HUMAN GW182 PARALOGS ARE THE CENTRAL ORGANIZERS FOR RNA-MEDIATED CONTROL OF TRANSCRIPTION

APPROVED BY SUPERVISORY COMMITTEE

_____
David R. Corey, Ph.D, Advisor

_____
Nicholas Conrad, Ph.D, Chairman

_____
David Mangelsdorf, Ph.D.

_____
William Kraus, Ph.D.

_____
James Brugarolas, M.D., Ph.D.

ACKNOWLEDGEMENTS

I would like to thank my wife Marisol for always being my rock through my quals, endless presentations, the tiring job search, and through my road to publication, both times. She has given me patience, strength, support, encouragement, and love for the last three and a half years and has made me a better person. I can't imagine doing this without her.

I would like to thank my parents, Jim and Jackie, for all the love, support and encouragement throughout my life. I could not have achieved my goals without them. I would also like to thank my mother and father-in law, Donna and Steve, for all the love and support during the time I have known them.

I would like to thank everyone in the Corey lab that has helped shape me into the scientist I am today. I would like to thank Dr. Keith Gagnon, Dr. Liande Li, and Dr. Masayuki Matsui for teaching me new techniques and experimental guidance. I would especially like to thank my friend and previous colleague Roya Kalantari for helping train and guide me on this journey in the Corey laboratory. We worked very closely together and she inspired my successful project.

I would like to thank my dissertation committee, Dr. David Mangelsdorf, Dr. Nicholas Conrad, Dr. James Brugarolas, and Dr. Lee Kraus for their guidance.

I would finally like to thank my graduate advisor Dr. David Corey. He has been an excellent mentor, teacher, and support system throughout my graduate career. I have seen myself grow in my critical thinking, writing, and presentation skills because of his guidance and these skills have helped me in my career path.

HUMAN GW182 PARALOGS ARE THE CENTRAL ORGANIZERS FOR RNA-MEDIATED CONTROL OF TRANSCRIPTION

by

JESSICA ANNE HICKS

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2017

HUMAN GW182 PARALOGS ARE THE CENTRAL ORGANIZERS FOR RNA-MEDIATED CONTROL OF TRANSCRIPTION

Jessica Anne Hicks, Ph. D.

The University of Texas Southwestern Medical Center at Dallas, 2017

Supervising Professor: David R. Corey, Ph.D.

RNA interference (RNAi) is an endogenous mechanism for regulating gene expression that can be manipulated for experimental or therapeutic purposes to knockdown protein expression. In the mammalian cell cytoplasm, small RNAs direct RNAi proteins to inhibit translation by regulating mRNA stability. In non-mammalian cell nuclei, RNAi proteins control gene transcription. Recently, in mammalian cell nuclei, the RNAi proteins were shown to control transcription and splicing. Despite what is known about RNAi factors in non-mammalian cell nuclei, the mechanisms of mammalian RNA-mediated nuclear regulation are not well understood and remain controversial, hindering the effective application of nuclear RNAi and blinding investigation of its natural regulatory roles.

Argonaute 2 (AGO2) and TNRC6A (GW182) are the core proteins of RNAi. To better understand RNAi-mediated nuclear gene regulation, my goal was to use semi-quantitative (SINQ) mass spectrometry analyses on purified protein complexes to build a protein interaction network of nuclear AGO2 and TNRC6A.

The stable interactions of AGO2 detected with this protocol were the TNRC6A, B, and C paralogs. While this did not reveal many proteins, it did provide a new direction to take with mass spectrometry analyses. Since the TNRC6 paralogs are stable interacting partners, analysis on those protein complexes were performed to reveal a new shell of nuclear RNAi interactions.

Mass spectrometry of TNRC6A protein complexes revealed these proteins are central to forming interactions between the RNAi machinery and many proteins involved in transcriptional regulation. TNRC6A interactions include the AGO proteins, CCR4-NOT complex, histone modifiers, and the mediator complex. In addition, novel interactions with four DNA damage repair proteins were identified, providing another direction for future investigation. Functional analysis revealed that TNRC6, AGO2, NAT10 (histone acetylation), WDR5 (H3K4me3), and MED14 (Mediator) proteins are involved in RNA-mediated COX-2 transcriptional activation. Taken together, the mass spectrometry and functional experiments provide evidence that these RNA-AGO-TNRC6 complexes act globally in cell nuclei to regulate transcription. These findings describe protein complexes capable of bridging RNA-mediated sequence-specific recognition of noncoding RNA transcripts with the regulation of gene transcription. The significance of my data is that it can lead to new advances in RNAi capabilities beyond canonical applications.

Table of Contents

ABSTRACT……...iv

PRIOR PUBLICATIONS……..xi

LIST OF FIGURES……..xii

LIST OF TABLES……xiv

LIST OF ABBEVIATIONS……..xv

CHAPTER 1—HISTORICAL BACKGROUND AND SIGNIFICANCE

1.1. Introduction……....1

1.2. Nuclear RNAi in non-mammalian systems……....2

1.3. Evidence of RNAi in mammalian cell nuclei……...3

1.4. The Argonaute Proteins……6

1.5. Functional Domains of the Argonaute Proteins……6

1.6. The TNRC6 (GW182) protein family……...7

1.7. The TNRC6 (GW182) Protein Family Functional domains……..8

1.8. RNA-mediated transcriptional silencing……...11

1.9. RNA-mediated transcriptional activation……..12

1.10. Activation in one context, repression in another: lessons from the progesterone receptor……13

1.11. Nascent transcripts are the target for RNA-mediated transcriptional regulation……14

1.12. RNAi factors are involved in transcriptional silencing and activation……...14

1.13. Lessons from COX-2 activation: filling in more mechanistic gaps……15

1.14. RNA-mediated regulation of alternative splicing……16

1.15. Endogenous roles for transcriptional regulation by RNAi……..18

1.16.Conclusions……19

1.17. Hypothesis……20

CHAPTER 2—POST-TRANSLATIONS MODIFICATIONS OF AGO2 AND TNRC6A

2.1. Introduction……..29

2.2. Post-translational modifications of RNAi factors……30

2.3. Purification of AGO2 in cytoplasmic and nuclear extracts for PTM analysis……33

2.4. Unbiased PTM search for FLAG and endogenous AGO2 modifications….....34

2.5. Cytoplasmic and nuclear FLAG-AGO2 PTMs……35

2.6. Identification of cytoplasmic and nuclear endogenous AGO2 PTMs……36

2.7. Identification of cytoplasmic and nuclear endogenous TNRC6A PTMs……37

2.8. Conclusions……37

2.9. Improving methods for identification of PTMs……38

CHAPTER 3—IDENTIFYING STABLE INTERACTING FACTORS OF AGO2

3.1. Introduction……..48

3.2. Prioritizing candidate interacting partners for AGO2……51

3.3. Nuclear interactions of AGO2……52

3.4. Focusing candidate identification using data from orthogonal analyses……...55

3.5. Stable interactions of TNRC6 protein family in the nucleus……56

3.6. Conclusion……...57

3.7. Overcoming challenges……58

3.8. Improving methods for candidate identifications……..60

CHAPTER 4—HUMAN GW182 PARALOGS ARE THE CENTRAL ORGANIZERS FOR RNA-MEDIATED CONTROL OF TRANSCRIPTION

4.1. Introduction……...69

4.2. Immunoprecipitation and mass spectrometry of TNRC6A complexes……71

4.3. Verifying TNRC6A protein interactions……73

4.4. AGO variants associate with nuclear TNRC6A……74

4.5. TNRC6A associates with components of the anaphase promoting complex……75

4.6. TNRC6A associates with proteins that modify histones……...76

4.7. TNRC6A associates with members of the CCR4-NOT complex……..76

4.8. TNRC6A associates with members of the mediator complex……77

4.9. Association of candidate proteins after fractionation……..78

4.10. Other proteins implicated by mass spectrometry……..79

4.11. Functional analysis: Impact on RNA-mediated transcriptional activation……..80

4.12. Summary of nuclear RNAi……...82

4.13. Previous lessons from RNA control of COX-2 activation……82

4.14. TNRC6 is the central organizing factor controlling nuclear RNAi……83

4.15. Conclusion……..85

4.16. Overcoming challenges…….....86

CHAPTER 5—METHODS

5.1. Cell culture……101

5.2. Extract preparation……...102

5.3. Western blots and extract analysis……102

5.4. Immunoprecipitation for mass spectrometry analysis……103

5.5. Affinity chromatography for AGO2 PTM analysis……104

5.6. Mass spectrometry data analysis……...105

5.7. Mass spectrometry data analysis for AGO2 and TNRC6A protein complexes……..106

5.8. Mass spectrometry data analysis for PTMs of AGO2 and TNRC6A……106

5.9. IBAQ analysis for stoichiometry……..107

5.10. Co-immunoprecipitations……...107

5.11. Chromatographic separations……...108

5.12. Double transfection assays……...109

5.13. Chromatin Immunoprecipitation (ChIP)……..109

CHAPTER 6—FUTURE DIRECTIONS AND CONCLUDING REMARKS

6.1. Introduction……....111

6.2. TNRC6A ChIP to determine endogenous gene targets……111

6.3. RNA-mediated Transcriptional Silencing and Activation of the Progesterone Receptor (PR)……112

6.4. Mass Spectrometry of nuclear TNRC6B and TNRC6C complexes……...113

6.5. Mass spectrometry of nuclear CNOT1 complexes……114

6.6. Interactions of the TNRC6A N-term and C-term domains……115

6.7. Localization of TNRC6A binding partners after knockdown……..116

6.8. The role of TNRC6 paralogs in DNA Damage Response (DDR)……..117

6.9. Concluding Remarks……119

BIBLIOGRAHPY……124

PRIOR PUBLICATIONS

Borghese CM, **Hicks JA**, Lapid DJ, Trudell JR, Harris RA. (2014). $GABA_A$ receptor transmembrane amino acids are critical for alcohol action: disulfide cross-linking and alkyl methanethiosulfonate labeling reveal relative location of binding sites. *J. Neurochem*. **128**. 363-375.

Warshauer JT, Lopez X, Gordillo R, **Hicks JA**, Holland WL, Anuwe E, Blankfard MD, Scherer PE, Lingvay I. (2015). Effect of Pioglitazone on Plasma Ceramides in Adults with Metabolic Syndrome. *Diabetes Metab. Res. Rev.* **31(7)**. 734-744.

Liu J, Hu J, **Hicks JA**, Prakash TP, Corey DR. (2015). Modulation of Splicing by Single-Stranded Silencing RNAs. *Nucleic Acid Ther.* **25(3)**. 113-120.

Kalantari R, **Hicks JA**, Li L, Gagnon KT, Sridhara V, Lemoff A, Mirzaei H, Corey DR. (2016). Stable Association of RNAi Machinery is Conserved Between the Cytoplasm and Nucleus of Human Cells. *RNA*. **22(7).** 1085-1098.

Hicks JA, Li L, Matsui M, Chu Y, Volkov O, Johnson KC, Corey DR. (2017). Human GW182 Paralogs are the Central Organizers for RNA-Mediated Control of Transcription. *Cell Reports*. In progress.

LIST OF FIGURES

Figure 1.1: RNA Interference in non-mammalian cell nuclei……...21

Figure 1.2. General structure of the four human argonaute proteins……...22

Figure 1.3. The role of TNRC proteins in miRNA gene silencing in the cytoplasm……...23

Figure 1.4. The general structure of TNRC6 proteins……..24

Figure 1.5. RNA-mediated transcriptional regulation in human cells……25

Figure 1.6. RNA-mediated activation of transcription at the COX-2 promoter……26

Figure 1.7. Two models of RNA-mediated alternative splicing……27

Figure 1.8. Model of R-loops cause transcriptional termination due to recruitment of RNAi factors……28

Figure 2.1. Representation of PTMs for AGO2 and TNRC6 RNAi factors……41

Figure 2.2. Analysis of cellular extracts and samples prior to mass spectrometry analysis of PTMs.....43

Figure 2.3. Example of a coomassie stain of an unsuccessful gel sample from an AGO2 purification……..47

Figure 3.1. Identification of candidates from nuclear immunoprecipitations (IPs)……..63

Figure 3.2. Co-immunoprecipitations of Candidate Interacting Factors……..64

Figure 3.3. Identification of candidates from whole cell immunoprecipitations (IPs)……65

Figure 3.4. Model of AGO2 interactions with TNRC6 paralogs in the nucleus……...66

Figure 3.5. Coomassie staining of immunoprecipitations prior to desalting……67

Figure 4.1. Experimental setup for mass spectrometry experiments……...88

Figure 4.2. Identification of candidate protein partners for TNRC6A……89

Figure 4.3. Stoichiometry analysis of mass spec samples with iBAQ values……91

Figure 4.4. Validating the interactions of candidate proteins in Turbonuclease-treated nuclear extracts……..92

Figure 4.5. Supporting data for candidate protein verification…….....94

Figure 4.6. Sequential size exclusion and anion exchange chromatography to detect potential complexes with TNRC6A……95

Figure 4.7. Functional roles for candidate proteins during RNA-mediated activation of *COX-2* expression……97

Figure 4.8. ChIP-qPCR showing enrichment of WDR5 at the COX-2 promoter upon addition of RNA12nc (25 nM)……98

Figure 4.9. Effects of Knockdown of ANAPC1 in HeLa Cells……..99

Figure 4.10. Model for TNRC6A interactions……100

Figure 6.1. Interacting proteins of CNOT1……120

Figure 6.2. Cellular localization of RNAi factors after TNRC6A/B/C knockdown……121

Figure 6.3. The role of RNAi factors in the DNA damage response (DDR)……122

LIST OF TABLES

Table 2.1. Table of characterized AGO2 PTMs for cytoplasmic function.....42

Table 2.2. Table of FLAG-AGO2 PTMs from triplicate samples.....44

Table 2.3. Table of endogenous AGO2 PTMs from one sample.....45

Table 2.4. Table of endogenous TNRC6A PTMs from duplicate (cytoplasm) and triplicate (nucleus) samples.....46

Table 3.1. Tables representing significant proteins in each nuclear AGO2 mass spectrometry sample analyzed by SAINT.....68

LIST OF ABBREVIATIONS

ABD	AGO binding domain
AE	Anion exchange
AGO	Argonaute
APC	Anaphase promoting complex
ASO	Antisense oligonucleotide
ChIP	Chromatin Immunoprecipitation
COX-2	Cyclooxygenase-2
DDR	DNA damage response
DSB	Double strand break
ER	Endoplasmic Reticulin
FPLC	Fast protein liquid chromatography
HA	Hemagglutin
IBAQ	Intensity based quantification
IP	Immunoprecipitation
mRNA	Messenger RNA
miRNA	MicroRNA
NES	Nuclear export signal
NLS	Nuclear localization signal
P body	Processing body
PCR	Polymerase chain reaction
PNA	peptide nucleic acid
PR	Progesterone receptor
PTM	Post-translational modification

RACE	Rapid amplification of cDNA ends
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNAseq	RNA sequencing
SAINT	Statistical analysis of the interactome
SD	Silencing domain
SE	Size exclusion
SINQ	Normalized spectral index quantitation
siRNA	Small interfering RNA
TNRC6/GW182	Trinucleotide repeat containing 6 protein family

Chapter 1
Historical Background and Significance

1.1. Introduction

RNA interference (RNAi) is universally recognized as a powerful mechanism to control protein translation in the cytoplasm of mammalian cells using microRNAs (miRNAs) and small interfering RNAs (siRNAs) to regulate the stability of target mRNA (Ipsaro and Joshua-Tor, 2015). miRNAs act post-transcriptionally, typically by forming imperfect base-pairs within the 3'-untranslated region of target mRNAs and blocking translation. siRNAs form perfect matches throughout mRNA targets and induce cleavage of target transcripts by the RNA Induced Silencing Complex (RISC) (Liu et al., 2004).

Small RNAs (siRNAs and miRNAs) are typically duplex RNAs, consisting of a guide strand complimentary to the RNA target sequence and a passenger strand. The guide strand is supplied as either a synthetic duplex RNA (Elbashir et al., 2001), chemically stabilized single-strand RNA (Lima et al., 2012), or through a naturally expressed within a cell as a miRNA (Ha and Kim, 2014; Hammond, 2015). During RNAi the guide strand RNA is loaded into the RISC, which consists of protein factors including Argonaute 2 (AGO2) (Liu et al., 2004; Meister et al., 2004; Rand et al., 2004) and TNRC6 proteins (Eulalio et al., 2008; Jakymiw et al., 2005), and directs the complex to target sequences through Watson-Crick base-pairing (Ui-Tei et al., 2012). When the central portion of the guide RNA is fully complimentary to its target, AGO2 (not AGO1,3,4) can cleave the target RNA (Liu et al., 2004). These features make RISC programmable and easily manipulated so that many different cellular RNA sequences can be targeted.

Additional factors known to be involved in cytoplasmic RNAi include TNRC6 proteins (GW182) (Eulalio et al., 2008; Nishi et al., 2013; Pfaff and Meister, 2013), Dicer (Bernstein et al., 2001; Ha and Kim, 2014), and TAR RNA binding protein (TRBP) (Daniels and Gatignol, 2012;

Takahashi et al., 2014; Wilson and Doudna, 2013). TNRC6 proteins bind and stabilize AGO-RNA complexes and localize them to RNA targets (Huntzinger and Izaurralde, 2011) and also recruit factors that contribute to the miRNA translational silencing process (Djuranovic et al., 2012). Dicer is an endoribonuclease that is responsible for processing pre-miRNAs to form mature, double-stranded miRNAs. Dicer and TRBP bind double-stranded miRNAs and assist loading of the guide strand into AGO.

Cytoplasmic RNAi has had important applications. For example, siRNAs are important tools for controlling the expression of specific genes in the laboratory and numerous siRNAs are being tested in clinical trials of gene silencing in cancer, infectious disease, genetic disorders, and other conditions (Watts and Corey, 2012; Yang and Zhang, 2012). Many years of experience with natural and synthetic RNAs has demonstrated the power of RNAi to drive recognition of RNA in the cytoplasm of mammalian cells. However, small RNAs are also found in mammalian nuclei, as are RNAi proteins like AGO2 and the TNRC6 proteins (Gagnon et al., 2014a; Matsui et al., 2015; Robb et al., 2005). The presence of both small RNAs and RNAi factors in nuclei suggests that RNA-mediated recognition may regulate RNA-dependent processes like transcription or splicing. Despite its potential impact, nuclear RNAi has remained an unexplored facet of gene regulation in mammalian cells.

1.2. Nuclear RNAi in non-mammalian systems

The potential for RNAi to regulate nuclear process in non-mammalian organisms has been recognized for over 20 years and has been extensively reviewed (Castel and Martienssen, 2013; Creamer and Partridge, 2011; Holoch and Moazed, 2015; Rogers and Chen, 2013).

Nuclear RNAi was first discovered in plants and shown to induce DNA methylation and silencing of sequences at gene promoters by AGO-small RNA complexes binding to nascent transcripts at the promoter and recruiting chromatin remodeling factors (Duan et al., 2015; Haag

and Pikaard, 2011; McCue et al., 2015; Mette et al., 2000; Underwood and Martienssen, 2015; Wassenegger et al., 1994; Wierzbicki et al., 2008; Zhu et al., 2013; Zilberman et al., 2003) **(Figure 1.1A)**. In yeast, AGO-small RNA complexes induce H3K9 methylation and silencing of centromeric repeats by binding to nascent transcripts and induction of histone modifications (Creamer and Partridge, 2011; Halic and Moazed, 2010; Holoch and Moazed, 2015; Hong et al., 2005; Horn et al., 2005; Jia et al., 2005; Reinhart and Bartel, 2002; Verdel et al., 2004; Volpe et al., 2002; Wu et al., 2012) **(Figure 1.1B)**. In C. *elegans*, AGO-small RNA complexes also induce H3K9 methylation and transcriptional silencing by binding to nascent transcripts and recruitment of histone methyltransferases (Burkhart et al., 2011; Burton et al., 2011; Grishok et al., 2001; Gu et al., 2012b; Gu et al., 2009; Guang et al., 2010; Guang et al., 2008; Holoch and Moazed, 2015; Ketting et al., 2001; Knight and Bass, 2001) **(Figure 1.1C)**.

The same general mechanism of transcriptional silencing is conserved between each organism; AGO-small RNA complexes bind to nascent transcripts and recruit other factors to induce silencing. This suggests the possibility that the same general mechanisms are present in mammalian cells.

1.3. Evidence of RNAi in mammalian cell nuclei

In spite of the significance of cytoplasmic RNAi in mammalian cells and the role of nuclear RNAi in non-mammalian organisms, the study of mammalian nuclear RNAi has remained understudied. The reason for this is because there has been skepticism from prominent RNAi researchers in the field since the first evidence of nuclear RNAi was presented in 2004 (Kawasaki and Taira, 2004; Morris et al., 2004). Before those reports were published, there were several reports and anecdotal evidence suggesting that RNAi was only a cytoplasmic process, making it harder for anyone to dispute what was the central dogma at the time. For example, there were two reports showing that siRNAs could not silence expression of pre-mRNA (Vickers et al., 2003;

Zeng and Cullen, 2002), so it was assumed that RNAi was only a cytoplasmic process. In addition, early microscopy of RNAi factors showed exclusive localization of AGO2 to cytoplasmic P bodies and in the endoplasmic reticulum (Ikeda et al., 2006; Stalder et al., 2013). However, based on what is known today, those early experiments ended up biasing results towards the cytoplasmic RNAi conclusion and needed to be performed more carefully. For example, siRNAs may not have been chosen carefully with the best controls or the immunofluorescence methods chosen could not efficiently penetrate the nucleus for antibody binding to nuclear proteins.

As mentioned previously, in 2004, a paper reported that duplex RNAs targeting gene promoters could repress gene expression in mammalian cells through DNA methylation and histone modification (Kawasaki and Taira, 2004; Morris et al., 2004). However, that report was met with criticism from RNAi researchers because the data were not convincing enough to challenge the assumptions about RNAi at the time.

Since then, several reports have been published suggesting that RNAi factors are present and active in mammalian cell nuclei. Two early reports suggested that small RNAs can silence nuclear localized 7SK RNA (Ohrt et al., 2012; Robb et al., 2005). Two reports identified tagged-AGO2 in purified cell nuclei (Robb et al., 2005; Rudel et al., 2008). In addition, fluorescence microscopy suggested that AGO2 is loaded with small RNAs in the cytoplasm and subsequently imported into the nucleus for silencing (Ohrt et al., 2008).

The mechanisms of nuclear import of RISC proteins became a subject of debate. There has been evidence that knockdown of an import protein, Importin 8, reduced AGO2 concentration in the nucleus, suggesting a role in RNAi factor nucleo-cytoplasmic shuttling (Wei et al., 2014; Weinmann et al., 2009). However, another report suggested that Importin 8 is directly involved in AGO nuclear import and that importin proteins have redundant functions in translocating AGO proteins (Schraivogel et al., 2015a). That report also suggested that TNRC6 proteins require an Importinα/β mechanism for nuclear transport. In another report, TNRC6A was discovered to

possess a nuclear localization and export signal (NLS/NES) and implicated to function in shuttling AGO-RNA complexes between the cytoplasm and nucleus (Nishi et al., 2013; Nishi et al., 2015). To date, AGO proteins have not been discovered to possess NLS and NES signals.

To address the skepticism in the field of nuclear RNAi, my laboratory published a report that provided multiple lines of evidence showing that RNAi factors are present and active in human cell nuclei (Gagnon et al., 2014a). They developed a stringent protocol to purify cell nuclei, significantly reducing cytoplasmic and ER contamination, and examined RNAi factor localization and function by microscopy, cellular fractionation and western blots, and activity assays.

To test for nuclear localization of RNAi factors, Immunofluorescence microscopy with high sensitivity 3-D imaging of multiple human cells lines was used to reveal that AGO2 and TNRC6A, as well as other RNAi factors, were present in both the cytoplasm and nucleus of cells. Cellular fractionation was a second method used to test for nuclear localization. AGO2, TNRC6A, and all other RNAi factors were present in both cytoplasmic and nuclear extracts of four different cell lines.

To test for nuclear RNAi activity, several assays were used. Duplex RNAs could knock down the nuclear localized RNA targets MALAT1 and NEAT1. 5'-RACE fragments of predicted size were detected, suggesting that each target RNA had been cleaved by the RNAi complex. Cleavage assays were also used in cytoplasmic and nuclear extracts showing that AGO2 can cleave RNA substrates in both compartments. When RNA substrate and the duplex RNA were added directly to the cell extract, they did not observe cleavage in nuclear extracts. However, when they were transfected into cells prior to activity assays, cleavage did occur. As previously reported, this suggests that the guide RNA is loaded into the RNAi complex in the cytoplasm and shuttled into the nucleus (Ohrt et al., 2008). Consistent with these observations, they reported that RNA loading factors Trax and Translin cannot be detected in cell nuclei.

1.4. The Argonaute Proteins

Argonaute (AGO) proteins are key components of RNA interference (RNAi) (Meister, 2013), which bind directly to miRNAs and siRNAs to form complexes that regulate gene expression. AGO proteins are conserved and are expressed ubiquitously in animals, plants, and yeast. There are four AGO variants in human cells (AGO1-4) (Schurmann et al., 2013). AGO2 is the only AGO protein to retain catalytic activity, and when combined with a siRNA can reconstitute RNAi in defined cell free systems (Liu et al., 2004). Roles for AGO1, AGO3, and AGO4 are not well defined, but they can bind to RNA and AGO1 and AGO3 have been shown to participate in gene silencing in the absence of AGO2 (Ruda et al., 2014).

AGO2 is essential for post-translational silencing of mRNA in the cytoplasm of mammalian cells and many studies have focused on this cytoplasmic role. However, AGO proteins are also active in the nucleus (Gagnon and Corey, 2012; Gagnon et al., 2014a; Schraivogel and Meister, 2014). AGO2 can regulate transcription (Chu et al., 2010; Matsui et al., 2013) and splicing (Ameyar-Zazoua et al., 2012; Liu et al., 2015) in human cells.

1.5. Functional Domains of the Argonaute Proteins

All eukaryotic AGO proteins have four distinct domains: N-terminal, PAZ, Mid, and PIWI (Swarts et al., 2014) **(Figure 1.2)**. The PAZ, Mid, and PIWI domains are the three functional domains of the Argonaute proteins. The PAZ domain binds to the 3' end of guide RNAs and is required for RNA-induced silencing complex (RISC) activation (Gu et al., 2012a; Lingel et al., 2004). The Mid domain anchors the 5' end of the guide RNA (Pfaff and Meister, 2013). The PIWI domain contains the RNase-H-like sequence motif (SLICER) responsible for RNA target cleavage in the Argonaute proteins (Ma et al., 2005; Parker et al., 2004; Song et al., 2004). In humans, only AGO2 is capable of cleavage because the slicer motif diverged with evolution (Tolia and Joshua-Tor, 2007).

The PIWI domain is also responsible for facilitating protein interactions. Both Dicer and Argonaute proteins function in selecting the active strand of siRNAs and miRNAs, and it has been demonstrated that there is a PIWI box motif that binds to one of the RNaseIII domains of Dicer (Tahbaz et al., 2004). The crystal structure of AGO2 revealed tandem tryptophan-binding pockets in the PIWI domain which define a likely interaction surface for tryptophan-rich proteins like the TNRC6 proteins (Schirle and MacRae, 2012).

1.6. The TNRC6 (GW182) protein family

The GW182 family of proteins were originally identified in human cells as an antigen recognized by autoimmune serum from a patient with motor and sensory neuropathy. It was recognized to form distinct bodies in the cytoplasm, which were termed GW bodies or P bodies (Eystathioy et al., 2002). They also discovered that GW182 proteins were phosphorylated with P^{32} incorporation assays.

There are three GW182 paralogs in vertebrates (TNRC6A/GW182, TNRC6B, TNRC6C), one ortholog in insects (GW182), and two orthologs in *C. elegans* (AIN-1 and AIN-2) (Eulalio et al., 2009d). In vertebrates, the TNRC6 paralogs are the most well studied binding partners of AGO2, with most of the focus on the AGO2-TNRC6A protein interaction (Pfaff and Meister, 2013).

The TNRC6 protein family was linked to function in the miRNA pathway through studies in which they were isolated either as interacting partners of the AGO proteins or as proteins required for miRNA-mediated gene silencing (Behm-Ansmant et al., 2006a; Behm-Ansmant et al., 2006b; Ding and Han, 2007; Liu et al., 2005; Meister et al., 2005). Those studies demonstrated these roles in *Drosophila Melanogaster* (Dm) cells, where there is only a single GW182 protein (Behm-Ansmant et al., 2006a; Eulalio et al., 2008; Rehwinkel et al., 2005). Demonstrating the essential role in other organisms has been difficult because of the existence of multiple paralogs with partially redundant functions. Human TNRC6A/GW182, TNRC6B, and TNRC6C are about

40% similar to each other in sequence identity, but all of their major domains are conserved. They are implicated to be redundant based on several observations: 1) These proteins associate with all four AGOs and have a common set of miRNA targets (Jakymiw et al., 2005; Lazzaretti et al., 2009; Lian et al., 2009; Liu et al., 2005; Meister et al., 2005; Takimoto et al., 2009; Zipprich et al., 2009), 2) Depleting each of these proteins separately only partially relieves silencing of miRNA targets (Chu and Rana, 2006; Jakymiw et al., 2005; Liu et al., 2005; Meister et al., 2005; Zipprich et al., 2009), and 3) Silencing is more inhibited when at least two of the proteins are inhibited (Huntzinger et al., 2010).

After mRNA target recognition by the RISC complex, GW182 proteins interact with PABP bound to the mRNA target poly-A tail (Fukaya and Tomari, 2011; Huntzinger et al., 2010; Jinek et al., 2010; Moretti et al., 2012; Zekri et al., 2009), and the PAN3 and CCR4-NOT deadenylase complexes (Braun et al., 2011; Braun et al., 2012, 2013; Chekulaeva et al., 2011; Fabian et al., 2011; Fabian et al., 2009; Mathys et al., 2014; Zekri et al., 2009). The assembly of this complex represses translation, most likely by displacing the binding of eIF4G to PABP and preventing the mRNA from being circularized for recognition by ribosomal subunits (Fabian and Sonenberg, 2012) **(Figure 1.3)**. The repressed mRNA is then deadenylated by PAN3-CCR4-NOT complexes and either stored or degraded (Behm-Ansmant et al., 2006a; Chen et al., 2009; Eulalio et al., 2009b; Eulalio et al., 2007; Eulalio et al., 2009d; Fabian et al., 2009; Piao et al., 2010; Rehwinkel et al., 2005; Wakiyama et al., 2007). Essentially, GW182 proteins recruit other protein complexes to the mRNA target for RNA-mediated silencing.

1.7. The TNRC6 (GW182) Protein Family Functional domains

TNRC6 proteins have two functional domains involved in silencing, the N-terminal AGO binding domain (ABD) and the bipartite silencing domain (SD) consisting of the Mid and C-terminal regions (Eulalio et al., 2009d) **(Figure 1.4A)**. All vertebrates and insects have several

conserved domains and motifs within those two functional domains: 1) GW-repeats (ABD), 2) UBA (Ubiquitin Associated)-like domain (SD), 3) Q rich region (SD), 4) CIM-1 motif (SD), 5) PAM2 motifs (SD), 6) RRM (RNA Recognition Motif) (SD) (Braun et al., 2013) **(Figure 1.4A)**. The UBA and RRM domains are predicted to be structured while the rest of the protein is predicted to be relatively unstructured, which is why no crystal structure has been resolved for this protein family. The unstructured regions of GW182 proteins are required for protein interactions and to mediate silencing activity (Fabian and Sonenberg, 2012). Nuclear localization and export signals (NLS/NES) are located between the GW and Q rich regions of each TNRC6 protein (Nishi et al., 2013).

Early co-Immunoprecipitation assays in D. *melanogaster* cells demonstrated that the N-terminal AGO binding domain mediates TNRC6 protein interactions with the AGO proteins (Behm-Ansmant et al., 2006a) and that it is required for P body localization in cytoplasmic gene silencing (Behm-Ansmant et al., 2006a; Eulalio et al., 2009a; Lazzaretti et al., 2009). In addition, the TNRC6 family of proteins have many WG/GW repeats (termed the AGO hook) in the AGO binding domain **(Figure 1.4B)**, and the tryptophan residues have been shown to interact with the tryp-binding pockets in the PIWI domain of AGO2 (Pfaff et al., 2013; Takimoto et al., 2009; Till et al., 2007). The exact GW repeats responsible for mediating that interaction were unclear for many years (Eulalio et al., 2009a; Lazzaretti et al., 2009; Till et al., 2007) and researchers debate whether there is a clear answer. However, recent evidence suggests that the first repeat of TNRC6B contributes to binding and that a minimum length of 10AA is required to place two tryptophan residues in the binding pocket of AGO2 (Pfaff et al., 2013). Three binding sites have been discovered for a TNRC6A-AGO2 interaction, two of which are conserved in TNRC6B (Takimoto et al., 2009). Since there are many more WG/GW repeats, it is likely that TNRC6 proteins provide a scaffolding platform for binding multiple Argonautes to assemble regulatory

complexes **(Figure 1.4C)**. Studies for the AGO-binding tryptophan residues on TNRC6C have not been performed.

The UBA-like domains conserved in insects and vertebrates have not been characterized in the context of TNRC6 proteins. These domains are distantly related to UBA domains, which are approximately 40 amino acids long and were first identified in proteins involved in ubiquitination (Buchberger, 2002). These domains do not play a specific role in silencing (Chekulaeva et al., 2009; Eulalio et al., 2009a; Lazzaretti et al., 2009; Zipprich et al., 2009).

The Q-rich region is a conserved region on the insect and vertebrate GW182 proteins. The large group of glutamine residues seems to play a role in P body localization in the presence of the AGO binding domain (ABD) (Behm-Ansmant et al., 2006a). However, this region is not strictly required for GW182 silencing activity (Eulalio et al., 2009a). This observation is significant because it shows that GW182 silencing activity does not correlate with the ability to localize to P bodies.

The RRM (RNA recognition motif) domain is highly conserved among GW182 proteins, and its presence hints at a possible interaction with RNA (Eystathioy et al., 2002). Whether GW182 proteins bind RNA is unknown, however the conserved RRM domain lacks typical RNA-binding features. For example, it lacks the conserved aromatic residues that in typical RRMs interact with RNA through stacking interactions, and it has no positively charged surface that is capable of mediating RNA binding (Eulalio et al., 2009c). A D. *melanogaster* GW182 mutant lacking the RRM was impaired in its silencing activity, but only for a subset of miRNA targets (Eystathioy et al., 2002).

The Mid and C-term regions of the bipartite Silencing Domain (SD) are required for the silencing activity of GW182 proteins because they participate in critical protein binding and scaffolding (Eulalio et al., 2009a; Lazzaretti et al., 2009; Yao et al., 2011; Zipprich et al., 2009). The SD is not highly conserved, with the exception of the RRM, PAM2 binding motif, and CIM1

binding motif. The PAM2 (PABP-interacting motif 2) motif mediates the GW182 protein interaction with the Poly-A binding protein (PABP) and confers direct binding to the MLLE domain of PABP (Derry et al., 2006; Fabian et al., 2009; Huntzinger et al., 2010; Zekri et al., 2009).

The silencing domain is known to contain several WG/GW repeats as well. However, instead of binding to AGO proteins, they form the tryptophan-containing CCR4-Not Interacting Motifs, CIM1 and CIM2 (only vertebrates), to mediate binding to members of the CCR4-NOT complex (Braun et al., 2011) **(Figure 1.4C)**. Two CCR4-NOT complex members, CNOT1 and CNOT9/RQCD1, have been shown to bind directly to TNRC6 proteins (Braun et al., 2011; Fabian et al., 2011; Mathys et al., 2014).

