INFORMATION TO USERS

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

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

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

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

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

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

MSY4, a sequence-specific RNA binding protein expressed during mouse spermatogenesis

by

Holly Gibs Davies

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2000

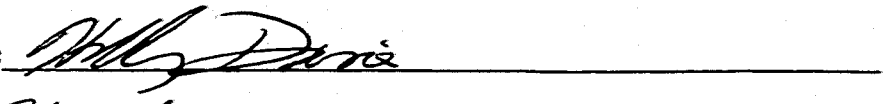
Program Authorized to Offer Degree: Genetics

UMI Number: 9964248

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Doctoral Dissertation

In presenting this thesis in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to UMI Dissertation Services, 300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature 

Date March 15, 2000

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Holly Gibs Davies

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

Robert E Braun

Robert Braun

Reading Committee:

Robert E Braun

Robert Braun

Stan Fields

Adam Geballe

Date: March 10, 2000

University of Washington

Abstract

MSY4, A SEQUENCE-SPECIFIC RNA BINDING PROTEIN EXPRESSED DURING MOUSE SPERMATOGENESIS

by Holly Gibs Davies

Chairperson of the Supervisory Committee: Associate Professor Robert E. Braun
Department of Genetics

Translational control is an important mechanism of gene regulation, especially in gametogenesis and early embryogenesis, where there are periods of time in which there is no transcription. Previously transcribed mRNAs must be translated at the correct time and in the correct sequence for development to be successfully completed. Regulatory elements have been identified within the 5' and 3' untranslated regions (UTRs) of mRNAs, but mechanisms of translational control are poorly understood. Transcription ceases midway through mouse spermatogenesis. Translational control of pre-existing mRNAs is necessary for the completion of spermatogenesis. This requires mRNA binding proteins to play a major role in spermatogenesis.

Protamine 1 (*Prm1*) and protamine 2 (*Prm2*) are among the genes that are under translational control during mouse spermatogenesis. *Prm1* is transcribed in round spermatids, but its mRNA is not translated until 2-8 days later in elongated spermatids. The protamines are small, arginine-rich proteins that function in the condensation of the nucleus during the later phase of spermatogenesis. Premature translation of *Prm1* causes premature nuclear condensation and sterility. Translation is controlled by *cis*-elements in the 3' UTR and several *cis*-and trans-elements have been described, but the mechanism of translational control is still not clear.

My dissertation focuses on one of the *cis*-elements within the *Prm1* 3' UTR and the proteins that bind to this region. I have characterized the composition of a previously described binding activity in testis extracts and cloned the gene *Msy4*, which encodes one of the protein components. MSY4 is highly expressed in the cytoplasm of spermatids, where translationally repressed mRNAs are stored. I have used the yeast three-hybrid system to define the Y box recognition sequence (YRS), which is the 7 nucleotide RNA binding site for MSY4 and a related protein, MSY2, and shown the necessity of the YRS for translational repression in transgenic mice. Ectopic expression of MSY4 later in mouse spermatogenesis causes dominant sterility, demonstrating the importance of MSY4 during mouse spermatogenesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The author wishes to thank my advisor Bob Braun, my supervisory committee Stan Fields, Adam Geballe, Ben Hall and my GSR. Bob Braun, Stan Fields and Adam Geballe for reading my dissertation. I have gotten so much help and support from my lab members. I'd like to thank the past lab members Mark Fajardo, Antoine H.F.M. Peters, and Madeleine Meagher, and the present lab members Jessica Nadler, Chuck Connolly, Bill Buaas, Flaviano Giorgini and Kathy Kafer. I would also like to thank my family, who have supported me throughout my graduate career and in everything I've done.

DEDICATION

The author wishes to dedicate this thesis to Adrian P. Spidle III.

INTRODUCTION

TRANSLATIONAL CONTROL

Translational control is an important mechanism of gene regulation, especially in gametogenesis and early embryogenesis (Curtis *et al.*, 1995). There are several advantages to regulation by translational control including sensitivity to levels of translation factors in the cell, potential for regulation in the absence of transcription, and the capacity to respond quickly to changing conditions and stimuli. Translational control can occur on a global level, regulating translation of the entire complement of mRNAs within a cell, or it can be very selective, affecting as little as a single species of mRNA. The *cis*-acting elements for translational control have been found in the 5' and 3' untranslated regions (UTRs) of regulated mRNAs (reviewed in Standart and Jackson, 1994). Even in cases where *cis*- and trans-elements have been identified, the mechanisms of translational repression and activation are poorly understood.

TRANSLATION

Translation consists of initiation, elongation, and termination, and is most frequently regulated at the initiation step. Eukaryotic mRNAs contain a 5' cap structure and a 3' poly(A) tail, which are important for translation. Ribosomes are recruited to the mRNA in a multi-step process that begins with the binding of eIF-4F, consisting of eIF4E, eIF4A, and eIF4G, to the cap. The 43S pre-initiation complex, which is composed of the 40S ribosomal subunit, tRNAmet, GTP and eIF2, is recruited to the

vicinity of the cap complex and then scans along the mRNA in an ATP-dependent manner until it reaches an AUG in a good context. Other initiation factors, eIF1A and eIF3, are needed to stablize the complex. Once the 40S subunit reaches the start codon, the 60S ribosomal subunit attaches to form the 80S ribosome and elongation begins. Elongation continues until the termination codon is reached. At this point the message and the peptide are released from the ribosome and the entire complex disassociates (reviewed in Hershey, 1991).

There is a cap-independent mechanism of translation initiation that was first identified in picornaviruses. These viral mRNAs have large and highly structured 5' UTRs that contain internal ribosome entry sites (IRES). No free 5' end is needed since a circular mRNA that contains an IRES can be translated (Chen and Sarnow, 1995). The IRES is an alternative method of recruiting the small ribosomal subunit. However, the initiation factors that have been identified for cap-dependent initiation are also needed for IRES initiation, in addition to other trans-factors such as the La autoantigen and poly-pyrimidine tract binding protein (PTB), which may be needed to stabilize the secondary structure of the IRES. IRESs are not unique to viruses and have been found in cellular mRNAs, such as *Drosophila Antennapedia* (Oh *et al.*, 1992) and the mRNA that encodes the immunoglobulin heavy chain binding protein BiP (Macejak and Sarnow, 1991). There are two classes of picornaviral IRESs that differ in the position of the AUG with respect to the IRES. Type II IRESs have the initiator codon at their 3' end, which is reminiscent of positioning by the Shine-Delgarno sequence in prokaryotic initiation

(Pestova et al. 1998). Type I elements require the small ribosomal subunit to scan 50-100 nt to reach the initiator AUG. A similar initiation mechanism called "hopping" has been reported in cauliflower mosaic virus 35S mRNA (Futterer *et al.*, 1993). In this discontinuous scanning mechanism the 5' UTR is highly structured, but is not an IRES.

In addition to the 5' cap, cellular mRNAs contain a poly(A) tail, which promotes efficient translation in general. The poly(A) tail seems to mediate its effect on translation through proteins bound to it and to the cap complex. The poly(A) tail is bound by poly(A) binding protein (PABP), which requires 12 nt to bind and occupies 25 nt (Sachs *et al.*, 1987). *PAB1* is essential in yeast and results of studies both *in vitro* and *in vivo* suggest it plays an important role in translation. In yeast PABP binds eIF4G (Tarun and Sachs, 1995), which is bound to eIF4E, which is bound to the cap. This circularization of mRNAs has been directly observed with atomic force microscopy (Wells *et al.*, 1998). The poly(A) tail has been suggested to stimulate translation at the joining of the 60S ribosomal subunit (Munroe and Jacobson 1990). In support of this, suppressor mutations that bypass a yeast mutation in *PAB1* affect the 60S subunit (Sachs and Davis, 1989). The binding of the 40S subunit to mRNA has also been reported as an effect of poly(A) (Tarun and Sachs, 1995). The PABP/eIF4G/eIF4F complex may exert its effect on translation by its ability to compete for limiting initiation complexes more effectively (Preiss and Hentze, 1998). Another protein that binds PABP has been reported in mammalian cells (Craig *et al.*, 1998). This PABP interacting protein (PAIP) has homology to eIF4G, but has no eIF4E binding site, and may tether PABP to the 5' end

via eIF4A. One hypothesis of how polyadenylation at the 3' end of the mRNA can affect initiation of translation at the 5' end suggests that poly(A) polymerase interacts with a cap-specific ribose methyltransferase required for cap binding of the initiation complex (Kuge and Richter, 1995).

KNOWN FEATURES OF TRANSLATIONAL CONTROL

There are many known *cis*-elements for translational control that reside in the untranslated regions of mRNAs. Interfering with the most efficient translation, as discussed above, can effect translational control. The 5' UTR can affect translation through modification of the cap, secondary structure, RNA-protein interactions, upstream open reading frames (uORFs), upstream start codons (uAUGs), and IRESs. The 3' UTR can affect translation through RNA-protein interactions, RNA-RNA interactions, cytoplasmic polyadenylation elements (CPEs), and poly(A) tail length. Although *cis*-elements have been identified within many UTRs, the underlying mechanisms of translational repression and derepression are generally poorly understood.

5' UNTRANSLATED REGIONS

One example of an mRNA that is under translational control that is well understood is ferritin. The translation of ferritin mRNA is controlled by steric hindrance of the 5' cap structure (Gray and Hentze, 1994). During iron-starvation conditions the iron response protein (IRP) binds to the iron response element (IRE), a short, stem-loop

structure located near the 5' cap. This binding prevents the initial stable interaction of the 40S ribosomal subunit with the cap, repressing translation of the message.

Upstream open reading frames (uORFs) are generally inhibitory to translation of the downstream ORF. The translation of the mRNA encoding the transcription factor GCN4 in yeast is regulated by four uORFs (reviewed by Hinnebusch, 1996). In non-starvation conditions the ribosome initiates synthesis at uORF1, reinitiates at uORF4, and fails to reinitiate at *GCN4*. During starvation conditions the ribosome initiates at uORF1, fails to reinitiate at uORF4, and reinitiates at *GCN4*. The rates of initiation and reinitiation are controlled by the amount of available initiation factors in the cell. *GCN4* is therefore translated when concentrations of initiation factors are low, when other transcripts are translated less frequently.

3' UNTRANSLATED REGIONS

OOGENESIS

Translational control is of particular importance in the germline, as there are periods of time in which there is no transcription (Figure I-1). During oogenesis, transcription ceases prior to oocyte maturation, requiring stored maternal mRNAs to regulate protein expression in oocyte maturation, fertilization and early embryogenesis (reviewed in Stebbins Boaz and Richter, 1997). The change from repression to translation of maternal mRNAs is correlated with a change in the poly(A) tail lengths. A set of transcripts, which includes mRNAs for *cyclin A1, B1, B2,* and *c-mos*, is repressed in

Xenopus oocytes. The mRNAs are processed normally in the nucleus before being transported to the cytoplasm, but are deadenylated and found in a translationally repressed state in ribonucleoprotein (RNP) particles after they reach the cytoplasm. At meiotic maturation or fertilization the messages are readenylated and then translated. The deadenylation and readenylation are mediated by a U-rich cytoplasmic polyadenylation element (CPE) and the nuclear polyadenylation signal, which are both found in the 3' UTR. The 3' UTRs are sufficient to confer polyadenylation-dependent translational control on a reporter mRNA (Sheets *et al.*, 1994).

This mechanism is conserved in mouse oocytes, as evidenced by the translational regulation of tissue-plasminogen activator (tPA) (Vassalli *et al.*, 1989). In primary mouse oocytes the *tPA* mRNA initially has a long poly(A) tail, which is then shortened to about 50 nt. While the *tPA* mRNA has a short tail, it is translationally repressed. At meiotic maturation the poly(A) tail is readenylated and the mRNA is translated. The translation is controlled by the CPE in the 3' UTR and a titratible trans-factor.

A protein that binds to the CPE, the CPE binding protein (CPEB), has been identified in several organisms, including *Xenopus* (McGrew and Richter, 1990), mouse (Gebauer and Richter, 1996), zebrafish (Bally-Cuiff et al. 1998), and surf clam (Walker *et al.*, 1999). In the surf clam *Spisula solidissima*, the CPEB protein p82 is phosphorylated by cdc2 and MAP kinases within ten minutes of fertilization and then migrates as 92kDa. p82 acts as a masking protein in clam oocytes and as a polyadenylation factor in fertilized eggs. It is tempting to propose that the

phosphorylation affects the RNA binding (Walker *et al.*, 1996 and Minshall *et al.*, 1999), but other results suggest phosphorylation does not affect RNA binding (Hake *et al.*, 1998; Walker *et al.*, 1996). At least one mRNA in *Xenopus* oocytes, *c-mos*, is polyadenylated before CPEB is phosphorylated (de Moor and Richter, 1997), demonstrating that polyadenylation is not dependent on the phosphorylation of CPEB. *Xenopus* CPEB has been found to interact with a new factor called maskin, which has similarity to eIF4E (Stebbins-Boaz *et al.*, 1999). The dissolution of the CPEB/maskin complex may result in the binding of CPEB to eIF4E, which binds eIF4G, and the translation of CPE containing mRNAs.

EMBRYOGENESIS

In *Drosophila* embryos the anterior/posterior axis is set up by the 3'UTR-mediated localization and translational repression of certain mRNAs (reviewed in Wickens *et al.*, 1996). *Nanos* mRNA (*nos*) is localized to the posterior end of the embryo, where it is translated and forms a gradient of nanos protein. The repression of *nos* mRNA is controlled by TCEs in its 3' UTR that are bound by Smaug (Smithberg et al.1996). Nanos protein represses the translation of *hunchback* mRNA, which is uniformly distributed, via two nanos response elements (NREs) in the *hunchback* 3' UTR. This produces a gradient of hunchback protein from the anterior to the posterior. Nanos is not thought to interact with *hunchback* mRNA directly, but through another protein, Pumilio.

The posterior morphogens, Bicoid and Oskar, are controlled in a similar manner. *Bicoid* mRNA (*bcd*) is localized to the anterior of the embryo, where it is translated and forms a gradient of bicoid protein. The repressed form of the *bcd* mRNA contains a short poly(A) tail of about 50 nt, whereas the translated form has about 150 nt (Sall'es *et al.*, 1994). Bicoid protein represses the translation of *caudal* mRNA, which is uniformly distributed, via elements in the 3' UTR of *caudal*. This produces a gradient of Caudal protein from the posterior to the anterior.

The anterior-posterior axis in *Drosophila* is dependent on the localization of Oskar protein in oocytes. If *oskar* mRNA (*osk*) is not translationally repressed, the protein accumulates throughout, specifying posterior body patterning throughout the embryo (Kim Ha *et al.*, 1995). During oogenesis *osk* mRNA is both localized to the posterior pole and translationally repressed in other areas. The translational repression is dependent upon sequences in the 3' UTR that include BREs (Bruno response elements), which are bound by Bruno protein. An *in vitro* translation system using *Drosophila* ovarian extracts can recapitulate this repression *in vitro*. The repression was shown to be dependent on the BREs and Bruno, but on neither the 3' poly(A) tail nor the 5' cap (Lie and Macdonald 1999). Better experiments to test the cap dependence await an IRES that is functional in *Drosophila* embryos. In the oocyte extracts, reporter mRNAs with or without a poly(A) tail were specifically repressed by Bruno protein. Consistent with this, *osk* mRNA has a poly(A) tail of only about 35 nt, and PABP is not present at high levels

in egg chambers where *osk* mRNA is localized to the posterior pole and translationally acitvated.

Although repression of *osk* mRNA appears to be independent of polyadenylation, activation of *osk* mRNA involves the *Drosophila* CPEB, Orb. Orb interacts with *osk* mRNA genetically and biochemically (Chang *et al.*, 1999). Oskar protein is not expressed at normal levels in Orb mutant ovaries, *osk* mRNA can be coimmunoprecipitated with Orb antibody, and crosslinked to Orb protein. Orb may promote *osk* mRNA translation in several ways. First, it may activate *osk* mRNA translation at the posterior pole of the oocyte through polyadenylation. Second, it may prevent degradation of *osk* mRNA. Third, it may displace Bruno or other repressors from the mRNA. In addition to the interaction of *osk* 3' UTR with Orb, derepression of *osk* mRNA at the posterior pole requires the 5' UTR (Gunkel *et al.*, 1998), suggesting a complex mechanism of the control of Oskar expression.

A group of male specific genes is required for dosage compensation in *Drosophila*, where the transcription of the male X chromosome is increased to match the transcription of the two female X chromosomes. Sex-lethal (Sxl) protein is expressed only in females, due to alternative splicing in males and females. Sxl then prevents the expression of *msl-2* in females by repressing its translation. There are Sxl binding sites of poly(U) tracts in both the 5' and 3' UTRs that contribute to repression. The repression can be recapitulated *in vitro* in *Drosophila* embryo extracts, and shown to be independent of the presence of the poly(A) tail. Sxl has RRM motif RNA-binding domains, which are

sufficient for translational repression. The Sxl RNA-binding sites in both the 5' and 3' UTRs reiterate the interaction between both ends of mRNA, but it's not clear whether Sxl interacts with the 5' and 3' UTRs as part of one mechanism of repression or if it interacts with each independently in redundant mechanisms.

In *C. elegans* the onset of spermatogenesis and the switch from spermatogenesis to oogenesis is controlled by the 3' UTR-mediated translational repression of *tra-2* and *fem-3*. Translational repression of *tra-2* is via sequences in the 3' UTR which are now called TGEs (tra-2 and GL1 elements) and used to be called DREs. There is a binding activity, called the DRF (direct repeat factor), in crude worm extracts that binds to the TGEs. Two tandem copies of the TGE were used as bait in the three-hybrid assay to clone the gene that encodes DRF (Jan *et al.*, 1999). GLD-1 (defective in germline development) was found to specifically bind the TGE. GLD-1 is a member of the STAR subfamily of KH domain RNA binding proteins that is only expressed in the cytoplasm of the germline and is required for oogenesis and spermatogenesis. Consistent with the three-hybrid results, GLD-1 specifically binds to TGEs *in vitro* and is part of the DRF. GLD-1 can repress a reporter mRNA with the *tra-2* 3' UTR both in worms and in yeast extracts. The mechanism of repression is unknown, but TGEs may control translation by regulating poly(A) tail length, since the presence of the TGEs is correlated with a short poly(A) tail (Jan *et al.*, 1997). Other possible mechanisms are that GLD-1 binds directly or indirectly to initiation factors, or GLD-1 may sequester or mask *tra-2* mRNA.

Whereas *tra-2* is repressed during *C. elegans* spermatogenesis, *fem-3* is repressed in the germline to switch from spermatogenesis to oogenesis. The *cis*-element, PME, was defined in the *fem-3* 3' UTR and a titratable factor that binds the PME is present in crude worm extracts (Ahringer and Kimble, 1991). Two copies of the PME were used in the three-hybrid system to clone the genes that encode the trans-factor FBF (Zhang *et al.*, 1997). FBF is a candidate for the *fem-3* repressor based on its specific RNA-binding in the three-hybrid system and its expression in the germline. Other results that support the role of FBF in the sex-determination pathway are that the loss of FBF by RNA interference results in masculinization of the germline and the epistatic analysis of *fbf* with other known genes in the pathway (Zhang *et al.*, 1997). FBF has homology to *Drosophila* Pumilio and defines a new Puf family of RNA-binding proteins. Like *Drosophila* Pumilio, FBF interacts with a *C. elegans* homolog to the *Drosophila* Nanos, NOS-3 (Kraemer *et al.*, 1999).

SPERMATOGENESIS

Spermatogenesis is a complex developmental process that is ongoing in adult males. During spermatogenesis a diploid cell develops into haploid mature spermatozoa that are capable of reaching and fertilization an egg. Translational control is very important during spermatogenesis, since there is no transcription during the latter phase of spermatogenesis (Figure I-1). In *Drosophila* the entire genetic material is dispensable after meiosis (Lindsley and Grell, 1969), although a few mRNAs are normally transcribed after meiosis.

The model for translational control in *Drosophila* spermatogenesis is the Mst(3)GCP family, which consists of 7 genes that encode structural proteins in the sperm tail. The genes are transcribed in primary spermatocytes, but translationally repressed until the spermatids are fully elongated. All seven genes have a conserved 12 nt sequence in the 5'UTR, called the translational control element (TCE). Both the sequence and the position of the TCE are conserved among family members, suggesting a steric hindrance mechanism. Consistent with this, displacement of the TCE leads to complete loss of translational regulation. There is a binding activity found only in testis extract that binds to the TCE (Kempe *et al.*, 1993). Several other male-specific transcripts contain the TCE in their 5' UTRs, although the position is not as well conserved. There are 2 TCE-like elements in the 5' UTR of *Mst89B* (Stebbings *et al.*, 1998), whose transcript is present in germline-derived cells throughout spermatogenesis. The 5' UTR of *janusB* is similar to the Mst(3)GCP family 5' UTRs and is responsible for translational repression (Yanicostas and Lepesant, 1990). *Mst33A* and *Mst59D* have also been reported to have TCEs in their 5' UTRs (Schafer *et al.*, 1995). This TCE may represent an important *cis*-element in many male-specific transcripts that are under translational control in *Drosophila* spermatogenesis.

MOUSE SPERMATOGENESIS

Mouse spermatogenesis is divided into three stages (Bellve, 1979). The first stage is spermatogonial proliferation, the second is meiosis, and the third is spermiogenesis. Spermiogenesis in mice takes about two weeks and can be divided into sixteen steps

(Oakberg, 1956). At step 9, midway through murine spermiogenesis, nuclear condensation begins, during which the histones are replaced by the transition proteins which are then replaced by the protamines. The nucleus is compacted to 1/10 its volume and transcription stops (Balhorn, 1982; Kierszenbaum and Tres, 1975; Monesi, 1964). Translational control of pre-existing mRNA is required to regulate protein levels in the absence of transcription.

In mouse spermatogenesis, *Prm1* is transcribed in step 7 round spermatids, but its mRNA is stored as an RNP particle in the cytoplasm for up to one week before translation begins in step 12 elongating spermatids (Kleene *et al.*, 1984). The 3' untranslated region (UTR) has been shown to be necessary and sufficient to confer *Prm1*-like translational control in transgenic mice (Braun *et al.*, 1989b). Further transgenic mouse and *in vitro* experiments have identified two *cis*-acting elements, the 3' most 62 nucleotides and the 5' most 37 nucleotides of the *Prm1* 3' UTR that are sufficient to confer *Prm1*-like translational control in transgenic mice (Braun, 1990; Fajardo *et al.*, 1994). The 3' element has been further defined as a 16 nt Z box (Zhong et al., submitted). Two additional *cis*-elements have been described, the Y and H boxes (Kwon and Hecht, 1991). The non-overlapping *cis*-elements suggest that there is complexity in the mechanism of *Prm1* translational control.

PROTEINS THAT BIND PRM1 *3' UTR RNA*

Identifying trans-acting factors is an important step in elucidating the mechanism of translational control. Several proteins that bind to the *Prm1* 3' UTR have been identified. A 40kDa protein that binds to the 3' most 62 nt was identified and the gene encoding this protein has been cloned (Lee *et al.*, 1996). This protein has been named PRBP, for protamine 1 RNA binding protein. PRBP binds nonspecifically to double stranded RNA and has been shown to repress translation *in vitro* nonspecifically. PRBP is present in the cytoplasm of round spermatids where *Prm1* mRNA is repressed and absent in elongating spermatids where *Prm1* mRNA is translated (Lee *et al.*, 1996). Mice that are nullizygous for *Tarbp2*, which encodes PRBP, are sterile and fail to activate the translation of *Prm1* in elongated spermatids (Zhong *et al.*, 1999b). *Tarbp2* $^{-/-}$ mice do not translate *Prm1* mRNA prematurely, and transgenes that are missing the PRBP binding site are under translational control, suggesting that PRBP is not the *Prm1* repressor, but may function in translational activation during spermatogenesis.

Kwon and Hecht (1991) defined two *cis*-acting regions in the *Prm2* 3'UTR, which they called the H and Y boxes, and identified a protein that binds to the Y box. Testis extracts that are enriched for this protein, testis-brain RNA-binding protein (TB-RBP), are able, in a cell-free translation system, to selectively repress a reporter gene linked to the *Prm2* 3'UTR or to the H and Y elements (Kwon and Hecht, 1993). TB-RBP is associated with actin and may function in the transport of specific mRNAs in the testis (Han *et al.*, 1995; Morales *et al.*, 1998).

Cytoplasmic fractions of germ cells contain a factor that binds in a sequence-specific manner to the 5' most 37 nt of *Prm1* 3' UTR (Fajardo *et al.*, 1994). This specific RNA-binding activity is detected as a closely spaced doublet of approximately 48 and 50kDa. My thesis work has focused on this binding activity and its RNA binding site within the *Prm1* 3' UTR. To clone the gene encoding this binding activity, the yeast three-hybrid system (SenGupta *et al.*, 1996) was used to screen a mouse testis cDNA library for genes that encode a protein that specifically binds to this RNA sequence. *Msy4*, a novel member of the Y box family of nucleic acid binding proteins was cloned (Davies *et al.*, 2000). Figure I-2 shows the *cis*- and trans- elements that have been identified in the *Prm1* 3' UTR.

The Y-box proteins were originally identified as transcriptional activators that bound to the Y-box DNA sequence (Deschamps *et al.*, 1992; Sakura *et al.*, 1988). This Y box sequence is not the same as the Y box sequence mentioned above in reference to Kwon and Hecht's work. Y-box family members include DNA-binding protein A (DBPA) and YB-1 in humans, FRGY1 and FRGY2 in frogs, and MSY1 and MSY2 in mice. All family members contain the highly conserved cold shock domain (CSD) that is 43% identical from *E. coli* to humans, and which contains RNP1 and RNP2 RNA binding domains. The small (30-85 amino acid (aa)) amino-terminal region is highly variable among family members. The carboxyl-terminus is more conserved and contains four alternating acidic and basic regions, although the actual amino acid sequences are not

highly conserved among family members. *Msy4* has the most similarity to human *DBPA* and is 85% identical at the nucleotide level.

Although these proteins were found to bind DNA specifically, they were originally thought to be nonspecific RNA-binding proteins. Y-box proteins are a major component of mRNPs and high concentrations of FRGY2 are able to translationally repress mRNAs (Bouvet and Wolffe, 1994). FRGY1 and FRGY2 can bind a hexanucleotide RNA specifically *in vitro* (Bouvet *et al.*, 1995) . This element is sufficient to confer specific translational repression on an mRNA in *Xenopus* oocytes by FRGY2 (Matsumoto *et al.*, 1996).

Msy4 is predominantly expressed in the testis and the protein is found in the cytoplasm of pachytene spermatocytes, round and elongating spermatids, but is absent in elongated spermatids. *In vitro* experiments with antibodies to MSY2 and MSY4 showed that both proteins are part of the previously described 48/50 kDa binding activity in mouse testis extracts. Using *in vitro* electrophoretic mobility shift assays (EMSAs) and the three-hybrid system, a 7 nt sequence that is specifcally bound by MSY2 and MSY4 was defined. This sequence, 5' UCCAUCA 3', is called the Y box recognition site (YRS). The 5' most 37 nt of *Prm1* 3' UTR is sufficient to confer *Prm1*-like translational control on a reporter gene in transgenic mice, and the YRS is necessary for this translational repression.

The function of MSY4 *in vivo* is unclear. In addition to the sequence-specific RNA binding seen in EMSAs and in the three-hybrid system, MSY4 can interact with all RNAs in a nonspecific manner in testis extracts. This suggests that MSY4 interacts with all mRNAs during its expression in mouse spermatogenesis. Its expression in mouse oocytes also suggests it plays a role in storage of maternal mRNAS. In transgenic mice, the expression of MSY4 later in spermatogenesis in elongated spermatids causes dominant male sterility. It's not known through which mRNAs MSY4 interacts to cause this defect in spermatogenesis. It may be through the protamines or other mRNAs.

