A Feedback Loop Couples Musashi-1 Activity to Omega-9 Fatty Acid Biosynthesis: A Dissertation

Carina C. Clingman

University of Massachusetts Medical School, carina.clingman@umassmed.edu

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

Part of the Amino Acids, Peptides, and Proteins Commons, Biochemistry Commons, and the Molecular Biology Commons

Recommended Citation


This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
A FEEDBACK LOOP COUPLES MUSASHI-1 ACTIVITY TO OMEGA-9 FATTY ACID BIOSYNTHESIS

A Dissertation Presented

by

Carina Cleveland Clingman

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

DOCTOR OF PHILOSOPHY

September 3, 2014

BIOCHEMISTRY AND MOLECULAR PHARMACOLOGY
A FEEDBACK LOOP COUPLES MUSASHI-1 ACTIVITY TO OMEGA-9 FATTY ACID BIOSYNTHESIS

A Dissertation Presented By

Carina Cleveland Clingman

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

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

___________________________________________
Alonzo H. Ross, Ph.D., Member of Committee

___________________________________________
Phillip D. Zamore, Ph.D., Member of Committee

___________________________________________
John Rinn, Ph.D., Member of Committee

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

___________________________________________
Kendall L. Knight, Ph.D., Chair of Committee

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

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

Biochemistry and Molecular Pharmacology
September 3rd, 2014
Acknowledgments

I would like to thank my thesis advisor Sean Ryder for making me feel welcome at UMass Medical School during my interview and ever since. His patience, understanding, and guidance have helped me become the scientist I am now. My lab members, both past and present, have also been incredibly supportive throughout my graduate school career. They provided a wonderful environment in which to learn, question, think, and do science. I would especially like to thank Ruth Zearfoss for her guidance. She has been my mentor since day one of my rotation, and she is still the first person I look to for scientific help or a friendly chat.

I also thank the friends I’ve made here on the East Coast. Leaving my support network in California was scary, but the friends I’ve made here are wonderful and supportive. Belinda Barbagallo deserves particular acknowledgment. Her support and sage advice during this process has been invaluable to my success, and her friendship has certainly saved me quite a bit of money in therapy. On that note, I’d also like to thank my wonderful dogs Boo Jones and Zoe, who need play time and tummy rubs whether or not my experiments are going well.

Finally, I would like to thank my thesis research advisory committee. Ken Knight has been a wonderful committee chairman, with kind and thoughtful advice and an open door whenever I needed it. Phil Zamore, Alonzo Ross, and Marc Freeman have given me wonderful guidance and provided a fun and supportive environment in which to craft my project. I am also grateful that John Rinn has agreed to join my defense committee as an outside member.
Abstract

All living creatures change their gene expression program in response to nutrient availability and metabolic demands. Nutrients and metabolites can directly control transcription and activate second-messenger systems. In bacteria, metabolites also affect post-transcriptional regulatory mechanisms, but there are only a few isolated examples of this regulation in eukaryotes. Here, I present evidence that RNA-binding by the stem cell translation regulator Musashi-1 (MSI1) is allosterically inhibited by 18-22 carbon \( \omega-9 \) monounsaturated fatty acids. The fatty acid binds to the N-terminal RNA Recognition Motif (RRM) and induces a conformational change that prevents RNA association. Musashi proteins are critical for development of the brain, blood, and epithelium. I identify stearoyl-CoA desaturase-1 as a MSI1 target, revealing a feedback loop between \( \omega-9 \) fatty acid biosynthesis and MSI1 activity. To my knowledge, this is the first example of an RNA-binding protein directly regulated by fatty acid. This finding may represent one of the first examples of a potentially broad network connecting metabolism with post-transcriptional regulation.
Table of Contents

Title Page i
Signature Page ii
Acknowledgments iii
Abstract iv
Table of Contents v
List of Figures and Tables vii
List of common abbreviations ix
Preface x

Chapter I: Introduction 1
Musashi is an RNA-binding protein essential for neural and epithelial development 2
Musashi family proteins protect stem and progenitor cell fate 3
Musashi-1 in gametogenesis and embryogenesis 12
Musashi-1 is essential for neural development 14
Msi1 promotes cancerous tumor growth 22
Metabolic regulation of gene expression 26
Riboswitches: metabolite-sensitive RNA elements 27
Identifying eukaryotic riboswitches 32
Metabolite-sensitive RNA-binding proteins 35
Identifying metabolite-sensing proteins with RNA-binding activity 46

Chapter II: The small molecule metabolite oleic acid is an allosteric inhibitor of Musashi-1 RNA-binding activity 49
Abstract 50
Introduction 51
Results 53
Small molecule screen to identify inhibitors of Musashi-1 53
Msi1 is specifically inhibited by 18-22 carbon ω-9 monounsaturated fatty acids 63
Oleic acid directly interacts with the RRM1 of Msi1 to affect inhibition 65
A model of the omega-9 fatty acid inhibition mechanism 81
Discussion 96
Methods 98

Chapter III: Musashi-1 and the lipid biosynthetic enzyme stearoyl-CoA desaturase participate in a feedback loop to regulate cellular proliferation 110
Abstract 111
Introduction 111
Results
Oleic acid treatment reduces oligodendrocyte progenitor cell proliferation 113
MSI1 regulates stearoyl-CoA desaturase 116
SCD lipid products are enriched upon MSI1 expression 123
Additional putative MSI1 regulatory targets involved in lipid metabolic processes 126

Discussion 126

Methods 134

Chapter 4: Concluding remarks and future directions 143
Biochemical characterization of Musashi-1 inhibition 144
18-22 carbon ω-9 fatty acids inhibit MSI1 RNA-binding activity 144
A model for the mechanism of inhibition 146
Biological characterization of Musashi-1 inhibition 149
A model for MSI1 in proliferation 149
A model for MSI1 in oligodendrocyte differentiation 155
Metabolic regulation 159

Appendix 1: Lipidomics protocol and data analysis 160

References 164
List of Figures and Tables

Figure 1.1: Pattern of MSI1 expression in the central nervous system..............................6
Figure 1.2: MSI1 and MSI2 are highly conserved.................................................................8
Figure 1.3: RNA-bound NMR structure of Msi1 RRM1 bound to RNA.............................11
Figure 1.4: Schematic representation of an oligodendrocyte..............................................16
Figure 1.5: Msi1 knockdown induces differentiation in CG-4 OPCs .................................20
Figure 1.6: TPP-regulated alternative splicing of the NMT1 gene....................................30
Figure 1.7: The structure of human ADAR2 bound to IP6 ....................................................39
Figure 1.8: The two forms of IRP1..........................................................................................44
Figure 2.1: His6-Msi1 purification and activity test...............................................................55
Figure 2.2: Small molecule screen.......................................................................................58
Table 2.1: Structure-activity relationship analysis demonstrates specificity of inhibition ..............................................................................................................60
Figure 2.3: Dissociation constant and CMC quantification ...............................................62
Figure 2.4: RNA binding specificity and inhibition by specific fatty acids is conserved in MSI2 ....................................................................................................................................67
Figure 2.5: Msi1 tryptophan fluorescence at 350 nm as a function of oleic and elaidic acid. ........................................................................................................................................................................69
Figure 2.6: The Msi1 tryptophan fluorescence spectrum changes as a function of inhibitor titration ..........................................................................................................................................................71
Figure 2.7: Musashi-1 RRM1 and RRM2 NMR structures..................................................74
Figure 2.8: CD and AUC characterization of oleic acid association. .....................................76
Figure 2.9: NMR characterization of Msi1 RRM1 ................................................................. 79
Figure 2.10: Docked model of Msi1 and oleic acid ............................................................... 83
Figure 2.11: Molecular dynamics studies of Msi1 with oleic acid ....................................... 86
Figure 2.12: MD model of Msi1 with oleic acid ................................................................. 89
Figure 2.13: MD comparison of RNA-bound and oleic acid-bound Msi1 ......................... 92
Figure 2.14: Mutational analysis of Msi1 RRM1 ............................................................... 95
Figure 3.1: Lipidomics analysis of CG4 oligodendrocyte progenitor cells ....................... 115
Figure 3.2: Fatty acid treatment of CG-4 OPCs ................................................................. 118
Figure 3.3: SCD1 is a putative Msi1 target ........................................................................ 120
Figure 3.4: Msi1 is a positive regulator of SCD ................................................................. 122
Figure 3.5: Lipidomics analysis of HEK293t cells ± MSI1 expression .............................. 125
Figure 3.6: Model of SCD regulation by MSI1 .................................................................. 129
Figure 4.1: Model of MSI1/SCD homeostasis in normal, proliferative cells ..................... 153
Figure 4.2: Model of MSI1 homeostasis in normal in OPCs ............................................. 157
List of common abbreviations

Musashi (MSI)
Homo sapiens Musashi-1 (hMSI1)
Mus musculus Musashi-1 (Msi1)
Drosophila melanogaster Musashi (Msi)
Xenopus laevis nervous system-specific RNP protein-1 (xNRP1) aka (xMsi1)
Danio rario Musashi-1 (zMsi1)
Homo sapiens Musashi-2 (MSI2)
Mus musculus Musashi-2 (Msi2)
RNA Recognition Motif (RRM)
Untranslated region (UTR)
Germline Development-2 (GLD2)
Germinal vesicle breakdown (GVBD)
Central nervous system (CNS)
Subventricular zone (SVZ)
Platelet-derived growth factor (PDGF)
Fibroblast growth factor-2 (FGF-2)
Doublecortin (DCX)
Hematopoietic stem cells (HSCs)
Mitochondrial proliferator-activated receptor (PPAR)
Phosphoinositide 3-kinase (PI3K)
Inositol trisphosphate (IP3)
Thiamine pyrophosphate (TPP)
Tryptophan RNA-binding attenuation protein (TRAP)
Oligodendrocyte progenitor cell (OPC)
Fluorescence electrophoretic mobility shift (F-EMSA)
Fluorescence polarization (FP)
Library of Pharmacologically Active Compounds (LOPAC)
Aurintricarboxylic acid (ATA)
Critical micelle concentration (CMC)
Coenzyme A (CoA)
N-acetyl-L-tryptophanamide (NATA)
Molecular dynamics (MD)
Stearoyl-CoA desaturase-1 (SCD1)
Stearoyl-CoA desaturase-1 (SCD)
False discovery rate (FDR)
Triacylglycerol (TAG)
Sterol responsive element binding protein (SREBP)
Saturated fatty acid (SFA)
Monounsaturated fatty acid (MUFA)
Oleate-activated transcription factor (OAF)
Preface

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

Part of Chapter I and one paragraph of Chapter IV were adapted from:

Chapters II and III were adapted from:

Laura M. Deveau performed the circular dichroism experiments and made part of figure 2.8. She also collected the NMR data in Chapter II.

Francesca Massi analyzed the NMR data collected by Laura M. Deveau and made part of figure 2.9. She also performed the MD simulations in Chapter II and made figures 2.11-2.13

Shivender MD Shandilya performed the molecular docking simulation in Chapter II.
Chapter I: Introduction
Musashi is an RNA-binding protein essential for neural and epithelial development

Stem cells have the capacity to replicate continually and differentiate into almost any cell type in the organism. Understanding the molecular mechanisms that maintain the pluripotent state of stem cells is essential to basic and pharmaceutical research. Stem cell research goals include the design of strategies to distinguish differentiated from pluripotent cells, specific induction of stem cell differentiation, and the potential for reprogramming patient-specific adult cells for research and treatment.

Regulation of gene expression at the translational level allows organisms to modulate the production of protein from mRNA according to specific developmental cues. Post-transcriptional and post-translational mechanisms enable organisms to regulate the expression of already-transcribed genes. Post-transcriptional regulation takes many forms, including capping, poly-A tail addition, splicing, translocation, mRNA stabilization or destabilization, sequestration, and editing. Some proteins, such as capping enzymes and splicing factors, act constitutively on most RNAs, while other proteins target a more limited set of specific transcripts to modulate gene expression according to precise cues. Understanding the
mechanisms by which these specific RNA-binding proteins recognize and regulate expression of their targets is essential to understanding development.

Musashi family proteins protect stem and progenitor cell fate

Musashi-1 (MSI1) is a dual RNA Recognition Motif (RRM) RNA-binding protein. MSI1 has a similar RNA binding domain to Drosophila melanogaster Musashi (Msi), which regulates asymmetric division of sensory organ precursor cells (Nakamura et al. 1994; Sakakibara et al. 1996). Msi was identified in a mutant fly with an external sensory bristle patterning defect (Nakamura et al. 1994; Sakakibara et al. 1996). Normal bristles consist of four cells: A neuron, a glial cell, and two external support cells called shaft and socket cells. In MSI mutant flies, the ratio of these cells is aberrant, leading to a phenotype in which some bristles contain multiple shaft cells so they appear to have double bristles (Nakamura et al. 1994). Msi regulates this asymmetric division by participating in a regulatory feedback loop with TTK69 and Notch (Okabe et al. 2001).

Mouse Msi1 was identified by its homology to Msi and Xenopus laevis nervous system-specific RNP protein-1 (xNRP1), also known as xMsi1 (Sakakibara et al. 1996). Human MSI1 (hMSI1) was later identified during the human genome project by its homology to Msi, xNRP1/xMsi, and mMsi1 (Good et al. 1998). The vertebrate Msi1 proteins are highly expressed in central nervous system (CNS) stem cells and in progenitor cells that give rise to glia and neurons. Expression is lost in differentiated glial and neuronal cells, suggesting that Msi1 plays a role in
neurogenesis (Figure 1.1) (Sakakibara et al. 1996; Good et al. 1998). Although Msi1 was first characterized in the nervous system, Msi1 is also expressed in epithelial stem and progenitor cells, where it promotes proliferation and prevents differentiation (Kayahara et al. 2003; Nishimura et al. 2003; Potten et al. 2003; Akasaka et al. 2005; Asai et al. 2005; Colitti and Farinacci 2009; Murayama et al. 2009). The importance of Msi1 in the epithelial cell lineage was further confirmed when Lan et al. demonstrated that Msi1+ mouse embryonic stem cells have the capacity to differentiate into neural and epithelial cells when hypodermically engrafted onto the backs of adult mice (Lan et al. 2010).

Several other MSI family proteins have also been identified in vertebrates and flies. Musashi-2 (Msi2) is 69% identical to Msi1 at the amino acid level and is expressed primarily in hematopoietic and leukemic stem cells (Figure 1.2) (Sakakibara et al. 2002; Kharas et al. 2010; Park et al. 2014). Additionally, Msi2 is required for embryonic stem cell self-renewal and pluripotency (Wuebben et al. 2012; Park et al. 2014). Zebrafish have two main splice variants of the Musashi 1 homolog (zMsi1) (Shibata et al. 2012). zMsi1 is expressed in neural tissue during development, and knockdown of zMsi1 results in CNS abnormalities (Shibata et al. 2012). Although Msi was the first MSI family protein identified in *Drosophila melanogaster*, recent genome sequencing efforts have revealed that RBP6 shares greater sequence homology with mouse Msi1 and human MSI1 (Siddall et al. 2012).
Figure 1.1
Figure 1.1: Pattern of MSI1 expression in the central nervous system. Blue cells indicate MSI1 expression pattern.
Figure 1.2: MSI1 and MSI2 are highly conserved. Sequence alignment shows 69% conservation between MSI1 and MSI2 across the entire protein, and 83% conservation within the RRM domains. Regions that correspond to α-helices, β-sheets, and intervening loops as defined by NMR spectroscopy for MSI1 RRM1 are diagramed above the alignment.
While RBP6 does not share Msi functionality in regulating asymmetric division of neural stem cells, it does appear to be expressed along with Msi in mouse germline cells (Siddall et al. 2012; Sutherland et al. 2013). A growing body of evidence suggests that MSI family proteins are involved in stem cell maintenance and differentiation, but little is yet known about the mechanisms by which MSI proteins regulate these processes.

In 2001, the Okano lab used in vitro SELEX experiments to identify aptamers that bind to Msi1 (Imai et al. 2001). The Msi1 consensus sequence, (G/A)U1-3AGU, was identified by eye in most SELEX aptamers (Imai et al. 2001). Msi1 RRM1 binds to RNA with greater affinity than RRM2, which is thought to stabilize the RNA interaction (Imai et al. 2001; Miyanoiri et al. 2003). NMR studies indicate that RRM1 is more dynamic than RRM2, possibly explaining the difference in apparent RNA affinities by the two RRM domains (Miyanoiri et al. 2003). The Katahira lab recently published the NMR structure of Msi1 RRM1 bound to the minimal motif GUAGU (Ohyama et al. 2011). The RNA-bound structure reveals that specific MSI1 residues interact with the RNA bases (Figure 1.3). As expected, most of the residues that interact with RNA are located on the beta sheet face, but some residues in the flexible loop regions also associate with RNA. Notably, the adenine (position 3) and guanine (position 4) bases stack with two of the three conserved phenylalanines found on the beta sheet face of most RRM domains, but the adenine is sandwiched by a third phenylalanine (F96) from a loop region. Additionally, the first guanine stacks on a tryptophan (W29) in the loop between β1 and α1 (Ohyama et al. 2011).
Figure 1.3: RNA-bound NMR structure of Msi1 RRM1 bound to RNA. Adapted from Oyahama et al, 2011.
Thermodynamic characterization of RNA mutants by Ruth Zearfoss has since revealed that sequence recognition by Msi1 is primarily driven by the presence of a UAG element (Zearfoss et al., submitted). This recent understanding of the MSI expression pattern and consensus binding sequence will be useful in the study of MSI family function in a number of systems, including gametogenesis, neurodevelopment, and cancers of the neural and epithelial lineages.

Musashi-1 in gametogenesis and embryogenesis

Gametogenesis is a precise cell cycling and differentiation process that results in the formation of a haploid gamete. During gametogenesis, gametes undergo periods of transcriptional inactivation, during which they must maintain temporal and spatial expression of genes required for proper development. Sequence-specific RNA-binding proteins assist in the translational control of mRNA transcripts that have been produced and stored in the gametes prior to transcriptional inactivation. Members of the Musashi family of RNA-binding proteins have been identified in germ cells of several organisms, where they assist in regulating the expression of genes critical for developmental.

Xenopus laevis Musashi (xMsi) expression in the reproductive system is localized to the oocyte, where it plays a critical role in oocyte maturation (Charlesworth et al. 2006). Polyadenylation assays show that xMsi1 translationally activates the proto-oncogene Mos, an essential regulatory factor in oocyte maturation. xMsi1 binds to the Mos 3’ UTR and associates with the poly-A
polymerase Germline Development-2 (GLD2) to promote cytoplasmic polyadenylation and translation (Charlesworth et al. 2006; Prasad et al. 2008; Cragle and Macnicol 2014). Interestingly, MSI-dependent Mos activation still occurs if xMsi1 is replaced by mouse Msi1 in *Xenopus* oocytes, indicating functional conservation across species (MacNicol et al. 2011).

MSI proteins are also crucial to meiotic progression during *Xenopus* gametogenesis (Gunter and McLaughlin 2011). Antisense ablation of both xMsi1 and xMsi2 in embryos halts meiotic progression and germinal vesicle breakdown (GVBD). Knockdown of either xMsi1 or xMsi2 yields a less-dramatic phenotype, with delayed GVBD and slower meiotic progression (Gunter and McLaughlin 2011).

MSI regulation during gametogenesis and embryogenesis in other systems has not been studied as comprehensively. Recent proteomics experiments indicate that MSI protein expression in oocytes is conserved among vertebrates and invertebrates (Lotan, et al. 2014). In mice, *in situ* hybridization and immunohistochemical analysis reveals that MSI proteins are expressed in the sertoli cells of the testis and the granulosa cells of the ovaries (Saunders et al. 2002). In *Drosophila* testis, loss of MSI disrupts the balance between germline stem cell renewal and differentiation, resulting in the premature differentiation (Siddall et al. 2006). Finally, there is evidence that the association between MSI proteins and GLD2 is conserved in mammals, which indicates that the mechanism of translational activation described in *Xenopus* oocytes may be widespread, although this has not yet been tested (Cragle and Macnicol 2014).
Musashi-1 is essential for neural development

Neural development relies on a complex series of temporal and spatial cues that enable precise differentiation of neural stem cells into highly specialized mature cells of the central and peripheral nervous system. These cells must accurately sense and relay information about the organism and its environment. In the central nervous system (CNS), which consists of the brain and spinal cord, oligodendrocytes form a myelin sheath by wrapping long cellular processes around neuronal axons (Figure 1.4). Myelin is an essential structure in the vertebrate nervous system that protects against neural degeneration and enables saltatory nerve impulse propagation. Oligodendrocyte maturation is a tightly regulated process. To become myelin-producing oligodendrocytes, progenitor cells must undergo a specific program of proliferation, migration, and differentiation, but we are still far from understanding the how these processes are regulated.

In the forebrain, oligodendrocyte progenitor cells originate from a neural stem cell population maintained in the subventricular zone (SVZ) (Levison and Goldman 1993). Mitogens and chemo-attractants such as platelet-derived growth factor (PDGF) and fibroblast growth factor-2 (FGF-2) promote oligodendrocyte progenitor cell proliferation and migration throughout the CNS (Li et al. 2009). When oligodendrocyte progenitor cells reach their target, signals in the local environment trigger differentiation (Li et al. 2009).
Figure 1.4

Zearfoss et al., BBA. 2008.
Figure 1.4: Schematic representation of an oligodendrocyte forming a myelin layer to wrap around a neuronal axon (left), and a cutout representation of the layered composition of the myelin membrane (right) (Adapted from Zearfoss, et al. 2008).
Regulation of the decision to proliferate or differentiate is likely complicated by the highly polarized nature of oligodendrocyte progenitor cells. These cells must sense changes in the extracellular environment and quickly respond by altering gene expression in regions far from the cell body. Post-transcriptional regulatory mechanisms allow cells to swiftly and regionally respond to environmental stimuli. It is therefore not surprising that several RNA binding proteins, including Quaking, Fragile X Mental Retardation Protein and MSI1, have been implicated in regulating oligodendrocyte differentiation and myelination (Ebersole et al. 1996; Wang et al. 2004).

