

RNA-binding and Metabolite Mediated Inhibition of Musashi1 Homologs

A Major Qualifying Project Report

Submitted to the Faculty of

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

In

Biology and Biotechnology

&

Biochemistry

By

Emily Johnson

April __, 2013

Abstract

Precise regulation of gene expression controls cell fate determination, and deregulation of gene expression is seen in cancer. The Musashi family of RNA-binding proteins regulates the translation of mRNA targets in neural and epithelial stem and progenitor cells. In neural and epithelial tumors, Musashi1 (MSI1) expression levels are correlated with malignancy and proliferation, and knockdown of MSI1 by RNAi leads to tumor regression. Investigation into an inhibitor of MSI1 has revealed oleic acid as a potent inhibitor of RNA-binding activity, but the mechanism and biological significance are unknown. The current study investigates differences in RNA-binding and inhibition constants across five MSI1 homologs from diverse species. Although MSI1 homolog sequences are highly conserved, individual amino acid differences exist and were hypothesized to yield differences in activity and to point towards functional differences and the biological significance of oleic acid inhibition.

Acknowledgements

I would like to thank Dr. Sean Ryder for allowing me to work in his lab at UMass Medical School and for his advice on project direction. Thank you to both Ryder Lab members Carina Clingman and Dr. Ruthie Zearfoss for help with cloning, protein purification, assay protocols and answering all of my questions. A special thank you to Ruthie for teaching me all of the basic techniques when I first joined the Ryder lab. Finally, I would like to thank all of the above and Professors Duffy and Heilman for their helpful discussion throughout this project and for their help and patience in editing.

Table of Contents

Abstract.....	i
Acknowledgements.....	ii
Table of Contents.....	iii
Introduction	1
Materials and Methods.....	12
Results.....	21
Discussion.....	29
Works Cited.....	32
Figures.....	36
Tables	44

Introduction

As the sophistication of scientific research progresses, alterations in signaling pathways are being discovered to have wide-spread implications in cancer development and tumor cell survival. Up-regulation of pathways promoting cellular proliferation, which are active in stem and progenitor cells, has been found to correlate with malignancy of tumors. Up-regulation of these pathways causes increased tumor cell proliferation; therefore, inhibition of these pathways is being investigated for future treatment. A more complete understanding of the mechanism and regulation of these pathways could lead to treatments for some of the 1.6 million people diagnosed with cancer each year (American Cancer Society, 2012).

As organisms develop, stem and progenitor cells first depend on these signaling pathways for survival. Cell populations are maintained throughout development by regulatory mechanisms that occur at the levels of transcription and translation. Post transcriptional regulation of target mRNA by RNA-binding proteins is known to regulate the translation of proteins involved in signaling pathways (Glisovic, Bachorik, Yong, & Dreyfuss, 2008). Musashi-1 (MSI1) is a member of the Musashi family of RNA-binding proteins, and is expressed in neural and epithelial stem and progenitor cells where it represses the translation of target mRNAs (Kaneko et al., 1999; Okano, Imai, & Okabe, 2002). MSI1 promotes cellular proliferation, and increased levels of MSI1 have been shown to promote growth and division of tumor cells (Figure 1) (Kanemura et al., 2001; Sanchez-Diaz, Burton, Burns, Hung, & Penalva, 2008; Seigel, Hackam, Ganguly, Mandell, & Gonzalez-Fernandez, 2007; Sureban et al., 2008; Wang et al., 2010). Knockdown of MSI1

leads to tumor cell death; therefore, inhibitors of MSI1 could lead to future treatments for tumors (Kanemura et al., 2001; Sanchez-Diaz et al., 2008; Seigel et al., 2007; Sureban et al., 2008; Wang et al., 2010). The Ryder Lab at the University of Massachusetts Medical School has discovered a small-molecule metabolite, oleic acid, which inhibits the RNA-binding activity of *Mus musculus* Musashi1 (mouse MSI1) (Clingman et al., *in review*). This was found in a small molecule screen for inhibitors of MSI1, and was one of only four hits out of over 30,000 compounds screened (Clingman et al., *in review*). In this study, four additional MSI1 homologs were tested for RNA-binding activity and inhibition by oleic acid in order to investigate possible future animal models and cell lines for testing the effect of oleic acid on tumor cells.

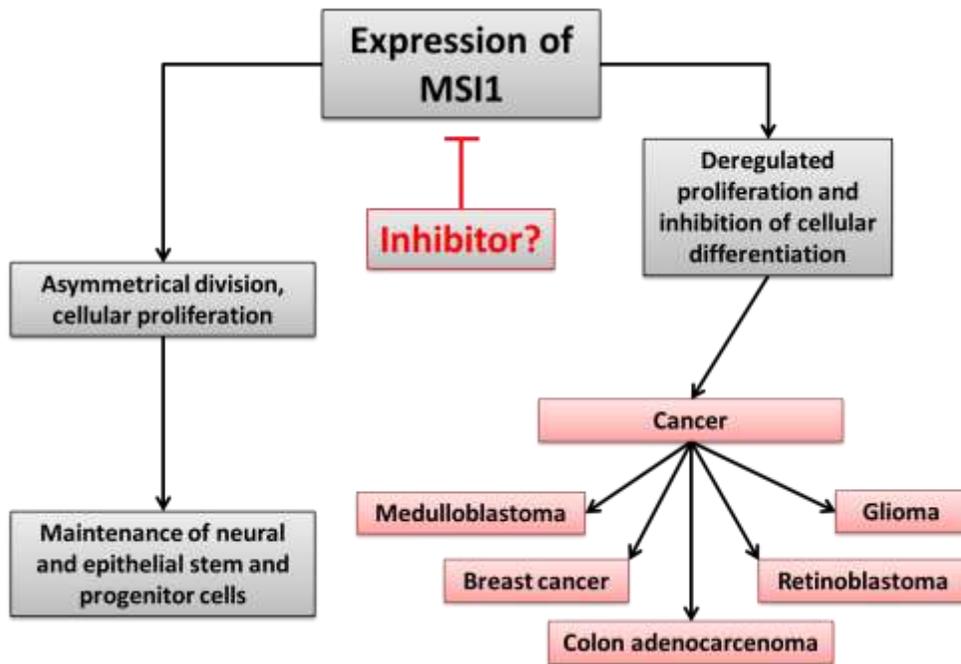


Figure 1: Consequences of the expression of MSI1 (Kanemura et al., 2001; Muto et al., 2012; Sakakibara, Nakamura, Satoh, & Okano, 2001; Sanchez-Diaz et al., 2008; Seigel et al., 2007; Sureban et al., 2008; Wang et al., 2008)

The Musashi RNA-binding protein family

The Musashi family of RNA-binding proteins consists of Musashi1 (MSI1) and Musashi2 (MSI2) in humans (Kaneko et al., 1999; Sakakibara et al., 1996, 2001). The Musashi family is characterized by two N-terminal RNA-recognition motifs (RRMs) (Sakakibara et al., 1996, 2001). RRMs facilitate specific recognition of RNA sequences and have high RNA-binding affinity (Maris, Dominguez, & Allain, 2005). MSI2 expression is seen specifically in proliferating cells in the ventricular and subventricular zones of the brain, and hematopoietic stem cells (De Andrés-Aguayo et al., 2011; Sakakibara et al., 2001). MSI1 is expressed in neural stem and progenitor cells, including committed oligodendrocyte progenitor cells and cells of the astrocyte lineage (Dobson, Zhou, Flint, & Armstrong, 2009; Kaneko et al., 1999; Sakakibara & Okano, 1997). MSI1 expression is lost in differentiated mammalian neuronal and epithelial cells, suggesting its role in proliferation and differentiation of stem and progenitor cells (Sakakibara et al., 1996). In flies (*Drosophila melanogaster*), Msi plays a role specifically in asymmetrical division (Sakakibara et al., 1996).

Fly Msi was the first Musashi protein to be identified (Sakakibara et al., 1996). Since, MSI1 homologs have been identified in human (MSI1), mouse (*Mus musculus*, MSI1), nematode (*Caenorhabditis elegans*, MSI-1), zebrafish (*Danio rerio*, msi1) and fly (Rbp6), among others (Good et al., 1998; Okabe, Imai, Kurusu, Hiromi, & Okano, 2001; Sakakibara et al., 1996; Shibata et al., 2012; Siddall et al., 2012; Yoda, Sawa, & Okano, 2000). The fly MSI1 homolog Rbp6 shares more sequence similarity to human MSI1 than does fly Msi (Siddall et al., 2012). Sequence similarities of the full length proteins, and more importantly of the functional RRMs, classify

Musashi as an evolutionarily conserved RNA-binding protein. Table 1 lists sequence similarities of MSI1 homologs, where conservative differences in amino acids count as similar amino acids. There is dramatically higher sequence similarity between the two RRM, highlighting their importance in the function of all homologs.

Table 1: Sequence similarity of MSI1 homologs

Organism	Abbreviation	Percent amino acid similarity to human MSI1	
		Full length protein	RRMs only
<i>Homo sapiens</i>	MSI1	-	-
<i>Mus musculus</i>	MSI1	99.72%	99.35%
<i>Danio rerio</i>	msi1	73.93%	95.91%
<i>Caenorhabditis elegans</i>	MSI-1	12.81%	85.89%
<i>Drosophila melanogaster</i>	Rbp6	11.65%	85.89%
<i>Drosophila melanogaster</i>	Msi	7.42%	44.44%

Musashi1 structure facilitates RNA-binding

The functional domain of mouse MSI1 is within the first RRM (RRM1) (Sakakibara et al., 1996).

This is suspected to be true of the MSI1 homologs in this study, based on the sequence similarities of the RRM domains (Table 1). Mouse MSI1 was previously used by the Ryder Lab, and is shown in Figure 2a to consist of 362 amino acids with RRM1 and RRM2 located at amino acids 21-100 and 110-189, respectively (Sakakibara et al., 1996). The N-terminal RRM (RRM1) has a higher binding affinity to the consensus RNA sequence than the C-terminal RRM (RRM2) (Miyanoiri et al., 2003). Figure 2b shows the RNA binding face of mouse MSI1 (left), which is composed of four beta-sheets (Miyanoiri et al., 2003). The positively charged beta-sheet of RRM1 allows for stronger interaction with the negative phosphate backbone of RNA than does the neutral beta-sheet of RRM2 (Miyanoiri et al., 2003). A difference in backbone dynamics of

the two beta-sheets makes RRM1's beta-sheet more flexible and capable of facilitating an induced fit during RNA recognition (Miyanoiri et al., 2003).

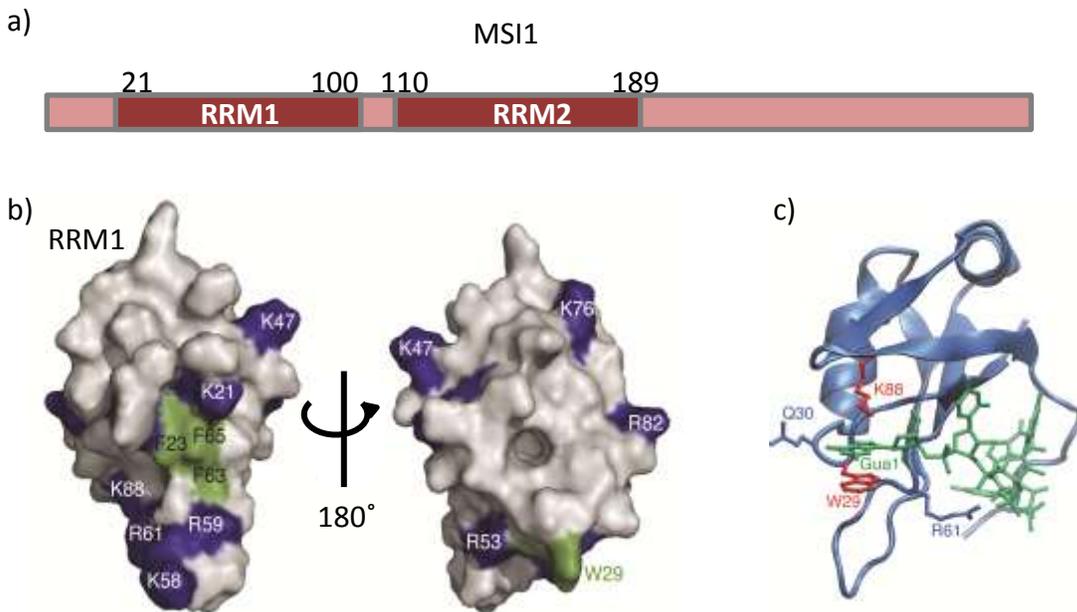


Figure 2: Mouse MSI1 structure; a) MSI1 RRMs; b) RRM1 only, right: RNA-binding face, left: hydrophobic cavity; c) MSI1 bound to RNA (5'-GUAGU-3') (Miyanoiri et al., 2003; Nagata et al., 1999; Ohyama et al., 2012)

RRM2 demonstrates weak RNA binding when measured by NMR, but no binding in gel-retardation experiments, suggesting the interference of salts and detergent with RNA binding (Nagata et al., 1999). When RRM2 is joined with RRM1, the RNA-binding affinity of RRM1 increases compared to its affinity when isolated (Nagata et al., 1999). Figure 2b also shows the alpha-helical face of mouse MSI1 (right), which has a hydrophobic cavity when viewed as a space filling model (Miyanoiri et al., 2003; Nagata et al., 1999). This hydrophobic cavity is hypothesized to be the docking site of the hydrophobic hydrocarbon tail of oleic acid (Clingman et al., *in review*).