1.8. RNA-mediated transcriptional silencing

As mentioned above, early reports suggested that duplex RNAs could silence expression at the transcriptional level by causing DNA methylation and histone modification (Kawasaki and Taira, 2004; Morris et al., 2004). However, those reports were met with heavy criticism and one report was retracted.

After those first publications, several other reports followed with more evidence of RNAi-dependent transcriptional gene silencing (TGS). For example, siRNAs were used to silence HIV-1 expression at the promoter level through an RNAi-dependent DNA methylation mechanism (Suzuki et al., 2005). Promoter-targeted siRNAs were also used to silence transcription of an artificial CCR5 promoter and endogenous RASSF1A promoter through histone modifications (Kim et al., 2006).

Before the first initial reports of transcriptional gene silencing came out, my laboratory had been studying the ability of peptide nucleic acid oligomers (PNAs) to inhibit progesterone receptor (PR) expression at the transcriptional level (Janowski et al., 2005b). While PNAs act in a completely different mechanism, my laboratory was interested in whether they could use siRNAs

to direct transcriptional silencing. Since they had already studied PR as a model system, they used that same system to observe that duplex RNAs targeting the gene promotor could also silence expression at the transcriptional level (Janowski et al., 2005a). Thorough experimentation demonstrated that the duplex RNAs were not targeting PR mRNA, seed sequence complementarity was required for silencing, RNA synthesis and RNA Pol II levels were decreased. This evidence was consistent with transcriptional silencing.

1.9. RNA-mediated transcriptional activation

While testing a series of duplex RNAs targeting the PR promoter, my laboratory observed that some of the RNAs resulted in a small increase in PR RNA and protein levels in T47D cells. It was reasoned that the small activation was mediated by RNAs, however T47D cells have a high basal level expression of PR, making this activation hard to distinguish. So, they decided to test their hypothesis in a cell line with low basal levels of PR expression.

Duplex RNAs were transfected into MCF7 cells, which is a cell line with a low basal level of PR expression. In contrast to their inhibitory data in T47D cells, they identified several duplex RNAs that were capable of activating PR expression at the transcriptional level (Janowski et al., 2007). In line with what they observed for silencing, activation required complimentary seed sequences, increased recruitment of RNA Pol II, increased pre-mRNA land mRNA levels. This evidence was consistent of transcriptional activation. My laboratory subsequently observed that duplex RNAs could activate transcription of two other physiologically relevant genes, LDL receptor (Matsui et al., 2010) and cyclooxygenase 2 (COX-2) (Matsui et al., 2013).

RNA-mediated transcriptional activation has also been reported by other laboratories. For example, small RNAs could activate three different genes, E-cadherin, p21, and VEGF, in three different human cell lines at the RNA and protein level (Junxia et al., 2010; Li et al., 2006b; Mao et al., 2008; Mao et al., 2010; Yang et al., 2008). They determined that this mechanism was

dependent on AGO2 and caused the loss of H3K9 methylation. Later studies went on to find that small RNAs also activate p53, PAR4, WT1, and NKX3 (Huang et al., 2010; Turunen et al., 2009) in non-human mammalian cells.

1.10. Activation in one context, repression in another: lessons from the progesterone receptor

As observed by my laboratory, duplex RNAs can act as either repressors or activators of transcription (Janowski et al., 2005a; Janowski et al., 2005b; Janowski et al., 2007). How can duplex RNAs activate transcription of a gene in one context and repress in another? It most likely can be explained by population of proteins involved in each mechanism.

One possible explanation is that some transcription factors have the potential to be both activators and repressors (Gow et al., 2014; Kamachi and Kondoh, 2013; Lee and Chiang, 2009; Mao and Byers, 2011; Schlereth et al., 2013; Shi et al., 1997; Stonestrom et al., 2015; Weth and Renkawitz, 2011; Yoshida et al., 2014). In the context of transcriptional repression, proteins can compete with activator binding sites, bind directly to activators to interfere with their activity, or bind to co-repressors to interfere with activator activity (Lee and Chiang, 2009; Shi et al., 1997). In the context of activation, the most common mechanism is the recruitment of activating factors that can either guide the transcription factor to promoters or bind to the promoter and recruit activating factors (Kamachi and Kondoh, 2013; Lee and Chiang, 2009; Shi et al., 1997; Stonestrom et al., 2015).

Another possible explanation is that several proteins have been found to activate in one cell type and repress in another (Mao and Byers, 2011; Shi et al., 1997). Given this information, we can predict that in the case of the PR promoter, the cellular context and basal level of expression influence the transcription factors present at the site of duplex RNA modulation.

1.11. Nascent transcripts are the target for RNA-mediated transcriptional regulation

The nuclear target for duplex RNAs has been a subject of debate among researchers in the field. How can they affect gene expression if they do not recognize mRNA? There are two clear possibilities, 1) the duplex RNAs bind directly to chromosomal DNA or 2) the mechanism involves RNA.

Many studies of transcription have revealed that a large portion of the genome is transcribed and that the transcripts overlap the 3' and 5' termini of genes. As was the case for my laboratory, they discovered that for PR (Schwartz et al., 2008), COX-2 (Matsui et al., 2013), and LDLR (Matsui et al., 2010), the target of our duplex RNAs were nascent transcripts that overlap the gene promoters. Another study using a reporter gene determined that transcriptional gene activation is dependent on local, nascent, anti-sense lcRNA transcripts (Zhang et al., 2014). Studies in yeast also discounted the possibility of siRNA-DNA base pairing in RNA-mediated heterochromatin formation. (Portnoy et al., 2016; Shimada et al., 2016).

Several experiments my laboratory performed supported the involvement of these RNAs in transcriptional regulation, 1) RNA immunoprecipitation of AGO2 revealed that the RNAi complex is recruited to promoter transcripts, 2) Antisense oligonucleotides designed to induce RNase H mediated cleavage of the transcript reversed RNA-mediated gene activation, 3) when the binding site for an RNA was changed by just one or two bases, activation or silencing was abolished, 4) duplex RNAs containing central mismatches to abolish AGO2 cleavage were just as effective for activation as fully complimentary RNAs. All of this evidence suggest that RNAi is involved and the nascent transcripts are the trigger for activation and that cleavage of the transcript is not necessary for this process.

1.12. RNAi factors are involved in transcriptional silencing and activation

Since duplex RNAs achieve transcriptional silencing and activation, and they are identical in structure to those used for cytoplasmic gene silencing (Elbashir et al., 2001), it makes sense that RNAi factors would be involved in these processes. However, few reports have been published to support this conclusion.

AGO proteins are the central factors for RNAi. There are four AGO proteins in mammalian cells (AGO1-4), and AGO2 is only variant capable of catalytic cleavage of RNA substrates (Liu et al., 2004; Meister, 2013; Meister et al., 2004; Rand et al., 2004). Because of their central role in silencing, my laboratory tested the involvement of AGO1-4 in transcriptional silencing and activation (Chu et al., 2010; Grishok et al., 2001; Janowski et al., 2006; Janowski et al., 2007; Matsui et al., 2010). Experiments only implicated AGO1 and AGO2 as having a role in silencing and AGO2 as have a role in activation **(Figure 1.5)**. They also observed that silencing expression of all of the TNRC6 paralogs (TNRC6A/B/C), not individual TNRC6 proteins, could reverse transcriptional activation of COX-2 (Matsui et al., 2013). One study implicated that AGO2 was required for transcriptional gene activation accompanied with histone modification (Zhang et al., 2014). Another study also implicated AGO1 as a critical factor for transcriptional silencing (Kim et al., 2006). Overall, it is clear that RNAi factors are crucial for RNA-mediated transcriptional regulation and they may be context dependent with regards to the promotor targeted.

1.13. Lessons from COX-2 activation: filling in more mechanistic gaps

How do the duplex RNAs and the RNAi complex regulate transcription? What other proteins are involved and how are they recruited to the promoter? In 2013, my lab published an extensive manuscript characterizing RNA-mediated activation at the COX-2 promoter in great detail. This work utilizing COX-2 as a model system answered more mechanistic questions about how duplex RNAs regulate transcription. They observed that duplex RNAs complementary to the COX-2 promoter activated the expression of COX-2 pre-mRNA, mRNA, and protein. Activation was dependent on seed sequence complementarity, expression of AGO2, and

expression of TNRC6A/B/C. Knockdown of just one TNRC6 paralog did not have any effect, suggesting their functions are redundant (Matsui 2013, *unpublished data*). They detected a transcript that overlapped the COX-2 promoter. Knock-down of this transcript ablated gene activation and addition of promoter-targeted RNA led to recruitment of AGO2 and TNRC6A to the target transcript. Addition of small RNA led to increased H3K4me3 and H4Ac histone modifications, as well as recruitment of WDR5.

Taken together, this prior work suggested that introduction of a promoter-targeted small RNA into cells could lead to binding of an AGO2-TNRC6-small RNA complex to a promoter transcript, recruitment of WDR5 and activating histone marks, enhanced binding of RNA polymerase, and dramatically upregulated levels of COX-2 expression **(Figure 1.6)**. Unanswered questions included the identity of other protein partners and the role of TNRC6 paralogs in modifying histones and transcriptional activation.

1.14. RNA-mediated regulation of alternative splicing

Alternative splicing is another nuclear process that leads to the production of many protein isoforms and is observed in >90% of all human genes (Keren et al., 2010). These protein isoforms have various functions and regulation of their production is necessary for cell function, growth, and development (Kalantari et al., 2016a). Splicing of pre-mRNA is controlled by protein splicing factors. Does RNAi play a role in this nuclear process as well?

Approximately 60% of disease-causing point mutations are related to defective splicing (Jensen et al., 2009). Antisense oligonucleotides (ASOs), single-stranded RNAs that act separately of the RNAi pathway, were the first nucleic acid molecules used to target alternative splicing of disease genes (Watts and Corey, 2012). For example, they were used to target the dystrophin pre-mRNA, which contain a premature stop codon leading to protein truncation causing Duchenne muscular dystrophy (DMD) (Aartsma-Rus et al., 2003; Arechavala-Gomeza et

al., 2007). ASOs targeted splice sites near intron/exon junctions causing the mutation to be skipped, correcting the reading frame to produce a partially functional protein. Similar splice alteration of the SMN2 protein, the protein responsible for spinal muscular atrophy, is another clinical approach to targeting disease genes (van Deutekom et al., 2007; Voit et al., 2014). ASOs act by binding pre-mRNA and blocking the binding of splicing factors, thus inhibiting their action on the site (Jarver et al., 2014).

My laboratory wanted to know whether we could use duplex RNAs and therefore the RNAi machinery to bind to splice sites and produce similar effects on splicing. The reason for this is that duplex RNAs are much more stable than single-stranded RNAs and can be used to manipulate endogenous RNAi protein complexes within cells. They first analyzed the effect of duplex RNA on splicing with an engineered model of a luciferase gene interrupted by a β-globin intron (Liu et al., 2012). By targeting the β-globin splice site, they were able to alter splicing of the retained intron and increase expression of active luciferase. They also observed that duplex RNAs could alter splicing of dystrophin and SMN2 similar to the action of ASOs. In a subsequent study, they were able to use single-stranded siRNAs (ss-siRNAs) mimicking the structure of ASOs to alter splicing as well (Liu et al., 2015). Both of these studies showed that alteration of splicing is dependent on expression of AGO2 and that it is recruited to splice junctions **(Figure 1.7A)**. The mechanism of action of these siRNAs is similar to that of ASOs, wherein the RISC complex competes for binding of the splice sites and displaces the splicing machinery.

Splicing can also be coupled to transcription in mammalian cells (Naftelberg et al., 2015) through the kinetics of RNA Polymerase and recruitment of splicing factors. Duplex RNAs targeting intronic or exonic sequences near alternative exons could affect splicing and inclusion of the exon through RNA Pol II pausing and was shown to require AGO1 and increase H3K9 and H3K27 methylation (Allo et al., 2009) **(Figure 1.7B)**. AGO1 and AGO2 have also been reported

to couple chromatin silencing with alternative splicing, although it is not clear based on that study (Ameyar-Zazoua et al., 2012). One report suggests that antisense transcripts and not just pre-mRNA, may be the trigger for splice-altering siRNAs (Batsche and Ameyar-Zazoua, 2015).

Given that RNAi and miRNAs are present in human cell nuclei, it seems possible that RNAi may play a role in regulating the splicing of endogenous genes. While no endogenous targets have been reported to date, my laboratory has shown through RNA sequencing that AGO2 bound species in human cell nuclei have many potential binding sites near exon/intron junctions, suggesting that AGO has a role in alternative splicing (Chu et al., 2015).

1.15. Endogenous roles for transcriptional regulation by RNAi

Do RNAi complexes function to endogenously regulate transcription through miRNA or dsRNA mediated mechanisms? Many studies focus on the exogenous or synthetic regulation of transcription through the use of duplex RNAs. However, miRNAs have also been discovered in the nuclei of mammalian cells. In fact, out of 456 miRNA species in whole cell extracts, 346 of them were identified in cell nuclei (Gagnon et al., 2014a). In addition, my laboratory demonstrated that miRNAs may play a role in splicing and other noncanonical functions in nuclei (Chu et al., 2015). RNA-circularization sequencing (RC-seq) identified that in contrast to cytoplasmic AGO mostly binding to the 3'-UTR of RNAs, nuclear AGO2 mostly binds to intron-exon junctions, suggesting a role in splicing.

My laboratory also identified miRNAs that regulate the PR and COX-2 promoters (Matsui et al., 2013; Younger and Corey, 2011). miRNA mimics were used to downregulate the expression of PR and the mechanism was found to involve AGO2 and H3K9 methylation. A miRNA, miR-589, was discovered to regulate the COX-2 promoter and miR-589 mimics were shown to increase the expression of COX-2 and downregulation showed a decreases in COX-2 expression, suggesting that miR-589 is an endogenous activating RNA. There have also been a

few reports showing RNAi-dependent miRNA-induced transcriptional regulation in mammals at E-cadherin, CSDC2, POLR3D, and CCNB1 promoters (Huang et al., 2012; Kim et al., 2008; Place et al., 2008), showing that miRNA promoter regulation is versatile in cells.

Recently, a new endogenous mechanism of RNAi-mediated transcriptional silencing has been proposed (Skourti-Stathaki et al., 2014). The Proudfoot lab suggested a mechanism where naturally forming R loops (RNA:DNA hybrids) cause transcriptional termination, synthesis of localized dsRNA and recruitment of RNAi factors. They showed that DICER, AGO1, and AGO2 are recruited to R loop sites and that induces repressive chromatin marks and heterochromatin formation **(Figure 1.8)**. This raises the possibility that R-loops, which naturally occur in cells, may more widely induce transcriptional silencing through the RNAi process.

Taken together, the evidence that RNAi factors are present in cell nuclei, miRNAs have been shown to effect transcription, and that R-loops cause RNAi-mediated transcriptional silencing leads us to believe that RNAi is playing an endogenous role in cells to regulate transcription and that more studies need to focus on the aspect of global RNAi regulation. This would open up another layer of RNAi-mediated gene regulation and lead to potential clinical outcomes.

1.16. Conclusions

Both miRNAs and RNAi factors exist in mammalian cell cytoplasm and nuclei. While most research has been directed to study cytoplasmic RNAi, there has been surprisingly little research dedicated to understanding what RNAi is capable of inside the nucleus. Evidence in non-mammalian systems shows a clear pattern of conserved nuclear RNAi mechanisms. RNAi factors are present in human cell nuclei and these factors are capable of silencing nuclear targets. We know that RNAi factors are also capable of regulating nuclear processes in human cells including transcription and splicing. Given the impact that cytoplasmic RNAi has on clinical research, it is

imperative to broaden our understanding of this novel gene regulatory mechanism of nuclear RNAi as it extends to transcription and splicing, rather than just translation.

1.17. Hypothesis

Both miRNAs and RNAi factors exist in mammalian cell nuclei. While it is well known that these factors can function in nuclei when presented with exogenous stimuli, it is still unclear what their endogenous roles are. AGO2 and TNRC6 proteins have been identified as key players in nuclear RNAi functions. However, to date, there have been no studies of their binding partners in human cell nuclei and how they differ relative to cytoplasmic binding partners. There also has not been a comprehensive study on how RNAi factors are regulated in their localization and functions between the cytoplasm and nucleus. The focus of this dissertation is to identify the protein binding partners of nuclear AGO2 and TNRC6A, and to identify post-translational modifications (PTMs) that may play a role in their regulation between the cytoplasm and nucleus. This will lead to a better understanding of the endogenous roles of these RNAi factors and shed light on a novel layer of gene regulation.

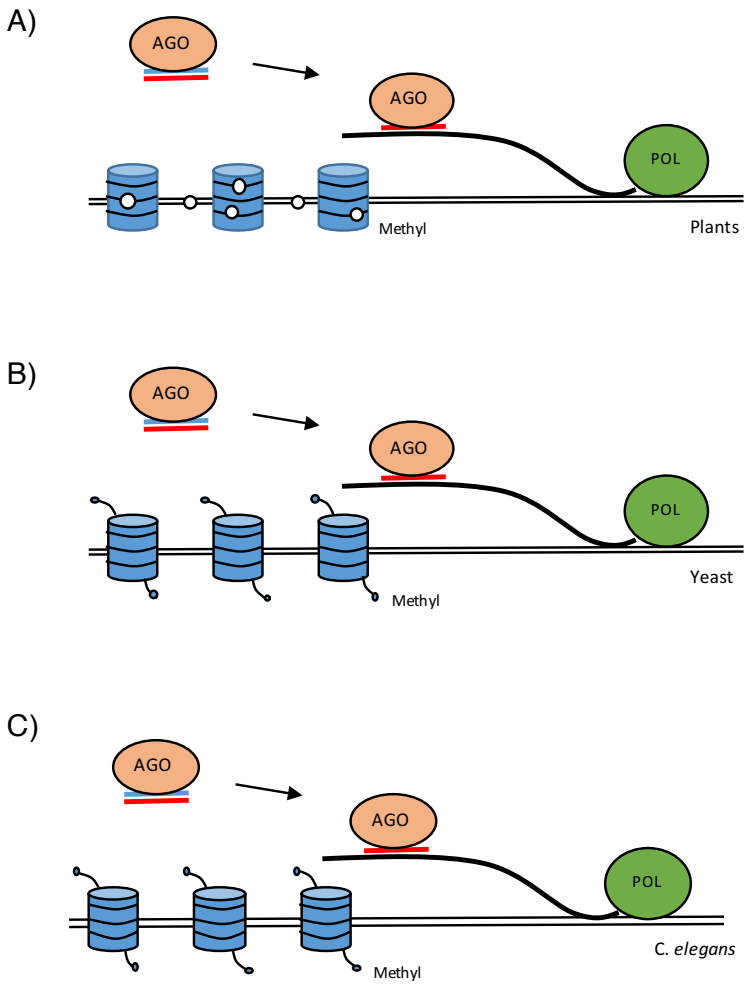


Figure 1.1: RNA Interference in non-mammalian cell nuclei

a) DNA methylation in plants. Polymerase transcribes nascent RNAs expressed at the promoter. Double-stranded RNAs targeting the nascent RNA associate with AGO proteins and the RISC complex and guide the complex to the target. Methylation factors localize to the site causing methylation and silencing

b) Histone methylation in Yeast. Polymerase transcribes nascent RNAs expressed at the promoter. Double-stranded RNAs targeting the nascent RNA associate with AGO proteins and the RISC complex and guide the complex to the target. Histone methylation complex localizes to the site causing methylation and silencing.

c) Histone methylation in worms. Polymerase transcribes nascent RNAs expressed at the promoter. Double-stranded RNAs targeting the nascent RNA associate with AGO proteins and the RISC complex and guide the complex to the target. Histone methylation complex localizes to the site causing methylation and silencing.

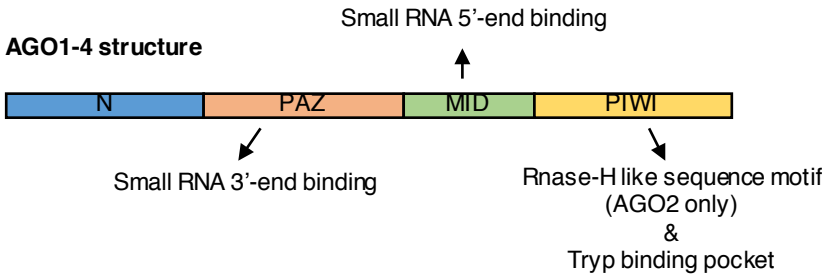


Figure 1.2. General structure of the four human argonaute proteins. Each has N-terminal, PAZ, MID and PIWI domains. The PAZ domain is responsible for binding RNAs at the 3'-end. The MID domain is responsible for binding RNAs at the 5'-end. The PIWI domain is responsible for cleavage activity (AGO2 only) and binding to WG/GW repeats.

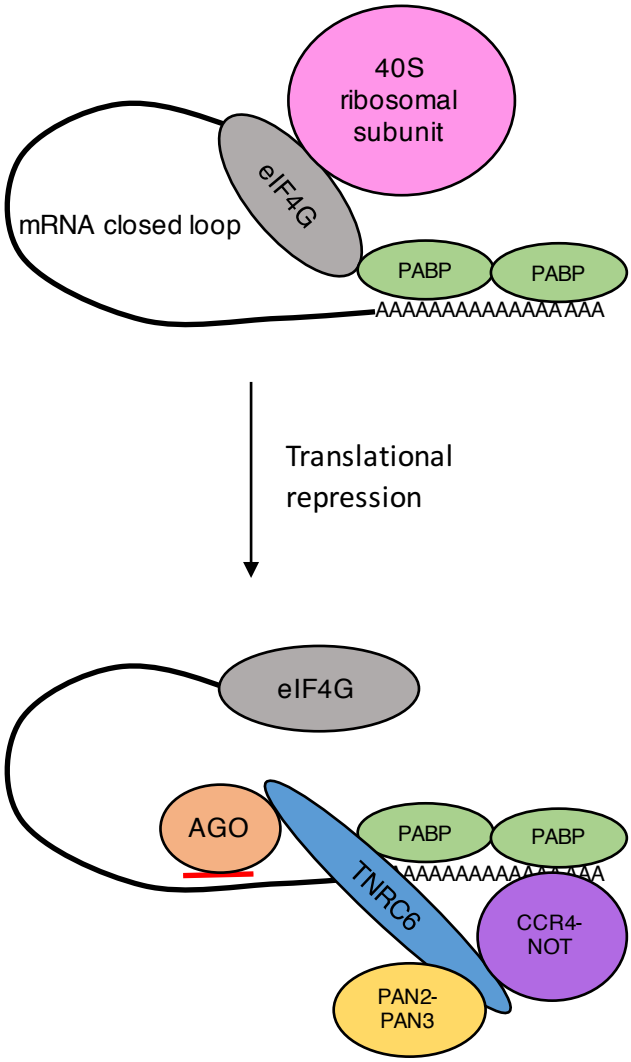
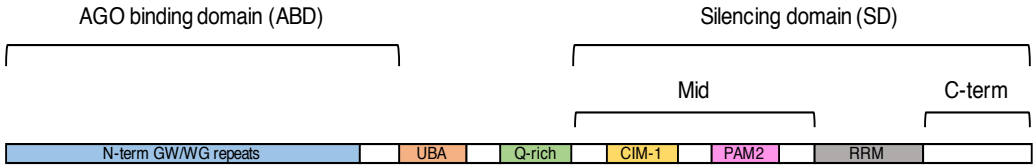


Figure 1.3. The role of TNRC proteins in miRNA gene silencing in the cytoplasm.

a) For normal translation to occur, eIF4G binds to PABP forming a closed loop mRNA structure. This allows for the recognition of the mRNA by the ribosome for translation

b) When miRNA gene silencing occurs, AGO-guide RNA-TNRC6 complexes recognize the mRNA target and TNRC6 proteins bind to PABP, blocking the binding site of eIF4G preventing the closed loop structure to form. This causes translational repression because the ribosome can no longer recognize the structure for translation. TNRC6 proteins also recruit the PAN2/3 and CCR4-NOT complexes to the poly-A tail for de-adenylation and degradation.

A)



B)

```
   1 mreleakatk dvernlsrdl vqeeeqlmee kkkkkddkkk keaaqkkate qkikvpeqik
  61 psvsqpqpan snngtstats tnnnakrata nnqqpqqqqq qqqpqqqqpq qqpqpqpqqq
 121 qpqqqpqalp ryprevpprf rhqehkqllk rgqhfpviaa nlgsavkvln sqsessaltn
 181 qqpqnngevq nsknqsdinh stsgshyens qrgpvsstsd sstncknavv sdlsekeawp
 241 sapgsdpela secmdadsas sseserniti masgntggek dglrnstglg sqnkfvvgss
 301 snnvghgsst gpwgfshgai istcqvsvda pesksessnn rmnawgtvss ssngglnpst
 361 lnsasnhgaw pvlennglal kgpvqsgssg iniqcstigq mpnnqsinsk vsggsthgtw
 421 gslqetcese vsgtqkvsfs gqpqnittem tgpnnttnfm tsslpnsgsv qnnelpssnt
 481 gawrvstmnh pqmqapsgmn gtslshlsng esksggsygt twgaygsnys gdkcsgpngq
 541 angdtvnatl mqpgvngpmg tnfqvntnkg ggvwesgaan sqstswgsgn gansggsrrg
 601 wgtpaqntgt nlpsvewnkl psnqhsndsa ngngktftng wksteeedqg satsqtneqs
 661 svwaktggtv esdgstestg rleekgtges qsrdrrkidq htllqsivnr tdldprvlsn
 721 sgwgqtpikq ntawdtetsp rgerktdngt eawgssatqt fnsgacidkt spngndtssv
 781 sgwgdpkpal rwgdskgsnc qqgweddsaa tgmvksnqwg nckeekaawn dsqknkqgwg
 841 dgqkssqgws vsasdnwget srnnhwgean kksssggsds drsvsgwnel gktssftwgn
 901 ninpnnssgw desskptpsq gwgdppksnq slgwgdsskp vsspdwnkqq divgswgipp
 961 atgkppgtgw lggpipapak eeeptgweep spesirrkme iddgtsawgd pskynyknvn
1021 mwnknvpngn srsdqqaqvh qlltpasais nkeassgsgw gepwgepstp attvdngtsa
1081 wgkpidsgps wgepiaaass tstwgsssvg pqalsksgpk smqdgwcgdd mplpgnrptg
1141 weeeedveig mwnsnssqel nsslnwppyt kkmsskglsg kkrrrergmm kggnkqeeaw
```

C)

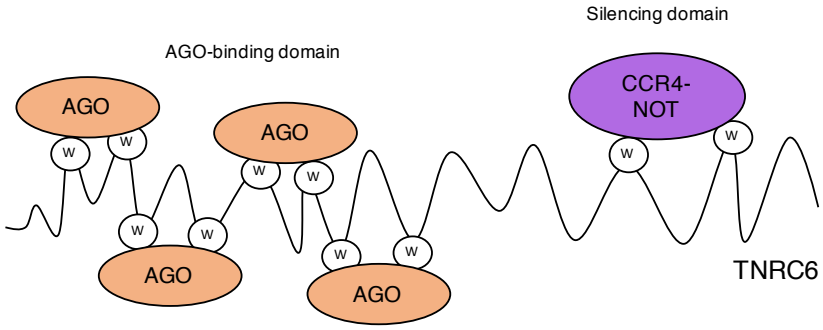


Figure 1.4. The general structure of TNRC6 proteins.

a) The domain organization that is conserved between all three paralogs of TNRC6. Each protein has an AGO-binding domain (ABD) at the N-terminal, UBA and Q-rich domains and a silencing domain (SD) consisting of the Mid and C-term regions with the CIM-1, PAM2 and RRM motifs.

b) The amino acid sequence of the AGO-binding domain (ABD) with the WG/GW repeats highlighted

c) A representation of the unstructured domains of TNRC6 proteins and the WG/GW repeats potentially binding to multiple AGO proteins (AGO-binding domain) and the CCR4-NOT complex (Silencing domain)

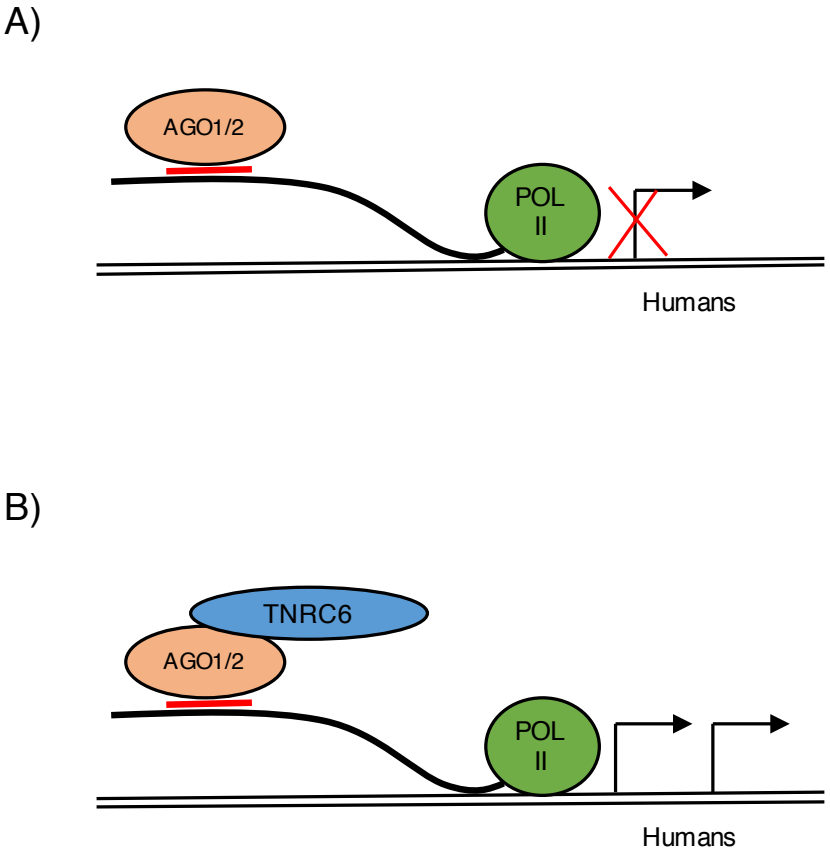


Figure 1.5. RNA-mediated transcriptional regulation in human cells.

a) RNA-mediated transcriptional silencing. PolII transcribes nascent RNAs expressed along the promoter. An AGO1/2-guide RNA complex binds to the target nascent transcript causing silencing of transcription.

b) RNA-mediated transcriptional activation. PolII transcribes nascent RNAs expressed along the promoter. An AGO1/2-guide RNA-TNRC6 complex binds to the target nascent transcript causing activation of transcription

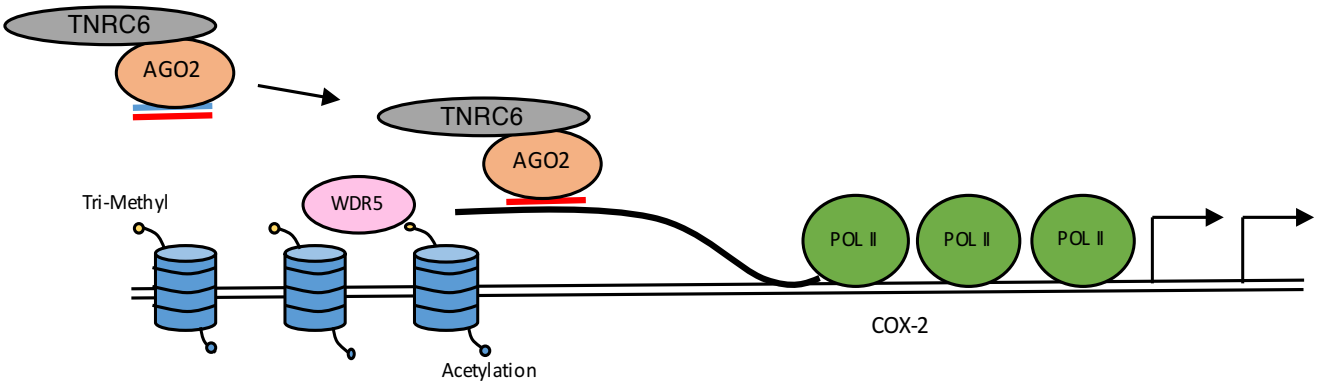


Figure 1.6. RNA-mediated activation of transcription at the COX-2 promoter. PolII transcribes nascent RNA expressed at the promoter. An AGO2-RNA12-TNRC6 complex binds to target nascent transcript. Histone modifying proteins, including WDR5, are recruited to the promoter causing an increase in H3K4me3 and H4ac and activation of transcription of the COX-2 gene.

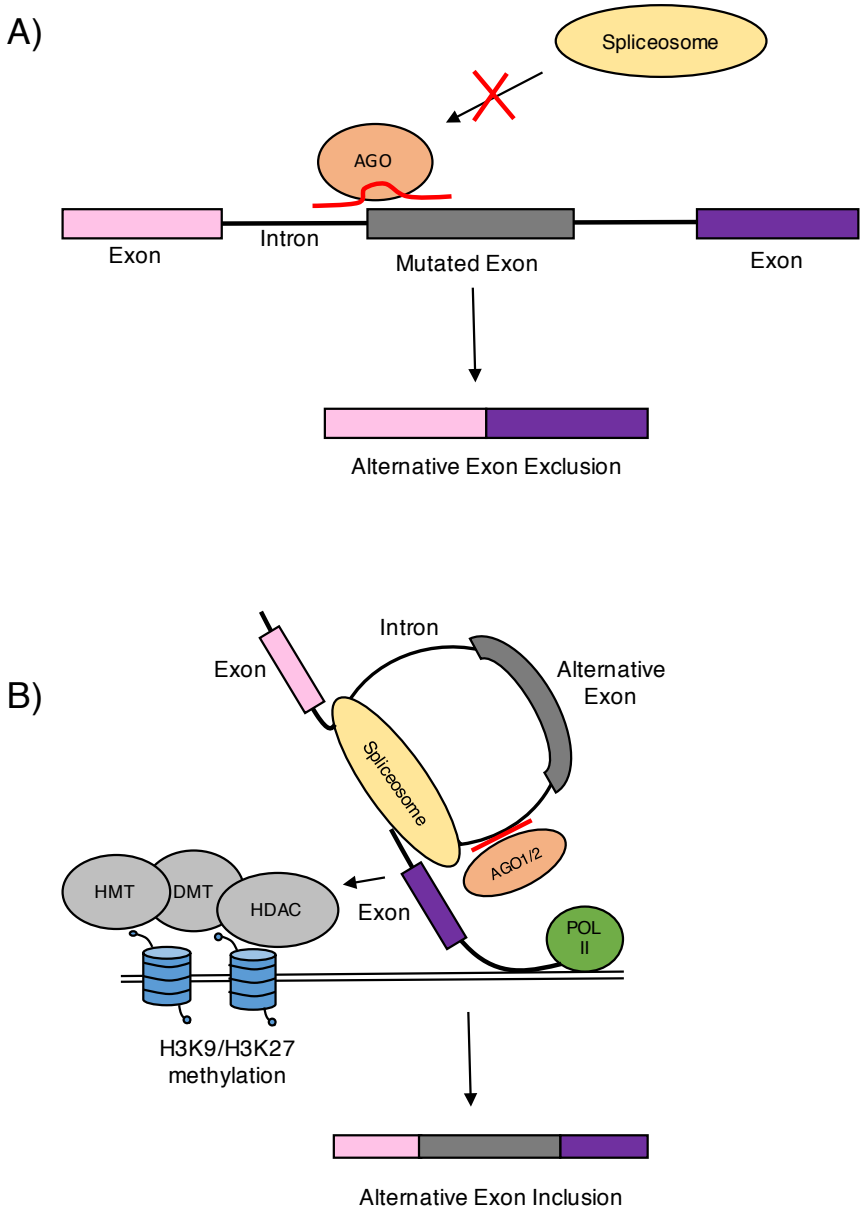


Figure 1.7. Two models of RNA-mediated alternative splicing.

a) RNA-mediated exon exclusion. An RNA with a central mismatch guides AGO proteins to its intron-exon target and prevents binding of the spliceosome. This causes the final product to exclude the exon that was targeted.

b) RNA-mediated exon inclusion. As PolII is transcribing RNA, an AGO-RNA complex targeting the intron binds to its target and causing pausing of transcription through H3K9/H3K27 methylation. The pausing of transcription provides enough time for the splicing machinery to include the alternative exon in the final product.