Figure I-1 Transcription and translation during gametogenesis and early embryogenesis in the mouse

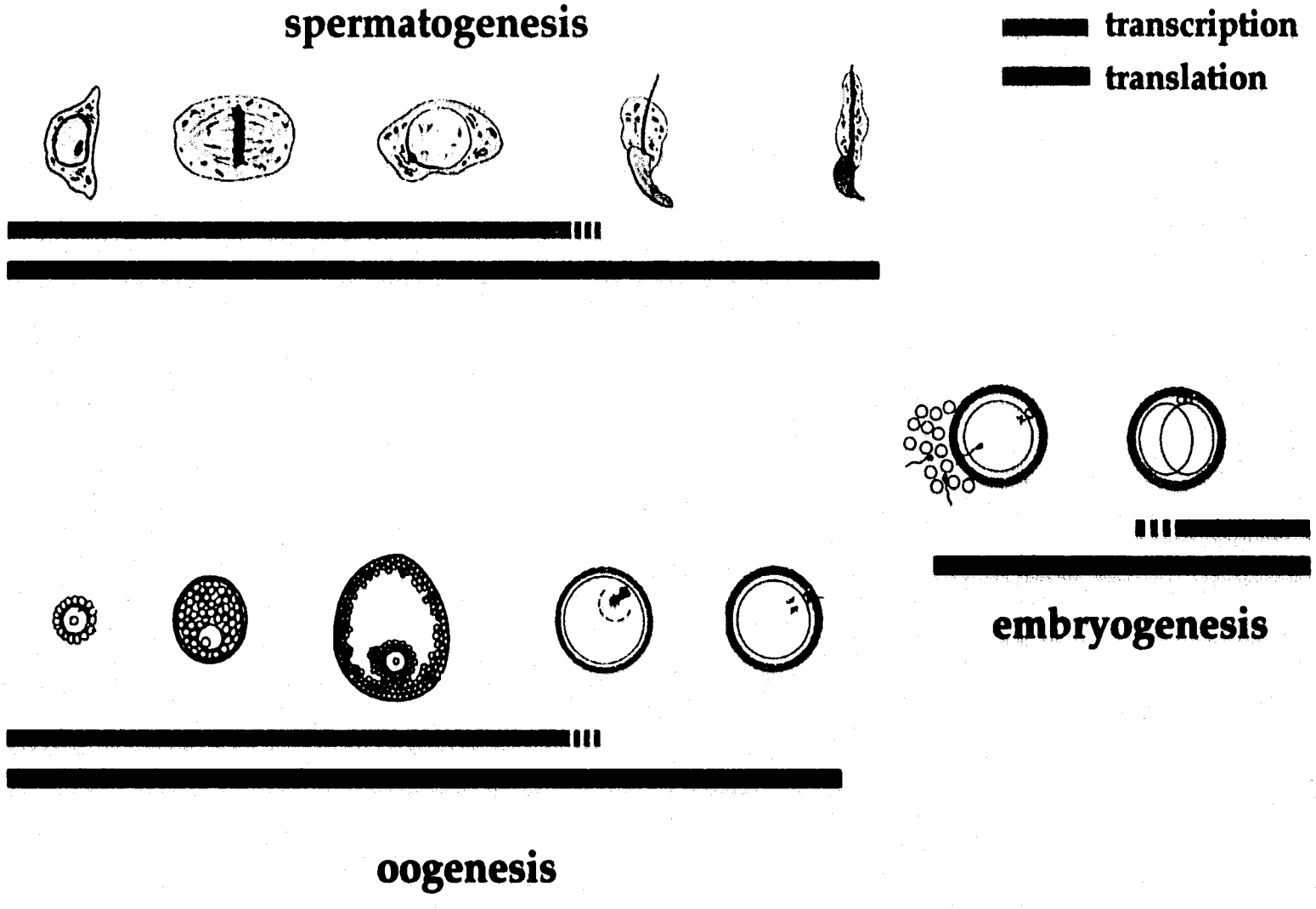
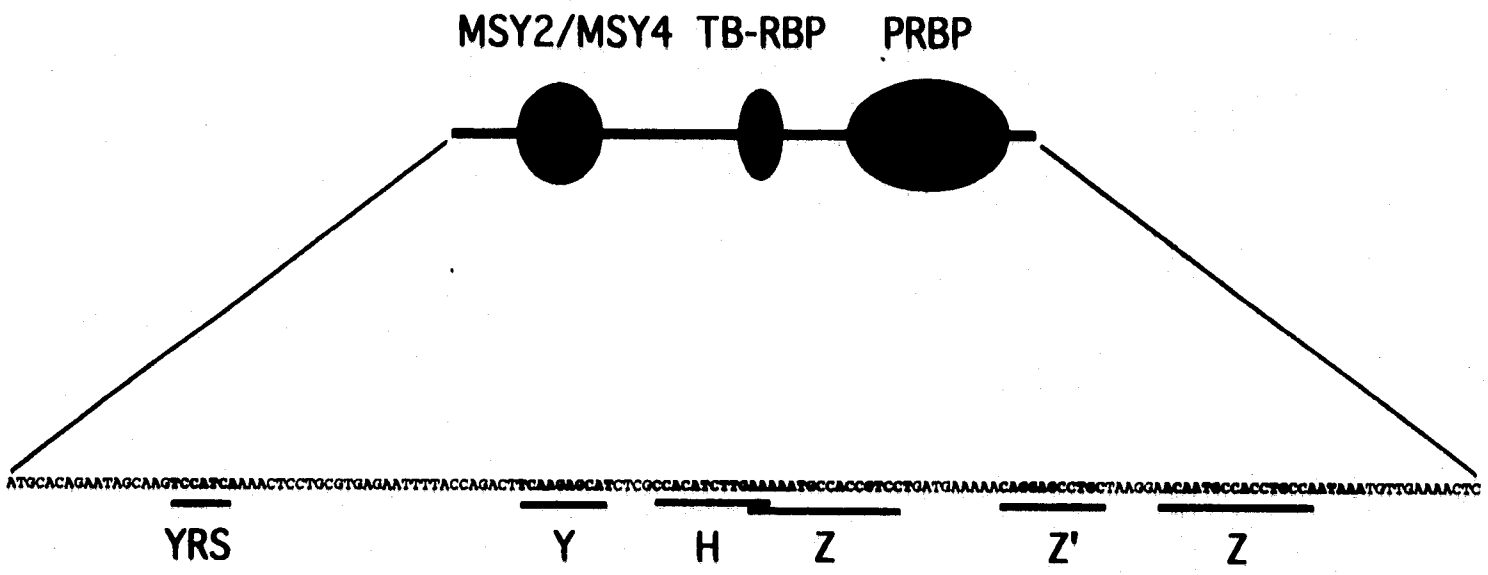

spermatogenesis
transcription
translation
embryogenesis
oogenesis

Figure I-2 *Cis*- and trans-elements in the protamine 1 3' untranslated region.


MSY2/MSY4
TB-RBP
PRBP
YRS
Y
H
Z
Z'
Z

CHAPTER 1: MATERIALS AND METHODS

THREE-HYBRID SELECTION AND SCREENING

A derivative of yeast strain L40 (*MATα, ura3-52, leu2-3, 112, hisΔ200, trp1Δ1, ade2, LYS2::(LexAop)-HIS3, ura3::(LexAop)-LacZ*), containing plasmids encoding the LexA DNA binding domain-MS2 coat fusion protein and the hybrid RNA pIII/MS2-2-*Prm1* $_{1\text{-}37}$ (wt) (SenGupta *et al.*, 1996), was transformed with a mouse testis cDNA Matchmaker library (Clontech, Palo Alto, CA). Transformants were plated onto synthetic media lacking leucine, uracil, tryptophan, and histidine. 5 mM 3-aminotriazole was used to select for higher levels of activation of *HIS3*. Approximately 7.5×10^6 transformants were screened. $HIS3^+$ prototrophs were assayed for β-galactosidase activity using 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) as a substrate in Z buffer in a filter assay (Breeden and Nasmyth, 1985).

Colonies were tested for RNA-dependence by assaying for β-galactosidase expression in cells cured of the plasmid that encodes the hybrid RNA, pIII/MS2-2-*Prm1* $_{1\text{-}37\ (wt)}$. Transformants were plated on synthetic media lacking leucine and tryptophan and containing 0.15% 5-fluoro-orotic acid (FOA) to select for the loss of *URA3*, and thus the loss of the hybrid RNA, then retested for β-galactosidase activity. Candidates that failed to activate *lacZ* following plasmid loss were tested for binding specificity by reintroducing hybrid RNA plasmids encoding *Prm1* $_{1\text{-}37\ (wt)}$, *Prm1* $_{1\text{-}37\ (mut)}$ or

Prm1$_{134\text{-}156}$ (Figure 2-1). Of 113 HIS3^{+} transformants, ten had β-galactosidase activity, and five were RNA-dependent. One transformant specifically bound the *Prm1*$_{1\text{-}37\ (wt)}$ RNA and not the *Prm1*$_{1\text{-}37\ (mut)}$ or *Prm1*$_{134\text{-}156}$.

YEAST THREE-HYBRID SYSTEM BINDING ANALYSIS

A derivative of the *Saccharomyces cerevisiae* L40 (*MATa, ura3-52, leu2-3, 112, hisΔ200, trp1Δ1, ade2, LYS2::(LexAop)-HIS3, ura3::(LexAop)-LacZ*) with an integrated fusion gene encoding the LexA DNA binding domain-MS2 coat fusion protein (SenGupta *et al.*, 1996) and containing the GAL4 activation domain fusion plasmid was transformed with the plasmid encoding the hybrid RNAs, pIII/MS2-2-*Prm1*$_{1\text{-}37}$. Transformants were plated onto synthetic media lacking leucine, uracil, tryptophan, and histidine. Interactions between the GAL4 activation domain fusion protein and the various hybrid RNAs were tested in triplicate by patching single transformant colonies onto synthetic media lacking leucine, uracil, tryptophan, and histidine, and containing 5 mM 3-aminotriazole. Interactions were also tested using β-galactosidase activity using 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) as a substrate in Z buffer in a filter assay.

GAL4 ACTIVATION DOMAIN FUSION PLASMIDS

Full-length MSY4 was cloned into pACTII. MSY4ΔN is the original clone from the three-hybrid screen in pGAD10 (Davies *et al.*, 2000). MSY4ΔC was isolated from the Matchmaker library (Clontech) by sequence homology. MSY4CSD was cloned by PCR

using the primers 5' CGCGGATCCCATGAGCGAGGCGGGCGAGGCCAC 3' and 5' CGCGGATCCGGTCAGCAGCATAGCGAC 3', then inserted into the BamHI site of pACT. MSY2 was cloned via PCR with the Matchmaker library as the template, using the primers 5' CGCGGATCCCAAGCCGGTGCTGGCAATCC 3' and 5' CGCGGATCCGAATCACTCCAGTATGGTG 3', then inserted into the BamHI site of pACT (Clontech) to be expressed as a fusion protein with the GAL4 activation domain. MSY1 was cloned by PCR with the Matchmaker library as the template, using the primers 5' CGCGGATCCCATGAGCAGCGAGGCCGAGACC 3' and 5' CGCGGATCCCTTACTCAGCCCCGCCCTGCTC 3', then cloned into the BamHI site of pACT.

Quantitative liquid β-galactosidase asays were done on duplicate cultures for each RNA hybrid transformant. Cultures were grown in liquid synthetic medium lacking leucine, uracil, tryptophan, and histidine to log phase (OD_{600} 0.5-1.0 and at least two doublings). Three 1 ml aliquots of cells were pelleted from each culture by centrifugation at 14,000 rpm in a microfuge. Cells were washed with 500 µl, then resuspended in 100 µl Z buffer and lysed by freeze/thaw in liquid nitrogen. 700 µl of Z buffer and 160 µl of freshly prepared o-nitrophenyl β-D-galactopyranoside (ONPG, 4mg/ml in Z buffer) was added to each reaction. After color development, 400 µl of 1M sodium carbonate was added to stop the reaction, the cell debris was pellet at 14,000 rpm for 10 minutes in a microfuge and the OD_{420} was read. Statistical analysis was performed using Microsoft Excel 5.0.

PLASMIDS

The vector encoding the hybrid RNA, pIII/MS2-2, carries the *URA3* marker and a unique SmaI site between an RNA polymerase III promoter and the sequence encoding the MS2 target RNA. To construct the hybrid RNAs used in this study complementary oligonucleotides encoding different regions of the *Prm1* 3' UTR were annealed and ligated into the SmaI site. These hybrid RNA constructs were generated by blunting the 5' overhanging EcoRI-BamHI ends of the $Prm1_{1\text{-}37}$ oligonucleotides by Klenow filling and cloning into the SmaI site of pIII/MS2-2. Constructs were verified by ^{35}S sequencing using Sequenase (USB, Cleveland, OH) or dye-terminator cycle sequencing (Perkin Elmer, Norwalk, CT). The LexA DNA binding domain-MS2 coat fusion protein is encoded by the plasmid pLexA-MS2, bearing the *TRP1* marker. The activation domain library was purchased from Clontech (Palo Alto, CA). The library plasmids carry the *LEU2* marker.

PHYLOGENY

The amino acid sequences were aligned using ClustalW (Thompson *et al.*, 1994). Phylogenies were estimated using a distance method, PROTDIST. The tree was built using a neighbor joining program, NEIGHBOR, and analyzed with SEQBOOT, all from the PHYLIP package of phylogeny programs (Felsenstein, 1993). Bootstrap analysis shows this tree is very highly supported. The use of bootstrapping in phylogeny estimation was introduced by Felsenstein (Felsenstein, 1985). It involves creating a new data set by sampling N characters randomly with replacement, so the new data set is the

same size as the original, but some characters have been left out and others duplicated. 100 new data sets were generated, analyzed by PROTDIST and NEIGHBOR. The tree shown is the consensus tree.

ISOLATION OF FULL-LENGTH CDNAS

Overlapping cDNAs were isolated from two libraries, a mouse germ cell cDNA library in λgt11 (M. Eddy, NIEHS) and a mouse testis cDNA Matchmaker library (Clontech), by plaque and colony hybridization, respectively, using the clone 3 partial cDNA as a probe (Sambrook *et al.*, 1989).

RNA ANALYSIS

Total RNA was isolated from dissected mouse tissues as previously described (Cathala *et al.*, 1983). RNA samples were electrophoresed in agarose-formaldehyde gels, transferred to nylon (hybond-N Amersham LIFE SCIENCES) and hybridized 15-20 hrs with radioactive (α-^{32}P) DNA probes prepared by random oligonucleotide-primed synthesis. The nylon membrane was washed at a final stringency of 0.1 X SSC and 0.5% SDS at 60°C and exposed to X-ray film.

ANTIBODIES

An antibody was raised against N14, a peptide in the amino-terminus of MSY4, whose amino acid sequence is KTTGGTTLPQAAADA. N14 was synthesized, purified by HPLC, and conjugated to BSA by the University of Washington Department of Pharmacology Protein Core. Antisera was raised by immunizing rabbits with N14. All rabbit work was done by R&R Rabbitry (Stanwood, WA). Affinity purified antibody to the BA2 region of p54 (Murray, 1994) and antibody to the purified p54/p56 proteins (Murray *et al.*, 1991) were kindly given to us by Mary Murray and affinity purified antibody to recombinant FRGY2 (Tafuri and Wolffe, 1992) was provided by Alan Wolffe. The α-BA2, α–p54/p56, and α-FRGY2 antibodies all crossreact specifically with MSY2. Polyclonal antibody to the HA tag was purchased from Clontech. Mouse monoclonal antibody to PRM2 is HUP2B (Stanker *et al.*, 1993).

IMMUNOHISTOCHEMISTRY

Immunohistochemistry was performed as previously described (Braun *et al.*, 1989b). Briefly, tissues were dissected from adult mice and fixed in Bouin's (testes) or Carnoy's (ovaries) overnight and embedded in paraffin. Sections were deparaffinized with xylene and rehydrated using standard procedures. Tissue sections were treated with primary antibody at overnight at 4°C or 2-3 hrs at room temperature. Biotinylated goat anti-rabbit IgG and streptavidin conjugated to horseradish peroxidase were used as recommended by the manufacturer (Zymed Laboratories, San Francisco, CA). Peroxidase

activity was visualized with the chromogen aminoethyl carbozole. Tissue sections were counterstained with hemotoxylin.

TISSUE EXTRACTS

Testis extracts were prepared as described (Dignam *et al.*, 1983) with the following modifications. Testes were dissected from mature mice and the tissue was put in 1 ml of buffer containing 10 mM KOAc, 10 mM POPSO pH 7.5, 1.5 mM MgOAc, 1 mm DTT and supplemented with protease inhibitors (1 mM PMSF, 1 μg/μl Pepstatin A, 1 μg/μl leupeptin) per gram of tissue. Cells were dounced using 25 strokes in a tight pestle. The lysate was pelleted at 3000 x g for 15 min in a fixed angle rotor. To the supernatant was added 1/10 volume of 0.3 M POPSO, 1.4 M KOAc, 1.5 mM MgOAc and glycerol to 20%.

Recombinant MSY4 was expressed in the BL21 strain of *E. coli* using the pET15-b/pLysS expression system (Novagen, Madison, WI). A cDNA clone encoding MSY4 was cloned into the pET15-b expression vector, and transformed into BL21. A single positive colony was used to inoculate 50 ml LB-ampicillin (100 μg/ml). At an OD_{600} of approximately 0.50 the culture was induced with a final concentration of 10 mM isopropyl β-D-thiogalactopyranoside (IPTG) for 2 hrs. Cells were harvested by centrifugation at 10,000 rpm for 10 min at 4°C. The cells were resuspended in testis extract buffer (10 mM KOAc, 10 mM POPSO pH 7.5, 1.5 mM MgOAc, 1 mm DTT) and

supplemented with protease inhibitors (1 mM PMSF, 1 μg/μl Pepstatin A, 1 μg/μl leupeptin) to approximately 5 mg/ml final concentration of protein extract.

IMMUNOPRECIPITATION

Co-immunoprecipitation of RNA was done with 50 μl of testis extract (approximately 12 mg/ml) incubated with 10 μl of rabbit sera and 200 μl of a 1:1 slurry of protein A agarose (Gibco-BRL) and lysis buffer (20 mM HEPES, 100 mM NaCl, 1.5 mM $MgCl_2$, and 0.5% NP40) for 2 hrs rocking at 4°C. The suspension was spun at 3000 x g for 4 min and the pellet was washed three times in 1 ml of cold lysis buffer. The final pellet was resuspended in 500 μl of solution D (20 mM HEPES, 20% v/v glycerol, 0.1 M KCl, 0.2 mM EDTA, and 0.5 mM DTT). The solution was treated with proteinase K at 0.2 μg/ml for 1 hr at 42°C, then the RNA was isolated and analyzed as described in **RNA Analysis**.

Co-immunoprecipitation of proteins was done as described above with the following modifications. 15 μl of testis extract was mixed with 1 μl of rabbit sera and 25 μl of slurry for 1 hr rocking at 4°C. The RNase treated testis extract was treated with RNase A at 0.5 μg/ml for 1 hr at room temperature prior to immunoprecipitation. The proteins were analyzed as in **Immunoblotting**.

RNASE H TREATMENT

RNA was treated with RNase H and 100 pmol oligo (dT) in 20 mM HEPES (pH 7.6), 50 mM KCL, 4 mM $MgCl_2$, 1 mM DTT and 0.5 μg BSA for 20 min at 37C. The RNA was isolated and analyzed as described in **RNA analysis**.

RNA PROBE PREPARATION

RNA probes were synthesized using SP6 bacteriophage polymerase using protocols suggested by the supplier (Promega, Madison, WI). *In vitro* transcription reactions were performed in 20 μl reaction volumes using 1 μg of linearized plasmid template, 500 μM ATP, UTP, and GTP, 25 μM CTP, and 50 μCi (α-^{32}P) CTP at 3000 ci/mmole (NEN Dupont, Boston, MA) for approximately 60 min. Samples were phenol-chloroform extracted, precipitated with ethanol, collected by microcentrifugation, and resuspended in 50 μl of H_2O.

RNA BINDING ASSAYS

RNA electrophoretic mobility shift assays (EMSA) were performed as described previously (Fajardo *et al.*, 1994) with the following modifications. RNA probes (1 μl, ~10^6 cpm) were combined with 1 μl of testis extract, in RNA binding buffer (20 mM HEPES (pH 7.6), 3 mM $MgCl_2$, 40 mM KCl, and 2 mM DTT) and incubated at room temperature for 20 min. The samples were then treated with 1 μl of RNAse T1

(Calbiochem, La Jolla, CA) for ten min at room temperature, then 2 µl of heparin (Sigma, St. Louis, MO) for another 10 min at room temperature. For the supershift analysis 1 µl of antibody was added after the RNA binding reaction and allowed to bind for 20 min at room temperature. The samples were combined with 5 µl of 50% glycerol and electrophoresed through a 4% polyacrylamide gel (60:1) for approximately 3 hrs at 4°C and 180V in a running buffer of 45 mM Tris (pH 8.3), 50 mM boric acid, and 1 mM EDTA. Gels were dried and visualized by autoradiography.

ULTRAVIOLET-CROSSLINKING

RNA binding was performed as described in **RNA Binding Assays**. Following the heparin treatment, the samples were placed on ice, with the tops of the microcentrifuge tubes open, and irradiated with ultraviolet light from a source located approximately 0.3 m away for 30 min. The irradiated samples were boiled in Laemmli buffer and analyzed by SDS-polyacrylamide gel electrophoresis using a 5% stacking gel and 10% resolving gel. Prestained molecular markers (Gibco-BRL Life Technologies, Rockville, MD) were used for size standards. The acrylamide gels were dried and visualized by autoradiography.

IMMUNOBLOTTING

Protein extracts were mixed with Laemmli buffer, boiled and electrophoresed in 8% SDS-polyacrylamide gels. The proteins were transferred to nitrocellulose (Gibco-BRL Life Technologies). After transfer the membrane was blocked for 30 min to several hrs at room temperature in 5% nonfat dry milk and phosphate buffered saline (BPBS) and then incubated overnight at 4°C with primary antibody at a 1:10,000 dilution. The membrane was washed once in BPBS with 0.05% Tween 20 and twice in BPBS, then incubated with secondary antibody conjugated to horseradish peroxidase (HRP) for several hrs at room temperature. After washing again as above, the HRP activity was detected using enhanced chemiluminescence (ECL) as described (Schneppenheim and Rautenberg, 1987). ECL reagent was prepared immediately prior to use by dissolving 40 mg of luminol (5-amino-2,3-dihydro-1,4-phthalazinedione) and 10 mg of 4-iodophenol in 1 ml of DMSO. Following the addition of 10 ml of 0.1 M Tris (pH 8.5), 5 ml of 5 M NaCl, 17 ml H_2O, and 125 µl H_2O_2, the membrane was incubated for 2 min and exposed to X-ray film.

POLYSOME ANALYSIS

Each testis was dissected from an adult mouse and homogenized in 1 ml homogenization buffer (100 mM NaCl, 1.5 mM $MgCl_2$, 20 mM POPSO (pH 7.5), and 1 mM PMSF). The nuclei and mitochondria were collected by centrifugation for 2 min at 12,000 x g, and the supernatant was layered over a 11 ml linear 15-50% (w/w) sucrose

gradient in lysis buffer and centrifuged in a Beckman SW40 rotor for 110 min at 36,000 rpm. The gradients were fractionated into 12 1 ml fractions using an Isco Density Gradient Fractionator (Model 185), while monitoring ultraviolet absorbance at 254 nm. As a control to verify mRNA association with polysomes, equivalent supernatants were prepared and centrifuged in sucrose gradients in buffer in which the $MgCl_2$ was replaced by 20 mM EDTA. The presence of EDTA causes mRNA and ribosomes to disassociate. Northern and western analysis was performed on each fraction. For western analysis 200 µl of each fraction was concentrated and analyzed as described in **Immunoblotting**. For northern analysis 500 µl of each fraction was treated with proteinase K at 0.2 ug/ml for 90 min at 55°C, and 100 µl was analyzed as described in **RNA analysis**.

TRANSGENIC CONSTRUCTS

A heterologous reporter was used to evaluate translational control function *in vivo* as previously described (Braun *et al.*, 1989b; Fajardo *et al.*, 1997; Zhong *et al.*, 1999). This reporter cassette contains 4.1kb of mouse *Prm1* 5' untranscribed sequence up to the transcriptional start site, chimeric 5'UTR of 159bp (91 bp of *Prm1* 5'UTR, 7bp of linker DNA, and 61bp of *hGH* 5'UTR) and the complete *hGH* coding sequence and introns (DeNoto *et al.*, 1981). Oligonucleotides that encode the *Prm1* $_{1\text{-}37\ (mutCAUC)}$ were inserted into the plasmid at a BamHI site 3' to the *hGH* ORF and 5' to the 5'-most 23nt of *Prm1* 3'UTR that contains the polyadenylation site. 140 bp of sequence downstream of the polyadenylation signal is also present to ensure proper 3' processing of the mRNA.

Prm1 sequences were used to express HA tagged MSY4 in elongated spermatids. The transgene contains 4.1 kb upstream *Prm1* sequence, 95 bp *Prm1* 5' UTR, 270 bp *Prm1* 3' sequence. *Msy4* was cloned into a BamHI site by PCR using the primers 5' CGCGGATCCATGAGCGAGGCGGGCGAGGCC 3' and 5' CGCGGATCCTCAAGCGTAGTCTGGGACGTCGTATGGGTACTCGGCACTGCTCT GTTCGG 3', which includes the sequence for the HA epitope in the 3' PCR primer and BamHI sites in both PCR primers. The PCR template was a cDNA with the complete *Msy4* ORF in pGAD10. The transgenic protein is missing 76 of 85 amino acids in the amino- terminus. The first six amino acids are present,76 are missing, then 3 amino acids before the CSD are present.

Transgenic mice were generated by microinjecting a purified DNA fragment at a concentration of 2 ng/μl in 10 mM Tris pH 7.5 and 0.25 mM EDTA into pronuclei of fertilized eggs dervived from FVB x FVB matings. Pseudopregnant B6 CBA foster females were used for oviduct implantation of eggs that survived microinjection. Transgenic animals were identified by PCR.

CHAPTER 2: CLONING AND CHARACTERIZATION OF *MSY4*

Holly G. Davies, Flaviano Giorgini, Mark A. Fajardo, and Robert E. Braun. A Sequence-Specific RNA Binding Complex Expressed in Murine Germ Cells Contains MSY2 and MSY4, Developmental Biology (accepted)

INTRODUCTION

Translational control is an important form of gene regulation, especially during gametogenesis and embryogenesis, in which there are periods of no transcription (Curtis *et al.*, 1995; Stebbins Boaz and Richter, 1997). During both oogenesis and spermatogenesis, repressed mRNAs are stored in ribonucleic protein (RNP) particles. In growing oocytes mRNA is synthesized and stored in mRNP particles in preparation for the suspension of transcription during oocyte maturation. These pre-existing mRNAs are translated to make the proteins necessary for the completion of oogenesis and the beginning of embryogenesis. Maternal mRNAs from the oocyte are needed at least until zygotic transcription begins. Transcription ceases in spermatogenesis during the third and final stage, spermiogenesis, in which haploid round spermatids differentiate into mature spermatozoa. Translation of repressed messages is required for the synthesis of numerous proteins required for flagellar and nuclear morphogenesis.

Among the genes that are regulated by translational control during spermatogenesis are protamine 1 (*Prm1*), protamine 2 (*Prm2*), and the transition proteins *Tnp1* and *Tnp2* (Kleene, 1996). The transition proteins and the protamines compact the chromatin during the terminal stage of spermatid differentiation. The protamines are transcribed in round spermatids, and their mRNA stored in mRNP particles in the cytoplasm for 2-8 days, until translated in elongated spermatids (Kleene *et al.*, 1984). Failure to delay the translation of the *Prm1* mRNA leads to precocious nuclear condensation and male sterility (Lee *et al.*, 1995). The mechanism of translational control of the protamines is unknown, but the *cis*-acting regulatory elements have been identified within the 3' untranslated region (UTR) (Braun *et al.*, 1989b). Two separate regions of the *Prm1* 3' UTR, the 3' most 62 nucleotides (nt) and the 5' most 37 nt ($Prm1_{1\text{-}37\ (wt)}$) (Figure 1 and Fajardo *et al.*, 1997) are sufficient for *Prm1*-like translational delay.

Several proteins that bind to the *Prm1* 3' UTR have been described. Among these are the Y box protein MSY2 (Gu *et al.*, 1998), a single-stranded DNA binding protein TB-RBP (Kwon and Hecht, 1993), a double-stranded RNA binding protein PRBP (Lee *et al.*, 1996), and a sequence-specific RNA binding activity of 48/50 kDa (Fajardo *et al.*, 1994). The molecular function of most of these proteins is unknown, although mice nullizygous for *Tarbp2*, which encodes PRBP, are sterile and have defects in the translational activation of the protamine mRNPs (Zhong *et al.*, 1999).