Msi1 is an essential regulatory factor involved in central nervous system development, as demonstrated through a variety of in vivo and cell-based assays (Sakakibara et al. 2002; Dobson et al. 2008). Quantitative analysis of mouse hippocampal gene expression patterns indicates that increased Msi1 expression is correlated with increased neural stem cell proliferation (Yagita et al. 2001). Additionally, Msi1 positive embryonic stem cells hypodermically engrafted into mice differentiate into a variety of tissue types, including epithelial and neural-like cells (Lan et al. 2010). Interestingly, MSI family involvement in neurodevelopment is highly conserved, with supporting data in Planaria, Drosophila, C. elegans, and all vertebrates assessed to date (Higuchi et al. 2008; MacNicol et al. 2008).

Msi1 null mice are born alive but die within 1-2 months of birth (Sakakibara et al. 2002). These mice display ataxia and hydrocephaly, a swelling of the brain.
Additionally, their brains are underdeveloped, and they have a pronounced expansion of progenitor cells from the ependymal layer of the neural tube. Embryonic neurospheres isolated from these mice contain fewer cellular lineages (Sakakibara et al. 2002).

In normal mice, imaging studies indicate that Msi1 is expressed in neural and glial progenitor cells, but not in mature neurons or glia (Figure 1.1) (Kaneko et al. 2000; Dobson et al. 2008). Msi1 knockdown in primary mouse oligodendrocyte progenitor cells results in increased caspase-3 production and apoptosis (Dobson et al. 2008). Elevated Msi1 expression in primary mouse oligodendrocyte progenitor cells blocks differentiation into mature oligodendrocytes (Dobson et al. 2008). My own unpublished data demonstrate that MSI1 reduction induces immortalized oligodendrocyte precursor cells to differentiate (Figure 1.5). Together, the Msi1 knockout mouse phenotypic and cell-based experiments indicate that Msi1 regulates neural progenitor and stem cell fate.

To date, only a few Msi1 targets have been identified and validated in neural tissue. The NOTCH antagonist NUMB was the first Msi1 target identified in vertebrates. Msi1 associates with consensus elements in the Numb 3’ UTR in vitro, and Numb mRNA co-immunoprecipitates with Msi1 protein from NIH 3T3 cells that have been transiently transfected with a HAT-tagged Msi1 construct. Finally, Msi1 negatively regulates luciferase protein expression from a reporter construct harboring the Numb 3’ UTR.
Figure 1.5

Control siRNA 10nM

Msi1siRNA 10nM

Control siRNA 10nM

Msi1siRNA 10nM
Figure 1.5: Msi1 knockdown induces differentiation in CG-4 OPCs. DIC (left) or fluorescent (right) images of CG-4 cells transfected with non-targeting control siRNA (top panels) or siRNA targeting Msi1 (bottom panels).
Reporter mRNA levels remain unaffected by Msi1 expression, indicating that Msi1 negatively regulates NUMB translation (Imai et al. 2001). NUMB inhibition is thought to be accomplished through a mechanism where the Msi1 C-terminus binds poly(A) binding protein, competing with eIF4G and preventing 80s ribosome assembly (Kawahara et al. 2008). NF-YA, a transcription factor that activates transcription of a Notch subunit, has also recently been identified as a putative Msi1 target (Lagadec et al. 2014). More work is needed to characterize the nature and extent of NF-YA regulation by Msi1.

MSI1 also translationally represses transcripts encoding the cell cycling antagonist p21WAF-1 (Battelli et al. 2006). Transient transfection of a Msi1 expression vector in HEK293T cells reduces p21WAF-1 protein levels but does not affect mRNA levels. Gel shift assays show that Msi1 interacts with a Msi1 binding site in the p21WAF-1 3′ UTR (Battelli et al. 2006). Additionally, Msi1 negatively regulates expression of a p21WAF-1 3′ UTR dsRED reporter in HEK293T cells (Battelli et al. 2006).

More recently, several putative Msi1 targets have been identified but not studied to the extent of NUMB and p21WAF-1. The immature neuronal migration factor doublecortin (DCX) has been identified as a MSI1 regulatory target (Horisawa et al. 2009). Msi1 binds to the 3′ UTR of DCX transcripts, and translationally represses expression of a DCX luciferase reporter in Neuro2A cells (Horisawa et al. 2009). Another putative Msi1 target is ROBO3, a protein involved in precerebellar neuron migration (Kuwako et al. 2010). Msi1 binds to ROBO3 transcripts and
upregulates ROBO3 protein levels in precerebellar neurons. Both Msi1 and ROBO3 knockout mice display similar midline crossing defects (Kuwako et al. 2010). Msi1 may also indirectly downregulate biogenesis of the microRNA let-7, an antagonist of stem cell proliferation (Kawahara et al. 2011). Together these Msi1 targets provide a link between Msi1 and its observed role as a factor involved in neurodevelopment and stem cell preservation.

**Msi1 promotes cancerous tumor growth**

Cancerous tumors often contain stem- and progenitor-like cells in which the proliferation checks and balances have malfunctioned. Proliferation-associated factors are increased in these cancers, while cell cycle regulatory factors are often decreased. Some of these factors are termed oncogenes for their apparent role in stimulating cancerous tumor growth. MSI1 was initially used as a common marker for neural and epithelial stem and progenitor cells, but recent evidence indicates that MSI expression is also elevated in cancerous tumors (Johnson et al. 2010; Levin et al. 2010). It is therefore not surprising that MSI1 has recently become a research focus as factor that likely supports tumorigenesis.

MSI1 is currently used as a marker of colorectal adenocarcinoma, medulloblastoma, ependymoma, endometrial carcinoma, esophageal adenocarcinoma, and oral squamous carcinoma (Nakano et al. 2007; Schulenburg et al. 2007; Gotte et al. 2008; Bobryshev et al. 2010; Ravindran and Devaraj 2012). RIP-ChIP analysis of MSI1-associated transcripts from HEK293T cells reveals 64 novel
putative regulatory targets (de Sousa Abreu et al. 2009). Of these, most are involved in cancer-related processes such as cell cycling, proliferation, differentiation, and apoptosis. Subsequent proteomics analysis indicates that these putative Msi1 targets undergo both positive and negative regulation in response to Msi1 overexpression (de Sousa Abreu et al. 2009).

A study with glioblastoma U87 tumor cell grafts showed that MSI1 expression increased with each successive tumor generation, indicating possible tumor cell dedifferentiation (Strojnik et al. 2006). Similarly, MSI1 mRNA is expressed at significantly higher levels in adenocarcinoma tissues than in normal colorectal mucosa tissues (Fan et al. 2010). MSI1 is expressed in 68% of primary breast tumors and in 100% of lymph node metastatic tumors that began as breast tumors (Wang et al. 2010).

MSI1 depletion in medulloblastoma and colorectal tumors results in decreased proliferation and increased apoptosis (Sanchez-Diaz et al. 2008; Sureban et al. 2008). MSI1 knockdown in MCF-7 and T47D breast tumor cells reduces cell survival and tumor xenograft growth (Wang et al. 2010). Conversely, overexpression of MSI1 in intestinal progenitor cells results in activation of the Wnt and Notch pathways and increased tumorigenesis in xenograph studies (Rezza et al. 2010). MSI1 also activates Notch signaling in endometrial carcinoma, where it has also been shown to regulate its cell cycling target P21WAF1 (Gotte et al. 2011).

A number of recent studies indicate that patient prognosis is correlated to MSI1 expression in tumors of neural and epithelial lineage. In patients with
malignant gliomas, poor prognosis is strongly correlated with MSI1 expression levels (Strojnik et al. 2007; Dahlrot et al. 2013). This correlation was also seen in patients with medulloblastoma tumors (Vo et al. 2012b). MSI1 expression is significantly higher in stage III adenocarcinoma tumors than in stage I-II tumors (Fan et al. 2010). Similarly, MSI1 levels are higher in tumors of the intestine than in normal intestinal mucosa (Kuang et al. 2013). Finally, poor patient prognosis is often correlated to the development of metastatic tumors. The frequency with which MSI is expressed in breast tumor tissue jumps from 68% of primary tumors to 100% of metastatic tumors assayed, providing compelling evidence for the link between patient survival and MSI1 expression (Wang et al. 2010).

MSI2 is primarily expressed in hematopoietic stem cells (HSCs), although some expression in neural stem and progenitor cells has been documented (Kharas et al. 2010; Lan et al. 2010). MSI2 knockdown results in reduced proliferation and increased apoptosis of HSCs in vivo (Kharas et al. 2010). Overexpression of human MSI2 promotes HSC cell cycling and proliferation in a mouse model (Kharas et al. 2010). MSI2 is highly expressed in myeloid leukemia cell models and depletion results in decreased proliferation and increased apoptosis (Kharas et al. 2010). Msi2 knockout mice survive up to a year but display defects in hematopoietic progenitor cell production that becomes more pronounced as they age (de Andres-Aguayo et al. 2011). Immature hematopoietic cells in these animals display increased differentiation as well as decreased proliferation (de Andres-Aguayo et al. 2011).
As with MSI1-expressing cancers, MSI2 expression levels are directly correlated with poor prognosis in myeloid and lymphoblastic leukemia patients (Kharas et al. 2010; Mu et al. 2013). Myeloid leukemia patients are classified into disease phases according to the severity of their symptoms and bone marrow composition. Phases include the almost asymptomatic chronic phase, the accelerated phase, and the blast crisis phase. Blast crisis behaves like an acute leukemia, usually resulting in patient death. Intriguingly, the samples from patients in blast crisis have low expression of the MSI1 target Numb (Ito et al. 2010). Loss of MSI2 restores Numb expression and attenuates the blast crisis phase, rendering MSI2 an attractive drug target in the quest for leukemia therapies (Ito et al. 2010).

Taken together, this body of previous research indicates that MSI family proteins may present a logical therapeutic target for various cancers and neurodegenerative diseases. However, regulation of MSI1 has not been extensively studied, so effective MSI1-targeting therapies have not yet been developed. While there is some correlation between expression of certain micro RNAs, post-transcriptional regulatory factors, transcription factors, and cyclin-dependent and mitogen-activated protein kinases, more work is needed to determine whether they regulate MSI1 activity (Vo et al. 2011; Arumugam et al. 2012; Vo et al. 2012a; Pasto et al. 2014). Interestingly, there is also evidence that MSI1, thyroid hormone, and the microtubule-associated protein tau are involved in a regulatory loop (Cuadrado et al. 2002). In Xenopus laevis, thyroid hormone administration induced xMsi1
expression in adult intestinal progenitor cells, but not in differentiated epithelial cells (Ishizuya-Oka et al. 2003). However, many of the mechanisms by which MSI proteins are regulated have not been established.

**Metabolic regulation of gene expression**

Metabolite homeostasis is required for normal cellular and systemic function, and the loss of homeostasis often leads to disease. A classic example is type I diabetes, where a precipitous change in blood glucose levels, caused by a failure in insulin signaling, leads to blindness, sores, infections on extremities, and nerve damage. To maintain homeostasis, organisms must sense and respond to changes in metabolic state by altering gene expression. However, the many of the mechanisms governing metabolite-mediated changes in gene expression remain a mystery.

There are numerous examples of metabolites regulating transcription factor activity (Wang et al. 1994; Kliewer et al. 1997). For example, the peroxisome proliferator-activated receptor (PPAR) family of nuclear hormone receptors bind directly to fatty acids and eicosanoids. Upon metabolite association, PPARs bind to specific DNA elements in promoters and act as transcription factors to regulate expression of a variety of genes, including several involved in lipid metabolism (Wang et al. 1994; Kliewer et al. 1997; Berger and Moller 2002). A number of potent synthetic PPAR ligands are used to treat diseases such as dyslipidemia and diabetes. In fact, metabolite responsive transcription factors are targeted in several therapeutic strategies, probably because of their importance in regulating gene
expression and their ability to bind to small molecule metabolite ligands (Zeng and Xie 2011).

Second messenger signaling also couples metabolic state to gene expression. Metabolites, hormones, gasses, and other small molecules bind and activate receptors on the cell surface to enact an intracellular signaling cascade that leads to the activation or repression of specific genes. An example of second-messenger signaling is the phosphoinositide 3-kinase (PI3K) pathway. PI3K catalyzes the formation of inositol trisphosphate (IP3) in response to a number of hormones and metabolites involved in cell growth and survival (Alcazar-Roman and Wente 2008; Kim et al. 2011). IP3 regulates genes involved in cell cycling and apoptosis via the mammalian target of rapamycin (mTOR) pathway (Kim et al. 2011). These complex signaling cascades have received extensive attention because activating or inhibiting metabolite and hormone surface receptors using small molecules can be a therapeutic strategy.

While bacteria widely use mechanisms of metabolic post-transcriptional regulation, there are relatively few examples in eukaryotes. Here, I will review some well-know bacterial metabolite sensors, as well as the few eukaryotic examples described to date.

**Riboswitches: metabolite-sensitive RNA elements**

Riboswitches are structured RNA elements often found in the 5′ untranslated region (UTR) of bacterial transcripts. They sense the concentration of specific
metabolites through direct interactions. Metabolite binding typically induces a change in riboswitch conformation that in turn alters transcription termination or translation initiation efficiency. As such, riboswitches provide a negative feedback loop that shuts down production of metabolic enzymes and related factors required to produce the associated metabolite.

The first riboswitch was described in 2002 by Breaker and coworkers, who reported that in *Escherichia coli*, transcripts of the cobalamin (vitamin B12) biosynthesis operon *btuB* directly bind cobalamin to induce a structural change that prevents ribosome binding (Nahvi et al. 2002). Since then, many riboswitches have been identified and categorized into classes based upon the type of ligand and the secondary structure formed upon association with that ligand. To date more than 20 classes of riboswitches sensitive to a variety of metabolites have been identified in bacteria. In contrast, only one class of riboswitch has been discovered in eukaryotes. A number of excellent reviews on bacterial riboswitches have been published recently (Winkler and Breaker 2005; Roth and Breaker 2009; Breaker 2012).

TPP Riboswitches in Eukaryotes

The thiamine-responsive TPP riboswitch is the most common bacterial riboswitch, and it has now been identified in plants, fungi, and archaea (Figure 1.6) (Kubodera et al. 2003; Cheah et al. 2007; Croft et al. 2007; Bocobza and Aharoni 2008). Hanamoto et al. discovered the first eukaryotic TPP riboswitch in the 5′ UTR
Figure 1.6
Figure 1.6: TPP-regulated alternative splicing of the NMT1 gene. A) Schematic of the 5' UTR intron structure. The thin line above the schematic denotes the predominant isoform in low TPP conditions. The thick line below the schematic denotes the pattern in high TPP conditions. B) The low TPP isoform contains a short 5' UTR with a single initiation codon, leading to efficient translation of NMT1. C) The high TPP isoform contains a longer 5’ UTR with two uORFs, leading to reduced translation initiation of the cognate NMT1 ORF. D) Chemical structure of TPP, sensed by the riboswitch motif in the 5' UTR intron.
of the *thiA* gene of the filamentous fungus *Aspergillus oryzae*, used in the production of sake (Kubodera et al. 2003). An intron is present in the *thiA* 5′ UTR that contains two motifs highly conserved in fungal thiamine biosynthesis genes. Reporter studies and northern analyses revealed that thiamine concentration controls the extent of 5′ intron splicing in a manner that depended on both conserved elements. The increase in unspliced 5′ UTR mRNA in the presence of thiamine correlates with a decrease in the expression of *thiA*. Subsequent studies by Breaker et al. identified sequences that match a TPP riboswitch consensus descriptor in thiamine biosynthesis genes from *Arabidopsis thaliana*, *Oriza sativa* (rice), *Poa secunda* (bluegrass), and the fungi *Neurospora crassa* and *Fusarium oxysporum* (Sudarsan et al. 2003). Structural probing revealed that the element from *Arabidopsis*, found in the *thiA* 3′ UTR, adopts a TPP riboswitch-like structure in the presence of thiamine, suggesting that the element comprises a bona fide riboswitch similar to those observed in bacteria.

In the case of the TPP riboswitch from the *Neurospora crassa* NMT1 5′ UTR (Figure 1.6), the riboswitch induces alternative splicing (Cheah et al. 2007). In the presence of elevated TPP, an upstream 5′-splice site is used leading to production of an mRNA with a longer 5′ UTR. The TPP riboswitch motif is present within the intron sequence that is normally spliced out. The longer mRNA includes an upstream open reading frame (uORF) that decreases translation of the downstream gene product by competing with the authentic translation start site. The data reveal
a mechanism whereby a structural change in the TPP riboswitch represses translation of the NMT1 gene via a change in the splice isoform ratio.

Additional studies of eukaryotic TPP riboswitches reveal a common role in modulating alternative mRNA splicing as a means of regulating gene expression (Thore et al. 2006; Croft et al. 2007; Thore et al. 2008). It is intriguing to note that the eukaryotic riboswitches work by modifying a eukaryote-specific process that occurs in the nucleus, rather than affecting transcription or translation initiation directly, as in bacteria. Although several examples of the TTP riboswitch have been characterized in eukaryotes, other riboswitch classes have not yet been identified. It is possible that they do not exist; however, it is likely that through refinement of bioinformatics and experimental techniques, additional examples will be found.

Identifying eukaryotic riboswitches

A number of groups have used riboswitch sequence and structure conservation to search bacterial genomes for additional riboswitches (Vitreschak et al. 2003; Abreu-Goodger et al. 2004; Barrick et al. 2004; Weinberg et al. 2007; Meyer et al. 2008). Merino and colleagues published an algorithm that sorted and paired operons with Clusters of Orthologous Groups (COGs) from the protein database (Abreu-Goodger et al. 2004; Abreu-Goodger and Merino 2005). Next they used successive rounds of Multiple EM for Motif Elicitation (MIME) followed by Motif Alignment and Search Tool (MAST) to identify over-represented motifs in each COG (Abreu-Goodger et al. 2004; Bailey et al. 2009). A subsequent publication described
the development of RibEx (riboswitch explorer), a tool that cross-references the genomes of non-redundant bacterial organisms to identify sequence motifs of putative riboswitches and other structural regulatory elements (Abreu-Goodger et al. 2004; Abreu-Goodger and Merino 2005). RibEx eliminates the use of COGs in favor of using 145 complete genomes in the initial alignment. Unfortunately, a similar tool does not yet exist for eukaryotes.

Riboswitch identification in eukaryotes has proven challenging due to the increased complexity of the regulatory mechanisms governing gene expression. In bacteria, riboswitch aptamer domains have high sequence conservation, while expression platforms can have disparate sequence, and the regulatory mechanisms are often inferred by their location. These expression platforms are usually positioned to block ribosome binding sites or prevent formation of transcription terminating stems. Identifying eukaryotic riboswitch expression platforms is complicated by a number of factors. First, eukaryotic genes are not grouped into operons. Second, riboswitches could theoretically work at the level of mRNA processing, nuclear export, stability, or translation. Eukaryotic riboswitches are therefore likely to have a heterogeneous nature at the structural and sequence levels. Third, most prokaryotic genomes are roughly 88% protein-coding, while higher organisms such as humans are estimated to have as little as 2% protein-coding sequence (Dinger et al. 2011). The increased size and complexity of eukaryotic genomes presents both challenges and possibilities for bioinformaticians. The large number of mRNA processing events identified in
eukaryotes, combined with the possibility that riboswitches may exist in non-coding RNAs to regulate gene expression in trans leads to almost limitless possibilities for eukaryotic riboswitch mechanisms. It is therefore not surprising that bioinformatics approaches have been difficult to develop for eukaryotic riboswitch identification.

Several groups have begun to search for non-coding RNAs within intergenic sequences of bacteria, archea, and eukaryotes. Although these searches are not specifically targeting riboswitch identification, it is possible that the methodology and the precedent set by this type of bioinformatic search may influence future riboswitch discovery in organisms other than bacteria. Breaker and colleagues have been optimizing clustering techniques to discover novel non-coding RNAs in bacteria (Weinberg et al. 2009). Similarly, a recent bioinformatic survey of the archaean *Pyrococcus abyssi* used clustering of sequence, primary structure, and secondary structure to identify conserved non-coding RNAs located in intergenic regions. This study revealed several elements that share features with the SAM-I and lysine bacterial riboswitches, although this remains to be experimentally tested (Phok et al. 2011). Although most recent work has focused on ribozyme discovery, the recent success in identifying large non-coding RNAs may enable future application of similar methods to the discovery of non-coding RNAs, including riboswitches, in eukaryotes.

Because only one riboswitch class has been identified in eukaryotes, it is not clear whether eukaryotic riboswitches are rare, or have not been discovered due to the increased complexity of higher organisms. If higher organisms truly lack of
riboswitches, perhaps proteins have taken over metabolite-sensing functionality. Numerous proteins regulate gene expression at the post-transcriptional level, and a handful of canonical RNA-binding proteins have demonstrated metabolite sensitivity. Additionally, over the past two decades several metabolic enzymes have demonstrated RNA-binding activity.

**Metabolite-sensitive RNA-binding proteins**

RNA-binding proteins alter gene expression by regulating pre-mRNA processing steps including alternative splicing, 5′ and 3′ end formation, and by controlling how the RNA sequence is edited. RNA binding proteins also regulate the subcellular location, stability, and translation efficiency of mature mRNAs. Recent work has shown that the activities of several RNA binding proteins are affected by the concentration of cellular metabolites, allowing cells to quickly respond to changes in the environment.

**Introduction to ADAR**

The increased complexity of higher organisms cannot be explained by the relatively small increase in the number of protein-coding genes relative to bacteria. Instead, genetic diversity can be attributed to an expansion in the number of gene products derived from each gene. Alternative pre-mRNA splicing and mRNA editing enable recoding of the information stored in the genome leading to production of multiple protein variants from a single gene.
One example of mRNA editing is the programmed conversion of select adenosine bases to inosine, catalyzed by enzymes termed adenosine deaminases that act on RNA (ADAR). ADARs were initially discovered in *Xenopus laevis* (Barraud and Allain 2012). Homologs have since been identified in most metazoa, but not in plants, fungi, or yeast (Barraud and Allain 2012). ADARs from different organisms are structurally similar; all contain several double-stranded RNA binding domains (dsRBDs) and a highly conserved C-terminal deaminase domain. ADARs deaminate specific adenosines to produce inosine at precise locations within an RNA sequence. Inosine pairs with cytidine and is therefore interpreted as a guanine base in biological settings. These editing events cause single-codon alterations, changes in alternative splicing, and modulation of mRNA stability (Farajollahi and Maas 2010).