The structure of mouse MSI1 in complex with the consensus RNA aptamer 5'-GUAGU-3' has recently been solved by NMR (Ohyama et al., 2012). It reveals the roles of aromatic stacking interactions and hydrogen bonding in RNA recognition (Figure 2c) (Ohyama et al., 2012). Stacking interactions between evolutionarily conserved tryptophan (W29) and phenylalanine (F23, F63, F65 and F96) residues of MSI1 and the aromatic bases and ribose rings of the RNA aptamer contribute to target recognition within RRM1 (Ohyama et al., 2012). Hydrogen bonds involving K88, K93, D91, V94 and K21 also contribute to RNA-binding (Ohyama et al., 2012). These specific residues facilitate binding of MSI1 to consensus binding sites found in the 3'-prime untranslated region (3'-UTR) of mRNA. It has been proposed by the Ryder Lab that binding of oleic acid shifts specific residues important in RNA recognition to allosterically inhibit RNA-binding (Clingman et al., *in review*).

MSI1 binds the 3'-UTR of target mRNA

The structure of MSI1 facilitates sequence specific binding to the 3'-UTR of mRNA transcripts involved in cell cycle regulation. The consensus sequence for MSI1 binding is (G/A)U₁₋₃AGU, where the first nucleotide can be either of the purines guanine or adenine (Imai et al., 2001). There is evidence that binding may be context dependent, specifically dependent on the formation of a stem-loop secondary structure in the target mRNA (Imai et al., 2001). MSI1 consensus sequences were found in most cases to be on the loop of a stem-loop structure, possibly facilitating recognition by MSI1 (Imai et al., 2001). After MSI1 binds the 3'-UTR of its target mRNA, it prevents the 80S ribosomal complex from forming and translating the mRNA into protein (Kawahara et al., 2008).

The mechanism for post transcriptional regulation of mRNA by MSI1 occurs through the inhibition of 80S ribosomal complex formation, which is necessary for translation in eukaryotes (Kawahara et al., 2008). In the absence of MSI1 (Figure 3a), poly(A) binding protein (PABP) binds the poly-A tail of the 3'-UTR in a target mRNA transcript (Kawahara et al., 2008). The eukaryotic initiation factor 4G (eIF4G) binds PABP to associate the mRNA with the 40S/eIF complex (Kawahara et al., 2008). This is followed by recognition of the start codon and binding of the 60S ribosomal subunit to form the 80S ribosomal complex. Translation of the mRNA then proceeds (Kawahara et al., 2008).

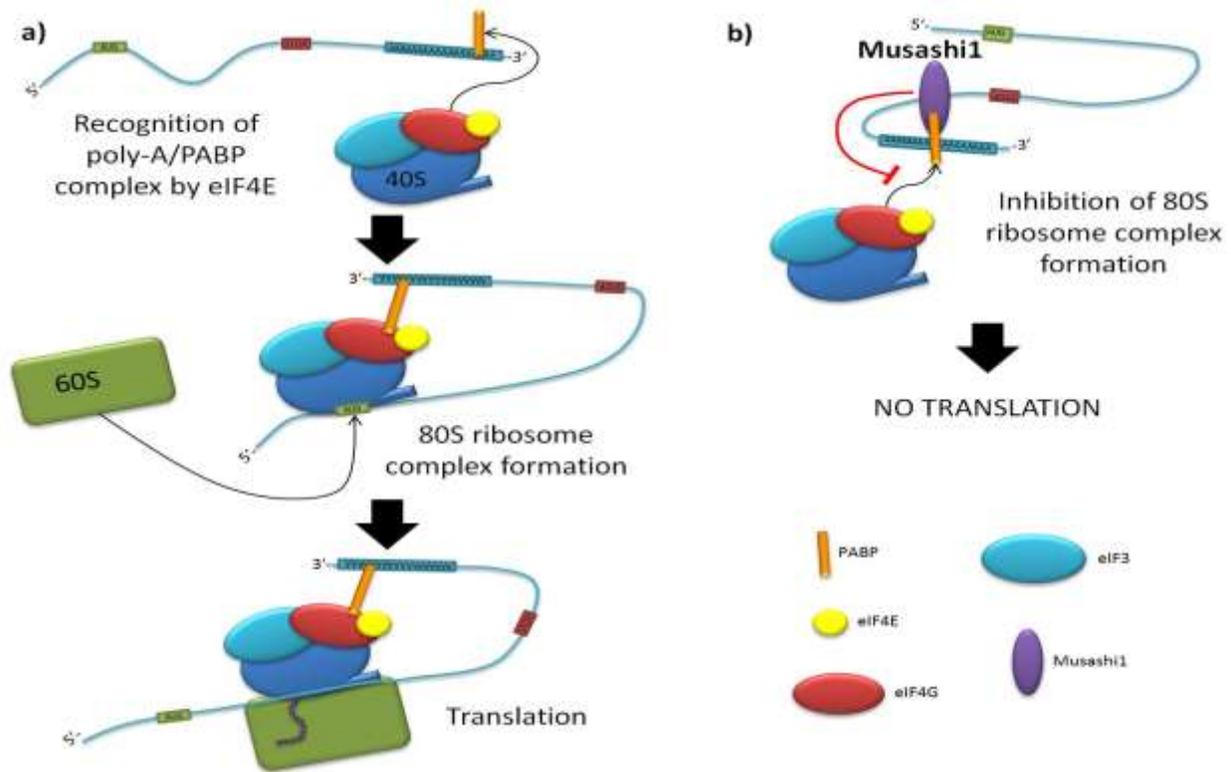


Figure 3: Binding of MSI1 inhibits formation of 80S ribosomal complex (a) Absence of MSI1: eIF4G recognizes PABP, 80S ribosomal complex formed; translation is initiated (b) Presence of MSI1: MSI1 blocks eIF4G recognition of PABP; translation is inhibited (figure adapted from Kawahara et al., 2008).

When MSI1 is present in the soma of the cell, MSI1 binds the 3'-UTR of the target mRNA and competes with eIF4G for interaction with PABP. Binding of MSI1 to PABP prevents binding of the 60S ribosomal subunit and translation of the target mRNA is inhibited (Figure 1b) (Kawahara et al., 2008). This mechanism allows MSI1 to repress translation of target mRNA.

Musashi1 promotes proliferation through translational repression

MSI1 may use the mechanism above to alter signaling pathways that lead to increased cellular proliferation, although this has not yet been directly linked. MSI1 may also inhibit translation by different mechanisms which have not been investigated. MSI1 consensus binding sequences are found on mRNAs encoding regulators of several signaling pathways, and more are expected to be discovered. For example, MSI1 represses the translation of *numb*, which is an inhibitor of the Notch and Hedgehog signaling pathways, which both promote cellular proliferation (Di Marcotullio et al., 2006; Imai et al., 2001). MSI1 represses the translation of p21^{WAF1}, leading to decreases in G₂/M arrest and apoptosis (Battelli, Nikopoulos, Mitchell, & Verdi, 2006; Sureban et al., 2008). MSI1 also inhibits translation of Dickkopf-3 (DKK3), an inhibitor of the Wnt pathway, also leading to cellular proliferation (Wang et al., 2008). The net effect of MSI1 on these signaling pathways is an increase in stem and progenitor cell proliferation. This is vital during development for maintaining populations of undifferentiated cells, but when unregulated, MSI1 promotes tumor cell proliferation. MSI1's involvement in the Notch signaling pathway has been implicated in tumor cell growth (Muto et al., 2012).

Musashi activates the Notch signaling pathway

One of the most well characterized pathways regulated by MSI1 is the Notch signaling pathway (Figure 4). MSI1 inhibits the translation of *numb*, an inhibitor of Notch cleavage (Imai et al.,

2001). Notch is a transmembrane protein that requires association with neighboring cell surface proteins (Delta or Jagged) for intracellular cleavage (Imai et al., 2001). After association, Notch is cleaved and binds RBP-J κ . This complex enters the nucleus and activates genes including *HES1*, *HEY2* and *NOTCH2*, which are genes known to increase cellular proliferation and inhibit differentiation (Imai et al., 2001). Therefore, MSI1 activates the Notch signaling pathway and promotes cellular proliferation.

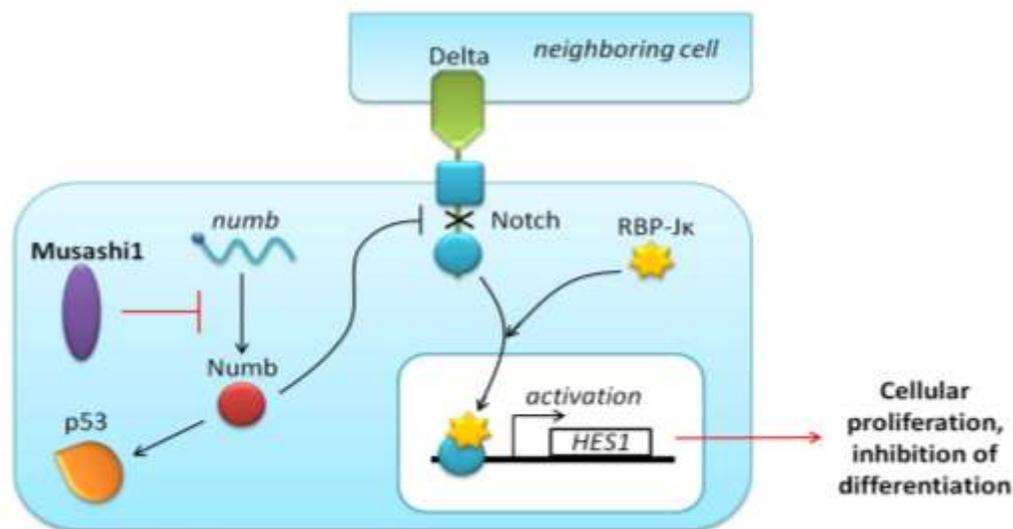


Figure 4: Activation of the Notch signaling pathway by MSI1. MSI1 represses translation of *numb*, preventing Numb from inhibiting Notch cleavage. Cleaved Notch combines with RBP-J κ to enter the nucleus and activate genes to increase cellular proliferation and inhibit differentiation (adapted from Colaluca et al., 2008; Imai et al., 2001; Sanchez-Diaz et al., 2008).

Numb has also been shown to inhibit degradation of the cell cycle inhibitor p53 (Figure 4) (Colaluca et al., 2008). In the presence of MSI1, Numb expression is decreased, leading to increased degradation of p53 (Sanchez-Diaz et al., 2008). Decreased levels of p53 have been correlated with cancer cell growth. Together, the activity of MSI1 has been linked to numerous

cancers arising from neural and epithelial stem and progenitor cells. Removal or inhibition of MSI1 from these signaling pathways would stop the activation of genes promoting cellular proliferation. Therefore, it is interesting to investigate inhibitors of MSI1, which could lead to a novel drug for decreasing tumor cell proliferation.

Up-regulation of MSI1 in tumor growth

As cells differentiate, MSI1 expression is depleted (Sakakibara et al., 1996). Increased levels of MSI1 were first correlated with tumor cell malignancy and proliferative activity in human gliomas (Kanemura et al., 2001). Elevated levels of MSI1 have since been discovered in astrocytomas, retinoblastomas, medulloblastomas, breast cancer, and colon adenocarcinoma (Figure 1) (Kanemura et al., 2001; Sanchez-Diaz et al., 2008; Seigel et al., 2007; Sureban et al., 2008; Wang et al., 2010). Depletion of MSI1 in gliomas, medulloblastomas, breast cancer and colon adenocarcinomas results in decreased proliferation and increased cell death of cancerous tissue (Muto et al., 2012; Sanchez-Diaz et al., 2008; Sureban et al., 2008; Wang et al., 2010).

