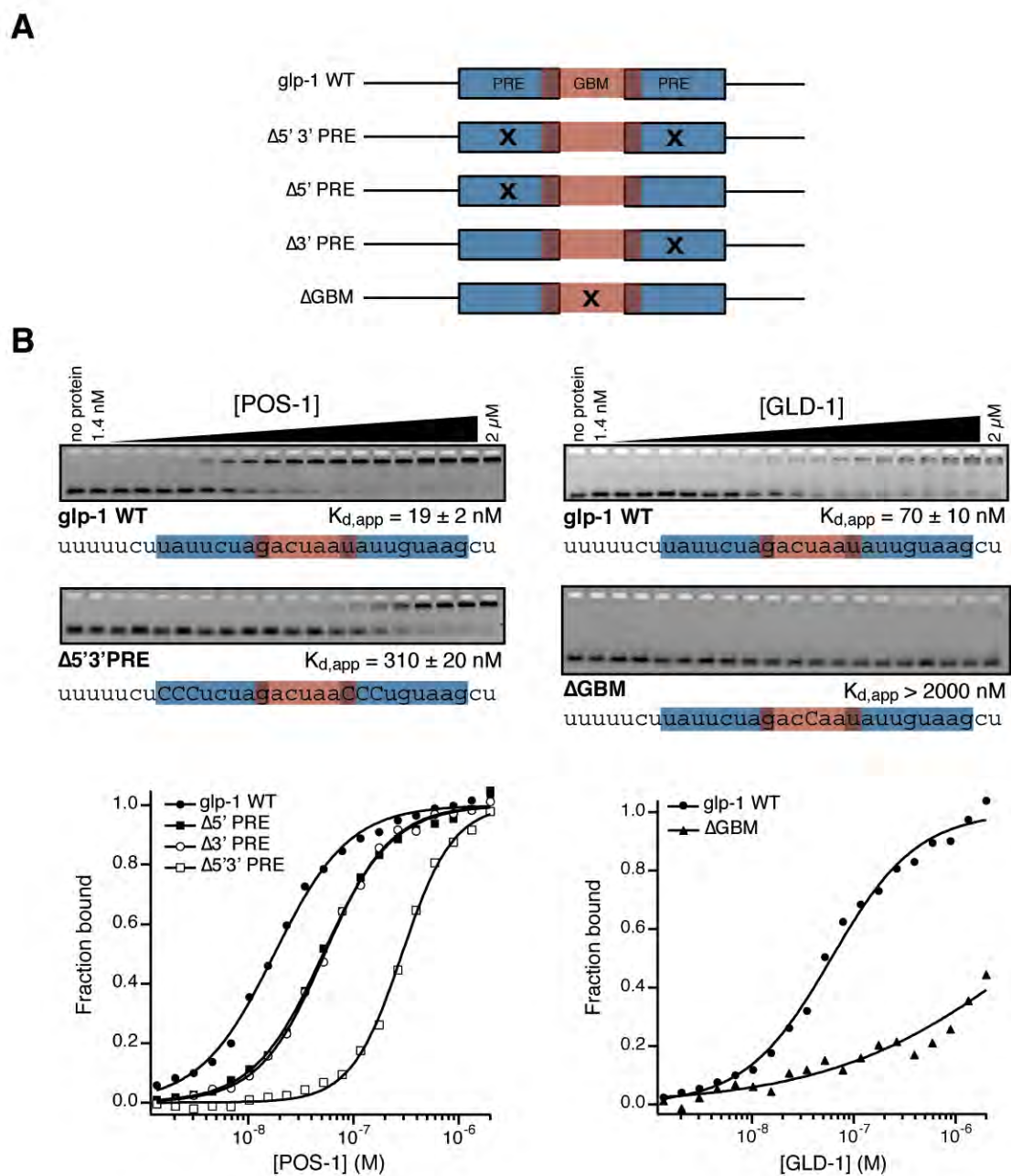


Figure 3.3

Figure 3.3. POS-1 and GLD-1 recognize the *glp-1* 3' UTR in a sequence-specific manner. A. Schematic of mutations in the *glp-1* 3' UTR. PREs are in blue, the GBM is in red, and mutated sites are denoted with an X. B. Fluorescent electrophoretic mobility shift assays with recombinant POS-1 (left) and GLD-1 (right) and fluorescently labeled fragments of the *glp-1* 3' UTR. Top, gel shift images. Bottom, quantifications and fits. POS-1 binds to the wild-type *glp-1* fragment with a $K_{d,app} = 19 \pm 2$ nM, mutation of either PRE individually reduces binding approximately 2.5-fold ($\Delta 5'$ PRE: $K_{d,app} = 54 \pm 4$ nM, $\Delta 3'$ PRE: $K_{d,app} = 52 \pm 1$ nM), while mutation of both reduces binding approximately 15-fold ($\Delta 5' 3'$ PRE: $K_{d,app} = 310 \pm 20$ nM). GLD-1 binds to the wild-type *glp-1* fragment with a $K_{d,app} = 70 \pm 10$ nM, and mutation of the GBM almost completely abrogates binding (Δ GBM: $K_{d,app} > 2000$ nM). Reported $K_{d,app}$ values are the mean \pm standard deviation of three independent replicates.

binding site dependent manner (Figure 3B).

The PREs are equivalent and independent

Given that only one of the PREs in the *glp-1* fragment perfectly matches the previously published consensus, we wanted to establish if POS-1 recognized each binding site. To test the contribution of each site individually, we designed oligonucleotides bearing mutations in only one of the PREs and performed EMSA experiments with recombinant POS-1. Mutation of either PRE resulted in a 2.5-fold reduction in binding compared to the wild-type sequence ($K_{d,app\ WT} = 19 \pm 2\text{ nM}$, $K_{d,app\ \Delta 5' PRE} = 54 \pm 4\text{ nM}$, $K_{d,app\ \Delta 3' PRE} = 52 \pm 1\text{ nM}$), demonstrating that each site is recognized by POS-1 with equivalent affinity (Figure 3.3).

To determine if the two equivalent PREs in the *glp-1* 3' UTR are independent, we analyzed the relationship between the macroscopic and microscopic dissociation constants. An RNA with two equivalent, independent binding sites for a protein where only one site is occupied at any given time should have a macroscopic dissociation constant that is two-fold tighter than the microscopic dissociation constants observed for either in isolation due to statistical effects. We observe a 2.5-fold decrease in affinity between the *glp-1* fragments with one intact PRE versus the wild-type sequence, suggesting that the two PREs are both equivalent and independent. This hypothesis is further supported by the unchanged Hill coefficients of the individual site mutants versus the wild-type sequence ($n_{WT} = 1.3 \pm 0.2$, $n_{\Delta 5' PRE} = 1.7 \pm 0.5$, $n_{\Delta 3' PRE} = 1.5 \pm 0.3$), as cooperative binding to the wild-type sequence is expected to increase the Hill coefficient.

POS-1 antagonizes GLD-1 binding to the glp-1 3' UTR

Given that there are two independent POS-1 binding sites that each overlap with the GBM, we hypothesized that POS-1 binding may inhibit GLD-1 binding to the *glp-1* 3' UTR. To test this hypothesis, we performed *in vitro* competition gel shift experiments with POS-1 and GLD-1. In these experiments, a range of concentrations of the competitor protein was titrated into a fixed, trace concentration of fluorescently labeled RNA, and a fixed subsaturating concentration of the other protein. The differently bound species of RNA were resolved from one another by electrophoresis on a native polyacrylamide gel.

When POS-1 was titrated into samples containing 400 nM GLD-1 and labeled RNA, we observed a decrease in the amount of GLD-1-RNA complex and corresponding formation of faster mobility POS-1-RNA complex (Figure 3.4). At high POS-1 concentration (>300 nM), non-specific POS-1 binding obscured visibility of residual GLD-1 complex (Figure 4A and B). In contrast, when GLD-1 is titrated into samples containing 100 nM POS-1, no GLD-1 complex is observed, even when GLD-1 is present at a concentration that is 10-fold greater than POS-1 (Figure 3.4). No evidence of a slower mobility species is apparent, suggesting that efficient ternary complex formation does not happen. Because the RNA contains two POS-1 binding sites that overlap with the single GLD-1 binding site, and because the apparent affinity of POS-1 is 3.5-fold tighter than the apparent affinity of GLD-1, this is the expected result if the proteins compete for binding to the RNA fragment. Similar results were obtained when the

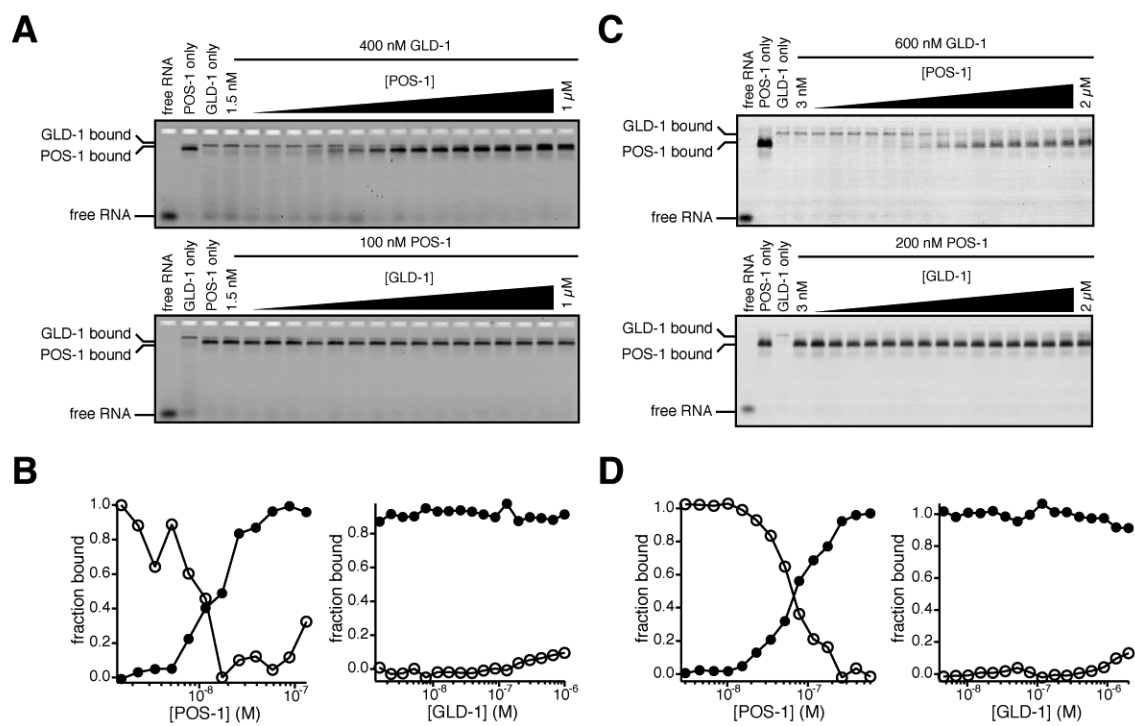


Figure 3.4

Figure 3.4. POS-1 antagonizes GLD-1 to the *glp-1* fragment. A. Gels of competition experiments with POS-1 and GLD-1. Top, POS-1 is titrated into a fixed concentration of GLD-1. Bottom, GLD-1 is titrated into a fixed concentration of POS-1. The mobilities of POS-1-bound RNA, GLD-1-bound RNA, and free RNA are labeled to the left of the gels. B. Quantifications of POS-1 1-206 competition experiments. Left, POS-1 is titrated into a fixed concentration of GLD-1. Right, GLD-1 is titrated into a fixed concentration of POS-1. In both plots, fraction of total RNA bound by each protein is plotted against the concentration of the titrated protein. Filled circles, POS-1 bound RNA. Empty circles, GLD-1 bound RNA. C-D. Competition experiments with POS-1-RBD and GLD-1. The gels in C are labeled as in A, and the plots in D are labeled as in B.

experiment was repeated with a shorter variant of POS-1 containing only the RNA-binding domain (POS-1-RBD, Figure 3.4). This construct binds to the *glp-1* fragment with similar affinity as the longer construct ($K_{d,app, POS-1-RBD} = 30 \pm 17$ nM, Figure 3.5), but provides greater resolution of the POS-1 and GLD-1 bound complexes. POS-1 efficiently competed with GLD-1 for binding to the *glp-1* fragment, while GLD-1 complex formation was strongly inhibited by POS-1. Together, the data show that POS-1 and GLD-1 do not simultaneously bind the *glp-1* 3' UTR fragment, and suggest that POS-1 could inhibit GLD-1 binding to the *glp-1* 3' UTR *in vivo*.