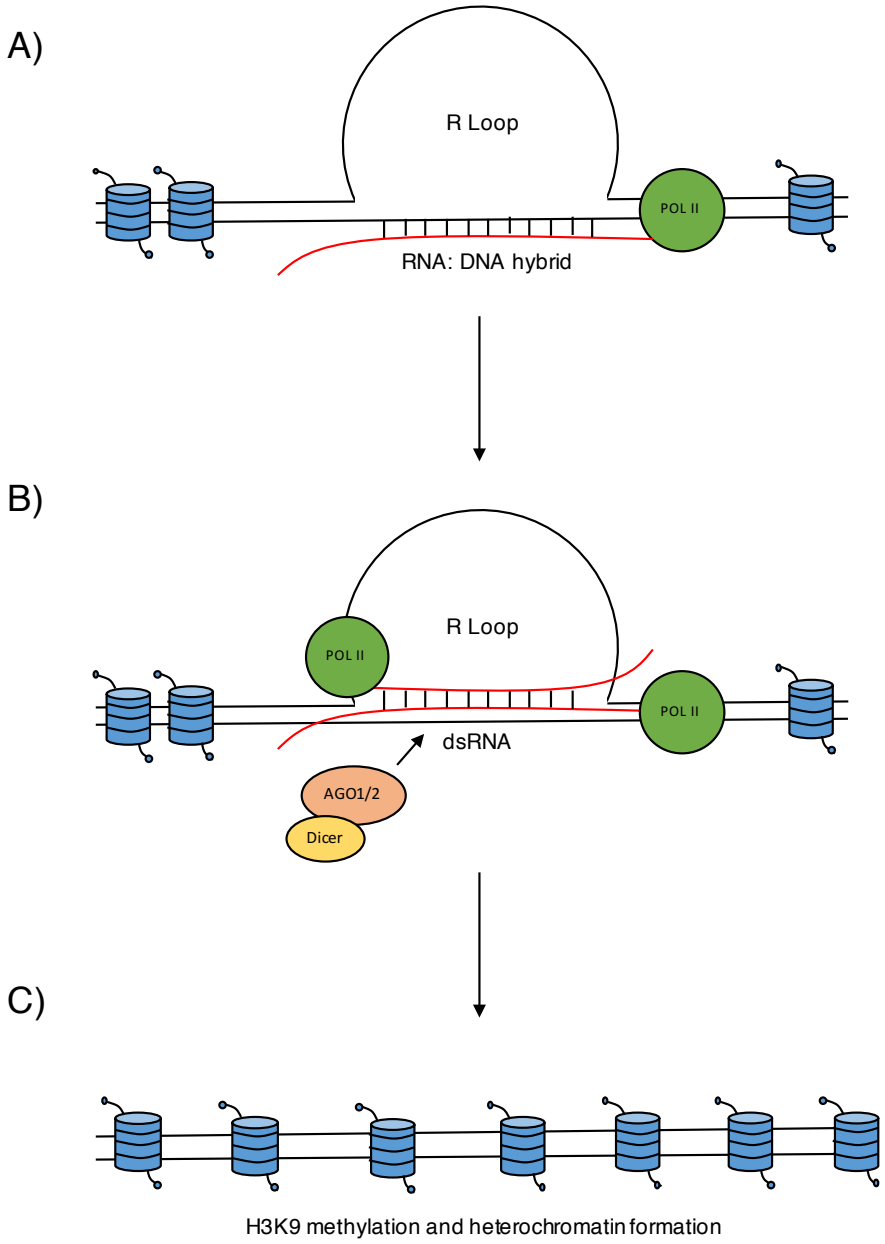


Figure 1.8. Model of R-loops cause transcriptional termination due to recruitment of RNAi factors.

a) A naturally occurring R loop forms an RNA:DNA hybrid as PolII is transcribing the sense strand of DNA.

b) As PolII transcribes the antisense strand of DNA, this causes the RNA:DNA hybrid to dissociate and form a double stranded RNA. The double stranded RNA recruits DICER and AGO proteins to the promoter.

c) The binding of RNAi factors causes the recruitment of histone modifiers for H3K9 methylation and heterochromatin formation, which results in transcriptional termination.

Chapter 2
Post-translational Modifications of AGO2 and TNRC6A

2.1. Introduction

Post-translational modifications (PTMs) are covalent processing events that change the properties of a protein and can determine its activity state, localization, turnover, and interactions with other proteins (Mann and Jensen, 2003). The most studied PTMs include phosphorylation, ubiquitination, methylation, acetylation, and glycosylation. Despite their great importance, studies on a large scale have been hindered by the lack of suitable methods, and many key modifications on proteins have only been discovered in the context of a specific biological process.

A typical unbiased approach to identify PTMs involves isolating a protein or proteins of interest and performing mass spectrometry on the peptides generated after digestion (Chicooree et al., 2015). The PTMs are measured by mass shifts caused by the modification and each modification has a unique value. Depending on the size and amino acid composition of the peptide harboring the PTM, this method may be able to determine the specific site of the modification. More recently, the development of high-resolution mass spectrometry (MS) methods has made it possible to quantify and broadly detect even low abundance PTMs and at least get a partial site-specific localization for each (Cox and Mann, 2011; Zhang et al., 2013).

Identifying PTMs via mass spectrometry relies heavily on how samples were prepared. The most successful studies will purify one protein of interest, which increases the sequence coverage of that protein and the number of modifications identified. However, many studies isolate all proteins that harbor a PTM of interest and that has biased the identification of modifications that are the most stable and abundant on proteins (Mann and Jensen, 2003). For example, since PTMs are usually lowly abundant and unstable in cells, most studies have used modification-specific antibodies (i.e., phospho-tyr, phospho-ser, ubiquitin, His-ubiquitin, acetyl-lys) to isolate all proteins harboring those PTMs, which provides more information on the global

modification levels and less peptide coverage of proteins of interest (Cox and Mann, 2011; Zhang et al., 2013). In addition, there are many examples of studies that bias detection of a particular class of modification through various methods, and a majority of these are phospho-proteomic studies (Han et al., 2008; Iliuk et al., 2010; Kweon and Hakansson, 2006; Li et al., 2008; Simon et al., 2008; Torres et al., 2005). These experiments are used to globally identify proteins harboring a PTM of interest and that does not provide optimal protein coverage or useful information for studying one protein. In addition, most studies focus on identifying one particular class of PTM on a protein of interest, usually phosphorylation, but many other modifications might be missed with this method. Therefore, the best way to approach a PTM study is to purify a protein of interest and use an unbiased PTM search.

2.2. Post-translational modifications of RNAi factors

Since AGO2 is a critical protein in the RNAi pathway, several studies have focused on identifying modification sites to better elucidate the function of the protein. To date, studies have identified 24 AGO2 PTM sites from cytoplasmic or whole cell extracts, which include phosphorylations, acetylations, sumoylations, and hydroxylations **(Figure 2.1A)**. Eleven of those sites were identified through large-scale proteomic studies not focused on AGO2 specifically and have never been verified directly.

Furthermore, regardless of TNRC6 critical function in RNAi, even less is known about the PTMs harbored by TNRC6 proteins. There have been whole cell phospho-proteome studies implicating various phosphorylation sites for each paralog **(Figure 2.1B)**, but those studies were not specifically analyzing TNRC6 proteins, making the identity less meaningful. There is one paper suggesting TNRC6A is ubiquitinated for protein degradation by TRIM65 (Li et al., 2014), but none of these studies verify the modification sites.

Only seven modification sites have been characterized for their functional alteration of AGO2 **(Table 2.1, boxes in Figure 2.1A)**: 1&2) Phosphorylation at S387 and S798 is necessary for AGO2 localization to P bodies, where RNAi components are stored (Lopez-Orozco et al., 2015; Zeng et al., 2008). 3) Phosphorylation at Y529 regulates AGO2 small RNA binding at the 5'-end (Rudel et al., 2011). 4&5) K402 sumoylation and P700 prolyl-4 hydroxylation regulate AGO2 stability and effective RNAi. (Qi et al., 2008; Sahin et al., 2014). 6) Phosphorylation at Y393 is responsible for inhibiting endogenous gene silencing during hypoxia and inducing senescence in cancer cells (Shen et al., 2013; Yang et al., 2014). 7) A phosphorylation cycle at S824-834 has been proposed to regulate AGO2 miRNA:target interactions and to maintain the global efficiency of miRNA silencing (Golden et al., 2017).

Most studies involving AGO2 PTM analysis use a tagged protein (Qi et al., 2008; Rudel et al., 2011; Yang et al., 2014). While tagged proteins facilitate easy purification, they can have altered folding and biological activity within cells. Additionally, all studies have only used whole cell or cytoplasmic extracts to purify AGO2 (Golden et al., 2017; Horman et al., 2013; Kim et al., 2011; Lopez-Orozco et al., 2015; Rudel et al., 2011; Sahin et al., 2014; Yang et al., 2014; Zeng et al., 2008). Furthermore, every study uses a biased approach looking for a particular class of PTM (Golden et al., 2017; Horman et al., 2013; Kim et al., 2011; Lopez-Orozco et al., 2015; Qi et al., 2008; Rudel et al., 2011; Sahin et al., 2014; Yang et al., 2014; Zeng et al., 2008), which can limit identification of important PTMs.

Since AGO2 and TNRC6 perform diverse functions within the cytoplasm and nucleus (Ameyar-Zazoua et al., 2012; Castel and Martienssen, 2013; Chu et al., 2010; Gagnon and Corey, 2012; Janowski et al., 2005a; Janowski et al., 2006; Janowski et al., 2005b; Janowski et al., 2007; Liu et al., 2012; Matsui et al., 2013; Matsui et al., 2010; Robb et al., 2005), it is important characterize those proteins in both compartments separately. It is important to overcome obstacles limiting our knowledge of PTMs that regulate AGO2 and TNRC6 trafficking and activity

in the nucleus. One way is to use an unbiased PTM search to identify and characterize PTMs of cytoplasmic and nuclear AGO2 and TNRC6A.

To identify post-translational modifications for AGO2, I performed antibody affinity chromatography and an unbiased PTM search through mass spectrometry for purified AGO2. I analyzed samples obtained by purifications using anti-FLAG antibody in a stable FLAG-AGO2 T47D cell line and an antibody that recognizes endogenous AGO2 in HeLa cells. Using two purification methods allowed me to crosscheck my results and rule out PTMs that may be artificial due to the overexpression of AGO2, because further characterization of PTMs requires use of a tagged AGO2 plasmid. In addition to that, I used two different cell lines (T47D and HeLa) for my purifications to crosscheck whether PTMs are cell-type specific. I also purified AGO2 from the cytoplasm and nucleus to compare across compartments. Analysis that relied on purification of FLAG-AGO2 identified more PTMs than endogenous AGO2.

Three AGO2 PTMs were identified in both compartments that had been characterized previously as only cytoplasmic (S387, S824-834 and Y529 phosphorylation). I also identified several novel possible PTMs for AGO2 including acetylations, methylations and ubiquitinations that are different in both the cytoplasm and nucleus. In addition, I identified four residues that could potentially harbor multiple types of modifications in different contexts.

In addition to identifying PTMs for AGO2, I identified PTMs for TNRC6A. For these studies, I performed large-scale immunoprecipitations for endogenous TNRC6A complexes from cytoplasmic and nuclear extracts and used an unbiased PTM search to identify a few novel PTMs. These include a phosphorylation site for nuclear TNRC6A, methylation sites for both nuclear and cytoplasmic TNRC6A, and acetylation sites for cytoplasmic TNRC6A. While this is a good start, affinity chromatography needs to be used to get better coverage of the TNRC6A sequence for PTM analysis.

2.3. Purification of AGO2 in cytoplasmic and nuclear extracts for PTM analysis

In this study, I enriched for FLAG-tagged AGO2 in a T47D stably expressing cell line and endogenous AGO2 in HeLa cells in parallel. The use of FLAG-tagged AGO2 is advantageous because the protein is highly expressed and interaction between the FLAG epitope and the antibody is efficient and maximizes the amount of AGO2 purified. One disadvantage is that introducing a tag or overexpressing a protein may alter AGO2 and produce artefactual PTMs. The advantage of using an anti-AGO2 antibody to purify endogenous AGO2 is that it is detected at normal levels in cells. Since further characterization of PTMs needs to be done using a FLAG-AGO2 plasmid, the use of both of these methods will streamline the identification of candidate PTMs for experimental validation.

Two different cell lines, FLAG-AGO2 T47D or HeLa, were used in these studies to rule out the possibility that the PTMs identified were cell-type specific. For the stably expressing FLAG-AGO2 T47D cell line, western analysis suggests that AGO2 is overexpressed about 4.7 fold in the cytoplasm and 1.9 fold in the nucleus relative to endogenous AGO2 in T47D cells **(Figure 2.2A)**. The cellular distribution of endogenously expressed AGO2 in T47D and HeLa cells is roughly 50% in the cytoplasm and 50% in the nucleus **(Figure 2.2B)**. However, the FLAG-AGO2 stable cell line has a distribution of roughly 67% in the cytoplasm and 33% in the nucleus **(Figure 2.2B)**.

Previous AGO2 PTM studies used extracts from whole cells or the cytoplasm (Golden et al., 2017; Horman et al., 2013; Kim et al., 2011; Lopez-Orozco et al., 2015; Qi et al., 2008; Rudel et al., 2011; Sahin et al., 2014; Yang et al., 2014; Zeng et al., 2008). My goal was to analyze cytoplasmic and nuclear PTMs separately using purified extracts from either compartment. Analysis of nuclear AGO2 requires rigorous isolation of cell nuclei. AGO2 is found in the endoplasmic reticulum (ER) contiguous with the nuclear envelope (Stalder et al., 2013) and my laboratory previously developed protocols for separating nuclei from ER (Gagnon et al., 2014b).

I used those same methods with the addition of two phosphatase inhibitors (sodium fluoride and sodium orthovanadate). Western blots demonstrate extract purity **(Figure 2.2C)**.

For purification of AGO2, antibody-conjugated beads were packed into a chromatography column and extracts were run through the column via gravity flow two times to improve the amount of AGO2 purified. The beads were washed three times with IP Wash buffer and bound AGO2 was further purified using cold SDS buffer, which will elute most of the proteins not directly associated with the antibody. SDS buffer heated to 95^{o}C was then used to elute the bound AGO2 off of the antibodies. Samples were desalted, concentrated and run on an SDS-PAGE gel. Coomassie staining and western blotting were routinely used to confirm the presence of sufficient AGO2 after purification to enable for adequate coverage for mass spectrometry analysis **(Figure 2.2D, E)**. Gel bands were excised, trypsinized and run through an HPLC column prior to mass spectrometry analysis.

2.4. Unbiased PTM search for FLAG and endogenous AGO2 modifications

Mass spectrometry datasets were analyzed in two ways to yield the broadest PTM identification. For the first analysis, each peptide spectra was run through an algorithm called ModLS (Modification Localization Scoring) to assess the confidence of the PTM assignments in both the identification and localization of the correct residue harboring the PTM (Baker et al., 2011; Chalkley and Clauser, 2012; Sunyaev et al., 2003). In addition to the analysis for specific PTMs in the standard analysis pipeline, an unrestricted PTM search was performed using the "Error tolerant" method in the Mascot software (Mann and Wilm, 1994; Yu et al., 2016). This method considers the possibility that any modification defined in the Unimod.org database may be present. This is a sensitive procedure that reports many putative modifications, however it is not very specific (i.e., artefactual modifications from sample prep). The proteomics core staff searches through the results manually and identifies the most biologically relevant PTMs in a

report. Together, these two methods produce the most accurate results for the identification of PTMs on a protein of interest. It should be noted that each PTM identified needs to be verified independently of mass spectrometry for true identification.

2.5. Cytoplasmic and nuclear FLAG-AGO2 PTMs

I began my analysis by examining the PTMs associated with FLAG-AGO2 purified from cytoplasmic and nuclear extracts. FLAG-AGO2 is much easier to purify than endogenous AGO2, making it a good starting point for data collection. For triplicate cytoplasmic and nuclear samples, mass spectrometry identified 85%-90% of the AGO2 sequence from the peptides in each sample. Ten PTMs were identified for cytoplasmic AGO2 and nine PTMs were identified for nuclear AGO2 in at least two of the three samples **(Table 2.2)**. Identified PTMs included acetylations, methylations, ubiquitinations, and phosphorylations. Both cytoplasmic and nuclear samples had four PTMs in common, meaning those likely don't lead to differential function or localization in both compartments.

Four PTMs in my samples have been previously identified **(Table 2.1, 2.2)**. S387 phosphorylation has been implicated in the localization of AGO2 to P bodies in the cytoplasm (Zeng et al., 2008), but it was also identified on nuclear AGO2 in this study. Y529 phosphorylation in cytoplasmic AGO2 samples has been previously characterized as a cytoplasmic modification and implicated to regulate AGO2 small RNA binding at the 5'-end (Rudel et al., 2011). The S824-834 phosphorylations identified on both cytoplasmic and nuclear AGO2 in my samples have been proposed to regulate AGO2 miRNA:target interactions in the cytoplasm (Golden et al., 2017) , and their presence on nuclear AGO2 is novel. S253 phosphorylation found on nuclear AGO2 has previously been identified from whole cell extracts and not characterized for functional alteration of AGO2 (Rudel et al., 2011).

The rest of the modifications identified are novel to the field. In fact, no other studies have identified methylations or acetylations for AGO2. Three residues identified as acetylated or methylated in my samples in either compartment were previously identified to harbor different modifications **(Figure 2.1A)**: 1) K402 sumoylation regulates AGO2 protein stability (Qi et al., 2008) and was found to be both methylated and acetylated in cytoplasmic samples 2&3) K248 and K726 ubiquitination were identified in a global ubiquitinome study (Kim et al., 2011). K248 was found to be methylated in the cytoplasm and both K248 and K726 were found to be acetylated in the nuclear samples. Another example of differential modification between compartments is the K355 acetylation in both compartments and ubiquitination of nuclear AGO2. This suggests the possibility that these residues harbor different PTMs in various contexts.

2.6. Identification of cytoplasmic and nuclear endogenous AGO2 PTMs

To ensure that the modifications observed for FLAG-AGO2 were not an artefact of the FLAG tag or overexpression of AGO2, I carried out analysis on endogenous AGO2 from cytoplasmic and nuclear extracts. I submitted one sample from both cytoplasmic and nuclear extracts and mass spectrometry identified 70-75% of the AGO2 sequence. The peptide sequence coverage is lower than FLAG-AGO2 samples because I was not able to purify an equivalent amount using the anti-AGO2 antibody. After analysis of the peptide spectra, four PTMs were identified for cytoplasmic AGO2 and three PTMs were identified for nuclear AGO2 **(Table 2.3)**. As observed in FLAG-AGO2 samples, S387 phosphorylation was identified in both the cytoplasm and nucleus and K248 methylation was observed in cytoplasmic samples **(Table 2.2)**. Another novel modification, R506 methylation, was identified in both compartments. Two novel methylations, R167 and R196, were identified for cytoplasmic and nuclear AGO2, respectively. It should be noted that arginine methylations have been implicated to have functional roles in

splicing, transcriptional control, and nuclear-cytoplasmic shuttling (Bedford and Clarke, 2009), so the differing methylation status between compartments would be interesting to follow up on.

2.7. Identification of cytoplasmic and nuclear endogenous TNRC6A PTMs

As mentioned previously, there have been no focused PTM analyses for TNRC6A to date, but only phospho-proteome studies. I collected PTM data for TNRC6A as a result of another proteomic study not specifically focused on identifying PTMs. Endogenous TNRC6A protein complexes were purified from cytoplasmic and nuclear extracts with an anti-TNRC6A antibody and submitted to the proteomics core for identification of proteins in the sample. A caveat to this method is that I did not get optimal coverage of the TNRC6A protein sequence because there are other proteins in the sample drowning out signal.

After initial data collection, the peptide spectra of TNRC6A were reanalyzed for modification sites and scored for ID and localization on the peptides. The cytoplasmic sample analyzed had 79% coverage of the TNRC6A sequence and the nuclear triplicate samples had 61-64% sequence coverage. These are lower than ideal because the purification was not performed specifically for TNRC6A. However, with these data, I was able to identify novel PTMs for TNRC6A in both compartments **(Table 2.4)**. Cytoplasmic modifications included three acetylations and a methylation. Nuclear data included a phosphorylation and methylation. There was no overlap between samples. While no PTMs have been specifically identified and characterized for TNRC6 proteins, here have been global phospho-proteome studies that identified similar T736/T738/S739 phosphorylation for TNRC6A (Brill et al., 2009; Sharma et al., 2014).

2.8. Conclusions

Overall, I was able to identify several novel PTMs for AGO2 and TNRC6A and they were also novel in a sense that the methylation and acetylation status have not been characterized for

either protein. I identified three previously characterized PTMs for cytoplasmic AGO2 and showed that S253 and S387 phosphorylation are also present on AGO2 isolated from the nucleus of cells. I also demonstrated that residues K248, K355, K402 and K726 can potentially harbor acetylations, ubiquitinations, and methylations in different cellular contexts within compartments or between compartments. While my endogenous AGO2 and TNRC6A datasets did not produce many confident PTMs, they were still novel datasets that identified PTMs on endogenous RNAi proteins and a good start with room for further improvements.

2.9. Improving methods for identification of PTMs

While these data are a good starting point and I identified several novel PTMs for both AGO2 and TNRC6A, I did not identify as many PTMs as I would expect and the data varied greatly from sample to sample. That is why I considered PTMs to be significant if they were identified in at least two of three samples for all datasets. Unfortunately, this is an inherent problem with PTM identification from mass spectrometry. Most PTMs are unstable after protein isolation and with my immunoprecipitation methods, I purify protein from all stages of the cell cycle and even from different locations in nuclear extracts, so this results in variable data. Despite the shortcomings of the method, there are ways to improve identification that could be implemented for future studies.

One of the biggest obstacles to overcome is to not only improve protein sequence coverage, but to increase the amount of purified protein to increase the odds that lowly abundant or unstable PTMs can be identified at higher levels. It should be noted that while 70-90% sequence coverage is good for PTM identification, more peptides harboring PTMs need to be isolated above the background noise. For my purifications, I started with 60mg of protein, which amounts to about 1 billion cells isolated from 40 15-cm dishes for cytoplasmic extracts and about 1.5 billion cells from 60 15-cm dishes for nuclear extracts. This was already a high amount of

protein to start out with and at the time, it did not seem feasible to double or even triple the amount used. So, for future studies, I would suggest using large-scale overnight IPs rather than antibody affinity chromatography. This would increase the time that the antibody conjugated beads have to capture AGO2 in the sample instead of just flowing the extract through the packed beads twice. The feasibility of this method is limited by the amount of material needed. It would require two to three times the amount of antibody and beads used. However, I believe it would improve identification of PTMs significantly.

Another step of the purification methods used that could be improved upon is the PAGE-gel purification of the protein before excising the band for mass spectrometry analysis. This step required significant attention and troubleshooting because early samples would diffuse through the gel **(Figure 2.3)** and could not be used for further analysis because the proteomics core has specific requirements for the size of the excised band to efficiently digest proteins. I determined that my samples contained too much salt to run normally on a gel, so I started desalting my samples with desalting columns before concentrating them to a volume that could be run on a gel. This worked for some samples **(Figure 2.2D)**, but other samples still diffused throughout the gel. That could be because there is too much protein in each sample to run down one lane on a gel. So, I started splitting up the sample into two gel lanes. This improved my coomassie stain signal, but the methodology still needs improvement. For future studies, I would propose splitting the sample into 3 or 4 gel lanes. That way, the proteomics core can efficiently digest my protein and pool the data to get improved protein coverage. The downside to this improvement is that it would cost twice the amount of money because each sample is analyzed separately before pooling the data.

A way to improve TNRC6A PTM identification is to use the same affinity chromatography method with the anti-TNRC6A antibody as used for AGO2 purification. Other than that, PTM identification could be improved by using another enzyme besides trypsin to digest the protein. I

noticed that TNRC6A does not have as many trypsin digestion sites as most proteins and that results in large peptides. Having many larger peptides results in lower sequence coverage because they cannot be efficiently identified in the mass spectrometer. Having larger peptides can also result in ambiguous localization of a PTM. If a large peptide is identified to harbor a PTM, but has several serine residues near to each other, it is nearly impossible to determine which serine is harboring the PTM identified on the peptide. This phenomenon is present in both my AGO2 and TNRC6A samples and represented as S824-S834 phosphorylation or T736/T738/S739 phosphorylation, respectively. The only way to get an accurate localization of the phosphorylation is to perform mutagenesis to each residue and determine whether it still harbors the modification. Using a different enzyme to digest the sample can result in completely different peptides that will likely be smaller and may split up these ambiguous modification sites. The proteomics core offers the use of different proteases, so those should be considered moving forward with this project.

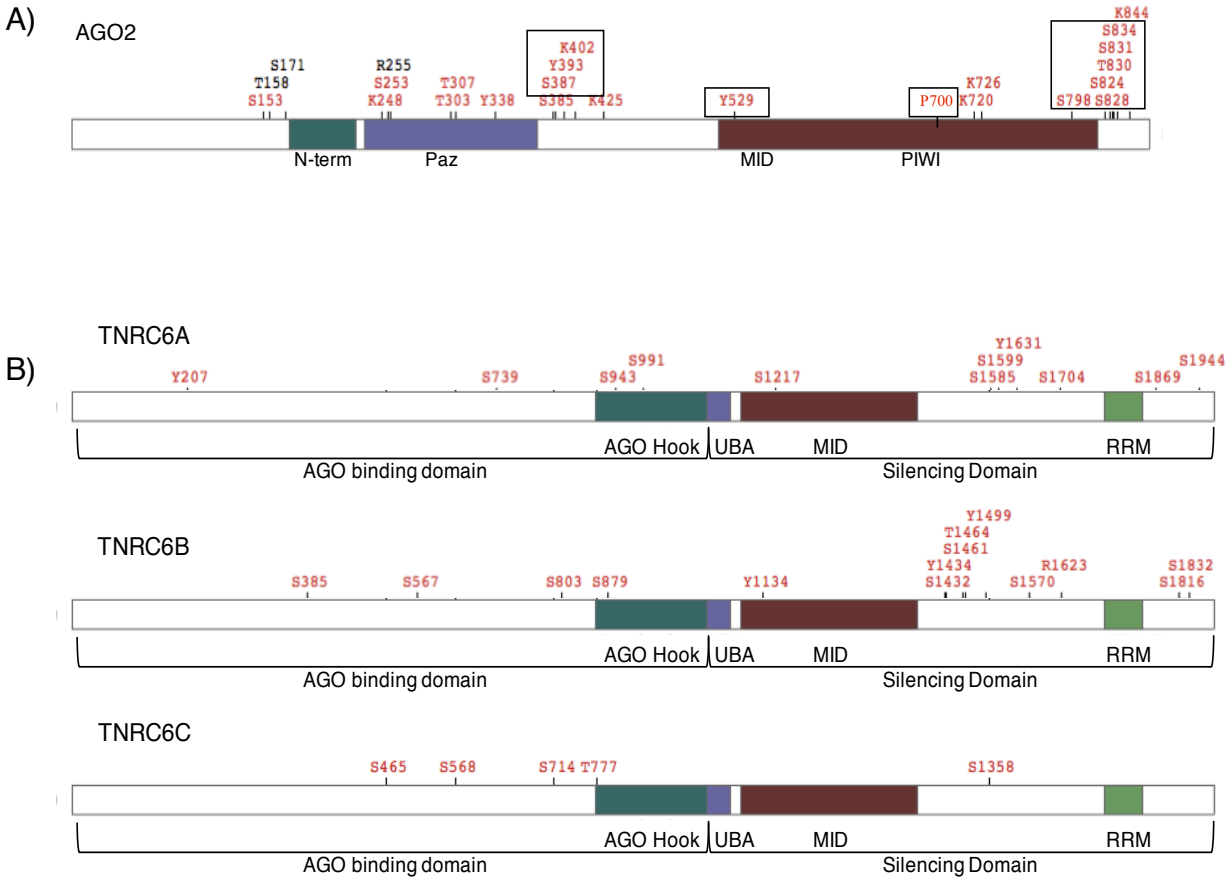


Figure 2.1. Representation of PTMs for AGO2 and TNRC6 RNAi factors.

a) AGO2 PTMs that have been discovered and cited more than once. These include phosphorylations, acetylations, sumoylations, and hydroxylations. Boxes indicate residues where modifications have been verified for their functional alteration of AGO2.

b) TNRC6 paralog PTMs reported through global phosphoproteome studies.

Modification	AGO domain	Function	Reference
S387 Phos	PAZ-MID	Localization to P bodies	Lopez-Orozco, 2015
S798 Phos	PIWI	Localization to P bodies	Zeng, 2008
Y529 Phos	MID	Binding to 5'-end of RNA	Rudel, 2011
K402 Sumo	PAZ-MID	Protein stability	Qi, 2008
P700 Hydrox	PIWI	Protein stability	Sahin, 2014
Y393 Phos	PAZ-MID	Inducing senescence in cancer cells	Shen, 2013; Yang, 2014
S824834 Phos	PIWI	Regulate binding to miRNA targets	Golden, 2017

Table 2.1. Table of characterized AGO2 PTMs for cytoplasmic function.

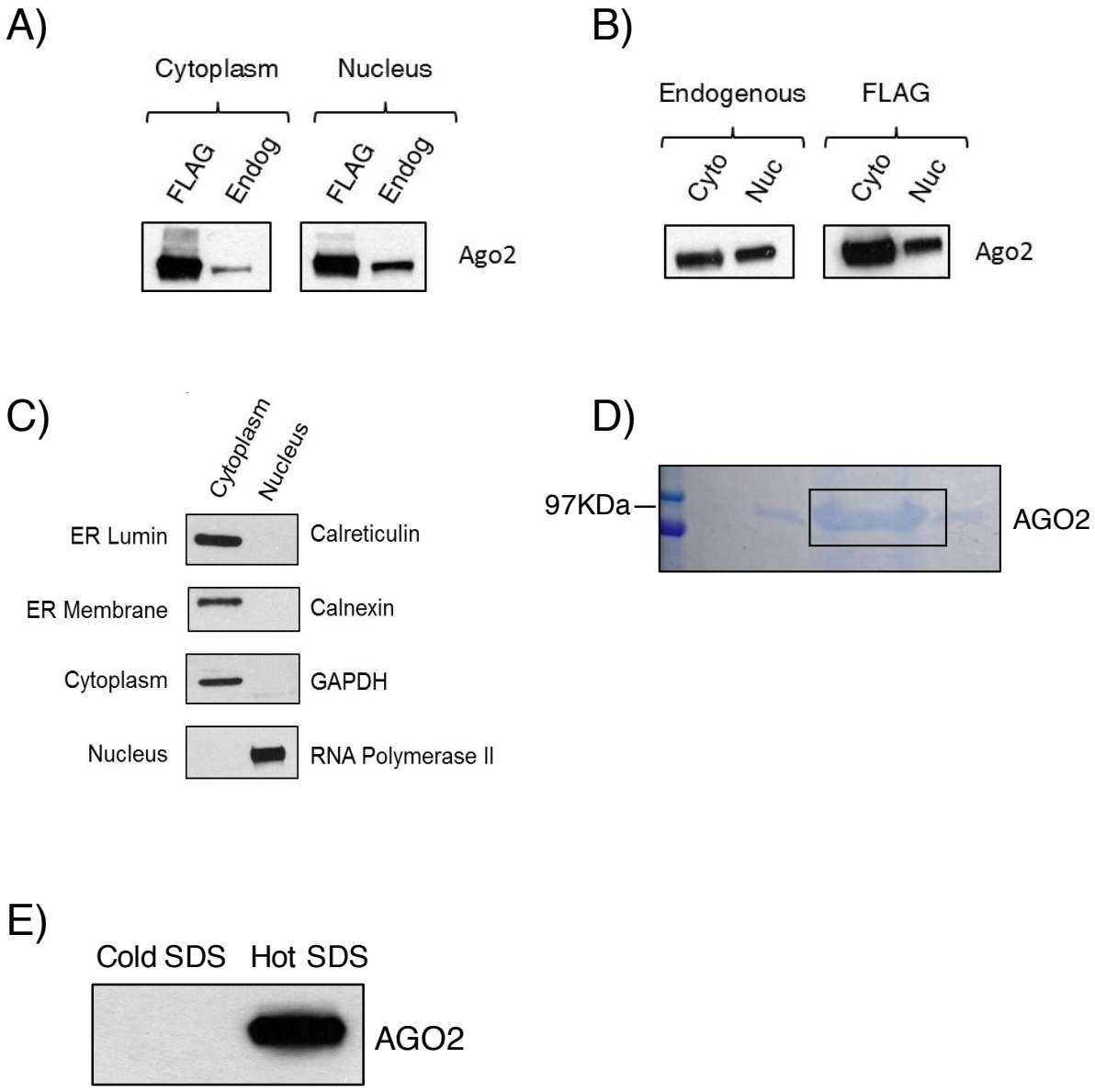


Figure 2.2. Analysis of cellular extracts and samples prior to mass spectrometry analysis of PTMs.

a) Western blot comparing levels of AGO2 in wild type and FLAG-AGO2 cell line.
b) Western blot comparing cellular distribution of AGO2 in wild type and FLAG-AGO2 cell line.
c) Western blot for extract purity.
d) Coomassie stain of purified AGO2 prior to submission to the proteomics core.
e) Western blot of both the cold and hot SDS elution of purified AGO2 prior to submission to the proteomics core.

FLAG AGO2 PTMs	
Cytoplasm	**Nucleus**
S387 phosphorylation*	S387 phosphorylation
K355 acetylation*	K355 acetylation
K720 acetylation*	K720 acetylation
S824-S834 phosphorylation*	S824-S834 phosphorylation
K402 acetylation	K248 acetylation
K248 methylation	K726 acetylation
Y529 phosphorylation	K355 ubiquitination
R583 dimethylation	K608 acetylation
K402 dimethylation	S253 phosphorylation
K313 mono methylation	

Table 2.2. Table of FLAG-AGO2 PTMs from triplicate samples. In order to be considered significant, the PTM must have high confidence and localization scores in at least two of the three samples.
*Represents common PTMs between samples

Endogenous AGO2 PTMs	
Cytoplasm	**Nucleus**
S387 phosphorylation*	S387 phosphorylation
R506 methylation*	R506 methylation
R167 methylation	R196 methylation
K248 methylation	

Table 2.3. Table of endogenous AGO2 PTMs from one sample. In order to be considered significant, the PTM must have high confidence and localization scores. *Represents common PTMs between samples.

Endogenous TNRC6A PTMs	
Cytoplasm	**Nucleus**
K226/236 acetylation	T736/T738/S739 phosphorylation
K729 acetylation	R1314 methylation
K915 acetylation	
K1363 methylation	

Table 2.4. Table of endogenous TNRC6A PTMs from duplicate (cytoplasm) and triplicate (nucleus) samples. In order to be considered significant, the PTM must have high confidence and localization scores in at least two of the three samples.