MSI1's involvement in different types of tumors is still being characterized. Two of the most thoroughly characterized cancers involving MSI1 are glial cell tumors and colorectal tumors (Kanemura et al., 2001; Muto et al., 2012). Thirty percent of all brain tumors are characterized as gliomas, arising from the supportive glial cells of the brain. As of 2012, approximately 206,429 people in the United States alone were living with a diagnosed glial brain tumor. Brain tumors are the second leading cause of cancer-related death in children under 20, and the second and fifth leading cause of cancer-related death in males and females over 20, respectively (American Brain Tumor Association, 2013). The number of newly diagnosed adults

with colorectal cancer in 2011 in the United States was 141,210. Also in 2011, 49,380 patients died of colorectal cancer in the United States alone (American Cancer Society, 2011).

Safer and more effective drug options are constantly being sought to inhibit tumor cell growth, many which target cellular growth and proliferation pathways. A small molecule inhibitor of MSI1 could alter these pathways and lead to decreased proliferation and increased death of cancerous tissue. The Ryder Lab in the Department of Biochemistry and Molecular Pharmacology at the University of Massachusetts Medical School has identified a small molecule metabolite, oleic acid, as an inhibitor of mouse MSI1 through a small molecule screen (Clingman et al., n.d.). Oleic acid is an omega-9 (ω -9) monounsaturated fatty acid that is absent in early development, and becomes enriched in postnatal human brains, an expression pattern opposite to MSI1 (Martínez & Mougan, 1998).

To investigate the conservation of RNA-binding affinity and inhibition of MSI1, a total of five MSI1 homologues from diverse species were cloned, expressed, purified and assayed. These proteins have varying percent similarities to human MSI1, with the greatest similarity present in the RNA-recognition motifs (RRMs) (Table 1). By studying the evolutionary conservation of binding and inhibition, this study aimed to provide insight into mechanistic properties of MSI1. It also aimed to investigate the possibility of different MSI1 homolog species that could serve as new models to explore the biological significance of MSI1 inhibition by oleic acid. For example, if fly Msi1 is inhibited by oleic acid *in vitro*, flies overexpressing Msi1 could be used as a model organism in which to test the effects of oleic acid on tumor cell growth.

Materials and Methods

Generation of expression constructs

Constructs containing genes for MSI1 homologs were created using restriction enzyme digestion followed by ligation. Restriction enzyme sites were inserted flanking the gene of interest by PCR using primers with restriction sites (Table 2). Following PCR, both target vector and gene of interest were digested with two restriction enzymes, followed by ligation of the gene of interest into the target vector. This created constructs that could be replicated in DH5 α *E. coli* and contained an inducible protein expression system, allowing protein expression in BL21 *E. coli*.

Insertion of restriction sites and amplification of insert using PCR

All inserts containing MSI1 homologs were generated from the commercially available cDNA clones listed in Table 2. To produce the truncated protein, primers (Table 2) were designed to insert restriction sites 42 base pairs upstream of RRM1 and 48 base pairs downstream of RRM2. Primer sets were designed to insert coding sequences in-frame. PCR reactions were done in 50 μ L [1x buffer (NEB), 10mM dNTPs, 100ng each of 3' and 5' primer, 40ng clone, 1 μ L PfuTurbo DNA polymerase (Stratagene)]. Product size was checked by gel electrophoresis and samples containing insert were purified using phenol/chloroform extraction followed by ethanol precipitation. Truncated expression constructs were designed with *D. rerio* msi1, *C. elegans* MSI-1, *D. melanogaster* Msi and *D. melanogaster* Rbp6. Mouse MSI1 protein containing only the two RRMs was provided by Carina Clingman. Constructs were abbreviated according to their scientific name, abbreviation and length (T=truncated). For example, *D. rerio* (zebrafish)

msi1 protein containing only the two RRM domains is abbreviated Dr_msi1_T. Proteins also follow this naming scheme.

Restriction digestion of inserts and vector

Restriction digestions of inserts and vector (pET-22HT) were done in 20 μ L using the appropriate enzyme-optimized 10x buffer (NEB), 500ng of DNA, 1 μ L of each restriction enzyme and 0.1 mg/mL BSA. For inserts and vectors with EcoR1 and Sal1 restriction sites, 10x EcoR1 buffer (NEB) was used. For inserts and vectors with Sac1 and Sal1 restriction sites, restriction digestion was done in two steps. Digestion with Sac1 was done in 10x buffer 1 (NEB). Reactions were then purified using phenol/chloroform extraction followed by ethanol precipitation. This was followed by digestion with Sal1 in 10x buffer 3 (NEB). Reactions were incubated at 37°C for one hour to allow for complete digestion. Product size was checked on an agarose gel.

The vector (pET-22HT) was subjected to restriction digestion with the same restriction enzyme pairs (EcoR1/Sal1 and Sal1/Sac1). Musashi inserts were ligated into the pET-22HT vector (Figure 5a) which contains a lac promoter, T7 polymerase promoter, lac operator, f1 origin of replication, ampicillin resistance gene, ColE1 origin of replication, and a polylinker containing restriction enzyme sites. This vector allowed for plasmid replication and expression of recombinant protein in different bacterial systems. The pET-22HT plasmid contains a 5'-histidine6-glycine tag (His6/Gly), and TEV protease site. The His6-tag is essential for protein purification.

In-gel ligation

In-gel ligation was used to avoid extracting DNA from gels. Restriction digested products were run on 1% low melting point agarose gels and bands were cut from the gel. In-gel ligation

reactions were done in 20 μ L [3 μ L of vector gel slice, 6 μ L of insert gel slice, 2 μ L of T4 DNA ligase (NEB) and 1x buffer with ATP (NEB)] and incubated overnight at room temperature (RT). To check for insertion of the Musashi1 gene homolog into pET-22HT, 100 μ L of *E. coli* (DH5 α) were transformed with 5 μ L of in-gel ligation product using the following steps. After gentle mixing, cells with in-gel ligation product were set on ice for 10 minutes; heat shocked for 1 minute and 45 seconds at 37°C, and then put on ice. Cells were added to 1mL of LB and shaken at 37°C for 1 hour to allow expression of the Amp^R gene. 200 μ L of transformed cells were plated on LB agar plates with 100 μ g/mL ampicillin and stored at 37°C overnight for growth. Single colonies were picked and grown overnight in 4mL of LB/Amp at 37°C to allow replication of the recombinant plasmid, which was isolated from cultures using a QIAprep Spin Miniprep Kit (Qiagen). Products were subjected to restriction digestion with restriction enzyme sets previously stated and analyzed by electrophoresis. Products were also sent to Elim Biopharmaceuticals for sequencing. Samples containing both vector and insert of the correct size without mutations were transformed into *E. coli* BL21 (DE3) cells for protein expression.

Transformation of *E. coli* BL21 cells, growth and induction

Test inductions

Rubidium chloride competent *E. coli* strain BL21 (DE3) cells were transformed with successful in-gel ligation products as described above for DH5 α *E. coli*. Test inductions were performed to confirm ability to induce protein expression and protein solubility. One colony was picked and grown in 5mL of LB with 100 μ g/mL ampicillin for one hour with shaking at 37°C. 1mL of starter culture was added to 50mL of LB with 100 μ g/mL ampicillin and grown to an OD of 0.5. Protein expression was induced with isopropyl- β -D-thiogalactoside (IPTG) at a final concentration of

1mM. IPTG displaces the lac repressor from the lac operator, first allowing expression of T7 polymerase from the host genome. T7 polymerase then can bind the T7 promoter on pET-22HT and transcribe the gene inserted into the polylinker site. Shaking at 37°C for an additional 3 hours allowed protein expression. Cells were pelleted by centrifugation at 5000xg for 30 minutes and stored at -20°C until use. Cells were lysed in Sample Buffer [100mM Tris-Cl (pH 6.8), 4% SDS, 0.2% bromophenol blue, 20% glycerol, 200mM β -mercaptoethanol (BME)]. Samples were run on an SDS-page gel to determine ability to induce protein expression and to determine solubility.

Protein expression from constructs

To prepare for expression of protein, one colony of BL21 *E. coli* containing the expression construct was picked and grown in a starter culture of 5mL of LB with 100 μ g/mL ampicillin for one hour with shaking at 37°C. 1mL of starter culture was added to 50mL of LB with 100 μ g/mL ampicillin and grown overnight with shaking at 37°C. 5mL were transferred to each of 2 1L flasks of LB with 100 μ g/mL ampicillin and grown at 37°C with shaking until cells reached the mid-log growth phase, determined by an OD reading between 0.6-0.8 at 600nm. Cultures were induced with 1mM IPTG and shaken at 37°C for 3 hours. Cells were pelleted as above and dry pellet was stored at -80°C until use. Samples from pre- and post-induction were saved to be run on a later SDS-page gel.

Protein purification

Recombinant MSI1 homologs were expressed and purified as previously described for His6-tagged Mouse MSI1 (Clingman et al., in review). Purification consisted of the three affinity columns described below.

Cell lysis and Ni-NTA column purification

Pelleted cells containing the recombinant MSI1 homolog proteins were thawed and resuspended in lysis buffer [300mM NaCl, 50mM NaH₂PO₄, 20mM Imidazole, 5mM β-mercaptoethanol (BME), cOmplete Mini EDTA-free protease inhibitor cocktail tablet (Roche)]. Cells were lysed using a microfluidizer. After lysis, soluble lysate was bound to a nickel-nitrilotriacetic acid (Ni-NTA) column (Thermo Scientific). This column selectively binds the His6-tag on protein. The column was washed with wash buffer (300mM NaCl, 50mM NaH₂PO₄, 50mM Imidazole, 5mM BME), and protein was eluted in 5mL fractions with elution buffer (300mM NaCl, 50mM NaH₂PO₄, 300mM Imidazole, 5mM BME). Samples were spotted on to Whatman paper and stained with Coomassie Brilliant Blue dye to determine the fractions containing protein (What-blot). Samples were further analyzed on an SDS-page gel to check the size and purity of the protein. Protein-containing fractions were combined and dialyzed overnight into MOPS buffer (pH 6.0) for a HiTrap SP cation exchange column (50mM MOPS pH 6.0, 20mM NaCl, 2mM DTT; GE Healthcare).

HiTrap SP cation exchange column

Dialyzed protein was bound to a HiTrap SP cation exchange column (S-column; GE Healthcare). The column was washed with low salt S buffer (50mM MOPS pH 6.0, 20mM NaCl, 2mM DTT). Protein was eluted by washing with an increasing concentration of salt over 2 hours in 8mL fractions by mixing low salt S buffer with increasing amounts of high salt S buffer (50mM MOPS pH 6.0, 2M NaCl, 2mM DTT). Presence of protein in fractions was determined using a What-blot. Fractions containing protein were combined and dialyzed overnight into buffer for a HiTrap Q anion exchange column (50mM Tris pH 8.8, 20mM NaCl, 2mM DTT; GE Healthcare).

HiTrap QP anion exchange column and concentration

The HiTrap QP column was run the same as the S column, using low salt Q buffer (50mM Tris pH 8.8, 20mM NaCl, 2mM DTT) and high salt Q buffer (50mM Tris pH 8.8, 2M NaCl, 2mM DTT).

Location of protein in fractions was determined by spotting samples onto Whatman paper and staining with Coomassie Brilliant Blue dye. Protein was dialyzed overnight into storage buffer (50mM Tris pH 8.0, 20mM NaCl, 2mM DTT). All saved samples were run on a 12% SDS-page.

Proteins were concentrated using VIVASPIN 20 10,000 MW cut-off spin concentrators (Sartorius Stedim Biotech). Concentration (c) was determined using Beer's law ($A_{280} = \epsilon \cdot l \cdot c$). Absorbance at 280nm was read using a Cary 50 Bio UV-Visible Spectrophotometer (Varian) with a path length (l) of 1cm. Extinction coefficients (ϵ) of recombinant proteins were calculated using the ProtParam tool available online from ExPASy. Purified proteins were stored at 4°C until use.

Determination of binding constants via direct titration assays

Direct titration to determine dissociation constant

Fluorescence polarization

Fluorescence polarization (FP) was used to quantify the binding affinity of the four recombinant Musashi1 proteins to RNA aptamers. This assay shines polarized light into samples and reads sample emission. Free RNA tumbles quickly in solution and emits depolarized light. RNA bound to protein tumbles more slowly and emits polarized light. RNA used for FP assays contained the MSI1 consensus binding site. The 12nt RNA aptamer "nrzr005" (UUUAUAGUUUUU-FI) contains a 3' fluorescein tag and was ordered pre-labeled from IDT and obtained from Dr. N. Ruth Zearfoss. This RNA aptimer contains one MSI1 binding site (underlined).