The 48/50 kDa sequence-specific RNA binding proteins are likely components of the protamine mRNPs. The proteins bind to a highly conserved sequence in the 5'-most

region of *Prm1* and *Prm2* 3' UTRs (Fajardo *et al.*, 1994) and (Giorgini et al., in preparation), and a transgenic mRNA that contains the binding site is translationally repressed *in vivo* (Fajardo *et al.*, 1997). The binding activity is present in the cytoplasm of round spermatids, where protamine mRNAs are repressed, and is absent in elongated spermatids, where protamine mRNAs are translated.

In this paper we describe the cloning of the gene that encodes the 48/50 kDa proteins, and show that the RNA binding activity present in testis extracts consists of two Y box proteins.

RESULTS

THREE-HYBRID CLONING OF MSY4

In hopes of cloning the gene(s) that encodes the previously described 48/50 kDa proteins that bind to the $Prm1_{1\text{-}37\ (wt)}$ RNA (Figure 2-1), I screened a library of mouse testis cDNAs linked to the GAL4 transcriptional activation domain using the yeast three-hybrid system (SenGupta *et al.*, 1996). In this assay, interaction of an RNA-binding protein with an RNA leads to transcriptional activation of a reporter gene in yeast (Figure 2-2). As bait I used a hybrid RNA, which included the first 37 nt of the *Prm1* 3' UTR, $Prm1_{1\text{-}37\ (wt)}$, and two copies of the MS2 coat protein recognition site.

Transformants were initially selected for HIS3 protrophy and positives were subsequently screened for β-galactosidase activity.

To identify RNA-dependent candidates I selected for loss of the plasmid encoding the hybrid RNA and rescreened for β-galactosidase activity. RNA-dependent candidates had β-galactosidase activity only when the hybrid RNA was present. Candidates that failed to activate *lacZ* after plasmid loss were tested for RNA binding specificity by reintroducing plasmids encoding hybrid RNAs with $Prm1_{1\text{-}37(wt)}$, and two negative controls, $Prm1_{1\text{-}37\ (mut)}$, or $Prm1_{134\text{-}156}$. The 48/50 kDa binding activity does not to bind to $Prm1_{1\text{-}37\ (mut)}$, an RNA that contains mutations in 8 of the nucleotides conserved between *Prm1* and *Prm2* in that region (Figure 2-1), or $Prm1_{134\text{-}156}$, an RNA consisting of the 3'-most 23 nt of the *Prm1* 3' UTR (data not shown). One transformant, clone 3, satisfied all criteria.

Figure 2-3A shows the binding specificity of clone 3 for several RNAs. *HIS3 /lacZ* expression was activated when the $Prm1_{1\text{-}37\ (wt)}$ was present, however, $Prm1_{1\text{-}37\ (mut)}$, and $Prm1_{1\text{-}37\ (C26A)}$, which contains a single nucleotide change within the binding site that abolishes protein binding (Figure 2-1), both eliminated activation. The plasmids encoding $Prm1_{134\text{-}156}$ or the MS2 binding sites alone (pIII/MS2-2) also failed to activate.

Clone 3 was sequenced and determined to be incomplete at both its 5' and 3' ends. To clone a full-length cDNA, clone 3 was used to screen a λgt11 mouse germ cell and Matchmaker testis cDNA libraries. Several overlapping clones containing the 5' UTR, open reading frame (ORF) and 3' UTR were characterized. The composite cDNA is 1693bp long and has an open reading frame which encodes a predicted protein of 358 amino acids (Figure 2-4A).

A Blast search of the NIH non-redundant database revealed that I cloned a new member of the Y box family of nucleic acid binding proteins. Members of this family include DBPA (Coles *et al.*, 1996; Sakura *et al.*, 1988) and YB-1 (Didier *et al.*, 1988) in humans, FRGY1 and FRGY2 (Deschamps *et al.*, 1992) in frogs, and MSY1 (Tafuri *et al.*, 1993) and MSY2 (Gu *et al.*, 1998) in mice (reviewed by Matsumoto and Wolffe, 1998). We refer to this new Y box member as MSY4.

All family members contain the highly conserved cold shock domain (CSD) that is 43% identical from *E. coli* to humans (Figure 2-4A, underlined). The structure of the CSD has been solved for bacterial members and is a five-stranded β-barrel (Schindelin *et al.*, 1994; Schindelin *et al.*, 1993). Within the CSD are the RNA binding motifs RNP-1 and RNP-2, present on the β2 and β3 strands, respectively (Figure 2-4A, double underlined). MSY4, and the other Y box proteins, contain a variable amino-terminus, the CSD, then a tail with four alternating basic and acidic regions. The amino acids in the tail are not highly conserved among family members, but the structure of alternating basic and acidic regions is conserved. The structure of several family members is shown in Figure 2-4B.

A phylogeny of Y box family members is shown in Figure 2-4C. MSY4 is closest to human DBPA, and is its mouse ortholog. MSY1 clusters with GFYP1, FRGY1, chicken YB-1, human YB-1, and rabbit p50, while MSY2 clusters with GFYP2 and FRGY2 (p54/p56).

MSY4 IS LOCALIZED TO THE CYTOPLASM OF GERM CELLS

By northern analysis, *Msy4* mRNA was most abundant in the testis (Figure 2-5, upper panel). Lower levels were detected in skeletal muscle and upon longer exposure low levels could be detected in all other tissues (H.D. unpublished). The membrane was reprobed with *GAPDH* as a control for levels of RNA in each lane (Figure 2-5, lower panel).

To determine if MSY4 is found in the cytoplasm of round spermatids during the time *Prm1* mRNA is translationally repressed, I raised an antibody to a peptide in the N-terminus of MSY4 (N14) and performed immunohistochemistry on sections of adult mouse testis. MSY4 protein was first detected in the cytoplasm of mid-pachytene cells (Figure 2-6B stage VIII tubule). The signal was strongest in late pachytene spermatocytes (Figure 2-6D, E and F) and early round spermatids (Figure 2-6B stages IV and V, and 2-6C), and persisted through nuclear elongation (Figure 2-6D and 2-6B stage XI). MSY4 was not detected in elongated spermatids (Figure 2-6B stages IV and V, and 2-6C). This expression of MSY4 is consistent with it functioning in the translational repression of *Prm1* mRNA in round spermatids.

The predominant proteins associated with masked messages in *Xenopus* oocytes, mRNP3+4, are Y box proteins (Darnbrough and Ford, 1981). Despite the low level of *Msy4* message in ovaries by northern analysis, α-MSY4 antibody specifically labeled the cytoplasmic compartment of both preantral and antral follicles in mouse ovaries (Figure 2-6H). The presence of MSY4 in the cytoplasm of male and female germ cells, which

both contain an abundance of masked messages, suggests that MSY4 is involved in mRNA storage.

THE 48/50 KDA ACTIVITY CONTAINS MSY4 AND MSY2

To determine if the previously described 48/50 kDa binding activity contains MSY4 we performed several experiments utilizing MSY4 antibody. The molecular weight of the 48/50 kDa activity was originally determined by crosslinking a radiolabeled *Prm1* 3' UTR RNA to the binding activity present in testis extracts, sizing the complexes on a denaturing polyacrylamide gel, and subtracting the molecular weight of the RNA component from the complexes (Fajardo *et al.*, 1994). The MSY4 polyclonal antibody specifically recognized proteins of 55/58 kDa in testis extract (Figure 2-7A, lane 1) and a protein of 55 kDa in a bacterial extract in which MSY4 had been expressed (lane 2). Although the predicted size of the protein encoded by the ORF is 38 kDa, it is usual for Y box proteins to run anomalously on SDS-PAGE gels (Deschamps *et al.*, 1992). The MSY4 antibody also recognized a protein of 60 kDa present in extracts prepared from transformed as well as untransformed bacterial cells.

The 48/50 kDa proteins in testis extract migrated at 53/55 kDa when UV-crosslinked to a radiolabeled *Prm1* $_{1\text{-}37(wt)}$ RNA (Figure 2-7C, lane 1). Recombinant MSY4 expressed in bacterial cells was also UV-crosslinked to radiolabeled *Prm1* $_{1\text{-}37(wt)}$ RNA (Figure 2-7B, lanes 2 and 3), but not to radiolabeled RNA with a point mutation, *Prm1* $_{1\text{-}37(C26A)}$ (Figure 2-7B, lanes 5 and 6). Both probes were seen to be intact on an acrylamide gel (data not shown). The UV-crosslinked species from testis extract were

dependent on the MSY4 Y box recognition sequence (YRS) within the *Prm1* $_{1\text{-}37\ (wt)}$ RNA, as the *Prm1* $_{1\text{-}37(mut)}$ did not form either complex (Figure 2-7C, lane 2). Antibody against MSY4 immunoprecipitated a portion of the UV-crosslinked complexes of 53/55 kDa from testis extract (Figure 2-7C, lane 5). Preimmune sera did not precipitate either complex (lane 7).

The 48/50 kDa complex can also be detected by an electrophoretic mobility shift assay (EMSA). We attempted to supershift the EMS complex with MSY4 antibody, but were unable to do so when using an RNA probe containing a single (sgl) copy of the binding site (Figure 2-7D, lane 3 arrow). However, an RNA containing two (dbl) copies of the MSY4 Y-box recognition sequence (YRS) also formed an EMS complex, and was supershifted by MSY4 antibody (Figure 2-7D, lane 6), but not by preimmune sera (lane 5). The ability to immunoprecipitate the UV-crosslinked complex, and supershift the EMS complex, strongly suggests that MSY4 is a component of the 48/50 kDa binding activity.

The failure to supershift all of the 48/50 kDa EMS complex (Figure 2-7D, lane 6), suggested that the complex contained protein(s) in addition to MSY4. Mouse testis extracts also contain MSY2, the ortholog of the *Xenopus* Y box protein FRGY2 (mRNP3+4), the major mRNP protein in oocytes. The similarity in the MSY2 and MSY4 cold shock domains suggested that MSY2 might also bind to the MSY4 YRS, and be contained in the 48/50 kDa complex. To test this possibility antibody against FRGY2, which crossreacts with MSY2 (Kwon *et al.*, 1993), was used to examine the 48/50 kDa

complex. Three different antibodies to FRGY2, all of which specifically crossreact with MSY2, were used, (Tafuri and Wolffe, 1992), (Murray, 1994) and (Murray *et al.*, 1991). α-FRGY2 detected a single band in mouse testis extract that migrated slightly lower than MSY4 (Figure 2-7A, lane 4). This antibody immunoprecipitated a portion of the UV-crosslinked complexes of 53/55 kDa from mouse testis extract (Figure 2-7C, lane 3). FRGY2 antibody was also able to supershift the EMS complex with both the *Prm1* $_{1\text{-}37(wt)}$ probe (Figure 2-7D, lane 7) and the longer probe with the additional MSY4 YRS (lane 8). Antibody that had been heated to destroy the antibody did not supershift the EMS (F.G. unpublished).

Since both α-MSY4 and α-MSY2 antibodies were able to immunoprecipitate a portion of the UV-crosslinked complexes and supershift a portion of the EMS complex, it suggests that the 48/50 kDa binding activity contains both MSY4 and MSY2. It further suggests that MSY4 and MSY2 may be part of the same complex. To test this possibility, co-precipitation experiments were performed. MSY4 antibody immunoprecipitated MSY2 from testis extract (Figure 2-E, lane 2) and FRGY2 antibody immunoprecipitated MSY4 (lane 5). Both interactions were RNA-dependent, as the treatment of the extract with RNase A abolished most of the interaction (lanes 3 and 6). MSY4 and MSY2 antibodies were also able to immunoprecipitate MSY4 and MSY2, respectively, both with and without the RNase A treatment (data not shown). Preimmune sera did not immunoprecipitate either MSY2 or MSY4 (data not shown).

The ability of MSY2 antibody to supershift the 48/50 kDa EMS complex suggests that MSY2 is contained within the complex, but does not necessarily show that MSY2 can bind the YRS sequence directly. To determine if MSY2 binds the YRS I cloned a cDNA encoding MSY2 and expressed it as a fusion protein with the GAL4 activation domain in the yeast three-hybrid system. Figure 2-3B shows the binding specificity of MSY2 for several RNAs. *HIS3 /lacZ* expression was activated when the $Prm1_{1\text{-}37\ (wt)}$ was present, however, $Prm1_{1\text{-}37\ (mut)}$, and $Prm1_{1\text{-}37(C22G)}$, which contains a single nucleotide change within the binding site that abolishes protein binding (Figure 2-1 and Figure 2-3A), both eliminated activation. The plasmid encoding $Prm1_{134\text{-}156}$ also failed to activate. Therefore, like MSY4, MSY2 is a sequence-specific RNA binding protein.

MSY4 ASSOCIATES WITH PRM1 RNA IN VIVO

To ascertain if MSY4 is associated with mRNPs *in vivo* we fractionated testis extract on a sucrose gradient and performed northern (Figure 2-8A) and western (Figure 2-8B) analysis on each fraction. The location of *Prm1* mRNA was used to define the fractions. *Prm1* was detected as a larger message in the mRNP fractions (Figure 2-8, lanes 2 and 3) and as a mixture of this larger message and shorter messages in the polysome fractions due to deadenylation (Figure 2-8A, lanes 5 and 6). The majority of MSY4 protein was detected in the mRNP fractions (Figure 2-8B, fractions 2 and 3), with a smaller amount in the monosomes (fraction 4), disomes (fraction 5) and polysomes (fractions 6 and 7). Equivalent gradients were run in the presence of EDTA, which

disassociates RNA from ribosomes. In the presence of EDTA all the *Prm1* mRNA and MSY4 protein were detected in the mRNP fractions (Figure 2-8A and B, lower panels).

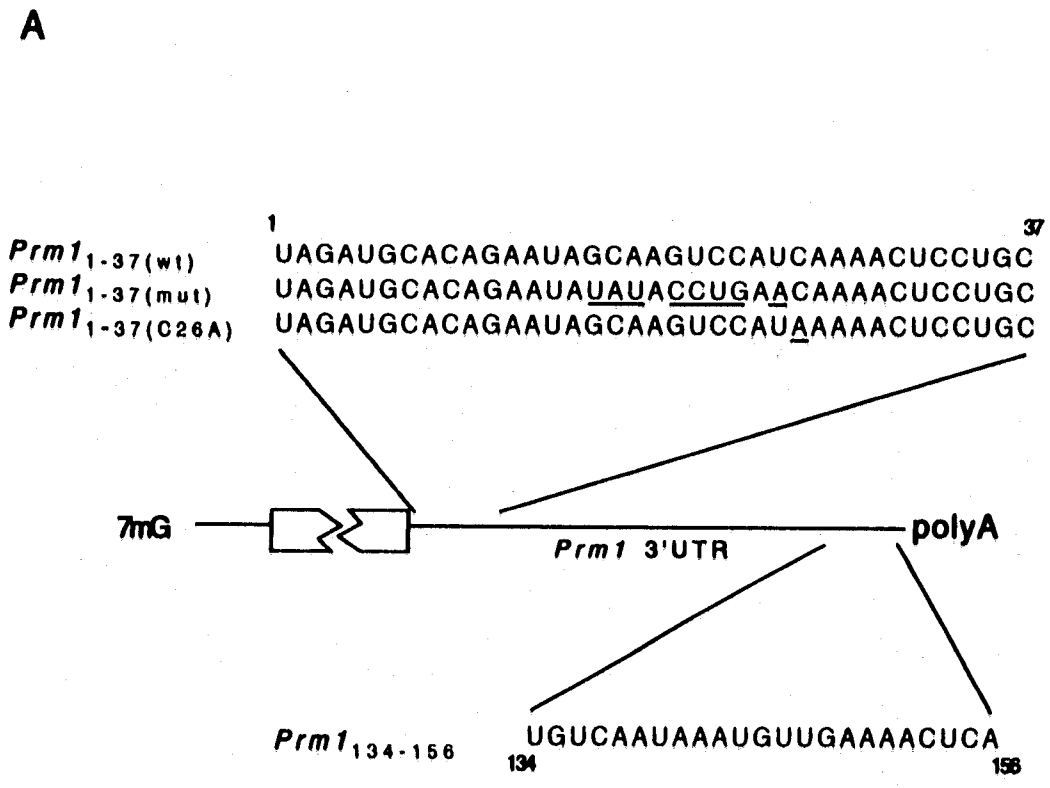
To determine if MSY4 associates with *Prm1* mRNA *in vivo*, I immunoprecipitated MSY4 from testis extract, isolated RNA, and probed for the presence of *Prm1* mRNA by northern analysis. *Prm1* RNA co-precipitated with MSY4 antibody (Figure 2-9A, lane 1), but not with preimmune sera (lane 2). *Prm1* mRNA is detected as a polyadenylated message in the repressed mRNP, and as a deadenylated message when it is being actively translated on polysomes (Figure 2-8A and Kleene, 1989).

To determine if MSY4 is associated with the long or short form of the *Prm1* mRNA, the RNA that was co-precipitated with the MSY4 antibody was analyzed by northern blotting along with total testis RNA that had been treated with oligo(dT) and RNase H to remove the poly(A) tails. As shown in Figure 2-9B, the α-MSY4 antibody co-precipitated the fully adenylated form of the *Prm1* message, strongly suggesting that MSY4 is bound to the repressed form of the *Prm1* message found in mRNP particles.

To determine if MSY4 was specifically bound to *Prm1* mRNA, I also probed for the presence of *Tenr* mRNA, an abundant message that is also under translational control (Schumacher *et al.*, 1995). I detected *Tenr* message in MSY4 immunoprecipitate (Figure 2-9A, middle panel, lane 2), demonstrating that MSY4 binds *Tenr* and *Prm1* mRNAs *in vivo*. To determine if other RNAs were also co-precipitated with α-MSY4 I probed for *GAPDH* and *Actin*, two mRNAs that are not known to be under translational control

during spermatogenesis. Surprisingly, both mRNAs were co-precipitated (Figure 2-9A and data not shown). While no specificity in which mRNAs were immunoprecipitated was observed, the mRNA immunoprecipitation was specific to the MSY4 antibody compared to preimmune sera. These data suggest that MSY4 may associate with many mRNAs *in vivo* (see Discussion), or that the immunoprecipitation conditions did not permit us to detect specific interactions.

Figure 2-1 (*A*) Sequences in the *Prm1* 3' untranslated region (UTR) used in this study. *Prm1* $_{1\text{-}37\,(wt)}$ is the 5'-most 37 nt of the *Prm1* 3' UTR. The altered nucleotide(s) in *Prm1* $_{1\text{-}37\,(mut)}$ and *Prm1* $_{1\text{-}37\,(C22G)}$ are underlined. *Prm1* $_{1\text{-}37\,(mut)}$ consists of mutations in 8 nt that are conserved between *Prm1* and *Prm2* in that region. *Prm1* $_{1\text{-}37\,(C26A)}$ contains a point mutation in which the C in position 26 has been changed to a A. The 3'-most 23 nt, *Prm1* $_{134\text{-}156}$, includes the polyadenylation signal. (*B*) Electrophoretic mobility shift analysis (EMSA) of the *Prm1* 3'UTR binding activity in mouse testis extracts. The probes are as described in (*A*). The upper arrows indicate the sequence-specific 48/50 kDa binding activity. The lower arrow indicates free probe.


A
1
37
Prm1 1-37(wt) UAGAUGCACAGAAUAGCAAGUCCAUCAAAACUCCUGC
Prm1 1-37(mut) UAGAUGCACAGAAUAUAUACCUGAACAAAACUCCUGC
Prm1 1-37(C26A) UAGAUGCACAGAAUAGCAAGUCCAUAAAAACUCCUGC
7mG
Prm1 3'UTR
polyA
Prm1 134-156 UGUCAAUAAAUGUUGAAAACUCA
134
156

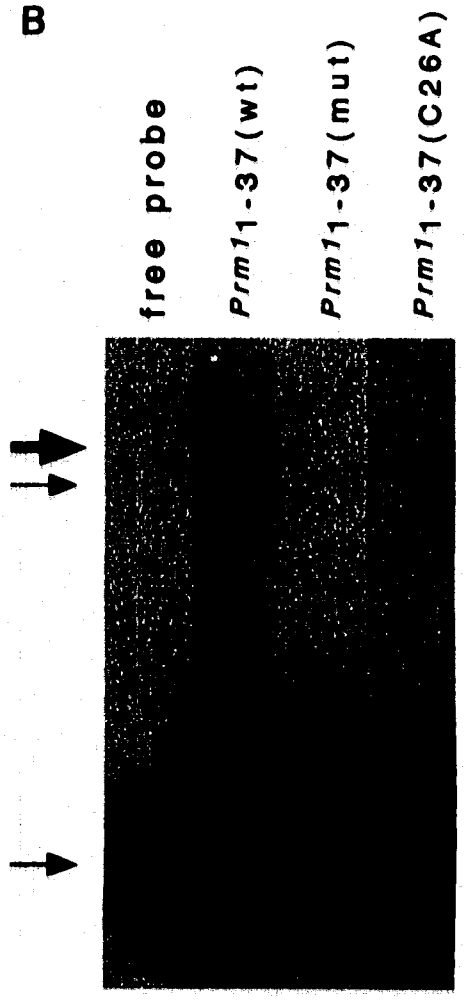

B
free probe
Prm11-37(wt)
Prm11-37(mut)
Prm11-37(C26A)

Figure 2-2 The three-hybrid screen. *Prm1* $_{1\text{-}37\,(wt)}$ was used as bait in a hybrid RNA with the bacteriophage MS2 coat protein RNA binding site to screen a mouse testis cDNA-GAL4 activation domain fusion library. Binding of a mouse testis protein to the *Prm1* $_{1\text{-}37\,(wt)}$ results in the transcriptional activation of *HIS3* and *lacZ* reporter genes.

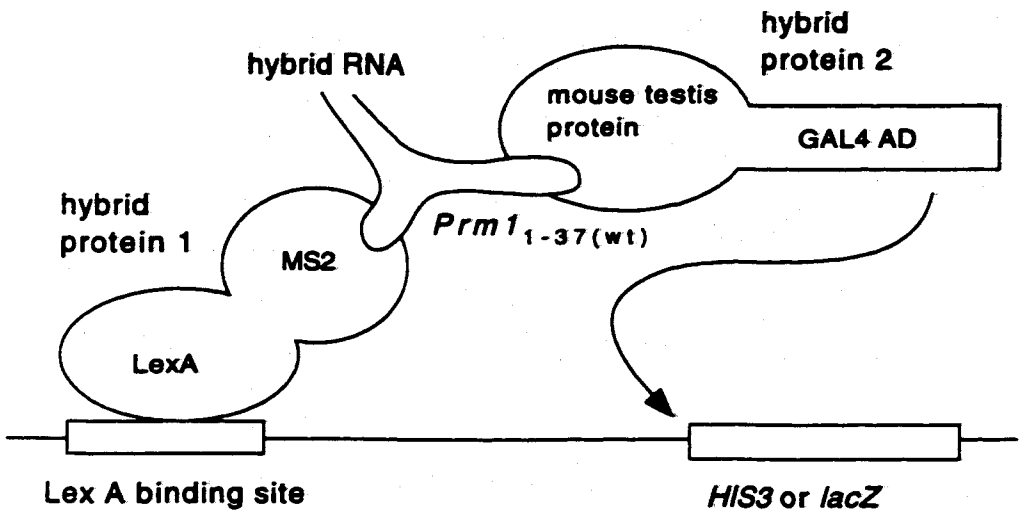
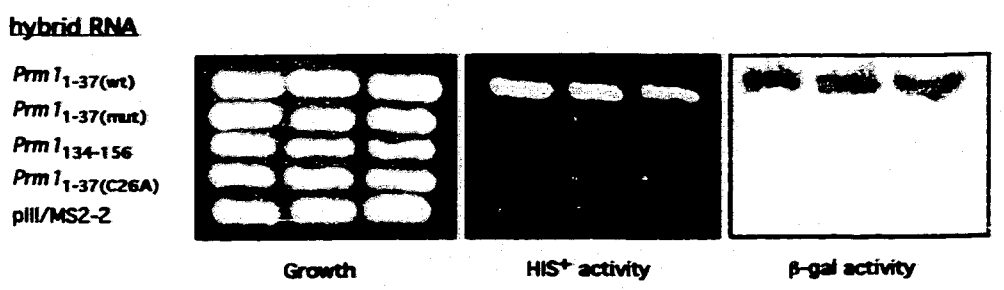

hybrid RNA
hybrid protein 2
mouse testis protein
GAL4 AD
hybrid protein 1
MS2
Prm1 1-37(wt)
LexA
Lex A binding site
HIS3 or lacZ

Figure 2-3 Sequence-specific RNA binding of MSY4 (*A*) and MSY2 (*B*) in the three-hybrid system. The yeast strains are able to grow on media lacking leucine, tryptophan, and uracil, selecting for the three plasmids used in the three-hybrid assay (left panels). MSY4 and MSY2 specifically bind only the *Prm1* $_{1\text{-}37\,(wt)}$, as shown by growth on media also lacking histidine (with 5 mM 3AT) (center panels) and by β-galactosidase activity in a filter assay (right panels). *Prm1* $_{1\text{-}37\,(wt)}$, *Prm1* $_{1\text{-}37\,(mut)}$, *Prm1* $_{134\text{-}156}$ and *Prm1* $_{1\text{-}37\,(C22G)}$ are as previously described. pIII/MS2-2 encodes the MS2 RNA binding site alone.

A MSY4


hybrid RNA
Prm1 1-37(wt)
Prm1 1-37(mut)
Prm1 134-156
Prm1 1-37(C26A)
pIII/MS2-2
Growth
HIS+ activity
β-gal activity

B MSY2

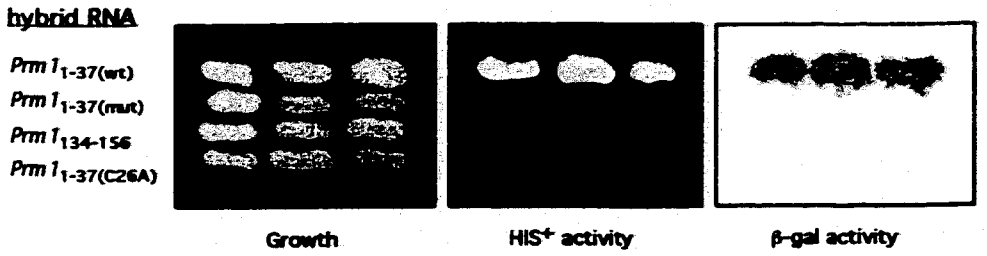
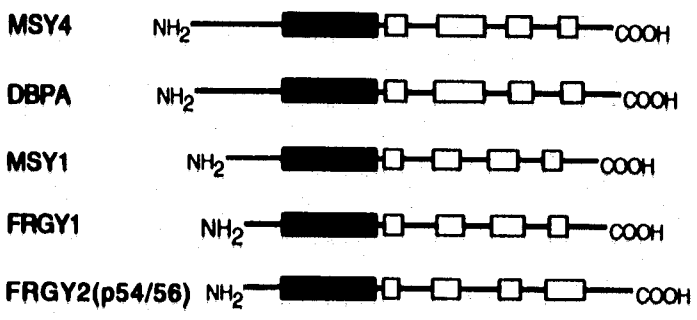

hybrid RNA
Prm1 1-37(wt)
Prm1 1-37(mut)
Prm1 134-156
Prm1 1-37(C26A)
Growth
HIS+ activity
β-gal activity

Figure 2-4 (*A*) The predicted amino acid sequence of MSY4. The cold shock domain (CSD) is underlined. The RNA binding domains within the CSD, RNP-1 and RNP-2, are double underlined. (*B*) Structure of MSY4 and other family members: human DBPA, mouse MSY1, and *Xenopus* FRGY1 and FRGY2. All contain the variable N-terminus, the highly conserved CSD (closed box), and the tail of alternating basic and acidic regions (open boxes and lines). Although the identity of the amino acids in the tail are not conserved, the basic structure of four alternating basic and acidic regions is conserved. (*C*) Phylogeny of Y box proteins from frog (FRGY1 and FRGY2), goldfish (GFYP1 and GFYP2), chicken (chicken YB1), mouse (MSY1 and MSY2), rabbit (p50), and human (YB1 and DBPA). *Aplysia* APY1was used as an outgroup. The tree was estimated from amino acid sequences using a distance method, thus the branch lengths are meaningful.