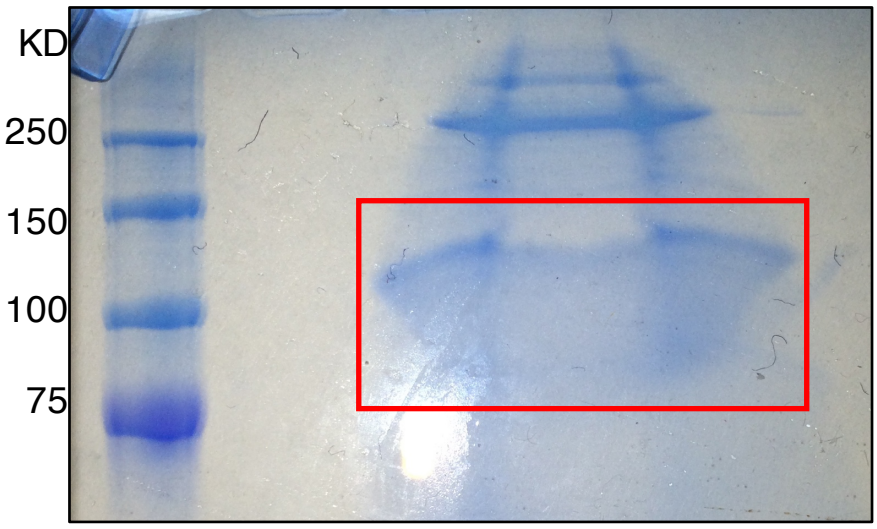


Figure 2.3. Example of a coomassie stain of an unsuccessful gel sample from an AGO2 purification. The sample diffused throughout the gel and became too big to submit to the proteomics core.

Chapter 3
Identifying Stable Interacting Factors of AGO2

3.1. Introduction

Argonaute (AGO) proteins are key components of RNA interference (RNAi) (Meister, 2013) and bind directly to miRNAs and siRNAs to form complexes that regulate gene expression. AGO proteins are conserved and are expressed ubiquitously in animals, plants, and yeast. There are four AGO variants in human cells (AGO1-4) (Schurmann et al., 2013). AGO2 is the only AGO protein to retain catalytic activity, and when combined with an siRNA, can reconstitute RNAi in defined cell free systems (Liu et al., 2004).

AGO2 is essential for post-translational silencing of mRNA in the cytoplasm of mammalian cells and many studies have focused on this cytoplasmic role. Additional factors known to be involved in cytoplasmic RNAi include TNRC6A (GW182) (Eulalio et al., 2008; Jakymiw et al., 2005; Nishi et al., 2013; Pfaff and Meister, 2013), Dicer (Bernstein et al., 2001; Ha and Kim, 2014), and TAR RNA binding protein (TRBP) (Daniels and Gatignol, 2012; Takahashi et al., 2014; Wilson and Doudna, 2013). TNRC6A influences the subcellular localization of AGO2 and is critical for miRNA-guided silencing. Dicer is an endoribonuclease that is responsible for processing pre-miRNAs to form mature, double-stranded miRNAs. Dicer and TRBP bind double-stranded miRNAs and assist loading of the guide strand into AGO.

While most attention has focused on RNAi in the cytoplasm, AGO proteins are also active in the nucleus (Gagnon and Corey, 2012; Gagnon et al., 2014a; Schraivogel and Meister, 2014; Schraivogel et al., 2015b). AGO2 can act in conjunction with guide RNAs to regulate transcription (Chu et al., 2010; Matsui et al., 2013) and splicing

(Ameyar-Zazoua et al., 2012; Liu et al., 2015) in human cells. While RNAi can be active in cell nuclei, there are also significant differences. Trax, Translin, and Hsp90, factors that assist guide strand loading onto AGO2, are absent in cell nuclei and loading does not occur in nuclear extracts (Gagnon et al., 2014a).

Understanding the role of RNAi in cell nuclei will require information about nuclear protein-protein interactions and how they compare with those occurring during cytoplasmic RNAi. While no study has specifically examined global nuclear interactions, several proteomic studies have investigated potential interactions of AGO2 in whole cell, cytoplasm, and chromatin (Ameyar-Zazoua et al., 2012; Carissimi et al., 2015; Frohn et al., 2012; Hock et al., 2007; Landthaler et al., 2008; Meister et al., 2005; Robb and Rana, 2007; Weinmann et al., 2009).

Tuschl and co-workers examined interactions of FLAG-tagged AGO1 and AGO2 (Meister et al., 2005). They used gel electrophoresis to purify specific bands and observed interactions with TNRC6B and MOV10. In a subsequent study, Meister used cells expressing FLAG-AGO1 or FLAG-AGO2 and immunoprecipitated tagged AGO1 and AGO2 from whole cell extracts (Hock et al., 2007). This study identified a large number of candidate interacting proteins including several involved in RNA metabolism. Robb and Rana took an alternative approach using a biotinylated siRNA to isolate Ago complexes prior to mass spectrometry (Robb and Rana, 2007). Two purified bands were analyzed and the proteins RHA and HSP 90b were identified. Tuschl subsequently examined interactions of FLAG-tagged Ago1-4 in whole cell extract (Landthaler et al., 2008). This study found a wide range of interacting proteins including Dicer, YBX1, RPS6, and RPL7.

Meister looked for common interacting factors of all 4 FLAG-tagged AGO proteins and identified Importin 8 (Weinmann et al., 2009). Finally, Meister's group used whole cell extract from mouse embryonic fibroblast cells and found that AGO2 can interact with RNA in the absence of miRNA (Frohn et al., 2012).

One study has focused on examining proteins within chromatin. Harel-Bellan and colleagues examined interactions between proteins with FLAG-AGO1 and FLAG-AGO2 and identified SRSF splicing factors as candidate interacting partners (Ameyar-Zazoua et al., 2012). Another recent study by Carissimi in whole cell extracts, reported interactions between endogenous AGO2 and SWI/SNF proteins normally found in the nucleus (Carissimi et al., 2015).

These prior studies have provided important information about candidate interacting proteins, but gaps remain in our understanding of factors that interact with AGO2 to facilitate function in both cytoplasm and nuclei. Modern mass spectrometry is sensitive and detects potential interactions with many proteins. Discriminating between real and artefactual interactions is a major challenge to accurate analysis.

One limitation of prior work is that most studies relied on over-expressed FLAG-tagged AGO2. The use of a tagged protein increases efficiency of immunoprecipitations, but using a tag affects interactions with other proteins, and overexpression can cause non-physiologic interactions and increase identification of false positives (Oeffinger, 2012). None of these previous studies compared interactions of FLAG-tagged AGO with endogenous AGO. Finally, most studies also did not use semi-quantitative or quantitative

techniques, further complicating the identification of candidates over background identifications that are not physiologically relevant.

The following is a summary of the project I was integrated into when I joined the laboratory. The project was initiated by a former graduate student, Roya Kalantari, and I joined early in the project's execution. We worked together to troubleshoot many problems encountered and after she left, I carried the manuscript through to publication.

To identify nuclear interactions of AGO2, we performed a semi-quantitative proteomic study. All studies were performed in duplicate or triplicate. To further reduce detection of artefactual proteins and efficiently prioritize candidates for experimental validation, we compared results from samples obtained after purification of FLAG-tagged AGO2 complexes and endogenous AGO2 complexes from nuclear extracts.

3.2. Prioritizing candidate interacting partners for AGO2.

We used mass spectral analysis to investigate interactions of AGO2. Large-scale proteomic analysis of interaction networks often leads to long lists of candidate proteins and it is important to prioritize the candidates for experimental validation. Therefore, our threshold values for significance were set higher than most studies.

For semi-quantitative proteomic analysis, we used the Normalized Spectral Index (SINQ) method (Griffin et al., 2010). SINQ is a label-free proteomic quantitation method. Normalized spectral counts for each protein are calculated using many factors, including the number of peptides identified for the protein, the weight given for the assignment of each peptide to the protein, the number of spectrum matches for each peptide found in

the protein, and the summed fragment ion-intensity for the assignment of an MS/MS spectrum to a peptide. This value is then normalized by the sum of all these values across all proteins in the data set, correcting for differences in protein loading between data sets. This result is then divided by the protein length, correcting for the propensity of longer proteins to produce a greater number of peptides than shorter proteins.

Normalized spectral counts can then be used to provide ratios between sample and control for each protein (Griffin et al., 2010; Trudgian et al., 2011a). Protein candidates characterized by a minimum spectral count of 5 as well as a minimum enrichment ratio of 5:1 for sample versus control were identified as top candidates. Data are presented as normalized spectral counts for each significant protein.

3.3. Nuclear interactions of AGO2

For these studies, we used two different cell lines to make nuclear extracts for mass spectrometry purifications. We used stably expressing FLAG-AGO2 T47D cells and regular T47D cells. We performed a parallel analysis with immunoprecipitations of overexpressed FLAG-AGO2 (FLAG antibody) and endogenous AGO2 (anti-AGO2 antibody) to identify the interaction network of nuclear AGO2. Overlap between these datasets represent significant interactions of AGO2.

After mass spectrometry of proteins associated with FLAG-AGO2, twenty-one proteins were identified as candidate interacting partners (**Figure 3.1A**). The top candidates possessed a wide array of functions, from apoptosis to RNA processing and included the RNAi factors Dicer, TNRC6A, TNRC6B, and TNRC6C. In contrast, the use

of anti-AGO2 antibody for the immunoprecipitation of endogenous AGO2 led to detection of fewer candidate proteins. Mass spec identified interactions with the RNAi factors AGO3, AGO1, TNRC6A, TNRC6B, and TNRC6C as well as the signal transduction protein LGALS1 (**Figure 3.1B).** LGALS1 was the least significant protein on the list.

The interaction with TNRC6A was confirmed by immunoprecipitation with anti-AGO2 antibody followed by western analysis (**Figure 3.2A**). The potential for interaction between AGO2 and and the other AGO proteins, AGO1 and AGO3, was less clear. The mass spectrometry datasets indicated a potential interaction with AGO1 and AGO3 after immunoprecipitation with an anti-AGO2 antibody (**Figure 3.1B**) but not when FLAG-AGO2 was recovered using an anti-FLAG antibody (**Figure 3.1A**). Co-immunoprecipitation using an anti-AGO2 antibody confirmed AGO1 and AGO3 as an interaction of endogenous AGO2 (**Figure 3.2A**). However, a similar analysis of proteins recovered after isolation of FLAG-AGO2 using an anti-FLAG antibody did not reveal association between AGO2 and AGO1, but did reveal an association between AGO2 and AGO3 (**Figure 3.2B**). Taken together, the mass spectral data and western analysis are most consistent with the conclusions that detection of an AGO1:AGO2 interaction is due to cross-reactivity between AGO1 and the anti-AGO2 antibody. However, the detection of a AGO3:AGO2 interaction is detected in both datasets and co-immunoprecipitations, which means it is likely a real interaction that was not detected in the FLAG-AGO2 purifications at the set threshold. It should be noted that the mass spectrometry data doesn't rule out interactions that aren't detected.

The majority of the factors identified by MS that immunoprecipitated with endogenous AGO2 were RNAi factors, however one potentially novel factor was also identified. LGALS1 is a protein that has known extracellular functions (Astorgues-Xerri et al., 2014) but was nonetheless identified in the nuclear MS data as the only non-RNAi candidate. It was not identified in the FLAG-AGO2 MS, and it was the least significant protein identified in endogenous AGO2 mass spec (**Figure 3.1B**). This suggests that if this interaction is biologically relevant it may be very weak. It was not co-purified with AGO2 through co-IPs (**Figure 3.2C**).

We also tested the potential for AGO2 interactions with the protein SMARCC1 that had been identified as an interacting factor in a study using whole cell extracts for mass spectrometry (Carissimi et al., 2015). SMARCC1 is an intriguing candidate because the protein is primarily nuclear and has functions in the SWI/SNF complex. Both the nuclear mass spectral analyses however, had not detected SMARCC1 as a candidate interacting protein. Co-IPs using anti-AGO2 antibody for the pull-down from nuclear extract did not identify an interaction between AGO2 and SMARCC1 (**Figure 3.2C**). SMARCC1 also failed to pull down AGO2 in a reciprocal Co-IP (**Figure 3.2D**). I also made whole cell extracts exactly as they did in their study and immunoprecipitated AGO2 complexes for mass spectrometry analysis. After careful analysis using the SINQ method, I did not observe SWI/SNF proteins in my significant dataset **(Figure 3.3).** One explanation of the results not supporting their datasets is that they used a different anti-AGO2 antibody, so it might be cross-reacting with non-specific proteins.

3.4. Focusing candidate identification using data from orthogonal analyses

A primary challenge confronting the use of mass spectrometry data is prioritizing the selection of candidates for further testing (Trapp et al., 2014). We adopted orthogonal approaches to crosscheck mass spectrometry data and restrict the number of candidate proteins for subsequent validation. One dataset was obtained using FLAG-based purification of FLAG-AGO2. The other dataset was based on samples analyzed after purification of endogenous AGO2 using anti-AGO2 antibody. Candidates identified in common by the two strategies became the best candidates for further validation.

FLAG purifications from all replicates yielded many more candidates than purifications using an anti-AGO2 antibody (**Figure 3.1A**). This outcome may be because the FLAG-AGO2 protein was overexpressed relative to endogenous AGO2. Alternatively, the FLAG tag may be forming independent interactions with cellular proteins or altering the interactions made by AGO2. I note that the orthogonal analysis, while generating a shorter list of highly ranked candidates, provides no direct evidence that proteins identified in FLAG-AGO pull downs but not endogenous-AGO pulldowns are less likely to be biologically relevant interactors. Indeed, some of the top candidates from both FLAG and anti-AGO purifications were the same proteins. Overall, orthogonal screening was useful in this study because it increased confidence in candidate selection for experimental validation, allowing resources to be directed more efficiently.

The evaluation of AGO1 as a candidate is an example of the advantage of incorporating orthogonal crosschecking into mass spectrometry. When purifying endogenous AGO2, we consistently identified AGO1 as an interacting partner. By

contrast, AGO1 was not identified as an interacting partner of FLAG-tagged AGO2. AGO2 and AGO1 have similar amino acid sequences, with 78.5% sequence homology. It is possible that the anti-AGO2 antibody used in these studies can also recognize AGO1 and be detected by mass spectrometry.

3.5. Stable interactions of TNRC6 protein family in the nucleus

When we compared nuclear extract data from the two isolation strategies we identified three common interacting factors, TNRC6A, TNRC6B, and TNRC6C (**Figure 3.1**). These data suggest that a stable RNAi complex of AGO2 and TNRC6 paralogs is conserved in the nucleus and can be readily detected **(Figure 3.4)**.

TNRC6A/GW182 is a well-known partner of AGO2. It is known to stabilize the AGO2-miRNA-mRNA interactions, recruit proteins for translational inhibition for miRNA silencing, and localize AGO2 to P-bodies in the cytoplasm (Pfaff et al., 2013; Pfaff and Meister, 2013). The three TNRC6 paralogs (A, B, and C) have some similarities - they each bind AGO2 and localize the complex to P-bodies (Lazzaretti et al., 2009). They all share the same domain organization and have two unstructured GW-repeat regions (Ding and Han, 2007). My laboratory has previously observed that simultaneous knockdown of all three TNRC6 proteins blocked RNA-mediated activation of cyclooxygenase-2 expression while knocking down TNRC6A alone was insufficient (Matsui et al., 2013).

Ui-Tei and colleagues have suggested that TNRC6 proteins direct AGO2 shuttling between the nucleus and cytoplasm (Nishi et al., 2013). AGO2 does not possess an obvious nuclear import sequence, and the nuclear localization and export sequences

within TNRC6 may be critical for guiding the complex (Nishi et al., 2013). Meister and colleagues, however, have recently suggested that nuclear transport of AGO2 and TNRC6 paralogs depend on different and possibly redundant transport pathways (Schraivogel and Meister, 2014). They have gone on to show that TNRC6 proteins along with AGO2 use the importin pathway, and that nuclear localization of both can be codependent (Schraivogel et al., 2015b).

3.6. Conclusion

miRNAs and RNAi factors are present in mammalian somatic cell nuclei (Ohrt et al., 2012; Robb et al., 2005). In the cytoplasm miRNAs and RNAi factors regulate gene translation (Hammond, 2015). The presence of RNAi factors in the nucleus suggests they might also have the ability to regulate uniquely nuclear processes like transcription or splicing. Preliminary RNA sequencing (RNAseq) of nuclear RNAs that interact with AGO2 supports this hypothesis by showing significant read clusters near intron-exon junctions and at gene promoters (Chu et al., 2015). The potential for nuclear regulation by RNAi factors is reinforced by observations that synthetic duplex RNAs in combination with AGO1 or AGO2 can be robust regulators of transcription (Chu et al., 2010; Huang et al., 2010; Janowski et al., 2005a; Li et al., 2006b; Matsui et al., 2013; Morris et al., 2004) and splicing (Allo et al., 2014; Allo et al., 2009; Liu et al., 2012; Liu et al., 2015).

Understanding the potential for RNAi in cell nuclei requires a more detailed knowledge of protein-protein interactions made by nuclear RNAi factors. In this study we focused on AGO2 because it is i) a central factor for RNAi activity (Hammond et al., 2001; Joshua-Tor and Hannon, 2011; Rand et al., 2004); ii) present in mammalian cell nuclei

(Gagnon et al., 2014a; Ohrt et al., 2012; Robb et al., 2005); and iii) can modulate RNA-mediated control of transcription (Chu et al., 2010; Huang et al., 2010; Janowski et al., 2005a; Li et al., 2006b; Matsui et al., 2013); iv) can modulate RNA-mediated control of splicing (Chu et al., 2015; Liu et al., 2012; Liu et al., 2015).

Mass spectrometry is a powerful tool for studying protein interactions but it is challenging to efficiently discriminate between high priority candidate interactions and background. Orthogonal methods for sample isolation allowed us to crosscheck results and prioritize candidate interactions. A stable complex between AGO2 and TNRC6 family paralogs is conserved in the nucleus. The finding that these interacting partners of the RNAi complex are conserved in the nucleus supports the hypothesis that nuclear RNAi has the capacity to drive recognition of nuclear RNA sequences and affect nuclear RNA metabolism and expression.

We do not exclude the possibility that other proteins form significant interactions with AGO2 because interactions that are less stable but biologically significant may be undetected by mass spectrometry. Important interactions with nuclear proteins that are mediated through TNRC6 family proteins may also contribute to function but might not be apparent when AGO2 is used as the bait protein. Now that AGO2 has been shown to form a limited number of stable interactions that can be detected by mass spectrometry, it may be useful to shift attention to the next shell of interacting partners to gain further insight into how RNAi factors function in splicing and transcriptional regulation.

3.7. Overcoming challenges

I was integrated into this project early in execution, so I was involved in optimizing the protocols. There are several problems we encountered along the way including problems with running samples of SDS-PAGE gels and mass spectrometry contaminants. Another problem encountered was addressing reviewer concerns about the quality of our mass spectrometry datasets.

While optimizing the immunoprecipitation protocol, we determined that the sample size needed to be increased 5-10 fold from what we originally started with. This created several downstream issues. We had to increase our elution volume to 250uL, which means that samples needed to be concentrated to run on an SDS-PAGE gel for sample preparation. The most common sample preparation and digestion method for mass spectral samples is to run samples on SDS-PAGE gels. To reduce sample volume to run on a gel, we used centrifugal protein concentrators. After this step was added to our protocol, we noticed that this affected the ability of the sample to run on SDS-PAGE gels **(Figure 3.5)**, most likely because there was too much salt in the sample. This made it impossible to cut out our samples for mass spectral analysis, so we had to add another step prior to concentration. The samples were run through desalting columns. This corrected the salt concentration in samples and allowed them to run properly on gels.

We also encountered a problem with our datasets that involved a mass spectral contaminant drowning out the signal of other proteins. In those samples, the noise from the contaminant was so high that we could not get efficient identification of other proteins in the samples. This started happening to all of our samples submitted, so I worked for two months to determine the source of the contaminant. I meticulously ran our protocols changing one variable at a time and finally determined the source of the contaminant was

the coomassie stain we were using prior to sample submission. This reagent was marketed as “mass spec safe”, so we started using it because it was cheaper. Another laboratory encountered the same contaminant in their samples, so that helped me narrow down the list of reagents to find the source. After two months of troubleshooting, we started using another coomassie stain and did not encounter any more of those problems.

A problem our manuscript encountered was concerns about the quality of our mass spectrometry datasets. Since our manuscript was heavy in mass spectrometry datasets, the reviewers wanted to be sure we were publishing high quality data, which has been a problem with many proteomic studies. So, to increase statistical rigor when prioritizing protein rankings, I worked with a statistician to re-analyze our datasets with another program called ‘Statistical Analysis of the INTeractome’ (SAINT) (Choi et al., 2012a; Choi et al., 2011; Choi et al., 2012b). This program assigns confidence scores to prey-bait interactions and a 1% false discovery rate (FDR) was used to generate a list of high-confidence protein interactors with AGO2 in each dataset **(Table 3.1)**. After careful analysis, the same highly ranked interactions were identified from this analysis method as with our original SINQ analysis method. This provided confidence in our datasets and strengthened our manuscript.

3.8. Improving methods for candidate identifications

One question the datasets bring up is why we detected so few proteins interacting with AGO2 and why the results didn’t match up with previous studies. While our significance level was set higher than most proteomics studies, lowering the cutoffs would not have added many more relevant proteins to the lists. Therefore, I think there are a

few ways to improve immunoprecipitation methods to detect more relevant AGO2 interacting proteins.

One way to improve detection of candidate proteins through mass spectrometry of protein complexes is to use crosslinking methods to convert non-covalent interactions between proteins into covalent bonds (Rappsilber, 2011). The artificially fused protein interactions withstand high salt washing and denaturing conditions and thus can be analyzed using methods that normally dissociate protein complexes. This would lead to the identification of many more candidate protein interactions compared to conventional purifications. This would also lead to the identification of less direct protein interactions that are perhaps bridged by direct interacting proteins. One disadvantage of using these methods is that it might induce non-specific interactions to form or might alter the bait protein in a way that the antibody cannot recognize the epitope. Our purifications of AGO2 did not employ crosslinking. The interactions detected must be stable enough to survive multiple washings and desalting. Since the datasets were limited for endogenous AGO2, further studies using cross-linking techniques could be used to fully explore more transitory interactions or interactions that might be made indirectly through TRNC6 paralogs rather than direct association with AGO2.

Another way to improve detection of potential interacting partners is to use a different anti-AGO2 antibody to compare with FLAG-AGO2 datasets. Recently, another AGO2 antibody from WAKO chemicals has been characterized by my laboratory as more efficient for immunoprecipitations. This antibody has also been characterized to recognize AGO2 under crosslinking conditions for RNA-seq experiments, so it can be

used in combination with cross-linking techniques for improved identification of protein partners.

Another way to improve detection in endogenous AGO2 samples is to scale up the amount of nuclear extract used for immunoprecipitations. For consistency, we used the same amount of extract for FLAG and endogenous AGO2 samples. However, AGO2 is overexpressed about two fold more in the FLAG-AGO2 nuclear extracts. This could explain why FLAG-AGO2 purifications detected more potential protein interactions than endogenous AGO2 purifications. To balance out the detection of protein interactions for endogenous extracts, I would suggest doubling the amount of extract used in combination with cross-linking and possibly use of another anti-AGO2 antibody.

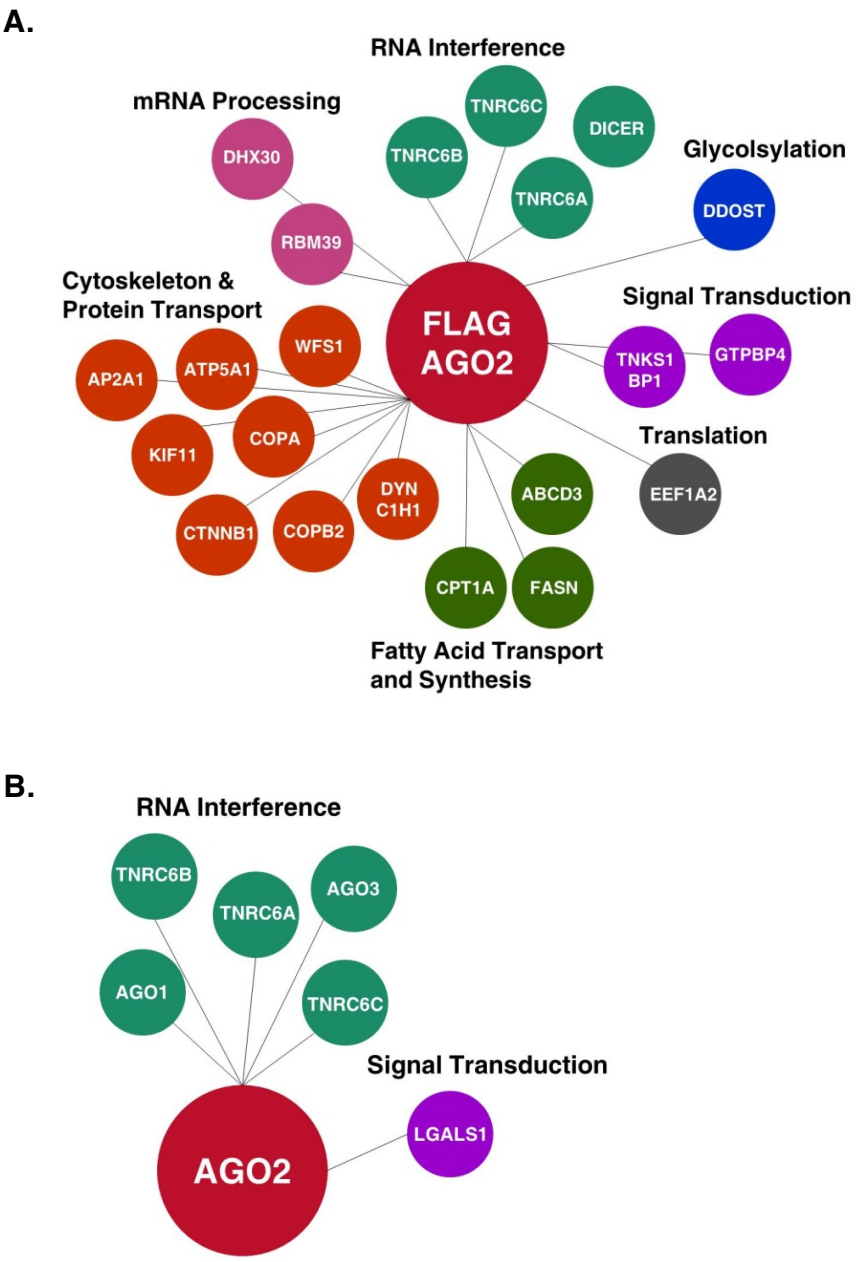


Figure 3.1. Identification of candidates from nuclear immunoprecipitations (IPs).

a) Bubble plot of candidates identified from FLAG-AGO2 IPs in the nucleus. Candidates are organized by function. n=2

b) Bubble plot of candidates identified from endogenous AGO2 IPs in the nucleus. Candidates are organized by function. n=3

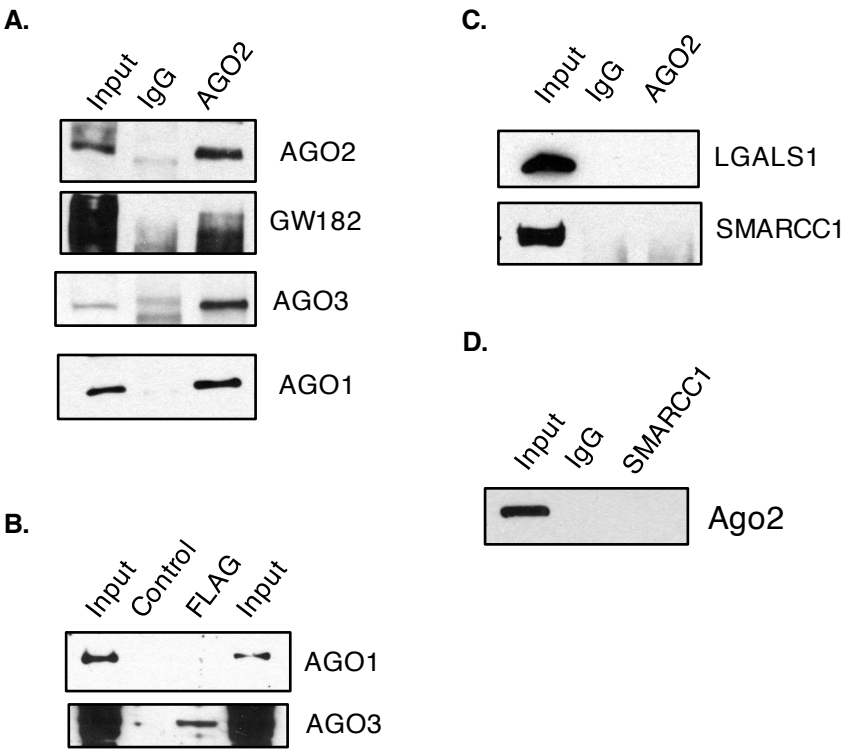


Figure 3.2. Co-immunoprecipitations of Candidate Interacting Factors.

a) Co-IP of AGO2 with RNAi factors from the nucleus. AGO2 was immunoprecipitated versus an IgG control and interacting factors were blotted for with antibodies by western blot.

b) Co-IP of FLAG-AGO2 with interacting factors from the nucleus identified by mass spectrometry. FLAG-AGO2 was immunoprecipitated from FLAG-AGO2 stable line cells versus a FLAG IP from endogenous T47D cells. Interacting factors were blotted for with antibodies by western blot.

c) Co-IP of AGO2 with potential interacting factor LGALS1 and previously identified interacting factor SMARCC1 from the nucleus. AGO2 was immunoprecipitated versus an IgG control and interacting factors were blotted for with antibodies by western blot.

d) Co-IP of SMARCC1 with AGO2 in nuclear extracts.

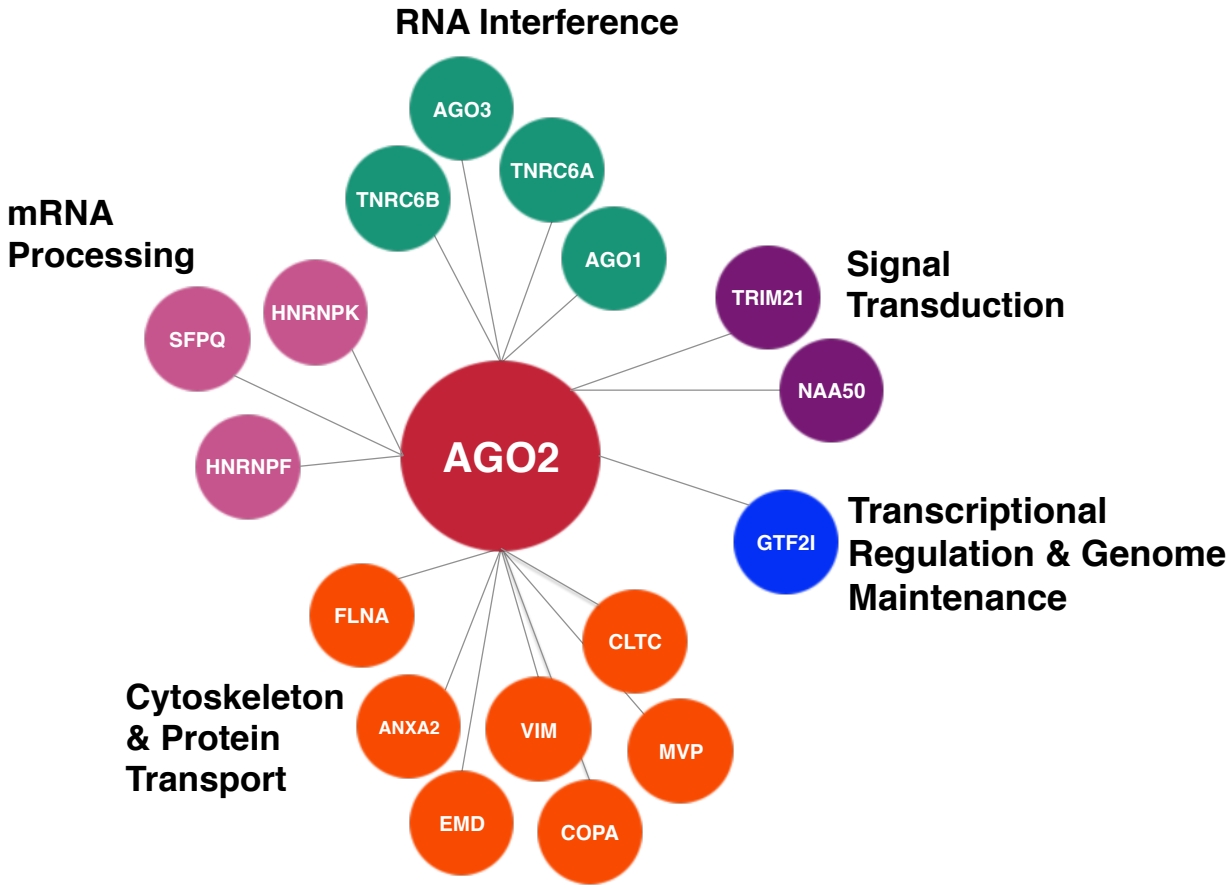


Figure 3.3. Identification of candidates from whole cell immunoprecipitations (IPs).
Bubble plot of candidates identified from endogenous AGO2 IPs from whole cell lysate. Candidates are organized by function. n=2

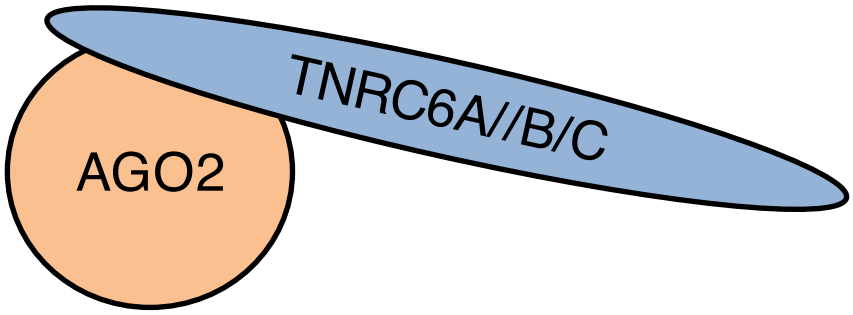


Figure 3.4. Model of AGO2 interactions with TNRC6 paralogs in the nucleus.

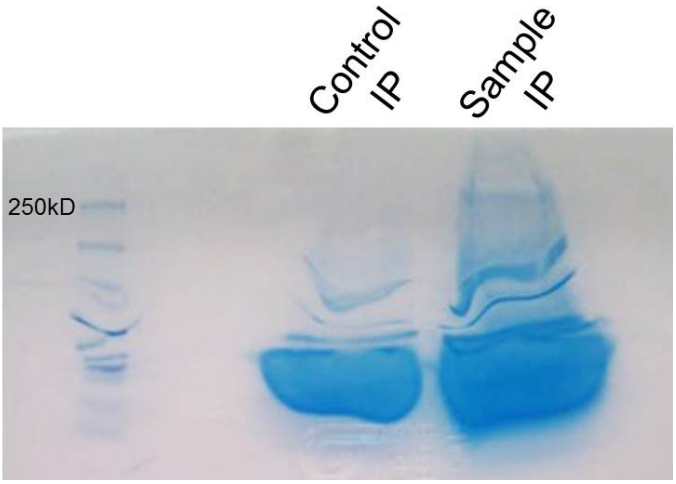


Figure 3.5. Coomassie staining of immunoprecipitations prior to desalting.
Control and sample immunoprecipitations were eluted and concentrated prior to running on the SDS-PAGE gel. Samples were run 15mm into the gel.

A.