RNA editing has been observed in pre-mRNA coding sequences, repetitive elements, and pri-miRNAs (Farajollahi and Maas 2010). Neurons use editing extensively to regulate specialized neurotransmitter receptors, ion channels, and surface protein isoforms (Maas et al. 2006). Mice stably expressing an shRNA that silences ADAR2 expression display increased neuronal sensitivity to restricted blood and nutrient supply and subsequent neuronal degeneration (Peng et al. 2006). ADAR activity is also reported to affect subcellular compartmentalization of certain mRNAs, and self-editing of ADAR mRNA is proposed to affect ADAR homo- and hetero-dimerization which in turn regulates mRNA editing efficiency (Rueter et al. 1999; Palladino et al. 2000; Sansam et al. 2003; Valente and Nishikura 2007;
ADAR editing is implicated in a number of disease states including schizophrenia, neuromuscular disorders, and certain cancers (Schmauss 2005; Maas et al. 2006; Peng et al. 2006).

The Bass group determined the high-resolution crystal structure of the catalytic domain of ADAR2 and discovered an inositol hexakisphosphate (IP6) molecule required for proper folding of the enzyme (Macbeth et al. 2005) (Figure 1.7). Subsequent experiments showed that an ADAR substrate is edited in wild type yeast expressing hADAR2, but not in yeast that lack the IPK1 gene and thus do not produce IP6 (Macbeth et al. 2005). The crystal structure revealed that IP6 is intimately associated with the protein, buried in an internal cavity lined with basic residues. IP6 was not added during purification or crystallization, indicating that endogenous IP6 co-purified with the protein from yeast (Macbeth et al. 2005). This suggests that the association between ADAR and IP6 is likely very tight, although to date the dissociation constant has not been measured.

It is probable that ADAR biogenesis is governed by intracellular IP6 concentration (Alcazar-Roman and Wente 2008). IP6 is abundant in healthy cells, where it is involved in signaling pathways including proliferation, differentiation, DNA repair, energy transduction, and RNA export (Kalam Shamsuddin and Bose 2012). IP6 serves as a phosphate donor in many signaling pathways. It is possible that IP6 levels influence the amount and stability of ADAR available to modify specific genes. This intriguing hypothesis requires additional biochemical and in vivo studies, especially in light of data that suggests that loss of ADAR editing
Figure 1.7
Figure 1.7: The structure of human ADAR2 bound to IP6. A) The surface of hADAR2 is rendered in mesh, revealing the deep internal cavity that coordinates IP6 (rendered in spheres). The structure was rendered from coordinate file 1ZY738. B) Chemical structure of IP6, sensed by ADAR.
activity can lead to cancer, while abundant IP6 can help prevent cancers (Maas et al. 2006; Kalam Shamsuddin and Bose 2012). Measurement of the extent of mRNA editing in the presence and absence of IP6, and determining the effect of editing on cellular physiology, will be necessary first steps towards demonstrating that ADAR is a biologically relevant sensor of IP6 concentration.

Metabolic enzymes with RNA-binding activity

Metabolic enzymes are integral to maintaining homeostasis because they catalyze the chemical reactions necessary to produce or use nutrients. Metabolic enzymes must be tightly regulated through feedback loops to ensure appropriate metabolic flux. Some metabolic enzymes have been reported to ‘moonlight’ as RNA-binding proteins, regulating gene expression at the post-transcriptional level in addition to catalyzing chemical reactions (Hentze and Preiss 2010). Hentze and Preiss termed this interplay between RNA, Enzymes and Metabolites the ‘REM’ phase of RNA regulation (Hentze and Preiss 2010). The following sections will introduce several examples of metabolic enzymes that act as RNA-binding proteins, and discuss the possibility that the REM network may be more extensive than is currently realized.

Cytosolic aconitase

Iron is an essential metabolite involved in processes including oxygen transport, cellular respiration and heme synthesis. Iron is also a cofactor for
numerous metalloenzymes involved in a wide range of biological processes. Failure to regulate intracellular iron levels results in iron deficiency or toxicity. Free iron catalyzes free radical generation, which can damage DNA, lipids, and proteins. Therefore, most intracellular iron exists in complex with enzymes and carrier proteins. Cytosolic aconitase is an iron-sensitive enzyme that conditionally doubles as an RNA-binding protein to regulate iron homeostasis (Klausner et al. 1993; Hentze and Kuhn 1996; Eisenstein 2000; Schneider and Leibold 2000; Theil and Eisenstein 2000; Cairo et al. 2002; Rouault 2002). Interestingly, the functionality of cytosolic aconitase in metabolic regulation was not understood until the discovery that iron-responsive protein-1 (IRP1) and cytosolic aconitase are the same protein with two distinct functions.

Aconitase is a tricarboxylic acid (TCA) cycle isomerase that catalyzes the conversion of citrate to isocitrate. Cytosolic aconitase (IRP1) has similar activity to that of mitochondrial aconitase. It was initially unclear why cells would contain two distinct copies of a functionally similar enzyme (Klausner and Rouault 1993). When intracellular iron is low, IRP1 binds iron-responsive RNA elements (IREs) in iron regulatory genes to modulate their expression. The iron regulatory genes include the iron storage factor ferritin, iron uptake factors like TfR and DMT-1, and the iron export factor ferroportin (Theil and Eisenstein 2000). When IRP1 was purified, sequenced, and cloned, it was found to have approximately 30% homology to *Saccharomyces cerevisiae* and porcine mitochondrial aconitase (Rouault et al. 1989; Neupert et al. 1990; Rouault et al. 1990; Hentze and Argos 1991; Rouault et al. 1991;
Hirling et al. 1992). Subsequent work showed that human recombinant IRP1 has aconitase activity in the presence of iron, and bovine cytosolic aconitase binds to IREs in low iron conditions (Kaptain et al. 1991; Haile et al. 1992; Kennedy et al. 1992). Further experiments from the Hentze and Klausner groups confirmed that cytosolic aconitase and IRP1 are the same protein, with distinct functional activities depending upon intracellular iron concentrations.

In the presence of high intracellular iron the 4Fe-4S catalytic cluster assembles to enable aconitase functionality (Dupuy et al. 2006). When intracellular iron decreases, the protein undergoes an allosteric change to reveal an iron responsive element (IRE)-binding site, enabling functionality as the RNA-binding protein IRP1 (Figure 1.8) (Walden et al. 2006). The IRE is a highly conserved hairpin present in the 3’ and 5’ UTRs of several genes related to iron homeostasis (Henderson et al. 1994). Interaction between IREs and IRP1 induces repression of ferritin mRNA translation and transferritin receptor mRNA stabilization (Butt et al. 1996).

How important is iron concentration sensing by IRP1? Interestingly, IRP1-/- mice develop normally without an apparent phenotype (Meyron-Holtz et al. 2004). In contrast, IRP2-/- mice display microcytic anemia, increased red cell protoporphyrin IX levels, and neurodegeneration, but are otherwise normal (LaVaute et al. 2001; Cooperman et al. 2005). While mice that lack IRP1 or IRP2 are viable, mice that lack both fail in embryogenesis prior to implantation (Smith et al. 2006). The mouse mutants reveal IRP activity is required for animal viability,
Figure 1.8
Figure 1.8: The two forms of IRP1. A) structure of IRP1 in the iron replete state adopts the canonical aconite fold. The protein structure is rendered as a cartoon, the 4Fe-FS cluster is rendered as spheres. The structure was rendered from coordinate file 2B3X58. B) In iron deficient state, IRP1 adopts an alternative conformation that binds to a specific stem loop RNA structure (red). The structure was rendered from coordinate file 3SNP59. C) Chemical structure of the 4Fe-4S cluster sensed by IRP1.
highlighting the importance of regulating iron homeostasis at the post-transcriptional level. They also suggest that IRP2 can compensate for loss of IRP1, while IRP1 is not sufficient to compensate for loss of IRP2 (Smith et al. 2006). Additional work is needed to understand the basis for this difference.

GAPDH

GAPDH has traditionally been labeled a housekeeping protein because it is involved in glycolysis. However, several studies indicate that its biological activity is more diverse (Sirover 1999; Nicholls et al. 2012). GAPDH has been shown to bind a range of RNA species, including mRNA, tRNA, rRNA and viral RNA, and may participate in such activities as RNA export and regulation of RNA stability (Ryazanov 1985; Singh and Green 1993; Dollenmaier and Weitz 2003). GAPDH directly associates with AU-rich elements (AREs) present in the 3' UTRs of number of RNA species (Nagy and Rigby 1995). In most cases, however, the physiological relevance of RNA-binding activity by GAPDH has not been validated.

One proposed role for GAPDH RNA-binding activity lies in the regulation of cytokine and endothelin expression. Cytokines and endothelins are small bioactive proteins that modulate the immune response and affect blood vessel constriction, respectively. GAPDH stabilizes mRNA encoding colony-stimulating factor 1 (CSF1), which is implicated in tumorigenesis (Lin et al. 2001). Conversely, GAPDH association promotes turnover of transcripts encoding the endothelial vasoconstrictor endothelin (ET) (Rodriguez-Pascual et al. 2008). The mechanism of
GAPDH-mediated regulation of mRNA stability is not known. There is growing evidence that GAPDH regulates lymphokine translation by binding lymphokine transcripts in polysomes (Nagy and Rigby 1995). Again, the mechanism of regulation is not known.

In vivo, GAPDH is predicted to exist in two conformations: an RNA-binding form that is not active in glycolysis, and an NAD+-binding form that is active in glycolysis (Nagy et al. 2000). The ratio of the two conformations may be regulated by the local concentration of NAD+, NADH, and ATP (Nagy et al. 2000; Arutyunova et al. 2003). Oxidation state is also predicted to differentially affect the RNA- or NAD+-binding activity of GAPDH (Nagy et al. 2000; Arutyunova et al. 2003). As such, GAPDH may sense the concentration of oxidized NAD+ in order to control expression of mRNA targets. More work is needed to understand the mechanism by which GAPDH converts from an active metabolic enzyme to an RNA-binding factor, and exactly how changes in metabolic state affect the various transcripts with which it associates.

**Identifying metabolite-sensing proteins with RNA-binding activity**

Recent advances in proteome-wide assays are enabling rapid identification of both metabolite-sensitive proteins with putative RNA-binding activity and RNA-binding proteins with putative metabolite-sensing activity. Several techniques that survey the RNA interactome and/or proteome can be used to assess changes upon alteration of metabolic state. These high-throughput approaches, combined with
For proteins with known RNA-binding activity, several methods are available to identify associated transcripts in cells. These include HITS-CLIP, PAR-CLIP, and iCLIP, each of which rely on crosslinking of the protein to associated RNAs in cells, followed by immunoprecipitation and deep sequencing to identify the associated RNA (Licatalosi et al. 2008; Hafner et al. 2010; König et al. 2011). To identify novel mRNA-binding proteins, Hentze, Krijgsveld, and colleagues adapted the HITS-CLIP and PAR-CLIP methods to monitor the entire mRNA interactome (Castello et al. 2012). In this technique, termed “interactome capture,” proteins are first cross-linked to associated RNAs in cells. Next, cells are lysed and oligo d(T) beads are used to capture polyadenylated RNA species. The protein interactome is then identified through mass spectrometry. In HeLa cells, over 300 novel RNA-binding proteins were identified by this approach, 46 of which are enzymes involved in intermediary metabolism. This indicates that metabolic enzymes that also bind to RNA are more widespread than previously thought. Because these enzymes function in intermediary metabolic pathways, they are logical candidates for coupling metabolite sensing to RNA regulation.

It will be more complicated to identify metabolite-responsive RNA-binding proteins that are not established metabolic enzymes. It is possible that by employing the methods outlined above while manipulating metabolite concentrations, some proteins will display altered RNA-binding activity and/or
target recognition. The most obvious candidates to use in this type of experiment are proteins that have already been shown to regulate genes involved in metabolite homeostasis. Proteins demonstrating changes in RNA-binding activity would then require \textit{in vitro} and \textit{in vivo} experimentation to determine the mechanisms employed to sense and respond to metabolic state.

In Chapter II, I outline a small molecule screen in which the intermediary metabolite oleic acid was found to inhibit RNA-binding activity of Msi1. This surprising result became the focus of my thesis research, in which I characterized the biochemical and biophysical nature of inhibition (Chapter II) and linked Msi1 activity to lipid homeostasis and metabolism (Chapter III). I will present models to explain the biological nature, mechanism, and consequences of fatty acid inhibition of Msi1 in Chapters III and IV.
Chapter II: The small molecule metabolite oleic acid is an allosteric inhibitor of Musashi-1 RNA-binding activity
Abstract

Stem and progenitor cells have the ability to differentiate into a wide variety of specialized cells and are often highly proliferative. These cells enable development in immature organisms, and cellular repair and regeneration in mature organisms. However, when the mechanisms governing progenitor cell proliferation fail, cancerous tumors form. In this chapter I present data for a small molecule screen and follow-up experiments, in which I show that RNA-binding by the stem cell translation regulator Musashi-1 (MSI1) is allosterically inhibited by 18-22 carbon ω-9 monounsaturated fatty acids. The fatty acid binds to the N-terminal RNA Recognition Motif (RRM) and induces a conformational change that prevents RNA association. Musashi proteins are critical for development of the brain, blood, and epithelium, and they are elevated in cancers of these lineages. MSI1 is therefore a promising therapeutic target in the ongoing search for cancer treatments.
Introduction

The RNA-binding protein Musashi-1 (MSI1) is expressed in stem and progenitor cells of neural and epithelial lineage. In the central nervous system, MSI1 is expressed in astrocytes and committed glial and neural progenitor cells, but not in mature neurons and oligodendrocytes (Figure 1.1) (Kaneko et al. 2000; Dobson et al. 2008). In histological studies of neural and epithelial tissues, MSI1 is routinely used as a marker for stem and progenitor cells (Johnson et al. 2010). Analysis in mice and primary cells shows that Msi1 regulates neural development. Msi1-/- knockout mice are uncoordinated, ataxic, develop hydrocephaly, and die within 1-2 months after birth (Sakakibara et al. 2002). Their brains are small, contain an expansion of early lineage progenitor cells, and display fewer mature cell types than normal (Sakakibara et al. 2002). Embryonic neurospheres cultured from Msi1-/- mouse brains have a reduced capacity to differentiate into mature neurons and oligodendrocytes (Sakakibara et al. 2002). In primary oligodendrocyte progenitor cells, Msi1 promotes progenitor cell survival and prevents differentiation into mature oligodendrocytes (Dobson et al. 2008). The phenotype and expression pattern reveal that Msi1 plays an early role in regulating neurogenesis and gliogenesis.

MSI1 contains two RRMs and is homologous to Drosophila melanogaster Musashi, a post-transcriptional regulatory protein that guides external sensory bristle patterning in flies (Sakakibara et al. 1996). In vitro SELEX experiments identified a series of aptamer sequences that bind to Msi1 (Imai et al. 2001). Visual
inspection identified a consensus sequence (G/A)U_{1,3}AGU that was present in most but not all of the aptamers. A number of Msi1 targets have been identified by co-immunoprecipitation, including NUMB, a repressor of NOTCH signaling. Numb transcripts harbor MSI1 consensus elements in the 3’ UTR (Imai et al. 2001). Msi1 interacts with the Numb 3’ UTR in vitro, and Numb mRNA co-immunoprecipitates with MSI1 in transiently transfected NIH 3T3 cells. Overexpression of MSI1 in NIH 3T3 cells decreases NUMB protein levels without affecting Numb mRNA and reduces the expression of a luciferase reporter in a 3’ UTR dependent manner (Imai et al. 2001). Together, the results show that MSI1 negatively regulates Numb mRNA translation. In contrast, xMsi1 acts as a translational activator in Xenopus laevis oocytes, where it modulates cell cycle progression by regulating mRNA encoding the proto-oncogene Mos (Charlesworth et al. 2006).

MSI1 also promotes proliferation of numerous cancers of the brain and epithelial tissues (Toda et al. 2001; Hemmati et al. 2003; Yokota et al. 2004; Sanchez-Diaz et al. 2008; Sureban et al. 2008). MSI1 depletion in medulloblastoma and colorectal tumors results in decreased proliferation and increased apoptosis (Sanchez-Diaz et al. 2008; Sureban et al. 2008). In colorectal tumors, MSI1 depletion is accompanied by inhibition of Notch-1 and upregulation of p21\textsuperscript{WAF1}, a MSI1 target involved in cell cycle regulation (Battelli et al. 2006; Sureban et al. 2008). Musashi-2 (MSI2) is 69% identical to MSI1 protein and is expressed in a partially overlapping set of tissues (Figure 1.2) (Sakakibara et al. 2002). MSI2 regulates hematopoiesis and is involved in acute myeloid leukemia (Ito et al. 2010; Kharas et al. 2010). In
myeloid leukemia cells, MSI2 is highly expressed, and depletion results in decreased proliferation and increased apoptosis (Kharas et al. 2010). The crisis phase of myeloid leukemia is marked by low NUMB expression (Ito et al. 2010). Loss of MSI2 restores NUMB expression and impairs the blast crisis phase of myeloid leukemia (Ito et al. 2010). Ultimately, MSI2 expression levels are directly correlated with poor prognosis in myeloid leukemia patients (Kharas et al. 2010).

Because of the importance of Musashi family proteins in stem and cancer cell proliferation, I sought to identify a small molecule inhibitor of MSI1 RNA-binding activity. Four inhibitors were identified in a screen of more than 30,000 compounds, one of which is the intermediary metabolite oleic acid. Here I present data to characterize the specificity and mechanism of oleic acid inhibition.

**Results**

*Small molecule screen to identify inhibitors of Musashi-1*

To screen for small molecule inhibitors of MSI1 RNA-binding activity, I developed an in vitro assay pipeline amenable to high throughput measurements. First, I tested the ability of a purified, his6-tagged Msi1 dual RRM construct (amino acids 7-192, Figure 2.1) to bind a fragment of a previously identified SELEX aptamer (CCCR005) (Imai et al. 2001) using two quantitative assays: fluorescence electrophoretic mobility shift (F-EMSA) and fluorescence polarization (FP, Figure 2.1) (Pagano et al. 2011). Msi1 binds with high affinity to the aptamer fragment, which contains two copies of the consensus sequence. Next, I optimized the FP assay
Figure 2.1

A

B

C

D

5'-AGC GUA GUA UUU AGU UCG-F1-3'

Kd, app (EMSA) = 14 ± 1 nM
Hill = 1.3 ± 0.2
Kd1, app (FP) = 14 ± 5 nM
Kd2, app (FP) = 422 ± 45 nM
Figure 2.1: His6-Msi1 purification and activity test. (A) Coomassie-stained SDS page gel shows that recombinant MSI1 is purified to greater than 95% over a 3-column purification protocol. (B-D) EMSA and FP of MSI1 binding to RNA aptamer CCCR005 (AGCGUUAGUUAAUUGUCG). EMSA data (red line) were fit to the Hill equation where all shifted species were fit as an aggregate. FP data (black line) were fit to a two-site binding model.
for use in high throughput screens. This assay was used to screen two small molecule libraries: the 1280-compound Sigma Library of Pharmacologically Active Compounds (LOPAC) and the 30,000-compound Chembridge library (Figure 2.2A, Table 2.1A). Inhibitors identified in the screen were validated by dose response measurements using both the FP and F-EMSA assays (Figure 2.2).