Mutations in the glp-1 3' UTR are specific for either POS-1 or GLD-1

To individually determine the regulatory contribution of direct binding of POS-1 or GLD-1 to the *glp-1* 3' UTR, we designed mutants that would exclusively affect either POS-1 or GLD-1 binding. To determine if the mutations designed to inhibit binding of either POS-1 or GLD-1 to the *glp-1* 3' UTR do not interfere with the binding of the other protein, we used quantitative EMSA to measure the affinity of POS-1 for the Δ GBM version of the *glp-1* fragment, and the affinity of GLD-1 for Δ 5', 3' PRE sequence. Mutation of the GBM does not change the apparent affinity of POS-1 ($K_{d,app WT} = 19 \pm 2$ nM, $K_{d,app \Delta GBM} = 21 \pm 1$ nM), indicating that single-nucleotide GBM mutation does not affect POS-1 binding. Mutation of both PREs results in a 3-fold increase in GLD-1 affinity ($K_{d,app WT} = 70 \pm 10$ nM; $K_{d,app \Delta PREs} = 20 \pm 4$ nM) (Figure 3.6), which is statistically significant ($p = 0.008$), suggesting that mutations in the POS-1 binding sites weakly increase the affinity of GLD-1 for the *glp-1* 3' UTR.

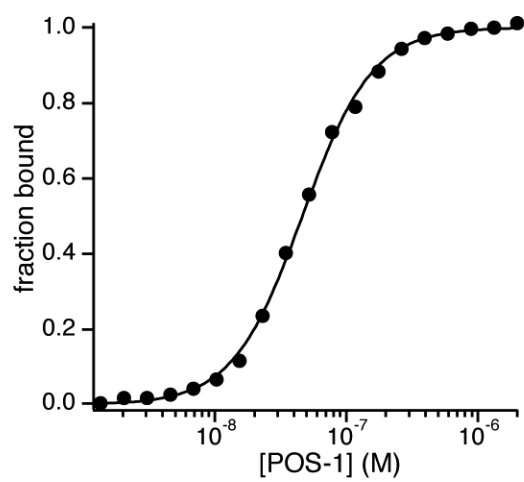
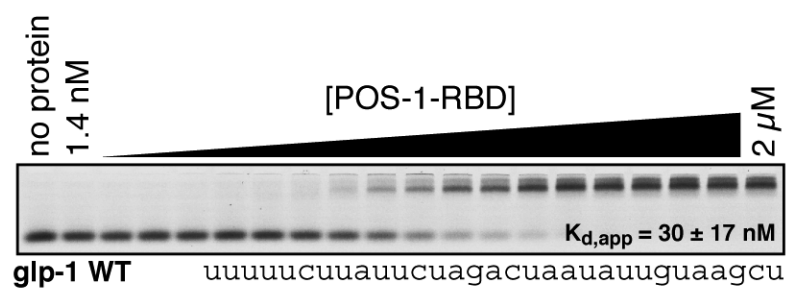


Figure 3.5

Figure 3.5. POS-1-RBD binds the *glp-1* fragment with a similar affinity to that of POS-1 1-206. Top, gel shift image labeled with the POS-1 concentrations (top) in each lane, and the RNA sequence beneath. The reported $K_{d,app}$ is the mean \pm the standard deviation of six separate replicates. Bottom, representative fit of gel shift data to the Hill Equation.

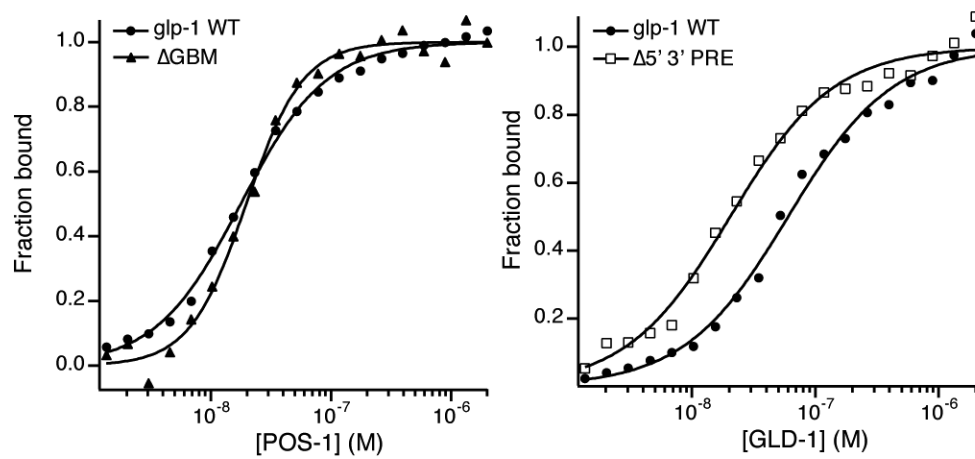


Figure 3.6

Figure 3.6. POS-1 or GLD-1 specific mutations in the *glp-1* 3' UTR. Left, representative EMSA fits for POS-1 binding to *glp-1* WT RNA (filled circles, $K_{d,app} = 19 \pm 2$ nM) and Δ GBM RNA (filled triangles, $K_{d,app} = 21 \pm 1$ nM). Right, representative EMSA fits for GLD-1 binding to *glp-1* WT RNA (filled circles, $K_{d,app} = 70 \pm 10$ nM) and Δ 5' 3' PRE RNA (open squares, $K_{d,app} = 20 \pm 4$ nM). Reported $K_{d,app}$ values are averages \pm one standard deviation of three independent replicates.

Both POS-1 and GLD-1 binding are required for repression of glp-1 translation in embryos

To determine if direct binding of POS-1 to the *glp-1* 3' UTR antagonizes GLD-1 and thus de-represses translation *glp-1* translation in early embryos, we generated green fluorescent protein (GFP) reporters carrying the wild-type *glp-1* 3' UTR or the mutant variations characterized *in vitro* (Table 3.2). To ensure that we were observing only the post-transcriptional regulation of *glp-1*, our reporters used the *mex-5* promoter, a germline promoter with a similar expression pattern to the *glp-1* promoter. The open reading frame of each reporter encodes a protein fusion of GFP and *C. elegans* histone 2B (H2B), which concentrates the fluorescent signal in the nucleus and facilitates cell identification. As the half-life of both GFP and H2B is long relative to oogenesis (Fränd et al. 2005), the open reading frame also contains the mouse ornithine decarboxylase PEST domain, which destabilizes the protein. To enable direct comparison of the reporter expression patterns resulting from different transgenic constructs, the transgenes were integrated site-specifically into chromosome II using Mos1-mediated single copy insertion of transgenes, or MosSCI (Frøkjær-Jensen et al. 2008) (Figure 3.7).

To determine the effects on reporter translation in embryos of disrupting POS-1 or GLD-1 binding to the *glp-1* 3' UTR, we dissected adult worms and observed live four-cell stage embryos. Endogenous GLP-1 is expressed only in the two anterior blastomeres at the four-cell stage (Evans et al. 1994). If direct binding of GLD-1 is required for translational repression, mutating the GBM should result in reporter expression in the posterior as well as the anterior of the four-cell stage embryo. If POS-1 antagonizes

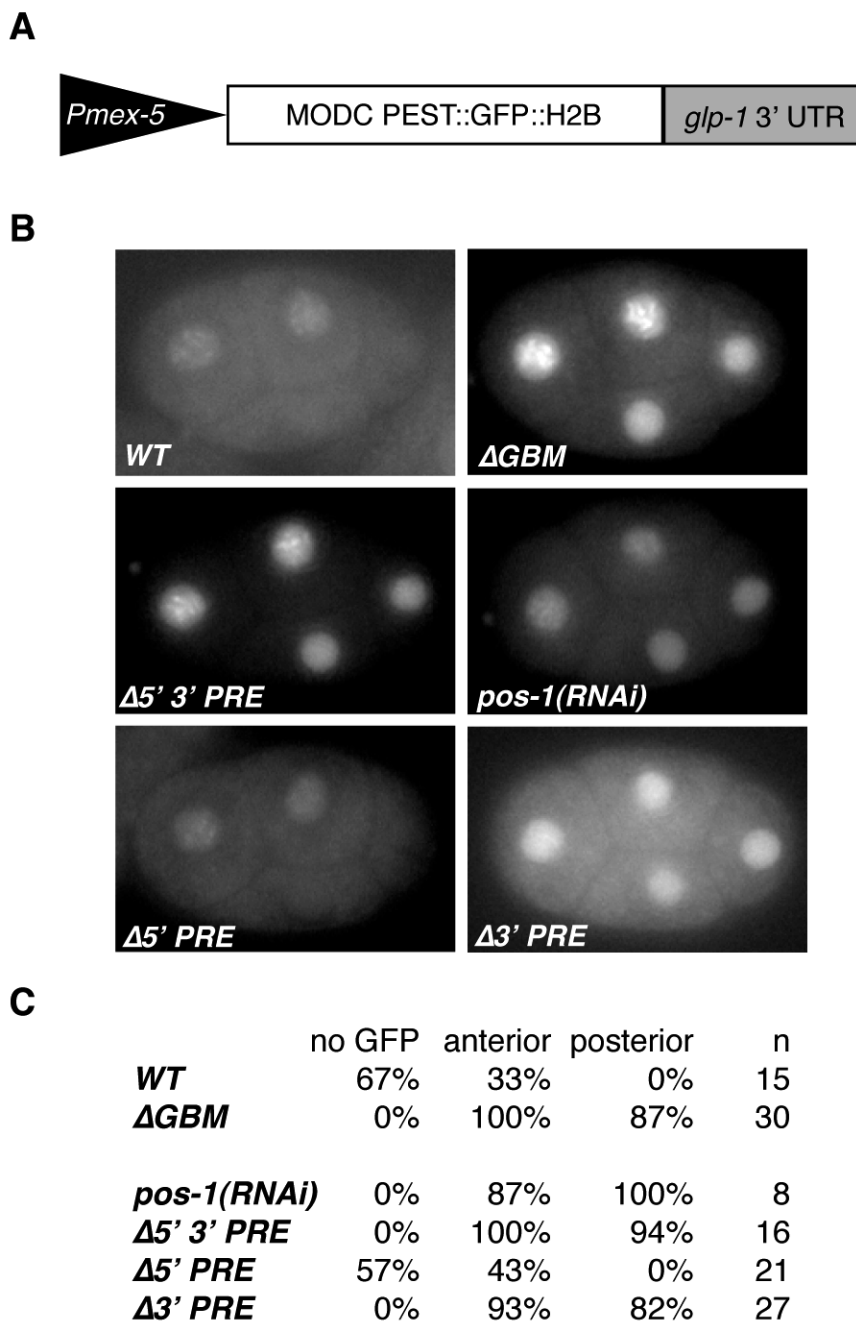


Figure 3.7

Figure 3.7. POS-1 and GLD-1 binding are independently required to repress a *gfp-1* 3' UTR reporter in embryos. A. Schematic of reporter constructs used in this study. B. Representative images of four-cell embryos with the listed reporter or experimental condition. C. Table of results. No GFP indicates the percentage of embryos with no detectable GFP, anterior denotes the percentage of embryos expressing GFP in the two anterior blastomeres, posterior denotes the percentage of embryos expressing GFP in the two posterior blastomeres, and n is the number of embryos observed.