FLAG AGO2 SAINT

Bait	Prey Gene	Spectral Counts	ctrlCounts	AvgP	BFDR
AGO2	FASN	19|26	0|1	1	0
AGO2	TNRC6C	12|14	0|0	1	0
AGO2	TNRC6A	7|10	0|0	1	0
AGO2	DHX30	6|12	0|1	1	0
AGO2	DICER	7|11	0|0	1	0
AGO2	USP10	3|4	0|0	1	0
AGO2	WFS1	9|18	1|3	0.99	0
AGO2	GTPBP4	6|8	1|1	0.99	0
AGO2	AGO3	3|3	0|0	0.99	0
AGO2	TNKS1BP1	33|51	2|12	0.98	0.01

B.

Endogenous AGO2 SAINT

Bait	Prey Gene	Spectral counts	ctrlCounts	AvgP	BFDR
AGO2	TNRC6A	109|68|52	1|0|1	1	0
AGO2	TNRC6C	79|81|69	0|0|0	1	0
AGO2	AGO1	72|58|56	0|0|0	1	0
AGO2	AGO3	80|75|55	0|0|0	1	0
AGO2	TNRC6B	116|94|106	0|0|0	1	0

Table 3.1. Tables representing significant proteins in each nuclear AGO2 mass spectrometry sample analyzed by SAINT.

a) FLAG nuclear AGO2 interactions. The top 10 proteins are shown in the table. AvgP is the average probability of a given bait-prey interaction across all replicates, as calculated by SAINT. FDR is the false discovery rate (FDR). For all samples, a cutoff of 1% FDR was used to represent true interactions of AGO2.

b) Endogenous AGO2 interactions. AvgP is the average probability of a given bait-prey interaction across all replicates, as calculated by SAINT. FDR is the false discovery rate (FDR). For all samples, a cutoff of 1% FDR was used to represent true interactions of AGO2.

Chapter 4
Human GW182 Paralogs are the Central Organizers for RNA-Mediated Control of Transcription

4.1. Introduction

MicroRNAs (miRNAs) and small interfering RNAs (siRNAs) can control gene translation by regulating the stability of mRNA in the cytoplasm (Ipsaro and Joshua-Tor, 2015). miRNAs act post-transcriptionally, typically by forming imperfect base-pairs within the 3'-untranslated region of mRNAs. siRNAs form perfect matches throughout mRNA and induce cleavage of target transcripts by argonaute 2 (AGO2) (Liu et al., 2004).

Cytoplasmic RNAi has had important applications. siRNAs are important tools for controlling the expression of specific genes in the laboratory. Numerous siRNAs are being tested in clinical trials of gene silencing in cancer, infectious disease, genetic disorders, and other conditions. Other clinical trials are underway using single-stranded oligonucleotides to block the function of naturally-occurring miRNAs or with RNA duplexes that mimic the structure of miRNAs (Matsui and Corey, 2016).

Small RNAs are also found in mammalian nuclei, as are key RNAi proteins like AGO2 and the GW182 paralogs TNRC6A, TNRC6B, and TNRC6C (Gagnon et al., 2014a; Matsui et al., 2015; Robb et al., 2005). The presence of both small RNAs and RNAi factors in nuclei suggests that RNA-mediated recognition may regulate RNA-dependent processes like transcription or splicing. Despite its potential impact, nuclear RNAi has remained an unexplored facet of gene regulation.

There have been reports that miRNAs can affect gene transcription in cell nuclei (Kim et al., 2008; Matsui et al., 2013; Place et al., 2008; Younger and Corey, 2011). Duplex RNAs have also been reported to affect gene splicing (Ameyar-Zazoua et al., 2012; Liu et al., 2012). While

these reports have built a strong case for nuclear RNAi function (Kalantari et al., 2016a), the detailed mechanism for transcriptional regulation has not been characterized, blocking progress towards understanding the broad significance of nuclear RNAi or the roles it might play in normal physiology and development.

AGO2 is an essential RNAi factor (Liu et al., 2004) that associates with small RNAs, assists recognition of complementary sequences, and induces cleavage of target RNAs when the match is fully complementary. AGO2 is essential for RNA-mediated regulation of transcription (Chu et al., 2010) and splicing (Liu et al., 2012).

Mass spectrometry of the partners of AGO2, however, revealed only limited insights (Kalantari et al., 2016b). The GW182 paralogs TNRC6A, TNRC6B, and TNRC6C and the AGO variant AGO3 were the only partner proteins detected. The narrow range of proteins that interact with AGO2 was insufficient to explain the observed functional control of small RNAs on transcription.

Because the proteomic analysis of AGO2 partners was inadequate to explain RNA-mediated transcriptional regulation, I examined the next shell of protein partners – those associated with TNRC6A. The GW182 family of proteins was first discovered as an autoantigen in patient serum (Eystathioy et al., 2002) and are characterized by glycine-tryptophan (GW) repeats within unstructured regions. TNRC6A and the closely related GW182 paralogs TNRC6B and TNRC6C are well known binding partners for cytoplasmic AGO2 (Pfaff et al., 2013; Yao et al., 2011). They stabilize AGO2-miRNA interactions and help localize AGO2 to cytoplasmic p-bodies.

TNRC6A is also found in cell nuclei (Gagnon et al., 2014a) and may be involved in shuttling AGO2 into the nucleus (Nishi et al., 2013; Nishi et al., 2015). TNRC6A is a multi-domain protein. Two domains play an essential role in silencing, the N-terminal GW-repeat domain and the bipartite silencing domain (Braun et al., 2013). These domains are embedded within

unstructured regions containing multiple glycine-tryptophan (GW) repeats that can form scaffolds for protein complexes.

A comprehensive study to identify protein interactions of TNRC6 proteins in cell nuclei has not been performed. However, there was a recent publication that used the FLAG-TNRC6-NES (nuclear export signal) mutant to pull out potential nuclear-specific interactions and identified them using mass spectrometry (Suzawa et al., 2017). They claimed to identify splicing proteins and RNA degradation proteins, including the CCR4-NOT complex as potential nuclear TNRC6A interactions. However, the proteins they focus on are lowly abundant on their lists and there are no validation experiments showing the interactions are real. In addition, most of the proteins they identify in their mass spectrometry experiments are clearly contamination from the endoplasmic reticulum (ribosomal proteins) and drown out useful signals from proteins that might be interactions in the nucleus.

In the present study, TNRC6 was chosen for proteomic analysis because of its demonstrated association with the critical RNAi factor AGO2, and its ability to be a scaffolding protein. TNRC6A, rather than TNRC6B or TNRC6C, was chosen for analysis because of the availability of an antibody suitable for isolating endogenous cellular protein from nuclear lysate for mass spectrometry (**Figure 4.1A**). TNRC6A is also the most abundant paralog found in HeLa cells (**Figure 4.1B**).

Here I report analysis of binding partners of TNRC6A in mammalian cell nuclei. I identify partners from multiple protein families, including mediator complex, histone modifiers, anaphase promoting complex, CCR4-NOT complex, and RNAi. My data build a mechanistic framework for understanding the action of small RNAs and RNAi factors in mammalian cell nuclei and suggests that TNRC6 proteins are central organizing factors in this process.

4.2. Immunoprecipitation and mass spectrometry of TNRC6A complexes

Nuclear extracts were prepared for large scale immunoprecipitations. All samples were treated with either DNase I or Turbonuclease to digest chromosomal DNA prior to immunoprecipitation. Turbonuclease, a broad spectrum RNase and DNase, was used to evaluate the influence of RNA on protein interactions and those samples were termed "Turbonuclease-treated". Samples treated with only DNase I were termed "DNase-treated".

Nuclear extracts were free of the cytoplasmic marker protein GAPDH and the ER marker Calreticulin (**Figure 4.1C**). After immunoprecipitation using anti-TNRC6A antibody and an anti-Rabbit IgG antibody as a control, the purified samples were analyzed by western blot to confirm the presence of TNRC6A and absence in the control (**Figure 4.1D**). Because a high depth of coverage was desired, all samples were required to have a visible band corresponding to the molecular weight for TNRC6A (**Figure 4.1E**). While TNRC6A knockout cells would also provide a useful reagent for comparisons, it was not straightforward to provide these due to the complications of chromosome duplications in HeLa cells.

Data was analyzed by semi-quantitative analysis using the normalized spectral index (SINQ) (Trudgian et al., 2011a). SINQ provides spectral counts and ratios between sample and control for each protein. To be deemed potentially significant, I required a minimum ratio of 4:1 for immunoprecipitated sample versus lysate control and at least six spectral counts be observed. Experiments were performed in duplicate or triplicate and the standards for potential significance needed to be met in all replicates. I identified thirty-eight top candidate proteins after turbonuclease treatment (**Figure 4.2A-C)** and thirty-three in samples treated with DNase I alone (**Figure 4.2D-F**). Significantly, in many cases, more than one member of a protein complex was identified. Most candidate proteins could be divided into three groups; RNAi factors, anaphase promoting complex, and proteins with known roles in transcription. The similarities of these two datasets are more important than their differences in this case. To better differentiate between RNase and DNase treatments, it would have been better to use RNase inhibitors in the DNase-

treated extracts. However, the fact that these complexes were identified in both datasets provides robust evidence that these are real interactions.

In addition to SINQ analysis, I performed stoichiometric analysis of protein interactions using the Intensity Based Quantification (iBAQ) method of Vermeulen and coworkers (Smits et al., 2013) (**Figure 4.3A-B)**. Using this approach in the RNase-treated dataset, one protein, TJP1, was present at a 15:1 ratio versus TNRC6A—an unlikely outcome (**Figure 4.3A**). TJP1 is not known to be a nuclear protein and, unlike all other candidate proteins, there is no obvious physiologic justification for it to associate with TNRC6A. The implausible ratio suggests this interaction is an artifact.

While stoichiometric ratios calculated by iBAQ should be considered estimates, it is interesting to note that The TNRC6A:AGO2 ratio was 1:2.2. This finding suggests that more than one AGO protein can bind simultaneously to TNRC6A (**Figure 4.3A-B**). This finding would be consistent with the presence of numerous tryptophan residues on the surface of TNRC6A and their potential to interact with the surface of AGO2 (**Figure 4.3C**). AGO3 was present at a 1:1 ratio (**Figure 4.3A-B**). This finding is also consistent with prior mass spectrometry using AGO2 as the bait protein that showed recovery of AGO3 (Kalantari et al., 2016b). Presumably, this previous result was mediated through the potential of AGO3 and AGO2 to bind simultaneously to TNRC6A. CNOT9 was identified at a 1:1 ratio, consistent with its known ability to bind TNRC6A (**Figure 4.3A-B**) (Mathys et al., 2014). Almost all other candidate proteins were present at less than 1:1, reflecting either partial association or loss of interactions during purification.

4.3. Verifying TNRC6A protein interactions

Representative proteins from key functional groups were chosen for closer analysis by immunoprecipitation in turbonuclease-treated extracts (**Figure 4.4**). In this analysis, I used immunoprecipitations with an anti-TNRC6A antibody elicited to a peptide sequence (575-625aa)

within the TNRC6A protein that was not found in TNRC6B or TNRC6C. To cross-validate the findings, I also used antibodies elicited to bind the candidate proteins. The data are summarized in **Figure 4.4G**. The TNRC6A antibody binds two isoforms of TNRC6A and that is verified via knockdown experiments (**Figure 4.5A).** I also used a second anti-TNRC6A antibody, elicited to a different peptide (1225-1275aa) unique to TNRC6A, for immunoprecipitations and confirmed interactions (**Figure 4.5B**).

Cross-validation with antibodies that recognize both potential interacting partners reduces the likelihood that a potential interaction is artefactual. Use of a second anti-TNRC6A antibody also supports the validity of the observed interactions. I chose the candidate proteins for validation by complementary antibodies because they are known core representatives from each functional cluster and because adequate antibodies were available for immunoprecipitation and detection by western analysis. Because adequate antibodies were not always available for all candidates, I do not claim to have verified interactions between TNRC6A and all candidates.

4.4. AGO variants associate with nuclear TNRC6A

Mass spectrometry identified interactions with AGO1, AGO2, and AGO3 in both sample sets. Regardless of whether RNase was added, AGO2 and AGO3 were among the proteins with the highest detected spectral counts and with high ratios for anti-TNRC6A samples versus controls (**Figure 4.2**). The association between AGO proteins and TNRC6A is consistent with the well-known partnership between AGO and TNRC6A/B/C and previous mass spectrometry using AGO2 as bait showing that the TNRC6 paralogs were the primary interacting partners for AGO2 in both cell cytoplasm and nucleus (Kalantari et al., 2016b).

Western analysis revealed that AGO2 was present in HeLa cell nuclei (**Figure 4.4A**). Immunoprecipitation with anti-TNRC6A antibody pulled down AGO2 (**Figure 4.4B-F**). Conversely, immunoprecipitation with anti-AGO2 antibody pulled down TNRC6A. Formation of

an AGO2:TNRC6A complex is significant because AGO2 has the potential to mediate sequence-specific binding to RNA while TNRC6A has the potential to be a scaffold and recruit other protein factors.

4.5. TNRC6A associates with components of the anaphase promoting complex

In RNase-treated samples, I identified interactions with components of the anaphase promoting complex (APC) (**Figure 4.2A-C**). The APC is a ubiquitin ligase that contributes to several cellular pathways by controlling ubiquitin-mediated proteolysis (Chang and Barford, 2014). APC is a multi-protein complex and several subunits appeared to interact with TNRC6A including ANAPC1, 4, 5, and 7. Cell cycle proteins CDC 16, 23, and 27 associate with APC and were also detected. I only detected ANAPC4 in samples only treated with DNase I (**Figure 4.2D-F**), suggesting that the interaction between APC and TNRC6A may be more transient or involves RNA.

APC and associated proteins control ubiquitin mediated proteolysis during the cell cycle and its association with TNRC6A suggests a degradation pathway for nuclear RNAi factor complexes. ANAPC1, the core component of APC, was chosen for further study because an antibody suitable for western analysis was available. ANAPC1 is a predominantly nuclear protein in HeLa cells (**Figure 4.4A**). Anti-TNRC6A antibody immunoprecipitated ANAPC1 while anti-AGO2 antibody did not (**Figure 4.4B**). These data are consistent with TNRC6A occupying a bridging position between AGO2 and ANAPC1. The converse immunoprecipitation was not done because the available anti-ANAPC1 antibody was not adequate.

I also observed the interaction in nuclear extracts that were treated with DNAse I only (**Figure 4.5C**). This finding suggests the reason the interaction wasn't detected in extracts that had been treated with only DNase I prior to mass spectrometry is because they might be more transient and the interaction was lost during purifications.

4.6. TNRC6A associates with proteins that modify histones

Mass spectrometry using turbonuclease-treated samples also identified interactions between TNRC6A and several proteins known to modify histones (**Figure 4.2**). The Mixed Lineage Leukemia (MLL) complex is an H3K4 methyltransferase (Li et al., 2016). I identified MLL complex subunits including MLL3 and MLL4 as well as associated factors including ASH2L, WDR5, RbB5, and AFF4 (Ruthenburg et al., 2006). I focused on WDR5 and RbBP5, both WD40 repeat proteins because they are the core proteins that interact to coordinate H3K4 tri-methylation activity (Dou et al., 2006).

An anti-TNRC6A antibody immunoprecipitated RbBP5 and an anti-RbBP5 antibody immunoprecipitated TNRC6A, but not AGO2 (**Figure 4.4C**). However, an anti-AGO2 antibody also immunoprecipitated RbBP5, so I cannot exclude that the interaction is mediated through AGO2. When an anti-WDR5 antibody was used for immunoprecipitation, I also detected interactions with TNRC6A but not AGO2 (**Figure 4.4C**). An association between WDR5 and RbBP5 was also observed.

I noted the potential for interactions with other proteins known to regulate transcription including histone acetylase NAT10 (Lv et al., 2003), PCF11 (Grzechnik et al., 2015), GATAD2A (Torchy et al., 2015), ZNF24 (Li et al., 2006a), and SWI/SNF proteins ARID1A (Dallas et al., 2000) and SMARCD1 (Phelan et al., 1999). Based on antibody availability I focused on NAT10 and observed that the anti-NAT10 antibody immunoprecipitated TNRC6A, but not AGO2 (**Figure 4.4C**). However, an anti-AGO2 antibody pulled out NAT10. I also observed interactions between NAT10, RbBP5, and WDR5 that had not been noted previously (**Figure 4.4C,G**).

4.7. TNRC6A associates with members of the CCR4-NOT complex

The CCR4-NOT complex plays an important role in cytoplasmic RNAi by associating with TNRC6 proteins (Chekulaeva et al., 2011) and contributing to deadenylation and translational repression (Chen et al., 2014; Mathys et al., 2014). CCR4-NOT has also long been recognized to play a role in transcriptional regulation (Collart, 2016; Kruk et al., 2011). For example, the CCR4-NOT complex has been shown to promote activating histone marks, H3K4 tri-methylation and H3 and H4 acetylation (Laribee et al., 2007; Mulder et al., 2007; Peng et al., 2008).

Mass spectrometry of nuclear lysate after immunoprecipitation with anti-TNCR6A antibody identified multiple components of CCR4-NOT complex including CNOT1, CNOT3, CNOT7, CNOT9, and CNOT10. (**Figure 4.2**). CNOT1 is a scaffolding protein that plays a central role in organizing the CCR4-NOT complex (Maillet et al., 2000). Both CNOT1 and CNOT9 (RQCD1) have been proposed to bind sites on the surface of TNRC6 protein (Fabian et al., 2011; Mathys et al., 2014), so I chose these proteins for further examination.

CNOT1 is predominantly expressed in HeLa cell cytoplasm but can also be readily detected in nuclei, while CNOT9 is primarily nuclear (**Figure 4.4A**). Anti-TNRC6A antibody pulled down CNOT1 and anti-CNOT1 antibody pulled down TNRC6A (**Figure 4.4D**). Anti-AGO2 antibody pulled down CNOT1, but not CNOT9 (**Figure 4.4D, 4.5D**). Immunoprecipitation with anti-CNOT9 antibody recovered TNRC6A but not AGO2 (**Figure 4.4E**).

4.8. TNRC6A associates with members of the mediator complex

The mediator complex is composed of twenty-one proteins, is essential for controlling transcription by RNA polymerase II (Kornberg, 2005), and can participate in either positive or negative regulation (Ansari and Morse, 2013). Regardless of whether turbonuclease was present in samples, I observed the interaction of TNRC6A and several different mediator subunits including MED12, MED13, MED1, and MED14 (**Figure 4.2**). An interaction with MED27 was

observed only after treatment with turbonuclease, and MED4, MED6, MED 17, and MED23 were observed in samples treated only with DNase I. Those interactions may be more transient and not easily detected with my purification methods.

MED14 was chosen as a representative subunit for further analysis for three reasons: 1) it was recently identified as both a structural and functional backbone of the mediator complex (Cevher et al., 2014); 2) it was identified regardless of treatment with turbonuclease; and 3) the anti-MED14 antibody was the most effective anti-mediator antibody available to us. Western analysis confirmed the nuclear localization of MED14 in HeLa cells (**Figure 4.4A**).

The anti-MED14 antibody pulled out TNRC6A but not AGO2 (**Figure 4.4F**). Anti-TNRC6A antibody pulled out MED14, while anti-AGO2 antibody did not. These data suggest that MED14 has a closer association with TNRC6A than AGO2. The anti-MED14 antibody also pulled down CNOT1 and ASH2L, suggesting the potential for additional interactions with proteins that regulate transcription.

4.9. Association of candidate proteins after fractionation

To further characterize nuclear complexes containing RNAi proteins, size exclusion chromatography followed by anion exchange chromatography was used to separate candidate proteins by size and charge (**Figure 4.6**). The samples were treated with RNase inhibitors to increase the potential for RNA-mediated interactions. Size exclusion chromatography of the nuclear extract revealed that TNRC6A was primarily identified in two fractions (**Figure 4.6A**). These fractions also contained AGO2, CNOT1, CNOT9, ANAPC1, NAT10, RbBP5, WDR5, MED14, and Histone H3. It should be noted that AGO2 is found in some fractions that do not contain TNRC6A, suggesting the possibility that it forms complexes independent of TNRC6A. Alternatively however, this result might be due to the AGO2-TNRC6A complex dissociating during purification and chromatography.

Size exclusion fraction C containing TNRC6A was then applied to a Mono-Q anion exchange column. TNRC6A was found primarily in two fractions (**Figure 4.6B**). These fractions also contained CNOT1, CNOT9, NAT10, RbBP5, WDR5, and MED14. While histone H3 had co-purified with TNRC6A after size exclusion, only a minimal amount co-purified after anion exchange. These data from sequential purifications are consistent with the conclusion that TNRC6A exists in complex with the proteins identified as candidate proteins by mass spectrometry and reciprocal immunoprecipitations. However, the dynamics of these complexes has not been studied, so I am not implying they are in one big complex together.

4.10. Other proteins implicated by mass spectrometry

Mass spectrometry suggested interactions between TNRC6A and two paraspeckle proteins, SFPQ and PSPC1. Paraspeckles are nuclear RNA bodies of unknown function that can affect protein localization and association with the RNA NEAT1 (Hirose and Nakagawa, 2012). SFPQ and PSPC1 were among the most strongly detected proteins obtained from samples that had been treated with DNase only (**Figure 4.2D-F**). They were not detected in samples treated with turbonuclease (**Figure 4.2A-C**), suggesting that their inclusion in complexes requires RNA. These data suggest that nuclear RNAi factors at least partially reside in paraspeckles.

I also observed four proteins involved in DNA repair, TP35BP1, CCAR2, and TNKS1BP1, and MDC1 (Francia et al., 2016; Lopez-Saavedra et al., 2016; Xu and Stern, 2003; Zou et al., 2015) in samples treated with Turbonuclease (**Figure 4.2A-C**). This finding is consistent with reports that RNAi factors can bind RNA during the DNA damage response (Francia et al., 2016; Li et al., 2012). Stoichiometric analysis indicated TP53BP1 and CCAR2 had approximately a 1:1.52 and 1:0.88 ratio, respectively, reinforcing the suggestion that these are good candidates for follow-up study (**Figure 4.3A-B**).

4.11. Functional analysis: Impact on RNA-mediated transcriptional activation

Immunoprecipitation confirmed the association between TNRC6A and highly ranked candidate proteins identified from mass spectrometry. To investigate the functional involvement of candidate proteins a colleague in my laboratory, Masayuki Matsui, tested their involvement in RNA-mediated activation of cyclooxygenase 2 (*COX-2*). *COX-2* was chosen because he previously showed that transcriptional activation can be triggered by miR-589 and RNA12nc. miR-589 is a miRNA with two partially complementary binding sites within a noncoding transcript that overlaps the *COX-2* promoter (Matsui et al., 2013). RNA12nc is a "miR-like" duplex with central mismatches designed to disable the cleavage activity of AGO2 but has overall complementary to a noncoding transcript that overlaps the *COX-2* promoter.

The levels of activation are robust, specific (sensitive to just one mismatch within the RNA duplex) and easily detectable (> 20-fold) making *COX-2* a sensitive system for examining expression (Matsui et al., 2013). The previous experiments had demonstrated that AGO2 was required for this process and addition of promoter-targeted RNA led to increased recruitment of RNA polymerase, increased expression of pre-mRNA, mRNA, and protein, and increased activating histone marks (H3K4me3 and H4Ac). As little as one mismatch in the seed sequence of the guide strand of a duplex RNA was sufficient to ablate RNA-mediated activation of gene expression. Antisense oligonucleotides designed to reduce expression of the noncoding RNA target strand also blocked RNA-mediated activation of gene expression.

To determine involvement of the candidate proteins, the first step was to transfect siRNAs into A549 lung cancer cells to block expression of a candidate gene. In a second transfection, promoter-targeted duplex RNA12nc known to activate expression of *COX-2* protein or mRNA (Matsui et al., 2013) was introduced into cells. For all experiments, a control duplex RNA, siGL2, was used that lacks complementarity to genes within the human genome and does not activate *COX-2* expression.

When the duplex RNA complementary to the *COX-2* promoter was added to A549 cells enhanced *COX-2* expression was observed (**Figure 4.7**). When all TNRC6 paralogs were inhibited, activation was not observed (**Figure 4.7A**), confirming functional involvement of TNRC6 in RNA-mediated gene activation. Inhibiting just one TNRC6 paralog was not adequate, suggesting overlapping contributions (Matsui et al., unpublished). Masayuki had previously observed that inhibiting AGO2 also reversed *COX-2* activation by RNA12nc (Matsui et al., 2013), so it is clear that the RNAi machinery is involved in this process.

Masayuki next examined roles for proteins associated with the regulation of transcription. Inhibiting expression of WDR5 reversed RNA-mediated activation of *COX-2* expression (**Figure 4.7B**) (Matsui et al., 2013). Inhibiting expression of NAT10 (**Figure 4.7C**) or MED14 (**Figure 4.7D**) also reversed RNA-mediated gene activation. These data are consistent with the immunoprecipitation data demonstrating association of WDR5, NAT10, and MED14 with TNRC6A (**Figure 4.4**). Chromatin immunoprecipitation demonstrated increased recruitment of WDR5 to the COX-2 promoter after addition of activating RNA (**Figure 4.8**).

By contrast to the reversal of gene expression after knocking down candidate proteins WDR5, NAT10, and MED14, inhibiting expression of the CCR4-NOT complex members CNOT1 (**Figure 4.7E**) or CNOT9 (**Figure 4.7F**) increased levels of COX-2 protein above what they would be from addition of activating duplex RNA alone. These data suggest that the CCR4-NOT complex is not required for RNA mediated transcriptional activation of *COX-2*. The CCR4-NOT complex is known to play a role in RNA-mediated translational silencing (Chekulaeva et al., 2011). Therefore, the increase of COX-2 protein levels may be related to the removal of regulation by a miRNA.

Masayuki also examined the functional role of the anaphase promoting complex component ANAPC1 (**Figure 4.7G**). When expression of the anaphase promoting complex subunit ANAPC1 was knocked down and RNA12nc added, expression of COX-2 protein

increased above the level produced by introducing the activating RNA alone. This result is consistent with the conclusion that anaphase promoting complex contributes to degradation of RNAi factors. However, when ANAPC1 was knocked down in HeLa cells, the levels of AGO2 and TNRC6A remained steady (**Figure 4.9**). The interactions of these factors need to be further examined to determine how ANAPC1 regulates COX-2 activation and why it interacts with TNRC6.

4.12. Summary of nuclear RNAi

The demonstrated ability of small RNAs to control splicing and transcription through nuclear RNAi in mammalian cells (Gao et al., 2014; Kalantari et al., 2016a; Weinberg and Morris, 2016) suggests an unanticipated layer of biological regulation that goes beyond the control exerted by textbook protein transcription and splicing factors.

For the regulation of splicing, nuclear RNAi can promote binding to critical RNA sequences (Kalantari et al., 2016a). The mechanism of action likely involves the blocking of splicing factors and is probably like the mechanism employed by antisense oligonucleotides that affect alternative splicing and that are being used in multiple advanced clinical trials (Havens and Hastings, 2016). For transcriptional control, nuclear RNAi is believed to function through recognition of nascent transcripts rather than direct binding to chromosomal DNA (Weinberg and Morris, 2016). The central mechanistic question has been how recognition of an RNA transcript can trigger a change in mRNA synthesis at a gene promoter. Defining the action of miRNAs, however, is not simple even for the miRNAs that function through post-transcriptional silencing in the cytoplasm – a process that has been the focus of much more research. My goal was to use mass spectrometry to build a better model of the proteins recruited to participate in nuclear RNAi.

4.13. Previous lessons from RNA control of COX-2 activation

In this chapter, I used RNA-mediated activation of COX-2 expression as a reporter of function because aspects of the mechanism had been characterized in detail (Matsui et al., 2013). In that earlier study, duplex RNAs complementary to the COX-2 promoter were shown to activate the expression of COX-2 pre-mRNA, mRNA, and protein. Activation was dependent on seed sequence complementarity, expression of AGO2, and expression of TNRC6. My lab detected a transcript that overlapped the COX-2 promoter. Knock-down of this transcript ablated gene activation and addition of promoter-targeted RNA led to recruitment of AGO2 and TNRC6A to the target transcript. Addition of small RNA led to increased H3K4me3 and H4Ac histone modifications, as well as recruitment of WDR5.

Taken together, this prior work suggested that introduction of a promoter-targeted small RNA into cells could lead to binding of an AGO2-TNRC6-small RNA complex to a promoter transcript, enhanced binding of RNA polymerase, and dramatically upregulated levels of COX-2 expression. Unanswered questions included the identity of other protein partners and the role of TNRC6.

4.14. TNRC6 is the central organizing factor controlling nuclear RNAi

In stark contrast to the narrow group of interacting partners for AGO2 (Kalantari et al., 2016b), using identical mass spectrometry methodology, I discovered many more candidate partners for TNRC6A (**Figure 4.2**). The finding that TNRC6A interacts with more proteins than AGO2 is consistent with the known ability of GW182 family members to act as scaffolding proteins (Yao et al., 2013; Pfaff et al., 2013). The finding that TNRC6A is a more promiscuous bait protein is also consistent with the expectation that detection by mass spectrometry will become weaker as interactions become more indirect. For example, while mass spectrometry using AGO2 as bait may readily detect GW182 family members, the proteins that directly bind the scaffolding protein GW182 may not be detected as candidates for association with AGO2 because the association

is not direct and are less able to survive protein purification and the conditions in the mass spectrometer.

In my previous AGO2 studies (Chapter 3) and present studies, stringent standards for candidate identification (duplicate or triplicate identification) further reduced the likelihood that indirect protein interactions would be identified. I made the decision to exclude lower probability candidates from the list to increase signal to noise focus on top candidates. Therefore, a caveat for the mass spectrometry data is that some members of AGO2/GW182 complexes may not be ranked as “detectable” under the conditions and candidate identification parameters used. Failure to definitively identify an interaction is not proof that the interaction (direct or indirect) does not occur.

Most of the proteins identified in my TNRC6A datasets clustered in six well-defined functional groups. These included RNAi factors (AGO1, AGO2, AGO3), members of the APC (ANAPC1, ANAPC4, ANAPC5, ANAPC7, ANAPC16, CDC16, CDC23, CDC27), members of the mediator complex (MED1, MED4, MED5, MED12, MED13, MED14, MED17, MED23, MED27), members of histone modifying complexes (MLL3, MLL4, WDR5, ASH2L, RbBP5, NCOA6, NAT10), members of the CCR4-NOT complex (CNOT1, CNOT3, CNOT7, CNOT9, CNOT10, CNOT11). The finding that interacting partners cluster into groups that define known protein complexes reinforces the likelihood that that the mass spectrometry is efficiently identifying legitimate protein partners.

Compared to AGO2, TNRC6A had many more experimentally verified interactions (**Figure 4.4G**). This finding is consistent with my mass spectrometry. Immunoprecipitation with anti-TNRC6A antibody revealed associations with AGO2, MED14, CNOT1, CNOT9, ASH2L, NAT10, RbBP5, ANAPC1, and WDR5. AGO2, by contrast, was associated only with TNRC6A, CNOT1, NAT10, and RbBP5, suggesting those interactions may be less direct. All proteins verified by co-immunoprecipitations eluted together in the same fractions after two chromatographic separations

(**Figure 4.6**). The functional data (**Figure 4.7**) further support a role for TNRC6. When all three paralogs are knocked down using RNAi, RNA-mediated activation of *COX-2* gene activation is reversed. Functional assays also support functional roles in gene activation for WDR5, NAT10, and MED14.

My mass spectrometry and immunoprecipitation data, along with the known role of GW182 family members as protein scaffolds, suggest that TNRC6A is a physical bridge between AGO2, the protein responsible for binding guide strand RNA and facilitating recognition of cell RNA targets, and proteins that affect transcription (**Figure 4.4G, Figure 4.10**). TNRC6A (and likely its TRNC6B/C paralogs), not AGO2, appears to be the central organizing factor bringing together AGO2, CCR4-NOT subunits, mediator proteins, and histone modifiers (**Figure 4.4G, Figure 4.10**). In previous studies my lab had shown that the AGO2:RNA guide complex binds to noncoding transcripts that overlap gene promoters (Matsui et al., 2013; Schwartz et al., 2008). My new data add to that model by providing experimental evidence for recruitment of multiprotein complexes with a demonstrated ability to impact gene transcription (**Figure 4.10).**

It should be noted that the model applies only to RNA-modulated control of COX-2, the focus of most of the experimental effort. There are many unanswered questions about how sequence-specific RNA-AGO2-TNRC6 protein complexes might affect other genes or how the composition of the complexes may change from one gene to the next. Also, while RNA-mediated control of gene expression can be a robust phenomenon, the extent to which it plays a role in endogenous gene regulatory pathways remains unclear.

4.15. Conclusion

Nuclear RNAi, regardless of whether it is controlling splicing, transcription, or some other nuclear process, would have distinct advantages as a mechanism for evolution because it would permit sequence specific control of gene expression by miRNAs. The evolution of a miRNA to

control transcription of a specific gene would likely be more straightforward than evolution of a protein. Synthetic RNAs have proven to be robust regulators of gene expression in nuclei and it seems reasonable to hypothesize that that evolutionary pressure should lead cells to use the mechanism. This study identifies proteins that may take part in protein complexes that associate with RNAi factors in cell nuclei to control transcription in conjunction with small RNAs. Many of these protein partners are involved in RNA transcription and provide obvious potential bridges between RNA recognition and the control of gene transcription.

4.16. Overcoming challenges

My biggest challenge working on this project was learning how to purify protein complexes through chromatographic separations. I needed to learn this new technique in a short amount of time to accelerate submission for publication. As challenging as this technique was, I believe the chromatography added significant value to my publication.

The main problems I encountered while optimizing protocols were low protein levels in my samples and loss of protein when concentrating samples to run for western blots. I started by using 800 million cell equivalents of nuclear extract (~40 15cm dishes). When those samples were run through a size exclusion column, the biggest peak obtained was 0.6 OD, which is low for fractionations. However, I still continued with the protocol of combining fractions from the first and second halves of the peak and ran those through the anion exchange column. Before running them on the column, separate samples were concentrated and run on western blots. TNRC6A and AGO2 could be detected in those fractions, but at low amounts. After anion exchange chromatography, the peaks obtained ranged from 0.1-0.3 OD for each fraction. After concentrating and running those fractions on a gel for western, TNRC6A and AGO2 could not be detected. Based on the low amount of protein measured in each fraction, I determined it was best to double the amount of extract used for the experiments. The proteins are spread out with

each separation technique that enough starting protein is needed to still be detected at the end of two separations.

After doubling the amount of starting protein (1.6 billion cell equivalents ~80 15cm dishes), the western blots for size exclusion were easy to obtain **(Figure 4.6A)**. However, TNRC6A and AGO2 were only slightly detected after anion exchange. Initially, I started by using TCA precipitation methods to concentrate the fractions, but after test experiments, it was clear that a majority of my sample was lost with this method. Instead of precipitation for each anion-exchange fraction, I used spin columns to concentrate samples. The use of spin columns significantly increased protein recovery after concentration and made it possible to obtain western blots for each protein of interest **(Figure 4.6B)**.