Four candidate inhibitors were identified. The weakest inhibitor was the Chembridge compound 7409829 ($K_{i,\text{app, FP}} = 15 \pm 2.8 \, \mu M$, $K_{i,\text{app, F-EMSA}} = 54 \pm 22 \, \mu M$). The most potent inhibitor was aurintricarboxylic acid (ATA, $K_{i,\text{app, FP}} = 230 \pm 30 \, nM$, $K_{i,\text{app, F-EMSA}} = 1.55 \pm 0.14 \, \mu M$), a compound that readily polymerizes in aqueous solution to form a polyanion. This compound has been identified in many high throughput small molecule assays as a non-specific inhibitor of protein-nucleic acid interactions (Lam et al. 1995). The next inhibitor was GW7647, a PPARα agonist (Berger and Moller 2002). PPARα is a nuclear hormone receptor that is activated by long chain unsaturated fatty acids (Gottlicher et al. 1992; Bocos et al. 1995; Forman et al. 1997; Kliewer et al. 1997). The $K_{i,\text{app}}$ for this compound was $6.5 \pm 0.4 \, \mu M$ by FP and $21 \pm 0.8 \, \mu M$ by F-EMSA. The final inhibitor was oleic acid, an eighteen-carbon monounsaturated fatty acid with one double bond located nine carbons from the aliphatic—omega—end of the molecule (18:1, ω-9). The apparent $K_i$ of this compound was $1.2 \pm 0.4 \, \mu M$ by FP and $1.4 \pm 0.7 \, \mu M$ by F-EMSA, and I observed decreased affinity for the RNA aptamer with increasing concentrations of oleic acid (Figure 2.2 and 2.3A). The $K_{i,\text{app}}$ is approximately two orders of magnitude below the
Figure 2.2

A

\[ \text{Aspartic acid} \]
\[ K_i \text{ app} = 1.55 \pm 0.14 \mu M \]
\[ \text{GW7647} \]
\[ K_i \text{ app} = 21 \pm 0.8 \mu M \]
\[ \text{Chembio 7409629} \]
\[ K_i \text{ app} = 54 \pm 2.2 \mu M \]

B

\[ \text{Olac acid (18:1 \&-9 trans)} \]
\[ K_i \text{ app} = 1.4 \pm 0.7 \text{ mM} \]
\[ \text{Eladic acid (18:1 \&-9 trans)} \]
Figure 2.2: Small molecule screen. (A) Assay scheme for the inhibitor screen and F-EMSA dose responses with hits identified from the small molecule screens (A) and oleic and elaidic acid (B). Each gel is one representative experiment of at least three independent experiments. No compound and no protein lanes identify the position of bound and free RNA migration, respectively.
**Table 2.1**

<table>
<thead>
<tr>
<th>Compound Name</th>
<th>CD</th>
<th>Ex trem Score</th>
<th>FP</th>
<th>K&lt;sub&gt;i&lt;/sub&gt; app (µM)</th>
<th>F-EMSA</th>
<th>K&lt;sub&gt;i&lt;/sub&gt; app (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chembridg 740629</td>
<td>9945</td>
<td>0.046</td>
<td>15 ± 2.1</td>
<td>54 ± 0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acyl-CoA thioester</td>
<td>25525</td>
<td>0.053</td>
<td>0.23 ± 0.03</td>
<td>15 ± 0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GW7647</td>
<td>3397731</td>
<td>−0.028</td>
<td>0.5 ± 0.4</td>
<td>21 ± 0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oleic Acid</td>
<td>445639</td>
<td>0.000</td>
<td>1.2 ± 0.4</td>
<td>1.4 ± 0.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th>Compound Name</th>
<th>Structure</th>
<th>Code</th>
<th>FP</th>
<th>K&lt;sub&gt;i&lt;/sub&gt; app (µM)</th>
<th>F-EMSA</th>
<th>K&lt;sub&gt;i&lt;/sub&gt; app (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oleic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>18:1 w-9</td>
<td>1.2 ± 0.4</td>
<td>1.4 ± 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eicosenoic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>20:1 w-9</td>
<td>1.2 ± 0.4</td>
<td>1.7 ± 0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enoic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>22:1 w-9</td>
<td>0.54 ± 0.2</td>
<td>0.82 ± 0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nervonic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>24:1 w-9</td>
<td>47 ± 30</td>
<td>23 ± 0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palmitoleic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>16:1 w-7</td>
<td>0.3 ± 0.5</td>
<td>13 ± 0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linoleic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>18:2 w-6,9</td>
<td>2.2 ± 0.2</td>
<td>1.2 ± 0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arachidonic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>20:4 w-6, 9, 12, 15</td>
<td>3.0 ± 0.2</td>
<td>1.1 ± 0.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olel-CoA</td>
<td><img src="image" alt="Structure" /></td>
<td>(18:1 w-9)</td>
<td>0.1 ± 0.3</td>
<td>1.0 ± 0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enoyl-CoA</td>
<td><img src="image" alt="Structure" /></td>
<td>(18:1 w-9)</td>
<td>4.1 ± 0.9</td>
<td>0.6 ± 0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retinoic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>(18:1 w-6)</td>
<td>No Inh</td>
<td>18 ± 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oleamide</td>
<td><img src="image" alt="Structure" /></td>
<td>(18:1 w-9)</td>
<td>No inh</td>
<td>No inh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythro erylate</td>
<td><img src="image" alt="Structure" /></td>
<td>(18:1 w-9)</td>
<td>No inh</td>
<td>No inh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-Methylumbelliferyl oleate</td>
<td><img src="image" alt="Structure" /></td>
<td>(18:1 w-9)</td>
<td>No inh</td>
<td>No inh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elaidic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>18:1 (trans)</td>
<td>No inh</td>
<td>No inh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stearic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>18:0</td>
<td>No inh</td>
<td>No inh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palmitic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>16:0</td>
<td>No inh</td>
<td>No inh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myristic acid</td>
<td><img src="image" alt="Structure" /></td>
<td>14:0</td>
<td>No inh</td>
<td>No inh</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.1: Structure-activity relationship analysis demonstrates specificity of inhibition. (A) Small molecule screen hits. Compound ID (CID) refers to each compound's LOPAC identification number. Screen scores were calculated by normalizing the polarization value of each compound to the no protein and no compound controls, as described in the supplemental methods. After the screen was complete, compounds that scored as hits were confirmed by FP and F-EMSA dose response experiments. Apparent inhibition constants (Ki, app) are the average and standard deviation of at least three independent experiments. (B) The code = carbon number : number of double bonds, followed by the position of the double bonds from the aliphatic end of the fatty acid. Where a fatty acid is modified, the parental fatty acid numerical code is given in parentheses for comparison purposes. FP and F-EMSA dose response results are reported as the average and standard deviation of at least three independent experiments.
Figure 2.3

A

B
Figure 2.3: Dissociation constant and CMC quantification. (A) Msi1 displays decreased affinity for an RNA aptamer upon addition of oleic acid. Apparent dissociation constants were determined by plotting fluorescence polarization as a function of Msi1 protein concentration and fitting the data to the Hill equation. 0 µM oleic acid: Kd, app = 16.3 ± 1.2 nM; 1 µM oleic acid: Kd, app = 18.1 ± 2.6 nM; 10 µM oleic acid: Kd, app = 40.5 ± 3.5 nM; 0 µM oleic acid: Kd, app > 2000 nM. (B) CMC determination by N-phenyl-1-naphthylamine (NPN) fluorescence in equilibration buffer (pH 8.0). Segmented linear regression was used to determine the breakpoint between baseline and micelle-associated NPN fluorescence. The value of the CMC presented is the average and standard deviation from three experiments.
critical micelle concentration (CMC) in equilibration buffer (pH 8.0) as measured by N-phenyl-1-naphthylamine (CMC ≥ 75 ± 8 μM, Figure 2.3B).

Oleic acid has been screened in 754 bioassays reported in the PubChem database. Of these, oleic acid scored as positive in 4% of the assays. It should be noted that this figure overestimates the hit rate because it includes multiple bioassays that target the same protein. For example, oleic acid scored as a positive in eleven bioassays targeting fatty acid binding proteins (FABP) and seven assays that target membrane-associated potassium channels. Both are known to be sensitive to fatty acids (Capaldi et al. 2006; Boland and Drzewiecki 2008). Other proteins responsive to oleic acid include fatty acid synthase, estrogen synthase, factor VIIa complex, and enterotoxin. It also scored positive in a screen for membrane permeant biomolecules. In total, only 14 unique proteins are responsive to oleic acid. This suggests that oleic acid inhibition is specific.

Msi1 is specifically inhibited by 18-22 carbon ω-9 monounsaturated fatty acids

Oleic acid is the most abundant fatty acid in body fat and is produced by mature oligodendrocytes during myelination (Martinez and Mougan 1998). Cells can produce almost any fatty acid by modifying existing fatty acids through metabolic pathways. Because oleic acid is structurally related to a large number of fatty acids, I obtained a library of fatty acids and analogs to assess the specificity of inhibition (Figure 2.2B, Table 2.1). First, I measured inhibition by longer omega-9
monounsaturated fatty acids using the FP and F-EMSA dose response assays.

Eicosenoic acid (20:1, \(\omega-9\)) inhibited Msi1 with a potency similar to oleic acid (\(K_{i,\text{app,FP}} = 1.2 \pm 0.4 \mu M, K_{i,\text{app,F-EMSA}} = 1.7 \pm 0.6 \mu M\)). Erucic acid (22:1, \(\omega-9\)) was a stronger inhibitor (\(K_{i,\text{app,FP}} = 640 \pm 150 \, \text{nM}, K_{i,\text{app,F-EMSA}} = 820 \pm 30 \, \text{nM}\)), and nervonic acid (24:1, \(\omega-9\)) inhibited more weakly (\(K_{i,\text{app,FP}} = 47 \pm 25 \mu M, K_{i,\text{app,F-EMSA}} = 23 \pm 8 \mu M\)).

Second, I assessed truncations and modifications of the aliphatic end of the fatty acid. Removing two carbons (palmitoleic acid, 16:1, \(\omega-7\)) had a moderate effect on inhibition (\(K_{i,\text{app,FP}} = 5.3 \pm 0.5 \mu M, K_{i,\text{app,F-EMSA}} = 12 \pm 0.9 \mu M\)). The presence of a hydroxyl group at carbon 12 (ricinoleic acid, 12-hydroxy-oleic acid) had a stronger effect; inhibition is barely detectable by FP and reduced 15-fold in F-EMSA measurements (\(K_{i,\text{app,F-EMSA}} = 18 \pm 9\)). Third, I assessed esterification or other modification of the carboxylate group. Oleamide, ethyl oleate, and 4-methylumbelliferyl oleate failed to inhibit RNA-binding. In contrast, the presence of a Coenzyme A (CoA) substituent was apparently tolerated, although some deacylation of acyl-CoA stocks was apparent by thin layer chromatography. Fourth, I assessed the requirement for the \(\omega-9\) double bond. Stearic (18:0), palmitic (16:0), and myristic (14:0) acids failed to inhibit Msi1 RNA-binding activity, indicating that the \(\omega-9\) double bond is required. Surprisingly, the orientation of the double bond is also critical. Elaidic acid (18:1, \(\omega-9\) \textit{trans}) has the same molecular weight as oleic acid, and nearly identical refractive index and molar aqueous solubility (~10 mM at pH 8.0), but its \(\omega-9\) double bond is \textit{trans} rather than \textit{cis}. Elaidic acid did not inhibit
Msi1 (Figure 2.2B). Linoleic (18:2, ω-9, ω-6) and arachidonic (20:4, ω-12, ω-9, ω-6, ω-3) polyunsaturated fatty acids also inhibit Msi1, but with a weaker apparent inhibition constant (Linoleic acid: $K_{i,\text{app},FP} = 2.2 \pm 0.2 \mu M$, $K_{i,\text{app},F-\text{EMSA}} = 1.2 \pm 0.03 \mu M$; Arachidonic acid: $K_{i,\text{app},FP} = 3.0 \pm 0.2 \mu M$, $K_{i,\text{app},F-\text{EMSA}} = 1.1 \pm 0.3 \mu M$). Together, this data indicates that omega-9 cis unsaturated fatty acids between 18 and 22 carbons specifically inhibit Msi1 RNA-binding activity and identify erucic acid as the most potent inhibitor (Table 2.1). MSI2 was inhibited with similar specificity (Figure 2.4).

*Oleic acid directly interacts with the RRM1 of Msi1 to affect inhibition*

In principle, inhibitory fatty acids could inhibit by interacting with either Msi1 or its RNA target. If the fatty acid bound to the RNA-binding domain of Msi1, the interaction might alter the local environment of its single tryptophan (W29), leading to a measurable change in the intrinsic fluorescence (Vivian and Callis 2001). To test this hypothesis, I titrated oleic acid or elaidic acid (trans-oleic acid) into recombinant Msi1 protein or n-acetyl-L-tryptophanamide (NATA) and measured tryptophan fluorescence. NATA is a compound that contains a single tryptophan as a control for non-specific association and fluorescence quenching. Titration of oleic acid, but not elaidic acid, strongly quenched tryptophan fluorescence and altered the emission intensity curve shape from 300 to 400 nm (Figure 2.5 and figure 2.6). However, neither compound affected the fluorescence of NATA. In my buffer system, the maximal emission for both NATA and Msi1 is 350
Figure 2.4: RNA binding specificity and inhibition by specific fatty acids is conserved in MSI2. (A) MSI2 binds the RNA aptamer CCCR005 with similar affinity to that of Msi1 by both FP and F-EMSA. The no protein control lane defines the position of free RNA. Data are the average and standard deviation of three independent experiments. (B) MSI2 is specifically inhibited by oleic acid and eicosenoic acid in FP and F-EMSA dose response experiments. No compound and no protein controls define the position of bound and free RNA respectively. Data are the average and standard deviation of three independent experiments.
Figure 2.5

Oleic Acid

K_d,app = 2.6 ± 1 μM

Elaidic Acid

Trp Fluorescence (Arbitrary Units)

[Oleic or Elaidic Acid] M
Figure 2.5: Msi1 tryptophan fluorescence at 350 nm as a function of oleic and elaidic acid. The $K_d$, $K_i$, and Hill parameters are the average and standard deviation of three independent replicates.
Figure 2.6
Figure 2.6: The Msi1 tryptophan fluorescence spectrum changes as a function of inhibitor titration. The tryptophan fluorescence emission spectrum is strongly quenched upon oleic acid (A) and erucic acid (C) but not elaidic acid (B) addition. (D) A tryptophan-containing control N-acetyl-tryptophanamide (NATA) does not change upon compound addition. Data from one representative experiment for each compound concentration are shown.
nm (Figure 2.6). I fit the emission at 350 nm to a quadratic, bimolecular association model in order to determine the apparent dissociation constant. The $K_{d,\text{app}}$ of oleic acid was $2.6 \pm 1 \mu M$, essentially identical to the $K_{i,\text{app}}$ determined by FP and F-EMSMA dose response experiments (Table 2.1, Figure 2.5A). These data are consistent with inhibition resulting from a direct association of oleic acid with the Msi1 protein.

The NMR structures of Msi1 RRM1 and RRM2 in the absence and presence of RNA show that both domains adopt the canonical RRM fold in which an anti-parallel beta sheet is buttressed by two alpha helices (Figure 2.7) (Nagata et al. 1999; Miyanoiri et al. 2003; Ohyama et al. 2011). In RRM domain proteins, the beta sheet surface typically forms the RNA-binding platform (Figure 2.7) (Kielkopf et al. 2004; Clery et al. 2008). In the RNA-bound structure, several conserved amino acids located on the beta sheet surface and loops make sequence-specific contacts to the RNA (Ohyama et al. 2011). Notably, W29 directly contacts RNA by stacking on the first purine nucleotide. Tryptophan fluorescence experiments reveal that oleic acid binding causes W29 fluorescence quenching (Figure 2.5), suggesting that binding changes the environment surrounding this amino acid. As such, fatty acid binding may inhibit Msi1 RNA-binding by an allosteric mechanism.

To assess whether oleic acid association induces a change in Msi1 secondary structure, a collaborator collected circular dichroism spectra as a function of fatty acid treatment. We observe a decrease in mean residue ellipticity centered around 220 nm upon treatment with oleic acid but not elaidic acid (Figure 2.8). To determine whether the spectral changes correspond to a change in oligomerization
Figure 2.7
Figure 2.7: Musashi-1 RRM1 and RRM2 NMR structures. Ribbon model of Msi1 RRM1 (top). Space-filling model of Msi1 RRM1 (middle) and RRM2 (bottom) (Nagata et al. 1999; Miyanoiri et al. 2003). Left, β-sheet surface, right, α-helical surface. Conserved phenylalanines and W29 are green. Lysine and arginine residues are blue. A hydrophobic pocket exists on the RRM1 α-helical surface.
Figure 2.8

A

B

C

Wavelength (nm)

Relative Concentration (AU)

Relative Concentration (AU)

Wavelength (nm)

S-Value

 MSI-1 only

 MSI-1 with 50 µM Oleic acid

 MSI-1 with 50 µM Elaidic acid
Figure 2.8: CD and AUC characterization of oleic acid association. (A) CD spectra of Msi1 RRM1 in the presence of oleic (top) or elaidic acid (bottom). (B) Difference spectra calculated for the far-UV CD spectra of Msi-1 collected at increasing oleic acid concentrations relative to the spectrum in the absence of oleic acid. Different colors correspond to different concentration of oleic acid: red is 3 μM, purple is 6 μM, blue is 12.5 μM, green is 25 μM and black is 50 μM. The inset shows the same difference spectra calculated for the CD spectra of Msi1 with increasing elaidic acid concentration relative to the spectrum collected in the absence of elaidic acid. The same color scheme applies to the different elaidic acid concentrations. (C) Envelope traces of the van Holde-Weischet analysis for analytical ultracentrifugation experiments of Msi1 alone (top), with oleic acid (middle) and with elaidic acid (bottom). The predominant species sediments where monomeric Msi1 would be expected, and there is no significant change in the sedimentation profile after addition of oleic or elaidic acid. Data are representative traces from one of three independent experiments.
state, I performed velocity sedimentation analytical ultracentrifugation experiments (Demeler et al. 2011). There was no significant change in the velocity sedimentation profile of Msi1 by van Holde-Weischet analysis after adding either oleic acid or elaidic acid, and the predominant species remains monomeric (Figure 2.8C) (Demeler and van Holde 2004; Demeler et al. 2011). These data indicate that oleic acid binding alters the secondary structure of Msi1, but this transition does not appear to induce aggregation at concentrations below the CMC.

The published NMR structures of Msi1 in its apo form show a hydrophobic cavity on the alpha-helical surface of RRM1, opposite the RNA-binding surface (Figure 2.7) and adjacent to W29 (Nagata et al. 1999; Miyanoiri et al. 2003). Msi1 RRM2 does not have this feature. I hypothesized that this cavity comprises the fatty acid binding site. To test this idea, I first asked whether RRM1 is sufficient for ω-9 fatty acid inhibition. I purified his6-tagged Msi1 RRM1 (amino acids 7-103) and used FP and F-EMSA to determine whether RRM1 is sufficient for both RNA binding and inhibition by oleic acid. RRM1 bound to the aptamer RNA with an apparent $K_d$ of 75.2 ± 10 nM by FP (Figure 2.9A). Addition of oleic acid but not elaidic acid inhibits RNA-binding activity with similar $K_{i, app}$ compared to the full RNA-binding domain (Figure 2.9B). Thus, RRM1 is sufficient for RNA recognition and fatty acid inhibition.

Next, I prepared a $^{15}$N labeled sample of RRM1 for NMR spectroscopy. Our collaborators collected an $^{15}$N-$^1$H HSQC spectrum, and titrated aptamer RNA to identify amide proton chemical shift differences associated with RNA binding (Figure 2.9C). The majority of the chemical shift differences map to the β-sheet
Figure 2.9

A

Macashi 1 RRM1

5'-AGCCGUAAGUAAUUAGUJCG-F3'-3'

$K_{d, app} = 7.5 \pm 10 \text{nM}$

[Graph showing polarization vs. [Protein] (M)]

B

$K_{d, app} = 6.8 \pm 2 \mu\text{M}$

[Graph showing polarization vs. [Cleic acid] (M)]

C

[Chemical shift plots with residues labeled]

D

[Graph showing $K_p$ vs. [CM] / [Meth] with lines and data points]

E

[Graph showing $l_{A0}$ vs. Residue number with data points and trend lines]

F

[Structural models of RRM1 Beta Sheet with labeled residues]
Figure 2.9: NMR characterization of Msi1 RRM1. (A-B) Recombinant His6-tagged Msi1 RRM1 binds RNA aptamer CCCR005 (A) and is inhibited by oleic acid (B). FP data is reported as the average and standard deviation of three independent experiments. (C) 2D 1H-15N HSQC spectra of Msi1 free (red) and bound to RNA (blue) at pH 7 and 25 °C. Selected resonances are labeled and enlarged in the insets. Crosspeaks are colored according to the increasing concentration of added RNA. (D-F) NMR titration of oleic acid (D) The cross peak intensities (I) from the 2D 1H-15N HSQC spectra of Msi1 are plotted for four representative residues (Q30 black, R82 green, K98 red, A100 blue) as a function of the concentration of oleic acid added to sample of 133 μM Msi1 at pH 7 and 25 °C. The intensities are normalized relative to the measured intensity in the absence of oleic acid, I0. The intensity ratios were fitted to straight lines. (E) The cross peak intensities for all assigned residues of Msi1 in the 2D NMR spectrum after adding 0.22 molar equivalents of oleic acid. I0 is the intensity without oleic acid addition and I0.22 is the intensity after addition of 0.22 molar equivalents of oleic acid. Each point represents the I0.22/I0 for a single residue. The average of all intensities at I0.22 (solid line) and ± 1 standard deviation (dashed lines) indicate that some residues lose more intensity than others. A number of residues display normalized intensity values that are higher (blue) or lower (red) than 1 standard deviation from average. (F) Space-filling model of Msi1 RRM1 (Nagata et al. 1999; Miyanoiri et al. 2003). Residues with intensities below 1 sd from the average are red. Residues with intensities higher than 1 sd from the average are blue.
surface and the loops, which is typical for RRM proteins. Notably, backbone amide protons corresponding to F23, G64, and F65, display large chemical shift changes upon RNA binding. In contrast, chemical shift changes on the α helical face are small. The data are consistent with the model in which the β-sheet recognizes RNA and binding involves a structural transition including F23, G64, and F65 (Ohyama et al. 2011). Next, oleic acid or elaidic acid was titrated into the sample of Msi1 RRM1 (Figure 2.9D-F). We expected to observe chemical shift changes associated with oleic acid binding to the protein. The addition of increasing amounts of oleic acid did not result in significant changes in chemical shifts but in considerable reduction of the peak intensities and in line broadening. The largest loss of signal upon addition of substoichiometric concentrations of oleic acid was observed for W29, Q30, L36, C49, L50, R53, S60, G62, V74, T89, and K98 (Figure 2.9E-F). The observed loss of signal is likely due to chemical exchange between the free and bound state. As expected, we detected a strong loss in signal intensity across the entire protein when the concentration of oleic acid exceeded the CMC (75 ± 8 µM, Figure 2.3B and Figure 2.9D-F). The observed general loss of Msi1 signal is probably due to precipitation of the protein-oleic acid complex, or it could be attributed to the interaction of Msi-1 with oleic acid micelles, or both. We note that a precipitant forms in the NMR tube at elevated oleic acid concentration. Because the CMC of oleic acid is below the concentration needed to make a saturated Msi1 sample (Figure 2.3B), we cannot form sufficient Msi1-oleic acid complex for NMR structural studies.
A model of the omega-9 fatty acid inhibition mechanism

Because the high concentrations needed for NMR spectroscopic studies precluded direct measurement of the Msi1 fatty acid interface, we performed computational docking calculations with eicosenoic acid or oleic acid with Msi1 RRM1 (Table 2.1, Figure 2.10A-B) (Friesner et al. 2004; Friesner et al. 2006). In the docked models, both compounds insert the ω-9 end into the cavity, form extensive contacts with the protein surface, and position their carboxy termini adjacent to a positively charged arginine side chain (R53 or R61, dependent on the model, Figure 2.10, Figure 2.9D-F). The docked model offers an explanation for the experimentally observed specificity of inhibition. Amide and ester derivatives lose the negative charge and cannot form charge-charge interactions with arginine (oleamide, ethyl oleate). Shorter omega-7 fatty acids will not fill the hole (palmitoleic acid). Modifications of the omega end limit insertion into the hole (ricinoleic acid). Eighteen carbons enable the carboxylate to reach R53 (oleic acid), (Figure 2.10). Twenty-four carbons position the group beyond the arginine (nervonic acid). Trans orientation of the double bond limits surface contact and prevents orientation of the carboxylate towards R53 (elaidic acid). Finally, saturation of the double bond would require a large entropic penalty in order to adopt the necessary conformation (stearic acid).