Table 3.2. Transgenic worm strains used in this study.

| <i>glp-1</i> 3' UTR variant | Strain identifier | Genotype |
|------------------------------------|--------------------------|---|
| WT | WRM5 | <i>sprSi5[Pmex-5::MODC PEST:GFP:H2B::glp-1 3'UTR cb-unc-119(+)]</i> II, <i>unc-119(ed3)</i> III |
| Δ GBM | WRM6 | <i>sprSi6[Pmex-5::MODC PEST:GFP:H2B::glp-1 3'UTR(ΔGBM) cb-unc-119(+)]</i> II, <i>unc-119(ed3)</i> III |
| Δ 5' PRE | WRM7 | <i>sprSi7[Pmex-5::MODC PEST:GFP:H2B::glp-1 3'UTR(Δ5' PRE) cb-unc-119(+)]</i> II, <i>unc-119(ed3)</i> III |
| Δ 3' PRE | WRM8 | <i>sprSi8[Pmex-5::MODC PEST:GFP:H2B::glp-1 3'UTR(Δ3' PRE) cb-unc-119(+)]</i> II, <i>unc-119(ed3)</i> III |
| Δ 5' 3' PRE | WRM9 | <i>sprSi9[Pmex-5::MODC PEST:GFP:H2B::glp-1 3'UTR(Δ5' 3' PRE) cb-unc-119(+)]</i> II, <i>unc-119(ed3)</i> III |

GLD-1 binding and thus leads to de-repression, neither anterior nor posterior expression of the reporter should be observed when the PREs are mutated. Expression of GFP in the anterior blastomeres of the four-cell stage is observed in 33% of embryos carrying the WT reporter, while 67% express no detectable GFP (n = 15) at this stage (Figure 3.7). The substantial fraction of transgene-bearing 4-cell stage embryos that lack detectable GFP fluorescence is likely due to the time required for maturation of the GFP chromophore (Reid and Flynn 1997), which is long relative to the duration of the four-cell stage. The anterior expression pattern matches that of endogenous GLP-1 protein, indicating that the *glp-1* 3' UTR is sufficient for appropriate patterning and the reporter mimics the expression pattern of endogenous GLP-1, as previously reported (Evans et al. 1994). Upon treatment with *pos-1(RNAi)*, 100% of embryos express GFP in all cells of the four-cell stage embryo (n = 8, Figure 3.7), suggesting that POS-1 protein plays an inhibitory role in the translational regulation of *glp-1*. POS-1 may have indirect effects on the translation of *glp-1*, so to investigate the requirement for POS-1 binding, we examined a reporter with both PREs mutated. Mutation of both PREs results in a similar expression pattern to embryos carrying the wild-type reporter treated with *pos-1(RNAi)* (94% express in posterior, n = 16, Figure 3.7), suggesting that the PREs are required for translational repression rather than de-repression of *glp-1*, and that POS-1 directly regulates *glp-1*. As the two PREs are equivalent and independent, we hypothesized that the two binding sites are redundant. To test this hypothesis and determine the individual contribution of each PRE, we generated transgenic strains bearing mutations in either the 5' or 3' PRE. Mutating the 5' PRE results in embryos exhibiting wild-type GFP

expression (anterior = 43%, no expression = 57%, n = 21, Figure 3.7), while mutating the 3' PRE results in ubiquitously expressed GFP at the four cell stage (82% express in posterior, n = 27, Figure 3.7). This suggests that despite the thermodynamic equivalence of the PREs *in vitro*, only the 3' PRE is required for POS-1 mediated translational repression of *glp-1*.

Given that POS-1 binding is essential for translational repression of *glp-1*, and that GLD-1 does not bind to the *glp-1* 3' UTR in the presence of POS-1 *in vitro*, we wanted to revisit the role of the GBM in embryos compared to the germline, where GLD-1 is present but POS-1 is not. To test this hypothesis, we generated a transgenic line carrying a mutation in the GBM and observed four-cell stage embryos. Embryos carrying this transgene express GFP in all cells of the early embryo (87% express in posterior, n = 30, Figure 3.7) suggesting that the GBM is also required for translational repression of *glp-1* in the embryo. This matches the previously published results for both the endogenous GLP-1 expression pattern in *gld-1(RNAi)* embryos (Marin and Evans 2003), as well as for reporters carrying mutations in the GRE (Marin and Evans 2003), confirming that GLD-1 acts through the GRE to repress *glp-1* translation in the embryo.

The 3' PRE mutations disrupt the GDE in the germline

The 3' PRE mutation lies entirely within the GDE, which is required for translational activation of *glp-1*, but results in delocalized expression in the embryo. We wanted to determine the effect of the PRE mutations on germline expression of the *glp-1* reporter. The gonads of live worms were observed by widefield fluorescence microscopy

(Figure 3.8). Worms carrying the wild-type reporter have strong GFP expression in the distal end of the germline, with diminishing expression in the syncytial region (95% strong distal expression, 85% weak syncytial expression, n = 20, Figure 3.8). This closely matches the expression of endogenous GLP-1 in the germline. Mutation of the 5' PRE has no effect on expression of GFP, consistent with the effects observed in embryos (86% strong distal expression, 89% weak syncytial expression, n = 28, Figures 6B, 6C). Mutation of both the 5' PRE and 3' PRE, or just the 3' PRE alone results in decreased reporter expression in the germline relative to the wild type reporter ($\Delta 5'$ PRE $3'$ PRE: 70% weak distal, 100% no syncytial, n = 27; $\Delta 3'$ PRE: 96% weak distal, 80% no syncytial, n = 25, Figure 3.8). This apparent decrease is not dependent on POS-1, as *pos-1(RNAi)* has no apparent effect on expression of the wild-type reporter in the germline (100% strong distal, 100% weak syncytial, n = 9, Figure 3.8). Taken together, this suggests that the 3' PRE mutation also disrupts the association of a factor required for activation of *glp-1* translation in the germline. The GBM is required for translational repression of *glp-1* in the syncytial region, as a reporter carrying a mutation in the GBM displays increased expression of GFP in the syncytial region of the germline (85% strong distal, 65% strong syncytial, n = 25, Figure 3.8).

SPN-4 does not directly activate glp-1 translation in the germline

The RRM-containing protein SPN-4 is a potential candidate for the activating factor that operates through the GDE in early embryos. SPN-4 is expressed throughout all cells of the embryo during the four cell-stage (Ogura et al. 2003), and GLP-1

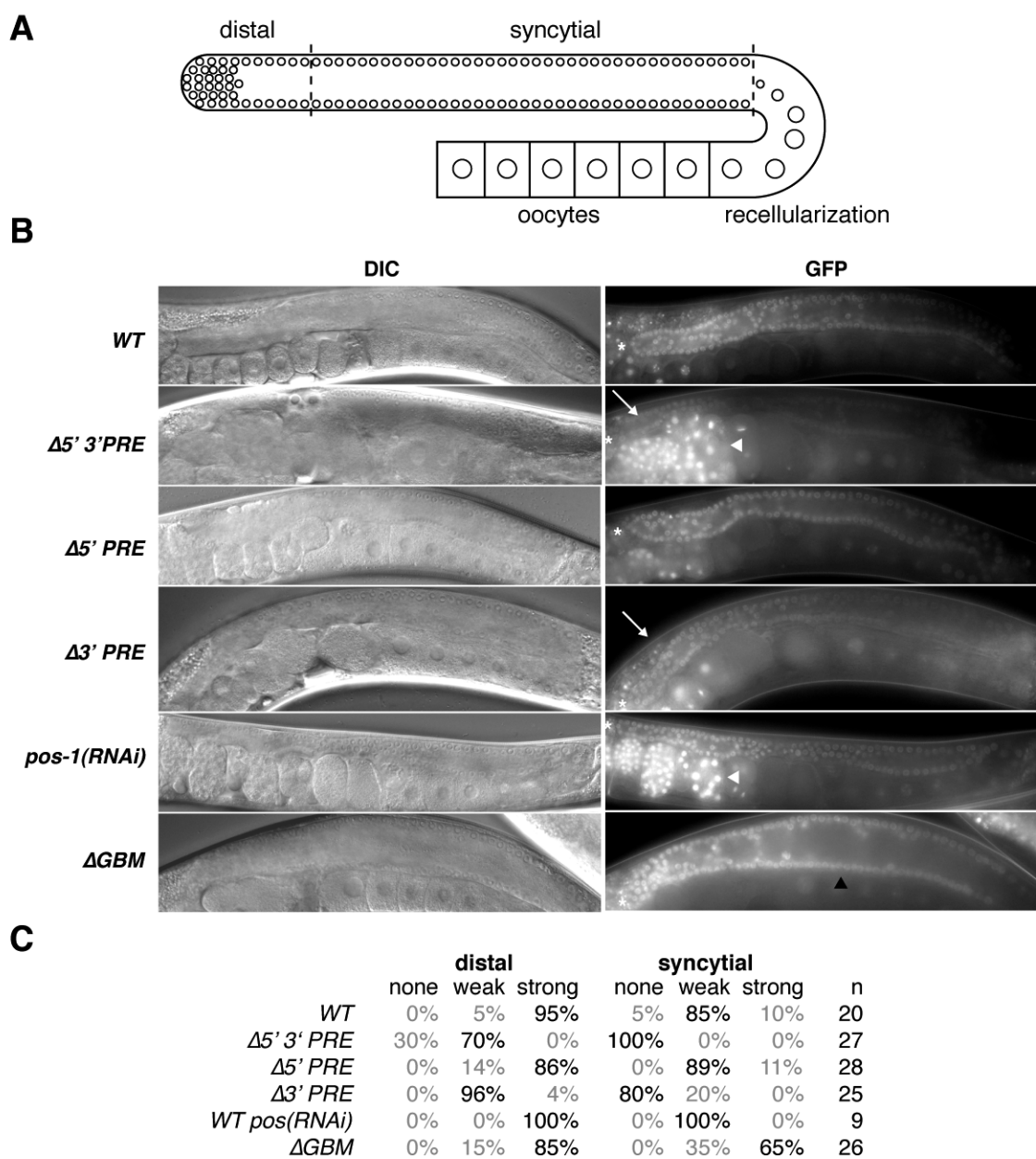


Figure 3.8

Figure 3.8. Mutation of the 3' PRE decreases *glp-1* reporter expression independently of POS-1. A. Schematic of the *C. elegans* gonad. Distal and syncytial regions are labeled. B. Representative differential interference contrast (DIC, left) and widefield fluorescence (GFP, right) images of the listed reporter strains or experimental condition. White asterisks mark the distal end of the gonad, white arrows mark decreased reporter expression in the distal region of $\Delta 5'$ 3' PRE and $\Delta 3'$ PRE reporter strains, white arrowheads mark increased reporter expression in the embryos of $\Delta 5'$ 3' PRE and *pos-1(RNAi)* treated WT reporter strains, and the black arrowhead marks increased reporter expression in the syncytium of the Δ GBM reporter strain. C. Table of results. Distal and syncytial refer to their respective regions of the gonad, and black percentages represent the most common observation for each strain in both the distal and syncytial regions of the gonad.

expression is undetectable in embryos lacking SPN-4 (Ogura et al. 2003). Also, SPN-4 and POS-1 compete for binding to the *nos-2* 3' UTR *in vitro* (Jadhav et al. 2008). To determine if SPN-4 activates expression of the *glp-1* 3'-UTR reporter, we observed the pattern of GFP expression in four cell embryos as a function of SPN-4 knockdown. Anterior expression of GFP is observed in 48% of untreated 4-cell embryos carrying the *glp-1* WT reporter (n = 28, Figure 3.9), while only 11% of *spn-4(RNAi)* 4-cell embryos had detectable GFP in the anterior cells (n = 38). This suggests that SPN-4 is indeed required to activate expression from the *glp-1* 3'UTR reporter in the early embryo. However, this activation is not mediated through the GDE, as SPN-4 knockdown has no effect on GFP expression in 4-cell embryos carrying a reporter bearing the $\Delta 3'$ PRE mutation ($\Delta 3'$ PRE: 100% of embryos have anterior and posterior GFP expression, n = 30; $\Delta 3'$ PRE *spn-4(RNAi)*: 93% of embryos have anterior GFP expression and 97% have posterior GFP expression, n = 30, Figure 3.9). Furthermore, purified SPN-4-RBD does not bind to either the *glp-1* fragment (Figure 3.9) or a longer RNA containing the entire *glp-1* GDE (data not shown) *in vitro*. Thus, SPN-4 is not likely to compete with POS-1 for binding to the GDE. Consistent with this interpretation, SPN-4 was previously observed to interact with a distant element of the *glp-1* 3'-UTR by yeast 3-hybrid. (Ogura et al. 2003).