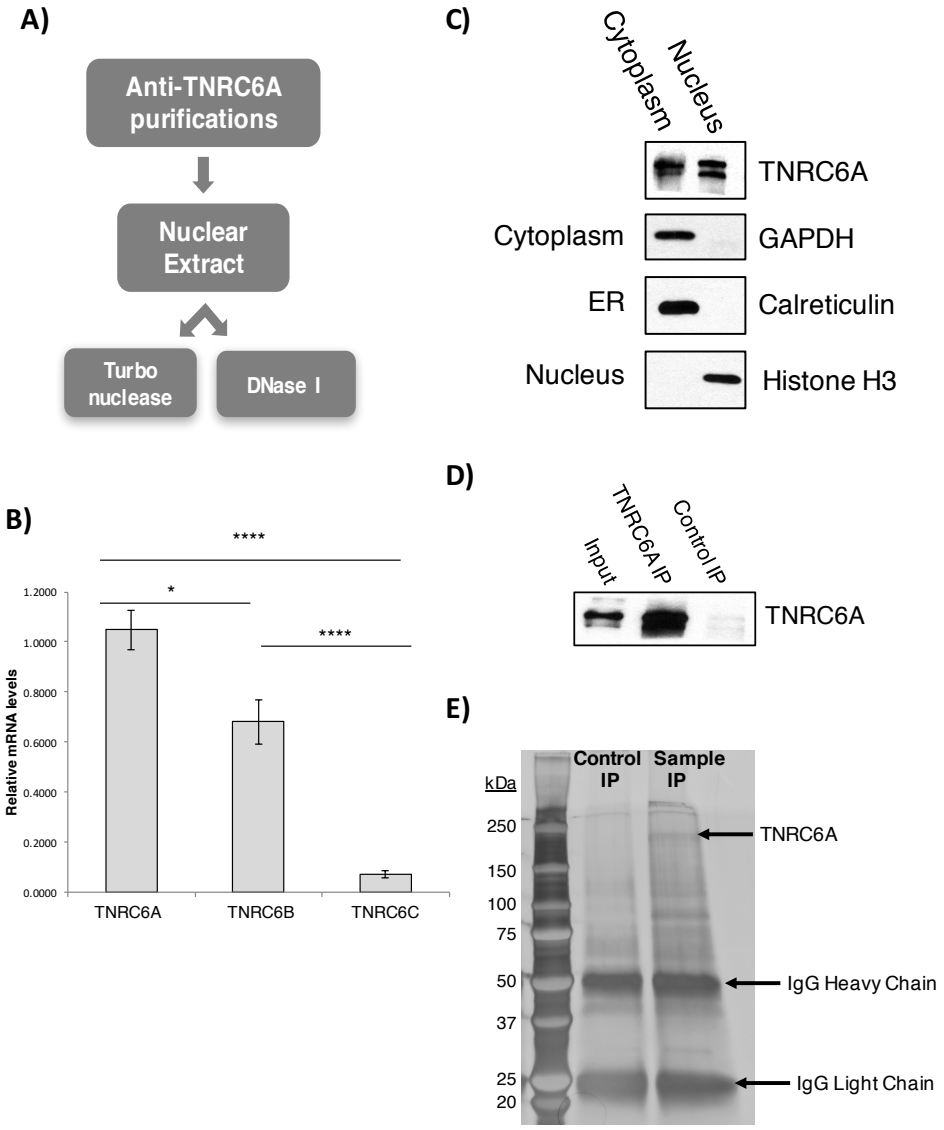
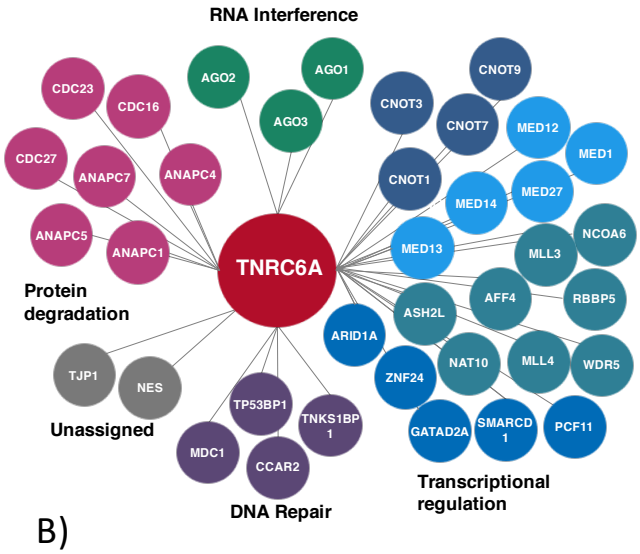


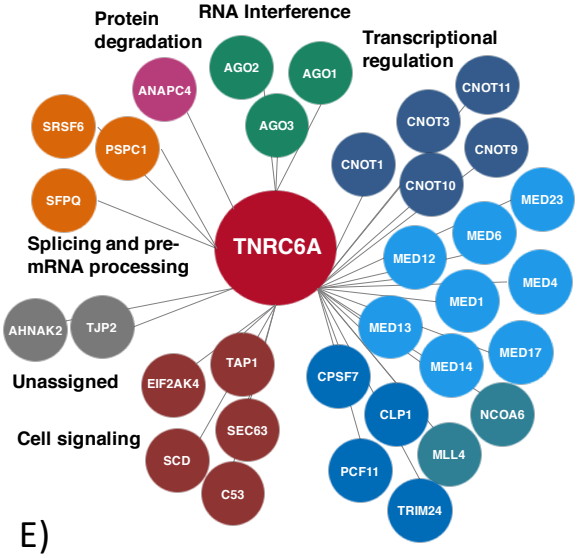
Figure 4.1. Experimental setup for mass spectrometry experiments.

a) Samples from purified nuclei are obtained with and without RNase treatment.
b) qPCR showing relative expression levels of TNRC6A, TNRC6B, and TNRC6C in HeLa cells (n=6). All data are presented as mean ± SE (standard error). *:P< 0.05; ****:P< 0.0001
c) Western analysis showing the purity of representative nuclear lysate that was subsequently used for mass spectrometry.
d) Detection of immunoprecipitated TNRC6A "bait" protein from nuclear lysate by western analysis.
e) Detection of immunoprecipitated TNRC6A "bait" protein after electrophoresis and silver staining.

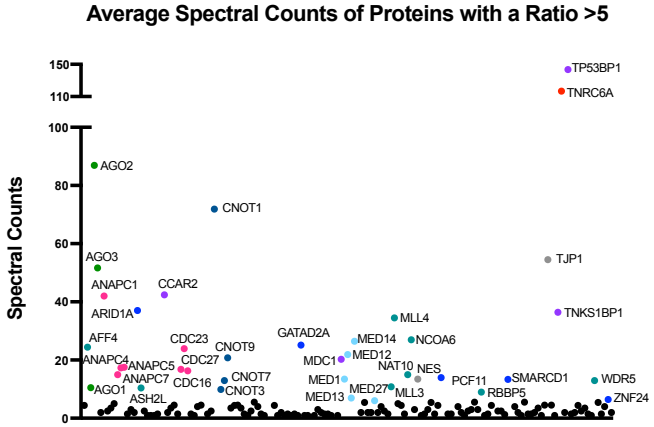
A) Turbonuclease (RNase & DNase)



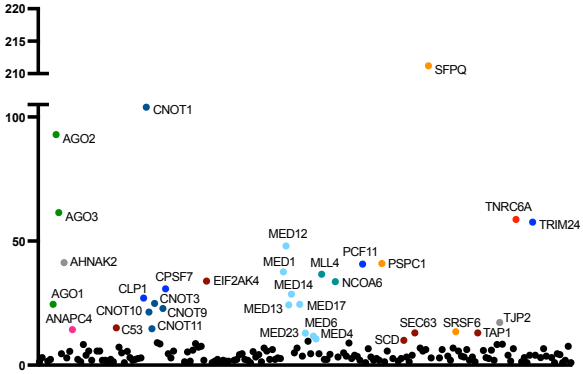
D) DNase Only



B)



E)



C)

Turbonuclease TNRC6A Interactions			
Protein	Average Ratio (Sample/Control	Protein	Average Ratio (Sample/Control
AGO3	Sample Only	CNOT7	102.08
MLL4	Sample Only	CNOT9	69.27
NCOA6	Sample Only	ANAPC7	54.65
CDC23	Sample Only	TJP1	30.84
AFF4	Sample Only	TNKS1BP1	25.74
MED12	Sample Only	TP53BP1	24.47
CDC27	Sample Only	PCF11	24.08
CDC16	Sample Only	ANAPC5	21.44
ANAPC4	Sample Only	ASH2L	20.77
RBBP5	Sample Only	MED1	14.02
CNOT3	Sample Only	ARID1A	13.61
MLL3	Sample Only	CCAR2	11.27
NES	Sample Only	WDR5	8.66
MED13	Sample Only	SMARCD1	8.32
ZNF24	Sample Only	ANAPC1	7.81
AGO1	Sample Only	CNOT1	7.67
MED27	Sample Only	GATAD2A	7.50
AGO2	418.44	MDC1	6.26
TNRC6A	227.19	NAT10	5.70
MED14	167.48		

F)

DNase TNRC6A Interactions			
Protein	Average Ratio (Sample/Control)	Protein	Average Ratio (Sample/Control)
AHNAK2	Sample Only	PCF11	13.86
AGO3	Sample Only	CNOT11	13.76
MED17	Sample Only	MED1	12.05
AGO1	Sample Only	TAP1	11.72
EIF2AK4	6862.78	SEC63	10.83
TNRC6A	1128.06	PSPC1	10.15
AGO2	519.64	CPSF7	9.94
NCOA6	352.61	MED12	9.56
CLP1	226.73	CNOT9	9.43
CNOT3	144.65	TRIM24	9.34
MLL4	59.51	C53	7.89
TJP2	53.97	ANAPC4	7.23
MED13	42.75	SFPQ	6.59
MED6	33.54	MED4	6.11
MED23	23.79	SCD	5.77
CNOT10	18.22	CNOT1	5.49
MED14	15.82	SRSF6	4.54

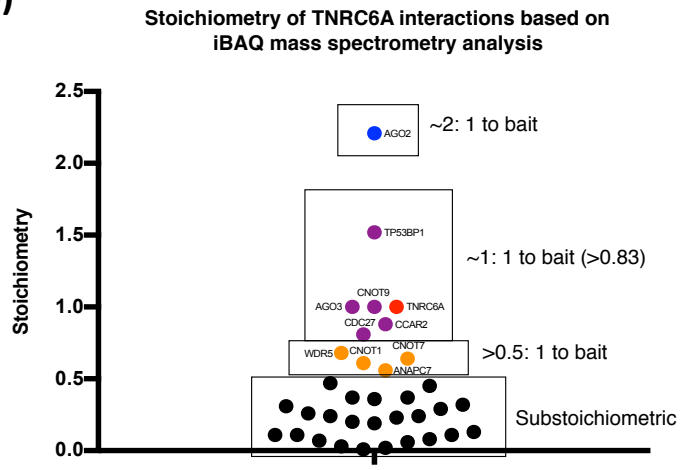
Figure 4.2. Identification of candidate protein partners for TNRC6A. Data was obtained after immunoprecipitation with anti-TNRC6A antibody and mass spectrometry. For parts (A), (B), and (C), samples were treated with turbonuclease, a broad-spectrum nuclease that has both RNase and DNase activity. For parts (D), (E), and (F), turbonuclease was not used to digest RNA, but DNASE I was used to digest chromatin.

a) Bubble plot of candidates. n=2
b) Scatterplot of spectral counts of all proteins with ratios >5 (cutoff). The significant proteins are colored corresponding to their category colors in (A). n=2
c) Sample/control ratios of significant candidate proteins. n=2
d) Bubble plot of candidates. n=3
e) Scatterplot of spectral counts of all proteins with ratios >4 (cutoff). The significant proteins are colored corresponding to their category colors in (D). n=3
f) Sample/control ratios of significant candidate proteins. n=3

A)

Gene names	Average stoichiometry ratio
TNRC6A	1.00
CDC23	0.45
CDC16	0.47
ANAPC4	0.29
ANAPC7	0.56
CDC27	0.81
ANAPC5	0.36
ANAPC1	0.37
NES	0.08
AGO2	2.21
AGO1	0.11
AGO3	1.00
CNOT3	0.32
CNOT1	0.61
CNOT7	0.64
CNOT9	1.00
MED14	0.24
MED12	0.11
MED13	0.03
MED27	0.13
MED1	0.10
NCOA6	0.23
MLL4	0.07
MLL3	0.02
RBBP5	0.37
NAT10	0.19
WDR5	0.68
AFF4	0.26
ASH2L	0.31
PCF11	0.11
ZNF24	0.20
GATAD2A	0.36
SMARCD1	0.21
ARID1A	0.15
MDC1	0.08
TNKS1BP1	0.20
TP53BP1	1.52
CCAR2	0.88
TJP1	15.6

B)



C)

```
  1 mreleakatk dvernlsrdl vqeeeqlmee kkkkkddkkk keaaqkkate qkikvpeqik
 61 psvsqpqpan snngtstats tnnnakrata nnqqpqqqqq qqqpqqqqpq qqpqpqpqqq
121 qpqqqpqalp ryprevpprf rhqehkqllk rgqhfpviaa nlgsavkvln sqsessaltn
181 qqpqnngevq nsknqsdinh stsgshyens qrgpvsstsd sstncknavv sdlsekeawp
241 sapgsdpela secmdadsas ssesserniti masgntggek dglrnstglg sqnkfvvgss
301 snnvghgsst gpwgfshgai istcqvsvda pesksessnn rmnawgtvss ssngglnpst
361 lnsasnhgaw pvlennglal kgpvgsgssg iniqcstigq mpnnqsinsk vsggsthgtw
421 gslqetcese vsgtqkvsfs qqpqnittem tgpnnttnfm tsslpnsgsv qnnelpssnt
481 gawrvstmnh pqmqapsqmn gtslshlsng esksggsygt twgaygsnys gdkcsgpnqq
541 angdtvnatl mqpgvngpmg tnfqvntnkg ggvwesgaan sqstswgsgn gansggsrrg
601 wgtpaqntgt nlpsvewnkl psnqhsndsa ngngktftng wksteeedqg satsqtneqs
661 svwaktggtv esdgstestg rleekgtges qsrdrrkidq htllqsivnr tdldprvlsn
721 sgwgqtpikq ntawdtetsp rgerktdngt eawgssatqt fnsgacidkt spngndtssv
781 sgwgdpkpal rwgdskgsnc qggweddsaa tgmvksnqwg nckeekaawn dsqknkqgwg
841 dgqkssqgws vsasdnwget srnnhwgean kksssggsds drsvsgwnel gktssftwgn
901 ninpnnssgw desskptpsq gwgdppksnq slgwgdsskp vsspdwnkqq divgswgipp
961 atgkppgtgw lggpipapak eeeptgweep spesirrkme iddgtsawgd pskynyknvn
1021 mwnknvpngn srsdqqaqvh qlltpasais nkeassgsgw gepwgepstp attvdngtsa
1081 wgkpidsgps wgepiaaass tstwgsssvg pqalsksgpk smqdgwcgdd mplpgnrptg
1141 weeeedveig mwnsnssqel nsslnwppyt kkmsskglsg kkrrrergmm kggnkqeeaw
```

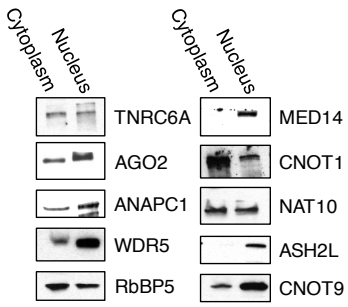
Figure 4.3. Stoichiometry analysis of mass spec samples with iBAQ values.

a) Stoichiometry values relative to TNRC6A (=1) levels. All iBAQ values for control (IgG) were subtracted from sample values. Turbonuclease samples were used in this case.

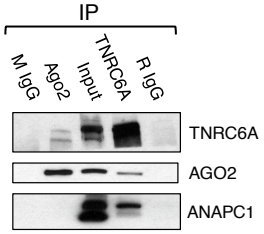
b) Histogram representing the stoichiometry ratios. TJP1 (=15.6) is excluded from this graph.

c) TNRC6A (isoform 1) N terminal AGO binding domain. Tryptophan (W) and Glycine-Tryptophan (GW) repeats are highlighted to indicate potential AGO binding sites. There are more GW repeats within the C terminal domain (not shown).

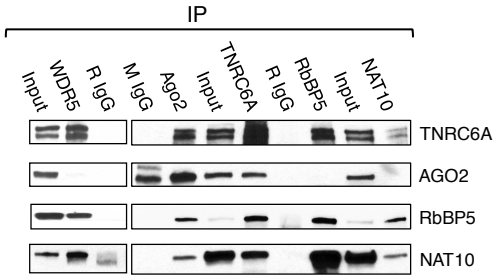
A)

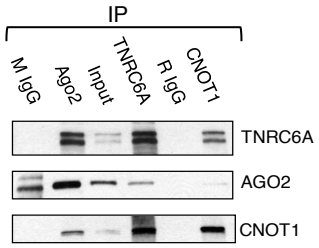


B) Interactions of the APC Complex

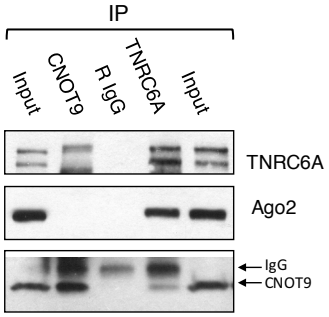


C) Interactions of Histone Modifiers

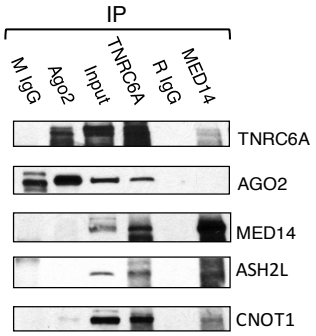




E) Interactions of the CCR4-NOT Complex



F) Interactions of the Mediator Complex



G) Summary of Co-Immunoprecipitations

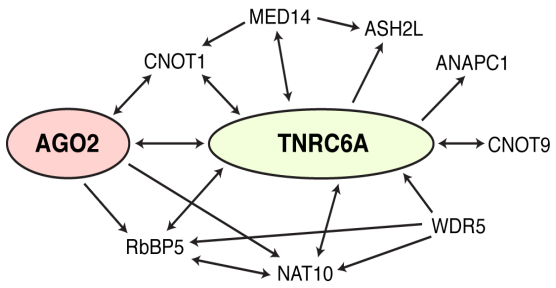


Figure 4.4. Validating the interactions of candidate proteins in Turbonuclease-treated nuclear extracts. Western analysis of candidate proteins after immunoprecipitations. These immunoprecipitations were performed in nuclear extracts that had been treated with turbonuclease to degrade RNA and chromatin. The bands in the Mouse IgG lane present in the AGO2 blots in C,D, and F are not the size of AGO2 and are caused by impurities in the Mouse IgG.

a) Distribution of candidate proteins between cytoplasm and purified nuclei. Co-immunoprecipitations representing
b) Interactions of ANAPC1, a member of the APC
c) Interactions of histone modifier proteins WDR5, RbBP5, and NAT10,
d) Interactions CCR4-NOT protein CNOT1
e) Interactions CCR4-NOT protein CNOT9
f) Interactions of MED14, a member of the mediator complex
g) Summary of Co-immunoprecipitation experiments.

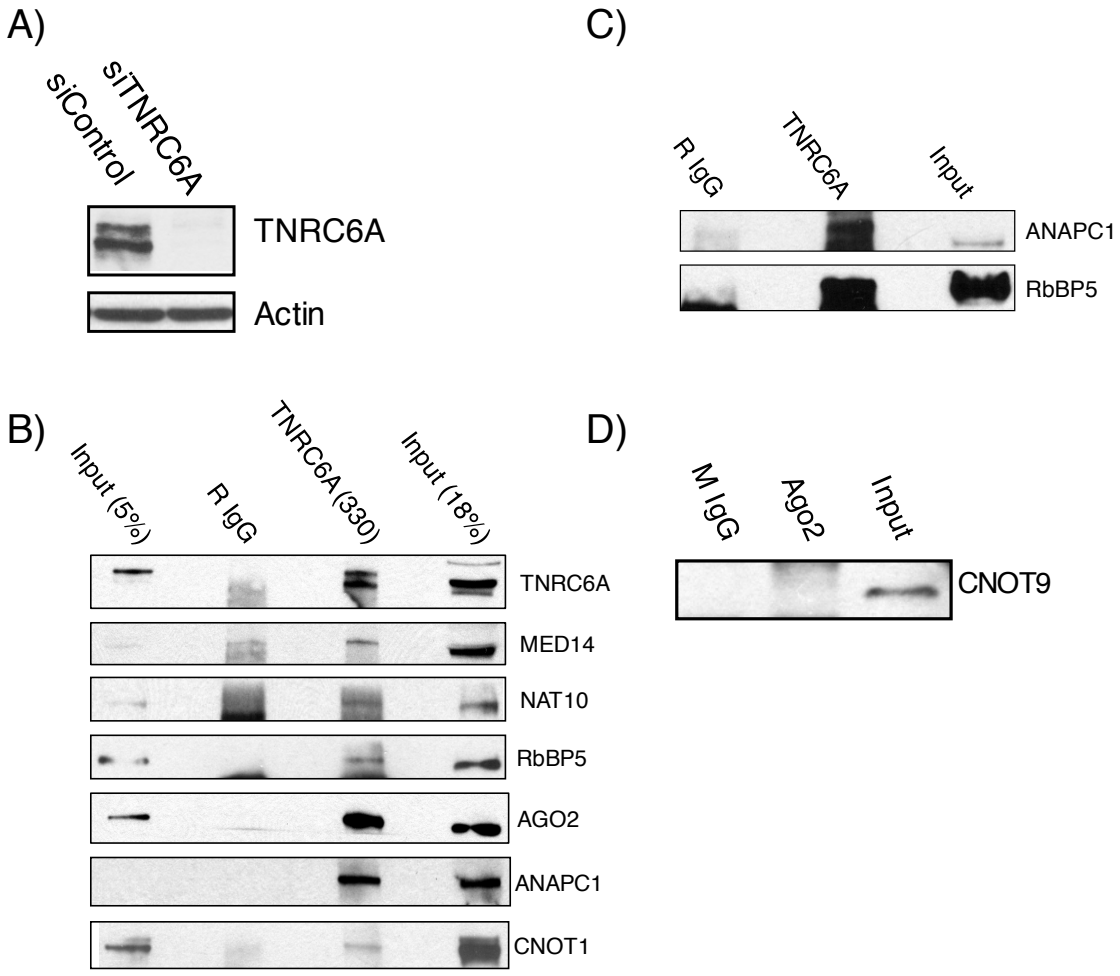


Figure 4.5. Supporting data for candidate protein verification.

a) Western blot of siRNA knockdown of TNRC6A (25nM concentration used).
b) Co-Immunoprecipitations in nuclear Turbonuclease extracts for candidate proteins using another TNRC6A antibody (Bethyl 302-330A) to verify interactions. The antibody does not IP as efficiently as the main TNRC6A antibody used. The inputs represent percentages of IP concentration.
c) Co-Immunoprecipitations in nuclear DNase-treated extracts for proteins that were not detected via mass spectrometry. RbBP5 is very abundant in nuclei and readily detected, so the input is usually overexposed compared to the IP lane.
d) Co-immunoprecipitation in nuclear Turbonuclease extract using the anti-AGO2 antibody and blotting for CNOT9.

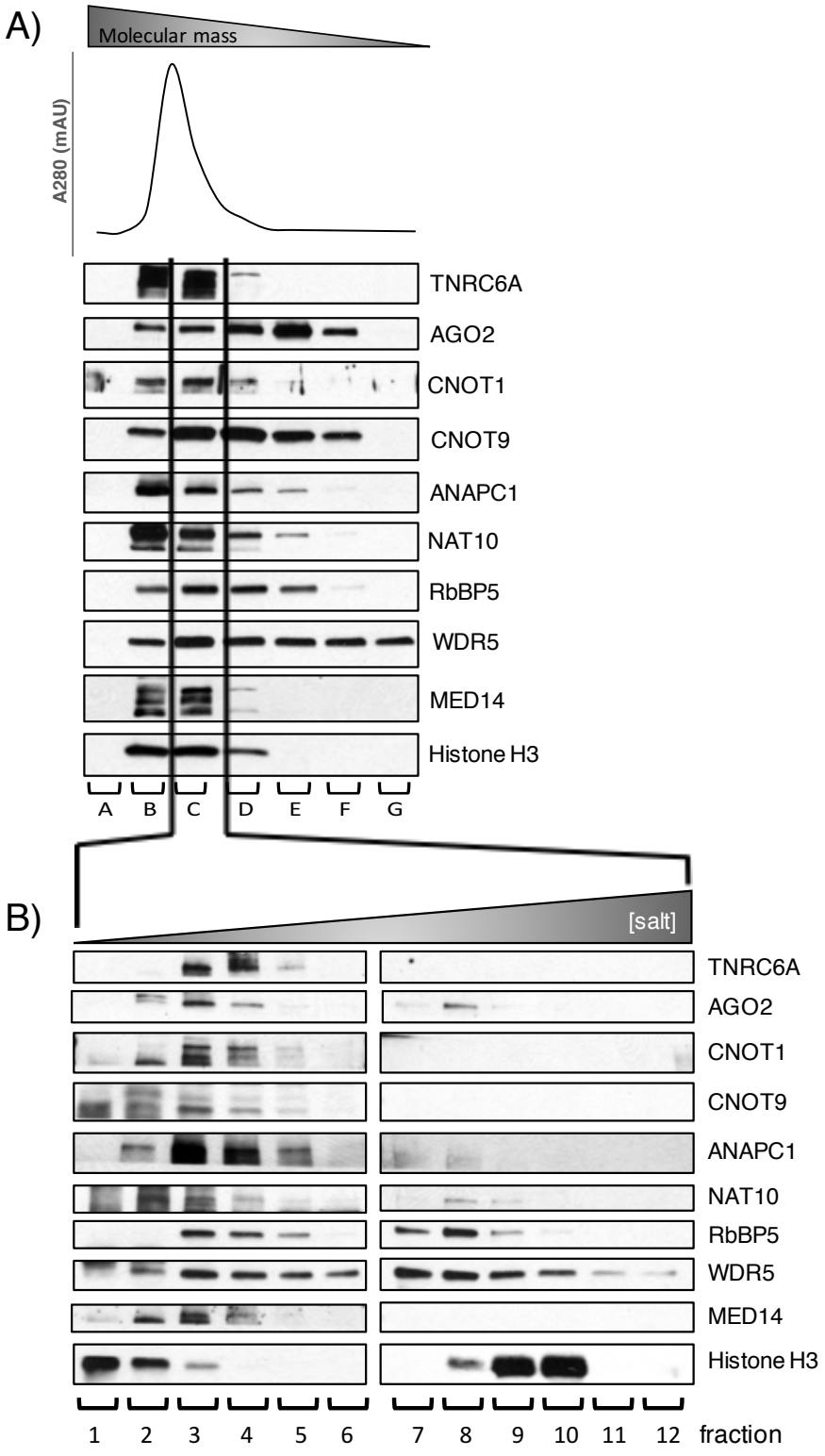

A)
Molecular mass
A280 (mAU)
TNRC6A
AGO2
CNOT1
CNOT9
ANAPC1
NAT10
RbBP5
WDR5
MED14
Histone H3
A B C D E F G
B)
[salt]
TNRC6A
AGO2
CNOT1
CNOT9
ANAPC1
NAT10
RbBP5
WDR5
MED14
Histone H3
1 2 3 4 5 6 7 8 9 10 11 12 fraction

Figure 4.6. Sequential size exclusion and anion exchange chromatography to detect potential complexes with TNRC6A. All FPLC was performed with nuclear extracts not treated with RNase.

a) Size exclusion chromatography fractions with decreasing size. Fraction C was used for subsequent fractionations.
b) Anion exchange chromatography of size exclusion fraction C with increasing salt concentration. Western blots were performed in each fraction for candidate proteins.

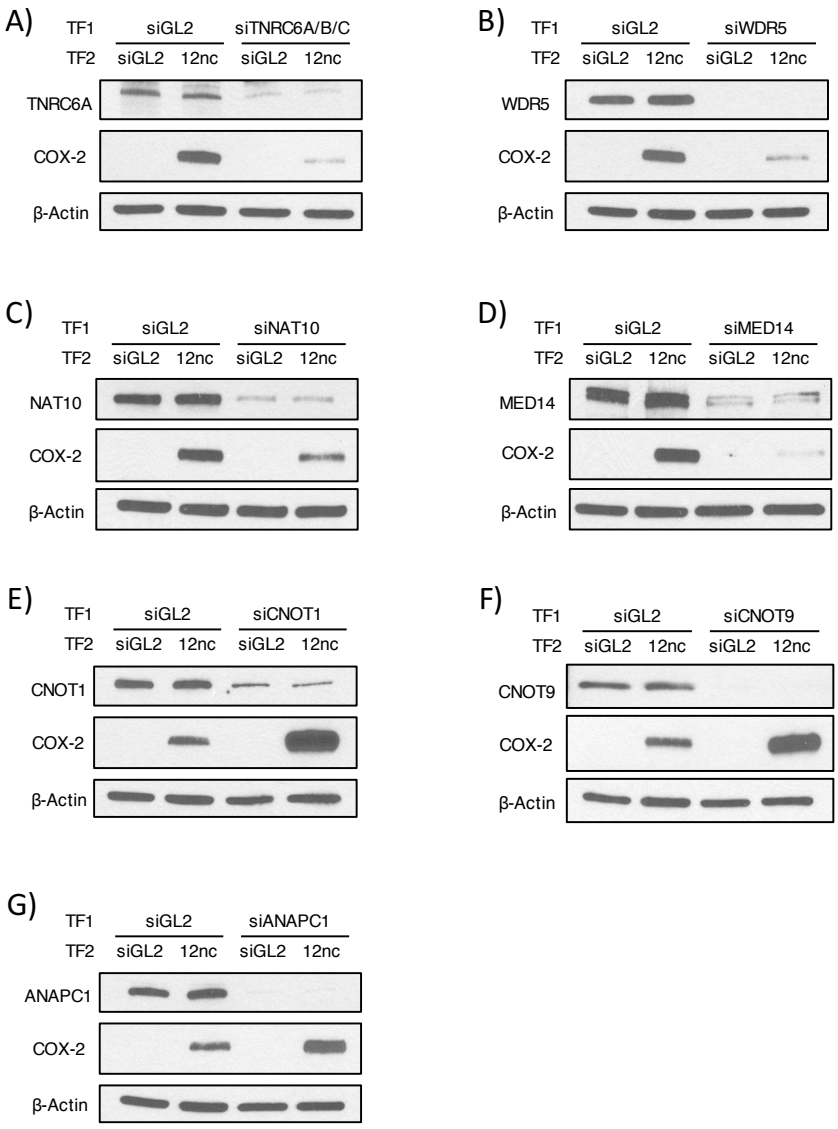


Figure 4.7. Functional roles for candidate proteins during RNA-mediated activation of *COX-2* expression. Western analysis of COX-2 protein expression in presence or after depletion of candidate proteins followed by treatment with siGL2 or RNA12nc. 25 nM duplex siRNA was used in all experiments. siGL2 is a control duplex RNA with minimal complementarity to other sequences within the human genome. TF1 = added during first transfection. TF2 = added during second transfection. Each represents effect of using duplex RNAs to deplete:

a) TNRC6A/B/C
b) WDR5
c) NAT10
d) MED14
e) CNOT1
f) CNOT9
g) ANAPC1

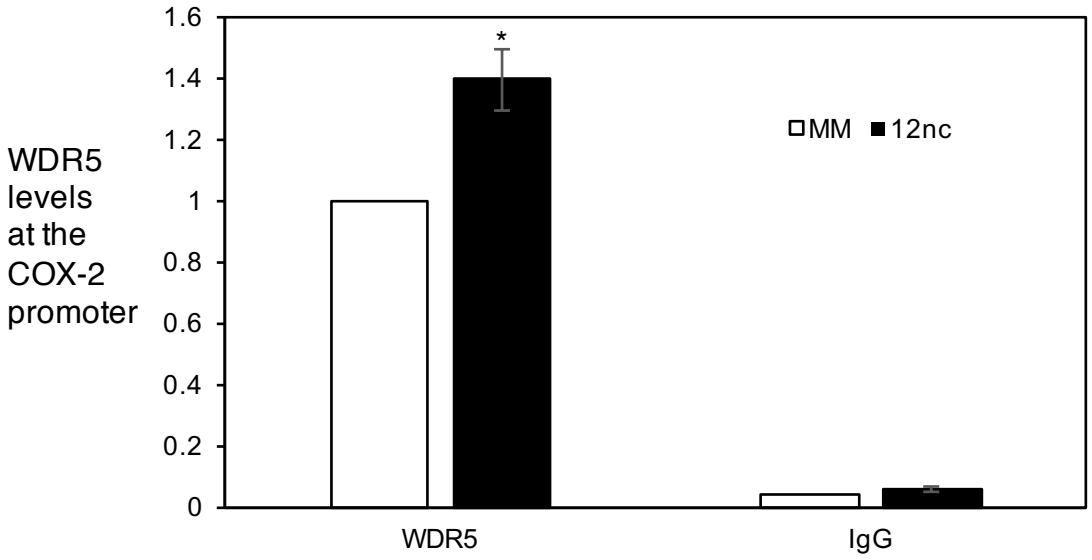


Figure 4.8. ChIP-qPCR showing enrichment of WDR5 at the COX-2 promoter upon addition of RNA12nc (25 nM).

PCR primers target the promoter region from -81 to +56, relative to COX-2 TSS. A commercial primer pair (Active Motif) targeting a gene desert on chr12 was used for data normalization. MM=mismatch control RNA duplex. Error bars are SD. *$p<0.05$ (t-test) relative to mismatch control. All experiments were done in A549 cells (n=3).

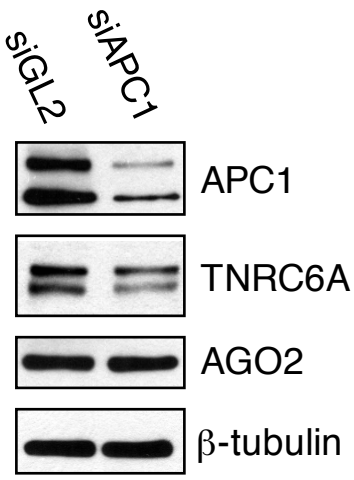


Figure 4.9. Effects of Knockdown of ANAPC1 in HeLa Cells.
siRNAs targeting GL2 (control) and ANAPC1 (APC1) were transfected at 25μm. Western blots were run with extracts of equal concentrations and APC1, TNRC6A and AGO2 were blotted for to determine levels after knockdown.