A

```
MSEAGEATTG GTTLPQAAAD APAAAPPDPA PKSPAASGAP QAPAPAALLA GSPGGDAAPG 60
PAPASSAPAG GEDAEKKVLA TKVLGTVKWF NVRNGYGFIN RNDTKEDVFV HQTAIKKNNP 120
RKYLRSVGDG ETVEFDVVEG EKGAEAANVT GPDGVPVEGS RYAADRRRYR RGYYGRRRGP 180
PRNYAGEEEE EGSGSSEGFE PPAADGQFSG ARNQLRRPQY RPPYRQRRFP PYHVGQTFDR 240
RSRVFPHPNR MQAGEIGEMS DGVPEGTQLQ AHRNPTYRPR FRRGPARPRP APAIGEAEDK 300
ENQQAANGPN QPSARRGFRR PYNYRRRSRP NAVSQDGKET KAGEAPTAPA PTEQSSAE   358
```

B



C

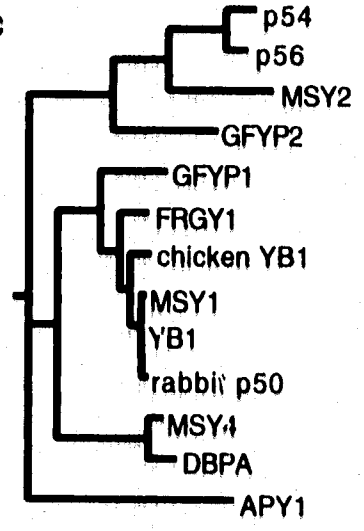


Figure 2-5 Northern analysis of *Msy4* expression in adult mouse tissues. 10 μg of RNA from each tissue was resolved on a 1.5% agarose-formaldehyde gel and transferred to a nylon membrane. The ^{32}P-labeled 1.2 kilobase (kb) clone 3 cDNA was used as a probe for *Msy4*. The size standards, in kb, are on the left (Gibco BRL-Life Technologies). The ^{32}P-labeled 1.3 kilobase (kb) glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) cDNA from chicken was used as a control (Dugaiczyk *et al.*, 1983). *GAPDH* is not detected at equal levels in all tissues, and is detected at higher levels in the brain, heart and kidney (Ambion, Austin, TX). Alternative forms are found in the testis (Mezquita *et al.*, 1998).

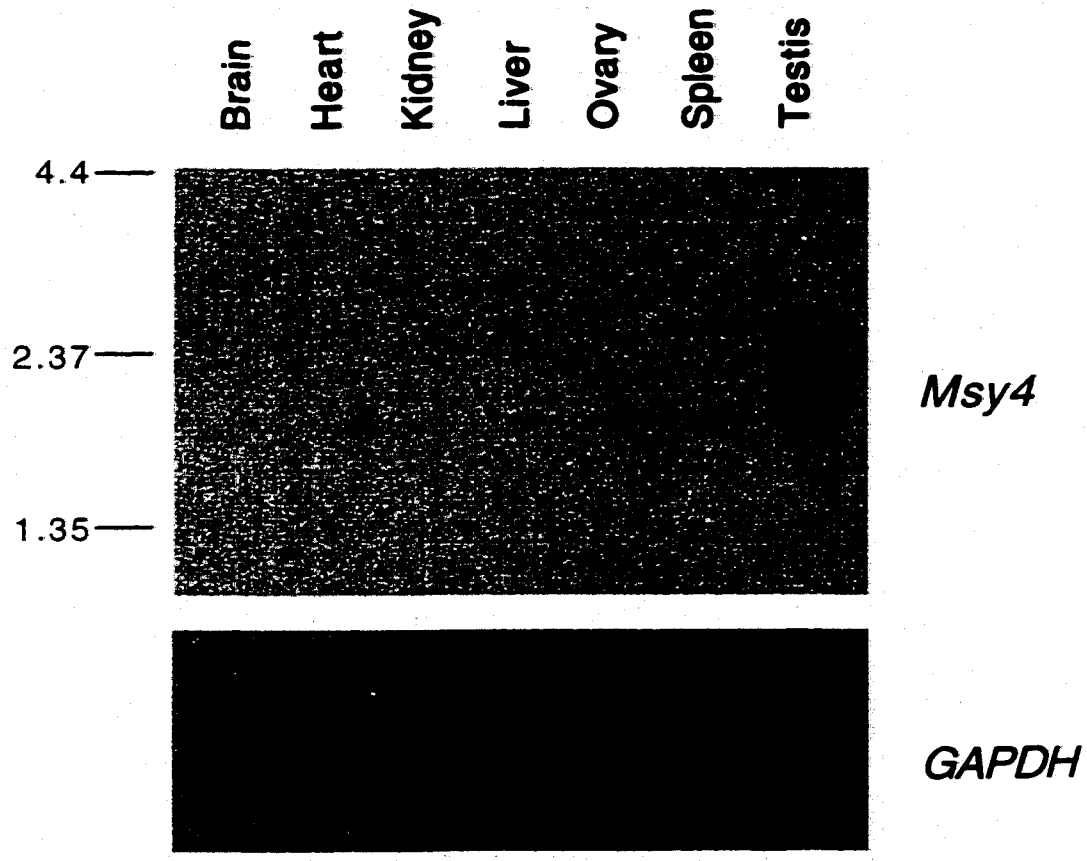
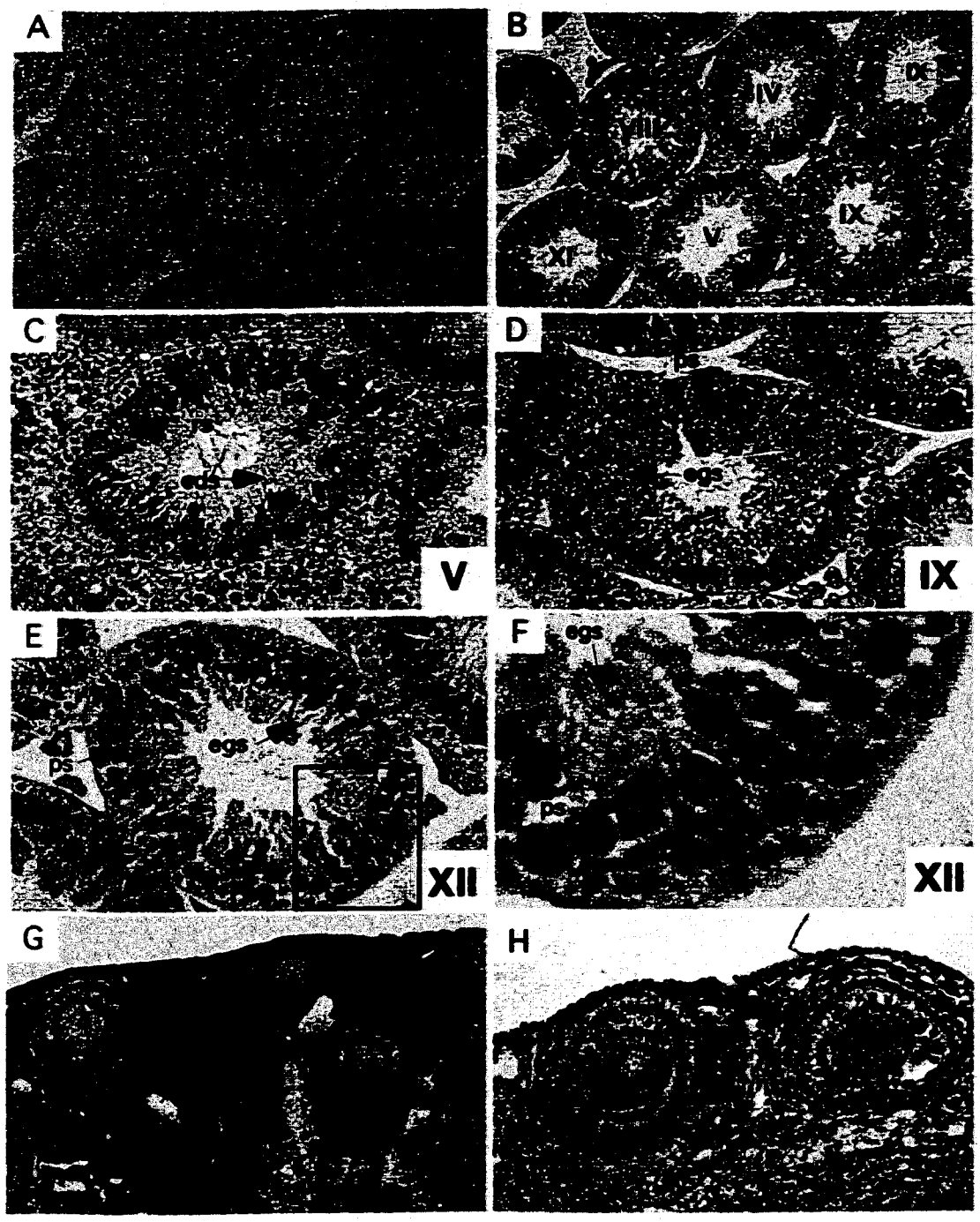

Brain
Heart
Kidney
Liver
Ovary
Spleen
Testis
4.4
2.37
1.35
Msy4
GAPDH

Figure 2-6 Immunohistochemical localization of MSY4 in adult mouse testis and ovary. The sections were counterstained with hemotoxylin. (*A-F*) Testis sections stained with (*A*) preimmune sera and (*B-F*) α-MSY4. (*F*) is an enlargement of the boxed area in (*E*). Roman numerals indicate the stage of the tubule (Russell et al., 1990). The germ cells are indicated as pachytene spermatocytes (ps), round spermatids (rs), elongating spermatids (egs) and elongated spermatids (eds). (*G* and *H*) Ovary sections stained with (*G*) preimmune sera and (*H*) α-MSY4. The arrows denote the oocytes within the follicles.


A
B
IV
DC
III
IX
V
XI
C
egs
V
D
egs
IX
E
egs
ps
XII
F
egs
ps
XII
G
H

4% native acrylamide gel, and visualized by autoradiography. *Prm1* $_{1\text{-}37\,(wt)}$ was used as the labeled probe (sgl= single binding site) in lanes 1,2,3 and 7. A probe with an additional YRS (dbl= double binding site) was used in lanes 4,5,6 and 8. (*E*) Co-immunoprecipitation of MSY2 and MSY4 from testis extracts of adult mice. Proteins were immunoprecipitated, resolved on an 8% SDS-PAGE gel, transferred to nitrocellulose, and probed with antibody to MSY2 (left panel, (Murray et al., 1991) or MSY4 (right panel). Extracts in lanes 3 and 6 had been treated with RNase A prior to immunoprecipitation. Lanes 1 and 4 are testis extracts that have not been immunoprecipitated. Size standards in kDa are shown on the right (Gibco BRL-Life Technologies).

Figure 2-7 The 48/50 kDa binding activity contains both MSY4 and MSY2. (*A*) Immunoblot of protein extracts from adult mouse testis, bacteria expressing MSY4, and untransformed bacteria. Aliquots of approximately 30 μg of each extract were resolved on a 8% SDS-PAGE gel, transferred to nitrocellulose, and probed with antibody to MSY4 or FRGY2 (Murray, 1994). Both antibodies specifically recognize either MSY4 or MSY2, respectively. The size standards, in kilodaltons (kDa), are indicated on the right (Gibco BRL-Life Technologies). (*B*) Specific binding of recombinant MSY4 to ^{32}P-labeled *Prm1* $_{1\text{-}37(wt)}$ RNA in lanes 2-3 compared to ^{32}P-labeled *Prm1* $_{1\text{-}37(C26A)}$ RNA in lanes 5-6. Lanes 1 and 4 are uninduced extracts, lanes 2 and 5 are extracts induced for 1 hour, and lanes 3 and 6 are extracts induced for 2 hours. 20 ug of each extract was resolved on a 10% SDS-PAGE gel and visualized by autoradiography. (*C*) Immunoprecipitation of testis proteins UV-crosslinked to ^{32}P-labeled RNA and resolved on a 10% SDS-PAGE gel. The complexes were visualized by autoradiography. Lanes 1 and 2 are the 48/50 kDa binding activity. Lanes 3 and 4 were immunoprecipitated with antibody to FRGY2 (Tafuri and Wolffe, 1992), 5 and 6 with antibody to MSY4, and 7 and 8 with preimmune sera. The *Prm1* $_{1\text{-}37\,(wt)}$ was used as the labeled probe in lanes 1,3,5, and 7. The *Prm1* $_{1\text{-}37\,(mut)}$ was used in lanes 2,4,6 and 8. (*D*) Supershift analysis of the 48/50 kDa binding activity bound to ^{32}P-labeled RNA. The arrow points to the EMS and the bracket indicates the supershift. MSY4 or MSY2 (Tafuri and Wolffe, 1992) antibody was added after the RNA binding reaction, the complexes were resolved on a

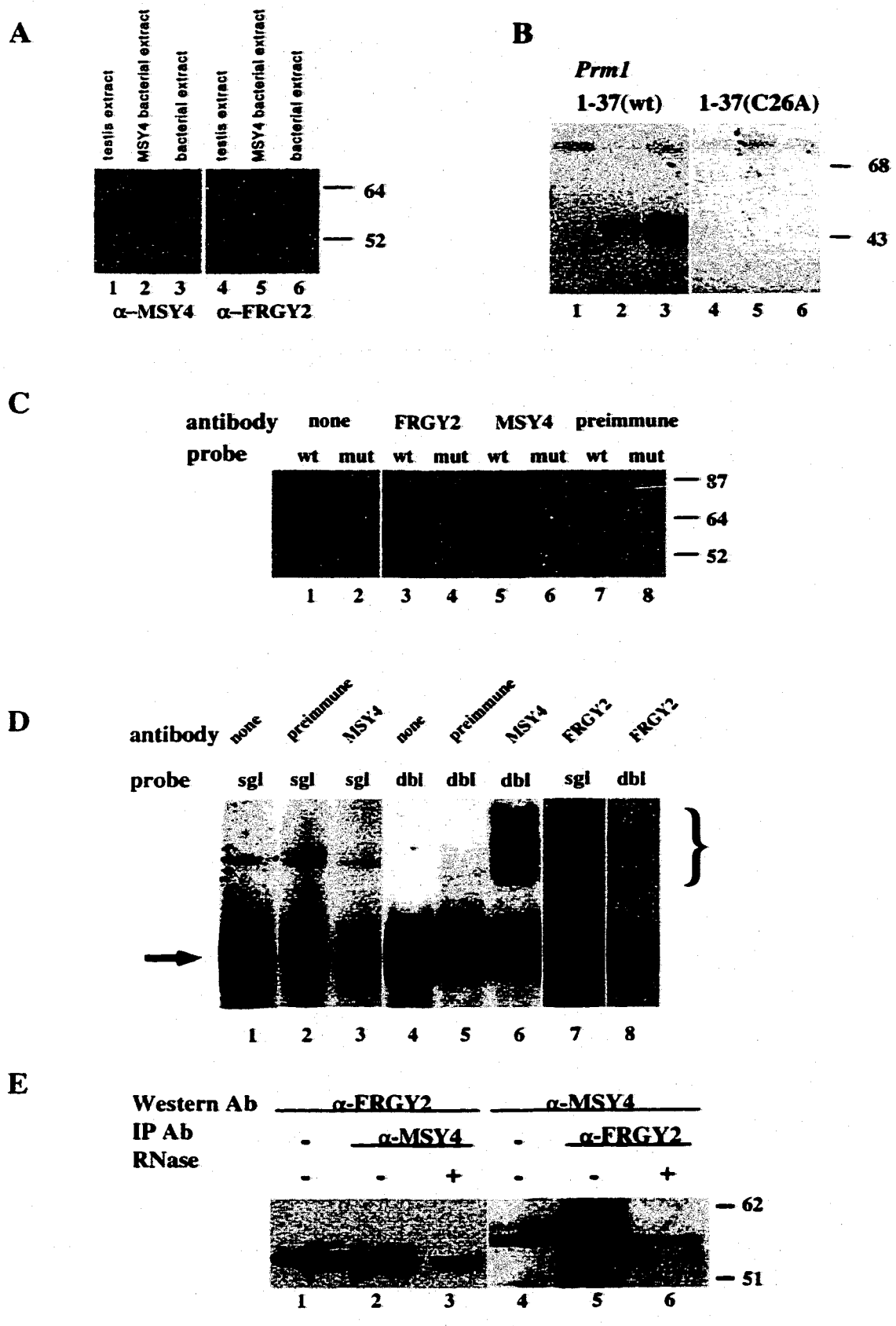

A
testis extract
MSY4 bacterial extract
bacterial extract
testis extract
MSY4 bacterial extract
bacterial extract
64
52
1 2 3 4 5 6
α–MSY4 α–FRGY2
B
Prm1
1-37(wt)
1-37(C26A)
68
43
1 2 3 4 5 6
C
antibody none FRGY2 MSY4 preimmune
probe wt mut wt mut wt mut wt mut
87
64
52
1 2 3 4 5 6 7 8
D
antibody none preimmune MSY4 none preimmune MSY4 FRGY2 FRGY2
probe sgl sgl sgl dbl dbl dbl sgl dbl
1 2 3 4 5 6 7 8
E
Western Ab α-FRGY2 α-MSY4
IP Ab - α-MSY4 - α-FRGY2
RNase - - + - - +
62
51
1 2 3 4 5 6

Figure 2-8 Distribution of *Prm1* mRNA and MSY4 in sucrose gradients. Testis extract from an adult mouse was sedimented on a 15-50% (w/w) sucrose gradient and 12 fractions were collected. Sample number 1 is the top of the gradient. Each fraction was divided for northern and western analysis. (*A*) Northern analysis on total RNA isolated from each fraction and hybridized with a probe specific to *Prm1*. The first lane is testis RNA from a wild-type mouse. The numbered lanes 1-12 refer to the collected fractions. (*B*) Western analysis on protein in each fraction, probed with MSY4 antibody. As a control for polysome association, the analysis was also performed on lysate prepared and sedimented in the presence of EDTA, which dissociates polysomes.

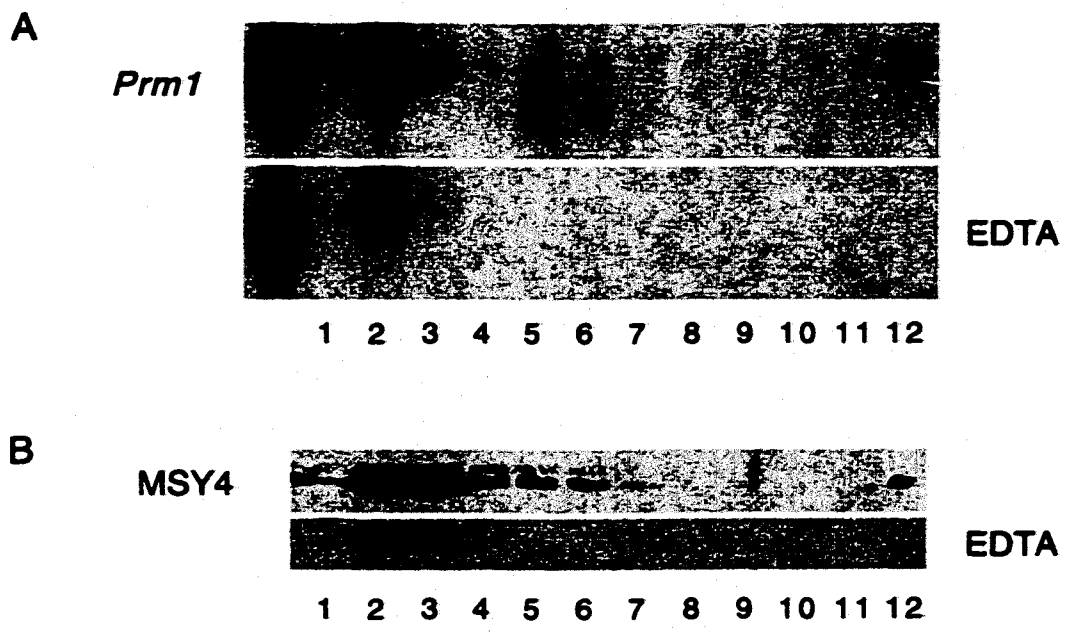
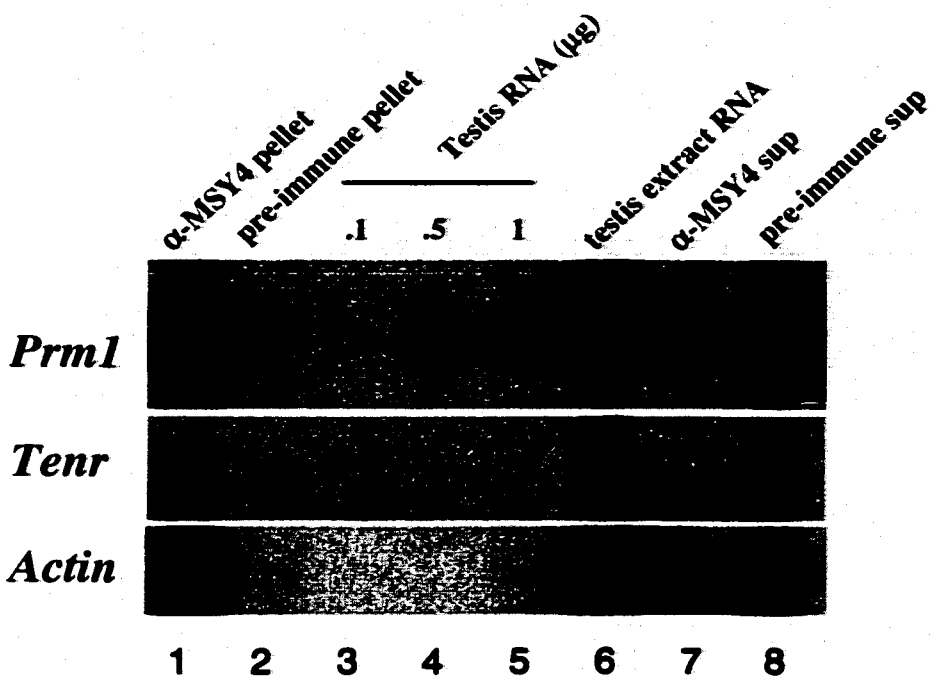
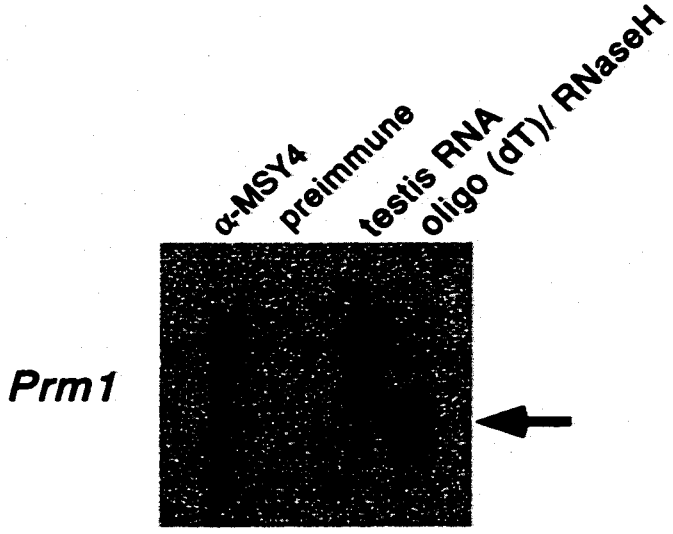

A
Prm1
EDTA
1 2 3 4 5 6 7 8 9 10 11 12
B
MSY4
EDTA
1 2 3 4 5 6 7 8 9 10 11 12

Figure 2-9 Co-immunoprecipitation of RNA with antibody to MSY4. *(A)*Testis extract from adult mice was incubated with MSY4 antibody or preimmune sera. RNA was isolated from the immunoprecipitate and the supernatant, resolved on a 2% agarose-formaldehyde gel, transferred to a nylon membrane and probed with ^{32}P-labeled *Prm1* cDNA (upper panel), stripped and then reprobed with ^{32}P-labeled *Tenr* cDNA (middle panel) and *Actin* cDNA (lower panel). RNA isolated from testis extract and dilutions of total testis RNA were used as a positive control. *(B)* RNase H treatment of the co-immunoprecipitated RNA. The analysis was as in *(A)*, with the following addition. RNA was isolated from testis extract and treated with oligo(dT) and RNase H to remove poly(A) tails. The arrow points to the deadenylated form of *Prm1*.

A


α-MSY4 pellet
pre-immune pellet
Testis RNA (μg)
.1
.5
1
testis extract RNA
α-MSY4 sup
pre-immune sup
Prm1
Tenr
Actin
1
2
3
4
5
6
7
8

B


α-MSY4
preimmune
testis RNA
oligo (dT)/ RNaseH
Prm1

DISCUSSION

I have used the yeast three-hybrid assay to clone a cDNA encoding a new member of the Y box family of RNA binding proteins. MSY4 binds to a conserved 7 nt sequence in the 3' UTR of the *Prm1* 3' UTR. This is the first report of a Y box protein binding to a specific site within an mRNA through a sequence-specific interaction. Using an antibody raised against a unique peptide from its amino terminus, we showed that MSY4 interacts with *Prm1* mRNA *in vivo* and that it constitutes part of a 48/50 kDa sequence-specific RNA binding activity present in murine testis extracts. The localization of MSY4 to the cytoplasm of round and elongating spermatids, the association of MSY4 with the fully-adenylated form of the *Prm1* mRNA in ribonucleoprotein particles, and the ability of the binding site for MSY4 to inhibit translation of a transgenic mRNA *in vivo* (Fajardo *et al.*, 1997), strongly implies that MSY4 interacts with *Prm1 in vivo*. The localization of MSY4 in growing oocytes suggests the possibility that it also functions in the control of maternal mRNA storage.

The Y box family of nucleic acid binding proteins was first identified as transcriptional activators that bind the Y box DNA sequence (CTGATTGGC/TC/TAA) (Deschamps *et al.*, 1992; Sakura *et al.*, 1988). Y box proteins were later shown to also bind RNA nonspecifically (Murray, 1994; Tafuri and Wolffe, 1993). The *Xenopus* Y box protein FRGY2 is the major component of mRNPs in oocytes (Darnbrough and Ford, 1981) and at high concentrations is able to repress translation *in vitro* (Bouvet *et al.*, 1995; Ranjan *et al.*, 1993; Richter and Smith, 1984). All Y box proteins contain the

highly conserved cold shock domain (CSD), which is 43% identical from bacteria to humans. The structure of the CSD in bacterial proteins has been solved (Schindelin *et al.*, 1994; Schindelin *et al.*, 1993) and consists of a five-stranded β-barrel. The RNA binding motifs, RNP1 and RNP2, are located on β2 and β3, respectively. The CSD has a different fold than the RNA recognition motif (RRM), within which the RNP1 and RNP2 motifs are usually found (Burd and Dreyfuss, 1994). RNP1 and RNP2 are thought to contact RNA as an open platform in a sequence nonspecific manner with RNA specificity imparted by other sequences within the RRM (Burd and Dreyfuss, 1994). Within vertebrate family members the 90 amino acid CSD domain is highly conserved, with any two members having greater than 95% amino acid identity. Vertebrate family members also contain a variable amino-terminus and a carboxyl-terminus consisting of four alternating basic and acidic regions. The carboxyl-terminus interacts with RNA nonspecifically (Bouvet *et al.*, 1995), mediates protein-protein interaction (Tafuri and Wolffe, 1992), and is necessary for the incorporation of FRGY2 into mRNPs *in vivo* (Matsumoto *et al.*, 1996).

MSY4 is most similar to human DBPA (92% at the amino acid level and 85% identical at the nucleotide level). *Msy4* is expressed in skeletal muscle as *DbpA* is, but unlike *DbpA*, *Msy4* is not highly expressed in heart (Kudo *et al.*, 1995). MSY4 shares 55% and 47% amino acid identity with mouse MSY1 and MSY2, respectively. The different murine Y box proteins contain unique amino-termini. They have their greatest

conservation within the CSD, and share a common structure within their carboxyl-termini consisting of four alternating basic and acidic regions.