Taken together, consistent with previous work, we conclude that translational control of *glp-1* requires both repressive and activating elements. GLD-1 binds to the GRE and is required for repression. POS-1 also binds to the GRE *in vitro*, but binding to the GRE has no effect on *glp-1* expression in worms. In contrast, POS-1 binds with

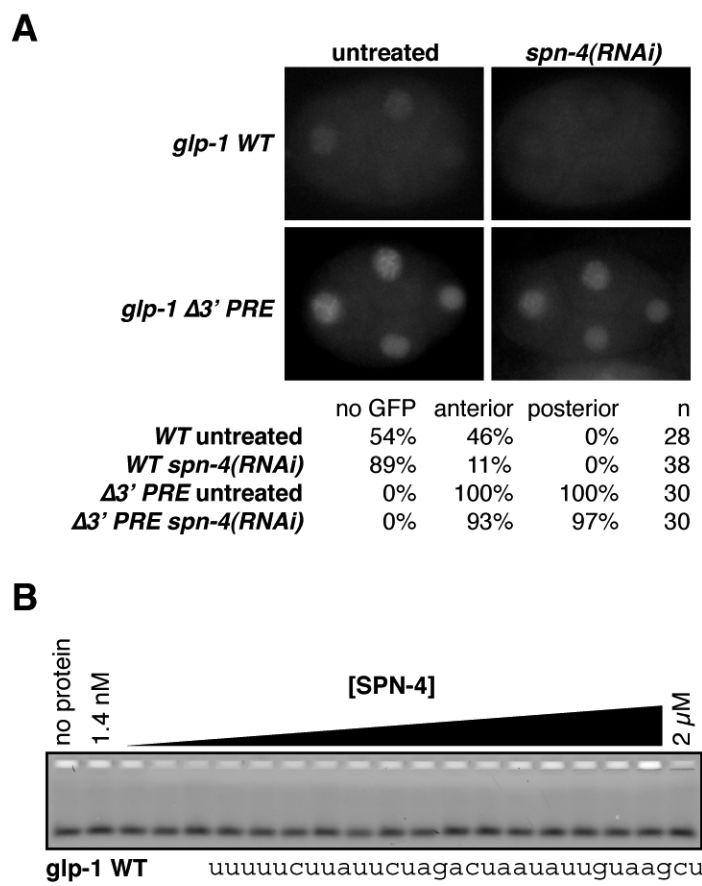


Figure 3.9

Figure 3.9. SPN-4 activates *glp-1* translation indirectly. A. Representative images of four-cell embryos carrying the listed reporter construct (left), and treated as described (top). Bottom, table of results. Anterior and posterior denote the percentage of 4-cell embryos with detectable GFP fluorescence in either the anterior or posterior cells, respective. None denotes the percentage of embryos with no detectable GFP fluorescence. B. Representative gel shift data for SPN-4-RBD. The protein concentration used in each lane is labeled above the gel.

equivalent affinity to the GDE *in vitro*, but binding is required for repression, not activation. We suggest that POS-1 represses translation by competing with an unidentified factor that binds to the GDE to promote translation. This would explain why two thermodynamically equivalent POS-1 binding sites, separated by only five nucleotides, contribute disparately to *glp-1* regulation, and why the PRE is necessary but not sufficient to confer regulation in worms.

DISCUSSION

Our data show that POS-1 and GLD-1 directly and independently repress *glp-1* translation in the early embryo. Independent regulation of *glp-1* may be a consequence of the differing spatial, temporal, and subcellular localization patterns of each protein. As POS-1 is expressed earlier than GLD-1 (Tabara et al. 1999; Jones et al. 1996), *glp-1* is likely regulated only by POS-1 during the one- and two-cell stages. Once GLD-1 translation begins, it may become the primary negative regulator of *glp-1*. As POS-1 antagonizes binding of GLD-1 to the *glp-1* 3' UTR, the handoff to GLD-1 mediated repression would likely require inactivation or turnover of POS-1. Alternatively, the differences in subcellular localization between POS-1 and GLD-1 may be indicative of different mechanisms of *glp-1* repression. In the germline, GLD-1 is present in P-body like granules (Noble et al. 2008), but it is unknown if the GLD-1 granules in the embryo are also P-bodies. As P-bodies are sites of mRNA decapping and turnover, GLD-1 may promote turnover of *glp-1* mRNA in the embryo. Thus, binding to either POS-1 or GLD-

1 may lead to different regulatory outcomes, and competition between these two proteins may be essential for controlling the rate of *glp-1* mRNA turnover.

Our data suggests that a third factor that promotes *glp-1* translation competes with both POS-1 and GLD-1 for binding to the GRE and GDE. One candidate is the atypical cytoplasmic poly(A) polymerase GLD-2, which is expressed throughout the germline and embryo and is required for development (Wang et al. 2002). GLD-2 lacks an RNA-binding domain, and requires the association of an RNA-binding protein to target the polymerase to specific transcripts and extend the poly(A) tail, activating translation. At least two such accessory factors (GLD-3 and RNP-8) are expressed throughout early embryos (Suh et al. 2006; Kim et al. 2010). When GLD-1 or POS-1 levels are low, for example in the anterior blastomeres ABa and ABp, GLD-2 may stimulate polyadenylation of *glp-1* transcripts, leading to translation activation. It is also possible that translational activation of *glp-1* could be mediated by sequence-specific factors that recognize the GDE but do not recruit GLD-2. More work is needed to identify the activating protein and dissect the mechanism of activation.

During early embryogenesis, numerous maternally supplied mRNAs encoding cell-fate specification factors are translationally regulated. Mis-expression of these factors can lead to mis-specification of cell fates, and ultimately embryonic lethality. Several RNA-binding proteins are present in different amounts in each blastomere of the early embryo. Competition between positive and negative regulatory factors would couple the amount of protein produced to the relative ratio of RNA-binding proteins in

each cell, providing a mechanism to drive cell specific expression of cell fate determinants.

Clusters of overlapping binding sites are a required component of this competition model. To identify clusters of binding sites, a detailed understanding of RNA-binding factor motifs is required. Intriguingly, the motifs recognized by POS-1, GLD-1, and a third embryonic RNA-binding protein, MEX-3, include partially overlapping elements (Farley et al. 2008; Wright et al. 2011; Pagano et al. 2009). Thus, clustering of binding sites is a natural outcome of their evolved specificity.

Another feature of the competition model is that RNA-binding proteins exist that do not recruit regulatory machinery, but instead antagonize other factors that do. We propose that POS-1 acts by this mechanism. Our model is supported by the inability of PREs to pattern a reporter (Farley et al. 2008), as well as the context dependence of the PRE required for *glp-1* regulation *in vivo*. PREs are present in over 40% of annotated 3' UTRs (Farley et al. 2008), but we predict many putative binding events serve no biological function. Thus, the apparent dichotomy between biologically relevant specificity and thermodynamic discrimination by RNA-regulatory factors could possibly be explained by antagonistic relationships between multiple proteins.

MATERIALS AND METHODS

Cloning and purification of POS-1, GLD-1 and SPN-4

DNA encoding amino acids 1-206 of POS-1 cloned into the protein expression vector pMAL-c2x (New England Biolabs) was graciously provided by Dr. Tom Evans (U. of Colorado Anschutz Medical Campus, Aurora, CO). The plasmid was transformed into *E. coli* strain BL21 (DE3). Protein expression was induced with 1 mM IPTG and 100 μ M Zn(OAc)₂. The cells were lysed using a microfluidizer, and the lysate was purified using an amylose column (New England Biolabs), followed by a Source Q column (GE Healthcare Life Sciences) and a HiPrep 16/60 Sephacryl S-200 column (GE Healthcare Life Sciences). After the final column, the protein was dialyzed into 25 mM Tris pH 8.0, 25 mM NaCl, 2 mM DTT, 100 μ M Zn(OAc)₂, concentrated to approximately 30 μ M, and used for experiments.

DNA encoding amino acids 50-135 of SPN-4 was cloned into the protein expression vector pHMTc (Ryder et al. 2004) and transformed into *E. coli* strain BL21 (DE3), and induced as POS-1 1-206 above, omitting the zinc acetate. Cells were lysed using a microfluidizer, and the lysate purified using an amylose column (New England Biolabs), followed by a HiTrap SP HP column (GE Healthcare Life Sciences) and a SourceQ column (GE Healthcare Life Sciences). Following the last column, the protein was dialyzed into 25 mM Tris pH 8.0, 25 mM NaCl, 2 mM DTT, concentrated to approximately 20 μ M, and used for experiments.

POS-1-RBD was expressed and purified from pHMTc-POS-1(80-180) as described in (Farley et al. 2008). GLD-1 was expressed and purified from pHMTc-GLD1(135-336) as described in (Ryder et al. 2004).

Fluorescent labeling of RNAs

RNA oligonucleotides were synthesized by Integrated DNA Technologies and fluorescently labeled on their 3' ends via periodate oxidation followed by reaction with fluorescein-5-thiosemicarbazide (Sigma). Unreacted label was purified away via G-25 spin column. A detailed protocol is available in (Pagano et al. 2011).

Fluorescent electrophoretic mobility shift assays

Fluorescent electrophoretic mobility shift assays were essentially performed and analyzed as described in (Pagano et al. 2011). Briefly, 2 nM fluorescently labeled RNA in equilibration buffer (50 mM Tris pH 8.0, 100 mM NaCl, 5 μ M Zn(OAc)₂, 0.01% IGEPAL CA-630, 0.01 mg/mL tRNA) was mixed with varying concentrations of either POS-1 or GLD-1 and equilibrated at room temperature for three hours. The protein-bound and free RNA was then resolved on a 1X TB native 5% polyacrylamide slab gel run at 120 volts for approximately one hour at 4 °C.

Competition assays were performed in essentially the same fashion, except a fixed concentration of POS-1 or GLD-1 was added to the labeled RNA in equilibration buffer to achieve approximately 70% bound RNA in the absence of competitor protein. Then, varying amounts of competitor protein were added to each reaction and incubated for at

least three hours. The reactions were then loaded onto a 1X TB native 5% polyacrylamide slab gel run at 120 volts for three hours at 4°C to resolve POS-1 and GLD-1 bound complexes. Gels were quantified using Image Gauge (Fuji), and the fraction of protein bound RNA was determined by quantifying the ratio of the background corrected pixel intensity of the protein-bound RNA relative to the sum of the background corrected pixel intensities of each RNA species. Two independent replicates of each competition experiment involving POS-1 were performed, and five independent replicates were performed with POS-1-RBD.