To further assess how fatty acid binding alters the dynamics of amino acids that contribute to RNA recognition, we performed molecular dynamics (MD)
Figure 2.10: Docked model of Msi1 and oleic acid. (A-B) Model of RRM1 bound to oleic acid (yellow) calculated by Schrödinger GLIDE (Friesner et al. 2004; Friesner et al. 2006). (C) Msi1 RRM1 overlays from two NMR models demonstrate the highly dynamic nature of loop 3. Arginine 53 (model 1) and arginine 61 (model 3) are predicted to interact with the carboxylate of oleic acid in the respective models.
simulations of the Msi1 RRM1 motif with and without oleic acid. The NMR structure of MS1 RRM1 served as the starting configuration for the free state (Nagata et al.
1999; Miyanoiri et al. 2003). Three different models derived by docking oleic acid provided the starting configurations for the oleic acid-bound state (Friesner et al.
2004; Friesner et al. 2006). Each of the four simulations was equilibrated for 1 ns and data were collected during a subsequent 30 ns trajectory. Upon oleic acid binding, Msi1 underwent a transition to a more open state characterized by an increase in solvent accessible surface area (SASA) and radius of gyration (Figure 2.11A). Direct visualization of the structure’s time evolution showed that binding of oleic acid is associated with stabilization of the C-terminus of α-helix 1, fraying of α-helix 2 at both the N- and C-termini and, in one trajectory, formation of an additional β-sheet in loop 5 (Figure 2.12). The computed probability of each residue to be in a secondary structural element (Figure 2.11C) supports our observations. Notably, the β-sheet that forms in loop 5 is absent in the Msi1 apo structure but is present in the structure of Msi1 bound to RNA, indicating that association with either RNA or inhibitor may induce secondary structural changes in loop 5 (Figure 2.12) (Ohyama et al. 2011).

Analysis of the structure and dynamics of Msi1 bound to oleic acid and its comparison to the RNA-bound structure suggests a mechanism of inhibition. Oleic acid binding stabilizes β-strand formation in loop 5 and alters its position relative to helix 1. The distance between the C-terminus of α-helix 1 and loop 5 in the free and oleic acid-bound state is greater than that observed in the RNA-bound state.
Figure 2.11
Figure 2.11: Molecular dynamics studies of Msi1 with oleic acid. (A-B) represent the normalized histograms of the protein solvent accessible surface area (SASA) and radius of gyration (rgyr) illustrating the structural transition undergone by the Msi1 upon oleic acid binding. (A) SASA distributions calculated from the MD trajectories of Msi1 bound to oleic acid and of Msi1 in the apo state are represented in red and black, respectively. (B) Radius of gyration distributions calculated from the MD trajectories of Msi1 bound to oleic acid and of Msi1 in the apo state are represented in red and black, respectively. (C) The probability of being in an α-helix or β-sheet is shown for each residue of Msi1. The probabilities calculated for each residue from the MD trajectories of Msi1 free and bound to oleic acid are shown in black and red, respectively. Oleic acid binding is associated with stabilization of the C-terminus of α-helix 1, fraying of α-helix 2, at both the N- and C-termini, extension of sheet 2, as well as the formation of an additional β-sheet at loop 5 (L5). (D-E) Normalized histograms of the distance between G35, located on α-helix 1, and L85, located on loop 5, calculated from the MD trajectories of oleic acid-bound Msi1 and of apo Msi1 are shown in red and black, respectively. The distribution of distances between the Cα atoms of G35 and L85 is depicted in (D). The distribution of distances between the Cα of G35 and the Cδ2 of L85 is shown in (E). The green lines show the values of these distances observed in the NMR structure of Msi1 bound to RNA (PDB ID 2RS2). In the oleic-bound state, loop 5 is restricted in approaching α-helix 1 due to the steric hindrance of the oleic acid. (F-G) Normalized histograms of two representative side chain distances of W29 and Q30. Histograms calculated over the
MD trajectories of oleic acid-bound Msi1 are shown in red, those of apo Msi1 in black. The distance between W29 Cε2 and Q30 Cγ is shown in (F). The distance between W29 Cζ2 and Q30 Nε2 is shown in (G). The green lines show these distances observed in the NMR structure of Msi1 bound to RNA (PDB ID 2RS2). In the oleic-bound state, the side chain of W29 is stacked against the side chain of Q30. This conformation of W29 is not observed in either the free or RNA-bound states of Msi1.
Figure 2.12

A Molecular Dynamics Model
RRM1 Alpha Helical Face

B Molecular Dynamics Model
RRM1 Beta Sheet Face
Figure 2.12: MD model of Msi1 with oleic acid. (A) Overlay of the oleic acid bound MD simulation (gray and red) with the apo-state NMR structure (gray and blue) (Ohyama et al. 2011). (B) RNA contact residues (red) in loop 5, helix 2, and strand 4 of the β-sheet are perturbed in oleic acid-bound molecular dynamics simulation (Ohyama et al. 2011).
(Figure 2.12, Figure 2.11). The presence of oleic acid blocks loop 5 from approaching the C-terminus of α-helix 1. When loop 5 is in the more open conformation, K88 is not in position to interact with the first purine nucleotide of the consensus, in this case Gua1. In addition, binding of oleic acid causes W29 to stack against the side chain of Q30. (Figure 2.13A-B). This observation is in agreement with the NMR signal intensity loss observed for certain residues during oleic acid titration (Figure 2.9D-F), as well as the strong quenching of Msi1 tryptophan fluorescence measured upon addition of oleic acid (Figure 2.5, Figure 2.6). W29 directly stacks with the first RNA nucleotide, stabilizing the interaction between Msi1 and RNA. Stacking is eliminated in the presence of oleic acid (Figure 2.13A-B). Finally, strand 4 is more flexible in the oleic acid-bound state than in free Msi1 (Figure 2.13A-B). Strand 4 contains several residues that directly contact RNA. The simulations suggest that oleic acid weakens these interactions. In contrast to our initial hypothesis, we do not observe a dramatic change in the dynamics of F65 in the presence of oleic acid over the time course of the simulations.

The circular dichroism and analytical ultracentrifugation data (Figure 2.8) are consistent with a net gain in secondary structure upon oleic acid binding, as predicted by the MD simulation (Figure 2.11 and Figure 2.13). In addition, the change in tryptophan fluorescence observed upon oleic acid but not elaidic acid treatment is consistent with the changes in tryptophan solvent exposure observed in the simulation.
Figure 2.13
Figure 2.13: MD comparison of RNA-bound and oleic acid-bound Msi1. (A-B) A representative snapshot from the MD simulation of Msi1 bound to oleic acid (white) compared to the Msi1-RNA NMR structure (blue) (Ohyama et al. 2011). Panel (A) shows the Gua 1 binding pocket. In the oleic-bound state, the open conformation of loop 5 (L5) orients K88 such that K88 cannot contact Gua 1. W29 is stacked against Q30 and unavailable for stacking against Gua 1. Interaction with the side chain of R61 stabilizes the conformation of W29 in the oleic-bound state. Panel (B) highlights the different conformations of residues that interact with Gua 1, Ura 2, Ade 3 and Gua 4; represented in grey, orange, red and purple, respectively. (C) Difference of the mean Lipari-Szabo order parameters by residue between the apo and oleic acid-bound states of Msi1. The Lipari-Szabo order parameters for the backbone NH bond vectors, S2, were calculated to quantify the backbone flexibility of the free and oleic acid-bound form of Msi1. The difference of the order parameters, ΔS2 = S2apo - S2MSI-OA, indicates that Msi1-oleic acid complex is more flexible than apo Msi1, with the few exceptions mostly observed at the N-terminus. The secondary structural elements are highlighted at the top. Error bars are calculated from the standard deviation among trajectories.
To test features of the model, I mutated residues predicted to interact with oleic acid (Figure 2.14) and measured the relative inhibition constant ($K_{rel, app}$) by dose response FP. I observed no change in fatty acid inhibition when Q30 was mutated to a glutamate ($K_{rel, app} = 0.9$, Figure 3-Figure 3 Supplement 1). Mutation of H83 to leucine also had no effect on fatty acid inhibition. Mutation of H83 to phenylalanine reduced inhibition by almost 2-fold. H83 lines the hydrophobic pocket. Mutation of this residue to phenylalanine is predicted to make the pocket narrower. Mutation of G64 to alanine reduced inhibition by 3.1 fold. G64 forms the floor of the hydrophobic pocket, mutation to an alanine is expected to make the pocket more shallow. Mutation of G35 to aspartate and glutamate results in a 2-fold and 7-fold reduction of inhibition, respectively. This is possibly due to the addition of a negative charge near the mouth of the hydrophobic cavity.

To test whether the R53 and/or R61, which are predicted to interact with the carboxylate end of oleic acid, are important for inhibition, I made single glutamate mutations at each position (R53E or R61E). Oleic acid inhibition of the R53E mutant is reduced 3-fold, while R61E is reduced 5-fold. Mutating both arginines (R53E/R61E) increases this effect to 35-fold. This indicates that R53 and R61 are both important for inhibition, and act in a partially redundant way. A triple mutant (R53E/R59E/R61E) weakens inhibition by a similar amount (27-fold). The R53E mutant has a small effect on RNA-binding, while the R61E mutant has a strong
### Figure 2.14

**A** RRM1 Beta Sheet Face  
RRM1 Alpha Helical Face

**B**

<table>
<thead>
<tr>
<th>Protein</th>
<th>RNA binding affinity</th>
<th>RNA binding inhibition by oleic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_{d_{app}}$ (nM)</td>
<td>Relative affinity</td>
</tr>
<tr>
<td>WT</td>
<td>3.0 ± 2</td>
<td>1.0</td>
</tr>
<tr>
<td>Q30E</td>
<td>2.3 ± 0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>G350</td>
<td>7.1 ± 0.4</td>
<td>2.4</td>
</tr>
<tr>
<td>G35E</td>
<td>30 ± 3</td>
<td>10.2</td>
</tr>
<tr>
<td>H83F</td>
<td>7.7 ± 0.9</td>
<td>2.5</td>
</tr>
<tr>
<td>H83L</td>
<td>6.5 ± 0.1</td>
<td>2.2</td>
</tr>
<tr>
<td>G64A</td>
<td>27 ± 2</td>
<td>9.0</td>
</tr>
<tr>
<td>R53E</td>
<td>9.0 ± 0.9</td>
<td>3.0</td>
</tr>
<tr>
<td>R81E</td>
<td>53 ± 8</td>
<td>17.8</td>
</tr>
<tr>
<td>R53E/R61E</td>
<td>97 ± 10</td>
<td>32.7</td>
</tr>
<tr>
<td>R53E/R59E/R61E</td>
<td>146 ± 10</td>
<td>49.11</td>
</tr>
</tbody>
</table>
Figure 2.14: Mutational analysis of Msi1 RRM1. (A) Space-filling model of Msi1 RRM1 with mutated residues colored. Orange residues indicate mutations with modest (<2 fold) or no reduction in inhibition by oleic acid. Red residues displayed a larger (>2 fold) reduction in inhibition by oleic acid. (B) Table of mutant RNA-binding and inhibition data. Data are the average of three independent fluorescence polarization experiments. All proteins were purified and tested during the same time period and using the same reagents, enabling comparison of wild type and mutant variants.
18-fold effect. The interaction of R61 with oleic acid would preclude its interaction with RNA, and thus may contribute to the mechanism of inhibition.

Together, the MD simulations and experimental data are consistent with an allostERIC model of inhibition, wherein fatty acid binding induces a change in conformation that modifies the secondary structure of RRM1 and perturbs the position of amino acids required for RNA recognition. Additional work will be necessary to determine whether fatty acid inhibition works via the mechanism suggested by the computational docking and MD simulations or through an alternative mechanism.

Discussion

The data presented in this chapter illustrate the development of a small molecule screen for inhibitors of the RNA-binding protein Msi1. In the pilot screen, four compounds inhibited RNA-binding activity in both the high-throughput screening phase and the secondary screening phase. Of these, oleic acid was of particular interest due to its biological activity as an intermediary metabolite.

MSI1 is implicated in neural and epithelial development, where it acts to protect the proliferative capacity of stem and progenitor cells (Imai et al. 2001). Our lab and the Armstrong lab have both observed a role for Msi1 in preventing differentiation of oligodendrocyte progenitor cells into mature oligodendrocytes (Dobson et al. 2008). MSI1 expression is lost as cells mature into their terminal
identity, so it is perhaps not surprising that MSI1 might play a role in preventing the differentiation process.

Myelin is an essential lipid-rich membrane produced by mature oligodendrocytes. It serves to insulate and protect neuronal axons, therefore enabling impulse propagation. Oleic acid is one of the most prevalent fatty acids used in the production of the myelin membrane. It is interesting to note that this fatty acid becomes more abundant in oligodendrocyte lineage cells as they differentiate and MSI1 expression and function is lost.

A number of diseases are caused by myelin lesions or demyelination, including multiple sclerosis and Guillain Barre syndrome. Therapies for these diseases are extremely limited, and a method for targeted re-myelination has yet to be found. If MSI1 inactivation in oligodendrocyte progenitor cells can induce differentiation into mature, myelin-producing oligodendrocytes, it would present an attractive therapeutic target for these diseases. More work must be done to characterize the role of MSI1 in these cells, including a detailed study on the morphology and viability of the differentiated oligodendrocytes.

Cancerous tumors often contain a population of highly-proliferative ‘cancer stem cells.’ These cells display traits of stem and progenitor cells, but lack the proliferative regulation found in normal stem and progenitor cells. MSI1 is highly expressed in cancerous cells of neural and epithelial lineage, and high MSI1 expression is correlated with metastasis and poor patient prognosis (Strojnik et al.)
Several labs have observed that MSI1 depletion in tumor cells leads to differentiation and decreased proliferation (Sanchez-Diaz et al. 2008; Sureban et al. 2008; Wang et al. 2010). MSI1 has therefore gained notice as an attractive therapeutic prospect in the ongoing search for cancer therapies. The small molecule screen described here may prove useful in identifying inhibitors of MSI1 for use in both cancer and myelin disease therapies.

Oleic acid, while interesting for its role as a biological metabolite, is likely not a useful therapeutic molecule. It is a building block for membrane phospholipids, energy storage molecules, and lipid signaling molecules, and is therefore not a molecule that can be specifically used to target MSI1. Further screening efforts will be necessary to yield a potent, non-biological inhibitor of MSI1 for use in myelin and cancer therapeutic research.

**Methods**

**Plasmids**

DNA encoding the mouse Msi1 RNA binding domain fragment (amino acids 7-192) was amplified from the mammalian gene collection (MGC) full-length ORF clone 100014969 (Invitrogen) using gene specific primers (forward primer: 5'-cgcgcggatccagccgctgcctgcc-3'; reverse primer: 5'-gcgcgaagcttccggacatcctctt-3'). This fragment was digested using BamHI and
HindIII restriction enzymes and subcloned into a modified version of pET-22b vector (Invitrogen) in which the pelB leader sequence was replaced with a His6-Gly tag followed by a TEV protease site to make pET-22HT-MSI1 (7-192). Mutant versions of the MSI1 RNA binding domain were prepared by site-directed mutagenesis using QuikChange (Stratagene). The human MSI2 RNA binding domain fragment (amino acids 8-193) was amplified from MGC full-length ORF clone 3505639 using gene specific primers (forward primer: 5′-cgcgcggatccggcacctcgggcagcgccaa-3′; reverse primer: 5′-gcgcgaagctttcatgggaacatgacttctttcg-3′). This fragment was cloned into the BamHI and HindIII restriction sites of pET-22HT to make pET-22HT-MSI2 (8-193). The MSI1 RRM1 plasmid pET-22HT-MSI1 (7-103) was prepared by site-directed mutagenesis using QuikChange (Stratagene) to replace M104 with an ochre stop codon. Full-length mouse Msi1 was amplified from MGC full-length ORF clone 100014969 (Invitrogen) using forward primer 5′-cgccgagcatggagaactgacgcgccca-3′ and reverse primer 5′-ccgggccggctcagtggtacccattggtgaa-3′. The resulting fragment was subcloned into pCDH-CMV-MCS-EF1-Puro (System Biosciences) using the BamHI and NotI restriction enzymes to make pCDH-CMV-MSI1(FL).

**Purification of recombinant proteins**

H6-TEV-MSI-1 (7-192), H6-TEV-MSI-1 (7-104), and H6-TEV-MSI-2 (8-294) were expressed and purified from *Escherichia Coli* BL21(DE3) cells. Liquid cultures grown at 37°C were induced for 3 hours during mid-log phase with 1 mM Isopropyl β-D-1-
thiogalactopyranoside (IPTG). Cells were pelleted, resuspended in lysis buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 20 mM Imidazole, 5 mM β-Mercaptoethanol (BME)), and lysed using a microfluidizer (IDEX Health and Science). Soluble lysate was applied to a Ni-NTA column (Qiagen), washed with wash buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 50 mM Imidazole, 5 mM BME), and eluted with elution buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 300 mM Imidazole, 5 mM BME). Fractions were analyzed by SDS-page and those containing recombinant MSI1 or MSI2 were pooled and dialyzed overnight into S buffer (50 mM MOPS pH 6.0, 20 mM NaCl, 2 mM DTT). Pooled fractions were applied to a HiTrap SP cation exchange column (GE Healthcare) and eluted using a gradient of 0.1M to 1 M NaCl in S Buffer. Fractions containing MSI1 or MSI2 were pooled and dialyzed overnight into Q buffer (50 mM Tris pH 8.8, 20 mM NaCl, 2 mM DTT) prior to loading a HiTrap Q anion exchange column (GE Healthcare). The protein was eluted over a gradient from 0.1 to 1 M NaCl in Q buffer. Fractions containing MSI were pooled and dialyzed using Spectra/Por 7 25 kD (MSI1) or 10 kD (MSI2) molecular weight cutoff tubing (Spectrum laboratories) overnight into storage buffer (50 mM Tris pH 8.0, 20 mM NaCl, 2 mM DTT). The yield of >95% pure MSI1 or MSI2 is typically 20 mg per liter of culture (Figure 2.1).

RNA sequences and labeling

Synthetic RNA oligonucleotides were ordered from IDT and 3′ end-labeled with fluorescein 5-thiosemicarbazide (Invitrogen) according to the method of Reines and Cantor (Reines and Cantor 1974; Pagano et al. 2007; Farley et al. 2008). Briefly,
5x10^{-10} \text{ mol of RNA were incubated with 100 mM NaOAc, pH 5.1 and 5 nmol of NaIO4 for 90 minutes at room temperature then ethanol precipitated. The RNA was resuspended in 50 \mu l of 1 mM fluorescein-5-thiosemicarbazide in 100 mM NaOAc, pH 5.1. After incubating overnight at 4 °C, the RNA was separated from unreacted label by ethanol precipitation and subsequently passaged through a Pierce centrifuge column packed with Spehadex G-25 resin (GE Healthcare). RNA CCCR005 was a truncated form of the SELEX aptamer S8-13 identified by the Okano lab (Imai et al. 2001). In preparation for the small molecule screen, fCCCR005 (AGCGUUAGUUAUUAGUUCG/36-FAM/) was ordered pre-labeled from IDT.}

**Fluorescence polarization and electrophoretic mobility shift assays**

Fluorescence polarization (FP), also known as fluorescence anisotropy, and Fluorescence electrophoretic mobility shift assays (F-EMSA) were used to measure the binding affinity of recombinant MSI1 to fluorescein-labeled RNA aptamers. Assays were conducted as described in Pagano, et al. (Pagano et al. 2011). Briefly, 2 nM fluorescein-labeled RNA was incubated with varying concentrations of recombinant purified MSI1 protein in equilibration buffer (37.5 mM Tris pH 8.0, 75 mM NaCl, 0.0075% igepal, 0.0075 mg/mL tRNA) for three hours. Fluorescence polarization was determined with a Victor V3 plate reader using a 480 ± 31 nm excitation filter and a 535 ± 40 nm emission filter. After measuring FP, the samples were mixed with 6x bromocresol green loading dye (0.15% (w/v) Bromocresol green, 30% (v/v) glycerol) and run on a 5% native polyacrylamide gel at 120V for
75 minutes at 4ºC. Wet gels were scanned with a Typhoon FLA 9000 Biomolecular imager (GE healthcare) using a 473 nm laser and a long-pass cut-off filter (510 nm). The fraction of bound RNA was determined by quantifying lower (free) and upper (bound) band intensities using MultiGauge and ImageGauge software (Fujifilm). For RNA sequences with two shifted species, the bands were quantified together.

Polarization values or the fraction of bound RNA were plotted as a function of protein concentration and fit to the Hill equation (1) to determine the apparent dissociation constant (Kd) and the apparent Hill coefficient (n). The upper (m) and lower (b) values were also fit in order to define the assay window.