Bouvet et al. (Bouvet *et al.*, 1995) used an *in vitro* selection amplification technique, SELEX (Tsai *et al.*, 1991), to show that FRGY1 and FRGY2 prefer to bind the 6 nt sequence 5' ACCAUC 3', however, specific *Xenopus* mRNAs containing this sequence have not been identified. While both the amino-terminus and carboxyl-terminus were found to contribute to the interaction with RNA, specific RNA binding was mapped to the RNP1 motif within the CSD. The similarity of the MSY4 binding site to the SELEX-defined binding site of the *Xenopus* FRGY1 and FRGY2 Y box proteins led us to determine if the murine ortholog of FRGY2, MSY2 (Gu *et al.*, 1998), is also a part of the testis 48/50 kDa binding activity. Using an antibody specific to MSY2, we showed that it is. Previous reports had shown that MSY2 binds RNA nonspecifically (Kwon *et al.*, 1993). However, the three-hybrid data shown here, and an extensive mutational analysis of the binding site *in vitro* and in the three-hybrid assay (Giorgini, *et al.*, in preparation), strongly support the conclusion that MSY2 and MSY4 bind a specific 7 nt RNA sequence, 5' UCCAUCA 3', perhaps through the RNP1 and RNP2 motifs in its CSD. Furthermore, bacterially expressed MSY4 is able to bind the $Prm1_{1\text{-}37}$ in a sequence-specific manner *in vitro* (Figure 2-7B). We refer to the binding site within the *Prm1* 3' UTR as the MSY2/MSY4 Y-box recognition sequence (YRS). The similarity of the FRGY1/FRGY2 and MSY2/MSY4 binding sites, and their cold shock domains, suggest the possibility that all Y box proteins bind a similar RNA sequence.

Although MSY4 binds the YRS specifically *in vitro* and in the three-hybrid assay, I did not detect association of MSY4 to specific mRNAs in testis extracts by immunoprecipitation. In addition to probes for specific mRNAs, I probed for poly(A)-containing RNA and determined by phosphoimaging that 75% of polyadenylated RNA was co-precipitated with anti-MSY4, compared to 25% with preimmune sera (data not shown). Some of the co-immunoprecipitation may be caused by association of protein with RNA during the preparation of the extracts and may not represent physiological interactions. However, the association of MSY4 with non-YRS-containing RNAs may also result from the nonspecific RNA binding properties of the Y box proteins *in vivo*. In addition to numerous translationally regulated messages, a surprisingly large fraction of mRNAs that are not under translational control are found in mRNPs in spermatogenic cells (Kleene, 1996; Schmidt *et al.*, 1999). The large percentage of mRNA in mRNPs may reflect constraints on the translational machinery in spermatogenic cells.

The translational repression properties of Y box proteins may be regulated by phosphorylation. Inhibitors of casein kinase II (CKII) are able to block translational repression (Braddock *et al.*, 1994) and CKII has been found in mRNPs (Cummings and Sommerville, 1988; Murray *et al.*, 1991; Yurkova and Murray, 1997). MSY4 has four potential CKII phosphorylation sites. Two forms of MSY4 are detected with our anti-MSY4 antibody suggesting the possibility that MSY4 is phosphorylated by CKII.

Y box proteins activate transcription from promoters that contain Y box sequences (Bienz, 1986). The presence of Y box elements in the promoter regions of

Prm1 (Zambrowicz *et al.*, 1993) and *Prm2* (Nikolajczyk *et al.*, 1995) suggest the possibility that MSY2 and MSY4 may also function in the transcriptional regulation of the protamine genes. MSY2 has been shown to bind to the *Prm2* Y box DNA sequence *in vitro* (Nikolajczyk *et al.*, 1995). The archetype of proteins which function in regulating transcription and translation is *Xenopus* TFIIIA (Pelham and Brown, 1980). TFIIIA stimulates transcription of the 5S RNA gene in growing oocytes and associates with the nascent 5S RNA to repress its translation in mature oocytes (Guddat *et al.*, 1990). TFIIIA interacts with DNA and RNA via separate zinc finger domains (Theunissen *et al.*, 1992). The function of TFIIIA, and perhaps some Y box proteins, as transcriptional activators, may enhance their association with particular mRNAs, via target sequences in the RNA, *in vivo*.

The MSY2/MSY4 YRS is contained within the first 37 nt of the *Prm1* 3' UTR. Transgenic mRNAs containing this sequence are translationally repressed in murine spermatids (Fajardo *et al.*, 1997). How MSY2 and MSY4 contribute to translational repression is unknown. The ability of Y box proteins to bind specific RNAs through their CSD, and nonspecifically via their basic and acidic repeats in the carboxyl-terminus, suggests the possibility that multiple MSY2/MSY4 molecules are bound to a single RNA. The co-precipitation of MSY2 and MSY4 from testis extracts supports this possibility. Binding to the YRS may nucleate the formation of a large RNA protein particle that masks the mRNA from the translational machinery.

CHAPTER 3: CHARACTERIZATION OF THE Y BOX RECOGNITION SEQUENCE

Flaviano Giorgini, Holly G. Davies, Kathy Kafer, and Robert E. Braun. The murine Y box proteins MSY2 and MSY4 bind a conserved sequence in the 3' untranslated region of protamine 1 mRNA (in preparation).

MSY2 and MSY4 belong to the Y box family of nucleic acid binding proteins, which were originally isolated by their ability to bind the DNA Y box sequence and activate transcription (Sakura *et al.*, 1988). Based on their presence in *Xenopus* oocyte mRNPs and their ability to bind RNA nonspecifically *in vitro*, they were thought to bind RNA *in vivo* exclusively in a nonspecific manner. FRGY1 and FRGY2 were found to bind a specific RNA sequence, 5' AACAUC 3', *in vitro* (Bouvet *et al.*, 1995) and specifically repress translation (Matsumoto *et al.*, 1996).

MSY2 and MSY4 are components of a 48/50 kDa RNA-binding activity in testis extracts (Davies *et al.*, 2000). This binding activity binds to a region in the *Prm1* 3' UTR that is sufficient to confer *Prm1*-like translational repression on a reporter gene in transgenic mice (Fajardo *et al.*, 1997). Like the endogenous protamines, the transgene is transcribed in round and early elongating spermatids but the mRNA is not translated until 2-8 days later in elongated spermatids.

MSY2 is the mouse ortholog of *Xenopus* FRGY2 and was cloned from an expression library screen with an antibody to FRGY2 (Gu *et al.*, 1998). MSY2 has only

been reported to bind RNA nonspecifically (Murray, 1994), although based on its homology to FRGY2, it could be expected to bind the same specific RNA sequence. *Msy4* was cloned from a mouse testis cDNA library using the yeast three-hybrid system with the first 37 nt of the *Prm1* 3' UTR as bait (Davies *et al.*, 2000). Both MSY2 and MSY4 are cytoplasmic proteins with similar expression profiles in mouse spermatogenesis, from mid-pachytene spermatocytes to early elongated spermatids (Davies *et al.*, 2000; Oko *et al.*, 1996; and H.D. unpublished data).

Previous work by other investigators has suggested that Y box proteins associate with translationally repressed messages and bind RNA nonspecifically. We have determined that MSY2 and MSY4 bind a highly conserved sequence present in the 3' UTR of *Prm1* mRNA.

RESULTS

DEFINITION OF THE MSY2/MSY4 BINDING SITE

The 48/50 kDa binding activity in testis extracts specifically binds to a region between nt 16 and 37 of the *Prm1* 3' UTR and between nt 85 and 104 of the *Prm2* 3'UTR (Fajardo *et al.*, 1994). There is a region of homology between these two sites, 5' CNANUCCAU 3'. Sequence comparison among several species revealed another two conserved nt on the 3' end (5' CNANUCCAUCA 3'). We have previously shown that an

8 nt mutation within this sequence eliminates the 48/50 kDa binding activity by EMSA and that the 48/50 kDa binding activity contains MSY2 and MSY4 (Davies *et al.*, 2000).

Flaviano Giorgini made point mutations within the conserved region to delineate the specific nt that are necessary for binding. He tested each point mutation by EMSA to see if it could be bound by MSY2/MSY4 (Figure 3-1). The *Prm1* $_{1\text{-}37}$ point mutations C17A, A19C, U21A, and C22A were bound by the complex, while U21G, C23A, A24C, U25G, A27G, C22G, and C26A all disrupted complex formation. This series of mutations defines a 7 nt binding site, 5' UCCAUCA 3', that we have termed the Y box recognition sequence (YRS). An RNA that contains the YRS in the context of the human growth hormone (*hGH*) 3' UTR was also tested (*hGH*-YRS). The *hGH* 3'UTR is not bound by MSY2/MSY4 (Fajardo *et al.*, 1994), but inclusion of the YRS was sufficient to permit binding.

COMPETITION OF WILD-TYPE RNA WITH MUTANT RNAS

Flaviano Giorgini carried out competition experiments to quantify the relative affinity of MSY2/MSY4 for the various RNAs compared to the *Prm1* $_{1\text{-}37(wt)}$. Cold RNA was preincubated with testis extract to allow binding of the RNA by MSY2/MSY4, followed by addition of ^{32}P-labeled *Prm1* $_{1\text{-}37(wt)}$ RNA and the binding reactions were subject to EMSA analysis. The competition experiments were performed with 0-, 25-, 50-, 100-, 300-, or 500-fold more cold competitor RNA to ^{32}P-labeled *Prm1* $_{1\text{-}37(wt)}$ RNA. All reactions were done in duplicate and the intensity of the EMSA complex was measured by phosphoimaging (example in Figure 3-2A). The duplicates were averaged and the

percentage competition calculated from comparison to no cold competitor. These percent competitions data were plotted as competition curves (Figure 3-2B). In general, *Prm1* $_{1\text{-}37}$ mutations that did not disrupt EMSA complex formation, such as C17A and A19C, were competent to compete wild-type RNA. Mutant RNAs that did disrupt EMSA complex formation did not effectively compete.

ANALYSIS OF MSY4 RNA-BINDING USING THE YEAST THREE-HYBRID SYSTEM

The RNA binding profile of MSY4 was also tested using the yeast three-hybrid system. Most of the mutant *Prm1* 3' UTR RNAs analyzed by EMSAs were tested for interactions with MSY4 using the yeast three-hybrid system. The three-hybrid system is a variation of the two-hybrid system, in which a hybrid RNA is used to link the two-hybrid proteins. If the protein of interest, fused to the GAL4 activation domain, can bind the RNA of interest, the reporter genes *HIS3* and *lacZ* are transcriptionally activated.

MSY4 interacts strongly with *Prm1* $_{1\text{-}37(wt)}$ and thus grows well on media lacking histidine and produces high levels of β-galactosidase (Figure 3-3A). The point mutations U21A, U21G, and C22A, as well as the *hGH*-YRS RNA, also interacted strongly with MSY4, allowing growth of the yeast on media lacking histidine and strong activation of the *lacZ* reporter gene. The point mutation A19C interacted with MSY4 in the three-hybrid assay, although weakly. In contrast, the RNA point mutations C22G, C23A, A24C, U25G, and C26A as well as the *Prm1* $_{1\text{-}37\,(mut)}$ RNA were not bound by MSY4.

Quantitative liquid β-galactosidase assays were performed to determine the relative affinity of MSY4 for various *Prm1* $_{1\text{-}37}$ mutant RNAs (Figure 3-3B). The *Prm1* $_{1\text{-}}$

$_{37\,(wt)}$ β-galactosidase activity was normalized to 1.0 unit. The liquid β-gal assays confirmed the filter assays, though quantitative differences between the mutants were detected.

Different domains of MSY4 were used in liquid β-galactosidase assays to determine which part of the protein was responsible for the specific RNA binding (Figure 3-5). Three *Prm1* $_{1\text{-}37}$ hybrid RNAs were used: wild-type, C22A, C22G, and the *hGH*-YRS RNA. MSY4ΔN, which is encoded by the original clone isolated in the three-hybrid screen (Davies *et al.*, 2000), full- length MSY4, MSY4ΔC, and the CSD were used. MSY4ΔN contains the CSD and the carboxyl-terminus tail, which contains four alternating regions of basic and acidic amino acids. MSY4ΔC contains the amino-terminus and the CSD, but is missing the carboxyl-terminus tail. All exhibited specific RNA binding to *Prm1* $_{1\text{-}37(wt)}$, C22A and *hGH*-YRS, but did not bind to C22G. Both MSY4ΔN and MSY4ΔC demonstrated specific RNA binding. These two domains share the CSD, which suggests the CSD is sufficient for the specific RNA binding. As predicted, the CSD alone specifically bound to *Prm1* $_{1\text{-}37(wt)}$, C22A and *hGH*-YRS, but did not bind to C22G. These results suggest the amino- and carboxyl-termini contribute to the general strength of the RNA binding, while the CSD alone is sufficient for specific RNA binding.

MSY4 is found with MSY2 in a specific-RNA binding complex in the mouse testis (Davies *et al.*, 2000). There are two alternatively spliced forms of *Msy2* mRNA, which are predicted to give rise to proteins with different amino-termini (Gu *et al.*, 1998).

A form of the protein that is amino-terminally deleted was tested for sequence-specific RNA binding in the three-hybrid assay using the same set of RNAs (Figure 3-4). MSY2 was found to bind the same RNAs as MSY4. A summary of the RNA binding profiles of MSY2/MSY4 by EMSA and MSY2 and MSY4 in the three-hybrid system is given in Table 3-1.

MSY2ΔN and MSY1, the other Y box protein that is known to be present in the mouse testis, were also tested for sequence-specific binding to *Prm1* $_{1\text{-}37}$ RNAs in the three-hybrid assay using quantitative β-galactosidase assays (Figure 3-5). *Msy1* was cloned from a λgt11 library by homology to *Frgy2* cDNA (Tafuri *et al.*, 1993). Although it was cloned by its similarity to *Frgy2*, MSY1 is more similar to FRGY1. Like the other Y box proteins, MSY1 showed sequence-specific binding in this assay. MSY1 bound *Prm1* $_{1\text{-}37(wt)}$, C22A and *hGH*-YRS, but did not bind to C22G. The sequence-specific RNA binding of MSY1, MSY2 and MSY4 to the YRS in the three-hybrid assay suggest all the family members bind a similar RNA sequence.

YRS FUNCTION IN VIVO

The YRS is within a sequence that is sufficient to confer *Prm1*-like translational control on a reporter gene in transgenic mice (Fajardo *et al.*, 1997). If the YRS is necessary for translational repression, it implies MSY2/MSY4 is also necessary. To test if the YRS is necessary for the translational repression seen in these transgenic mice, I generated transgenic mice using the *Prm1-hGH* reporter cassette with a mutant YRS. The transgene encodes a chimeric reporter mRNA consisting of the *Prm1* 5' UTR,

hGH coding sequence, and a 3' UTR containing *Prm1* $_{t\text{-}37}$ with a 4 nt mutation (23CAUC26 to 23ACGA26) in the YRS fused to the 3'-most 23 nt, which contains the nuclear polyadenylation signal (Figure 3-6). Flav had demonstrated that each of the four point mutations was not bound by MSY2/MSY4 in EMSAs (Figure 3-1 and Table 3-1). I then showed that the 4 nt mutation in the *Prm1* $_{t\text{-}37}$ RNA was not bound by MSY2/MSY4 in an EMSA (Figure 3-6). Both RNA probes were intact and the mutant RNA probe did not inhibit the binding of the wild-type probe (data not shown).

Three lines of mice, 3505, 3514, and 3515, were generated and analyzed. The expression of each transgenic line was analyzed by northern blot. Line 3505 expressed the transgene at a lower level than lines 3514 and 3515, but all three lines expressed the transgene at high levels (data not shown). The developmental regulation of a transgene can be studied in the testis from a single adult mouse because spermatogenesis is ongoing in the adult testis. Germ cells at different stages of development can be identified histologically by their morphological characteristics and predictable association with cells at other stages of development. To determine if the transgene is regulated like the endogenous protamines, adult testes were analyzed by immunohistochemistry. At least two mice were analyzed from each line. Immunohistochemistry with *hGH* antibody showed expression of transgenic protein in round spermatids, starting at step 7 (Figure 3-7A). The protein is detected in the acrosome, as described earlier (Braun *et al.*, 1989b) and (Fajardo *et al.*, 1997). hGH protein is detected in the cytoplasm of step 12 elongating spermatids (Figure 3-7B) and continued to be detectable in the acrosome of elongated

spermatids (Figure 3-7C) . The early expression of *hGH* shows this transgene is not under translational control and the expression is similar to other *Prm1-hGH* transgenes that are not under translational control (Braun *et al.*, 1989b; Zhong *et al.*, submited).

In some stage XII tubules there is *hGH* signal in the cytoplasm of a few meiotic cells distributed around the tubule (data not shown). It is not clear if this signal is hGH protein, since *hGH* RNA is not detected in any germ cells before step 8 by *in situ* analysis (HD unpublished). If this is early expression of the transgene in some pachytene cells, it suggests the promoter is active early in some cells.

Figure 3-1 Mutagenesis of the MSY2/MSY4 binding site within the *Prm1* 3' UTR (*Prm1* $_{1-37}$) (A) The first 37 nt of the wild-type *Prm1* 3' UTR sequence are shown, with the defined binding site (YRS) above, and mutations in conserved nt below. (B) EMSA analysis with testis extracts showing presence or absence of MSY2/MSY4 binding *Prm1* $_{1-37}$ variants. (C) Schematic diagram depicting *Prm1* 3' UTR nt 16-27. Mutations which do not disrupt MSY2/MSY4 binding are shown below, while mutations that eliminate MSY2/MSY4 binding are shown above.

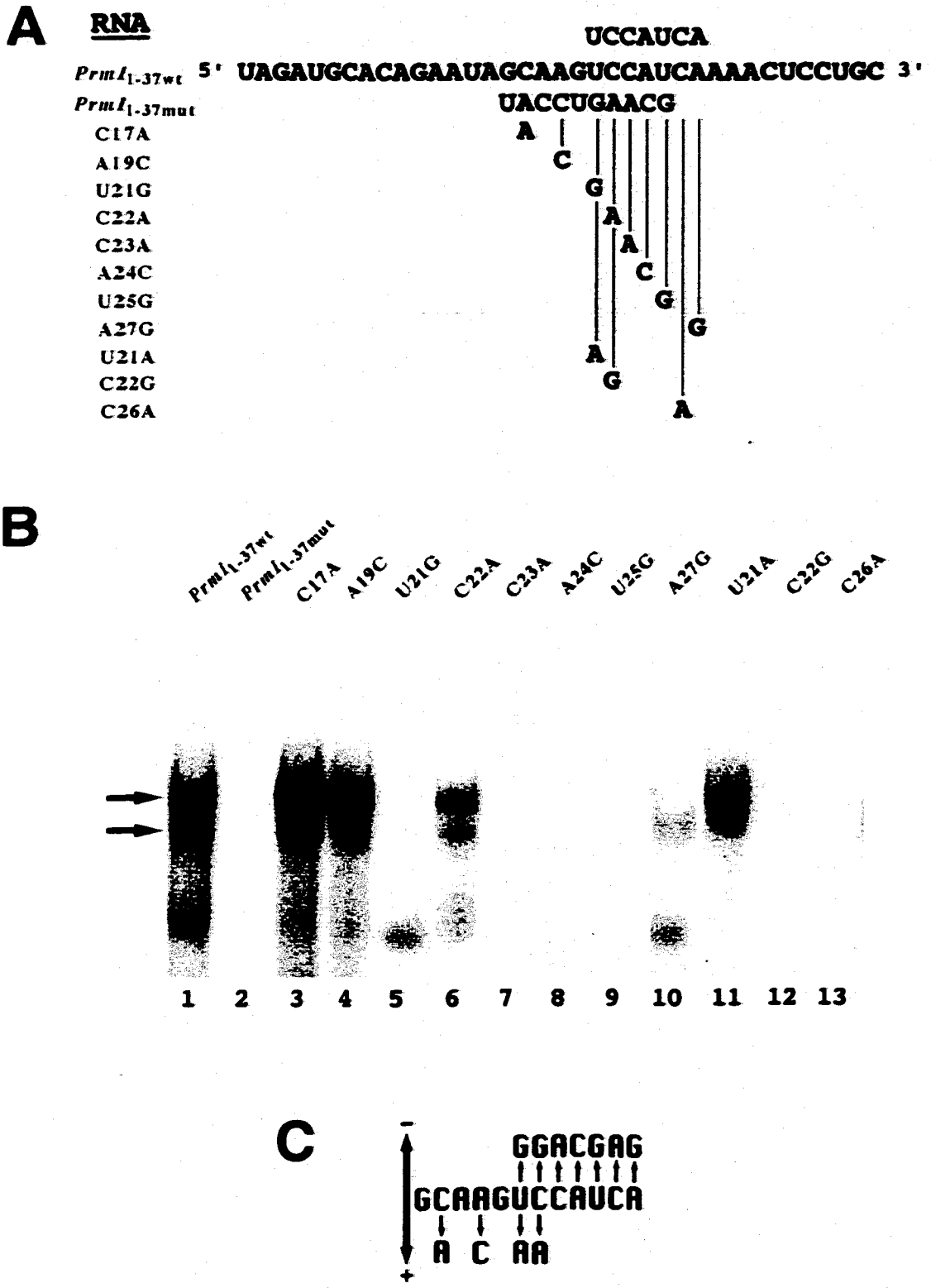
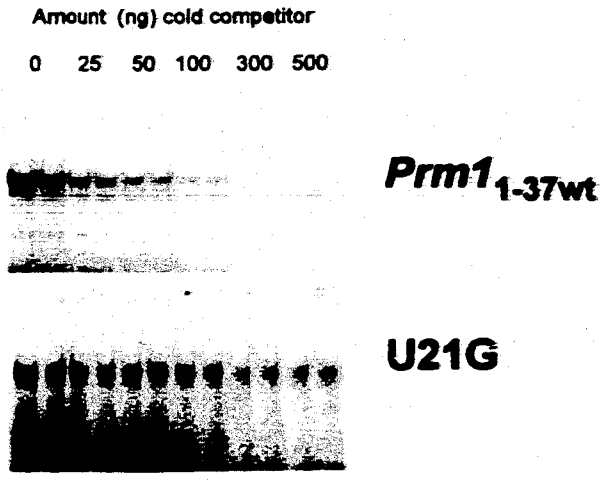

A
RNA
UCCAUCA
PrmI1-37wt 5' UAGAUGCACAGAAUAGCAAGUCCAUCAAAACUCCUGC 3'
PrmI1-37mut
UACCUGAACG
C17A
A19C
U21G
C22A
C23A
A24C
U25G
A27G
U21A
C22G
C26A
B
PrmI1-37wt
PrmI1-37mut
C17A
A19C
U21G
C22A
C23A
A24C
U25G
A27G
U21A
C22G
C26A
1 2 3 4 5 6 7 8 9 10 11 12 13
C
GGACGAG
GCAAGUCCAUCA
A C AA

Figure 3-2 Competition of *Prm1* $_{1\text{-}37(wt)}$ by *Prm1* $_{1\text{-}37}$ mutants. (A) EMSA analysis showing the ability of cold *Prm1* $_{1\text{-}37(wt)}$ to compete hot *Prm1* $_{1\text{-}37(wt)}$ binding, while U21G, which is not bound by MSY2/MSY4, does not compete *Prm1* $_{1\text{-}37(wt)}$ binding. Each quantity of cold competitor was done in duplicate. (B) Competition curves of wild-type *Prm1* $_{1\text{-}37}$ RNA by mutant *Prm1* $_{1\text{-}37}$ RNAs. Percent of wild-type EMSA complex formation is plotted vs. amount of mutant *Prm1* $_{1\text{-}37}$ RNA competitor.

A


Amount (ng) cold competitor
0 25 50 100 300 500
Prm1 1-37wt
U21G

B

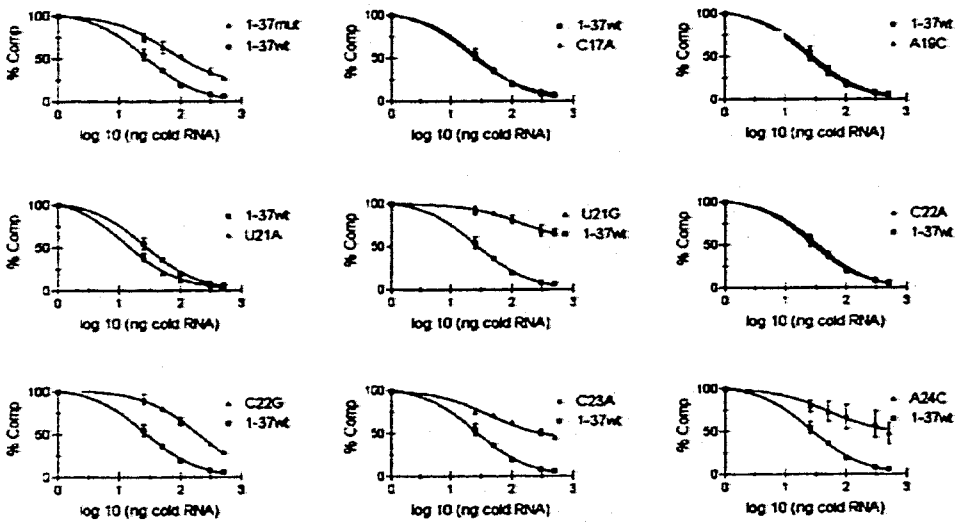
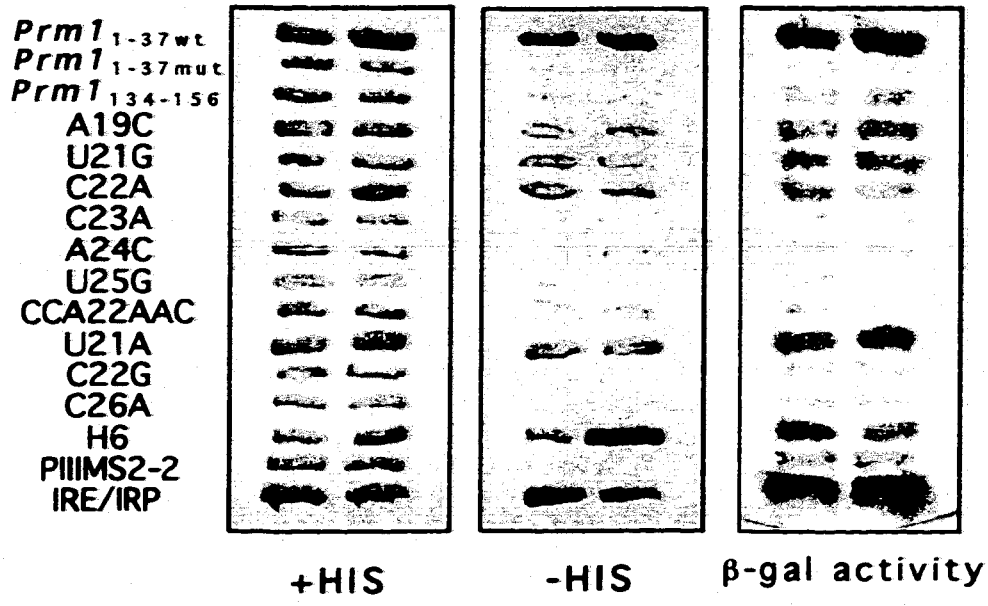

% Comp
log 10 (ng cold RNA)
1-37mut
1-37wt
C17A
A19C
U21A
C22A
C22G
C23A
A24C

Figure 3-3 Yeast three-hybrid analysis of MSY4 RNA binding. (A) The yeast strains are able to grow on media lacking leucine, tryptophan, and uracil, selecting for the three plasmids used in the three-hybrid assay (left panels). MSY4 specifically binds RNAs, as shown by growth on media also lacking histidine (with 5 mM 3AT) (center panels) and by β-galactosidase activity in a filter assay (right panels). All RNAs are as previously described. The iron response element (IRE) and its binding protein (IRP) were used as a positive control. (B) Quantitative liquid β-galactosidase assays using ONPG as the substrate. RNAs are as described previously. β-galactosidase activity is normalized to *Prm1* $_{1\text{-}37(wt)}$. The boxes represent standard error and the lines represent 95% confidence intervals. The shaded boxes are β-galactosidase activities that are not significantly different from each other.