Cloning of reporter constructs

The *glp-1* 3' UTR was amplified via PCR using Elongase (Invitrogen) from worm genomic DNA using primers that added the attB2R and attB3 sites to the 5' and 3' ends of the product, respectively. The PCR product was then cloned into pDONR2RP3 using BP Clonase II (Invitrogen). Site-specific mutations in the *glp-1* 3' UTR were introduced via Quickchange mutagenesis using Pfu Turbo. Each of the resulting variants of the *glp-1* 3' UTR was then used in a multi-site gateway reaction with plasmids bearing the *mex-5* promoter (pCM1.111) and MODC PEST::GFP::H2B ORF (pBMF2.7) to generate constructs for integration. The gateway reaction was catalyzed with LR Clonase II Plus (Invitrogen), and the promoter::ORF::3' UTR fusions were cloned into the MosSCI integration vector pCFJ151.

Generation and verification of transgenic strains

Single copy integrated transgenic worms strains were generated by MosSCI (Frøkjær-Jensen et al. 2008). Plasmids bearing the transgene to be integrated were microinjected into the gonads of young adult worms of strain EG4322 along with pharyngeal- and body wall-expressed mCherry markers and a constitutive germline-expressed Mos1 transposase. Prior to injection, worms were maintained at 15 °C on NGM agar plates seeded with *Comamonas* (DA1877). Worms were propagated for two generations, and screened for successful integration by checking for wild-type movement without expression of the mCherry extra-chromosomal array markers. Putative integrants were confirmed by PCR using a transgene specific primer and a worm genome specific primer. These PCR products were then sequenced to validate the mutations in the reporter's 3' UTR.

Imaging of fluorescent reporter strains

Prior to imaging, worms were maintained at 25 °C for at least 24 hours to promote GFP folding. Embryos were obtained by dissecting adult worms and then mounted on 2% agarose pads, and whole worms were paralyzed with 0.4 mM levamisole and mounted on 2% agarose pads. Both DIC and GFP images were collected using a Zeiss Axioskop microscope.

RNAi knockdown

Embryos were harvested from adult worms by treatment with 0.5 N NaOH and 2% Clorox bleach, washed twice with water, and then transferred to NGM plates seeded

with bacteria expressing dsRNA targeting either POS-1 or SPN-4. Worms were maintained at 25 °C and imaged as described above.

ACKNOWLEDGEMENTS

The authors would like to thank Rob Stefani for technical assistance with the competition gel shifts, Dr. Craig Mello for guidance and microinjection training, Dr. Nick Rhind for graciously sharing equipment, and Dr. Ruth Zearfoss and Ebru Kaymak for critical comments and discussion. We also thank Dr. Tom Evans for providing the POS-1 expression plasmid. This work was supported by NIH Grant GM081422 to S.P.R.

Chapter IV

The ramifications of promiscuous RNA

binding specificity

The goal of the research presented within this dissertation is to identify the direct mRNA regulatory targets of the *C. elegans* CCCH-type tandem zinc finger protein POS-1 by determining its *in vitro* RNA-binding specificity and using this information to predict *in vivo* regulatory targets. Embryos lacking POS-1 fail to complete gastrulation, and the arrested embryos exhibit severe defects in three spatially distinct tissue types: pharynx, intestine, and germline precursors. Identifying the regulatory targets of POS-1 could provide insight on the biological pathways that guide cell fate specification events during embryogenesis as well as the mechanism of maintaining the distinction between germline and somatic identities during embryogenesis.

Based on the results presented here, POS-1 recognizes a degenerate sequence that is present in the 3' UTRs of approximately 40% of all genes in the *C. elegans* genome. Two POS-1 binding sites are located within 5 nucleotides of each other within the 3' UTR of *glp-1*, a known regulatory target of POS-1. These binding sites overlap with two previously characterized regulatory elements within the *glp-1* 3' UTR: the *glp-1* repression element (GRE) and *glp-1* de-repression element (GDE). POS-1 recognizes each of these sites with equivalent affinity *in vitro*, but only the 3' POS-1 binding site is required for translational repression in embryos. Mutation of both sites or just the 3' POS-1 binding site results in decreased reporter expression in the germline that is independent of POS-1, suggesting that a factor required to activate translation binds to the *glp-1* 3' UTR through this site. It is therefore possible that POS-1 regulates its targets not by recruiting regulatory machinery, but instead by interfering with the association of other RNA-binding proteins that do.

In addition to sites for POS-1, this region of the *glp-1* 3' UTR contains a binding site for GLD-1, which is required to repress translation of *glp-1* in the distal germline as well as the posterior of the early embryo. GLD-1 acts through the GRE, and mutation of this site within the context of a reporter results in expanded reporter expression in the gonad as well as reporter expression within the posterior of the early embryo. This, together with the requirement for POS-1 and the unidentified activating factor, indicates that multiple RNA-binding proteins each require this fragment of the *glp-1* 3' UTR for different regulatory pathways.

POS-1 specificity

One potential model for how POS-1 regulates a limited subset of targets involves competition with other RNA-binding proteins. POS-1 itself may not recruit translational repression machinery, but instead interfere with the association of translation-activating factors. This model is supported by the two POS-1 binding sites within the *glp-1* 3' UTR. POS-1 binds to each of these sites with an equivalent affinity *in vitro*, but only the 3' POS-1 site is required for translational repression of a reporter *in vivo*. The 3' POS-1 site is coincident with the GDE, mutations of which result in either less or no expression of a reporter relative to a wild-type sequence. Another RNA-binding protein required for translational activation of *glp-1* may act through the GDE, which overlaps with the 3' POS-1 binding site, but not the 5' POS-1 binding site. If this is the case, binding of POS-1 to the 3' site would prevent the association of the activating factor and serve to repress translation of *glp-1*. In this scenario, POS-1 would not need to recruit any negative

regulatory machinery to the *glp-1* 3' UTR to elicit translational repression. Instead, binding to a site required by another regulatory factor and inhibiting its association would suffice.

As a result of this model, the functionality of POS-1 binding sites would depend on which, if any, factors associated nearby. Many of the predicted POS-1 binding sites throughout *C. elegans* 3' UTRs may be non-functional because binding of POS-1 to these sites would not interfere with another regulatory factor. The requirement that POS-1 binding competes with another RNA-binding protein for regulation would greatly increase the target specificity of POS-1, as only a limited subset of POS-1 sites would overlap with or be in close proximity to those of another protein. Understanding which POS-1 binding sites are functional regulatory elements requires the identification of which factors POS-1 competes with for binding to specific sites across *C. elegans* 3' UTRs.

There are two possible groups of RNA-binding proteins that POS-1 may compete with to regulate its targets: the atypical cytoplasmic poly(A) polymerases, and the other CCCH-type tandem zinc finger proteins that are present in *C. elegans*. Poly(A) tails are generally required for translation, as poly(A) binding protein (PABP) recruits the translation initiation factor EIF4G, which binds to the 5' cap binding protein EIF4E and permits the assembly of the 80S ribosome at the 5' end of the transcript. Removal or shortening of the poly(A) tail prevents PABP binding, which in turn prevents the formation of the translation initiation complex and results in translational repression. mRNAs transcribed in the germline that are destined for translation in the embryo are

frequently translationally repressed by this mechanism. Shortened poly(A) tails can be extended in the cytoplasm by cytoplasmic poly(A) polymerases, which permits translation. Two cytoplasmic poly(A) polymerases have been identified in *C. elegans*: GLD-2 and GLD-4 (Wang et al. 2002; Schmid et al. 2009). Both of these proteins are expressed in the germline and in developing oocytes, and each is required for meiotic progression in the distal end of the germline. Unlike most poly(A) polymerases, GLD-2 and GLD-4 contain only a catalytic domain and require the association of an RNA-binding protein to be targeted to mRNA. GLD-2 associates with either the KH-domain protein GLD-3 or the RNA-recognition motif containing protein RNP-8 (Wang et al. 2002; Kim et al. 2010), while GLD-4 associates with either GLD-3 or the novel P-granule component GLS-1 (Rybarska et al. 2009). The mRNAs that either poly(A) polymerase are targeted to changes based upon the associated specificity factor. Any of these poly(A) polymerases or specificity factors may be required to activate the translation of *glp-1* or other potential POS-1 targets in the germline or early embryo. As the RNA-binding specificities of GLD-3, GLS-1, and RNP-8 are unknown, it is difficult to determine if any of these factors have potentially overlapping specificities with POS-1.

Another group of candidate factors that may have positive regulatory roles and compete with POS-1 for binding to mRNA targets are the other members of the *C. elegans* CCCH-type tandem zinc finger family of RNA-binding proteins. The *C. elegans* genome contains 16 CCCH-type TZF proteins (reviewed by Kaymak et al. 2010). 8 of these proteins are required for oogenesis or embryogenesis, while the functions of the other 8 are largely unknown. Members of the CCCH-type TZF family display a range of

distinct but related sequence specificities: the mammalian homologs TTP and Tis11d recognize the sequence UAUUUAUU with high affinity and specificity (Hudson et al. 2004; Brewer et al. 2004), POS-1 recognizes the sequence $UAU_{2-3}RDN_{1-3}G$ with high affinity, and MEX-5 is capable of binding to U-rich sequences (Pagano et al. 2007). It is possible that one of the other *C. elegans* CCCH-type TZF proteins recognizes a similar, but not identical sequence to that of POS-1. Given this similarity in specificity, there would be a significant number of sites that could be recognized by both proteins. If one of the other CCCH-type tandem zinc finger proteins recruits positive regulatory machinery, it could serve as a competitive activator of POS-1 bound mRNAs. One potential TZF protein is DCT-13, which partially suppresses the *gld-1(-)* phenotype of a mitotically proliferating germline tumor (Pinkston-Gosse and Kenyon 2007). As mitotic proliferation of germ cell precursors depends upon GLP-1 activity, tumor suppression may be mediated by a decrease in the levels of GLP-1. The level of GLP-1 expression in either a *dct-13(-)* or *gld-1(-);dct-13(-)* background has not been investigated, and could prove to implicate DCT-13 in *glp-1* translational activation.

It is also possible that neither of these groups of factors acts to activate the translation of *glp-1* or other mRNAs that are also bound by POS-1. The 3' half of the POS-1 recognition sequence is only weakly specific ($UAU_{2-3}R$ versus $DN_{1-3}G$), which may facilitate overlapping binding sites with a variety of RNA-binding proteins with differing specificities. The *glp-1* 3' UTR presents a system in which these competing factors – if any – can be identified. Using the reporter strains already generated over the course of this dissertation research, nucleotide capture experiments can be performed that

specifically recover reporter mRNAs (by designing capturing oligos that are complimentary to the coding sequence of GFP). If the samples are crosslinked via UV or formaldehyde treatment prior to recovery, the suite of proteins associated with the reporter mRNA can be recovered with it, and these can be identified via proteomic methods. Comparing the ensemble of proteins associated with the reporter bearing the wild-type *glp-1* 3' UTR or a version carrying a mutation in the 3' PRE would permit the identification of proteins that differentially associate with the *glp-1* 3' UTR as a function of the 3' PRE. Once identified, *in vitro* biochemical analysis could be performed to verify that the putative RNA-binding proteins associate with the *glp-1* 3' UTR in a POS-1 binding site dependent manner. This approach would permit the identification of new *glp-1* 3' UTR associated factors in a model independent manner, and could offer new insights on the mechanism of POS-1 mediated translational repression.

Translational regulation of *glp-1*

The data presented in this dissertation support the hypothesis that *glp-1* is translationally repressed via two separate pathways during *C. elegans* embryogenesis. One pathway requires POS-1 and is mediated through the 3' PRE in the *glp-1* 3' UTR, while the other pathway requires GLD-1 and is mediated through the SBE. Mutations in either binding site result in reporter expression in the posterior of the four-cell embryo, demonstrating that each protein is required for translational repression in the posterior cells. *In vitro* competition experiments suggest that POS-1 competes with GLD-1 for binding to the *glp-1* 3' UTR. Furthermore, the footprint of GLD-1 when bound to RNA is

more than 20 nucleotides, making it extremely unlikely that both proteins can bind simultaneously *in vitro* or *in vivo*. A co-regulatory complex of POS-1 and GLD-1 requires that both proteins bind together, and is thus not probable. On the other hand, if POS-1 and GLD-1 are redundant translational repressors of *glp-1*, mutation of either the SBE or 3' PRE should not be sufficient to cause de-repression of a reporter as the non-mutated site could still mediate repression. Neither of these models is correct, as both POS-1 mediated repression and GLD-1 mediated repression are required in the embryo.