The FP assay was conducted as previously described (Pagano, Clingman, & Ryder, 2011). Assay conditions were modified following suggestions from Dr. Phil Zamore. Original assay conditions

used 1x EMSA buffer (50mM Tris pH 8.0, 100mM NaCl, 0.01mg/mL tRNA, 0.01% Igepal CA-630) to dilute RNA to a 2nM final concentration, while protein dilution series were made with water (final high protein concentration 2 μ M). Modified assay conditions used 1x EMSA buffer plus bovine gamma globulin (BGG; 0.1mg/mL BGG) to dilute RNA and protein to the same concentrations. Protein samples consisted of 23 points of a two-thirds dilution series plus a control sample without protein. The reactions were incubated covered in a 96-well plate for 2 hours. All experiments were done in triplicate.

Samples were read for fluorescence polarization using a Victor 3 plate reader (Perkin Elmer; emission filter: 535 \pm 40, excitation filter: 480 \pm 31, measurement time: 1.0s). Each well was read 5 times. The data was analyzed by fitting the resulting polarization versus concentration graph to the Hill equation (equation 1). Equation 1 includes fluorescence polarization (ϕ), base (b), max (m) and high protein concentration (P_t). This equation is used to determine the dissociation constant (K_d) and the Hill coefficient (n , proportional to the slope of the sigmoid curve) for RNA aptimers containing one MSI1 binding site.

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

Fluorescence-electrophoretic mobility shift assay

Samples from the FP assay were analyzed by fluorescence electrophoretic mobility shift assay (F-EMSA) after FP was read. Bromocresol green loading dye (1x concentration: 6% glycerol, 0.03% bromocresol green) was added to each sample. Samples were run on a 5% native polyacrylamide gel for 75 minutes at 120V in chilled 1x TBE buffer. Gels were imaged with a Typhoon FLA 9000 Biomolecular imager (GE Healthcare) (laser: 473nm, long-pass cut-off filter:

510nm). ImageGuage software (Fujifilm) was used to determine the fraction of RNA bound to protein. The intensity of the upper (bound) and lower (free) bands were quantified relative to background to determine the fraction of recombinant protein bound to RNA. Data were graphed as fraction bound versus protein concentration. Curves were fit to the Hill equation as previously described using equation 1 where ϕ is fraction bound (Ryder, Recht, & Williamson, 2008).

Stoichiometric binding assay

To confirm the stoichiometry of protein-RNA binding, a two-thirds dilution series of protein was made in 1x EMSA buffer with 15 points, plus a negative control without protein. The high protein concentration was 24 μ M. RNA was added to protein with a final concentration of 4 μ M unlabeled RNA and 2nM fluorescein-labeled RNA. All dilutions were made with 1x EMSA buffer. Protein-RNA mixtures were equilibrated for 3 hours at room temperature and quantified by FP and F-EMSA as previously described. Data was graphed as fluorescence polarization or fraction bound versus molar equivalents (RNA:MSI1). Data was fit to both a piecewise function, which fits data to two straight lines and finds the intersection point, and to equation 2, which describes protein-RNA complex formation (Rambo & Doudna, 2004). ϕ is fluorescence polarization or fraction bound, r is the ratio RNA:protein, and n is the stoichiometric equivalence point.

$$\phi = \frac{r+K_d+n-\sqrt{(r+K_d+n)^2-4rn}}{2n} \quad (2)$$

Dose response assays with omega-9 fatty acids

Dose response assays were used to determine the inhibition of RNA-binding by MSI1 homologs with omega-9 (ω -9) monounsaturated fatty acids, as previously discovered by the Ryder Lab in

a small molecule screen with mouse MSI1. Experiments were performed using a modified FP and F-EMSA protocol. A two thirds dilution series of compound (oleic or eliadic acid) was prepared in DMSO, and then brought to a final high concentration of 384 μ M of compound using 1x EMSA buffer without BGG. The concentration at which the protein was 80% bound in the FP direct titration assay was used and held constant in all wells. The final concentration of fluorescein-labeled RNA was 2nM in all wells. The mixture equilibrated at room temperature for 1 hour. FP and F-EMSA data were collected as described above in direct titration assays. The half maximal inhibitory concentration (IC_{50}) was calculated using equation 3, including polarization or fraction bound (ϕ), base (b), max (m), protein concentration (P) and Hill coefficient (n).

$$\phi = b + (m - b) \left(\frac{1}{1 + \left(\frac{IC_{50}}{[P]} \right)^n} \right) \quad (3)$$

The Lin and Riggs conversion was used to correct for the K_d of the specific Musashi1 protein and for the concentration of RNA and protein used in the experiment. Equation 4 was used to calculate the apparent inhibition constant ($K_{i,app}$) and includes the protein-specific dissociation constant (K_d), concentration of protein (P) and concentration of RNA (R) (Lin & Riggs, 1972; Ryder & Williamson, 2004)).

$$K_{i,app} = \frac{2(K_d)(IC_{50})}{2P - R - 2K_d} \quad (4)$$

Results

Cloning of MSI1 homologs

To investigate the binding and inhibition of MSI1 homologs, recombinant proteins first had to be cloned, expressed and purified. To clone MSI1 homolog genes into the expression vector, the primers listed in Table 2 were used to add restriction sites flanking the two RRMs. Next, both the MSI1 homolog gene and the target vector (pET-22HT) were cut with restriction enzymes and MSI1 homolog genes were successfully ligated into the target vector (data not shown). Figure 5a shows details about the pET-22HT vector used. This vector allowed for the addition of an N-terminal His6-tag (Figure 5b) to the MSI1 homolog protein, which was essential during protein purification. Figure 5c shows that overall cloning scheme used for generation of expression constructs.

Purification of recombinant Musashi homologs

After cloning, MSI1 homolog protein was successfully expressed through induction with IPTG, which used the lac repressible system shown on the pET-22HT plasmid in Figure 5a. MSI1 homologs were then purified away from all cellular components for use in assays. Expression levels and purification efficiency varied between proteins. Truncated proteins containing the two RRMs were all soluble, and were successfully expressed and purified using three affinity columns.

Location of protein during purification was monitored closely, as the four MSI1 homologs had not been previously purified by the Ryder Lab. Figure 6 shows an example SDS-PAGE gel with samples of truncated His6-tagged zebrafish *msi1* (Dr_ *msi1*_T, 22 kD) throughout stages of purification. In Figure 6a, the presence of protein was detected three hours post-induction with

IPTG. The Ni-NTA spin column bound the His6-tag of the recombinant protein, but loss of about 50% of protein in the flow-through could indicate that some protein did not contain a His6-tag or that the column did not have a large enough capacity to bind all recombinant protein.

After Ni-NTA column purification, protein was run over an S cation exchange column to remove contaminating proteins. In Figure 6b, the presence of Dr_msi1_T after dialysis against MOPS buffer (pH 6.0) is seen, which is the buffer necessary for use with the S column. This column is negatively charged at pH 6.0. Proteins were expected to be positively charged at pH 6.0 based on their theoretical isoelectric point (pI), causing them to stick to the column until eluted with a high salt buffer. Protein eluted from the S column in fractions 8-14, and was undetectable in the wash and flow-through when spotted onto Whatman paper and stained with Coomassie Brilliant Blue dye (not shown). The S column successfully removed the 15 kD band present in fractions 1 and 2 of the Ni-NTA spin column.

The final column used was a Q anion exchange column, which removed RNA bound to protein and free RNA. After the S column, the protein was then dialyzed against Tris pH 8.8 buffer in preparation for purification with the Q column, as shown in Figure 6b. The Q column is positively charged at pH 8.8, and the MSI1 homologs were calculated to be negatively charged at pH 8.8 based on their predicted pI. Protein eluted from the Q column in fractions 6-11 with the addition of a high salt buffer. A large amount of protein was also observed in the flow-through; therefore, both flow-through and fractions 6-11 from Dr_msi1_T purification were saved and tested for activity.

Due to differences in predicted pIs of the MSI1 homologs, proteins varied in their affinity for the different ion exchange columns. Ce_MSI-1_T and Dm_Rbp6_T bound to both the S and Q columns and were able to be eluted in fractions, which contained active protein. *Drosophila* Msi and zebrafish MSI1 bound to the S column, but did not bind tightly to the positively charged Q column. This is due to the high predicted pI of the two proteins calculated using ProtParam available online from ExPASy. The pH of Q column buffers could ideally be adjusted to give the proteins a negative overall charge. The pH must be above the pI of the protein for this to be achieved. However, the Q column buffers were already pH 8.8 and could not be raised without danger of denaturing the protein in the basic conditions. Therefore both fractions and flow-through from the Q column were tested by FP and F-EMSA to determine the active fraction, data from which will be shown later.

Recombinant proteins bind RNA aptimer with varying affinity

It was necessary to determine the affinity of the MSI1 homologs for their target RNA by examining the apparent dissociation constants of the interactions (K_d). This was done using an assay called direct titration in which protein is titrated into a constant concentration of RNA, which is present in trace amounts. Samples from this assay were analyzed in two different ways. Fluorescence polarization (FP) and fluorescence-electrophoretic mobility shift assay (F-EMSA) were used and are explained in detail in Figure 7. FP quantifies the amount of free RNA in solution by reading the polarization of emitted light. F-EMSA quantifies the fraction of bound RNA based on the principal that RNA bound to protein will migrate slower through a native gel matrix than free RNA. Both methods yielded a value for the dissociation constant (K_d) for RNA-

binding of protein. Before collecting data using these methods, assay conditions had to be optimized.

Assay optimization for Hill coefficient reduction

It was observed that using original assay conditions yielded a steep transition on the binding curves for both FP and F-EMSA, like the plot of Dr_msi1_T in Figure 8a. A steep transition is quantified by a high Hill coefficient. This steep transition suggested that at least one of two events was occurring: 1) cooperative binding, or 2) loss of protein at low concentrations. A stoichiometric binding assay was used to investigate the first possible event.

The Hill coefficient (n) gives a value for the steepness of the binding curve. This coefficient indicates the number of ligand binding sites present on the protein. Hill coefficients above 1 indicate binding cooperativity, where the protein consists of more than one binding site and binding at the first site facilitates binding at the second (Weiss, 1997). MSI1 has been previously shown to contain only one RNA-binding site. With original assay conditions, a Hill coefficient of 2 or greater was obtained for all MSI1 homologs, which is contradictory to previous data for mouse MSI1. To determine the stoichiometry of MSI1 binding to RNA, a stoichiometric binding assay was used. Data for Mm_MSI1_T indicate a 1:1 ratio by both FP and F-EMSA when fit to a piecewise function and equation 2 describing RNA-protein complex formation (Figure 9). This lead to the investigation of the effect of assay conditions on the Hill coefficient.

BGG prevents protein loss at low concentrations

Loss of protein at low concentrations due to protein binding to the assay plate was suggested by Dr. Phil Zamore of UMass Medical School. This event would shift the sigmoid binding curve to the right at low protein concentrations, increasing the curve steepness and consequently

increasing the Hill coefficient. The addition of the protein BGG to buffer reduced the Hill coefficient in all proteins tested. For example, Dr_msi1_T experienced a decrease from $n = 2.01$ to $n = 1.17$ for one FP replicate after the addition of BGG to buffer (Figure 8a, b). It is thought that BGG acts to coat the plate wells and prevent loss of MSI1 homologs at low protein concentrations. After optimization of assay conditions, all four MSI1 homologs, plus mouse MSI1, were tested for RNA-binding activity and inhibition of activity by ω -9 monounsaturated fatty acids. The binding properties of all proteins were tested with a 12 nucleotide fluorescein-labeled RNA containing one MSI1 binding site (5'-UUUAUAGUUUUU-3'). Direct titration was performed first, followed by dose response, and both were analyzed by FP and F-EMSA.

The dissociation constants were found in order to be able to compare the RNA-binding properties of the different MSI1 homologs. An example of an active protein binding to RNA is shown in Figure 8b with Dr_msi1_T collected from the flow-through of the Q column. The K_d for this replicate was 31.1nM. The FP data for Dr_msi1_T protein that eluted in fractions 6-11 of the Q column is shown in Figure 8c. This protein showed weaker binding (61.4nM) and may have been insoluble or aggregated at high concentrations, resulting in a decrease in RNA binding signified by a decrease in polarization of emitted light. Dr_msi1_T protein from the Q column flow-through was therefore used for all subsequent experiments. Data from all direct titration FP experiments done with the addition of BGG can be seen in Table 3.