A


Prm1 1-37wt
Prm1 1-37mut
Prm1 134-156
A19C
U21G
C22A
C23A
A24C
U25G
CCA22AAC
U21A
C22G
C26A
H6
PIIIMS2-2
IRE/IRP
+HIS
-HIS
β-gal activity

B

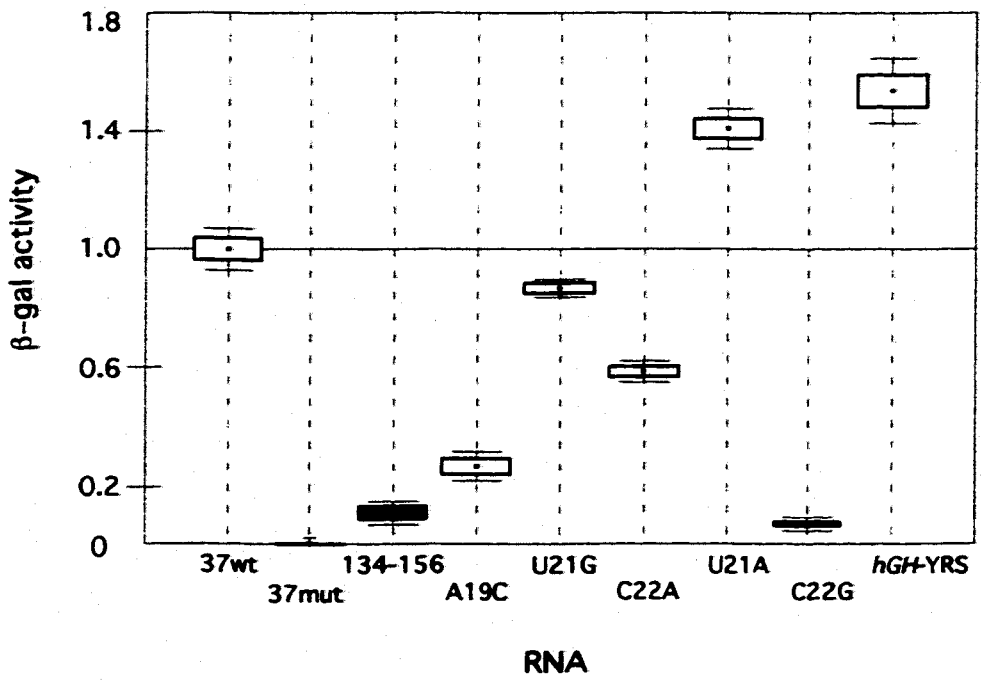

1.8
1.4
1.0
0.6
0.2
0
β-gal activity
37wt
37mut
134-156
A19C
U21G
C22A
U21A
C22G
hGH-YRS
RNA

Figure 3-4 Yeast three-hybrid analysis of MSY2 RNA binding. The yeast strains are able to grow on media lacking leucine, tryptophan, and uracil, selecting for the three plasmids used in the three-hybrid assay (left panel). MSY2 specifically binds RNAs, as shown by β-galactosidase activity in a filter assay (right panel). All RNAs are as previously described.

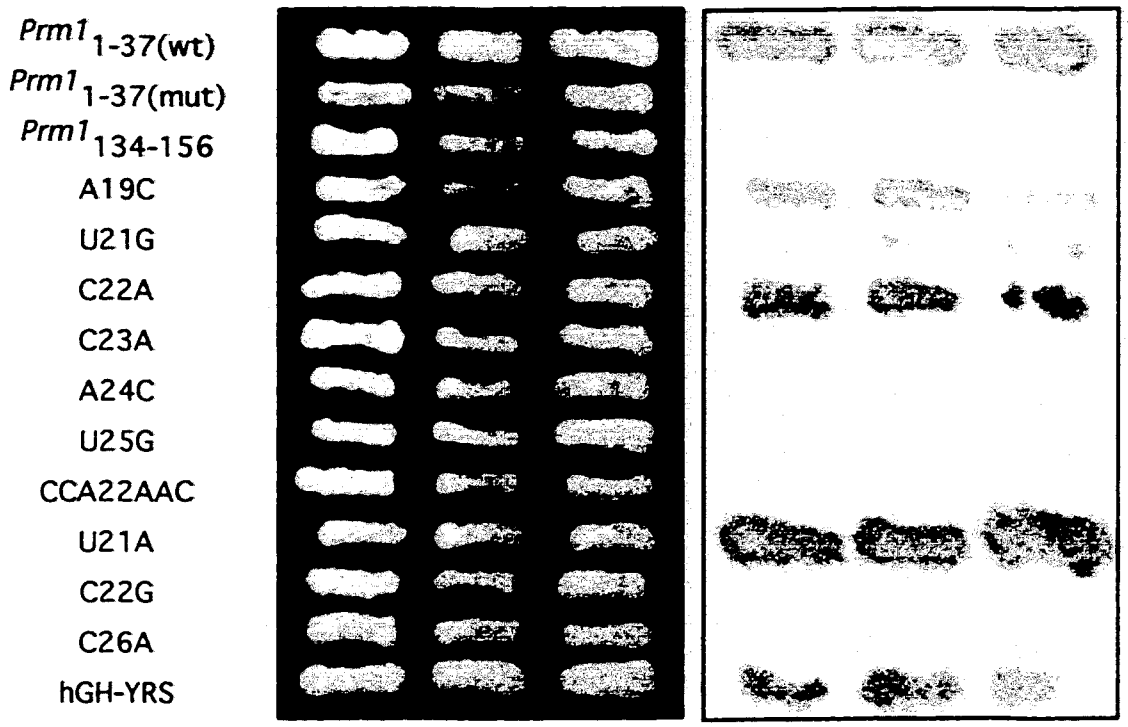

Prm1$_{1\text{-}37(wt)}$
Prm1$_{1\text{-}37(mut)}$
Prm1$_{134\text{-}156}$
A19C
U21G
C22A
C23A
A24C
U25G
CCA22AAC
U21A
C22G
C26A
hGH-YRS

Figure 3-5 Sequence-specific RNA binding of MSY4 domains, MSY2, and MSY1 in the three-hybrid assay. Quantitative liquid β-galactosidase assays using ONPG as the substrate and RNAs *Prm1* $_{1\text{-}37(wt)}$, C26A, C26G, and *hGH*-YRS as previously described. β-galactosidase activity is normalized to MSY4ΔN with *Prm1* $_{1\text{-}37(wt)}$. Bars represent standard error.

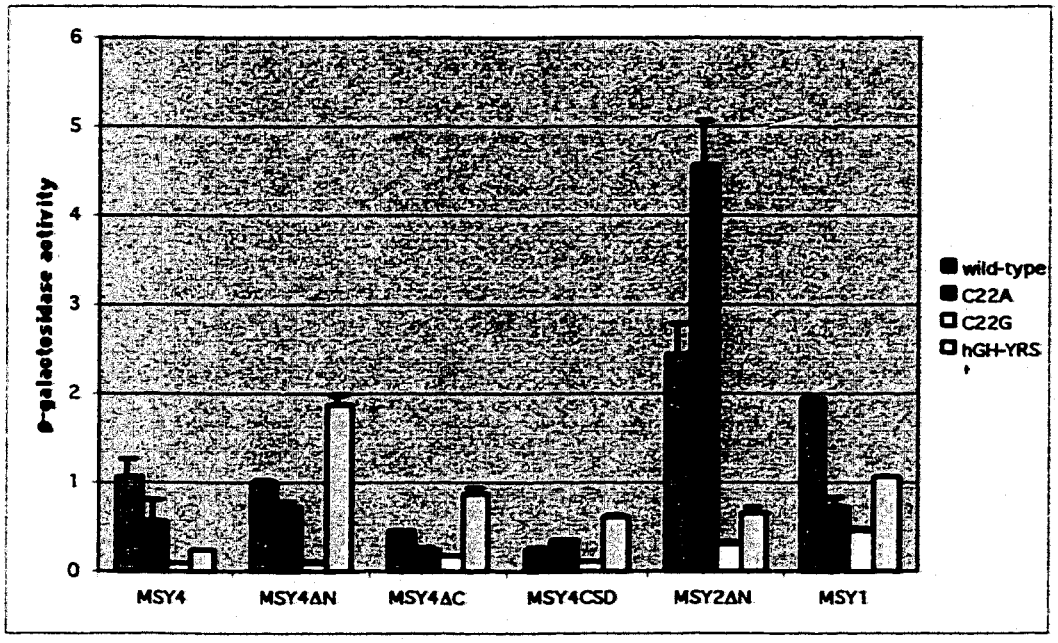
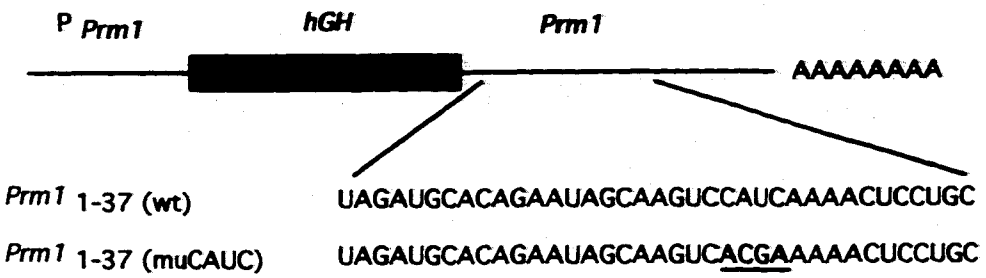

β-galactosidase activity
6
5
4
3
2
1
0
MSY4
MSY4ΔN
MSY4ΔC
MSY4CSD
MSY2ΔN
MSY1
wild-type
C22A
C22G
hGH-YRS

Figure 3-6 Structure and design of the *Prm1-hGH-Prm1* $_{1\text{-}37(mu)}$ transgene. (A) The transgene includes the *Prm1* 5' UTR, *hGH* coding region, and a deletion variant of the *Prm1* 3' UTR, which is missing nt 38-133. The wild-type version of the transgene, reported in (Fajardo *et al.*, 1997), is shown above, with the mutant version shown below. The mutant nt within the YRS are underlined. (*B*) EMSA with testis extracts showing the binding of MSY2/MSY4 to *Prm1* $_{1\text{-}37(wt)}$, but not to the mutant version used in the transgene.


P Prm1
hGH
Prm1
AAAAAAAA
Prm1 1-37 (wt)
UAGAUGCACAGAAUAGCAAGUCCAUCAAAACUCCUGC
Prm1 1-37 (muCAUC)
UAGAUGCACAGAAUAGCAAGUCACGAAAAACUCCUGC

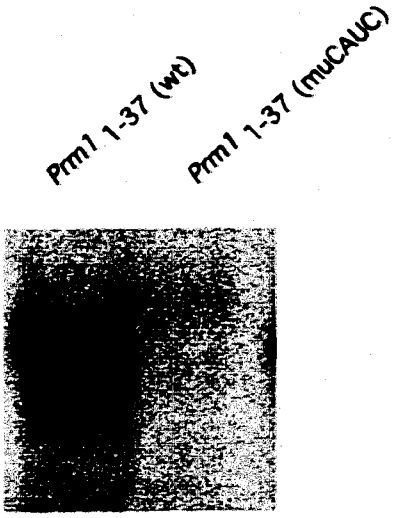

Prm1 1-37 (wt)
Prm1 1-37 (muCAUC)

Figure 3-7 Immunohistochemistry of *Prm1-hGH-Prm1* $_{1\text{-}37(mu)}$ transgene.

Immuhistochemisty detected hGH protein in (A) a stage IX tubule, (B) a stage XII tubule, and (C) a stage V tubule. Sections were counterstained with hemotoxylin. The germ cells are indicated as round spermatids (rs), elongating spermatids (egs), and elongated spermatids (eds).

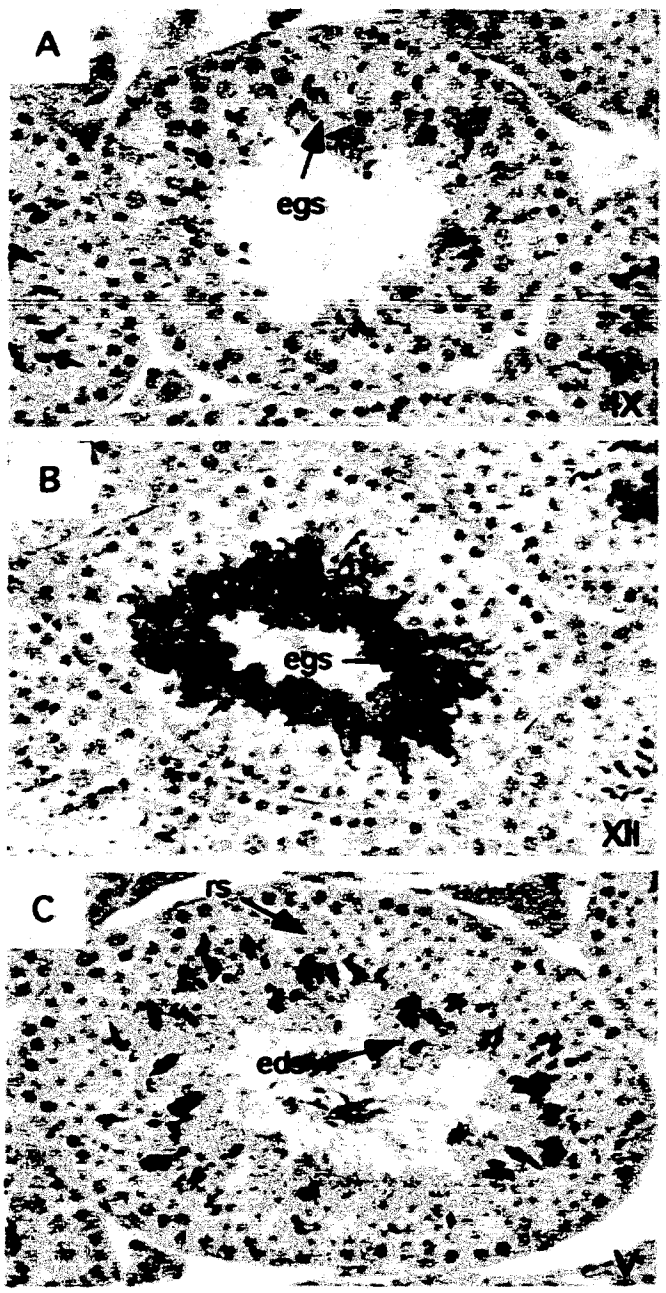

A
egs
X
B
egs
XII
C
rs
edsp
V

Table 3-1 Comparison of the *Prm1* $_{1\text{-}37}$ RNA binding by MSY2/MSY4 in testis extract using EMSAs and MSY2 and MSY4 binding in the three-hybrid assay.

	EMSA	Three Hybrid	
RNA	MSY2/MSY4	MSY4	MSY2
Prm1 $_{1\text{-}37(wt)}$	+	+	+
Prm1 $_{1\text{-}37(mut)}$	-	-	-
Prm1 $_{134\text{-}156}$	-	-	-
A19C	+	+	+
U21G	-	+	+
U21A	+	+	+
C22G	-	-	-
C22A	+	+	+
C23A	-	-	-
A24C	-	-	-
CCA22AAC	-	-	-
U25G	-	-	-
C26A	-	-	-
hGH-YRS	+	+	+

DISCUSSION

Comparative sequence analysis and mutagenesis were used to define a 7nt sequence, 5' UCCAUCA 3' (the YRS), that is present in the *Prm1* 3'UTR and binds the murine Y box proteins MSY2 and MSY4. In a novel application of the three-hybrid system, the sequence-specificity of murine Y box proteins for the YRS was confirmed. This agreement of the three-hybrid system with the EMSA shows the potential for the three-hybrid as a tool for the analysis of RNA binding. The highly conserved cold shock domain (CSD) was shown to be sufficient for the specific-RNA binding of MSY4 in the three-hybrid assay.

The YRS, 5' UCCAUCA 3', is very similar to the FRGY2 site, 5' AACAUC 3', defined *in vitro* by Selex (Bouvet *et al.*, 1995). This site has not been found in a *Xenopus* mRNA that could be a target of FRGY2. MSY2/MSY4 can bind the FRGY2 YRS in the context of the *Prm1* $_{1\text{-}37}$ (FG unpublished). The RNA binding profiles of the murine Y box proteins and the frog Y box proteins do not appear to be the same, although they are quite similar. The same changes in the mouse and frog YRS affect the binding of the proteins differently. For example, changes in the FRGY2 YRS equivalent to U21G, C23A, and A24C, which all disrupt MSY2/MSY4 binding, do not affect FRGY2 binding. On the other hand, a change equivalent to C26A in the FRGY2 YRS, which disrupts MSY2/MSY4 binding, also disrupts FRGY2 binding (Bouvet *et al.*, 1995). Mutations that mimic the FRGY2 YRS, U21A and C22A in which the first 2 nt are changed to A,

are bound by MSY2/MSY4. MSY1 was also seen to specifically bind to the YRS in the three-hybrid assay. Despite variations in binding patterns, it seems that Y box family members can bind a conserved subset of sequences.

A 47 kDa protein present in rat testis extracts has been shown to bind a conserved 8 nt site contained within the 3' UTR of rat transition protein 2 (*Tnp2*) (Schlicker *et al.*, 1997). Deletion of this site or mutation of 2 nt within this site eliminates RNA binding as assayed by EMSA. Comparison of this site to the mouse YRS reveals 6/7 conserved nt. This protein is present in the cytoplasm of pachytene spermatocytes and round spermatids. Though it is not known if this activity is a Y box protein, the similarities of the binding sites, the size of the protein, and the protein expression pattern are suggestive of a Y box protein functioning in a manner analogous to MSY2 and MSY4. This protein may be the rat ortholog of FRGY2 that can be detected in rat testis with FRGY2 antibody (Oko *et al.*, 1996).

Experiments on FRGY2 have shown that the CSD is required for sequence-specific RNA binding, while nonspecific RNA binding of the carboxyl-terminal tail is required for stable association of FRGY2 into mRNPs (Matsumoto *et al.*, 1996). We show that the MSY4 CSD is sufficient for the sequence-specific binding to *Prm1* $_{1\text{-}37}$ RNA and the tail contributes to the RNA binding. This is consistent with the similar RNA binding profiles of different Y box proteins, since the CSD is the most highly conserved domain in the protein.

Prm1 $_{1-37}$ has been shown to be sufficient to confer translational repression on a reporter mRNA in transgenic mice (Fajardo *et al.*, 1997). In this study, we show the YRS is necessary for this repression. A 4 nt mutation in the YRS, which is not bound by MSY2/MSY4, causes the transgene to be translated early, rather than being translationally delayed. Although this demonstrates the importance of the YRS *in vivo*, it does not show that MSY2/MSY4 proteins are needed. The specific binding of MSY2/MSY4 to the YRS RNA both by EMSAs and in the three-hybrid assay suggests they interact with the *Prm1-hGH-Prm1* $_{1-37}$ transgene, but not the transgene containing a mutant YRS. The *Prm1-hGH-Prm1* $_{1-37}$ transgene may not be repressed by the same mechanism as the endogenous protamines. We have seen that the YRS alone is not sufficient for translational repression in the context of the full length *Prm1* 3' UTR in which the 3' *cis* element is mutated (Zhong *et al.*, submitted). This suggests that the *Prm1-hGH-Prm1* $_{1-37}$ transgene may be repressed by an alternative mechanism, which is dependent on the YRS, and perhaps on MSY2/MSY4.

We have undertaken to establish the null phenotype of MSY4 by targeted disruption of the locus in mice and the gain of function phenotype of expressing transgenic MSY4 later in spermatogenesis in elongated spermatids. The effects of absent MSY4 or ectopic MSY4 on the *Prm1-hGH-Prm1* $_{1-37}$ transgenes may show a more direct link between the YRS and the protein *in vivo*. These experiments will also provide insight into the role of MSY4 *in vivo* and what mRNAs it interacts with.

It is clear that the original picture of Y box proteins as only specific DNA binding proteins has changed. Y box proteins may bind all mRNAs, however, the discovery of MSY2 and MSY4 as sequence-specific RNA binding proteins suggests an additional role of Y box proteins as factors important for targeting of specific mRNAs into mRNP particles.

CHAPTER 4: FUNCTION OF MSY4 *IN VIVO*

INTRODUCTION

Spermatogenesis is a complex developmental process that is ongoing in adults. Midway through mouse spermatogenesis transcription ceases, resulting in the need for genes to be transcribed earlier and regulated by translational control. This suggests that RNA binding proteins play an important role in the completion of spermatogenesis.

Protamine 1 (*Prm1*) and protamine 2 (*Prm2*) are among the genes that are under translational control in mouse spermatogenesis. These genes are transcribed in round spermatids and their mRNAs stored as repressed mRNPs in the cytoplasm for 2-8 days before being translated in elongated spermatids (Kleene *et al.*, 1984). We have described a novel member of the Y box family of nucleic acid binding proteins, *Msy4*, which was cloned by its ability to specifically bind to the *Prm1* 3' UTR in the yeast three-hybrid system (Davies *et al.*, 2000).

Y box proteins were originally described by their ability to specifically bind to the DNA Y box element in promoters (Sakura *et al.*, 1988). They also are the major components of mRNPs in *Xenopus* oocytes (Darnbrough and Ford, 1981). A role for them in spermatogenesis has been emerging, where they have been found to bind DNA in the 5' regulatory regions of genes (Nikolajczyk *et al.*, 1995) and (Schwatzbauer et al. 1998) and RNA in the 3' regulatory regions (Davies *et al.*, 2000). In addition to MSY4,

two other Y box proteins have been detected in mouse spermatogenesis. Both MSY1 and MSY2 were originally identified by their similarity to FRGY2. Y box proteins contain a variable amino-terminus, the highly conserved cold shock domain, and a tail of four alternating basic and acidic domains. The cold shock domain contains the RNA binding motifs RNP1 and RNP2, and is sufficient for specific RNA binding, while the tail contributes to RNA binding ((Matsumoto *et al.*, 1996; Giorgini *et al.*, in preparation).

MSY4 is first detected in the cytoplasm of mid-pachytene spermatocytes, becomes expressed at a higher level in late pachytene and early round spermatids, and persists until step 12 elongating spermatids (Davies *et al.*, 2000). MSY4 is a specific-RNA binding protein that is expressed in the cytoplasm of germ cells that contain repressed mRNAs, suggesting it plays an important role in the completion of spermatogenesis and may interact with repressed mRNAs during that time, including the protamines, which are actively translated when MSY4 disappears. If the disappearance of MSY4 at this time is important, then its continued expression may negatively affect spermatogenesis.

I generated transgenic mice that expressed MSY4 later in spermatogenesis. The ectopic expression of MSY4 in elongated spermatids caused dominant male sterility. These transgenic mice are an important tool for ascertaining the role of MSY4 in spermatogenesis.

RESULTS

TRANSGENE DESIGN AND EXPRESSION

A transgene was constructed that placed the *Msy4* cDNA between the *Prm1* 5' and 3' sequences to control its expression. These sequences have been used in transgenes to express hGH (Braun *et al.*, 1989b) and GFP (Schmidt *et al.*, 1999) in mouse spermatogenesis. MSY4 was tagged with the small and highly immunoreactive hemagglutinin (HA) epitope tag on its carboxyl-terminus for immunodetection. Most of the codons encoding the amino-teminus of MSY4 are deleted in this transgene. The amino-terminus is not necessary for the specific RNA-binding of MSY4 (Davies *et al.*, 2000). The peptide within the amino-terminus that is recognized by the MSY4 antibody is not present in the transgenic protein. Thus, the MSY4 antibody recognized only the endogenous MSY4 and the HA antibody recognized the transgenic MSY4HA protein. One founder male was analyzed. A line has been established from a founder female, and her son has been analyzed. Another female founder line is being established.

To demonstrate the expression of both the endogenous MSY4 and the transgenic MSY4HA, immunohistochemistry was performed on serial sections with antibodies to MSY4 and HA. The stage VI tubule shows normal expression of MSY4 (figure 4-1A) in the cytoplasm of round spermatids and expression of transgenic MSY4HA in the cytoplasm of elongated spermatids (figure 4-1B). Neighboring tubules also showed MSY4 in pachytene spermatocytes and round spermatids, while MSY4HA was detected in elongated spermatids of the same tubules. The immunohistochemistry reveals that the

transgenic MSY4HA was successfully expressed in the cytoplasm of elongated spermatids. Even though only half of the haploid spermatids contain the transgene DNA, because the mouse is hemizygous for the transgene, all the haploid spermatids express MSY4HA protein. Haploid spermatids are functionally diploid due to intercellular bridges that allow the sharing of cytoplasmic factors (Braun *et al.*, 1989a).

STERILITY

Hemizygous transgenic males are sterile. They are able to form copulation plugs in females, but no pups have ever resulted. Sperm counts show the amount of sperm to be less than 10% of wild-type levels.

In order to examine the mutant phenotype, testes from wild-type and transgenic adult males were investigated histologically (Figure 4-2). In the transgenic male, elongated spermatids were visible within the tubules. Step 16 elongated spermatids are normally released into the lumen in stage VIII tubules. In the wild-type stage IX tubule, step 9 elongating spermatids were seen and there were no elongated spermatids (Figure 4-2A and C). In contrast, the transgenic stage IX tubule clearly had both step 9 elongating spermatids and retained elongated spermatids (Figure 4-2B and D). As expected from the retained spermatids in the tubules, there were very few spermatozoa in the cauda epididymis from the transgenic male (Figure 4-2F), compared to wild-type (Figure 4-2E).

PROTAMINE TRANSLATION

MSY4 was identified by its ability to specifically bind to a conserved region of the *Prm1* 3' UTR mRNA *in vitro*, so the expression of *Prm1* and *Prm2* was examined in the transgenic mice. Sucrose gradient sedimentation was used to evaluate the polysome distrubution of *Prm1* mRNA in a transgenic male. The optical density profile shows noticeable peaks of ribosomal subunits, monosomes, disomes, trisomes, and higher molecular weight polysomes (Figure 4-3). Testis extract was fractionated on a sucrose gradient and I performed northern analysis on each fraction (Figure 4-3 lower panel). The RNA from fraction 2 was degraded, but *Prm1* was detected as a larger message in mRNP fraction 3 (lane 3) and as a mixture of this larger message and shorter messages in the polysome fractions due to deadenylation (lanes 5 and 6). Due to the shortness of *Prm1* mRNA, it is seen on disomes (lane 5) and trisomes (lane 6), but not on larger polysomes (lanes 7-12). This analysis showed that the expression of MSY4HA in elongated spermatids did not completely prevent the translation of *Prm1* mRNA.

To ascertain the expression of PRM2 in the transgenic mice, the testis was analyzed by immunohistochemistry. In the wild-type mouse, PRM2 was detected in the nucleus of elongated spermatids (Figure 4-4C). Consistent with the location of *Prm1* mRNA in the polysome gradient, which suggested the protamines were translated, PRM2 was detected in some elongated spermatids in the transgenic mouse (Figure 4-4D). The expression pattern was different from the expression in the wild-type tubules. This can be seen at a lower magnification, comparing the expression in wild-type (Figure 4-4A) and transgenic (Figure 4-4B) tubules. High expression was seen in wild-type elongated

spermatids, which was not seen in transgenic elongated spermatids. The lower magnification view contained tubules that have elongated spermatids, including the tubule seen in Figure 4-4D. PRM2 continued to be detected in the spermatoza in the caput epididymis of the wild-type mouse (Figure 4-4E). Most of the small number of spermatozoa in the caput epididymis of the transgenic mouse did not have detectable PRM2 (Figure 4-4F). The abnormal pattern of PRM2 expression could be a direct or indirect effect of MSY4HA expression.

The location of MSY4 and MSY4HA within the polysome gradient was also inspected by western analysis (Figure 4-3 upper panels). Endogenous MSY4 is mostly seen in the mRNP fractions, as previously reported (Davies *et al.*, 2000). The transgenic MSY4HA was also seen predominantly in mRNP fractions 2 and 3. This suggests both the endogenous and transgenic MSY4 protein is mostly complexed in mRNPs with repressed mRNAs.

Figure 4-1 Expression of endogenous and transgenic MSY4 by immunohistochemistry. Serial sections of the same stage VI tubule from a founder male are shown. (A) MSY4 detected with MSY4 antibody. (B) MSY4HA detected with HA antibody. MSY4HA is missing the epitope in the amino-terminus that is recognized by the MSY4 antibody. Germ cells are described as pachytene spermatocytes (ps), round spermatids (rs), elongating spermatids (egs), and elongated spermatids (eds).