One possibility for the requirement of both proteins is that each is required in a separate sub-cellular location within the embryo. The reporter constructs used in these experiments contains histone 2B, which directs GFP to the nucleus regardless of where in the cell translation takes place. Immunofluorescence of either POS-1 or GLD-1 reveals that the two proteins have different expression patterns within the cells in which they are expressed. POS-1 is distributed throughout the cytoplasm of the cells in which it is expressed (Tabara et al. 1999), while GLD-1 is predominantly localized to cytoplasmic granules (Jones et al. 1996). This localization is observed in both embryos and the syncytial region of the germline, where GLD-1 is also expressed and serves to translationally repress multiple targets. These cytoplasmic granules contain numerous mRNAs as well as other RNA-binding proteins. Two proteins in particular are required for the stability of the cytoplasmic granules in the germline: CGH-1 and CAR-1 (Noble et al. 2008). Mutation of the gene encoding either one of these proteins results in the disappearance of cytoplasmic granules in the syncytium of the germline, as well as a significant decrease in the expression level of mRNAs that are normally contained within

the granules (Noble et al. 2008; Scheckel et al. 2012). mRNAs targeted to these granules are translationally repressed but also stabilized, suggesting that these granules are for the storage of mRNAs required in embryogenesis. GLD-1 may serve to translationally repress mRNAs that are localized to these storage bodies. *glp-1* mRNA is contained within the granules in the germline, and the *glp-1* 3' UTR is sufficient for localization (Noble et al. 2008). A similar mechanism may be taking place in the embryo, generating two distinct pools of *glp-1* mRNA: one within the cytoplasm, and one within cytoplasmic granules. POS-1 may be the primary factor required for translational repression within the cytoplasm, while GLD-1 may be the primary repressor in granules, which would make each protein necessary for translational repression within embryos.

POS-1 and GLD-1 are just two of the RNA-binding proteins required to translationally repress *glp-1* in the germline and early embryo. PUF-5, PUF-6, and PUF-7 are required in the region of the gonad where oocytes recellularize (Lublin and Evans 2007), OMA-1/2 may be required in maturing oocytes, and MEX-3 is required in one-cell embryos (Pagano et al. 2009). The specificity of MEX-3 is known, and most members of the PUF family of proteins recognize a sequence that contains a UGU trinucleotide with an UA dinucleotide close downstream. Investigating the *glp-1* 3' UTR for these sequences reveals an approximately 100 nucleotide fragment that contains two MEX-3 sites, three candidate PUF-protein sites, and the POS-1 and GLD-1 sites previously described (Figure 4.1). In addition, this fragment is well conserved across closely related nematode species, suggesting that it may be a functional regulatory element. The RNA-binding specificity of OMA-1/2 is unknown, but preliminary experiments suggest that it

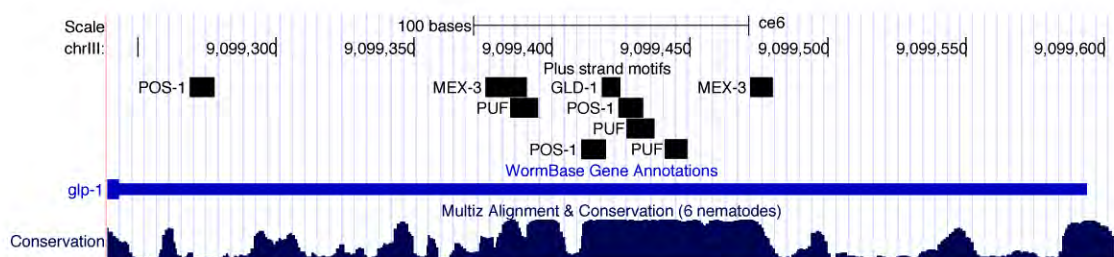


Figure 4.1

Figure 4.1. A cluster of binding sites is present in the *glp-1* 3' UTR. An image from the UCSC genome browser with predicted binding sites for POS-1, GLD-1, MEX-3, and PUF-like consensus sequences annotated. Top, binding site predictions. Middle, *glp-1* 3' UTR annotation. Bottom, nucleotide conservation as determined by the PhastCons algorithm. Higher bars within the histogram indicate greater conservation.

too is a direct translational repressor of *glp-1* and that it acts through an unidentified sequence within this cluster of binding sites (E. Kaymak, unpublished data). Clusters of sites for multiple RNA-binding proteins may be indicative of a functional regulatory element, especially in light of the proposed mechanism of POS-1-mediated translational repression. Identifying clusters of binding sites requires an understanding of the RNA-binding specificity of multiple RNA-binding proteins, as well as a method for identifying individual binding sites within the transcriptome.

Genome-wide identification of functional regulatory elements

The method currently used for identifying binding sites for a given RNA-binding protein relies on an experimentally determined consensus sequence. Two methods are frequently used to generate this consensus: identifying motifs in common among a population of sequences known to bind to the RNA-binding protein in question (using computational tools such as MEME or Cosmo), or measuring the affinities of a series of point mutations of a short, high-affinity binding sequence, followed by setting a threshold for significant mutations. In the second approach, the consensus sequence is defined as the set of all mutations that weaken binding less than the chosen threshold. Each of these approaches identifies the most common or highest affinity binding sites, but RNA-binding proteins can also interact with weak or cryptic sites. For example, the POS-1 consensus failed to identify one of the two high-affinity sites within the *glp-1* 3' UTR because of the presence of two changes relative to the consensus: a C in place of a purine, and a single nucleotide reduction in the length of the binding site. Based on the affinity

measurements described in Chapter II, the C substitution decreases affinity, while the reduction in length increases affinity by a comparable amount. A simple consensus search will miss variations in binding sites such as this.

In at least one case, multiple weak binding sites for an RNA-binding protein in close proximity can elicit a similar regulatory response to a single strong binding site for that same protein. FBF is a direct translational repressor of *cki-2*, which represses the *C. elegans* cyclin E homolog *cye-1* and thus promotes mitosis in the distal end of the germline (Kalchauer et al. 2011). The *cki-2* 3' UTR contains four FBF binding sites within 100 nucleotides. Three of these sites have an affinity that is 10-fold weaker than that observed for FBF in complex with a functional regulatory element from the *gld-1* 3' UTR, while the fourth has an equivalent affinity to that observed for the *gld-1* binding site. Mutation of only the highest affinity FBF site within the context of a reporter results in no detectable change relative to the wild-type sequence, while mutation of all four results in de-repression of the reporter in the distal end of the germline. This implies that all four sites are required for translational repression, not just the highest affinity site. Thus, weak sites must also be taken into consideration when attempting to identify regulatory elements within the 3' UTRs of translationally regulated mRNAs.

To identify both weak and strong sites, I have developed an algorithm based on the pairwise alignment algorithm that uses the quantitative *in vitro* measurements made for panels of point mutations for POS-1, GLD-1, and MEX-3. By determining which nucleotides do not match the optimal nucleotide at each position and summing the measured $\Delta\Delta G^\circ$ for each mismatch, an estimate of the affinity for any arbitrary sequence

can be made. This permits the identification of both weak and strong sites throughout the transcriptome. This method assumes that the effects of individual mutations are independent, and that the effect of multiple mutations is equivalent to the sum of the effects of the individual mutations. This assumption is valid for GLD-1 (Wright et al. 2011). The affinities of 46 different seven-nucleotide sequences were measured. These sequences were the most enriched motifs among a population of sequences recovered by RIP-CHIP from worm extract against GLD-1. This population of motifs was sufficiently diverse to provide multiple opportunities to measure the effect of a point mutation within different sequence contexts. For nearly all point mutations measured, the relative effect of the mutation remained constant regardless of the sequence background it was found in. Mutations may not be independent for other RNA-binding proteins, but this is at least a reasonable hypothesis with regards to GLD-1.

Using this method, clusters of binding sites for POS-1, GLD-1, and MEX-3 can be identified throughout the entire *C. elegans* transcriptome and subsequently tested for regulatory activity via the generation of transgenic reporter strains. Clusters of binding sites may be more informative than individual binding sites alone, thus enabling the identification of regulatory regions of translationally regulated transcripts. This will enable a greater understanding of the mechanism and function of POS-1 and other RNA-binding protein mediated translational regulation in *C. elegans*, and this strategy could conceivably be extended to other species as well.

Bibliography

- Allen MA, Hillier LW, Waterston RH, Blumenthal T. 2011. A global analysis of *C. elegans* trans-splicing. *Genome Res* **21**: 255–264.
- Ariz M, Mainpal R, 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.
- Artavanis-Tsakonas S, Rand MD, Lake RJ. 1999. Notch signaling: cell fate control and signal integration in development. *Science* **284**: 770–776.
- Austin J, Kimble J. 1987. *glp-1* is required in the germ line for regulation of the decision between mitosis and meiosis in *C. elegans*. *Cell* **51**: 589–599.
- Barton MK, Kimble J. 1990. *fog-1*, a regulatory gene required for specification of spermatogenesis in the germ line of *Caenorhabditis elegans*. *Genetics* **125**: 29–39.
- Batchelder C, Dunn MA, Choy B, Suh Y, Cassie C, Shim EY, Shin TH, Mello C, Seydoux G, Blackwell TK. 1999. Transcriptional repression by the *Caenorhabditis elegans* germ-line protein PIE-1. *Genes Dev* **13**: 202–212.
- Baugh LR, Hill AA, Slonim DK, Brown EL, Hunter CP. 2003. Composition and dynamics of the *Caenorhabditis elegans* early embryonic transcriptome. *Development* **130**: 889–900.
- Bernstein D, Hook B, Hajarnavis A, Opperman L, Wickens M. 2005. Binding specificity and mRNA targets of a *C. elegans* PUF protein, FBF-1. *RNA* **11**: 447–458.
- Berry LW, Westlund B, Schedl T. 1997. Germ-line tumor formation caused by activation of *glp-1*, a *Caenorhabditis elegans* member of the Notch family of receptors. *Development* **124**: 925–936.
- Beuck C, Szymczyna BR, Kerkow DE, Carmel AB, Columbus L, Stanfield RL, Williamson JR. 2010. Structure of the GLD-1 homodimerization domain: insights into STAR protein-mediated translational regulation. *Structure* **18**: 377–389.
- Bezares-Calderón LA, Becerra A, Salinas LS, Maldonado E, Navarro RE. 2010. Bioinformatic analysis of P granule-related proteins: insights into germ granule evolution in nematodes. *Dev Genes Evol* **220**: 41–52.
- Blackshear PJ, Lai WS, Kennington EA, Brewer G, Wilson GM, Guan X, Zhou P. 2003. Characteristics of the interaction of a synthetic human tristetraprolin tandem zinc finger peptide with AU-rich element-containing RNA substrates. *J Biol Chem* **278**:

19947–19955.