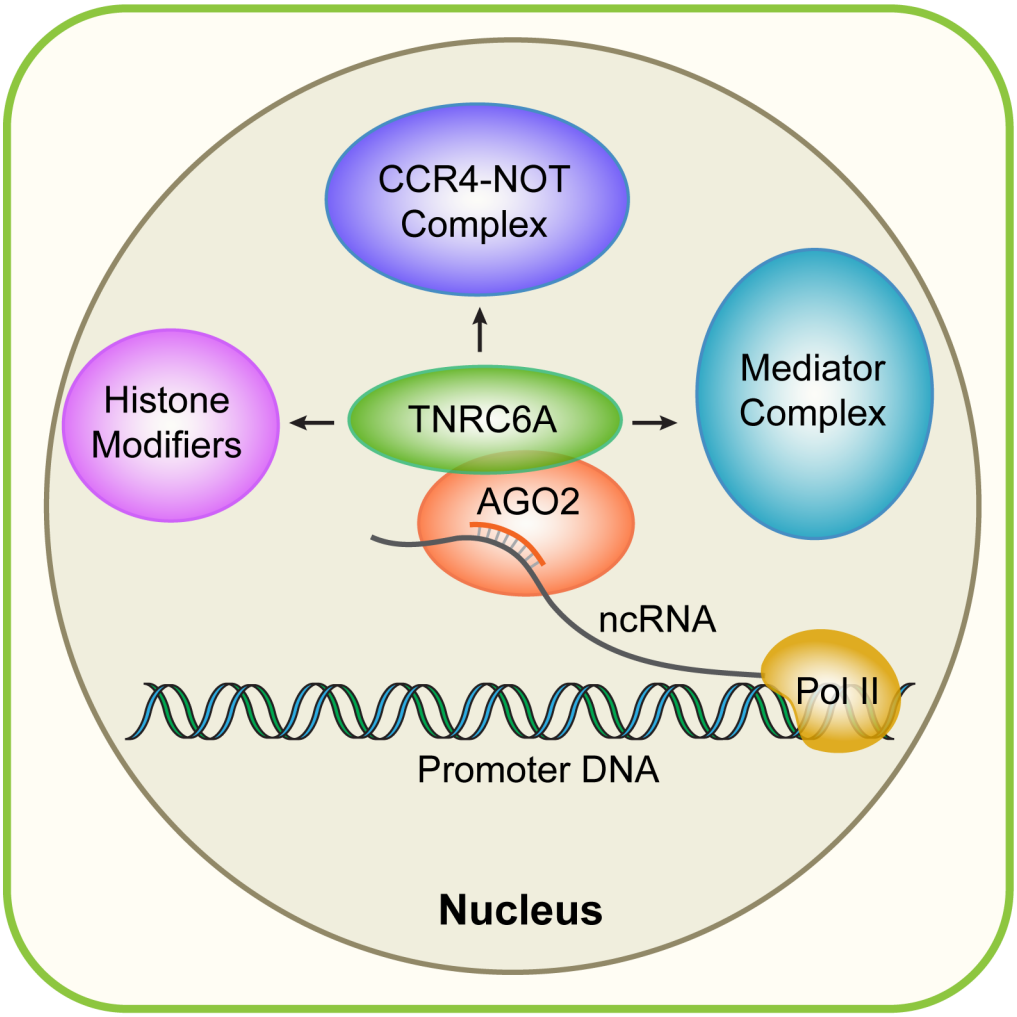


Figure 4.10. Model for TNRC6A interactions.
TNRC6A is bound to AGO2 and can recruit other proteins to affect gene transcription. TNRC6A, not AGO2, is the central organizing factor for nuclear RNAi and the hub for most protein interactions. Small RNA:AGO2 complexes bind sequence specifically to a noncoding transcript at a gene promoter. TNRC6A binds to AGO2 and organizes complexes that have the potential to contain mediator complex, histone modifier proteins, and the CCR4-NOT complex. This model shows potential interactions and should not be taken to imply that the interaction must occur simultaneously or that multiple different complexes with differing protein compositions may exist.

Chapter 5
Methods

5.1. Cell culture

HeLa cells (ATCC) were cultured in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 5% fetal bovine serum (FBS) and 0.5% non-essential amino acids (NEAA). A549 cells (ATCC) were cultured in F-12K media supplemented with 10% FBS. T47D cells (ATCC) were cultured in RPMI medium supplemented with 10% FBS, 0.5% NEAA, 20 μg/mL insulin, 10 mM pH 7.0-7.6 HEPES, and 1 mM sodium pyruvate. T47D cells stably expressing FLAG-HA-tagged AGO2 were cultured identically to T47D cells but media supplemented with 0.2 mg/mL G418. All cells were grown at 37°C in 5% CO2.

The FLAG-HA-tagged AGO2 stable cell line was developed by transfecting T47D cells with the (corrected) pIRESnew-FLAG/HA Ago2 plasmid (Addgene,10822) and Lipofectamine 2000 (Life Technologies, 11668019). Cells were treated with 500 μg/mL G418 for two weeks. Colonies were picked with 5% trypsin and transferred to 24-well dishes. Colonies were tested for expression of FLAG-HA AGO2 by western blot with anti-HA and anti-AGO2 antibodies. Colony labeled 'Colony 1' was used for this study.

5.2. Extract preparation

Purification of cytoplasmic and nuclear extract for mass spectrometry, Co-IP and western blotting from HeLa or T47D cells was performed as described previously (Gagnon et al., 2014) with a few modifications. Cytoplasmic extract was collected by lysing cells with Hypotonic Lysis (HL) Buffer (10mM Tris, pH 7.4, 10mM NaCl, 3mM $MgCl_2$, 0.3% NP-40) at 1mL/150mm dish. The nuclear pellet was spun down at 500xg for 5 minutes. The supernatant was saved as the cytoplasmic extract and 140mM KCl and 10% glycerol were added. The nuclear pellet was washed with 10mL

cold Hypotonic Lysis buffer 3X for 5 minutes. Nuclear extract was collected by adding Nuclear Lysis (NL) Buffer (20mM Tris, pH 7.4, 150mM KCl, 3mM $MgCl_2$, 0.5% NP-40, 10% Glycerol) at a concentration of 0.5mL/150mm dish. A 27-gauge needle was used to lyse the nuclei in addition to NL buffer. For all extracts we either used Turbonuclease (general RNase and DNase) or DNaseI to break up the chromatin to pull off any potential protein complexes. For "Turbonuclease" extracts, Turbonuclease was added at a 1:10000 dilution (Accelagen, N0103M) to the extract and incubated for 30min at room temperature. For "DNase" treated nuclear extracts, DNaseI was added at a 1:20 dilution (Worthington, LS006353) and incubated at 37°C for 30min. Extract was spun down at 1000xg for 5 minutes and supernatant was flash frozen and stored for analysis.

For chromatography experiments, nuclear extracts were prepared differently to preserve RNA-bridged protein interactions and effectively digest the chromatin to pull off protein complexes. After adding Nulear Lysis Buffer (NLB), nuclear extracts were sonicated at 20% power for 20 seconds each 3X. 20 units of RNase Inhibitors (Promega) were added and incubated with extract at 37^{o}C for 30minutes. Extract was spun down at 1000xg for 5 minutes and supernatant was flash frozen and stored for analysis.

5.3. Western blots and extract analysis

Western blots to determine the purity of nuclear and cytoplasmic fractions from HeLa and T47D cells were performed as before (Gagnon et al., 2014). For comparing nuclear and cytoplasmic fractions by Western blot, the same cell equivalents of extract were separated by electrophoresis (1/2 the volume of nuclear extract per 1 volume of cytoplasmic). Blocked Western blot membranes (Hybond-C Extra, GE Healthcare Life Sciences) were incubated with the following primary antibodies over night at 4^{o}C in PBST (PBS + 0.05% TWEEN-20) + 5% milk with rocking: anti-TNRC6A at 1:4000 (Bethyl, A302-329A), anti-Calreticulin at 1:1000 (Cell Signalling, 2891S), anti-

Histone H3 at 1:10000 (Abcam, ab1791), anti-GAPDH at 1:600 (Abcam, ab9484), anti-Ago2 at 1:1000 (Abcam, ab57113).

Western blots of immunoprecipitation and chromatography samples were performed as described above. For each gel, we focused on performing western blots for various sizes of proteins. For the TNRC6A project, the identity of key proteins determined how long the gel is run and the two TNRC6A isoform bands may not be resolved consistently. The following primary antibodies were used for immunoprecipitation analysis: TNRC6A (Bethyl, A302-329A); CNOT1 (Proteintech, 14276-1-AP); NAT10 (Proteintech, 13365-1-AP); ASH2L (Cell Signaling, 5019); ANAPC1 (Bethyl, A301-653A); MED14 (Abcam, ab72141); RQCD1/CNOT9 (Proteintech, 22503-1-AP); RbBP5 (Abcam, ab52084); WDR5 (Cell signaling, 13105); AGO2 western (WAKO, 018-22021). Depending on the size of the protein or conditions, different peroxidase-linked secondary antibodies were used: Anti-Mouse IgG (Jackson immuno research), Anti-Rabbit IgG (Jackson immuno research), Anti-Rabbit IgG Heavy Chain (Abcam ab77902), Anti-Rabbit IgG Light Chain (Abcam, ab99697).

For TNRC6 paralog expression analysis, RNA was extracted from HeLa cells using the Trizol method. qPCR was performed. For each gene, 2 primer sets targeting two different regions unique to the specific gene were used, with 6 biological replicates (n=6) for each primer set. The relative expression level for each gene was calculated based on the average of the two primer sets for each gene, with 6 replicates for each primer set.

5.4. Immunoprecipitation for mass spectrometry analysis

FLAG conjugated resin (Sigma, A2220) was incubated with extract from FLAG-Ago2 stable line cells or incubated with extract from T47D cells lacking FLAG-Ago2 as the control. For immunoprecipitations of endogenous Ago2 in T47D extracts, either anti-Ago2 (Abcam, ab57113) or mouse IgG antibody (Millipore, 12-371) were cross-linked to Protein A Plus/ Protein G resin.

For immunoprecipitations of endogenous TNRC6A in HeLa extracts, either anti-TNRC6A (Bethyl, A302-329A) or Rabbit IgG antibody (Millipore, 12-370) were cross-linked to Protein A Plus/ Protein G resin (Millipore, IP05). Antibody was incubated with resin (65 μg/mL resin) in coupling buffer (10mM Sodium Phosphate, pH 7, 75mM NaCl) with rocking at room temperature for 2hrs. Resin was washed 3X with coupling buffer. 0.4mM DSS was added to resin and incubated with rocking for 30min at room temperature. The reaction was quenched with Tris-HCl, pH 8.0 and resin was resuspended in IP-EQ buffer (50mM Tris-HCl, pH 7.4, 2mM $MgCl_2$, 150mM KCl, 0.05% NP-40) in a 33% mixture for use in experiments.

For each extract, a starting protein amount of 22mg was pre-cleared using Rabbit IgG conjugated Agarose (Sigma, A2909), rocking at 4^{o}C for 30 minutes. Extracts were transferred to 1mL of specific-antibody bound resin mixture and incubated over night at 4^{o}C with rocking. Resin was washed 6X with IP_{300} Buffer (20mM Tris-HCl, pH 7.4, 2mM $MgCl_2$, 300mM NaCl, 0.05% NP-40) and proteins were eluted with 300 μl of 1X SDS loading buffer. To eliminate excess salt, elutions were run through Illustra Nap-5 desalting columns (GE Healthcare, 17-0853-02). 200ul H_2O was added to samples before desalting for a final volume of 500ul. Samples were eluted from Nap-5 columns with 1mL H_2O. To concentrate samples, elutions were run through Amicon Ultra-0.5mL centrifugal protein concentrators (Millipore, UFC901024). Elutions were run 15mm into SDS-PAGE gels, Coomassie stained with GelCode Blue (Thermo, 24592), and each lane was cut into three 5mm slices. The gel slices were submitted to the UT Southwestern Proteomics Core for further analysis.

5.5. Affinity chromatography for AGO2 PTM analysis

Empty poly-prep chromatography columns (Bio-rad #7311550) were packed with 2ml of antibody conjugated beads. For FLAG-AGO2 extracts, 2ml of FLAG-resin (Sigma, A2220) was used. For endogenous AGO2 extracts, 2ml of resin conjugated to Abcam AGO2 antibody (described above)

was packed into the column. Resin were washed 2X with IP_{EQ} buffer (50mM Tris-HCl, pH 7.4, 2mM $MgCl_2$, 150mM KCl, 0.05% NP-40).

Nuclear extracts from FLAG-AGO2 T47D cells and endogenous HeLa cells were prepared as described above. At 4°C, nuclear extracts were loaded into the columns and run through via gravity flow. Flow through was collected and run through the column a second time to optimize the amount of AGO2 bounds to beads. After extract flow through, IP_{300} Buffer (20mM Tris-HCl, pH 7.4, 2mM $MgCl_2$, 300mM NaCl, 0.05% NP-40) was used to wash the resin 3X via gravity flow. After washes, bound AGO2 and complexes was further purified with 2ml of a cold 1X SDS buffer wash. After cold SDS buffer wash, elutions were performed with 2ml of 95°C 1X SDS buffer. The elution was collected, desalted and concentrated as described above. Samples were run on an SDS-PAGE gel and bands were excised for trypsin digest, HPLC and mass spectrometry analysis.

5.6. Mass spectrometry data analysis

Gel bands were digested overnight with trypsin (Promega) following reduction and alkylation with DTT and iodoacetamide (Sigma–Aldrich). Following solid-phase extraction cleanup with Oasis HLB plates (Waters), the resulting samples were analyzed by LC/MS/MS using an Orbitrap Elite, Q Exactive, or Q Exactive Plus mass spectrometer (Thermo Electron) coupled to an Ultimate 3000 RSLC-Nano liquid chromatography system (Dionex). Samples were injected onto either a 180μm i.d., 15-cm long column packed in-house with a reverse-phase material ReproSil-Pur C18-AQ, 1.9μm resin (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany), or a 75μm i.d., 50-cm long Easy Spray column (Thermo). Peptides were eluted with a gradient from 1-28% buffer B over 40 min (180 μm column) or 60 min (75 μm column). Buffer A contained 2% (v/v) acetonitrile (ACN) and 0.1% formic acid in water, and buffer B contained 80% (v/v) ACN, 10% (v/v)

trifluoroethanol, and 0.08% formic acid in water. The mass spectrometer acquired up to 20 MS/MS spectra for each full spectrum acquired.

Raw MS data files were converted to a peak list format and analyzed using the central proteomics facilities pipeline (CPFP), version 2.0.3 (Trudgian and Mirzaei, 2012; Trudgian et al., 2010). Peptide identification was performed using the X!Tandem (Craig and Beavis, 2004) and open MS search algorithm (OMSSA) (Geer et al., 2004) search engines against the human protein database from Uniprot, with common contaminants and reversed decoy sequences appended (Elias and Gygi, 2007). Fragment and precursor tolerances of 20 ppm and 0.5 Da were specified, and three miscleavages were allowed. Carbamidomethylation of Cys was set as a fixed modification and oxidation of Met was set as a variable modification.

5.7. Mass spectrometry data analysis for AGO2 and TNRC6A protein complexes.

Each sample was run in duplicated or triplicate. Label-free quantitation of proteins across samples was performed using SINQ normalized spectral index Software (Trudgian et al., 2011b). SINQ is a semi-quantitative mass spec analysis that uses spectral counts (peptide abundance) and ratios of sample vs. control cutoffs to identify proteins in a complex mixture. The standard spectral count cutoff for SINQ analysis is 3 or higher, however, we used a cutoff of at least 4 spectral counts in all samples. A second layer of analysis used was through the ratios of sample to control. For this, a cutoff of at least 4:1 was set for all replicates. For proteins to be considered significant in each treatment, they had to appear above the cutoffs set for all replicates. Raw datasets can be found MassIVE database (ID:MSV000081227).

5.8. Mass spectrometry data analysis for PTMs of AGO2 and TNRC6A

Each sample was run in duplicated or triplicate. Mass spectrometry datasets were analyzed in two ways to yield the broadest PTM identification. For the first analysis, each peptide spectra

was run through an algorithm called ModLS (Modification Localization Scoring) to assess the confidence of the PTM assignments in both the identification and localization of the correct residue harboring the PTM (Baker et al., 2011; Chalkley and Clauser, 2012; Sunyaev et al., 2003). In addition to the analysis for specific PTMs in the standard analysis pipeline, an unrestricted PTM search was performed using the "Error tolerant" method in the Mascot software (Mann and Wilm, 1994; Yu et al., 2016). This method considers the possibility that any modification defined in the Unimod.org database may be present. This is a sensitive procedure that reports many putative modifications, however it is not very specific (i.e., artefactual modifications from sample prep). The proteomics core staff searches through the results manually and identifies the most biologically relevant PTMs in a report. Together, these two methods produce highly confident results for the identification of PTMs on a protein of interest.

5.9. IBAQ analysis for stoichiometry

For turbonuclease TNRC6A mass spectrometry samples, the proteomics core re-analyzed the raw data using the MaxQuant software package to obtain the intensity-based absolute quantification (iBAQ) values for each significant protein. The iBAQ intensities accurately represent the relative abundance of all proteins identified in a sample. I subtracted the background iBAQ values (IgG control) from each TNRC6A IP sample value and used them to calculate the stoichiometry of each protein interaction relative to the bait protein, TNRC6A. This analysis was used in Vermeulen et al., 2013, *NAR.*

5.10. Co-immunoprecipitations

For co-immunoprecipitations, 300 μg of starting protein was used per sample. Extract was precleared with rabbit or mouse IgG conjugated resin (30 μL of 50% slurry) for 30min at 4^0C. For FLAG IPs, 30 μL of a 50% slurry of FLAG conjugated resin was added to the extract after preclear

and incubated for 2.5-3hr. Samples were then incubated with 3 μg of antibody overnight at 4^{o}C. 45 μl of Protein A plus/ Protein G resin (Millipore, IP05) was added and incubated 1-2 hrs. Samples were washed 4X with IP Wash Buffer (20mM Tris-HCl, pH 7.4, 3mM $MgCl_2$, 400mM NaCl, 0.05% NP-40, 0.01%SDS) with 5min incubations and eluted with 75^{o}C 1X SDS for 5 minutes. Samples were loaded onto 4-20% SDS-PAGE gels and Western blots were performed. Antibodies used: GW182 (Bethyl, A302-329A); CNOT1 (Proteintech, 14276-1-AP); NAT10 (Proteintech, 13365-1-AP); ASH2L (Cell Signaling, 5019); ANAPC1 (Bethyl, A301-653A); MED14 (Abcam, ab72141); RQCD1/CNOT9 (Proteintech, 22503-1-AP); RbBP5 (Abcam, ab52084); WDR5 (Cell signaling, 13105); AGO2 IP (Abcam, ab57113); AGO1 (Cell Signaling, D84G10); LGALS1 (Cell Signaling, 5418S); SMARCC1 (Abcam,ab72502).

5.11. Chromatographic separations

For both Size-exclusion (SE) and anion-exchange (AE) chromatography, extracts were prepared as described above with a few changes. The extracts were sonicated at 20% power for three 20 second pulses and subsequently treated with 200 units of SUPERase-In (Ambion) for 1hour at room temperature. This was to more efficiently pull off protein complexes associated with chromatin and to preserve interactions that include RNA. Forty-one dishes (~820 million cell equivalents) were used and concentrated down to 2mL for SE.

For SE chromatography, extracts were filtered and injected onto a Superdex200 HiLoad 16/60 column (Amersham Pharmacia) pre-equilibrated with FPLC buffer (20mM Tris, pH 7.4, 150mM NaCl, 3mM $MgCl_2$, 5% glycerol), and separated by FPLC. Protein was eluted off the column with FPLC buffer and collected at 4^{o}C. Eluted fractions were assayed by western blot. For subsequent tandem fractionation by anion exchange (AE), size-exclusion fractions were diluted and concentrated, treated again with SUPERase-In, and injected into a Mono-Q FPLC column (Amersham Pharmacia) equilibrated with FPLC buffer at 0.1M NaCl. Elution was performed at

4°C with a linear gradient from 0.1M to 1M NaCl FPLC buffer. Fractions were collected and snap-frozen in liquid nitrogen. Western blot analysis was performed as described above.

5.12. Double transfection assays.

siRNAs (siTNRC6ABC, siWDR5, siCNOT1, siANAPC1, siASH2L, siNAT10, siMED14, siCNOT9, and control siRNA (siGL2) were reverse transfected into A549 cells at 25 nM using Lipofectamine RNAiMAX. Two days after transfection, cells were trypsinized, split at the ratio of 1:2 in another six-well plate, and reverse transfected with siCtrl or RNA12nc at 25 nM. Cells were harvested 3 days after the second transfection for western blot. Twenty micrograms of total protein were analyzed by SDS-PAGE (4–20% TGX gels (Bio-Rad)). Proteins were transferred to nitrocellulose membrane (Amersham Protran Supported 0.45 μm NC). After blocking the membrane with 5% non-fat dry milk/TBST, the membrane was incubated with primary antibodies at the following dilution ratio: anti-COX-2 (Cayman Chemical; 1:800), anti-TNRC6A(GW182) (Bethyl Labs; 1:8,000), anti-WDR5 (Cell Signaling Technology; 1:1,000), anti-CNOT1 (Protein Tech; 1:1,000), anti-ANAPC1 (1:2,000), anti-ASH2L (Cell Signaling Technology; 1:2,000), anti-NAT10 (Protein Tech; 1:1,000), anti-MED14 (Abcam; 1:500), anti-CNOT9 (Protein Tech; 1:2,000), anti-β-actin (Sigma-Aldrich; 1:15,000). Then the membrane was incubated with HRP-conjugated anti-mouse IgG or anti-rabbit IgG (Jackson ImmunoResearch; 1:1,000–15,000) secondary antibody. Proteins were visualized using SuperSignal West Pico Chemiluminescent Substrate. Sequences of the siRNAs are listed in **Sup. Table 7**.

5.13. Chromatin Immunoprecipitation (ChIP)

A549 cells (~3,000,000 cells per 15-cm dish) were reverse-transfected with RNA12nc or mismatch oligomer (25 nM) on day 0 with two dishes for each treatment. Media was changed on day 2. On day 3, cells were crosslinked with 1% formaldehyde for 10 min and quenched with

glycine solution (1.25 M in PBS). Cells were collected by scraping and nuclei were isolated using hypotonic lysis buffer (4 ml; 10 mM Tris-HCl pH 7.5, 10 mM NaCl, 3 mM MgCl2, 0.5% (v/v) NP-40). Nuclei were lysed in lysis buffer (1 ml; 1% SDS, 10 mM EDTA, 50 mM Tris-HCl pH 8.1, 1xRoche protease inhibitors cocktail, RNasin Plus RNase inhibitor (Promega)) and sonicated (3 pulses, 20%, 20 s, Ultrasonic Homogenizer Model 150V/T). Nuclear lysate (100 μl), 900 μl IP buffer (0.01% SDS, 1.1% TX-100, 1.2 mM EDTA, 16.7 mM Tris-HCl pH 8.1, 167 mM NaCl, 1xRoche protease inhibitors cocktail, RNasin Plus RNase inhibitor) and 3 μg of rabbit IgG or WDR5 antibody (Bethyl, A302-429A) were rotated overnight. Next day, 40 μl of protein A/G magnetic beads (bimake.com, B23201) was added to the above mixture to capture the antibody-protein-DNA complex. The whole mixture was rotated at room temperature for 50 min. The beads were extensively washed with low salt buffer (containing 150 mM NaCl, 2x), high salt buffer (containing 500 mM NaCl, 2x), LiCl buffer and TE buffer (2x). The antibody-protein-DNA complex was eluted twice with 250 μL of elution buffer (1% SDS, 0.1 M $NaHCO_3$). NaCl solution (final 200 mM) was added to the mixture and cross-linking was reversed by heating the mixture at 65°C for 3 h. RNA was removed by treating mixture with RNase A at 30 °C for 30 min. Protein was digested by incubating with proteinase K (20 μg; invitrogen) at 42°C for 45 min, followed by phenol-chloroform extraction and ethanol precipitation. DNA pellet was dissolved in 100 μl of nuclease free water and analyzed by qPCR. In the data analysis of a WDR5-ChIP sample, the initial Ct value at COX-2 promoter (from -81 to +56) and a gene desert on Chr12 (Active Motif, human negative control primer set 1, 71001) were normalized to input (1%). Then the normalized Ct value at COX-2 promoter was normalized again to normalized Ct value obtained at the gene desert. By comparing the values from samples 12nc and MM, a number for fold change can be obtained. Each ChIP experiment is repeated for at least three times.

Chapter 6
Future Directions and Concluding Remarks

6.1. Introduction

The work within this dissertation shows that key RNAi factors, AGO2 and TNRC6A, are present at promoters and form protein interactions with many proteins involved in transcriptional regulation in cell nuclei. While this work takes the first step towards understanding the mechanisms by which these RNAi factors act in cell nuclei, further work to understand the endogenous functions and mechanisms are needed. For example, my laboratory should perform experiments to identify the promoters that these RNAi factors endogenously regulate. In addition, they should investigate the potential non-transcriptional functions of RNAi factors such DNA repair.

6.2. TNRC6A ChIP to determine endogenous gene targets

My work with TNRC6A introduces the idea that TNRC6 paralogs recruit proteins to promoters for RNA-mediated transcriptional regulation. Validated mass spectrometry identification of endogenous protein interactions of nuclear TNRC6A show that it interacts with many protein complexes involved in transcriptional regulation. In addition, the use of an artificial system of COX-2 RNA-mediated gene activation provides functional evidence suggesting that AGO2, TNRC6A, and transcriptional proteins are involved in this process. The power of these experiments together suggests that this idea can be expanded into global transcriptional regulation by RNAi proteins. Moving forward, it is important to identify the genes RNAi factors AGO2 and TNRC6A are responsible for regulating in cell nuclei.

The easiest way to identify these genes is to perform chromatin immunoprecipitation sequencing (ChIP-seq). However, performing these studies for endogenous levels of proteins with anti-TNRC6A or anti-AGO2 antibodies is not feasible at this time because there are no

commercially available antibodies suitable for ChIP. Colleagues in my laboratory have tested many antibodies for both proteins and determined that the crosslinking conditions used to crosslink the proteins to DNA changes the epitopes of AGO2 or TNRC6A in a way that the antibodies can no longer efficiently IP the protein. An alternative to endogenous ChIP-seq is to use tagged AGO2 or TNRC6A cell lines. The tag would make immunoprecipitation and candidate gene identification much easier. However, a caveat to using this methodology is that the tagged protein is overexpressed and the tag itself might alter natural cellular functions of the proteins. The ChIP-seq datasets from tagged AGO2 and TNRC6A cell lines can be compared to determine whether these proteins localize to the same genes or act independently of each other. Although this method uses tagged and overexpressed proteins, it is a good starting point to provide insights into what genes RNAi factors are naturally regulating in cell nuclei.

6.3. RNA-mediated Transcriptional Silencing and Activation of the Progesterone Receptor (PR)

The work within this dissertation provides evidence that AGO2 and TNRC6 proteins are necessary for RNA-mediated activation of COX-2 transcription and that other proteins such as NAT10, WDR5 and MED14 are recruited to the promoter for activation of transcription. However, this model is promoter-specific and more work needs to be done to describe RNAi transcriptional regulation at different promoters and more global regulation as described in the previous section.

My laboratory's previous research showed that siRNAs can be used to silence or activate transcription of both progesterone receptor transcripts (PR-A and PR-B) in T47D cells (Janowski et al., 2005a) and MCF7 cells (Janowski et al., 2007), respectively. They also showed that AGO2 is involved in that process and the target RNAs are nascent transcripts (Chu et al., 2010) and not the promoter DNA itself. This work was novel in showing that RNAi factors were involved in

transcriptional regulation, although the only RNAi factors tested or function in this system were the AGO proteins.

After showing TNRC6A interacts with the CCR4-NOT complex and transcription factors in cell nuclei and the involvement of these factors in RNA-mediated transcriptional activation of COX-2, the next step would be to test the involvement of these factors in the PR RNA-mediated silencing/activation model system. These two systems provide an example of differential promoter responses due to cellular context and have been characterized by my lab in great detail. It is imperative to test whether TNRC6 proteins contribute to these processes, especially in an RNA-mediated silencing context rather than activating. This has the potential to show versatility of these TNRC6 complexes for different functions and might reveal the involvement of different factors in each context. This might also reveal similar mechanisms for RNA-mediated activation in two systems.

6.4. Mass Spectrometry of nuclear TNRC6B and TNRC6C complexes

At the time of project execution, the best commercially available antibody for the TNRC6 proteins was the Bethyl TNRC6A antibody. Therefore, I chose to pursue mass spectrometry experiments with that antibody. More recently, antibodies for the other two paralogs have been developed for IPs (TNRC6B-Abcam, TNRC6C-Bethyl). Although it is unpredictable whether these antibodies would work for large-scale IP experiments, I propose to use them to isolate TNRC6B and C complexes for mass spectrometry analysis. These datasets would be verified through Co-IPs and used to compare the similarities and differences of each paralog in terms of their interactions. If these proteins are in fact endogenously redundant, the same complexes should be identified as interactions. Based on their protein structure homology, I predict this will be the case. However, they also might have a few differences in their interactions and these differences could be further studied.

6.5. Mass spectrometry of nuclear CNOT1 complexes

CNOT1 is a known binding partner of TNRC6A (Fabian et al., 2011) and was identified as a strong interacting protein in my TNRC6A mass spectrometry datasets. CNOT1 is also the core protein of the CCR4-NOT complex and a scaffolding protein like TNRC6A (Maillet et al., 2000). Identifying the complexes that can be formed from these two scaffolding proteins binding together would be significant to the field. This would provide a new shell of protein interactions in these TNRC6A complexes. Given this is an interesting direction for my project, I attempted CNOT1 mass spectrometry experiments that needs improvement in the future.

To date, there is only one commercially available antibody for CNOT1 that is suitable for IPs. However, I discovered that antibody does not IP under the chemical crosslinking conditions used to crosslink the antibody to resin. That is a critical step in my large-scale IP protocol. Without the crosslinking, some of the antibody will dissociate from the resin during the harsh washing conditions and proteins will be lost. Given that caveat, I performed duplicate IPs in nuclear extracts without crosslinking the antibody to resin and submitted the samples for mass spectrometry. A western blot for a fraction of the sample was used to check the samples prior to submission **(Figure 6.1A)**. Although CNOT1 is present in the sample IP lanes, the signal was not as strong as it should have been for a high quality sample. That became apparent when the mass spectrometry datasets were analyzed. CNOT1 did not have optimal spectral counts or ratios and this lead to the identification of few proteins **(Figure 6.1B).** Despite the dataset quality, it should be noted that this was a good starting point because other CCR4-NOT proteins were identified as well as some involved in transcriptional regulation, splicing/RNA processing, and DNA repair. Some proteins identified overlap with the TNRC6A datasets, including TNKS1BP1, TRIM24 and the CCR4-NOT proteins. In the future, the laboratory should consider another antibody for these experiments once one becomes available.

6.6. Interactions of the TNRC6A N-term and C-term domains

The proposed model of TNRC6 function in RNA-mediated transcriptional regulation is that the N-terminal domain (AGO binding domain) of TNRC6 binds to the AGO-RNA complexes bound to target RNA at the promoter and the C-terminal domain (Silencing Domain) of TNRC6 forms other interactions with proteins responsible for transcriptional regulation. The general idea of this model is the same for TNRC6 cytoplasmic function (Yao et al., 2011). AGO2 binding is required to recruit TNRC6 proteins to cytoplasmic mRNA targets and the C-terminal domain recruits proteins involved in translational silencing. In fact, it was shown that the N-terminal (AGO binding) domain cannot silence mRNA targets independent of the C-terminal (Silencing) domain, but the C-terminal domain fused to mRNA targets had full silencing capabilities (Lazzaretti et al., 2009). This further supports the concept that the C-terminal domain recruits factors to the RNA targets for downstream functions. To further support our model of RNA-mediated transcriptional regulation, experiments differentiating the N-term from the C-term functions of TNRC6 proteins should be performed in cell nuclei.

The most straightforward experiments to determine the protein interactions of each TNRC6A domain would be to transiently express tagged N-term and C-term domains of TNRC6A in cells. Similar Co-IP experiments as in chapter 4 can be performed including AGO2, CNOT1, CNOT9, RbBP5, WDR5, NAT10, MED14 and ANAPC1. The N-terminal domain is predicted to exclusively interact with AGO2 and the C-terminal domain is predicted to interact with all other proteins. Unbiased experiments include performing mass spectrometry on purified nuclear N-term and C-term TNRC6A complexes and comparing datasets.

Another experiment would be to activate COX-2 expression with RNA12nc as described in chapter 4 (Matsui et al., 2013) and transfect either full length, N-term or C-term fragments of

TNRC6A or other mutants of TNRC6A lacking silencing domain regions such as the CIM1/2 GW/WG motifs and the PAM2 and RRM regions. It would be predicted that transfection of full length TNRC6A would further enhance transcriptional activation. The N-term fragment would likely not enhance transcriptional activation because although it binds to AGO, it does not have any other binding sites to recruit proteins to the promoter. The C-term fragment would likely not enhance transcriptional activation either because it cannot be recruited to the promoter by binding AGO. The TNRC6A silencing domain deletions may be shown to have no effect or enhance transcriptional activation of COX-2 depending on what regions bind proteins to recruit to the promoter. These experiments would provide more information on the mechanism by which TNRC6A recruits transcription factors to the COX-2 promoter and which regions of the protein are responsible for interactions.

6.7. Localization of TNRC6A binding partners after knockdown

It has previously been shown that TNRC6A may act as a nucleo-cytoplasmic shuttling protein for AGO2 (Nishi et al., 2013; Nishi et al., 2015). Microscopy experiments showed that TNRC6A mutants with deleted nuclear export signal (NES) retain AGO2 in cell nuclei and the opposite is also true for the TNRC6A-NLS mutant. The work within this dissertation shows that TNRC6A interacts with other proteins that are not exclusively localized to cell nuclei. Those proteins include ANAPC1, WDR5, RbBP5, CNOT1 and CNOT9 (**Figure 4.4A**). This opens up the question of whether TNRC6A shuttles any of these proteins into cell nuclei and if they are shuttled in complexes between the compartments, especially the CCR4-NOT proteins since they have already been discovered as direct binding partners. Furthermore, this also open up the question of whether the localization patterns of these proteins within cell nuclei are influenced by the presence of TNRC6 proteins.

Preliminary experiments have been performed to show expression and localization patterns of AGO2 with TNRC6A/B/C knockdown in whole cell, cytoplasmic and nuclear extracts (**Figure 6.2**). As seen with whole cell extracts, knockdown of TNRC6A/B/C results in an increase of AGO2 protein expression by 50%. The same pattern is observed in cytoplasmic and nuclear extracts as well. AGO2 expression levels go up by 42% and 61%, respectively. In cytoplasmic extracts, an efficient TNRC6A knockdown was observed, however in nuclear extracts, only a knockdown efficiency of 34% was achieved. This means that TNRC6A is retained in cell nuclei under knockdown conditions. The same phenomenon was observed with AGO2 knockdowns in our laboratory (Matsui et al., 2015). AGO2 in cell nuclei only had a knockdown efficiency of 45%. Interestingly, AGO2 knockdown in whole cell, cytoplasm and nuclear extracts increased the protein levels of TNRC6A, especially in whole cell and cytoplasmic extracts. These data suggest that the expression of these two RNAi factors are somehow linked. Given the observations of other protein interactions of TNRC6A in this dissertation, it would be interesting to observe the changes of expression levels of proteins like CNOT1, CNOT9, WDR5, RbBP5, NAT10 and ANAPC1 under TNRC6 knockdown conditions.

After initial observations of protein expression through cellular fractionations, microscopy should be used to observe changes in localization patterns of these proteins as well. This would help determine whether TNRC6 knockdown changes localization patterns within subcellular compartments. Another way to observe changes is to transiently transfect the TNRC6A-NES and TNRC6A-NLS mutants into cells and observe whether the overexpression of TNRC6A in either compartment influences localization of interacting factors.

6.8. The role of TNRC6 paralogs in DNA Damage Response (DDR)

Recently, a role for non-coding RNAs has been implicated as important for the DNA damage response (DDR) (d'Adda di Fagagna, 2014; Wei et al., 2012). The model for this is that

active transcription around the double-stranded break (DSB) forms nascent transcripts at the site which then form long double-stranded RNAs (dsRNAs), processed into dsRNAs, and the dsRNAs targeting the DSB site recruit DNA damage response (DDR) proteins for repair, including MDC1 and 53BP1 (**Figure 6.2**) (d'Adda di Fagagna, 2014). The roles of DICER, DROSHA and AGO2 have been implicated as necessary for this process (Francia et al., 2016; Francia et al., 2012; Gao et al., 2014). Knockdown or inactivity of these proteins inhibited DSB repair and caused a loss of recruitment of important DDR proteins such as MDC1 and 53BP1.

The data presented in this dissertation show potential TNRC6A interactions with the DNA damage response mediator proteins, MDC1 and 53BP1, and two other proteins involved in DSB repair, TNKS1BP1 (Zou et al., 2015) and CCAR2 (Magni et al., 2014; Magni et al., 2015) (**Figure 4.3**). Although these interactions have not been verified via co-immunoprecipitations, it should be of interest to our laboratory to explore these interactions further. Given what is known about the role of RNAi in the DDR pathway and what we know about TNRC6-dependent recruitment of factors (ie., MDC1 and 53BP1) to the sites of RNAi regulation, it would be interesting to test whether TNRC6 proteins have a role in the DNA damage response.