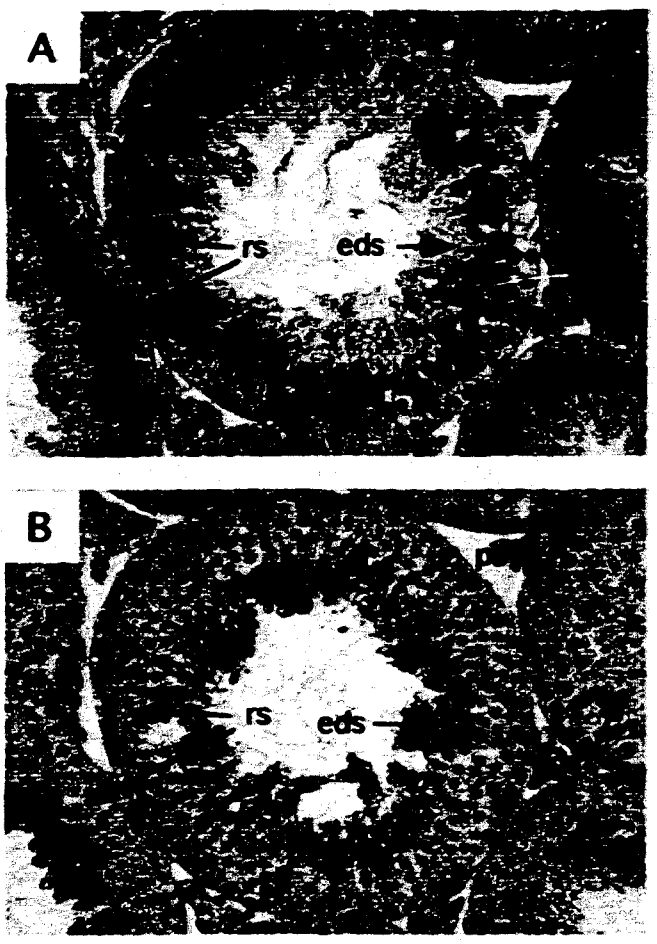

A
rs
eds
B
rs
eds

Figure 4-2 Sections from wild-type (A, C, and E) and transgenic (B, D, and F) testis (A-D) and cauda epididymis (E and F) showing the histology of each animal. (C) and (D) are enlargements of the boxes areas in (A) and (B), respectively. Spermatids are described as elongating spermatids (egs) and elongated spermatids (eds). The sections are stained with PAS and hematoxylin. Roman numerals indicate the stage of the tubule.

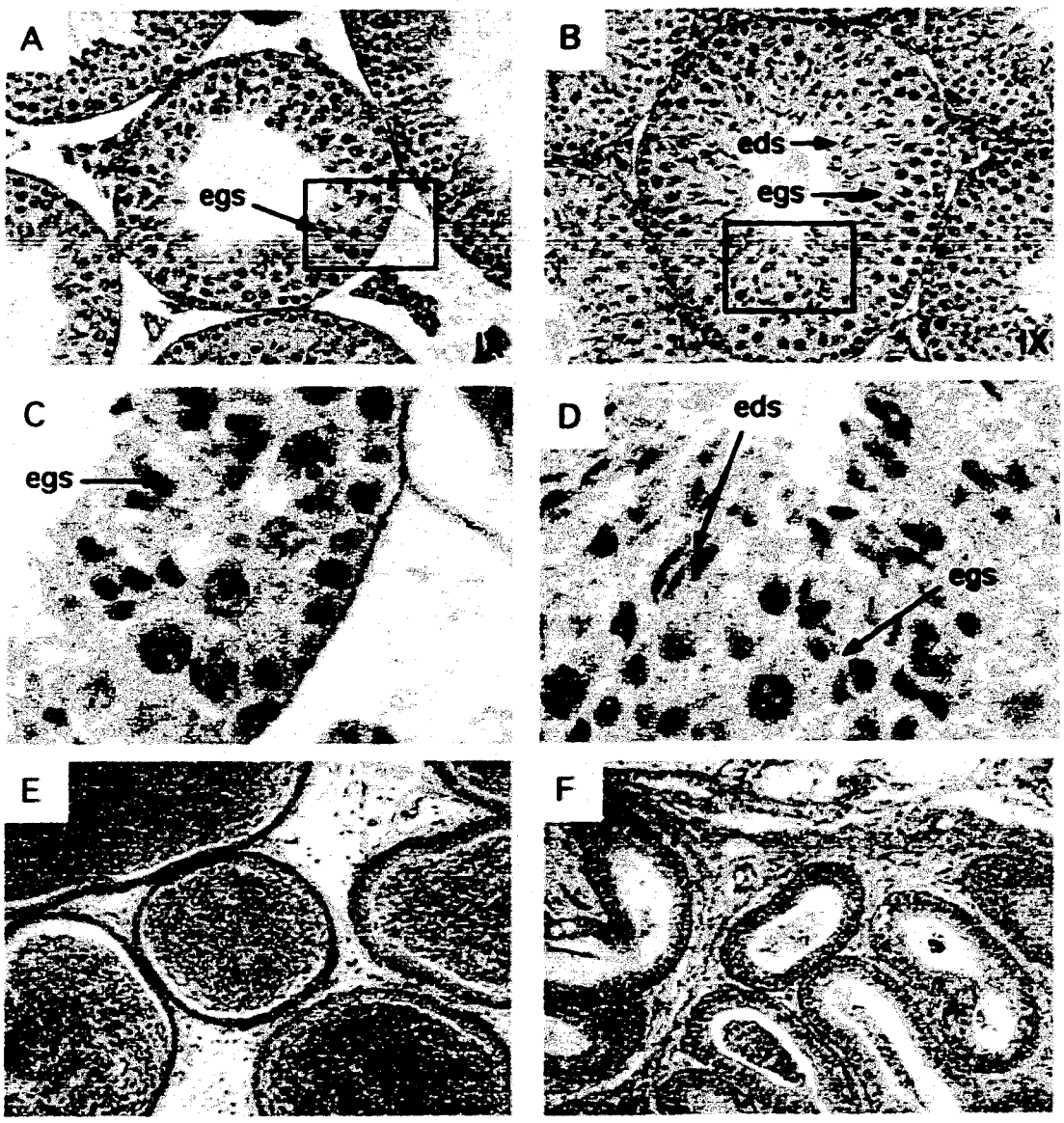

A
egs
B
eds
egs
1X
C
egs
D
eds
egs
E
F

Figure 4-3 Distribution of *Prm1* mRNA, MSY4 and MSY4HA in a sucrose gradient. Testis extract from an adult male was sedimented on a 15% sucrose (w/w) gradient, the absorbance at 254 nm was recorded, and twelve fractions were collected. Sample one is the top of the gradient. Each fraction was divided for northern or western analysis. Prominent peaks are labeled on the A_{254} graph. Single ribosome subunits are labeled 40S and 60S, monosomes as 80S/M, disomes as D, and trisomes as T. Protein from each fraction was run on a 9% SDS denaturing gel, transferred to nitrocellulose, and probed with antibody to MSY4 (upper panel), then antibody to HA (middle panel). The location of the 51 kDa standard is indicated on the right. Total RNA was isolated from each fraction and run on a 1.5% agarose-formaldehyde gel, transferred to a nylon membrane, and hybridized to a probe specific to *Prm1* (lower panel). The RNA from fraction 2 was degraded.

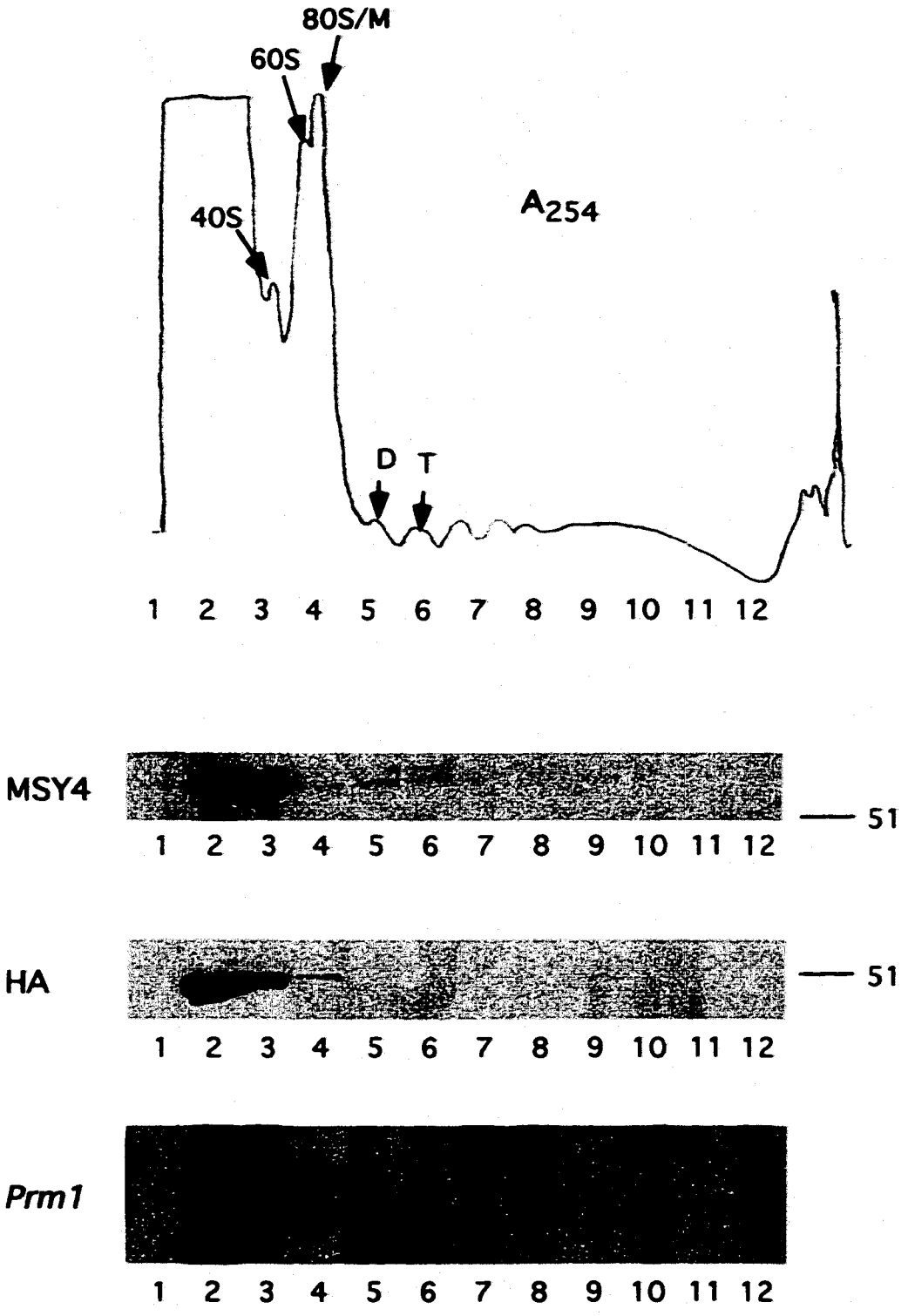
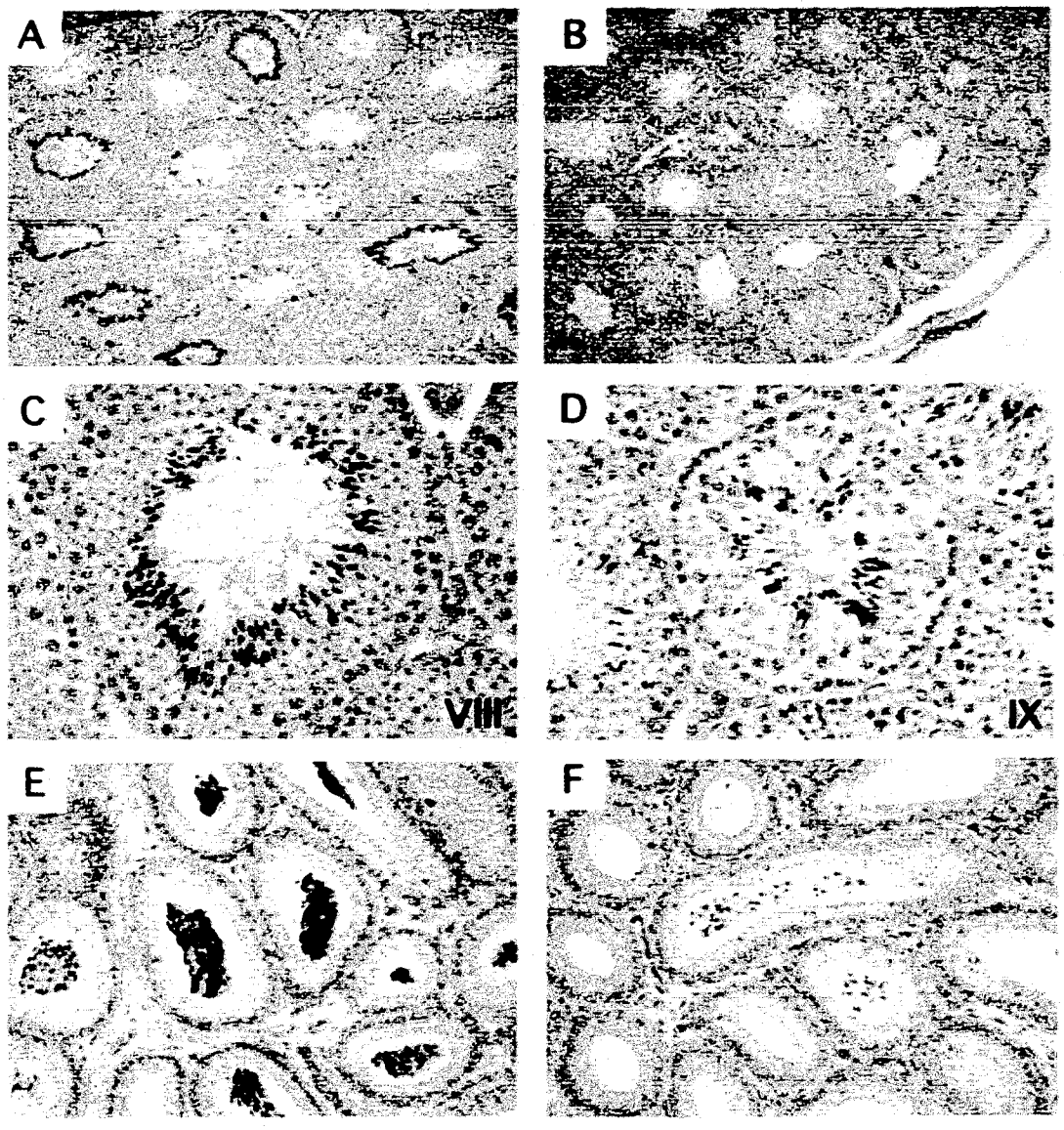

80S/M
60S
40S
A254
D
T
1 2 3 4 5 6 7 8 9 10 11 12
MSY4
51
1 2 3 4 5 6 7 8 9 10 11 12
HA
51
1 2 3 4 5 6 7 8 9 10 11 12
Prm1
1 2 3 4 5 6 7 8 9 10 11 12

Figure 4-4 Expression of *PRM2* in wild-type (A, C, E) and transgenic (B, D, F) testsis (A-D) and caput epididymis (E and F) by immunohistochemistry with an antibody to *PRM2*. Sections were counterstained with hematoxylin. Roman numerals indicate the stage of the tubule.


A
B
C
VIII
D
IX
E
F

DISCUSSION

MSY4 is normally expressed in spermatogenesis from mid-pachytene spermatocytes through elongating spermatids. The ectopic expression of MSY4 in elongated spermatids results in dominant male sterility. This demonstrates that MSY4 levels must decrease at the correct time for the successful completion of spermatogenesis. The phenotype gives us insight into the role the endogenous MSY4 plays earlier in spermatogenesis. Further experiments with MSY4 both *in vivo* and *in vitro* will identify RNA targets of MSY4 and elucidate the role of MSY4 in mouse spermatogenesis.

The obvious phenotype of the transgenic mice is the lack of spermiation and retention of elongated spermatids. The retained spermatids are most likely drawn back into the tubule and phagocytosed by the Sertoli cells (Russell *et al.*, 1990). This is not the primary defect, but an effect of abnormal spermiogenesis. The primary phenotype is most likely in early elongated spermatids where MSY4 begins to be expressed ectopically. Because of retained spermatids from the previous round, it is difficult to analyze the early elongated spermatids in adults. Further analysis of prepubertal males, that have not completed the first round of spermatogenesis, will be valuable. Polysome analysis to compare the amount of *Prm* mRNA on polysomes in the transgenic male compared to a wild-type male of the same age will show if *Prm* is recruited to polysomes at wild-type levels in the transgenic mice. Immunohistochemistry on a testis of a prepubertal transgenic male compared to a nontransgenic male will also show if PRM is expressed

normally in early elongated spermatids. The expression of protamines can also be examined by acid urea gel electrophoresis.

Electron microscopy is being used to investigate the mutant phenotype. The resolution of electron microscopy is great enough to reveal subcellular details, such as whether the spermatid nuclei have condensed properly.

MSY4 specifically binds the YRS, a 7 nt sequence within the first 37 nt of the *Prm1* 3' UTR (Davies *et al.*, 2000). *Prm1* $_{1\text{-}37}$ is sufficient to confer *Prm1* like translational control on a reporter gene in transgenic mice (Fajardo *et al.*, 1997), and the YRS is necessary (Giorgini *et al.*, in preparation). The *Prm1-hGH-Prm1* $_{1\text{-}37(wt)}$ and *Prm1-hGH-Prm1* $_{1\text{-}37(mut)}$ transgenes will be used to make a connection between MSY4 and the YRS *in vivo*. A male from each *Prm1-hGH-Prm* $_{1\text{-}37}$ transgenic line has been crossed to a female that expresses the *Prm1-Msy4-Prm1* transgene. If MSY4 specifically interacts with the YRS *in vivo*, then the ectopic expression of MSY4 in elongated spermatids will only affect the *Prm1-hGH-Prm1* $_{1\text{-}37}$ transgene that contains a wild-type YRS and not the transgene that contains the mutant YRS.

The successful expression of MSY4HA in elongated spermatids using the *Prm1* transgenic cassette argues against MSY4 directly repressing *Prm1* mRNA through sequences in the 3' UTR. As the transgenic MSY4HA protein was produced, it would repress the translation of the transgene mRNA, which contains the YRS in its 3' UTR. It is paradoxical that the transgenic MSY4HA is translated in elongated spermatids as expected, whereas the endogenous PRM2 is translated abnormally in elongated

spermatids. Both the transgene and the endogenous protamines are translationally repressed correctly in round spermatids, but the transgene appears to be more translationally active in elongated spermatids. The obvious differences between the two mRNAs is the lack of introns in the *Prm1-Msy4-Prm1* transgene and the differences in length. The *Msy4* cDNA that was used is approximately 900 bp, compared to 300 bp for *Prm2*. In addition the sequence of the coding regions and the mRNA expression levels are different. The *Prm1* 3' UTR was not able to confer *Prm1*-like translational control on *lacZ* in transgenic mice (RB, unpublished data), demonstrating that not all open reading frames act the same in the transgenic cassette. It was later discovered that *lacZ* has a cryptic splice site. The lack of introns, size of the message, amount of mRNA, or other elements withing the coding region, may allow the *Prm1-Msy4-Prm1* transgene to more effectively compete for the translational machinery.

A subsequent transgenic experiment would be to utilize the *hGH* 3' UTR, which is not sufficient to confer *Prm1*-like translational control on a reporter gene in transgenic mice (Braun et al. 1989). We would expect the ectopic MSY4 expressed from that transgene, *Prm1-Msy4-hGH*, to be found in later round, elongating and elongated spermatids. Since the *hGH* 3' UTR does not have a YRS, the transgenic protein would also not affect its own translation.

CHAPTER 5: SUMMARY AND FUTURE DIRECTIONS

Spermatogenesis is a complex developmental process that can be studied in adult animals. It can be divided into three phases, mitotic proliferation, meiosis, and spermiogenesis. Spermiogenesis is the differentiation of haploid round spermatids into mature spermatozoa that are capable of reaching and fertilizing an egg. Transcription ceases midway through spermiogenesis, so translational control of pre-existing mRNAs is necessary for the completion of spermiogenesis. This requires RNA binding proteins to play a major role in spermatogenesis.

The protamines are among the genes that are under translational control in spermatogenesis. The two protamine genes are transcribed in round spermatids, the mRNA stored in the cytoplasm for 2-8 days, then translated in elongated spermatids. Nuclear condensation involves the replacement of the histones with the transition proteins which are then replaced by the protamines. If *Prm1* is not translationally delayed, the nucleus prematurely condenses and dominant sterility results (Lee *et al.*, 1995). Several *cis*- and trans-elements have been identified, but the mechanism of translational control is still not known.

I have focused on the first 37 nt of the *Prm1* 3' UTR, which were identified as sufficient for translational control, and the proteins that bind to this region. Using the *Prm1* $_{1\text{-}37}$ RNA in a three-hybrid screen, I cloned *Msy4*, a novel member of the Y box family of nucleic acid binding proteins. MSY4 is expressed in the testis in pachytene

spermatocytes and round and elongating spermatids and in the cytoplasm of oocytes. This expression pattern suggests MSY4 interacts with stored mRNAs. We used electrophoretic mobility shift assays (EMSAs) and the three-hybrid system to define the site within *Prm1* $_{1\text{-}37}$ bound by MSY4, which we have named the Y box recognition sequence (YRS). We also showed that a previously described binding complex in testis extract contains MSY4 and a related protein, MSY2. These proteins are clearly sequence-specific RNA binding proteins, contrary to previous results for MSY2 using other RNA sequences (Murray, 1994). The YRS is necessary for the translational repression seen in the *Prm1-hGH-Prm1* $_{1\text{-}37}$ transgene. MSY4 was shown to have RNA targets in spermatids, since ectopic expression of MSY4 in elongated spermatids causes dominant male sterility. There are still questions that remain about the function of MSY4 *in vivo*, the interaction between MSY4 and the YRS, and how *Prm1* mRNA is translationally regulated. There are several models of translational repression via *cis*-elements in the 3' UTR. One model is that the repression is mediated through the presence or absence of a poly (A) tail, which interacts with the 5' end of the mRNA. A second model is that a protein that binds the *cis*-element in the 3' UTR interacts directly or indirectly with the 5' cap or binding proteins. A third model is that the mRNA is completely bound up by proteins and masked from the translational machinery. The masking may be nucleated by a specific protein.

TRANSGENES

Several transgenic mouse experiments utilizing *hGH* as a reporter have been performed to analyze the *cis*-elements that are required for *Prm1* translational control.

These began with the seminal experiments that showed the *cis*-elements resided in the 156 nt 3' UTR. *Prm1* $_{1-37}$ and *Prm1* $_{93-156}$ were both shown to be sufficient for the translational control. Further mutational analysis of each region defined the YRS as necessary within *Prm1* $_{1-37}$ and the Z box within the *Prm1* $_{69-156}$. *Prm1-hGH-Prm1* $_{1-156}$ with a mutant Z box was not under translational control, demonstrating that the YRS alone is not sufficient. These results suggest that the Z box is the essential *cis*-element, while the YRS, and by extension the proteins that bind it, have another function. Their primary function may be in stabilizing the mRNA and protecting it from RNases by its location in an mRNP. A secondary effect may be keeping the repressed mRNA unavailable to the translational machinery. If the *Prm1-hGH-Prm1* $_{1-156(muZ)}$ transgene were partially under translational control, we wouldn't be able to detect it. We can detect a small amount of early expression from transgenes that are mostly repressed. However, if a transgene was largely translated early, a smaller proportion of transgenic mRNA being translated later rather than earlier would not be detected.

The 3' UTRs of *Prm1* and *Prm2* are very similar, but not identical. The 3' end of the *Prm1* 3' UTR has two copies of the Z box, with a complementary region in between. This region has the potential to form two different stem loop structures. However, the stem loop structures are not necessary for translational control, since a transgene that contains the region with a compound mutation that abolishes the possible stem loop structures is under translational control (Zhong *et al.*, submitted). This result is consistent with the *Prm2* 3' UTR, which doesn't contain the potential stem loop structures, but is

under translational control. Both 3' UTRs have the same region of the YRS, H, Y and Z boxes, with roughly the same spacing among the elements. This may reflect a necessity for several *cis*-elements to effect the correct expression of the protamines. Consistent with this, all the transgenes with a portion of the 3' UTR deleted that are under translational control are leaky, with some early expression of the transgene.

SPECIFIC VS. NONSPECIFIC RNA BINDING BY MSY4

MSY2 and MSY4 specifically bind the YRS-containing RNAs by EMSA and in the three-hybrid assay, which conflicts with the data from the immunoprecipitation experiments, where MSY4 was seen to interact with all mRNAs in testis extract. Association of MSY4 with RNA during the preparation of the testis extract may not represent physiological interactions. There may be indirect interactions among proteins and mRNAs and these large complexes may have been immunoprecipitated in these experiments. Thus, some of the mRNA seen to be immunoprecipitated by MSY4 antibody may not have been bound directly by MSY4. MSY4 was seen to interact with mRNAs by immunoprecipitation that it does not interact with *in vivo*, such as *in vitro* transcribed *hGH* mRNA that was added to the testis extract and immunoprecipitated with MSY4 antibody. Other mRNAs that were immunoprecipitated with MSY4 antibody are mostly found on polysomes, while MSY4 is mostly in mRNP fractions. *Actin* mRNA is mostly on polysomes in the testis (Kleene, 1996), so MSY4 shouldn't interact with *Actin* mRNA *in vivo*, but it was seen to interact by immunoprecipitation. MSY2 has been seen to bind nonspecifically *in vitro* (Murray, 1994), but those experiments were not done with

RNAs that contain the YRS target. The data suggest MSY2 and MSY4 can bind RNA nonspecifically, but much prefer to bind the YRS *in vitro* and in the three-hybrid assay. This may reflect different functions of Y box proteins in mouse spermatogenesis. They may have one role in binding all RNA and another role in binding specific RNAs. This is consistent with a model of *Prm1* translational repression in which MSY2 or MSY4 specifically binds to the YRS, then recruit additional proteins, including MSY2 and MSY4 to bind to the mRNA nonspecifically and mask it.

MSY4 FUNCTION

Flaviano Giorgini is establishing the null phenotype of MSY4 by targeted disruption of the locus in mice. The mutation is being constructed as a conditional allele, which can be specifically targeted in individual tissues. This will enable the investigation of phenotypes in individual tissues in isolation. While we will have to wait for the results, there are several possible phenotypes we would expect. *Msy4* is highly expressed in the testis, and although it is expressed in low levels in other tissues, we are concentrating on the possible testis phenotypes. Since there are several other Y box proteins in the mouse testis, there may be no phenotype if the function of MSY4 is redundant with other Y box proteins. Even if they have redundant functions, the mutant mouse lacking MSY4 will likely have less total Y box protein, which may affect the targets of the Y box proteins. One possible role for MSY4 is protection of mRNAs. If MSY4 protects mRNAs from degradation, then without MSY4 those mRNAs will be degraded and never translated. Another possible function of MSY4 is translational repression. If MSY4 is a translational

repressor, then without MSY4 its mRNA targets will be translated early. There may be effects in tissues other than the testis. Since MSY4 is also expressed in the cytoplasm of oocytes, like the *Xenopus* family members, there may be an ovarian phenotype. Again the predicted phenotype is that oocytes without MSY4 will be defective in some aspect of mRNA storage or stabilty. The transgenic mouse line that expresses MSY4 ectopically in elongated spermatids will also address the function of MSY4 *in vivo*.

TRANSCRIPTION LINK

A link between transcriptional activation of a gene in the nucleus and the translational repression of its mRNA in the cytoplasm has been seen for FRGY2 in *Xenopus* oocytes. In general, mRNA that is injected into *Xenopus* oocytes is translated, despite the high concentration of masking proteins. For example, Histone H1 mRNA synthesized *in vitro* and injected into the oocyte cytoplasm is translated, but mRNA synthesized *in vivo* is translationally competent, but masked in the cytoplasm (Bouvet and Wolffe, 1994). Similar results have been seen for other endogenous and exogenous mRNAs. Antibodies to FRGY2 and caesin kinase II, when injected into the nucleus, relieve translational repression (Braddock *et al.*, 1994) and (Sommerville and Ladomery, 1996a), further demonstrating that nuclear activities are needed. Although MSY4 is only detected in the cytoplasm, other cytoplasmic Y box proteins have been shown to be transcriptional activators. Some Y box proteins may be predominantly in the cytoplasm, but present in the nucleus at low levels or for short periods of times. FRGY2 can be

detected at low levels in the oocyte nucleus and on lampbrush chromosome loops in *Xenopus* oocytes (Sommerville and Ladomery, 1996b).