F-EMSA was also used to calculate the K_d of recombinant proteins based on quantification of bound and free RNA. F-EMSA results can be seen in Table 3. The K_d found by F-EMSA is generally weaker (larger) than the K_d found by FP. This is most likely due to dissociation of the

RNA-protein complex during the running of the gel. Figure 10a shows a F-EMSA gel done with truncated *C. elegans* MSI-1 (Ce_MSI-1_T). The amount of bound RNA increases with increasing protein concentration. Data was analyzed as above and fit to the Hill equation (equation 1). This yielded a value for the K_d and a Hill coefficient. K_d values in Table 3 are the means of three independent replicates with standard deviations. Dm_Msi_T ($K_{d, FP} = 360 \pm 12.8\text{nM}$) showed 14-fold weaker binding to the one-site RNA aptimer than Mm_MSI1_T ($K_{d, F-EMSA} = 26.0 \pm 1.06\text{nM}$) when tested by FP. Dm_Msi_T ($K_{d, F-EMSA} = 313 \pm 13.0\text{nM}$) showed 10-fold weaker binding than Mm_MSI1_T ($K_{d, F-EMSA} = 28.4 \pm 5.30\text{nM}$) when tested by F-EMSA. Dm_Rbp6_T ($K_{d, FP} = 20.8 \pm 1.36\text{nM}$), Mm_MSI1_T ($K_{d, FP} = 26.0 \pm 1.06\text{nM}$) and Dr_msi1_T ($K_{d, FP} = 26.2 \pm 1.42\text{nM}$) all showed similar K_d values of around 26nM when tested by FP. The Ce_MSI-1_T ($K_{d, FP} = 61.7 \pm 4.24\text{nM}$) showed 2-fold weaker binding by FP and 3-fold weaker binding by F-EMSA than Mm_MSI1_T. Data from direct titration assays was then used to determine the conditions for dose response assays.

Binding of MSI1 homologs to RNA is inhibited by oleic acid

Dose response assays were performed to determine the strength of oleic acid inhibition of RNA-binding in order to compare the inhibition of all MSI1 homologs. Direct titration data was used to determine the concentration of a specific protein at which it was 80% bound to RNA, and this concentration was held constant in dose response assays. RNA concentration was also held constant, while fatty acid was titrated into samples.

Dose response experiments were analyzed similarly to direct titration experiments using FP and F-EMSA data with a modified Hill equation (equation 3). Hill equations for dose response experiments included a term for the half maximal inhibitory concentration (IC_{50}), which is the

concentration of inhibitor required for 50% inhibition of RNA-protein complex formation. Figure 11a shows the FP data with a constant concentration of Dm_Rbp6_T and one-site RNA, and titrated oleic acid. Oleic acid inhibited binding of Rbp6 to RNA with an IC_{50} of $9.32\mu\text{M}$ in this specific replicate. The apparent inhibition coefficient ($K_{i, \text{app}}$) for Rbp6 was $5.08\mu\text{M}$, which takes into consideration the previously measured K_d of the specific protein using the Lin and Riggs conversion (equation 4). Quantification of the gel in Figure 11b is shown in Figure 11c. An increase in curve steepness is apparent between FP and F-EMSA data. Table 4 summarizes all dose response data from both FP and F-EMSA experiments. Dm_Msi_T ($IC_{50, \text{FP}} = 29.9 \pm 1.68\mu\text{M}$; $IC_{50, \text{F-EMSA}} = 19.2 \pm 1.24\mu\text{M}$) had a 2-fold weaker mean IC_{50} compared to Mm_MSI1_T ($IC_{50, \text{FP}} = 16.0 \pm 0.94\mu\text{M}$; $IC_{50, \text{F-EMSA}} = 7.19 \pm 0.25\mu\text{M}$) by both FP and F-EMSA. A 3-fold weaker mean $K_{i, \text{app}}$ for Dm_Msi_T ($K_{i, \text{app, FP}} = 33.6 \pm 1.88\mu\text{M}$) was seen compared to Mm_MSI1_T ($K_{i, \text{app, FP}} = 9.69 \pm 0.57\mu\text{M}$) by FP. A 5-fold weaker $K_{i, \text{app}}$ was seen in Dm_Msi_T ($K_{i, \text{app, F-EMSA}} = 21.5 \pm 1.39\mu\text{M}$) than Mm_MSI1_T ($K_{i, \text{app, F-EMSA}} = 4.35 \pm 0.15\mu\text{M}$) by F-EMSA. No significant differences were seen between other homologs. This data suggests differences in the effectiveness of oleic acid inhibition of RNA-binding specifically with Dm_Msi_T.

All MSI1 homologs were also tested with eliadic acid, another ω -9 monounsaturated fatty acid. Eliadic acid has a *trans* double bond, and oleic acid has a *cis* double bond. Figure 11d shows the gel from Dm_Rbp6 with titrated eliadic acid. No inhibition of RNA-binding activity is seen. Interestingly Ce_MSI-1_T was inhibited by both oleic acid ($K_{i, \text{app, F-EMSA}} = 3.82 \pm 0.14\mu\text{M}$) and eliadic acid ($K_{i, \text{app, F-EMSA}} = 40.9 \pm 14.9\mu\text{M}$) by F-EMSA (Figure 11e, f). However, inhibition of Ce_MSI-1_T by oleic acid was not seen by FP. This 10-fold difference in inhibition by eliadic acid could be tested further to determine if it is biologically relevant.

Evolutionary conservation of MSI1 homologs

After seeing differences in RNA-binding and inhibition of the MSI1 homologs, protein sequences were aligned and examined to find differences in amino acid sequence which could contribute to these differences in activity. Figure 12 shows the sequence alignment for mouse MSI1 and the four MSI1 homologs. Amino acids known to participate in aromatic stacking interactions with RNA were written in green (Ohyama et al., 2012). The residues in purple represent amino acids known to participate in hydrogen-bonding with RNA (Ohyama et al., 2012). All of these residues are highly conserved within RRM1, which is the RNA-recognition motif that interacts with high affinity to RNAs. Two amino acid differences were seen in Dm_Msi_T, which are boxed in red. Dm_Msi_T contains a histidine instead of a valine (V94H) which can no longer participate in hydrogen bonding with RNA. Dm_Msi_T also contains a threonine instead of a phenylalanine (F96T), which cannot participate in aromatic stacking interactions. These differences are hypothesized to cause variability in RNA affinity and may suggest different mRNA targets for Dm_Msi_T *in vivo*.

The Ryder Lab has predicted five amino acids to be involved in oleic acid inhibition of mouse MSI1. These residues are highlighted in yellow in Figure 12. Arginine 61, glycine 64 and phenylalanines 63 and 65 are seen to be conserved in all MSI1 homologs. In both fly MSI1 homologs (Msi and Rbp6), arginine 53 is changed to lysine (R53K), but this does not correlate with observed differences in inhibition. Arginine and lysine are both positively charged amines and may have similar function in MSI1 homologs. Since all MSI1 homologs showed inhibition by oleic acid, it will be possible to use these homologs for further study of this activity *in vivo*.

Discussion

The RNA-binding protein MSI1 is overexpressed in cancers arising from neural and epithelial stem and progenitor cells. It has been shown that knocking down MSI1 by RNAi causes tumor cell regression. Alternatively, it is hypothesized that an inhibitor of MSI1 could have the same effect on tumor cells. An inhibitor of MSI1 could provide a future treatment to inhibit tumor cell proliferation and lead to tumor regression. The Ryder Lab at the University of Massachusetts Medical School has found oleic acid to be a strong inhibitor of mouse MSI1 *in vitro*. This study has shown that MSI1 homologs from diverse species are also inhibited by oleic acid *in vitro*. It was also shown that MSI1 homologs have different RNA-binding and inhibition constants.

Differences in RNA-binding of MSI1 homologs were seen and highlight the importance of specific amino acids. Fly Msi exhibited the weakest RNA-binding activity and includes the amino acid differences V94H and F96T when compared to the other MSI1 homologs tested (Figure 12). All other residues known to be involved in RNA-binding were conserved in fly Msi. The variations observed in the RNA-binding affinity, specifically in fly Msi, suggest that mRNA targets may differ. To investigate the possibility of increased affinity of fly Msi for different RNA sequences, Dr. Ruthie Zearfoss of the Ryder Lab has recently tested fly Msi's affinity for slightly different RNA sequences. This has shown that fly Msi exhibits higher affinity for an RNA aptamer of different sequence, suggesting that fly Msi may have different mRNA targets (unpublished data). It would be interesting to further investigate this difference. Experiments could be performed to investigate the effects of amino acid sequence on the specificity of RNA binding in MSI1 homologs. Mutation of the amino acids H94 and T96 in fly Msi to the

evolutionarily conserved valine and phenylalanine residues, respectively, could be done to determine if binding to the specific RNA used in this study is improved.

The observed conservation of oleic acid inhibition across the diverse species of MSI1 homologs tested may provide new models to explore its biological significance in the future. Conservation of amino acids (R61, F63, G64 and F65) across the five MSI1 homologs tested strengthens the hypothesis proposed by the Ryder Lab about their involvement in RNA-binding inhibition.

Further study into the mechanism of inhibition of MSI1 by oleic acid can be directed towards these amino acids. To determine the importance of the individual residues, mutant proteins of any of the MSI1 homologs tested could be created in which these residues are mutated to functionally different residues. Significantly different inhibition constant values would indicate the specific residue's importance in recognition and binding of oleic acid. This could help to elucidate the mechanism of oleic acid inhibition of MSI1.

The next step for assessing the significance of oleic acid inhibition of MSI1 RNA-binding activity is to investigate the effect of oleic acid *in vivo*. Since all MSI1 homologs tested exhibited inhibition by oleic acid, a model organism expressing any of these homologs of MSI1 could be created to test the effect of oleic acid. For example, zebrafish could be used as a model organism. If tumor cell development could be induced in neural or epithelial cells in the zebrafish, then the tumors could be treated with oleic acid to investigate oleic acid's ability to inhibit tumor cell proliferation. Alternatively, cell lines expressing one of the MSI1 homologs could be cultured and treated with oleic acid, and proliferation rates could be monitored. This experiment could be done in *D. melanogaster* Schneider 2 (S2) cells, which express fly Msi1.

Proliferation rates could be compared between treatment with oleic acid, elaidic acid and RNAi Msi knockdown in both cell lines and animal models. Additionally, Western blots to determine the expression of proteins like Numb could be performed on cell extracts before and after treatment with oleic acid, elaidic acid and RNAi. MSI1 would normally inhibit the translation of *numb* mRNA into Numb protein. If oleic acid is inhibiting MSI1 *in vivo*, proteins like Numb would still be expressed.

Further investigation into the inhibition of tumor cell proliferation with oleic acid in cell lines or animal models expressing MSI1 homologs from this study could yield a treatment for cellular proliferation in tumors. MSI1 has been shown to be upregulated in many cancers arising from neural and epithelial stem and progenitor cells. Therefore, inhibition of MSI1 with oleic acid could be a potential treatment for any cancer arising from these tissues. The preservation of oleic acid inhibition of RNA-binding activity across evolutionarily diverse MSI1 homologs discovered in this study provides a foundation for future research into the biological relevance of this conserved property.

Works Cited

- Battelli, C., Nikopoulos, G. N., Mitchell, J. G., & Verdi, J. M. (2006). The RNA-binding protein Musashi-1 regulates neural development through the translational repression of p21WAF-1. *Molecular and cellular neurosciences*, *31*(1), 85–96. doi:10.1016/j.mcn.2005.09.003
- Clingman, C. C., Deveau, L. M., Hay, S. A., Genga, R. M., Shandilya, M. D., Massi, F., & Ryder, S. P. (n.d.). Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite.
- Colaluca, I. N., Tosoni, D., Nuciforo, P., Senic-Matuglia, F., Galimberti, V., Viale, G., Pece, S., et al. (2008). NUMB controls p53 tumour suppressor activity. *Nature*, *451*(7174), 76–80. doi:10.1038/nature06412
- De Andrés-Aguayo, L., Varas, F., Kallin, E. M., Infante, J. F., Wurst, W., Floss, T., & Graf, T. (2011). Musashi 2 is a regulator of the HSC compartment identified by a retroviral insertion screen and knockout mice. *Blood*, *118*(3), 554–64. doi:10.1182/blood-2010-12-322081
- Di Marcotullio, L., Ferretti, E., Greco, A., De Smaele, E., Po, A., Sico, M. A., Alimandi, M., et al. (2006). Numb is a suppressor of Hedgehog signalling and targets Gli1 for Itch-dependent ubiquitination. *Nature cell biology*, *8*(12), 1415–23. doi:10.1038/ncb1510
- Dobson, N. R., Zhou, Y., Flint, N. C., & Armstrong, R. C. (2009). Musashi1 RNA-binding protein regulates oligodendrocyte lineage cell differentiation and survival. *Glia*, *56*(3), 318–330. doi:10.1002/glia.20615.Musashi1
- Glisovic, T., Bachorik, J. L., Yong, J., & Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. *FEBS letters*, *582*(14), 1977–86. doi:10.1016/j.febslet.2008.03.004
- Good, P., Yoda, a, Sakakibara, S., Yamamoto, a, Imai, T., Sawa, H., Ikeuchi, T., et al. (1998). The human Musashi homolog 1 (MSI1) gene encoding the homologue of Musashi/Nrp-1, a neural RNA-binding protein putatively expressed in CNS stem cells and neural progenitor cells. *Genomics*, *52*(3), 382–4. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9790759>
- Imai, T., Tokunaga, A., Yoshida, T., Hashimoto, M., Mikoshiba, K., Weinmaster, G., & Nakafuku, M. (2001). The Neural RNA-Binding Protein Musashi1 Translationally Regulates Mammalian numb Gene Expression by Interacting with Its mRNA. *Molecular Cell Biology*, *21*(12), 3888–3900. doi:10.1128/MCB.21.12.3888
- Kaneko, Y., Sakakibara, S., Imai, T., Suzuki, A., Nakamura, Y., Sawamoto, K., Ogawa, Y., et al. (1999). Musashi1: an evolutionally conserved marker for CNS progenitor cells including neural stem cells. *Developmental neuroscience*, *22*, 139–53. doi:17435