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

For RNA sequences where a bi-phasic transition was observed, the FP data were fit using a two-site model (equation 2) to determine both apparent dissociation constants (Kd1 and Kd2) and the fraction (F) of signal that corresponds to each transition.

\[
\phi = \left[ F(m - b) \left( \frac{P_t}{P_t - K_{d1}} \right) \right] + \left[ (1 - F)(m - b) \left( \frac{P_t}{P_t + K_{d2}} \right) \right] + b
\]  \hspace{1cm} (2)
**Small molecule screen**

The small molecule screen was performed at the UMass Medical School Small Molecule Screening core facility using a variation of the FP assay described above. 100 nM recombinant MSI1 protein and 2 nM fluorescein-labeled RNA aptamer CCCR005 were added to each assay well of 384-well black plates (Corning) using a µFill liquid dispenser (BioTek). Compounds dissolved in DMSO from the LOPAC and Chembridge libraries of small molecules were spotted in each assay well to a final concentration of 384 µM using a Tecan genesis workstation 150. Sixty-four wells of each plate were reserved for controls, including thirty-two wells that included no protein and no compound (free RNA), and 32 wells that included protein, RNA, and DMSO (no compound). Plates were equilibrated at 25°C prior to collecting polarization and fluorescence intensity data for each well using a Victor V2V plate reader (Perkin Elmer). The Z’ score, a measure of signal-to-noise, was calculated for each plate using the average (µ) and standard deviation (σ) of the control wells (equation 3).

\[
Z' = 1 - \left[ \frac{3(\sigma_1 + \sigma_2)}{\mu_1 - \mu_2} \right]
\]  

(3)

Plate reads with a Z’ of < 0.5 were repeated. The average Z’ of all plates was 0.7 ± 0.2. The polarization values (mP) of each well were normalized against the assay
window using the mean polarization values of the no protein control (µ₂) and no
compound control (µ₁) wells to generate an assay score (equation 4).

\[
\text{Score} = \frac{mP - \mu_2}{\mu_1 - \mu_2}
\]  

(4)

Hits were classified as wells with a score of 0.1 or less where the fluorescence
intensity remained within 2-fold of the control average to eliminate false positives
due to compound fluorescence or quenching.

**Dose response experiments**

Dose response experiments to assess inhibition activity were performed using a
modified FP and F-EMSA protocol. A constant concentration of sub-saturating
protein was equilibrated with 2 nM fluorescein-labeled RNA and varying
concentrations of compound in equilibration buffer. The FP and F-EMSA data were
collected as above and fit to a sigmoidal dose response equation to determine the
IC₅₀ (equation 5).

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

(5)
The apparent inhibition constant was calculated using the Lin and Riggs conversion (equation 6), which corrects for the equilibrium dissociation constant of MSI1 for the labeled RNA as well as the concentration of labeled RNA and protein used in the experiment (Lin and Riggs 1972; Ryder and Williamson 2004).

\[ K_{i,\text{app}} = \frac{2(K_d)(IC_{50})}{2P - R - 2K_d} \]  

(6)

**Intrinsic tryptophan fluorescence assay**

To directly assay the association of fatty acids and MSI1 protein, 6 µM MSI1 or N-acetyl-tryptophanamide (NATA) was incubated with varying concentrations of compound. Equilibrated reactions were excited at 280 nm, then steady-state fluorescence emission spectra were recorded between 295 and 400 nm in one nm intervals using a T-format Fluorolog fluorimeter (Horiba). 350 nm emission data were normalized and fit to a quadratic bimolecular association curve (equation 7) to determine the apparent dissociation constant, where C is the total compound concentration, P is the total protein concentration, and m and b represent the maximal and minimal signal, respectively.

\[ \phi = b + (m - b) \left[ \frac{C + P + K_d - \sqrt{(C + P + K_d)^2 - 4(CP)}}{2C} \right] \]  

(7)
In-silico Docking analysis

The ligand molecules were downloaded from the PubChem database (pubchem.ncbi.nlm.nih.gov) and prepared for docking using the LigPrep module in Maestro (Schrödinger, LLC). The target protein structures were downloaded from the PDB (www.rcsb.org) and prepared for docking in Maestro (Schrödinger, LLC) using the Protein preparation wizard. Glide (Schrödinger, LLC) was used to generate the receptor grid for subsequent docking and scoring the docked ligands in Standard Precision (SP) mode. The pose with best Glide score from each ligand/receptor docking run was selected for further analysis (Friesner et al. 2004; Halgren et al. 2004; Friesner et al. 2006).

NMR

Labeling with \(^{15}\)N was performed by growing cells in isotopically enriched M9 medium, 1g \(^{15}\)NH\(_4\)Cl per liter. 2D \(^1\)H-\(^{15}\)N HSQC spectra were collected using samples of U-\(^{15}\)N MSI1 in 90% H\(_2\)O/10% D\(_2\)O buffer solution of 50 mM Tris at pH 7.0. 2D \(^1\)H-\(^{15}\)N HSQC spectra were collected for each incremental addition of the unlabeled ligand (either aptamer RNA, oleic acid, or elaidic acid) to the \(^{15}\)N MSI1 sample to determine amide proton chemical shift changes upon titration of the ligand. All experiments were performed at 600 MHz on a Varian Inova spectrometer equipped with a triple-resonance cold probe at 298° K. Data processing was performed using NMRPipe (Delaglio et al. 1995) and sparky (Goddard and Kneller) software.
Molecular dynamics simulation

We performed molecular dynamics (MD) simulations of the RRM1 domain of MSI1 (residues 20-96) free and bound to oleic acid. We modeled the unknown structure of MSI1 bound to oleic acid by starting from the NMR solution structure of MSI1 RRM1 (Miyanoiri et al. 2003; Ohyama et al. 2011) and docking the oleic acid ligand using the GLIDE software package from Schrödinger, LLC (Friesner et al. 2006), followed by energy minimization and equilibration. All structures were solvated and neutralized in a TIP3P water box (Jorgensen et al. 1986). Energy minimization and MD simulations were subsequently carried out using the NAMD software package (Phillips et al. 2005) and using the version 27 CHARMM potential energy function (MacKerell et al. 1998). The particle mesh Ewald method (Darden et al. 1993; Essmann et al. 1995) was used to treat electrostatic interactions and periodic boundary conditions were applied throughout. The SHAKE algorithm (Ryckaert et al. 1977) was applied throughout the simulation to constrain the hydrogen atom bond lengths at their equilibrium values and an integration time step of 2 fs was used. After an initial energy minimization, the system was simulated in the isothermal-isobaric ensemble. Non-bonded interactions were calculated every time step using a cut-off distance of 12 Å. After equilibration in the isothermal-isobaric ensemble, an additional stage of equilibration was performed in the microcanonical ensemble. We then collected three independent 30 ns constant-NVE production runs of MSI-1 bound to oleic acid and one for MSI-1 in the free state at an average
temperature of 298° K.

Visualization and secondary structure analysis was performed in VMD, using the STRIDE method (Frishman and Argos 1995).

In addition, we employed a number of measures to characterize the structural and dynamical changes between the free and bound state of MSI-1, including the radius of gyration, the solvent accessible surface area (SASA), distance between backbone and/or side chain atoms of residues G35-L85 and W29-Q30, and Lipari and Szabo order parameters (Lipari and Szabo 1982; Humphrey et al. 1996). These quantities were calculated using VMD (Humphrey et al. 1996) software along with bespoke programs previously described elsewhere (Morgan and Massi 2010).

**Circular dichroism**

Far-UV circular dichroism (CD) spectra were collected using 10 µM MSI1 in 50 mM Tris pH 8.0, 100 mM NaCl and 0.1% TFE on a Jasco-810 spectropolarimeter (Jasco Inc., Easton, MD). Spectra were collected from 215-260 nm in a 0.2 cm path length quartz cuvette using a scan rate of 20 nm min⁻¹ and a response time of 8 s. The sample temperature for all CD measurements was maintained at 293° K.

**Analytical ultracentrifugation**

Sedimentation velocity analyses were conducted using a Beckman Optima XL-I analytical ultracentrifuge in the University of Massachusetts Medical School Ultracentrifuge Facility. Data were analyzed with UltraScan III version 1.0 (Demeler
et al. 2011). Sedimentation velocity experiments were performed with 45 μM MSI1 in storage buffer (50 mM Tris-HCl pH 8.0, 20 mM NaCl, 2 mM DTT). Measurements were made at 20°C using an AN60ti rotor at 20000 rpm and 280 nm in intensity mode. Partial specific volumes were estimated on the basis of peptide sequence with UltraScan and found to be 0.7280 cm3/g for the MSI1 RBD. Data were analyzed by two-dimensional spectrum analysis with simultaneous removal of time and radially invariant noise (Demeler et al. 2009; Brookes et al. 2010). Noise and diffusion-corrected, model-independent sedimentation coefficient distributions were generated using the enhanced van Holde-Weischet analysis (Demeler et al. 1997; Demeler and van Holde 2004).

**Critical micelle concentration determination**

The critical micelle concentration (CMC) for oleic acid in equilibration buffer was determined using N-phenyl-1-naphthylamine (NPN), a compound that fluoresces when sequestered into micelles (Hagihara et al. 2002; Hasegawa et al. 2008). 25 μM NPN was incubated with varying amounts of oleic acid for 30 minutes. Fluorescence intensity was determined with a Victor V3 plate reader using a 355 nm excitation filter and a 460 nm emission filter. Fluorescence intensity data were plotted as a function of fatty acid concentration. The CMC was determined by a two state segmented linear regression to identify the breakpoint between baseline fluorescence and fluorescence caused by NPN association with micelles.
Chapter III: Musashi-1 and the lipid biosynthetic enzyme stearoyl-CoA desaturase participate in a feedback loop to regulate cellular proliferation.
Abstract

Gene expression and metabolism are coupled at numerous levels. Cells must sense and respond to nutrients in their environment, and specialized cells must synthesize metabolic products required for their function. Pluripotent stem cells have the ability to differentiate into a wide variety of specialized cells. How metabolic state contributes to stem and progenitor cell proliferation and differentiation is not understood. Data from the previous chapter showed that RNA-binding by Musashi-1 is allosterically inhibited by 18-22 carbon ω-9 monounsaturated fatty acids. Musashi proteins are critical for development of the brain, blood, and epithelium. In this chapter I present data indentifying the lipogenic enzyme stearoyl-CoA desaturase-1 (SCD1) as a MSI1 target. SCD products are used in essential cellular processes including proliferation, revealing a feedback loop between ω-9 fatty acid biosynthesis and MSI1 activity.

Introduction

During development, stem and progenitor cells give rise to the cells that comprise mature tissues and organs in response to a myriad of temporal and spatial regulatory signals. After development, small populations of these immature cells remain to repair damage and replace cells that die or are selectively culled. Cancerous tumors often result when the proliferation machinery in these immature cells becomes deregulated. Although the factors controlling proliferation and
differentiation in these cells are not yet understood, the dual RRM RNA-binding protein MSI1 likely plays a role in these processes.

MSI1 is highly expressed in neural and epithelial progenitor cells, where it promotes proliferation and prevents differentiation (Kayahara et al. 2003; Nishimura et al. 2003; Potten et al. 2003; Akasaka et al. 2005; Asai et al. 2005; Colitti and Farinacci 2009; Murayama et al. 2009). MSI1 is also highly expressed many in cancers of neural and epithelial origin (Strojnik et al. 2006; Fan et al. 2010; Wang et al. 2010). MSI1 depletion results in decreased proliferation, increased apoptosis, and reduced tumor xenograft growth (Sanchez-Diaz et al. 2008; Sureban et al. 2008; Wang et al. 2010). Additionally, poor patient prognosis is correlated with high MSI1 expression in tumors of neural and epithelial lineage (Strojnik et al. 2007; Wang et al. 2010; Vo et al. 2012b; Dahlrot et al. 2013).

These data implicate MSI1 as a logical therapeutic target for various cancers and diseases of neural and epithelial lineage tissues. However, the mechanisms that govern MSI1 expression are not yet known. There is evidence that certain micro RNAs, post-transcriptional regulatory factors, transcription factors, and cyclin-dependent and mitogen-activated protein kinases may regulate MSI1 activity, but it is not clear whether regulation is direct or indirect (Vo et al. 2011; Arumugam et al. 2012; Vo et al. 2012a; Pasto et al. 2014). The data presented in chapter II show that MSI1 is allosterically inhibited by a class of intermediary metabolites, including oleic acid. This data provides a link between the post-transcriptional regulation of proliferation and metabolic state.
In the current chapter, I present data that further strengthen the link between MSI1-mediated post-transcriptional regulation and metabolism. Here, the lipid biogenic enzyme stearoyl-CoA desaturase (SCD) is implicated as a novel MSI1 regulatory target. SCD is the rate-limiting enzyme in the production of omega-9 unsaturated fatty acids and is therefore essential for proliferation, energy storage, and lipid signaling. I propose a feedback loop involving SCD and MSI1 to regulate proliferative homeostasis in healthy neural and epithelial lineage cells.

**Results**

*Oleic acid treatment reduces oligodendrocyte progenitor cell proliferation*

To understand how fatty acids regulate MSI1, I investigated the effect of treating the immortalized rat oligodendrocyte progenitor cell line CG-4 with oleic acid (Louis et al. 1992). CG-4 cells maintain normal precursor cell morphology, can be readily transfected, and can be induced to differentiate into mature oligodendrocytes by withdrawal of growth factors (Franklin et al. 1995) (Louis et al. 1992). I observed that CG-4 cells undergo a lipid profile shift upon differentiation, as detected by quantitative lipidomics mass spectrometry (Figure 3.1). As with primary OPCs, Msi1 is expressed strongly in the precursor state but decreases upon differentiation (OPCs: 1.0 ± 0.3, oligodendrocytes: 0.15 ± 0.01, $p$-value = 0.0035). Treatment of CG-4 OPCs with oleic acid strongly inhibited the rate of cell proliferation (Figure 3.2), matching the published phenotype of Msi1 knock down in primary OPCs (Dobson et al., 2008). In contrast, treating HEK293T cells—which do
Figure 3.1
Figure 3.1: Lipidomics analysis of undifferentiated and differentiated CG4 oligodendrocyte progenitor cells (A) Scatter plot of lipidomics data for differentiated vs undifferentiated CG4 cells. Data are reported as nMoles per million cells. Red data points indicate lipids that are significantly different after differentiation (FDR = 5%). (B) Volcano plot of lipidomics data. Dashed lines denote fold-changes of ±1.5 and ±3. Red data points indicate lipids that are significantly different after differentiation (FDR = 5%).
not express Msi1—with the same concentration of oleic acid had little or no effect on the rate of proliferation. Treatment with stearic acid, which does not inhibit MSI1, had no effect on proliferation rate in either cell type (Figure 3.2). The data suggest that reduction of Msi1 activity by oleic acid limits proliferation, but there is a possibility that the fatty acid modulates other cellular pathways that contribute to proliferation.

**MSI1 regulates stearoyl-CoA desaturase**

Because metabolic homeostasis is often regulated through feedback loops, it is possible that MSI1 might control the expression of enzymes required to make long chain monounsaturated fatty acids. In humans, non-dietary oleic acid is produced from stearic acid by SCD. Expression of SCD is tightly controlled at the transcriptional, post-transcriptional, and post-translational levels. In rodents, there are four SCD genes, *Scd1*-4. I used FP and F-EMSA assays to test the ability of recombinant Msi1 to bind each of seven putative consensus sites present in the *Scd1* 3′ UTR (Figure 3.3). Msi1 binds to all seven sites by FP and F-EMSA (Figure 3.3B). Site 2 and site 7 bind to Msi1 with comparable affinity to the selected aptamer. The slightly decreased binding affinity measured for these 7 fragments is likely due to the fact that aptamer CCCR005 contains two Msi1 binding sites, while each *Scd1* 3′ UTR fragment contains only a single Msi1 binding site. Transcripts encoding SCD1 co-immunoprecipitated with Msi1 from CG-4 cell extracts using two independent antibodies (Figure 3.4A). Similar results were obtained with *Numb* mRNA, a positive
Figure 3.2: Fatty acid treatment of CG-4 OPCs (A) Dose response for oleic acid treatment in cell culture. HEK293T (dashed) and CG-4 (solid) cell proliferation as a function of increasing oleic acid (red = 50 µM, grey = 5 µM black = 0 µM). The data is the average and standard deviation of at least three biological replicates. (B) HEK293T (dashed) and CG-4 (solid) cell proliferation as a function of oleic acid or stearic acid treatment (red = treated, black = untreated). The data is the average and standard deviation of at least three biological replicates.
Figure 3.3

A

B

<table>
<thead>
<tr>
<th>Name</th>
<th>Mouse Scd1 3 UTR fragment</th>
<th>Sequence</th>
<th>PP Kd, app (nM)</th>
<th>F-EMSA Kd, app (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCCR008</td>
<td>1</td>
<td>GGCAGUGUGACC</td>
<td>79 ± 7</td>
<td>141 ± 8</td>
</tr>
<tr>
<td>CCCR006</td>
<td>2</td>
<td>CCCAGAAGUGUGAGAUU</td>
<td>35 ± 9</td>
<td>54 ± 10</td>
</tr>
<tr>
<td>CCCR009</td>
<td>3</td>
<td>GCCAGUGAGUAC</td>
<td>71 ± 4</td>
<td>65 ± 4</td>
</tr>
<tr>
<td>CCCR012</td>
<td>4</td>
<td>UUGAUGAGUGUGUCUG</td>
<td>80 ± 7</td>
<td>157 ± 10</td>
</tr>
<tr>
<td>CCCR011</td>
<td>5</td>
<td>GGAGAUUGCAAGA</td>
<td>380 ± 10</td>
<td>250 ± 20</td>
</tr>
<tr>
<td>CCCR010</td>
<td>6</td>
<td>AUUCUGUGAGCCUG</td>
<td>185 ± 8</td>
<td>94 ± 7</td>
</tr>
<tr>
<td>CCCR013</td>
<td>7</td>
<td>CCAGUCGGUGUUGCUACA</td>
<td>36 ± 11</td>
<td>36 ± 7</td>
</tr>
</tbody>
</table>
Figure 3.3: SCD1 is a putative Msi1 target. (A) There are seven MSI1 consensus sites in the 3’-UTR of Scd-1 mRNA. The Kd, app is the average and standard deviation of at least three experiments. (B) MSI1 binds each of 7 SCD1 3’ UTR fragments containing the MSI1 consensus sequence (G/A)U1-3AGU with variable affinity. FP and F-EMSA binding experiments are reported as the average and standard deviation of three independent experiments.
Figure 3.4

A

B

C

D

Full Length SCD
Proteolysis products

Normalized SCD Level

p<0.018

Fold Change

SCD1 mRNA

Fold Change

Meth

Meth
Figure 3.4: Msi1 is a positive regulator of SCD. (A) Scd-1 transcripts co-immunoprecipitate with anti-Msi1 antibodies. The data were quantified using a FUJI FLA-5000 imager. (B) Western analysis of SCD expression in HEK293T cells. The data were quantified using the LICOR Odyssey system relative to non specific bands (** and *, Figure 6-figure 6 supplement 2B) to control for loading. The average and standard deviation of at least three independent experiments is shown. (C) The SCD1 antibody is specific for full-length SCD1 and the associated proteolysis products. Two independent shRNA constructs reduced full-length SCD1 and associated proteolysis products compared to non-transfected or control shRNA transfected HEK293T cells. (D) qRT-PCR of SCD1 mRNA. The data are the mean and SD of the fold-change in SCD1 mRNA from 5 independent experiments. Each experiment comprised 3 technical replicates normalized to tubulin or GTF2i.
control. Overexpression of Msi1 in HEK293T cells, which do not endogenously express MSI1, increased the amount of SCD and SCD proteolysis products by $2.5 \pm 0.35$ fold ($p$-value = 0.018, Figure 3.4B-C). I was unable to detect SCD expression in the rat CG-4 cells using available antibodies. qRT-PCR analysis of HEK293T mRNA shows a slight but non-significant increase in SCD1 transcript abundance upon Msi1 overexpression (Figure 3.4D).

*SCD lipid products are enriched upon MSI1 expression*

Next I sought to assay downstream effects of SCD regulation by MSI1 in cells. Because SCD plays an integral role in lipogenesis, it is reasonable to expect that changes in SCD activity would result in a shift in the cellular lipid profile. To this end, I performed quantitative lipidomics mass spectrometry analysis. OPC proliferation halts upon MSI1 knockdown (Dobson et al. 2008). For this reason I performed these experiments in HEK293T cells with and without Msi1 expression. I observed statistically significant changes ranging from 2 to 32-fold in 54 of the 312 lipids assayed (FDR = 0.05, Figure 3.5A-B). 50 of 54 significant data points displayed positive changes, indicating that Msi1 expression stimulates production of certain lipids. 38 of the significantly changing lipids fell into one of three cholesterol and triacylglycerol (TAG) classes: cholesterol esters, 16:0 acyl-containing TAG, and 16:1 acyl-containing TAG (Figure 3.5C-D). Cholesterol esters and TAGs are made using SCD products, and have been shown to be more abundant in mice overexpressing
Figure 3.5: Lipidomics analysis of HEK293t cells ± MSI1 expression. Source data are included in Figure 6 source data 1. (A) Volcano plot of lipidomics data. Dashed lines denote fold-changes of ±1.5 and ±3. Red data points indicate lipids that are significantly changed upon Msi1 expression. (B) Scatter plot of lipidomics data. Data are reported as nMoles per million cells. Red data points indicate lipids that are significantly changed upon Msi1 expression (FDR = 0.05). (C) Fold-changes of the total cholesterol esters and two TAG classes in which 38 of the 54 significantly changing lipids are categorized. Each class changes significantly with MSI1 overexpression (P < 0.05). (D) Fold-changes for the four lipids that comprise the total cholesterol esters class. All display significant changes with Msi1 expression (FDR = 0.05).
SCD1, and less abundant in liver SCD1 knockout mice (Miyazaki et al. 2000; Attie et al. 2002). This indicates that SCD activity is upregulated in the presence of Msi1. Together, the binding data, co-immunoprecipitation, western blots, and lipidomics profiling results show that stearoyl-CoA desaturase expression is increased in the presence of MSI1, possibly through direct association with consensus motifs in the 3′ UTR of SCD transcripts.

Additional putative MSI1 regulatory targets involved in lipid metabolic processes

There are 2107 annotated 3′ UTRs for genes involved with lipids, cholesterol, fatty acid biosynthesis, and activity of the lipid-responsive transcription factor sterol responsive element binding protein (SREBP). Of these, 1275 contain at least one MSI1 binding element. While the presence of a binding element does not confer regulation, MSI1 regulation of lipid metabolism may be more complex than simple direct regulation of SCD. Indeed, a survey of 64 genes that associate with MSI1 in a RIP-ChIP experiment reveals that 8 are annotated to be involved with lipid metabolic processes (de Sousa Abreu et al. 2009).

Discussion

The data presented in chapter II reveal that long chain ω-9 fatty acids between 18-22 carbons in length are allosteric inhibitors of MSI1 RNA binding activity. In this chapter I present data implicating SCD, the enzyme that catalyzes the ω-9 desaturation, as a MSI1 regulatory target. The results are consistent with a
model where MSI1 controls cellular proliferation through a feedback loop that includes SCD and its enzymatic products, such as oleic acid (Figure 3.6).