Prm1 and *Prm2* contain Y boxes in their promoters that are bound by Y box proteins in mouse testis extracts (Nikolajczyk *et al.*, 1995). This suggests a model where the Y box proteins first bind to *Prm1* DNA in the nucleus, are targeted to the nascent RNA transcript, and transported to the cytoplasm already bound. One proposed experiment to test the interaction between transcription and translation of *Prm1* is to make a *Prm1-hGH-Prm1* transgene that contains a mutant version of the Y box in the *Prm1* promoter. The effect of the Y box on the transcription of *Prm1* has not been directly tested, but transgenes without this region are transcribed, demonstrating that this region is not essential for transcription (Zambrowicz *et al.*, 1993). If the *Prm1*-like translational control is dependent on the interaction of MSY4 or other Y box proteins with the Y box in the promoter, then the mutation in the Y box would abolish the translational control.

IN VITRO STUDIES

An *in vitro* testis extract translation system would be of use in addressing questions of mechanism. Translation extracts from yeast, *Drosophila* ovaries and *Drosophila* embryos have been valuable (Iizuka et al. 1994; Lie and Macdonald 1999; Gebauer *et al.*, 1999). Experiments in *Drosophila* have shown the importance of which tissues are used for the extracts, since the specific translation of *oskar* mRNA that is seen in *Drosophila* was recapitulated in ovarian extracts, but not in embryonic extracts or

rabbit reticulocyte extracts (Lie and Macdonald 1999). One basic advantage to an *in vitro* translation system is the ability to look at the dependence on the 5' cap and 3' poly(A) tail. Reporter mRNAs can be synthesized *in vitro* with or without either modification and tested to measure the effects of capping and polyadenylation.

Chuck Connolly in our lab tried to establish an *in vitro* testis extract translation system, but was unable to do so. One important consideration is to make an extract from testes that do not have the activating factor, such as from a prepubertal mouse or a mouse from a transgenic line in which spermatogenesis is arrested (Lee *et al.*, 1995) to recapitulate *Prm1* like repression. His goal was to use the system to identify translational repressors, but the extract could also be used to test the function of candidate proteins. He was unable to detect the translation of an added reporter mRNA in an extract that had been treated with micrococcal nuclease to degrade endogenous mRNAs. He was able to detect ^{35}S-methionine labeled proteins in an extract that was not nuclease treated. This demonstrates that elongation of endogenous mRNAs was occurring in the extract, but does not allow us to make any conclusions about initiation. The use of a more sensitive reporter, such as luciferase, may allow the detection of the translation of the added reporter mRNA in an extract that is not micrococcal nuclease treated. Different 3' UTRs would then be used to attempt to recapitulate the specific translational repression of *Prm1* that is seen in spermatogenesis. If an extract is seen to support the specific translational repression of *Prm1*, then purified candidate proteins can be added. Candidate proteins include PRBP, MSY2 and MSY4 or purified fractions from cytoplasmic extracts or

polysome gradients. Experiments could also be done with testis extracts that are missing candidate proteins, either through immunodepletion or by using mutant mice that are missing that protein.

Xenopus oocytes may be a more tractable system to study possible functions of mouse Y box proteins. FRGY2 has been seen to specifically repress reporter mRNAs in *Xenopus* oocytes (Bouvet and Wolffe, 1994). Since *Xenopus* oocytes can support repression by another Y box family member, they are likely to have any other factors that MSY4 needs for its function. The same reporter mRNAs and candidate proteins that were discussed for the testis extract above would be used in *Xenopus* oocytes. Either the mRNAs for the reporters and protein(s) can be injected into the oocyte or the mRNP can be reconstituted *in vitro* with the reporter mRNA and recombinant protein and then the complex injected into oocytes to assay mRNA stability and translation. Experiments in *Xenopus* oocytes also allow the inquiry into a link between transcription and translation.

Tissue culture is another system to study the function of MSY4. One advantage to tissue culture is that the protein of interest can be expressed at high levels in a stable cell line before transfection of reporter mRNAs. Preliminary experiments in NIH3T3 cells suggests that the YRS is sufficient for specific translational repression in the context of *Prm1* $_{1\text{-}37}$. NIH3T3 cells contain MSY4 at a level that is lower than in mouse testis, but detectable by western analysis (HD unpublished results). These results are consistent with the transgenic results, but the tissue culture system is more amenable to manipulation. Further experiments with cell lines that over express MSY4 will show if MSY4 can

amplify the YRS-dependent repression. In addition cell lines could be established from mice lacking *Msy4*, to investigate the function of MSY4.

Experiments with internal ribosome entry sites (IRESs) can test the dependence of translational repression on the 5' cap. Chuck Connolly in our lab has tried to investigate the mechanism of *Prm1* translational repression using transgenes that contain IRESs. Initial data support the model of cap-dependent repression, since an IRES interferes with the usual repression mediated by the *Prm1* 3'UTR. It's too early to make definite conclusions from this work, however, since control experiments are still being completed.

BIBLIOGRAPHY

Ahringer, J., and Kimble, J. (1991). Control of the sperm-oocyte switch in Caenorhabditis elegans hermaphrodites by the fem-3 3' untranslated region. *Nature* **349,** 346-8.

Balhorn, R. (1982). A model for the structure of chromatin in mammalian sperm. *J Cell Biol* **93,** 298-305.

Bellve, A. R. (1979). The molecular biology of mammalian spermatogenesis. *In* "Oxford Reviews of Reproductive Biology" (C. A. Finn, Ed.), Vol. 1, pp. 159-261. Oxford University Press, London and New York.

Bienz, M. (1986). A CCAAT box confers cell-type-specific regulation on the Xenopus hsp70 gene in oocytes. *Cell* **46,** 1037-42.

Bouvet, P., Matsumoto, K., and Wolffe, A. P. (1995). Sequence-specific RNA recognition by the Xenopus Y-box proteins. An essential role for the cold shock domain. *J Biol Chem* **270,** 28297-303.

Bouvet, P., and Wolffe, A. P. (1994). A role for transcription and FRGY2 in masking maternal mRNA within Xenopus oocytes. *Cell* **77,** 931-41.

Braddock, M., Muckenthaler, M., White, M. R., Thorburn, A. M., Sommerville, J., Kingsman, A. J., and Kingsman, S. M. (1994). Intron-less RNA injected into the nucleus of Xenopus oocytes accesses a regulated translation control pathway. *Nucleic Acids Res* **22,** 5255-64.

Braun, R. E. (1990). Temporal translational regulation of the protamine 1 gene during mouse spermatogenesis. *Enzyme* **44,** 120-128.

Braun, R. E., Behringer, R. R., Peschon, J. J., Brinster, R. L., and Palmiter, R. D. (1989a). Genetically haploid spermatids are phenotypically diploid. *Nature* **337,** 373-376.

Braun, R. E., Peschon, J. J., Behringer, R. R., Brinster, R. L., and Palmiter, R. D. (1989b). Protamine 3'-untranslated sequences regulate temporal translational control and subcellular localization of growth hormone in spermatids of transgenic mice. *Genes Dev* **3,** 793-802.

Breeden, L., and Nasmyth, K. (1985). Regulation of the yeast HO gene. *Cold Spring Harb Symp Quant Biol* **50,** 643-50.

Burd, C. G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265,** 615-21.

Cathala, G., Savouret, J. F., Mendez, B., West, B. L., Karin, M., Martial, J. A., and Baxter, J. D. (1983). A method for isolation of intact, translationally active ribonucleic acid. *Dna* **2,** 329-35.

Chang, J. S., Tan, L., and Schedl, P. (1999). The Drosophila CPEB homolog, orb, is required for oskar protein expression in oocytes. *Dev Biol* **215,** 91-106.

Chen, C. Y., and Sarnow, P. (1995). Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* **268,** 415-7.

Coles, L. S., Diamond, P., Occhiodoro, F., Vadas, M. A., and Shannon, M. F. (1996). Cold shock domain proteins repress transcription from the GM-CSF promoter. *Nucleic Acids Res* **24,** 2311-7.

Craig, A. W., Haghighat, A., Yu, A. T., and Sonenberg, N. (1998). Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* **392,** 520-3.

Cummings, A., and Sommerville, J. (1988). Protein kinase activity associated with stored messenger ribonucleoprotein particles of Xenopus oocytes. *J Cell Biol* **107,** 45-56.

Curtis, D., Lehmann, R., and Zamore, P. D. (1995). Translational regulation in development. *Cell* **81,** 171-178.

Darnbrough, C. H., and Ford, P. J. (1981). Identification in Xenopus laevis of a class of oocyte-specific proteins bound to messenger RNA. *Eur J Biochem* **113,** 415-24.

Davies, H. G., Giorgini, F., Fajardo, M. A., and Braun, R. E. (2000). A sequence-specific RNA binding complex expressed in murine germ cells contains MSY2 and a novel protein, MSY4. *(submitted)* .

de Moor, C. H., and Richter, J. D. (1997). The Mos pathway regulates cytoplasmic polyadenylation in Xenopus oocytes. *Mol Cell Biol* **17,** 6419-26.

DeNoto, F. M., Moore, D. D., and Goodman, H. M. (1981). Human growth hormone DNA sequence and mRNA structure: Possible alternative splicing. *Nucleic Acids Res* **9,** 3719-3730.

Deschamps, S., Viel, A., Garrigos, M., Denis, H., and le Maire, M. (1992). mRNP4, a major mRNA-binding protein from Xenopus oocytes is identical to transcription factor FRG Y2. *J Biol Chem* **267,** 13799-802.

Didier, D. K., Schiffenbauer, J., Woulfe, S. L., Zacheis, M., and Schwartz, B. D. (1988). Characterization of the cDNA encoding a protein binding to the major histocompatibility complex class II Y box. *Proc Natl Acad Sci U S A* **85,** 7322-6.

Dignam, J. D., Lebovitz, R. M., and Roeder, R. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11,** 1475-1489.

Dugaiczyk, A., Haron, J. A., Stone, E. M., Dennison, O. E., Rothblum, K. N., and Schwartz, R. J. (1983). Cloning and sequencing of a deoxyribonucleic acid copy of glyceraldehyde-3-phosphate dehydrogenase messanger ribonulceic acid isolated from chicken muscle. *Biochemistry* **22,** 1605-13.

Fajardo, M. A., Butner, K. A., Lee, K., and Braun, R. E. (1994). Germ cell-specific proteins interact with the 3' untranslated regions of *Prm-1* and *Prm-2* mRNA. *Dev Biol* **166,** 643-653.

Fajardo, M. A., Haugen, H. S., Clegg, C. H., and Braun, R. E. (1997). Separate elements in the 3' untranslated region of the mouse protamine 1 mRNA regulate translational repression and activation during murine spermatogenesis. *Dev Biol* **191,** 42-52.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783-791.

Felsenstein, J. (1993). PHYLIP (Phylogeny Inference Package). Distributed by author, Seattle, WA.

Futterer, J., Kiss-Laszlo, Z., and Hohn, T. (1993). Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* **73,** 789-802.

Gebauer, F., Corona, D. F., Preiss, T., Becker, P. B., and Hentze, M. W. (1999). Translational control of dosage compensation in Drosophila by Sex- lethal: cooperative silencing via the 5' and 3' UTRs of msl-2 mRNA is independent of the poly(A) tail. *Embo J* **18,** 6146-54.

Gebauer, F., and Richter, J. D. (1996). Mouse cytoplasmic polyadenylylation element binding protein: an evolutionarily conserved protein that interacts with the cytoplasmic polyadenylylation elements of c-mos mRNA. *Proc Natl Acad Sci U S A* **93,** 14602-7.

Gray, N. K., and Hentze, M. W. (1994). Iron regulatory protein prevents binding of the 43S translation pre- initiation complex to ferritin and eALAS mRNAs. *Embo J* **13,** 3882-91.

Gu, W., Tekur, S., Reinbold, R., Eppig, J. J., Choi, Y. C., Zheng, J. Z., Murray, M. T., and Hecht, N. B. (1998). Mammalian male and female germ cells express a germ cell-specific Y-box protein, MSY2. *Biology Of Reproduction* **59,** 1266-1274.

Guddat, U., Bakken, A. H., and Pieler, T. (1990). Protein-mediated nuclear export of RNA: 5S rRNA containing small RNPs in xenopus oocytes. *Cell* **60,** 619-28.

Gunkel, N., Yano, T., Markussen, F. H., Olsen, L. C., and Ephrussi, A. (1998). Localization-dependent translation requires a functional interaction between the 5' and 3' ends of oskar mRNA. *Genes Dev* **12,** 1652-64.

Hake, L. E., Mendez, R., and Richter, J. D. (1998). Specificity of RNA binding by CPEB: requirement for RNA recognition motifs and a novel zinc finger. *Mol Cell Biol* **18,** 685-93.

Han, J. R., Yiu, G. K., and Hecht, N. B. (1995). Testis/brain RNA-binding protein attaches translationally repressed and transported mRNAs to microtubules. *Proc Natl Acad Sci U S A* **92,** 9550-4.

Hershey, J. W. (1991). Translational control in mammalian cells. *Annu Rev Biochem* **60,** 717-55.

Hinnebusch, A. G. (1996). "Translational control of GCN4: gene-specific regulation by phosphorylation of eIF2." Cold Spring Harbor Laboratory Press,

Jan, E., Motzny, C. K., Graves, L. E., and Goodwin, E. B. (1999). The STAR protein, GLD-1, is a translational regulator of sexual identity in Caenorhabditis elegans. *Embo J* **18,** 258-69.

Jan, E., Yoon, J. W., Walterhouse, D., Iannaccone, P., and Goodwin, E. B. (1997). Conservation of the C.elegans tra-2 3'UTR translational control. *Embo J* **16,** 6301-13.

Kempe, E., Muhs, B., and Schafer, M. (1993). Gene regulation in Drosophila spermatogenesis: analysis of protein binding at the translational control element TCE. *Dev Genet* **14,** 449-59.

Kierszenbaum, A. L., and Tres, I. L. (1975). Structural and transcriptional features of the mouse spermatid genome. *J Cell Biol* **65,** 258-270.

Kim Ha, J., Kerr, K., and Macdonald, P. M. (1995). Translational regulation of *oskar* mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell* **81,** 403-412.

Kleene, K. C. (1996). Patterns of translational regulation in the mammalian testis. *Mol Reprod Dev* **43,** 268-81.

Kleene, K. C., Distel, R. J., and Hecht, N. B. (1984). Translational regulation and deadenylation of a protamine mRNA during spermiogenesis in the mouse. *Dev Biol* **105,** 71-79.

Kraemer, B., Crittenden, S., Gallegos, M., Moulder, G., Barstead, R., Kimble, J., and Wickens, M. (1999). NANOS-3 and FBF proteins physically interact to control the sperm- oocyte switch in caenorhabditis elegans [In Process Citation]. *Curr Biol* **9,** 1009-18.

Kudo, S., Mattei, M. G., and Fukuda, M. (1995). Characterization of the gene for dbpA, a family member of the nucleic-acid-binding proteins containing a cold-shock domain. *Eur J Biochem* **231,** 72-82.

Kuge, H., and Richter, J. D. (1995). Cytoplasmic 3' poly(A) addition induces 5' cap ribose methylation: implications for translational control of maternal mRNA. *Embo J* **14,** 6301-10.

Kwon, Y. K., and Hecht, N. B. (1991). Cytoplasmic protein binding to highly conserved sequences in the 3' untranslated region of mouse protamine 2 mRNA, a translationally regulated transcript of male germ cells. *Proc Natl Acad Sci U S A* **88,** 3584-3588.

Kwon, Y. K., and Hecht, N. B. (1993). Binding of a phosphoprotein to the 3' untranslated region of the mouse protamine 2 mRNA temporally represses its translation. *Mol Cell Biol* **13,** 6547-57.

Kwon, Y. K., Murray, M. T., and Hecht, N. B. (1993). Proteins homologous to the Xenopus germ cell-specific RNA-binding proteins p54/p56 are temporally expressed in mouse male germ cells. *Dev Biol* **158,** 99-100.

Lee, K., Fajardo, M. A., and Braun, R. E. (1996). A testis cytoplasmic RNA-binding protein that has the properties of a translational repressor. *Mol Cell Biol* **16,** 3023-3034.

Lee, K., Haugen, H. S., Clegg, C. H., and Braun, R. E. (1995). Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. *Proc Natl Acad Sci U S A* **92,** 12451-12455.

Lindsley, D. L., and Grell, E. H. (1969). Spermiogenesis without chromosomes in Drosophila melanogaster. *Genetics* **61,** Suppl.

Macejak, D. G., and Sarnow, P. (1991). Internal initiation of translation mediated by the 5' leader of a cellular mRNA [see comments]. *Nature* **353,** 90-4.

Matsumoto, K., Meric, F., and Wolffe, A. P. (1996). Translational repression dependent on the interaction of the Xenopus Y-box protein FRGY2 with mRNA. Role of the cold

shock domain, tail domain, and selective RNA sequence recognition. *J Biol Chem* **271,** 22706-12.

Matsumoto, K., and Wolffe, A. P. (1998). Gene regulation by Y-box proteins: coupling control of transcription and translation. *Trends In Cell Biology* **8,** 318-323.

McGrew, L. L., and Richter, J. D. (1990). Translational control by cytoplasmic polyadenylation during Xenopus oocyte maturation: characterization of cis and trans elements and regulation by cyclin/MPF. *Embo J* **9,** 3743-51.

Mezquita, J., Pau, M., and Mezquita, C. (1998). Several novel transcripts of glyceraldehyde-3-phosphate dehydrogenase expressed in adult chicken testis. *J Cell Biochem* **71,** 127-39.

Monesi, V. (1964). Ribonucleic acid synthesis during mitosis and meiosis in the mouse testis. *J Cell Biol* **22,** 521-532.

Murray, M. T. (1994). Nucleic acid-binding properties of the Xenopus oocyte Y box protein mRNP3+4. *Biochemistry* **33,** 13910-7.

Murray, M. T., Krohne, G., and Franke, W. W. (1991). Different forms of soluble cytoplasmic mRNA binding proteins and particles in Xenopus laevis oocytes and embryos. *J Cell Biol* **112,** 1-11.

Nikolajczyk, B. S., Murray, M. T., and Hecht, N. B. (1995). A mouse homologue of the Xenopus germ cell-specific ribonucleic acid/deoxyribonucleic acid-binding proteins p54/p56 interacts with the protamine 2 promoter. *Biol Reprod* **52,** 524-30.

Oakberg, E. F. (1956). A description of spermiogenesis in the mouse and its use in analysis of the cycle of the seminiferous epithelium and germ cell renewal. *Am J Anat* **99,** 391-413.

Oh, S. K., Scott, M. P., and Sarnow, P. (1992). Homeotic gene Antennapedia mRNA contains 5'-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes Dev* **6,** 1643-53.

Oko, R., Korley, R., Murray, M. T., Hecht, N. B., and Hermo, L. (1996). Germ cell-specific DNA and RNA binding proteins p48/52 are expressed at specific stages of male germ cell development and are present in the chromatoid body. *Mol Reprod Dev* **44,** 1-13.

Pelham, H. R., and Brown, D. D. (1980). A specific transcription factor that can bind either the 5S RNA gene or 5S RNA. *Proc Natl Acad Sci U S A* **77,** 4170-4.

Preiss, T., and Hentze, M. W. (1998). Dual function of the messenger RNA cap structure in poly(A)-tail- promoted translation in yeast. *Nature* **392,** 516-20.

Ranjan, M., Tafuri, S. R., and Wolffe, A. P. (1993). Masking mRNA from translation in somatic cells. *Genes Dev* **7,** 1725-36.

Richter, J. D., and Smith, L. D. (1984). Reversible inhibition of translation by Xenopus oocyte-specific proteins. *Nature* **309,** 378-80.

Russell, L. D., Ettlin, R. A., SinhaHikim, A. P., and Clegg, E. D. (1990). "Histological and Histopathological Evaluation of the Testis."

Sachs, A. B., and Davis, R. W. (1989). The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. *Cell* **58,** 857-67.

Sachs, A. B., Davis, R. W., and Kornberg, R. D. (1987). A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. *Mol Cell Biol* **7,** 3268-76.

Sakura, H., Maekawa, T., Imamoto, F., Yasuda, K., and Ishii, S. (1988). Two human genes isolated by a novel method encode DNA-binding proteins containing a common region of homology. *Gene* **73,** 499-507.

Sall'es, F. J., Lieberfarb, M. E., Wreden, C., Gergen, J. P., and Strickland, S. (1994). Coordinate initiation of Drosophila development by regulated polyadenylation of maternal messenger RNAs. *Science* **266,** 1996-9.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning, A Lab Manual, 2nd edition." Cold Spring Harbor Press, Cold Spring Harbor, NY.

Schafer, M., Nayernia, K., Engel, W., and Schafer, U. (1995). Translational control in spermatogenesis. *Dev Biol* **172,** 344-52.

Schindelin, H., Jiang, W., Inouye, M., and Heinemann, U. (1994). Crystal structure of CspA, the major cold shock protein of Escherichia coli. *Proc Natl Acad Sci U S A* **91,** 5119-23.

Schindelin, H., Marahiel, M. A., and Heinemann, U. (1993). Universal nucleic acid-binding domain revealed by crystal structure of the B. subtilis major cold-shock protein. *Nature* **364,** 164-8.

Schlicker, M., Reim, K., Schluter, G., and Engel, W. (1997). Specific binding of a 47-kilodalton protein to the 3' untranslated region of rat transition protein 2 messenger ribonucleic acid. *Biol Reprod* **56,** 697-706.

Schmidt, E. E., Hanson, E. S., and Capecchi, M. R. (1999). Sequence-independent assembly of spermatid mRNAs into messenger ribonucleoprotein particles. *Mol Cell Biol* **19,** 3904-15.

Schneppenheim, R., and Rautenberg, P. (1987). A luminescence Western blot with enhanced sensitivity for antibodies to human immunodeficiency virus. *Eur J Clin Microbiol* **6,** 49-51.

Schumacher, J. M., Lee, K., Edelhoff, S., and Braun, R. E. (1995). Distribution of TENR, an RNA Binding Protein, in a Lattice-Like Network within the Spermatid Nucleus in the Mouse. *Biol of Reprod* **52,** 1274-1283.

SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S., and Wickens, M. (1996). A three-hybrid system to detect RNA-protein interactions in vivo. *Proc Natl Acad Sci U S A* **93,** 8496-501.

Sheets, M. D., Fox, C. A., Hunt, T., Vande Woude, G., and Wickens, M. (1994). The 3'-untranslated regions of c-mos and cyclin mRNAs stimulate translation by regulating cytoplasmic polyadenylation. *Genes Dev* **8,** 926-38.

Sommerville, J., and Ladomery, M. (1996a). Masking of mRNA by Y-box proteins. *Faseb J* **10,** 435-43.

Sommerville, J., and Ladomery, M. (1996b). Transcription and masking of mRNA in germ cells: involvement of Y-box proteins. *Chromosoma* **104,** 469-78.

Standart, N., and Jackson, R. J. (1994). Regulation of translation by specific protein/mRNA interactions. *Biochimie* **76,** 867-79.

Stanker, L. H., Wyrobek, A., McKeown, C., and Balhorn, R. (1993). Identification of the binding site of two monoclonal antibodies to human protamine. *Mol Immunol* **30,** 1633-1638.

Stebbings, L., Grimes, B. R., and Bownes, M. (1998). A testis-specifically expressed gene is embedded within a cluster of maternally expressed genes at 89B in Drosophila melanogaster. *Dev Genes Evol* **208,** 523-30.

Stebbins Boaz, B., and Richter, J. D. (1997). Translational control during early development. *Crit Rev Eukaryot Gene Expr* **7,** 73-94.

Stebbins-Boaz, B., Cao, Q., de Moor, C. H., Mendez, R., and Richter, J. D. (1999). Maskin is a CPEB-associated factor that transiently interacts with eIF- 4E [In Process Citation]. *Mol Cell* **4,** 1017-27.

Tafuri, S. R., Familari, M., and Wolffe, A. P. (1993). A mouse Y box protein, MSY1, is associated with paternal mRNA in spermatocytes. *J. Biol. Chem.* **268,** 12213-12220.

Tafuri, S. R., and Wolffe, A. P. (1992). DNA binding, multimerization, and transcription stimulation by the *Xenopus* Y box proteins in vitro. *New Biol* **4,** 349-359.

Tafuri, S. R., and Wolffe, A. P. (1993). Selective recruitment of masked maternal mRNA from messenger ribonucleoprotein particles containing FRGY2 (mRNP4). *J Biol Chem* **268,** 24255-61.

Tarun, S. Z., Jr., and Sachs, A. B. (1995). A common function for mRNA 5' and 3' ends in translation initiation in yeast. *Genes Dev* **9,** 2997-3007.

Theunissen, O., Rudt, F., Guddat, U., Mentzel, H., and Pieler, T. (1992). RNA and DNA binding zinc fingers in Xenopus TFIIIA. *Cell* **71,** 679-90.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22,** 4673-80.

Tsai, D. E., Harper, D. S., and Keene, J. D. (1991). U1-snRNP-A protein selects a ten nucleotide consensus sequence from a degenerate RNA pool presented in various structural contexts. *Nucleic Acids Res* **19,** 4931-6.

Vassalli, J. D., Huarte, J., Belin, D., Gubler, P., Vassalli, A., O'Connell, M. L., Parton, L. A., Rickles, R. J., and Strickland, S. (1989). Regulated polyadenylation controls mRNA translation during meiotic maturation of mouse oocytes. *Genes Dev* **3,** 2163-71.

Walker, J., Dale, M., and Standart, N. (1996). Unmasking mRNA in clam oocytes: role of phosphorylation of a 3' UTR masking element-binding protein at fertilization. *Dev Biol* **173,** 292-305.

Walker, J., Minshall, N., Hake, L., Richter, J., and Standart, N. (1999). The clam 3' UTR masking element-binding protein p82 is a member of the CPEB family. *Rna* **5,** 14-26.

Wells, S. E., Hillner, P. E., Vale, R. D., and Sachs, A. B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* **2,** 135-40.

Wickens, M., Kimble, J., and Strickland, S. (1996). "Translational control of developmental decisions." Cold Spring Harbor Laboratory Press,

Yanicostas, C., and Lepesant, J. A. (1990). Transcriptional and translational cis-regulatory sequences of the spermatocyte-specific Drosophila janusB gene are located in the 3' exonic region of the overlapping janusA gene. *Mol Gen Genet* **224,** 450-8.

Yurkova, M. S., and Murray, M. T. (1997). A translation regulatory particle containing the Xenopus oocyte Y box protein mRNP3+4. *J Biol Chem* **272,** 10870-6.

Zambrowicz, B. P., Harendza, C. J., Zimmermann, J. W., Brinster, R. L., and Palmiter, R. D. (1993). Analysis of the mouse protamine 1 promoter in transgenic mice. *Proc Natl Acad Sci USA* **90,** 5071-5075.

Zhang, B., Gallegos, M., Puoti, A., Durkin, E., Fields, S., Kimble, J., and Wickens, M. P. (1997). A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line. *Nature* **390,** 477-84.

Zhong, J., Peters, A. H., Lee, K., and Braun, R. E. (1999). A double-stranded RNA binding protein required for activation of repressed messages in mammalian germ cells. *Nat Genet* **22,** 171-4.

Zhong, J., Peters, A. H. F. M., Lee, K., Liggitt, D. H., Abraham, N., Bell, J. C., Sonenberg, N., and Braun, R. E. (1999b). Growth retardation and perinatal lethality in Prbp deficient mice is not caused by hyper-activation of PKR. *(submitted)* .

VITA

HOLLY GIBS DAVIES

University of Washington

2000

Holly Davies was born to Charlie and Sally Davies on October 11, 1971 in Glen Cove, New York. In 1976 the family moved back to Sea Cliff, New York, where the family has lived for several generations. As the name implies, Sea Cliff is on the water, and Holly and her brother Andy had a wonderful childhood hanging around the beach and sailing. After being graduated from North Shore High School in 1989, Holly enrolled at Cornell University to study biology. She concentrated in genetics and development and received a B.S. in January of 1993. Along the way she met her future husband, Adrian, in a class on ecological genetics. Adrian really wanted to study salmon, so in 1994 Holly moved to Seattle to start graduate school and joined the lab of Dr. Robert Braun.