- Kanemura, Y., Mori, K., Sakakibara, S., Fujikawa, H., Hayashi, H., Nakano, A., Matsumoto, T., et al. (2001). Musashi1, an evolutionarily conserved neural RNA-binding protein, is a versatile marker of human glioma cells in determining their cellular origin, malignancy, and proliferative activity. *Differentiation*, 68, 141–52. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11686236>
- Kawahara, H., Imai, T., Imataka, H., Tsujimoto, M., Matsumoto, K., & Okano, H. (2008). Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP. *The Journal of cell biology*, 181(4), 639–53. doi:10.1083/jcb.200708004
- Lin, S.-Y., & Riggs, A. D. (1972). Zac Repressor binding to Non-operator DNA : Detailed Studies and a Comparison of Equilibrium and Rate Competition Methods. *Journal of Molecular Biology*, 72, 671–690.
- Maris, C., Dominguez, C., & Allain, F. H.-T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS journal*, 272(9), 2118–31. doi:10.1111/j.1742-4658.2005.04653.x
- Martínez, M., & Mougan, I. (1998). Fatty acid composition of human brain phospholipids during normal development. *Journal of neurochemistry*, 71(6), 2528–33. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9832152>
- Miyanoiri, Y., Kobayashi, H., Imai, T., Watanabe, M., Nagata, T., Uesugi, S., Okano, H., et al. (2003). Origin of higher affinity to RNA of the N-terminal RNA-binding domain than that of the C-terminal one of a mouse neural protein, musashi1, as revealed by comparison of their structures, modes of interaction, surface electrostatic potentials, and backbone . *The Journal of biological chemistry*, 278(42), 41309–15. doi:10.1074/jbc.M306210200
- Muto, J., Imai, T., Ogawa, D., Nishimoto, Y., Okada, Y., Mabuchi, Y., Kawase, T., et al. (2012). RNA-binding protein Musashi1 modulates glioma cell growth through the post-transcriptional regulation of Notch and PI3 kinase/Akt signaling pathways. *PLoS One*, 7(3), e33431. doi:10.1371/journal.pone.0033431
- Nagata, T., Kanno, R., Kurihara, Y., Uesugi, S., Imai, T., Sakakibara, S., Okano, H., et al. (1999). Structure , Backbone Dynamics and Interactions with RNA of the C-terminal RNA-binding Domain of a Mouse Neural RNA-binding Protein , Musashi1. *Journal of Molecular Biology*, 287, 315–330.
- Ohyama, T., Nagata, T., Tsuda, K., Kobayashi, N., Imai, T., Okano, H., Yamazaki, T., et al. (2012). Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. *Nucleic acids research*, 40(7), 3218–31. doi:10.1093/nar/gkr1139
- Okabe, M., Imai, T., Kurusu, M., Hiromi, Y., & Okano, H. (2001). Translational repression determines a neuronal potential in Drosophila asymmetric cell division. *Nature*, 411(6833), 94–8. doi:10.1038/35075094

- Okano, H., Imai, T., & Okabe, M. (2002). Musashi: a translational regulator of cell fate. *Journal of cell science*, *115*(7), 1355–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11896183>
- Pagano, J. M., Clingman, C. C., & Ryder, S. P. (2011). Quantitative approaches to monitor protein – nucleic acid interactions using fluorescent probes. *RNA*, *17*(1), 14–20. doi:10.1261/rna.2428111.readout
- Rambo, R. P., & Doudna, J. a. (2004). Assembly of an active group II intron-maturase complex by protein dimerization. *Biochemistry*, *43*(21), 6486–97. doi:10.1021/bi049912u
- Ryder, S. P., Recht, M. I., & Williamson, J. R. (2008). Quantitative analysis of protein-RNA interactions by gel mobility shift. *Methods in molecular biology (Clifton, N.J.)*, *488*, 99–115. doi:10.1007/978-1-60327-475-3_7
- Ryder, S. P., & Williamson, J. R. (2004). Specificity of the STAR / GSG domain protein Qk1 : Implications for the regulation of myelination. *RNA*, *10*(9), 1449–1458. doi:10.1261/rna.7780504.cross
- Sakakibara, S., Imai, T., Hamaguchi, K., Okabe, M., Aruga, J., Nakajima, K., Yasutomi, D., et al. (1996). Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. *Developmental biology*, *176*(2), 230–42. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8660864>
- Sakakibara, S., Nakamura, Y., Satoh, H., & Okano, H. (2001). Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. *The Journal of neuroscience*, *21*(20), 8091–107. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11588182>
- Sakakibara, S., & Okano, H. (1997). Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *17*(21), 8300–12. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9334405>
- Sanchez-Diaz, P. C., Burton, T. L., Burns, S. C., Hung, J. Y., & Penalva, L. O. F. (2008). Musashi1 modulates cell proliferation genes in the medulloblastoma cell line Daoy. *BMC cancer*, *8*(280). doi:10.1186/1471-2407-8-280
- Seigel, G. M., Hackam, A. S., Ganguly, A., Mandell, L. M., & Gonzalez-Fernandez, F. (2007). Human embryonic and neuronal stem cell markers in retinoblastoma. *Molecular vision*, *13*(May), 823–32. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2768758&tool=pmcentrez&rendertype=abstract>

- Shibata, S., Umei, M., Kawahara, H., Yano, M., Makino, S., & Okano, H. (2012). Characterization of the RNA-binding protein Musashi1 in zebrafish. *Brain Research*, 1462, 162–73. doi:10.1016/j.brainres.2012.01.068
- Siddall, N. a, Kalcina, M., Johanson, T. M., Monk, A. C., Casagrande, F., Been, R. P., McLaughlin, E. a, et al. (2012). Drosophila Rbp6 is an orthologue of vertebrate Msi-1 and Msi-2, but does not function redundantly with dMsi to regulate germline stem cell behaviour. *PLOS one*, 7(11). doi:10.1371/journal.pone.0049810
- Sureban, S. M., May, R., George, R. J., Dieckgraefe, B. K., McLeod, H. L., Ramalingam, S., Bishnupuri, K. S., et al. (2008). Knockdown of RNA binding protein musashi-1 leads to tumor regression in vivo. *Gastroenterology*, 134(5), 1448–58. doi:10.1053/j.gastro.2008.02.057
- Wang, X.-Y., Penalva, L. O., Yuan, H., Linnoila, R. I., Lu, J., Okano, H., & Glazer, R. I. (2010). Musashi1 regulates breast tumor cell proliferation and is a prognostic indicator of poor survival. *Molecular cancer*, 9, 221. doi:10.1186/1476-4598-9-221
- Wang, X.-Y., Yin, Y., Yuan, H., Sakamaki, T., Okano, H., & Glazer, R. I. (2008). Musashi1 modulates mammary progenitor cell expansion through proliferin-mediated activation of the Wnt and Notch pathways. *Molecular and cellular biology*, 28(11), 3589–99. doi:10.1128/MCB.00040-08
- Weiss, J. N. (1997). The Hill equation revisited: uses and misuses. *FASEB*, 11, 835–841.
- Yoda, A., Sawa, H., & Okano, H. (2000). MSI-1, a neural RNA-binding protein, is involved in male mating behaviour in *Caenorhabditis elegans*. *Genes to cells*, 5(11), 885–895. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11122376>

Figures

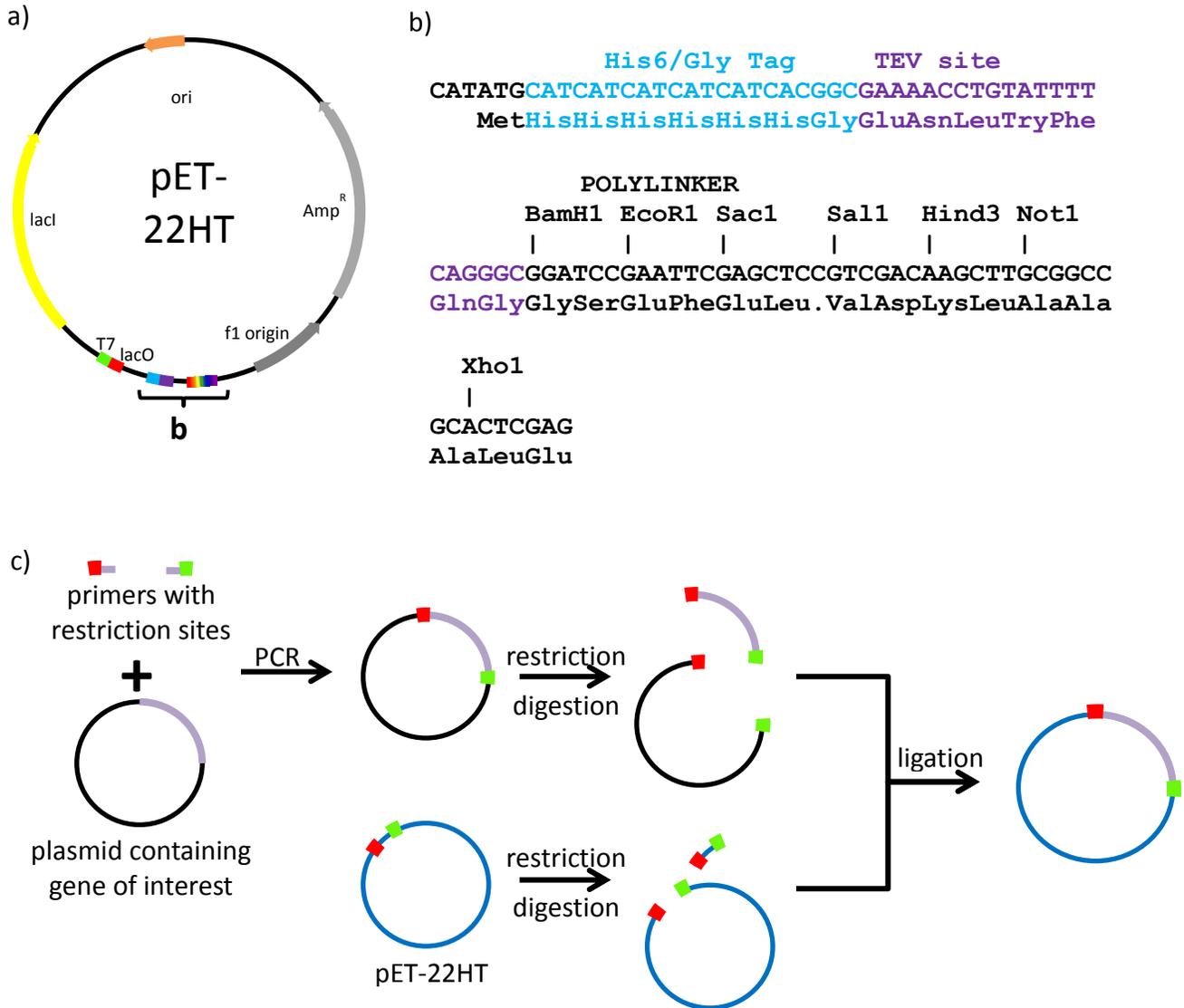


Figure 5: Cloning scheme used for generation of expression constructs containing MSI1 homologs: a) The pET-22HT plasmid used as cloning vector. Contains a *lacI* gene and *lac* operon (*lacO*) for inducible expression, a T7 RNA polymerase promoter, bacterial *f1* origin for expression of the ampicillin resistance gene (Amp^R), origin of replication (*ori*), N-terminal histidine6/glycine tag with TEV protease site and a polylinker containing multiple restriction enzyme sites for insertion of a gene of interest; b) His6/Gly-tag and polylinker region of pET-22HT showing available restriction sites. *EcoR1* and *Sac1* were used as 5' restriction sites to insert genes of interest in frame; c) Primers designed with restriction sites were used in PCR to insert two restriction sites flanking the RRM of the MSI1 homolog gene. Both vector pET-22HT and MSI1 homolog were subjected to restriction digestion, then ligated together to produce expression constructs.