Oleic acid is the precursor for synthesis of longer chain fatty acids such as eicosanoic, erucic, and nervonic acid, signaling molecules such as arachidonic acid, endocannabinoids, and prostaglandins, and membrane phospholipids. Oleic acid is abundant in the lipid-rich myelin membranes produced by mature oligodendrocytes (Martinez and Mougan 1998), and CG4 OPCs show global changes in the lipid profile upon differentiation (Figure 3.1). OPCs express MSI1, but mature myelinating oligodendrocytes do not (Dobson et al. 2008). In the oligodendrocyte lineage, oleic acid and MSI1 levels are anti-correlated, which suggests a possible biological role for MSI1 inhibition by fatty acids (Martinez and Mougan 1998; Dobson et al. 2008). Another MSI1 inhibitor identified in the small molecule screen presented in chapter II is the PPARα agonist GW7647. PPARs are nuclear hormone receptor proteins that regulate cellular processes including fatty acid metabolism, differentiation, cell cycling, and inflammation (Berger and Moller 2002). PPARγ agonists accelerate oligodendrocyte maturation (De Nuccio et al. 2011). Both GW7647 and oleic acid function as PPAR agonists, and both inhibit MSI1 RNA-binding activity, suggesting that MSI1 and PPARs possibly regulate gene expression in a reciprocal fashion.

These data implicate the metabolic enzyme SCD as regulatory target of MSI1. The two human isoforms of SCD, SCD1 and SCD5, catalyze the conversion of saturated fatty acids (SFAs) into monounsaturated fatty acids (MUFAs) (Zhang et al. 1999; Beiraghi et al. 2003; Minville-Walz et al. 2010). MUFAs are then used in the
Figure 3.6: Model of SCD regulation by MSI1, and subsequent downstream consequences of SCD activity changes.
synthesis of numerous lipids, including phospholipids, di- and triacylglycerols, cholesterol esters, and signaling molecules such as eicosanoids. SCD is therefore an essential enzyme to normal cellular proliferation, metabolism, and signaling. It is also possible that MSI1 may also be regulating other factors involved in lipid, fatty acid and cholesterol metabolism. The Penalva group identified 64 putative MSI1 targets using a RIP-ChIP analysis in HEK293T cells (de Sousa Abreu et al. 2009). Of these putative targets, eight are annotated to play a role in lipid, fatty acid or cholesterol metabolic processes. These include three variants of the fatty acid elongase ELOVL5, two variants of the glycolipid biosynthetic factor PIGF, two variants of Lamin B receptor (LBR), and aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase (AASDHPT).

SCD has been implicated in a number of disease states, including obesity, diabetes, hyperlipidemia, and cancer. Obese and diabetic animals produce abnormally high levels of TAGs and cholesterol esters (Coleman and Lee 2004). These energy storage molecules are also elevated in a hyperlipidemia mouse model that overexpresses SCD (Attie et al. 2002). Conversely, SCD1 knockout mice display impaired biosynthesis of both TAGs and cholesterol esters (Miyazaki et al. 2000). SCD is upregulated in tumor cells, including colonic and esophageal carcinoma (Li et al. 1994). Cancerous cells use newly synthesized lipids primarily for phospholipid production, which are used in new membranes (Swinnen et al. 2000). Additionally, new evidence suggests that SCD1 activates the oncogenic Akt and AMPK signaling pathways (Scaglia and Igal 2008; Scaglia et al. 2009).
SCD regulates SFA and MUFA homeostasis and therefore plays an integral role in lipid signaling pathways. SFAs serve as proinflammatory factors by acting as ligands for immune receptors such as those in the Toll-like receptor family (Shi et al. 2006; Nguyen et al. 2007). MUFA products of SCD are further modified to become polyunsaturated fatty acids, such as arachidonic acid, which are converted into eicosanoids by cyclooxygenases (James et al. 2000). Eicosanoids serve as proinflammatory and immune signaling molecules (Liu et al. 2011). Recent data also links SCD to Wnt signaling through mediation of the palmitic/palmitoleic acid conversion, as active Wnt proteins require conjugation of palmitoleic acid (Rios-Esteves and Resh 2013). The balance between SFAs and MUFAs must be tightly controlled, and while SCD is not the only enzyme involved in the process, it does serve as an essential gatekeeper in the conversion of dietary SFA to MUFA.

Several examples of functional interactions between transcription factors and fatty acids have been published. For example, fatty acids associate with membrane-bound SREBP to inhibit a cleavage event that produces the activated form of SREBP, a transcription factor involved in lipid homeostasis (Wang et al. 1994). Free fatty acids have been shown to stimulate proinflammatory cytokine expression while decreasing anti-inflammatory cytokine expression in adipocytes (Bradley et al. 2008). Interestingly, treatment of various tumor lines with oleic acid results in transcriptional inhibition of the Her-2/neu (erbB-2) oncogene through PEA3, although the precise mechanism of regulation remains unknown (Menendez et al. 2006). In yeast cells, fatty acids bind oleate-activated transcription factors.
(OAFs) to effect transcription of genes responsible for fatty acid metabolism, glucose metabolism, stress response, and other related processes (Gurvitz and Rottensteiner 2006). Although numerous transcription factors are regulated by fatty acids, MSI1 is the first example of a fatty acid-responsive RNA binding protein.

Integration of metabolite sensing and post-transcriptional regulation is widespread in bacteria, but comparatively few examples have been found in eukaryotes (Winkler and Breaker 2003; Roth and Breaker 2009). Riboswitches regulate gene expression at the RNA level in response to intermediary metabolites. Small molecule metabolites bind mRNA transcripts, usually in the 5′ UTR, to induce a structural change that interferes with the transcriptional or translational machinery (Winkler and Breaker 2005). In bacteria, riboswitches sensitive to a number of metabolites have been characterized, including guanine, adenine, coenzyme B12, glycine, lysine, and thiamine pyrophosphate (TPP), among others (Grundy et al. 2003; Mandal et al. 2003; Sudarsan et al. 2003; Vitreschak et al. 2003; Mandal and Breaker 2004; Mandal et al. 2004). Although most riboswitches have been identified within the bacterial mRNA 5′ UTR, TPP riboswitches, have also been identified in plants and fungi (Kubodera et al. 2003; Sudarsan et al. 2003). While riboswitches comprise an essential cis-regulatory mechanism in bacteria, other mechanisms of coupling metabolic state to post-transcriptional gene regulation may occur in eukaryotes.

An increasing number of protein-metabolite-RNA interactions have now been identified in bacteria and eukaryotes. One example is the *bacillus subtilis*
tryptophan RNA-binding attenuation protein (TRAP). Increasing intracellular tryptophan levels induce TRAP multimerization, which enables mRNA recognition and subsequent translational repression of targets (Gollnick et al. 1990; Antson et al. 1995; Babitzke et al. 1995; Yakhnin et al. 2004). Intriguingly, several metabolic enzymes have been proposed to “moonlight” as RNA-binding proteins in “RNA / Enzyme / Metabolite (REM)” networks (Ciesla 2006; Hentze and Preiss 2010; Castello et al. 2012). A notable example is cytosolic aconitase, which demonstrates mutually exclusive enzymatic and RNA-binding functionality, depending upon cellular iron levels (Hentze and Argos 1991; Rouault et al. 1991). A number of metabolite-sensitive enzymes have also demonstrated RNA-binding activity, including GAPDH, glutamate dehydrogenase, thymidylate synthase, and dihydrofolate reductase (Ryazanov 1985; Chu et al. 1991; Chu et al. 1993; Preiss et al. 1993; Dollenmaier and Weitz 2003). This growing body of research suggests that metabolite-mediated post-transcriptional regulation is much more prevalent than previously thought.

The results presented here show that MSI1 N-terminal RRM1 acts as a metabolite sensor, the first example of such activity observed for the most abundant RNA-binding motif in eukaryotic genomes. A survey of RRM structures in the protein data bank reveals several with a surface cavity on the alpha helical face, which may comprise metabolite-binding pockets. I predict that a network of RNA regulatory proteins act as metabolite sensors, possibly replacing the bacterial riboswitch regulation that appears to have been largely lost in eukaryotes.
Methods

Plasmids

DNA encoding the mouse Msi1 RNA binding domain fragment (amino acids 7-192) was amplified from the mammalian gene collection (MGC) full-length ORF clone 100014969 (Invitrogen) using gene specific primers (forward primer: 5′-cgccgcggatccagccggccctcgcctcccc-3′; reverse primer: 5′-gcgcgaagcttcggggacatcacctcctttg-3′). This fragment was digested using BamHI and HindIII restriction enzymes and subcloned into a modified version of pET-22b vector (Invitrogen) in which the pelB leader sequence was replaced with a His6-Gly tag followed by a TEV protease site to make pET-22HT-MSI1 (7-192).

Purification of recombinant proteins

H6-TEV-MSI-1 (7-192) was expressed and purified from Escherichia Coli BL21(DE3) cells. Liquid cultures grown at 37ºC were induced for 3 hours during mid-log phase with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were pelleted, resuspended in lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 20 mM Imidazole, 5 mM β-Mercaptoethanol (BME)), and lysed using a microfluidizer (IDEX Health and Science). Soluble lysate was applied to a Ni-NTA column (Qiagen), washed with wash buffer (50 mM NaH₂PO₄, 300 mM NaCl, 50 mM Imidazole, 5 mM BME), and eluted with elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 300 mM Imidazole, 5 mM BME). Fractions were analyzed by SDS-page and those containing recombinant MSI1
or MSI2 were pooled and dialyzed overnight into S buffer (50 mM MOPS pH 6.0, 20 mM NaCl, 2 mM DTT). Pooled fractions were applied to a HiTrap SP cation exchange column (GE Healthcare) and eluted using a gradient of 0.1M to 1 M NaCl in S Buffer. Fractions containing MSI1 or MSI2 were pooled and dialyzed overnight into Q buffer (50 mM Tris pH 8.8, 20 mM NaCl, 2 mM DTT) prior to loading a HiTrap Q anion exchange column (GE Healthcare). The protein was eluted over a gradient from 0.1 to 1 M NaCl in Q buffer. Fractions containing MSI were pooled and dialyzed using Spectra/Por 7 25 kD (MSI1) or 10 kD (MSI2) molecular weight cutoff tubing (Spectrum laboratories) overnight into storage buffer (50 mM Tris pH 8.0, 20 mM NaCl, 2 mM DTT). The yield of >95% pure MSI1 or MSI2 is typically 20 mg per liter of culture (Figure 2.1).

**RNA sequences and labeling**

Synthetic RNA oligonucleotides were ordered from IDT and 3’ end-labeled with fluorescein 5-thiosemicarbazide (Invitrogen) according to the method of Reines and Cantor (Reines and Cantor 1974; Pagano et al. 2007; Farley et al. 2008). Briefly, 5x10^{-10} mol of RNA were incubated with 100 mM NaOAc, pH 5.1 and 5 nmol of NaIO₄ for 90 minutes at room temperature then ethanol precipitated. The RNA was resuspended in 50 µl of 1 mM fluorescein-5-thiosemicarbazide in 100 mM NaOAc, pH 5.1. After incubating overnight at 4 °C, the RNA was separated from unreacted label by ethanol precipitation and subsequently passaged through a Pierce centrifuge column packed with Spehadex G-25 resin (GE Healthcare).
Fluorescence polarization and electrophoretic mobility shift assays

Fluorescence polarization (FP), also known as fluorescence anisotropy, and Fluorescence electrophoretic mobility shift assays (F-EMSA) were used to measure the binding affinity of recombinant MSI1 to fluorescein-labeled RNA aptamers. Assays were conducted as described in Pagano, et al. (Pagano et al. 2011). Briefly, 2 nM fluorescein-labeled RNA was incubated with varying concentrations of recombinant purified MSI1 protein in equilibration buffer (37.5 mM Tris pH 8.0, 75 mM NaCl, 0.0075% igepal, 0.0075 mg/mL tRNA) for three hours. Fluorescence polarization was determined with a Victor V3 plate reader using a 480 ± 31 nm excitation filter and a 535 ± 40 nm emission filter. After measuring FP, the samples were mixed with 6x bromocresol green loading dye (0.15% (w/v) Bromocresol green, 30% (v/v) glycerol) and run on a 5% native polyacrylamide gel at 120V for 75 minutes at 4°C. Wet gels were scanned with a Typhoon FLA 9000 Biomolecular imager (GE healthcare) using a 473 nm laser and a long-pass cut-off filter (510 nm). The fraction of bound RNA was determined by quantifying lower (free) and upper (bound) band intensities using MultiGauge and ImageGauge software (Fujifilm). For RNA sequences with two shifted species, the bands were quantified together.

Polarization values or the fraction of bound RNA were plotted as a function of protein concentration and fit to the Hill equation (1) to determine the apparent dissociation constant (K_d) and the apparent Hill coefficient (n). The upper (m) and lower (b) values were also fit in order to define the assay window.
\[ \phi = b + (m - b) \left[ \frac{1}{1 + \left( K_d / \left[ P_i \right] \right)^n} \right] \]  

(1)

**Cell-based experiments**

**Cell lines and culture**

CG-4 rat oligodendrocyte progenitor and B104 neuroblastoma cells were a gift from Lynn Hudson and were cultured as previously described (Louis et al. 1992). CG-4 cells were maintained in the progenitor state in 30% B104 conditioned media. HEK293T cells were cultured in Dulbecco’s modified Eagle’s medium (Invitrogen) with 10% fetal bovine serum (Atlanta Biologicals Inc., Lawrenceville, GA), 100 units/ml penicillin, and 100 µg/ml streptomycin (Invitrogen).

**Proliferation assays**

100,000 HEK293T or CG-4 cells were seeded into each well of 6-well culture plates. 24 h after seeding, cells were treated with either 50 µM oleic acid in EtOH or EtOH only as a control. 24, 48, and 72 h after fatty acid treatment, cells were assayed for proliferative changes using the Cyquant direct cell proliferation assay (Invitrogen) according to the manufacturers instructions using a Victor V3 plate reader (PerkinElmer). The treatment concentration was chosen after a dose response for toxicity and efficacy across a range defined by established protocols for neuroblastoma cells (Figure 3.2A) (Di Loreto et al. 2007).
**RNA co-immunoprecipitation**

For each experiment, $10^6$ CG-4 cells were crosslinked for 10 minutes with 0.1% formaldehyde in PBS then scraped into lysis buffer (1% (w/v) SDS, 10 mM EDTA, 1x EDTA-free protease inhibitors). Lysate was diluted in 10X IP buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1, 167 mM NaCl) with 8 units SUPERase-In (Ambion). Immunoprecipitation was performed using rabbit polyclonal anti-Musashi 1 (Abcam ab21628) or rabbit polyclonal anti-GFAP (Abcam ab7260) conjugated to Dynabeads protein G (Invitrogen) according to the manufacturer’s instructions. After proteinase K digestion (NEB) and heat-induced crosslink reversal, RNA was phenol-chloroform extracted, ethanol precipitated and resuspended in 20 µL nuclease-free water. RNA was then treated with DNase (Ambion Turbo DNAfree) to remove any DNA contamination. RNA was reverse transcribed and amplified using a SuperScript® III One-Step RT-PCR System with Platinum®Taq (Invitrogen) and transcript-specific fluorescein-labeled primers for scd1 and numb. PCR products were separated from free primers on a 5% native polyacrylamide gel, imaged on a Fuji FLA-5000 imager, and quantified using Image Gauge software.

**MSI1 overexpression and qRT-PCR**

Twenty four hours prior to transfection, $10^6$ HEK293T cells were seeded in 10 cm culture plates. Cells were transfected with 2 µg pCDH-CMV-MSI1(FL) or empty
pCDH-CMV-MCS-EF1-Puro using Effectine reagent according the manufacturer's instructions (Qiagen). 24 h after transfection, transfected cells were selected for with 2 µg/mL puromycin (Sigma). RNA and protein were harvested 72 h after transfection. To confirm MSI1 overexpression, RNA was collected using Trizol reagent (Invitrogen) followed by DNAs smear treatment (Turbo DNAfree, Ambion) according to the provided instructions. RNA yield was determined by spectrophotometry and quantitative RT-PCR was performed with an Opticon thermal cycler (Bio-Rad) using an iScript One-Step RT-PCR Kit with SYBR Green (Bio-Rad) according to the manufacturer’s instructions. All assays were performed in triplicate. Data were analyzed by sigmoidal curve fitting according to the method of Rutledge (Rutledge 2004) and normalized to GTF2i or Tubulin.

**SCD western blots**

Cells were lysed with SDS lysis buffer (1% SDS, 10 mM EDTA, 1x EDTA-free protease inhibitors) and protein concentration was determined by the bicinchoninic acid assay method (BCA, Pierce) according to the manufacturer protocol. Samples were boiled for 5 minutes after the addition of 6x Laemmli buffer (9% SDS and 60% glycerol, 375 mM Tris-HCl pH 6.8, 0.015% Bromophenol blue, 12% β-mercaptoethanol). Proteins were separated for 2 h on a 12% SDS-polyacrylamide gel using a Bio-Rad mini-PROTEAN electrophoresis apparatus, transferred to low-fluorescence PVDF membranes (Millipore) for 2 h at 4°C at a constant 70 V in 25 mM Tris, 150 mM glycine, and 10% (v/v) methanol transfer buffer, blocked with
Odyssey Blocking Buffer (OBB, LI-COR Biosciences) for 1 h at room temperature, and probed with 1:1000 dilution of mouse monoclonal anti-SCD antibody (Abcam ab19862) in OBB overnight at 4ºC. After washing with PBS-T, membranes were incubated with fluorescent goat anti-mouse secondary antibodies (LI-COR Biosciences, 926-32210) at 1:15000 in OBB for 1 h at room temperature. Signal was detected using an Odyssey Infrared Imaging System (LI-COR Biosciences). To ensure antibody specificity, shRNA knockdown of SCD1 was performed with two independent shRNA constructs (Open Biosystems V3LHS_305870 and V3LHS_305872) and compared with non-silencing shRNA construct (Open Biosystems RHS4346).

**Total lipid extraction**

Cells were removed from plates with 0.25% Trypsin-EDTA (Gibco) and the total cell number was obtained using a hemocytometer. Each sample contained 1-2x10⁶ cells. After pelleting by centrifugation, cells were resuspended in 200 µL H₂O, immediately followed by 250 µL methanol and 125 µL chloroform/0.01% Butylated hydroxytoluene (BHT). Samples were shaken vigorously for 30 seconds, then an additional 250 µL chloroform/0.01% BHT and 250 µL H₂O was added. After 30 seconds of additional shaking, samples were centrifuged for 5 minutes at 4º C. The chloroform layer was removed to a new tube, and the cell sample was treated with 250 µL of chloroform/0.01% BHT, vigorous shaking, centrifuging and subsequent chloroform layer removal twice more. Combined chloroform layers were washed
with 500 µL KCL, then 500 µL H₂O. Insoluble layers were transferred to glass vials and dried under argon. Samples were sent to the Kansas lipidomics research center for analysis by mass spectrometry as previously described (Welti et al. 2002). 5 independent biological replicates were analyzed for each treatment group. See Appendix I.

**Lipidomics data analysis**

The lipid profile data were acquired at Kansas Lipidomics Research Center (KLRC). Instrument acquisition and method development at KLRC was supported by NSF grants MCB 0455318, MCB 0920663, DBI 0521587, DBI 1228622, Kansas INBRE (NIH Grant P20 RR16475 from the INBRE program of the National Center for Research Resources), NSF EPSCoR grant EPS-0236913, Kansas Technology Enterprise Corporation, and Kansas State University. Lipidomics data were reported as nMoles / 1x10⁶ cells. After removing data points for samples with values below the limit of detection for the instrumentation (0.02 nMoles / 1x10⁶ cells), all data points for which no signal was detected in either treatment group were removed. A pseudocount of 0.002 nMoles / 1x10⁶ cells was added to all remaining data points. Outliers were eliminated using a Q test. Significant differences between treatments for each lipid class were determined using a Student’s t-test and an FDR of 0.05. See Appendix I.

**Critical micelle concentration determination**
The critical micelle concentration (CMC) for oleic acid in equilibration buffer was
determined using N-phenyl-1-naphthylamine (NPN), a compound that fluoresces
when sequestered into micelles (Hagihara et al. 2002; Hasegawa et al. 2008). 25 µM
NPN was incubated with varying amounts of oleic acid for 30 minutes. Fluorescence
intensity was determined with a Victor V3 plate reader using a 355 nm excitation
filter and a 460 nm emission filter. Fluorescence intensity data were plotted as a
function of fatty acid concentration. The CMC was determined by a two state
segmented linear regression to identify the breakpoint between baseline
fluorescence and fluorescence caused by NPN association with micelles.
Chapter 4: Concluding remarks and future directions
The initial goal of the research presented here was to identify a small molecule inhibitor of the RNA-binding protein MSI1. The small molecule screen and subsequent characterization reveal that the intermediary metabolite oleic acid inhibits MSI1 RNA-binding activity. Additionally, I identified a novel putative MSI1 target, stearoyl-CoA desaturase (SCD), which couples MSI1 to metabolic regulation of lipid production in proliferating and differentiating cells. However, I am left with a number of questions both biochemical and biological in nature. The mechanism of regulation is not yet completely defined, and the structure of the bound complex remains unsolved. Additionally, the biological significance of a regulatory feedback loop between MSI1 and SCD is not yet understood, and we have not yet characterized any additional natural fatty acid derivatives that may inhibit MSI1 activity in cells.

**Biochemical characterization of Musashi-1 inhibition**

18-22 carbon ω-9 fatty acids inhibit MSI1 RNA-binding activity

The inhibitor screen reported here consisted of over 30,000 compounds across two libraries. 4 potential inhibitors were confirmed through secondary screening FP and F-EMSA assays. Of these, two were of biological interest – GW7647 and oleic acid - because they both activate PPAR nuclear hormone proteins, which are metabolically regulated transcription factors. GW7647 was not a particularly strong inhibitor. Oleic acid was a better inhibitor, as determined by its Ki, and is abundant in myelin, which is produced by mature oligodendrocyte cells in the
central nervous system. MSI1 is abundant in oligodendrocyte progenitor cells, but not in mature myelinating oligodendrocytes. It was therefore conceivable that as oleic acid levels rise during differentiation, oleic acid could be acting to turn off MSI1 function during the transition to maturity.