- Boag PR, Nakamura A, Blackwell TK. 2005. A conserved RNA-protein complex component involved in physiological germline apoptosis regulation in *C. elegans*. *Development* **132**: 4975–4986.
- Bowerman B, Draper BW, Mello CC, Priess JR. 1993. The maternal gene *skn-1* encodes a protein that is distributed unequally in early *C. elegans* embryos. *Cell* **74**: 443–452.
- Boyd L, Guo S, Levitan D, Stinchcomb DT, Kemphues KJ. 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.
- Brewer BY, Malicka J, Blackshear PJ, Wilson GM. 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.
- Carballo E, Lai WS, Blackshear PJ. 1998. Feedback inhibition of macrophage tumor necrosis factor- α production by tristetraprolin. *Science* **281**: 1001–1005.
- Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, et al. 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**: 1393–1406.
- Ciosk R, DePalma M, Priess JR. 2006. Translational regulators maintain totipotency in the *Caenorhabditis elegans* germline. *Science* **311**: 851–853.
- Clifford R, Lee MH, Nayak S, Ohmachi M, Giorgini F, Schedl T. 2000. FOG-2, a novel F-box containing protein, associates with the GLD-1 RNA binding protein and directs male sex determination in the *C. elegans* hermaphrodite germline. *Development* **127**: 5265–5276.
- Crittenden SL, Bernstein DS, Bachorik JL, Thompson BE, Gallegos M, Petcherski AG, Moulder G, Barstead R, Wickens M, Kimble J. 2002. A conserved RNA-binding protein controls germline stem cells in *Caenorhabditis elegans*. *Nature* **417**: 660–663.
- Crittenden SL, Troemel ER, Evans TC, Kimble J. 1994. GLP-1 is localized to the mitotic region of the *C. elegans* germ line. *Development* **120**: 2901–2911.
- Cuenca AA, Schetter A, Aceto D, Kemphues K, Seydoux G. 2003. Polarization of the *C. elegans* zygote proceeds via distinct establishment and maintenance phases. *Development* **130**: 1255–1265.

- D'Agostino I, Merritt C, Chen P-L, Seydoux G, Subramaniam K. 2006. Translational repression restricts expression of the *C. elegans* Nanos homolog NOS-2 to the embryonic germline. *Dev Biol* **292**: 244–252.
- Detwiler MR, Reuben M, Li X, Rogers E, Lin R. 2001. Two zinc finger proteins, OMA-1 and OMA-2, are redundantly required for oocyte maturation in *C. elegans*. *Dev Cell* **1**: 187–199.
- Doniach T. 1986. Activity of the sex-determining gene *tra-2* is modulated to allow spermatogenesis in the *C. elegans* hermaphrodite. *Genetics* **114**: 53–76.
- Draper BW, Mello CC, Bowerman B, Hardin J, Priess JR. 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, Overbeek R. 1997. Searching for patterns in genomic data. *Trends Genet* **13**: 497–498.
- Eckmann CR, Crittenden SL, Suh N, Kimble J. 2004. GLD-3 and control of the mitosis/meiosis decision in the germline of *Caenorhabditis elegans*. *Genetics* **168**: 147–160.
- Etemad-Moghadam B, Guo S, Kemphues KJ. 1995. Asymmetrically distributed PAR-3 protein contributes to cell polarity and spindle alignment in early *C. elegans* embryos. *Cell* **83**: 743–752.
- Evans TC, Crittenden SL, Kodoyianni V, Kimble J. 1994. Translational control of maternal *glp-1* mRNA establishes an asymmetry in the *C. elegans* embryo. *Cell* **77**: 183–194.
- Farley BM, Pagano JM, Ryder SP. 2008. RNA target specificity of the embryonic cell fate determinant POS-1. *RNA* **14**: 2685–2697.
- Farley BM, Ryder SP. 2008. Regulation of maternal mRNAs in early development. *Crit Rev Biochem Mol Biol* **43**: 135–162.
- Fehon RG, Kooch PJ, Rebay I, Regan CL, Xu T, Muskavitch MA, Artavanis-Tsakonas S. 1990. Molecular interactions between the protein products of the neurogenic loci Notch and Delta, two EGF-homologous genes in *Drosophila*. *Cell* **61**: 523–534.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806–811.

- Francis R, Barton MK, Kimble J, Schedl T. 1995a. *gld-1*, a tumor suppressor gene required for oocyte development in *Caenorhabditis elegans*. *Genetics* **139**: 579–606.
- Francis R, Maine E, Schedl T. 1995b. Analysis of the multiple roles of *gld-1* in germline development: interactions with the sex determination cascade and the *glp-1* signaling pathway. *Genetics* **139**: 607–630.
- Frand AR, Russel S, Ruvkun G. 2005. Functional genomic analysis of *C. elegans* molting. *PLoS Biol* **3**: e312.
- Frøkjaer-Jensen C, Davis MW, Hopkins CE, Newman BJ, Thummel JM, Olesen S-P, Grunnet M, Jorgensen EM. 2008. Single-copy insertion of transgenes in *Caenorhabditis elegans*. *Nat Genet* **40**: 1375–1383.
- Ghildiyal M, Zamore PD. 2009. Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**: 94–108.
- Gibert MA, Starck J, Beguet B. 1984. Role of the gonad cytoplasmic core during oogenesis of the nematode *Caenorhabditis elegans*. *Biol Cell* **50**: 77–85.
- Glisovic T, Bachorik JL, Yong J, Dreyfuss G. 2008. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* **582**: 1977–1986.
- Gomes JE, Encalada SE, Swan KA, Shelton CA, Carter JC, 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.
- Grant B, Hirsh D. 1999. Receptor-mediated endocytosis in the *Caenorhabditis elegans* oocyte. *Mol Biol Cell* **10**: 4311–4326.
- Gumienny TL, Lambie E, Hartwig E, Horvitz HR, Hengartner MO. 1999. Genetic control of programmed cell death in the *Caenorhabditis elegans* hermaphrodite germline. *Development* **126**: 1011–1022.
- Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp A-C, Munschauer M, et al. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**: 129–141.
- Hansen D, Wilson-Berry L, Dang T, Schedl T. 2004. Control of the proliferation versus meiotic development decision in the *C. elegans* germline through regulation of GLD-1 protein accumulation. *Development* **131**: 93–104.

- Hasegawa E, Karashima T, Sumiyoshi E, Yamamoto M. 2006. *C. elegans* CPB-3 interacts with DAZ-1 and functions in multiple steps of germline development. *Dev Biol* **295**: 689–699.
- Henderson ST, Gao D, Lambie EJ, Kimble J. 1994. lag-2 may encode a signaling ligand for the GLP-1 and LIN-12 receptors of *C. elegans*. *Development* **120**: 2913–2924.
- Hillier LW, Reinke V, Green P, Hirst M, Marra MA, Waterston RH. 2009. Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. *Genome Res* **19**: 657–666.
- Hirsh D, Oppenheim D, Klass M. 1976. Development of the reproductive system of *Caenorhabditis elegans*. *Dev Biol* **49**: 200–219.
- Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE. 2004. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat Struct Mol Biol* **11**: 257–264.
- Hunter CP, 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.
- Jadhav S, Rana M, Subramaniam K. 2008. Multiple maternal proteins coordinate to restrict the translation of *C. elegans* nanos-2 to primordial germ cells. *Development* **135**: 1803–1812.
- Jones AR, Francis R, 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.
- Jungkamp A-C, Stoeckius M, Mecenas D, Grün D, Mastrobuoni G, Kempa S, Rajewsky N. 2011. In vivo and transcriptome-wide identification of RNA binding protein target sites. *Mol Cell* **44**: 828–840.
- Kalchauer I, Farley BM, Pauli S, Ryder SP, Ciosk R. 2011. FBF represses the Cip/Kip cell-cycle inhibitor CKI-2 to promote self-renewal of germline stem cells in *C. elegans*. *EMBO J* **30**: 3823–3829.
- Kaye JA, Rose NC, Goldsworthy B, Goga A, L'Etoile ND. 2009. A 3'UTR pumilio-binding element directs translational activation in olfactory sensory neurons. *Neuron* **61**: 57–70.
- Kaymak E, Wee LM, Ryder SP. 2010. Structure and function of nematode RNA-binding

- proteins. *Curr Opin Struct Biol* **20**: 305–312.
- Keene JD. 2007. RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* **8**: 533–543.
- Keene JD, Komisarow JM, Friedersdorf MB. 2006. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat Protoc* **1**: 302–307.
- Kelly WG, Xu S, Montgomery MK, Fire A. 1997. Distinct requirements for somatic and germline expression of a generally expressed *Caenorhabditis elegans* gene. *Genetics* **146**: 227–238.
- Kershner AM, Kimble J. 2010. Genome-wide analysis of mRNA targets for *Caenorhabditis elegans* FBF, a conserved stem cell regulator. *Proc Natl Acad Sci USA* **107**: 3936–3941.
- Kim KW, Wilson TL, Kimble J. 2010. GLD-2/RNP-8 cytoplasmic poly(A) polymerase is a broad-spectrum regulator of the oogenesis program. *Proc Natl Acad Sci USA* **107**: 17445–17450.
- Kimble J, Crittenden SL. 2007. Controls of germline stem cells, entry into meiosis, and the sperm/oocyte decision in *Caenorhabditis elegans*. *Annu Rev Cell Dev Biol* **23**: 405–433.
- Kimble J, Hirsh D. 1979. The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Dev Biol* **70**: 396–417.
- Kimble JE, White JG. 1981. On the control of germ cell development in *Caenorhabditis elegans*. *Dev Biol* **81**: 208–219.
- Koh YY, Wang Y, Qiu C, Opperman L, Gross L, Tanaka Hall TM, Wickens M. 2011. Stacking interactions in PUF-RNA complexes. *RNA* **17**: 718–727.
- Kraemer B, Crittenden S, Gallegos M, Moulder G, Barstead R, Kimble J, Wickens M. 1999. NANOS-3 and FBF proteins physically interact to control the sperm-oocyte switch in *Caenorhabditis elegans*. *Curr Biol* **9**: 1009–1018.
- Kuersten S, Goodwin EB. 2003. The power of the 3' UTR: translational control and development. *Nat Rev Genet* **4**: 626–637.
- Kuwabara PE, Okkema PG, Kimble J. 1992. *tra-2* encodes a membrane protein and may mediate cell communication in the *Caenorhabditis elegans* sex determination pathway.

- Mol Biol Cell* **3**: 461–473.
- Lai WS, Carballo E, Strum JR, Kennington EA, Phillips RS, Blackshear PJ. 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 WS, Carrick DM, Blackshear PJ. 2005. Influence of nonameric AU-rich tristetraprolin-binding sites on mRNA deadenylation and turnover. *J Biol Chem* **280**: 34365–34377.
- Lasko P. 2000. The drosophila melanogaster genome: translation factors and RNA binding proteins. *J Cell Biol* **150**: F51–6.
- Lee M-H, Schedl T. 2006. RNA-binding proteins. *WormBook* 1–13.
- Lee MH, Schedl T. 2001. Identification of in vivo mRNA targets of GLD-1, a maxi-KH motif containing protein required for *C. elegans* germ cell development. *Genes Dev* **15**: 2408–2420.
- Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**: 174–187.
- Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, Clark TA, Schweitzer AC, Blume JE, Wang X, et al. 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**: 464–469.
- Lublin AL, Evans TC. 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.
- Lunde BM, Moore C, Varani G. 2007. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8**: 479–490.
- Maduro M, Pilgrim D. 1995. Identification and cloning of unc-119, a gene expressed in the *Caenorhabditis elegans* nervous system. *Genetics* **141**: 977–988.
- Maduro MF. 2010. Cell fate specification in the *C. elegans* embryo. *Dev Dyn* **239**: 1315–1329.
- Maduro MF, Meneghini MD, Bowerman B, Broitman-Maduro G, Rothman JH. 2001. Restriction of mesendoderm to a single blastomere by the combined action of SKN-1

- and a GSK-3 β homolog is mediated by MED-1 and -2 in *C. elegans*. *Mol Cell* **7**: 475–485.
- Mango SE, Thorpe CJ, Martin PR, Chamberlain SH, Bowerman B. 1994. Two maternal genes, *apx-1* and *pie-1*, are required to distinguish the fates of equivalent blastomeres in the early *Caenorhabditis elegans* embryo. *Development* **120**: 2305–2315.
- Marin VA, Evans TC. 2003. Translational repression of a *C. elegans* Notch mRNA by the STAR/KH domain protein GLD-1. *Development* **130**: 2623–2632.
- McCarter J, Bartlett B, Dang T, Schedl T. 1999. On the control of oocyte meiotic maturation and ovulation in *Caenorhabditis elegans*. *Dev Biol* **205**: 111–128.
- McNally KL, McNally FJ. 2005. Fertilization initiates the transition from anaphase I to metaphase II during female meiosis in *C. elegans*. *Dev Biol* **282**: 218–230.
- Mello CC, Draper BW, Priess JR. 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 CC, Kramer JM, Stinchcomb D, Ambros V. 1991. Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences. *EMBO J* **10**: 3959–3970.
- Mello CC, Schubert C, Draper B, Zhang W, Lobel R, Priess JR. 1996. The PIE-1 protein and germline specification in *C. elegans* embryos. *Nature* **382**: 710–712.
- Merritt C, Rasoloson D, Ko D, Seydoux G. 2008. 3' UTRs are the primary regulators of gene expression in the *C. elegans* germline. *Curr Biol* **18**: 1476–1482.
- Mickey KM, Mello CC, Montgomery MK, Fire A, Priess JR. 1996. An inductive interaction in 4-cell stage *C. elegans* embryos involves APX-1 expression in the signalling cell. *Development* **122**: 1791–1798.
- Miller MA, Nguyen VQ, Lee MH, Kosinski M, Schedl T, Caprioli RM, Greenstein D. 2001. A sperm cytoskeletal protein that signals oocyte meiotic maturation and ovulation. *Science* **291**: 2144–2147.
- Moore GP. 1975. The RNA polymerase activity of the preimplantation mouse embryo. *J Embryol Exp Morphol* **34**: 291–298.
- Moore GP, Lintern-Moore S. 1978. Transcription of the mouse oocyte genome. *Biol Reprod* **18**: 865–870.