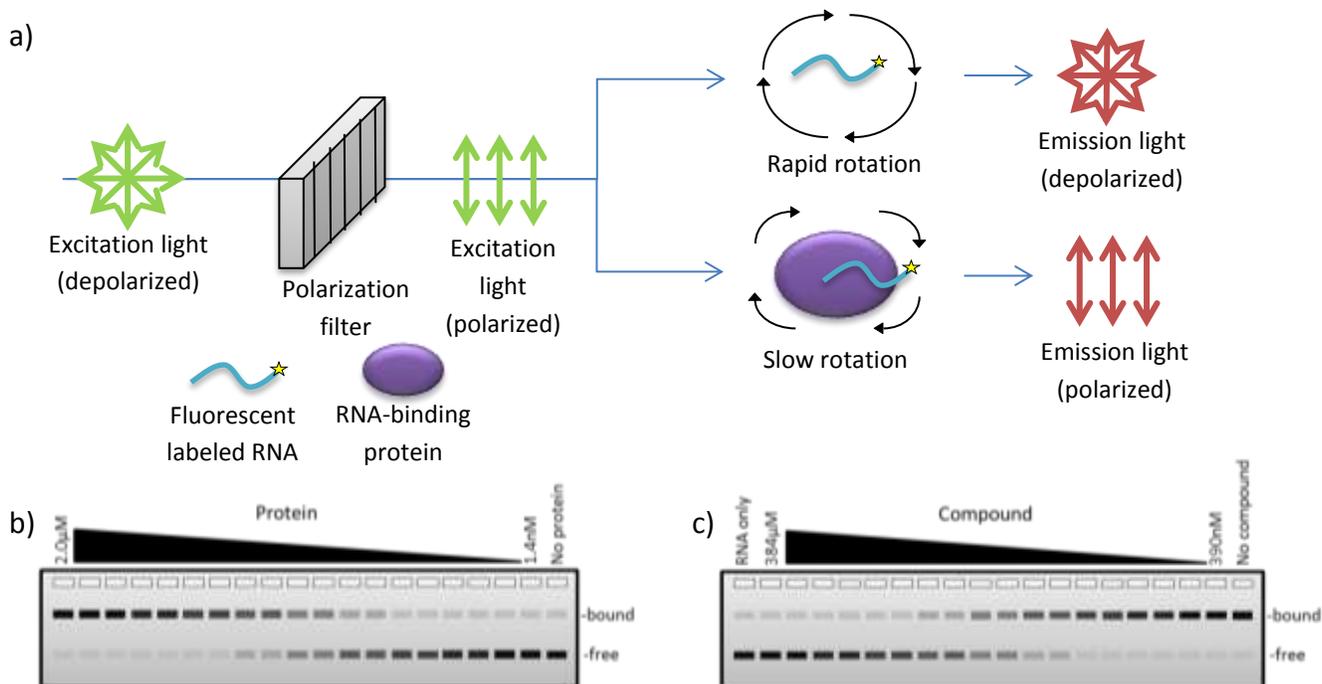


Figure 6: Theory of assays used: a) Fluorescence polarization quantifies the amount of polarized light emitted from samples. Light is passed through a polarization filter into samples from a direct titration with varying ratios of RNA and protein. If an individual RNA is not bound to protein, it tumbles rapidly in solution and will emit depolarized light. If an individual RNA is bound to protein, it tumbles slowly in solution and will emit polarized light. This allows for the quantification of the relative amount of bound versus free RNA in solution; b) Fluorescence-electrophoretic-mobility shift assay (F-EMSA) quantifies the amount of free versus bound RNA in a sample. The direct titration experiment titrates protein into a fixed concentration of RNA. Gels contain a “no protein” control sample to verify position of the free RNA band; c) F-EMSA was used to analyze dose response assays in which compound was titrated into a fixed amount of RNA and protein. An “RNA only” control verifies position of free RNA band and a “No compound” control verifies position of bound RNA.

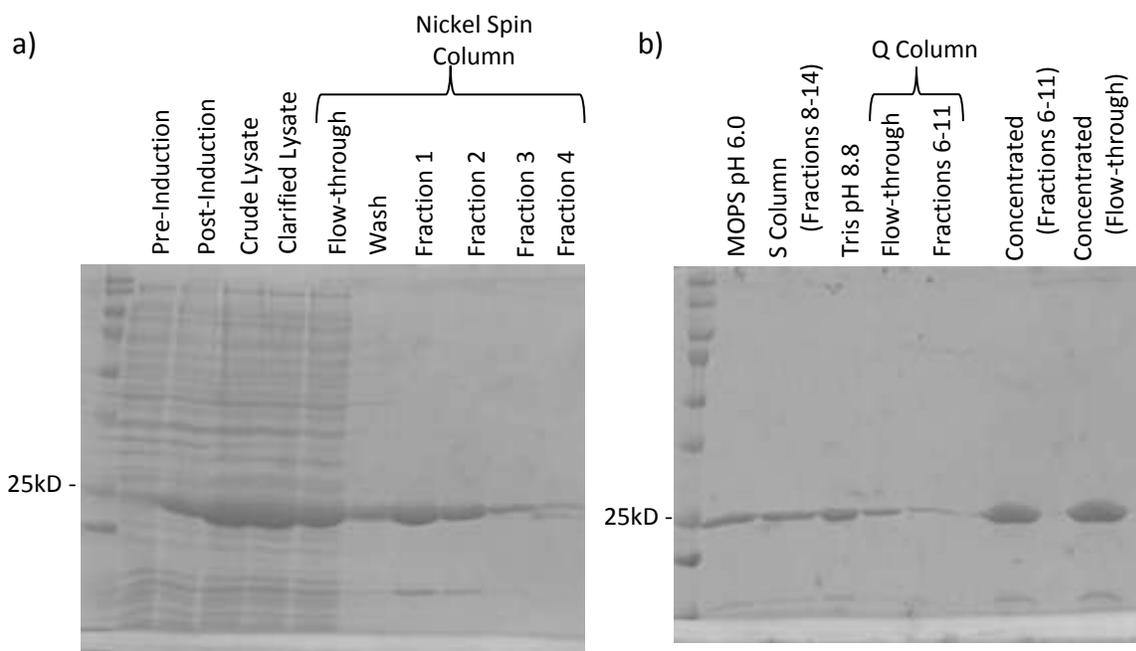


Figure 7: Example of purification of *D. rerio* msi1 (Dr_msi1_T): a) induced expression of Dr_msi1_T; presence in clarified lysate indicates solubility; presence in flow-through and wash of Ni-NTA spin column indicates protein loss; protein elutes in fractions 1-3 of Ni-NTA spin column with 15kD band, b) Dr_msi1_T present after dialysis against MOPS (pH 6.0) buffer in preparation for HiTrap SP cation exchange column (S column); S column removes 15kD band. Presence in Tris (pH 8.8) buffer in preparation for HiTrap Q anion exchange column (Q column). Dr_msi1_T present in both flow-through and fractions 6-11 of Q column. Concentration of protein assures no contamination not seen in more dilute samples is present.

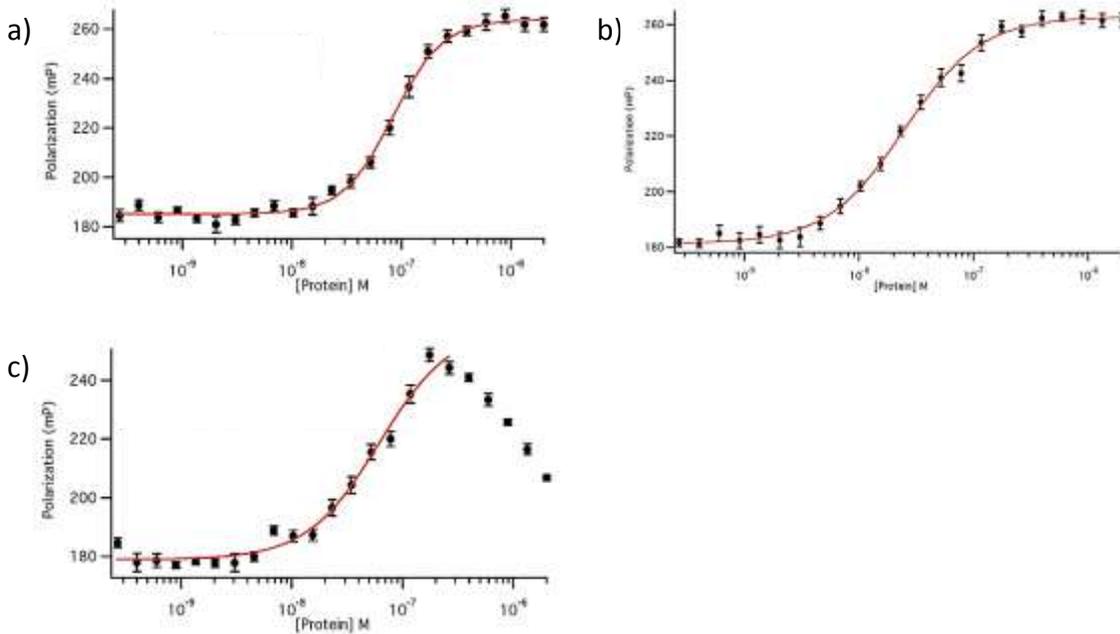


Figure 8: Examples of fluorescence polarization plots with Dr_msi1_T: a) flow-through from Q column without bovine gamma globulin (BGG) [$K_d = 83.6 \pm 2.88$ nM, Hill = 2.01]; b) flow-through from Q column in 1x buffer plus BGG [$K_d = 31.1 \pm 1.53$ nM, Hill = 1.17]; c) fractions from Q column in 1x buffer plus BGG; decrease in polarization at high concentrations suggests loss of protein due to insolubility; therefore, Dr_msi1_T collected from fractions was not used in further assays.

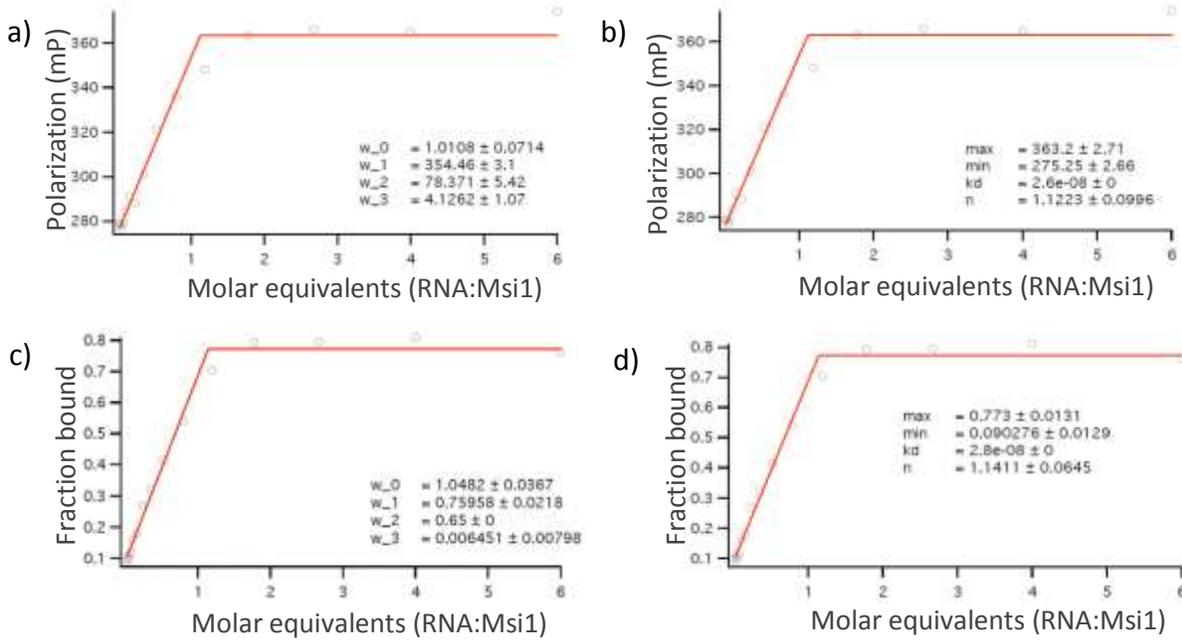


Figure 9: Stoichiometric binding of truncated mouse MSI1 (Mm_MSI1_T) reveals 1:1 binding to RNA: a) Piecewise fit of FP fits data to two straight lines and takes the intercept to be the stoichiometric binding coefficient (n) ($n = 1.08 \pm 0.07$); b) Equation 3 fit of FP ($n = 1.23 \pm 0.10$); c) Piecewise fit of F-EMSA ($n = 1.05 \pm 0.04$); d) Equation 3 fit of F-EMSA ($n = 1.14 \pm 0.06$).