A survey of fatty acids and fatty acid derivatives revealed that a number of structural features are important for inhibition. Modification of the carboxylate is not tolerated, indicating that the charge and/or size of that group is important. Altering the number of carbons between the double bond and the omega end of the fatty acid molecule reduced inhibition, which indicates a length requirement for the omega tail. Finally, changing the orientation or saturating the cis double bond eliminates inhibition, indicating that a kink in the fatty acid structure is essential for inhibitory function. The list of fatty acids tested was not exhaustive, so it is possible that some unidentified modifications to the fatty acid would be tolerated.

Oleoyl-CoA and erucyl-CoA are abundant and soluble in cells, so I sought to test whether these molecules are effective inhibitors. Oleoyl-CoA and erucyl-CoA both inhibited MSI1 RNA-binding activity, but thin layer chromatography revealed the stocks contained some free oleic acid and erucic acid. I therefore could not rule out the possibility that inhibition was due to association with the free fatty acid. Other derivatives that should be tested include 2-hydroxy oleic acid (2OHOA) and Iloprost. 2OHOA is in clinical trials as an anti-cancer drug for cancers of the epithelial and neural lineage (Teres et al. 2012). Preliminary results indicate that glioma tumors treated with 2OHOA show decreased proliferation and increased
differentiation (Teres et al. 2012). An additional hydroxyl at C2 may even strengthen the electrostatic properties of the carboxylate group. Iloprost is a prostaglandin I2 analog that has been shown to induce differentiation in oligodendrocyte progenitor cells (Takahashi et al. 2013). Structurally, it is less clear how Iloprost might associate with MSI1, but it is a fatty acid derivative and knowing whether or not it inhibits MSI1 will add to our understanding of the mechanism of inhibition. Finally, endocannabinoids and eicosanoids are signaling molecules derived from fatty acids that play an essential role in brain and organ development. While there are many molecules that fall into these classes, strategic testing should be conducted to determine if they have Msi1 inhibitory properties.

A model for the mechanism of inhibition

Biochemical and biophysical characterization of the association between MSI1 and oleic acid revealed a model for the mechanism of inhibition. There is a hydrophobic binding pocket located on the alpha-helical face of RRM1. This binding pocket is evolutionarily conserved in all MSI family proteins for which structures have been solved. The sequence of RRM1 is highly conserved among all MSI family proteins, which indicates that the binding pocket structure is likely also conserved. Modeling studies suggest that oleic acid docking with this binding pocket is thermodynamically favorable. The double bond allows 18-22 carbon ω-9 fatty acids to bend around MSI1 to enable interaction of the carboxylate group with one of
several basic residues on the flexible loop region connecting the alpha helical and beta sheet faces of RRM1.

The structure-activity relationship studies support this model, as modification of the carboxylate, omega-end length, or presence/orientation of the double bond are not tolerated. Circular dichroism and tryptophan fluorescence experiments indicate that inhibition occurs through direct interaction of MSI1 with the fatty acid, which induces a change in the secondary structure of the protein. Analytical ultracentrifugation indicates that the majority of the protein remains monomeric, ruling out apparent inhibition by aggregation or oligomerization of protein.

Attempts to solve the oleic acid-bound MSI1 structure by NMR were unsuccessful. The protein behaves well for NMR, and has been used in NMR studies previously (Kurihara et al. 1997; Nagata et al. 1999; Miyanoiri et al. 2003; Ohyama et al. 2011). However the CMC of oleic acid in our buffer system is 75 µM, which precludes use in NMR because concentrations must be higher to obtain a saturated binding spectrum at good resolution. We did observe peak broadening and signal loss upon oleic acid titration, with the greatest signal losses mapping to key residues in the areas predicted to bind oleic acid in our docked model. These include W29, Q30, L36, C49, L50, R53, S60, G62, V74, T89, K98, which are computationally predicted to either interact directly or undergo a structural change upon oleic acid association. The observed loss of intensity is likely due to protein precipitation in
the presence of high concentrations of oleic acid, association with oleic acid micelles, or a combination of the two.

To obtain a bound structure, it would be interesting to attempt an NMR method using detergents that have been used previously to solve structures with insoluble ligands and/or proteins. These detergents create a non-reactive micellar environment in which the protein/ligand can be suspended. For instance, the detergent ($^{2}$H$_{38}$)dodecylphosphocholine (DCP) has been used to study the structures of several insoluble membrane proteins in NMR studies. There are a number of additional detergents that can also be screened for appropriate NMR conditions. It would be worth trying some of these techniques, as a bound structure would enable detailed mechanistic study and development of inhibitor derivatives and protein mutants to separate RNA-binding from oleic acid-binding functionality.

To further probe the mechanism of inhibition in lieu of a bound NMR structure, Francesca Massi used molecular dynamics simulations to investigate the probable structural changes that MSI1 undergoes upon oleic acid binding. Her results further support a model in which oleic acid binds to the alpha helical face of MSI1 RRM1, but indicate a surprising amount of flexibility and structural reorganization when oleic acid is bound. This simulation indicates that residues involved in RNA binding are pulled away from the RNA-binding beta sheet face. Many of these demonstrated signal intensity loss in the NMR titration of oleic acid, which supports the molecular dynamics simulation prediction that they may be undergoing a structural change.
Mutational analysis of several residues predicted to be important for oleic acid binding further supports the computational model and simulation. Residues that surround the opening (G35) and comprise the hydrophobic pocket (G64 and H83) on the alpha helical face are important for fatty acid inhibition. Additionally, residues predicted to interact with the carboxylate of the inhibitory fatty acids (R53 and R61) are also required to maintain inhibition. Further mutational analysis will be required to identify a separation-of-function mutant MSI1 that maintains RNA-binding ability but is immune to the inhibitory activity of 18-22 carbon $\omega$-9 unsaturated fatty acids. A true separation-of-function mutant will enable in-vivo experimentation to determine whether fatty acids are regulating Msi1 in cells. Generating such a mutant will likely be accomplished through multiple strategic mutations that preserve RNA-binding structural properties but perturb the structural properties that allow oleic acid association. A high-resolution NMR structure of oleic acid bound to MSI1 will undoubtedly help guide mutational efforts by illuminating the interactions that stabilize fatty acid binding.

**Biological characterization of Musashi-1 inhibition**

*A model for MSI1 in proliferation*

MSI1 has been implicated as a cellular proliferation factor in both healthy and cancerous cells of neural and epithelial lineages. Overexpressing MSI1 prevents differentiation, while knockdown results in decreased proliferation and increased rates of apoptosis in both healthy and tumorous cells. The biochemical results
presented here indicate that MSI1 is inhibited by a specific class of fatty acids that are abundant in proliferating cells and in oligodendrocytes, which are a specialized type of differentiated non-proliferative cell.

Treatment of MSI1-expressing oligodendrocyte progenitor cells with oleic acid results in decreased proliferation, while treatment with non-inhibitory stearic acid does not decrease proliferation. HEK cells, which do not express MSI1, respond to neither oleic acid nor stearic acid treatment. While these data do not demonstrate that MSI1 is inhibited by oleic acid in cells, the phenotype is in agreement with that of MSI1 knockdown in oligodendrocyte progenitor cells by RNAi. These experiments should be repeated in a HEK background with a mutant form of MSI1 that is unresponsive to fatty acid inhibitors. Luciferase reporters that contain the 3’ UTR of a MSI1-responsive gene, such as Numb, would provide more convincing data to determine whether fatty acids are directly inhibiting MSI1 regulatory activity.

The data presented here also indicate that MSI1 positively regulates SCD, the enzyme that produces ω-9 unsaturated fatty acids by inserting a double bond in 16 and 18 carbon saturated fatty acids. Stearic acid is the most prevalent substrate of SCD, and the product of stearic acid desaturations is oleic acid. Oleic acid can be modified into longer-chain fatty acids and/or incorporated into phospholipids for membranes, energy storage molecules such as cholesterol esters and triacylglycerols, or further desaturated into polyunsaturated fatty acids for signaling molecules.
Lipidomics analysis of MSI1-expressing HEK293T cells indicates that downstream products of SCD are specifically upregulated. These are predominantly cholesterol esters and triacylglycerols, both of which are more abundant in mice overexpressing SCD1 and less abundant in liver SCD1 knockout mice. Cholesterol esters and triacylglycerols are energy storage molecules essential in highly proliferative cells. Interestingly, cancerous cells display elevated SCD expression, and SCD knockdown results in decreased proliferation and increased apoptosis – a pattern startlingly similar to that of MSI1.

Together, the data presented here indicate a model in which MSI1 and SCD exist in a feedback loop to maintain homeostasis of proliferative epithelial and neural progenitor cells (Figure 4.1). Loss of homeostasis could contribute to the increased proliferation observed in tumors with increased MSI1 and/or SCD abundance. There are two obvious scenarios for this loss of homeostasis. 1) Loss of MSI1/SCD homeostasis drives tumorigenesis: MSI1 may be positively regulated by an unknown factor, resulting in positive regulation of SCD and therefore increased production of ω-9 unsaturated fatty acids. Assuming MSI1 is abundant enough, the increase in unsaturated fatty acid concentration might be insufficient to inhibit MSI1 and restore homeostasis. 2) Loss of MSI1/SCD homeostasis is a by-product of tumorigenesis: The machinery that incorporates ω-9 unsaturated fatty acids into membranes, energy storage, and signaling molecules might be more active, thereby relieving feedback inhibition of MSI1 activity and subsequently stimulating SCD
Figure 4.1

18-22:1 ω-9 monounsaturated fatty acids

Proliferation
Energy storage
Signaling
Figure 4.1: Model of MSI1/SCD homeostasis in normal, proliferative cells.
expression. This scenario would likely occur in cancerous cells where proliferative checks and balances have already been disrupted.

Understanding whether MSI1 is driving or reacting to tumorigenesis will be an essential step in characterizing its role in cancer. To date, no direct regulators of MSI1 expression have been identified, but there is some evidence for their existence. In colorectal tumors, MSI1 expression is upregulated upon NOTCH3 stimulation, and inhibition of NOTCH3 signalling results in elevated levels of NUMB, a protein that is normally translationally repressed by MSI1 (Pasto et al. 2014). Additionally, the absence of MSI1 protein in mature neural and epithelial cells indicates that transcription and/or translation are being repressed.

Identification of candidate transcriptional and post-transcriptional regulators, while not trivial, can be accomplished using the growing body of sequencing data. We can use genomic information to identify potential transcription factor and post-transcriptional regulatory factor binding sites. Depending upon the candidate regulatory factor, data sets may be available to elucidate their binding targets, gene expression changes upon knockdown, or the ribosome profile changes upon knockdown. Candidate targets can also be tested in vitro using biochemical binding assays and in vivo using reporter assays. To design efficient and effective anti-cancer therapies, we must first understand the pathways that lead to tumorigenesis. MSI1 may be a good therapeutic target regardless of whether it is
driving or supporting tumorigenesis, but we ultimately must understand whether it plays a driving or supporting role in tumorigenesis.

A model for MSI1 in oligodendrocyte differentiation

MSI1 is required for proper neural development. MSI1 knockout mice display abnormal brain development, and embryonic neurospheres cultured from these mice display fewer mature cell types (Sakakibara et al. 2002). MSI1 depletion in OPCs results in differentiation, reduced proliferation, and increased apoptosis (Dobson et al. 2008). Only two MSI1 regulatory targets have been characterized in mammalian neural tissue, although recent RIP ChIP studies have identified new putative targets that require in vitro and in vivo characterization. MSI1 translationally represses Numb, an inhibitor of NOTCH signaling, and P21WAF1, a cell cycling inhibitor. Repression of these targets by MSI1 may be involved in preventing oligodendrocyte progenitor cell differentiation, but the regulatory mechanisms governing MSI1 expression levels remain unknown.

This research supports a novel hypothesis for the regulation of MSI1 levels in oligodendrocyte progenitor cells. 18-22 carbon ω-9 unsaturated fatty acids are the most abundant fatty acids in the myelin membrane produced by mature oligodendrocytes. MSI1 is expressed in oligodendrocyte progenitor cells, but not in mature, myelinating oligodendrocytes. However, the MacNicol lab has shown that MSI1 protein may linger during differentiation, indicating a need to inactivate the existing active protein (MacNicol et al. 2008). The data herein indicate that MSI1
Figure 4.2

- **MSH1**
  - **P21WAF-1** translation
    - **p21**
      - Cell Cycle
  - m-numb translation
    - **numb**
      - Notch Signaling
      - 18-22:1ω-9 monounsaturated fatty acids
      - Differentiation
      - Proliferation
Figure 4.2: Model of MSI1 homeostasis in normal OPCs.
may be inactivated by the same fatty acids that become abundant as the differentiating cells prepare to produce myelin.

This hypothesis proposes a "switch" mechanism to drive differentiation, but again, there may be two scenarios in which the "switch" is flipped, depending upon whether MSI1 expression is turned off to help initiate differentiation, or as a result of differentiation (Figure 4.2). 1) Oleic acid levels in proliferating progenitor cells may serve to prevent over-activity of MSI1, thereby maintaining proliferation homeostasis. If MSI1 is downregulated in preparation for differentiation, the switch would be flipped, enabling the existing free fatty acids to inactivate the remaining MSI1 protein so differentiation can proceed. This scenario assumes that MSI1 expression is turned off as differentiation is initiated. 2) Oleic acid levels may be increased in preparation for differentiation, thereby flipping the switch and inactivating existing MSI1 protein. This model assumes that MSI1 expression is turned off after the differentiation machinery is activated.

Testing these models is difficult with existing technology. Cell samples would need to be precisely staged, using fluorescent markers for progenitor cells (A2B5 and O4) and mature oligodendrocytes (MBP and O1). Immature oligodendrocytes that are undergoing the transition may be positive for both markers. Lipidomics mass spectrometry analysis will reveal how the lipid profile changes at each stage, indicating whether fatty acid levels increase during or after differentiation. qRT-PCR and Western blot analysis will reveal how the MSI1 expression profile changes at each stage, indicating whether MSI1 RNA and protein levels decrease during or after
the transition. Acquiring enough cellular material, especially for the differentiated oligodendrocyte population, will prove challenging. Additionally, lipidomics analysis will not provide a sensitive reading for free fatty acids, but the downstream products can be a good indicator of fatty acid levels in cells.

**Metabolic regulation**

Changes in metabolic state induce compensatory changes in gene expression. In Eukaryotes, transcriptional control of gene expression in response to metabolic changes has been widely studied. Metabolites including fatty acids, vitamins, alcohols, amino acids, and nucleotides and derivatives are known to regulate transcription. It is surprising that more is not known about the effect that intermediary metabolites have on post-transcriptional regulation.

In the introduction I highlighted several examples of post-transcriptional mechanisms that respond to intermediary metabolites. These examples may represent unique, non-pervasive mechanisms of post-transcriptional regulation. However, it is likely that new systems-level approaches will identify additional examples. If metabolic control of post-transcriptional regulation is indeed widespread in eukaryotes as it is in bacteria, there will be broad implications for the development of novel therapeutics that target metabolite sensitive RNA-binding proteins involved in disease processes. The Musashi-1/fatty acid/SCD feedback loop described here may provide the necessary link to spur additional research into post-transcriptional regulation of metabolic processes.
Appendix 1: Lipidomics protocol and data analysis
Note: I recommend contacting the Kansas Lipidomics Research Center (KLRC) to discuss your application. The director and lab manager were extremely helpful. They provided suggestions for experimental design, suggested application-specific sample preparation techniques, and assisted with the data analysis.

**Total lipid extraction and preparation for lipidomics analysis**

1) Prepare and treat cells – must prepare at least 5 plates/flasks per treatment for proper statistical analysis.

2) Collect cells from plates or flasks.
   a. Use preferred method, such as 0.25% Trypsin-EDTA (Gibco)

3) Count total cell number from each plate using a hemocytometer.

4) Loosely pellet cells (800xG for 2 minutes) and aspirate media.

5) Resuspend pellet in 200 µL H₂O per 1-2x10⁶ cells.

6) Add 250 µL methanol and 125 µL chloroform/0.01% Butylated hydroxytoluene (BHT).

7) Shake samples vigorously for 30 seconds.

8) Add an additional 250 µL chloroform/0.01% BHT and 250 µL H₂O.

9) Shake samples vigorously for 30 seconds.

10) Centrifuge for 5 minutes at 4º C at maximum speed.

11) Remove the chloroform layer and transfer it to a new tube.

12) Add an additional 250 µL of chloroform/0.01% BHT to the cell lysate tubes.
13) Shake samples vigorously for 30 seconds.
14) Centrifuge for 5 minutes at 4º C at maximum speed.
15) Remove the chloroform layer and transfer it to the tube containing the chloroform layer from step 10.
16) Repeat steps 11-14 once more. You should have a tube with the combined chloroform layers from three total chloroform extractions. This will contain the total cellular lipids from each sample.
17) Add 500 µL KCL to the combined chloroform layers.
18) Centrifuge for 5 minutes at 4º C at maximum speed.
19) Remove the chloroform layer to a new tube.
20) Add 500 µL H₂O to chloroform layer.
21) Centrifuge for 5 minutes at 4º C at maximum speed.
22) Remove the chloroform layer to a glass vial.
23) Dry samples under argon.
24) Pack samples in dry ice for shipping.

**Lipidomics data analysis**

1) Use the data set from the KLRC in which Lipidomics data are reported as nMoles per 1x10⁶ cells.
2) Remove data points for samples with values below the limit of detection for the instrumentation (0.02 nMoles / 1x10⁶ cells).
3) Remove data points for which no signal was detected in both treatments.
4) Add a pseudocount of 0.002 nMoles / 1x10⁶ cells to all remaining data points to enable fold-change calculations.
5) Eliminate outliers from within each set of 5 data points using a Q test (Prism).

6) Use a Student’s t-test and an FDR of 0.05 to identify lipids that change significantly.
References


Brookes E, Demeler B, Rosano C, Rocco M. 2010. The implementation of SOMO (SOlution MOdeller) in the UltraScan analytical ultracentrifugation data analysis


Demeler B, Gorbet G, Zollars D, Brookes E, Cao W, Dubbs B. 2011. UltraScan III:
UltraScan-III, a comprehensive analysis software for analytical ultracentrifugation
experiments. The University of Texas Health Science Center at San Antonio,
Dept. of Biochemistry.

Demeler B, Saber H, Hansen JC. 1997. Identification and interpretation of complexity in

Demeler B, van Holde KE. 2004. Sedimentation velocity analysis of highly

Di Loreto S, D'Angelo B, D'Amico MA, Benedetti E, Cristiano L, Cinque B, Cifone MG,
Ceru MP, Festuccia C, Cimini A. 2007. PPARbeta agonists trigger neuronal
differentiation in the human neuroblastoma cell line SH-SY5Y. *Journal of cellular
physiology* **211**: 837-847.

Dinger ME, Gascoigne DK, Mattick JS. 2011. The evolution of RNAs with multiple

regulates oligodendrocyte lineage cell differentiation and survival. *Glia* **56**: 318-
330.

dehydrogenase with secondary and tertiary RNA structural elements of the
hepatitis A virus 3' translated and non-translated regions. *J Gen Virol* **84**: 403-
414.


Franklin RJ, Bayley SA, Milner R, Ffrench-Constant C, Blakemore WF. 1995. Differentiation of the O-2A progenitor cell line CG-4 into oligodendrocytes and
astrocytes following transplantation into glia-deficient areas of CNS white matter.

*Glia* **13**: 39-44.


Menendez JA, Papadimitropoulou A, Vellon L, Lupu R. 2006. A genomic explanation connecting "Mediterranean diet", olive oil and cancer: oleic acid, the main monounsaturated fatty acid of olive oil, induces formation of inhibitory "PEA3


Phok K, Moisan A, Rinaldi D, Brucato N, Carpousis AJ, Gaspin C, Clouet-d’Orval B.  
2011. Identification of CRISPR and riboswitch related RNAs among novel 

Sakakibara S, Okano H. 2003. Identification of a putative intestinal stem cell and 

Prasad CK, Mahadevan M, MacNicol MC, MacNicol AM. 2008. Mos 3’ UTR regulatory 
differences underlie species-specific temporal patterns of Mos mRNA 
cytoplasmic polyadenylation and translational recruitment during oocyte 

Preiss T, Hall AG, Lightowlers RN. 1993. Identification of bovine glutamate 

Ravindran G, Devaraj H. 2012. Aberrant expression of CD133 and musashi-1 in 
preneoplastic and neoplastic human oral squamous epithelium and their 

Reines SA, Cantor CR. 1974. New fluorescent hydrazide reagents for the oxidized 3’-
terminus of RNA. *Nucleic Acids Res* **1**: 767-786.

of the putative gut stem cell marker Musashi-1 induces tumorigenesis through 

Rios-Esteves J, Resh MD. 2013. Stearoyl CoA desaturase is required to produce active, 


Scaglia N, Chisholm JW, Igal RA. 2009. Inhibition of stearoylCoA desaturase-1
inactivates acetyl-CoA carboxylase and impairs proliferation in cancer cells: role

Scaglia N, Igal RA. 2008. Inhibition of Stearoyl-CoA Desaturase 1 expression in human

Schmauss C. 2005. Regulation of serotonin 2C receptor pre-mRNA editing by serotonin.

Schneider BD, Leibold EA. 2000. Regulation of mammalian iron homeostasis. *Curr Opin

CD44-positive colorectal adenoma cells express the potential stem cell markers
musashi antigen (msi1) and ephrin B2 receptor (EphB2). *The Journal of
pathology* 213: 152-160.

Shi H, Kokoeva MV, Inouye K, Tzameli I, Yin H, Flier JS. 2006. TLR4 links innate


Siddall NA, Kalcina M, Johanson TM, Monk AC, Casagranda F, Been RP, McLaughlin
EA, Hime GR. 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