- Munro E, Nance J, Priess JR. 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.
- Nadarajan S, Govindan JA, McGovern M, Hubbard EJA, Greenstein D. 2009. MSP and GLP-1/Notch signaling coordinately regulate actomyosin-dependent cytoplasmic streaming and oocyte growth in *C. elegans*. *Development* **136**: 2223–2234.
- Neves A, English K, Priess JR. 2007. Notch-GATA synergy promotes endoderm-specific expression of ref-1 in *C. elegans*. *Development* **134**: 4459–4468.
- Newport J, Kirschner M. 1982. A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. *Cell* **30**: 687–696.
- Noble SL, Allen BL, Goh LK, Nordick K, Evans TC. 2008. Maternal mRNAs are regulated by diverse P body-related mRNP granules during early *Caenorhabditis elegans* development. *J Cell Biol* **182**: 559–572.
- Ogura K-I, Kishimoto N, Mitani S, Gengyo-Ando K, 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.
- Opperman L, Hook B, DeFino M, Bernstein DS, Wickens M. 2005. A single spacer nucleotide determines the specificities of two mRNA regulatory proteins. *Nat Struct Mol Biol* **12**: 945–951.
- Pagano JM, Clingman CC, Ryder SP. 2011. Quantitative approaches to monitor protein-nucleic acid interactions using fluorescent probes. *RNA* **17**: 14–20.
- Pagano JM, Farley BM, Essien KI, Ryder SP. 2009. RNA recognition by the embryonic cell fate determinant and germline totipotency factor MEX-3. *Proc Natl Acad Sci USA* **106**: 20252–20257.
- Pagano JM, Farley BM, McCoig LM, Ryder SP. 2007. Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. *J Biol Chem* **282**: 8883–8894.
- Pasquinelli AE. 2012. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat Rev Genet* **13**: 271–282.
- Pinkston-Gosse J, Kenyon C. 2007. DAF-16/FOXO targets genes that regulate tumor growth in *Caenorhabditis elegans*. *Nat Genet* **39**: 1403–1409.

- Praitis V, Casey E, Collar D, Austin J. 2001. Creation of low-copy integrated transgenic lines in *Caenorhabditis elegans*. *Genetics* **157**: 1217–1226.
- Priess JR. 2005. Notch signaling in the *C. elegans* embryo. *WormBook* 1–16.
- Priess JR, Schnabel H, Schnabel R. 1987. The *glp-1* locus and cellular interactions in early *C. elegans* embryos. *Cell* **51**: 601–611.
- Rambo RP, Doudna JA. 2004. Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry* **43**: 6486–6497.
- Reese KJ, Dunn MA, Waddle JA, 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.
- Reid BG, Flynn GC. 1997. Chromophore formation in green fluorescent protein. *Biochemistry* **36**: 6786–6791.
- Riddle DL, Blumenthal T, Meyer BJ, Priess JR, Riddle DL, Blumenthal T, Meyer BJ, Priess JR. 1997. *Introduction to C. elegans*. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY).
- Riley KJ, Yario TA, Steitz JA. 2012. Association of Argonaute proteins and microRNAs can occur after cell lysis. *RNA* **18**: 1581–1585.
- Rocheleau CE, Downs WD, Lin R, Wittmann C, Bei Y, Cha YH, Ali M, Priess JR, Mello CC. 1997. Wnt signaling and an APC-related gene specify endoderm in early *C. elegans* embryos. *Cell* **90**: 707–716.
- Rybarska A, Harterink M, Jedamzik B, Kupinski AP, Schmid M, Eckmann CR. 2009. GLS-1, a novel P granule component, modulates a network of conserved RNA regulators to influence germ cell fate decisions. *PLoS Genet* **5**: e1000494.
- Ryder SP, Frater LA, Abramovitz DL, Goodwin EB, Williamson JR. 2004. RNA target specificity of the STAR/GSG domain post-transcriptional regulatory protein GLD-1. *Nat Struct Mol Biol* **11**: 20–28.
- Scheckel C, Gaidatzis D, Wright JE, Ciosk R. 2012. Genome-Wide Analysis of GLD-1-Mediated mRNA Regulation Suggests a Role in mRNA Storage. *PLoS Genet* **8**: e1002742.
- Schisa JA, Pitt JN, Priess JR. 2001. Analysis of RNA associated with P granules in germ cells of *C. elegans* adults. *Development* **128**: 1287–1298.

- Schmid M, Kütchler B, Eckmann CR. 2009. Two conserved regulatory cytoplasmic poly(A) polymerases, GLD-4 and GLD-2, regulate meiotic progression in *C. elegans*. *Genes Dev* **23**: 824–836.
- Schubert CM, Lin R, de Vries CJ, Plasterk RH, Priess JR. 2000. MEX-5 and MEX-6 function to establish soma/germline asymmetry in early *C. elegans* embryos. *Mol Cell* **5**: 671–682.
- Seydoux G, Dunn MA. 1997. Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of *Caenorhabditis elegans* and *Drosophila melanogaster*. *Development* **124**: 2191–2201.
- Seydoux G, Mello CC, Pettitt J, Wood WB, Priess JR, Fire A. 1996. Repression of gene expression in the embryonic germ lineage of *C. elegans*. *Nature* **382**: 713–716.
- Shimada M, Yokosawa H, Kawahara H. 2006. OMA-1 is a P granules-associated protein that is required for germline specification in *Caenorhabditis elegans* embryos. *Genes Cells* **11**: 383–396.
- Shirayama M, Soto MC, Ishidate T, Kim S, Nakamura K, Bei Y, van den Heuvel S, Mello CC. 2006. The Conserved Kinases CDK-1, GSK-3, KIN-19, and MBK-2 Promote OMA-1 Destruction to Regulate the Oocyte-to-Embryo Transition in *C. elegans*. *Curr Biol* **16**: 47–55.
- Sönnichsen B, Koski LB, 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.
- Stinchcomb DT, Shaw JE, Carr SH, Hirsh D. 1985. Extrachromosomal DNA transformation of *Caenorhabditis elegans*. *Mol Cell Biol* **5**: 3484–3496.
- Strome S, Wood WB. 1983. Generation of asymmetry and segregation of germ-line granules in early *C. elegans* embryos. *Cell* **35**: 15–25.
- Strome S, Wood WB. 1982. Immunofluorescence visualization of germ-line-specific cytoplasmic granules in embryos, larvae, and adults of *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* **79**: 1558–1562.
- Subramaniam K, 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.
- Suh N, Crittenden SL, Goldstrohm A, Hook B, Thompson B, Wickens M, Kimble J.

2009. FBF and its dual control of *gld-1* expression in the *Caenorhabditis elegans* germline. *Genetics* **181**: 1249–1260.
- Suh N, Jedamzik B, Eckmann CR, Wickens M, Kimble J. 2006. The GLD-2 poly(A) polymerase activates *gld-1* mRNA in the *Caenorhabditis elegans* germ line. *Proc Natl Acad Sci USA* **103**: 15108–15112.
- Sulston JE, Horvitz HR. 1977. Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol* **56**: 110–156.
- Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* **100**: 64–119.
- Swanger SA, Bassell GJ. 2011. Making and breaking synapses through local mRNA regulation. *Curr Opin Genet Dev* **21**: 414–421.
- Tabara H, Hill RJ, Mello CC, Priess JR, Kohara Y. 1999. *pos-1* encodes a cytoplasmic zinc-finger protein essential for germline specification in *C. elegans*. *Development* **126**: 1–11.
- Tadros W, Lipshitz HD. 2009. The maternal-to-zygotic transition: a play in two acts. *Development* **136**: 3033–3042.
- Thorpe CJ, Schlesinger A, Carter JC, Bowerman B. 1997. Wnt signaling polarizes an early *C. elegans* blastomere to distinguish endoderm from mesoderm. *Cell* **90**: 695–705.
- Wang L, Eckmann CR, Kadyk LC, Wickens M, Kimble J. 2002. A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. *Nature* **419**: 312–316.
- Wang X, Zamore PD, Hall TM. 2001. Crystal structure of a Pumilio homology domain. *Mol Cell* **7**: 855–865.
- Wang Y, Opperman L, Wickens M, Hall TMT. 2009. Structural basis for specific recognition of multiple mRNA targets by a PUF regulatory protein. *Proc Natl Acad Sci USA* **106**: 20186–20191.
- Ward S, Argon Y, Nelson GA. 1981. Sperm morphogenesis in wild-type and fertilization-defective mutants of *Caenorhabditis elegans*. *J Cell Biol* **91**: 26–44.
- Ward S, Carrel JS. 1979. Fertilization and sperm competition in the nematode *Caenorhabditis elegans*. *Dev Biol* **73**: 304–321.

- Wickens M, Bernstein DS, Kimble J, Parker R. 2002. A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet* **18**: 150–157.
- Wolke U, Jezuit EA, Priess JR. 2007. Actin-dependent cytoplasmic streaming in *C. elegans* oogenesis. *Development* **134**: 2227–2236.
- Worthington MT, Pelo JW, Sachedina MA, Applegate JL, Arseneau KO, Pizarro TT. 2002. RNA binding properties of the AU-rich element-binding recombinant Nup475/TIS11/tristetraprolin protein. *J Biol Chem* **277**: 48558–48564.
- Wright JE, Gaidatzis D, Senften M, Farley BM, Westhof E, Ryder SP, Ciosk R. 2011. A quantitative RNA code for mRNA target selection by the germline fate determinant GLD-1. *EMBO J* **30**: 533–545.
- Yochem J, Greenwald I. 1989. *glp-1* and *lin-12*, genes implicated in distinct cell-cell interactions in *C. elegans*, encode similar transmembrane proteins. *Cell* **58**: 553–563.
- Zamore PD, Bartel DP, Lehmann R, Williamson JR. 1999. The PUMILIO-RNA interaction: a single RNA-binding domain monomer recognizes a bipartite target sequence. *Biochemistry* **38**: 596–604.
- Zamore PD, Williamson JR, Lehmann R. 1997. The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* **3**: 1421–1433.
- Zhang B, Gallegos M, Puoti A, Durkin E, Fields S, Kimble J, Wickens MP. 1997. A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line. *Nature* **390**: 477–484.