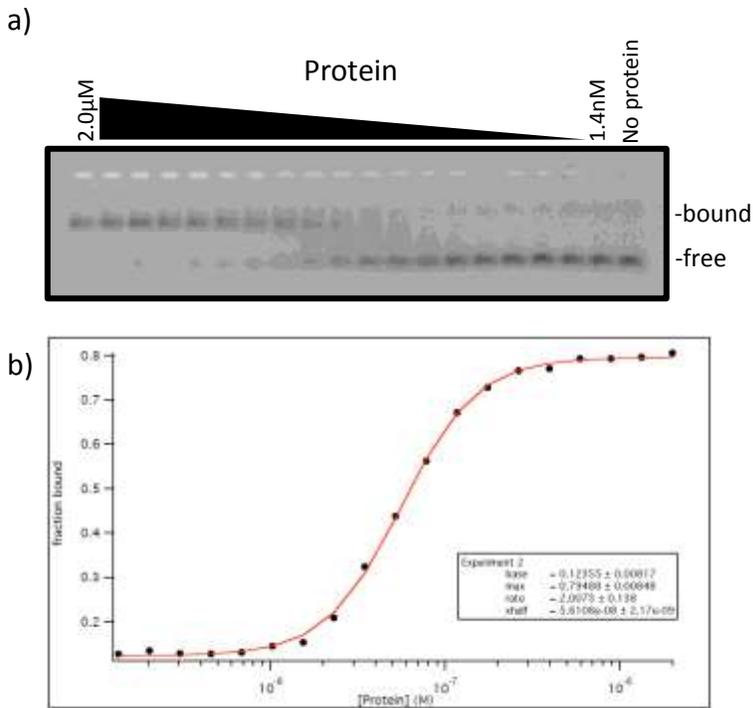


Figure 10: F-EMSA with *C. elegans* MSI-1 (Ce_MSI-1_T): a) Acrylamide gel shows distinction between bound and free RNA; b) Quantification of gel in (a) as a plot of fraction bound (RNA) versus protein concentration (M) ($K_d = 56.1 \pm 2.17$ nM; Hill coefficient = 2.01).

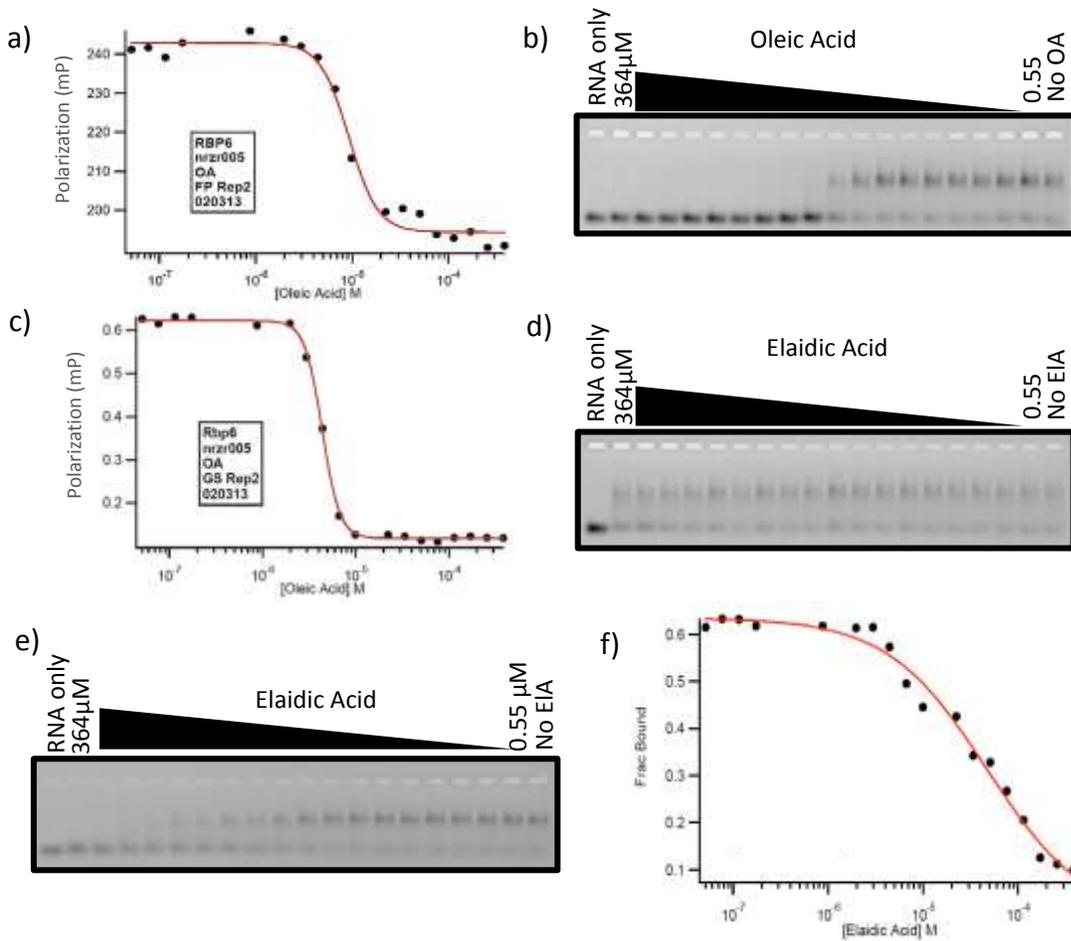


Figure 11: Dose response by FP with *D. melanogaster* Rbp6 (Dm_Rbp6_T) showed no inhibition, while weak inhibition was seen with *C. elegans* MSI-1 (Ce_MSI-1_T): a) Fluorescence polarization versus oleic acid concentration of Dm_Rbp6_T ($IC_{50} = 9.3\mu M$, $K_{i, app} = 5.08\mu M$, Shape = 2.9); b) F-EMSA 5% acrylamide gel with oleic acid and Dm_Rbp6_T; c) Quantification of gel in (b) ($IC_{50} = 4.37\mu M$, $K_{i, app} = 5.18\mu M$, Shape = 4.6); d) F-EMSA 5% acrylamide gel with elaidic acid and Dm_Rbp6_T (no inhibition); e) F-EMSA 5% acrylamide gel with elaidic acid and Ce_MSI-1_T; f) Quantification of gel in (e) ($IC_{50} = 37.0\mu M$, $K_{i, app} = 40.9\mu M$, Shape = 0.88).

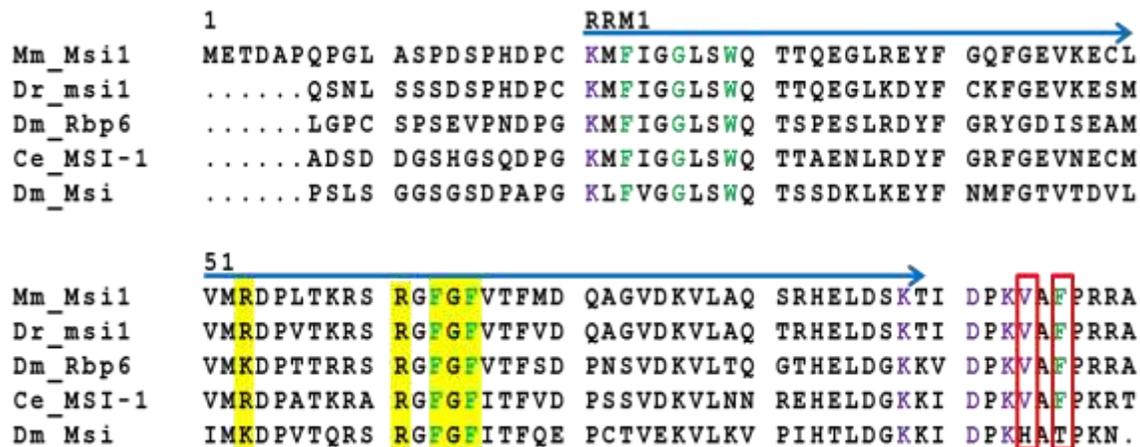


Figure 12: Alignment of amino acid sequences of MSI1 homolog RRM1. Green residues participate in aromatic stacking interactions with RNA, purple residues participate in hydrogen-bonding with RNA, residues in red boxes highlight amino acid differences in fly Msi (Dm_Msi), and residues highlighted in yellow are predicted by the Ryder Lab to contribute to oleic acid inhibition. Weaker binding of Dm_Msi_T to RNA was observed and may be due to amino acid differences boxed in red.

Tables

Table 2: Clones, primers and restriction sites used in creation of MSI1 homolog expression constructs. Primer sequences were designed to insert restriction sites flanking the RRM2 of MSI1 homolog genes and to insert stop codons 16 amino acids after RRM2.

Species	Construct Abbreviation	Clone containing insert	N-terminal restriction site	C-terminal restriction site	Included amino acids
<i>D. rerio</i>	Dr_msi1_T	pME185-FL3	EcoR1	Sal1	32-217
<i>C. elegans</i>	Ce_MSI-1_T	pDONR201	EcoR1	Sal1	7-192
<i>D. melanogaster</i>	Dm_Msi_T	LD31631	Sac1	Sal1	162-346
<i>D. melanogaster</i>	Dm_Rbp6_T	RE25373	EcoR1	Sal1	16-201
<i>M. musculus</i>	Mm_MSI1_T	Obtained from C. Clingman			7-192

Construct Abbreviation	Primer sequence	
Dr_msi1_T	Forward	GGGG GAATTC CAAAGTAACCTGTCCTCC
	Reverse	GGGG GTCGAC UU CTA TGGTGACATCACTTCCTT
Ce_MSI-1_T	Forward	GGGG GAATTC GCTGATTCCGATGACGGG
	Reverse	GGGG GTCGAC UU CTA CGGAAGCATGACCTCCTT
Dm_Msi_T	Forward	GGGG GAGCTC CCCAGCCTGAGCGGAGGC
	Reverse	GGGG GTCGAC UU CTA CGGTGTGACTGCTTCCTT
Dm_Rbp6_T	Forward	GGGG GAATTC CTGGGACCCTGTTCCCCC
	Reverse	GGGG GTCGAC UU CTA CGGCAGCATCACTTCCTT
Mm_MSI1_T	Obtained from C. Clingman	

Table 3: Direct titration FP and F-EMSA data for mouse MSI1 and MSI1 homologs. K_d is the binding constant at which protein is 50% bound to RNA and was determined from fitting FP and F-EMSA data to equation 1.

Protein	FP		F-EMSA	
	Mean K_d	Mean Hill	Mean K_d	Mean Hill
Dm_Rbp6_T	20.8 ± 1.36nM	1.42	32.4 ± 5.25nM	1.77
Mm_MSI1	26.0 ± 1.06nM	1.35	28.4 ± 5.30nM	1.44
Dr_msi1_T	26.2 ± 1.42nM	1.32	24.7 ± 3.73nM	1.72
Ce_MSI-1_T	61.7 ± 4.24nM	1.32	94.0 ± 2.00nM	1.98
Dm_Msi_T	360 ± 12.8nM	2.16	313 ± 13.0nM	4.37

Table 4: Dose response FP and f-EMSA data for MSI1 and MSI1 homologs with the ω -9 monounsaturated fatty acid: oleic acid.

Protein	Compound	Method	Mean IC_{50}	Mean $K_{i, app}$	Mean Shape
Dm_Rbp66_T	Oleic Acid	FP	16.4 ± 0.33 μ M	8.91 ± 0.18 μ M	2.50
Dm_Rbp6_T	Oleic Acid	F-EMSA	7.39 ± 0.21 μ M	9.00 ± 0.26 μ M	3.90
Mm_MSI1	Oleic Acid	FP	16.0 ± 0.94 μ M	9.69 ± 0.57 μ M	2.42
Mm_MSI1	Oleic Acid	F-EMSA	7.19 ± 0.25 μ M	4.35 ± 0.15 μ M	4.40
Dr_msi1_T	Oleic Acid	FP	14.3 ± 0.51 μ M	3.87 ± 0.16 μ M	2.38
Dr_msi1_T	Oleic Acid	F-EMSA	7.52 ± 0.12 μ M	2.02 ± 1.23 μ M	4.51
Ce_MSI-1_T	Oleic Acid	FP	6.40 ± 0.49 μ M	3.37 ± 2.55 μ M	5.67
Ce_MSI-1_T	Oleic Acid	F-EMSA	3.45 ± 0.13 μ M	3.82 ± 0.14 μ M	4.98
Dm_Msi_T	Oleic Acid	FP	29.9 ± 1.68 μ M	33.6 ± 1.88 μ M	2.28
Dm_Msi_T	Oleic Acid	F-EMSA	19.2 ± 1.24 μ M	21.5 ± 1.39 μ M	2.67