Experiments using either TNRC6A/B/C knockdown or Tagged-TNRC6A-overexpressed cells could be used to test involvement in DDR. Double-strand breaks can be induced by irradiation (IR) and markers for DSB repair such as γH2AX (Phosphorylated histone at DSB sites) can be visualized via fluorescent microscopy. If TNRC6 proteins are involved in the DDR, it would be expected that under knockdown conditions, the response would be inhibited and vice versa for overexpression. To further test for whether TNRC6 proteins are involved in recruiting the DDR mediator proteins, MDC1 and 53BP1, co-localization microscopy under TNRC6 knockdown and overexpression conditions can be used to see whether these proteins are localized to sites of DDR after IR using γH2AX as a marker. As predicted, if TNRC6 proteins act as recruiting factors

for proteins involved in DDR, the co-localization would be reduced under knockdown conditions and vice versa for overexpression. These experiments could answer how dsRNAs targeting DSB sites recruit DDR proteins and could provide another general mechanism of AGO2-RNA-dependent TNRC6 recruitment of proteins for RNAi regulation.

6.9. Concluding Remarks

The experiments descried above, along with the work described in this dissertation will provide a detailed picture of RNAi in mammalian cell nuclei. By using relatively unbiased strategies to build an interaction network of TNRC6A, the interaction partners can be broken down into different aspects of nuclear activity, including transcription and DNA damage. In this dissertation, I have described a novel mechanism in which TNRC6 proteins recruit transcription factors to promoters for RNA-mediated regulation. The role of TNRC6 proteins in nuclear activities has remained elusive even though they are the main binding partners for AGO proteins and act as scaffolding proteins. My project has provided more insights into another layer of gene regulation. These data can lead to new advances in RNAi capabilities far beyond what it is currently applied to.

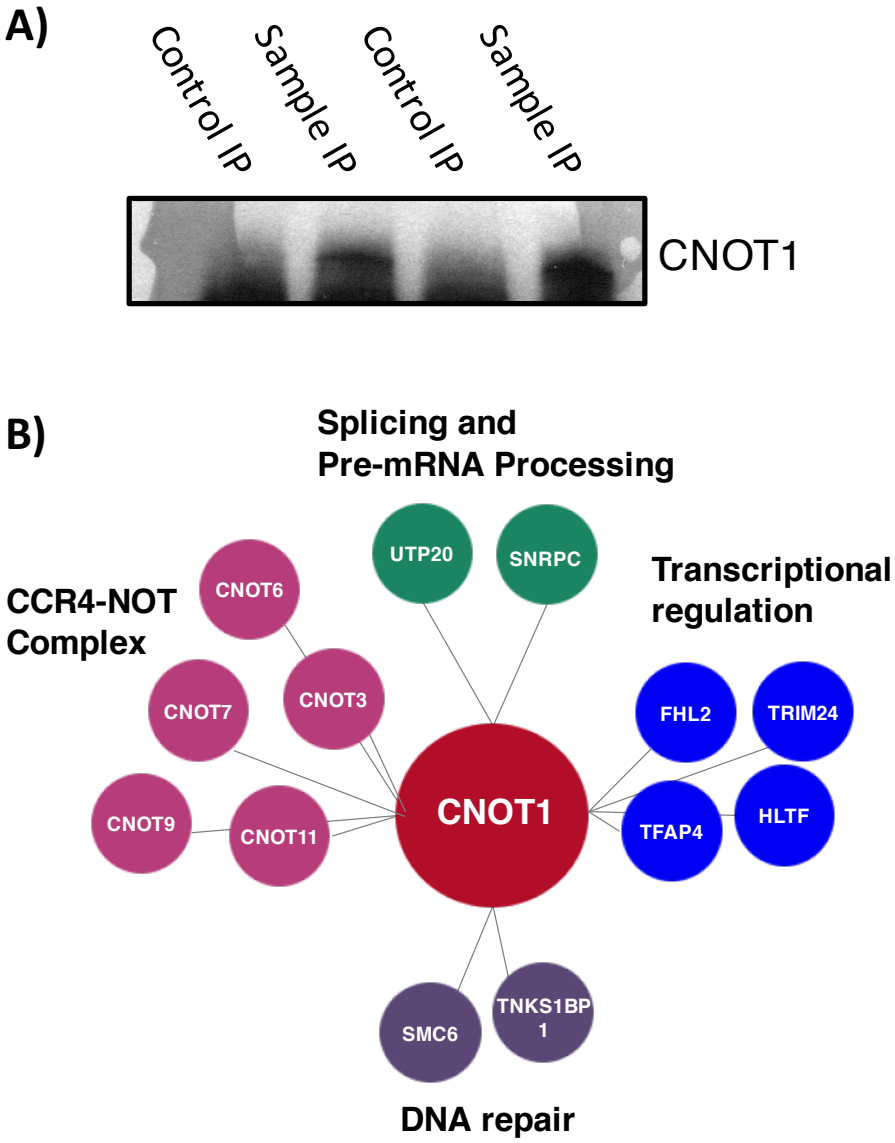


Figure 6.1. Interacting proteins of CNOT1.

a) Western blot of mass spectrometry duplicate CNOT1 sample IPs.
b) Bubble plot of CNOT1 interacting proteins identified by mass spectrometry. Spectral counts >3, Ratio >2

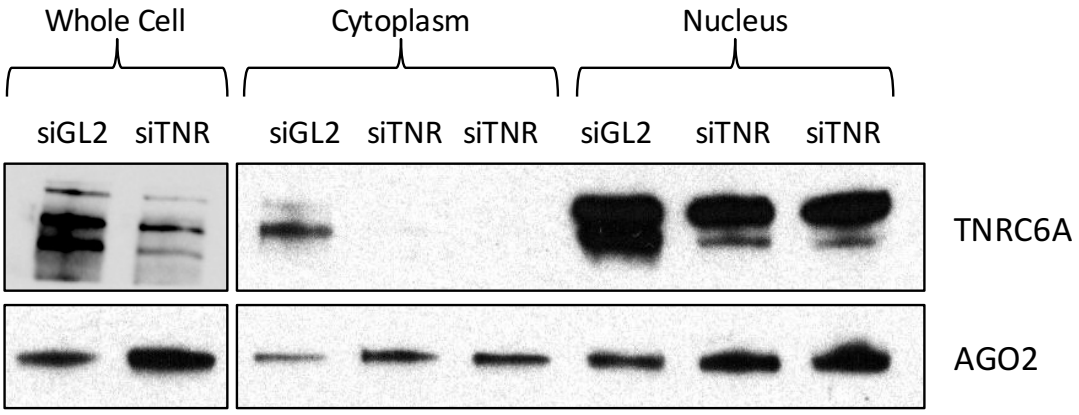


Figure 6.2. Cellular localization of RNAi factors after TNRC6A/B/C knockdown.
Transfection of control (siGL2) and TNRC6A/B/C siRNAs in either whole cell, cytoplasmic or nuclear extracts. Duplicates of knockdowns were used for cytoplasmic and nuclear extracts.

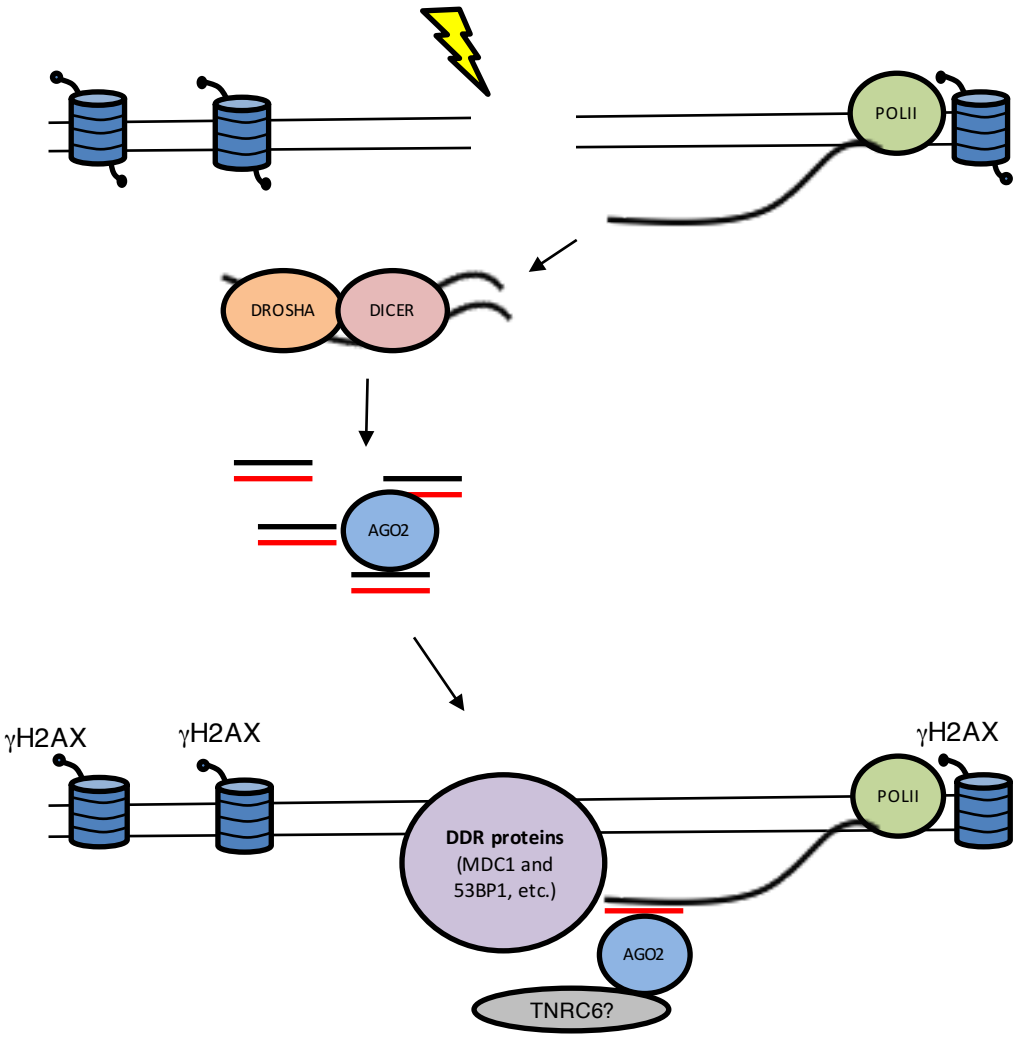


Figure 6.3. The role of RNAi factors in the DNA damage response (DDR).
A double-stranded break occurs in DNA, histones are phosphorylated with gH2AX mark, and nascent transcripts form long double-stranded RNAs (dsRNAs) to be processed by DICER/DROSHA. After processing into dsRNAs, AGO2 and RISC complex binds to dsRNAs and is guided to the target nascent transcript. After RISC binding, the DNA damage response (DDR) proteins, including MDC1 and 53BP1, are recruited for repair.

BIBLIOGRAPHY

Aartsma-Rus, A., Janson, A.A., Kaman, W.E., Bremmer-Bout, M., den Dunnen, J.T., Baas, F., van Ommen, G.J., and van Deutekom, J.C. (2003). Therapeutic antisense-induced exon skipping in cultured muscle cells from six different DMD patients. Human molecular genetics *12*, 907-914.

Allo, M., Agirre, E., Bessonov, S., Bertucci, P., Gomez Acuna, L., Buggiano, V., Bellora, N., Singh, B., Petrillo, E., Blaustein, M.*, et al.* (2014). Argonaute-1 binds transcriptional enhancers and controls constitutive and alternative splicing in human cells. Proc Natl Acad Sci U S A *111*, 15622-15629.

Allo, M., Buggiano, V., Fededa, J.P., Petrillo, E., Schor, I., de la Mata, M., Agirre, E., Plass, M., Eyras, E., Elela, S.A.*, et al.* (2009). Control of alternative splicing through siRNA-mediated transcriptional gene silencing. Nat Struct Mol Biol *16*, 717-724.

Ameyar-Zazoua, M., Rachez, C., Souidi, M., Robin, P., Fritsch, L., Young, R., Morozova, N., Fenouil, R., Descostes, N., Andrau, J.C.*, et al.* (2012). Argonaute proteins couple chromatin silencing to alternative splicing. Nat Struct Mol Biol *19*, 998-1004.

Ansari, S.A., and Morse, R.H. (2013). Mechanisms of Mediator complex action in transcriptional activation. Cellular and molecular life sciences : CMLS *70*, 2743-2756.

Arechavala-Gomeza, V., Graham, I.R., Popplewell, L.J., Adams, A.M., Aartsma-Rus, A., Kinali, M., Morgan, J.E., van Deutekom, J.C., Wilton, S.D., Dickson, G.*, et al.* (2007). Comparative analysis of antisense oligonucleotide sequences for targeted skipping of exon 51 during dystrophin pre-mRNA splicing in human muscle. Human gene therapy *18*, 798-810.

Astorgues-Xerri, L., Riveiro, M.E., Tijeras-Raballand, A., Serova, M., Neuzillet, C., Albert, S., Raymond, E., and Faivre, S. (2014). Unraveling galectin-1 as a novel therapeutic target for cancer. Cancer treatment reviews *40*, 307-319.

Baker, P.R., Trinidad, J.C., and Chalkley, R.J. (2011). Modification site localization scoring integrated into a search engine. Molecular & cellular proteomics : MCP *10*, M111.008078.

Batsche, E., and Ameyar-Zazoua, M. (2015). The influence of Argonaute proteins on alternative RNA splicing. Wiley Interdiscip Rev RNA *6*, 141-156.

Bedford, M.T., and Clarke, S.G. (2009). Protein arginine methylation in mammals: who, what, and why. Mol Cell *33*, 1-13.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006a). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes & development *20*, 1885-1898.

Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. (2006b). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. Cold Spring Harbor symposia on quantitative biology *71*, 523-530.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Braun, J.E., Huntzinger, E., Fauser, M., and Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. Mol Cell *44*, 120-133.

Braun, J.E., Huntzinger, E., and Izaurralde, E. (2012). A molecular link between miRISCs and deadenylases provides new insight into the mechanism of gene silencing by microRNAs. Cold Spring Harb Perspect Biol *4*.

Braun, J.E., Huntzinger, E., and Izaurralde, E. (2013). The role of GW182 proteins in miRNA-mediated gene silencing. Advances in experimental medicine and biology *768*, 147-163.

Brill, L.M., Xiong, W., Lee, K.B., Ficarro, S.B., Crain, A., Xu, Y., Terskikh, A., Snyder, E.Y., and Ding, S. (2009). Phosphoproteomic analysis of human embryonic stem cells. Cell stem cell *5*, 204-213.

Buchberger, A. (2002). From UBA to UBX: new words in the ubiquitin vocabulary. Trends in cell biology *12*, 216-221.

Burkhart, K.B., Guang, S., Buckley, B.A., Wong, L., Bochner, A.F., and Kennedy, S. (2011). A pre-mRNA-associating factor links endogenous siRNAs to chromatin regulation. PLoS genetics *7*, e1002249.

Burton, N.O., Burkhart, K.B., and Kennedy, S. (2011). Nuclear RNAi maintains heritable gene silencing in Caenorhabditis elegans. Proc Natl Acad Sci U S A *108*, 19683-19688.

Carissimi, C., Laudadio, I., Cipolletta, E., Gioiosa, S., Mihailovich, M., Bonaldi, T., Macino, G., and Fulci, V. (2015). ARGONAUTE2 cooperates with SWI/SNF complex to determine nucleosome occupancy at human Transcription Start Sites. Nucleic Acids Res *43*, 1498-1512.

Castel, S.E., and Martienssen, R.A. (2013). RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet *14*, 100-112.

Cevher, M.A., Shi, Y., Li, D., Chait, B.T., Malik, S., and Roeder, R.G. (2014). Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. Nat Struct Mol Biol *21*, 1028-1034.

Chalkley, R.J., and Clauser, K.R. (2012). Modification site localization scoring: strategies and performance. Molecular & cellular proteomics : MCP *11*, 3-14.

Chang, L., and Barford, D. (2014). Insights into the anaphase-promoting complex: a molecular machine that regulates mitosis. Current opinion in structural biology *29*, 1-9.

Chekulaeva, M., Filipowicz, W., and Parker, R. (2009). Multiple independent domains of dGW182 function in miRNA-mediated repression in Drosophila. Rna *15*, 794-803.

Chekulaeva, M., Mathys, H., Zipprich, J.T., Attig, J., Colic, M., Parker, R., and Filipowicz, W. (2011). miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. Nat Struct Mol Biol *18*, 1218-1226.

Chen, C.Y., Zheng, D., Xia, Z., and Shyu, A.B. (2009). Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. Nat Struct Mol Biol *16*, 1160-1166.

Chen, Y., Boland, A., Kuzuoglu-Ozturk, D., Bawankar, P., Loh, B., Chang, C.T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. Mol Cell *54*, 737-750.

Chicooree, N., Unwin, R.D., and Griffiths, J.R. (2015). The application of targeted mass spectrometry-based strategies to the detection and localization of post-translational modifications. Mass spectrometry reviews *34*, 595-626.

Choi, H., Glatter, T., Gstaiger, M., and Nesvizhskii, A.I. (2012a). SAINT-MS1: protein-protein interaction scoring using label-free intensity data in affinity purification-mass spectrometry experiments. Journal of proteome research *11*, 2619-2624.

Choi, H., Larsen, B., Lin, Z.Y., Breitkreutz, A., Mellacheruvu, D., Fermin, D., Qin, Z.S., Tyers, M., Gingras, A.C., and Nesvizhskii, A.I. (2011). SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nature methods *8*, 70-73.

Choi, H., Liu, G., Mellacheruvu, D., Tyers, M., Gingras, A.C., and Nesvizhskii, A.I. (2012b). Analyzing protein-protein interactions from affinity purification-mass spectrometry data with SAINT. Current protocols in bioinformatics / editoral board, Andreas D Baxevanis [et al] *Chapter 8*, Unit8 15.

Chu, C.Y., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. PLoS biology *4*, e210.

Chu, Y., Wang, T., Dodd, D., Xie, Y., Janowski, B.A., and Corey, D.R. (2015). Intramolecular circularization increases efficiency of RNA sequencing and enables CLIP-Seq of nuclear RNA from human cells. Nucleic Acids Res *43*, e75.

Chu, Y., Yue, X., Younger, S.T., Janowski, B.A., and Corey, D.R. (2010). Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. Nucleic Acids Res *38*, 7736-7748.

Collart, M.A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip Rev RNA *7*, 438-454.

Cox, J., and Mann, M. (2011). Quantitative, high-resolution proteomics for data-driven systems biology. Annual review of biochemistry *80*, 273-299.

Craig, R., and Beavis, R.C. (2004). TANDEM: matching proteins with tandem mass spectra. Bioinformatics *20*, 1466-1467.

Creamer, K.M., and Partridge, J.F. (2011). RITS-connecting transcription, RNA interference, and heterochromatin assembly in fission yeast. Wiley Interdiscip Rev RNA *2*, 632-646.

d'Adda di Fagagna, F. (2014). A direct role for small non-coding RNAs in DNA damage response. Trends in cell biology *24*, 171-178.

Dallas, P.B., Pacchione, S., Wilsker, D., Bowrin, V., Kobayashi, R., and Moran, E. (2000). The human SWI-SNF complex protein p270 is an ARID family member with non-sequence-specific DNA binding activity. Molecular and cellular biology *20*, 3137-3146.

Daniels, S.M., and Gatignol, A. (2012). The multiple functions of TRBP, at the hub of cell responses to viruses, stress, and cancer. Microbiology and molecular biology reviews : MMBR *76*, 652-666.

Derry, M.C., Yanagiya, A., Martineau, Y., and Sonenberg, N. (2006). Regulation of poly(A)-binding protein through PABP-interacting proteins. Cold Spring Harbor symposia on quantitative biology *71*, 537-543.

Ding, L., and Han, M. (2007). GW182 family proteins are crucial for microRNA-mediated gene silencing. Trends in cell biology *17*, 411-416.

Djuranovic, S., Nahvi, A., and Green, R. (2012). miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science (New York, NY) *336*, 237-240.

Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L., Allis, C.D., and Roeder, R.G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol *13*, 713-719.

Duan, C.G., Zhang, H., Tang, K., Zhu, X., Qian, W., Hou, Y.J., Wang, B., Lang, Z., Zhao, Y., Wang, X*., et al.* (2015). Specific but interdependent functions for Arabidopsis AGO4 and AGO6 in RNA-directed DNA methylation. The EMBO journal *34*, 581-592.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494-498.

Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nature methods *4*, 207-214.

Eulalio, A., Helms, S., Fritzsch, C., Fauser, M., and Izaurralde, E. (2009a). A C-terminal silencing domain in GW182 is essential for miRNA function. Rna *15*, 1067-1077.

Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008). GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. Nat Struct Mol Biol *15*, 346-353.

Eulalio, A., Huntzinger, E., Nishihara, T., Rehwinkel, J., Fauser, M., and Izaurralde, E. (2009b). Deadenylation is a widespread effect of miRNA regulation. Rna *15*, 21-32.

Eulalio, A., Rehwinkel, J., Stricker, M., Huntzinger, E., Yang, S.F., Doerks, T., Dorner, S., Bork, P., Boutros, M., and Izaurralde, E. (2007). Target-specific requirements for enhancers of decapping in miRNA-mediated gene silencing. Genes & development *21*, 2558-2570.

Eulalio, A., Tritschler, F., Buttner, R., Weichenrieder, O., Izaurralde, E., and Truffault, V. (2009c). The RRM domain in GW182 proteins contributes to miRNA-mediated gene silencing. Nucleic Acids Res *37*, 2974-2983.

Eulalio, A., Tritschler, F., and Izaurralde, E. (2009d). The GW182 protein family in animal cells: new insights into domains required for miRNA-mediated gene silencing. RNA *15*, 1433-1442.

Eystathioy, T., Chan, E.K., Tenenbaum, S.A., Keene, J.D., Griffith, K., and Fritzler, M.J. (2002). A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. Molecular biology of the cell *13*, 1338-1351.

Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F*., et al.* (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. Nat Struct Mol Biol *18*, 1211-1217.

Fabian, M.R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J.T., Svitkin, Y.V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J.A*., et al.* (2009). Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. Mol Cell *35*, 868-880.

Fabian, M.R., and Sonenberg, N. (2012). The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC. Nat Struct Mol Biol *19*, 586-593.

Francia, S., Cabrini, M., Matti, V., Oldani, A., and d'Adda di Fagagna, F. (2016). DICER, DROSHA and DNA damage response RNAs are necessary for the secondary recruitment of DNA damage response factors. Journal of cell science *129*, 1468-1476.

Francia, S., Michelini, F., Saxena, A., Tang, D., de Hoon, M., Anelli, V., Mione, M., Carninci, P., and d'Adda di Fagagna, F. (2012). Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature *488*, 231-235.

Frohn, A., Eberl, H.C., Stohr, J., Glasmacher, E., Rudel, S., Heissmeyer, V., Mann, M., and Meister, G. (2012). Dicer-dependent and -independent Argonaute2 protein interaction networks in mammalian cells. Molecular & cellular proteomics : MCP *11*, 1442-1456.

Fukaya, T., and Tomari, Y. (2011). PABP is not essential for microRNA-mediated translational repression and deadenylation in vitro. The EMBO journal *30*, 4998-5009.

Gagnon, K.T., and Corey, D.R. (2012). Argonaute and the nuclear RNAs: new pathways for RNA-mediated control of gene expression. Nucleic acid therapeutics *22*, 3-16.

Gagnon, K.T., Li, L., Chu, Y., Janowski, B.A., and Corey, D.R. (2014a). RNAi factors are present and active in human cell nuclei. Cell reports *6*, 211-221.

Gagnon, K.T., Li, L., Janowski, B.A., and Corey, D.R. (2014b). Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. Nature protocols *9*, 2045-2060.

Gao, M., Wei, W., Li, M.M., Wu, Y.S., Ba, Z., Jin, K.X., Li, M.M., Liao, Y.Q., Adhikari, S., Chong, Z.*, et al.* (2014). Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. Cell research *24*, 532-541.

Geer, L.Y., Markey, S.P., Kowalak, J.A., Wagner, L., Xu, M., Maynard, D.M., Yang, X., Shi, W., and Bryant, S.H. (2004). Open mass spectrometry search algorithm. J Proteome Res *3*, 958-964.

Golden, R.J., Chen, B., Li, T., Braun, J., Manjunath, H., Chen, X., Wu, J., Schmid, V., Chang, T.C., Kopp, F.*, et al.* (2017). An Argonaute phosphorylation cycle promotes microRNA-mediated silencing. Nature *542*, 197-202.

Gow, C.H., Guo, C., Wang, D., Hu, Q., and Zhang, J. (2014). Differential involvement of E2A-corepressor interactions in distinct leukemogenic pathways. Nucleic Acids Res *42*, 137-152.

Griffin, N.M., Yu, J., Long, F., Oh, P., Shore, S., Li, Y., Koziol, J.A., and Schnitzer, J.E. (2010). Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. Nature biotechnology *28*, 83-89.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell *106*, 23-34.

Grzechnik, P., Gdula, M.R., and Proudfoot, N.J. (2015). Pcf11 orchestrates transcription termination pathways in yeast. Genes & development *29*, 849-861.

Gu, S., Jin, L., Huang, Y., Zhang, F., and Kay, M.A. (2012a). Slicing-independent RISC activation requires the argonaute PAZ domain. Current biology : CB *22*, 1536-1542.

Gu, S.G., Pak, J., Guang, S., Maniar, J.M., Kennedy, S., and Fire, A. (2012b). Amplification of siRNA in Caenorhabditis elegans generates a transgenerational sequence-targeted histone H3 lysine 9 methylation footprint. Nature genetics *44*, 157-164.

Gu, W., Shirayama, M., Conte, D., Jr., Vasale, J., Batista, P.J., Claycomb, J.M., Moresco, J.J., Youngman, E.M., Keys, J., Stoltz, M.J*., et al.* (2009). Distinct argonaute-mediated 22G-RNA pathways direct genome surveillance in the C. elegans germline. Mol Cell *36*, 231-244.

Guang, S., Bochner, A.F., Burkhart, K.B., Burton, N., Pavelec, D.M., and Kennedy, S. (2010). Small regulatory RNAs inhibit RNA polymerase II during the elongation phase of transcription. Nature *465*, 1097-1101.

Guang, S., Bochner, A.F., Pavelec, D.M., Burkhart, K.B., Harding, S., Lachowiec, J., and Kennedy, S. (2008). An Argonaute transports siRNAs from the cytoplasm to the nucleus. Science (New York, NY) *321*, 537-541.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol *15*, 509-524.

Haag, J.R., and Pikaard, C.S. (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol *12*, 483-492.

Halic, M., and Moazed, D. (2010). Dicer-independent primal RNAs trigger RNAi and heterochromatin formation. Cell *140*, 504-516.

Hammond, S.M. (2015). An overview of microRNAs. Advanced drug delivery reviews.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science (New York, NY) *293*, 1146-1150.

Han, G., Ye, M., and Zou, H. (2008). Development of phosphopeptide enrichment techniques for phosphoproteome analysis. The Analyst *133*, 1128-1138.

Havens, M.A., and Hastings, M.L. (2016). Splice-switching antisense oligonucleotides as therapeutic drugs. Nucleic Acids Res *44*, 6549-6563.

Hirose, T., and Nakagawa, S. (2012). Paraspeckles: possible nuclear hubs by the RNA for the RNA. Biomolecular concepts *3*, 415-428.

Hock, J., Weinmann, L., Ender, C., Rudel, S., Kremmer, E., Raabe, M., Urlaub, H., and Meister, G. (2007). Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. EMBO reports *8*, 1052-1060.

Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. Nat Rev Genet *16*, 71-84.

Hong, E.J., Villen, J., Gerace, E.L., Gygi, S.P., and Moazed, D. (2005). A cullin E3 ubiquitin ligase complex associates with Rik1 and the Clr4 histone H3-K9 methyltransferase and is required for RNAi-mediated heterochromatin formation. RNA biology *2*, 106-111.

Horman, S.R., Janas, M.M., Litterst, C., Wang, B., MacRae, I.J., Sever, M.J., Morrissey, D.V., Graves, P., Luo, B., Umesalma, S.*, et al.* (2013). Akt-mediated phosphorylation of argonaute 2 downregulates cleavage and upregulates translational repression of MicroRNA targets. Molecular cell *50*, 356-367.

Horn, P.J., Bastie, J.N., and Peterson, C.L. (2005). A Rik1-associated, cullin-dependent E3 ubiquitin ligase is essential for heterochromatin formation. Genes & development *19*, 1705-1714.

Huang, V., Place, R.F., Portnoy, V., Wang, J., Qi, Z., Jia, Z., Yu, A., Shuman, M., Yu, J., and Li, L.C. (2012). Upregulation of Cyclin B1 by miRNA and its implications in cancer. Nucleic Acids Res *40*, 1695-1707.

Huang, V., Qin, Y., Wang, J., Wang, X., Place, R.F., Lin, G., Lue, T.F., and Li, L.C. (2010). RNAa is conserved in mammalian cells. PLoS One *5*, e8848.

Huntzinger, E., Braun, J.E., Heimstadt, S., Zekri, L., and Izaurralde, E. (2010). Two PABPC1-binding sites in GW182 proteins promote miRNA-mediated gene silencing. The EMBO journal *29*, 4146-4160.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet *12*, 99-110.

Ikeda, K., Satoh, M., Pauley, K.M., Fritzler, M.J., Reeves, W.H., and Chan, E.K. (2006). Detection of the argonaute protein Ago2 and microRNAs in the RNA induced silencing complex (RISC) using a monoclonal antibody. Journal of immunological methods *317*, 38-44.

Iliuk, A.B., Martin, V.A., Alicie, B.M., Geahlen, R.L., and Tao, W.A. (2010). In-depth analyses of kinase-dependent tyrosine phosphoproteomes based on metal ion-functionalized soluble nanopolymers. Molecular & cellular proteomics : MCP *9*, 2162-2172.

Ipsaro, J.J., and Joshua-Tor, L. (2015). From guide to target: molecular insights into eukaryotic RNA-interference machinery. Nat Struct Mol Biol *22*, 20-28.

Jakymiw, A., Lian, S., Eystathioy, T., Li, S., Satoh, M., Hamel, J.C., Fritzler, M.J., and Chan, E.K. (2005). Disruption of GW bodies impairs mammalian RNA interference. Nature cell biology *7*, 1267-1274.

Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Hardy, D., Shames, D.S., Minna, J.D., and Corey, D.R. (2005a). Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs. Nature chemical biology *1*, 216-222.

Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Nordsell, R., Shames, D.S., Minna, J.D., and Corey, D.R. (2006). Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat Struct Mol Biol *13*, 787-792.

Janowski, B.A., Kaihatsu, K., Huffman, K.E., Schwartz, J.C., Ram, R., Hardy, D., Mendelson, C.R., and Corey, D.R. (2005b). Inhibiting transcription of chromosomal DNA with antigene peptide nucleic acids. Nature chemical biology *1*, 210-215.

Janowski, B.A., Younger, S.T., Hardy, D.B., Ram, R., Huffman, K.E., and Corey, D.R. (2007). Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. Nature chemical biology *3*, 166-173.

Jarver, P., O'Donovan, L., and Gait, M.J. (2014). A chemical view of oligonucleotides for exon skipping and related drug applications. Nucleic acid therapeutics *24*, 37-47.

Jensen, C.J., Oldfield, B.J., and Rubio, J.P. (2009). Splicing, cis genetic variation and disease. Biochemical Society transactions *37*, 1311-1315.

Jia, S., Kobayashi, R., and Grewal, S.I. (2005). Ubiquitin ligase component Cul4 associates with Clr4 histone methyltransferase to assemble heterochromatin. Nature cell biology *7*, 1007-1013.

Jinek, M., Fabian, M.R., Coyle, S.M., Sonenberg, N., and Doudna, J.A. (2010). Structural insights into the human GW182-PABC interaction in microRNA-mediated deadenylation. Nat Struct Mol Biol *17*, 238-240.

Joshua-Tor, L., and Hannon, G.J. (2011). Ancestral roles of small RNAs: an Ago-centric perspective. Cold Spring Harb Perspect Biol *3*, a003772.

Junxia, W., Ping, G., Yuan, H., Lijun, Z., Jihong, R., Fang, L., Min, L., Xi, W., Ting, H., Ke, D.*, et al.* (2010). Double strand RNA-guided endogeneous E-cadherin up-regulation induces the apoptosis and inhibits proliferation of breast carcinoma cells in vitro and in vivo. Cancer science *101*, 1790-1796.

Kalantari, R., Chiang, C.M., and Corey, D.R. (2016a). Regulation of mammalian transcription and splicing by Nuclear RNAi. Nucleic Acids Res *44*, 524-537.

Kalantari, R., Hicks, J.A., Li, L., Gagnon, K.T., Sridhara, V., Lemoff, A., Mirzaei, H., and Corey, D.R. (2016b). Stable association of RNAi machinery is conserved between the cytoplasm and nucleus of human cells. Rna *22*, 1085-1098.

Kamachi, Y., and Kondoh, H. (2013). Sox proteins: regulators of cell fate specification and differentiation. Development *140*, 4129-4144.

Kawasaki, H., and Taira, K. (2004). Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature *431*, 211-217.

Kawasaki, H., Taira, K., and Morris, K.V. (2005). siRNA induced transcriptional gene silencing in mammalian cells. Cell cycle (Georgetown, Tex) *4*, 442-448.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet *11*, 345-355.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes & development *15*, 2654-2659.

Kim, D.H., Saetrom, P., Snove, O., Jr., and Rossi, J.J. (2008). MicroRNA-directed transcriptional gene silencing in mammalian cells. Proc Natl Acad Sci U S A *105*, 16230-16235.

Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat Struct Mol Biol *13*, 793-797.

Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J*., et al.* (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol Cell *44*, 325-340.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans. Science (New York, NY) *293*, 2269-2271.

Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. Trends in biochemical sciences *30*, 235-239.

Kruk, J.A., Dutta, A., Fu, J., Gilmour, D.S., and Reese, J.C. (2011). The multifunctional Ccr4-Not complex directly promotes transcription elongation. Genes & development *25*, 581-593.

Kweon, H.K., and Hakansson, K. (2006). Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis. Analytical chemistry *78*, 1743-1749.

Landthaler, M., Gaidatzis, D., Rothballer, A., Chen, P.Y., Soll, S.J., Dinic, L., Ojo, T., Hafner, M., Zavolan, M., and Tuschl, T. (2008). Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA *14*, 2580-2596.

Laribee, R.N., Shibata, Y., Mersman, D.P., Collins, S.R., Kemmeren, P., Roguev, A., Weissman, J.S., Briggs, S.D., Krogan, N.J., and Strahl, B.D. (2007). CCR4/NOT complex associates with the proteasome and regulates histone methylation. Proc Natl Acad Sci U S A *104*, 5836-5841.

Lazzaretti, D., Tournier, I., and Izaurralde, E. (2009). The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins. RNA *15*, 1059-1066.

Lee, A.Y., and Chiang, C.M. (2009). Chromatin adaptor Brd4 modulates E2 transcription activity and protein stability. J Biol Chem *284*, 2778-2786.

Li, A., Wei, G., Wang, Y., Zhou, Y., Zhang, X.E., Bi, L., and Chen, R. (2012). Identification of intermediate-size non-coding RNAs involved in the UV-induced DNA damage response in C. elegans. PLoS One *7*, e48066.

Li, J., Chen, X., Yang, H., Wang, S., Guo, B., Yu, L., Wang, Z., and Fu, J. (2006a). The zinc finger transcription factor 191 is required for early embryonic development and cell proliferation. Experimental cell research *312*, 3990-3998.

Li, L.C., Okino, S.T., Zhao, H., Pookot, D., Place, R.F., Urakami, S., Enokida, H., and Dahiya, R. (2006b). Small dsRNAs induce transcriptional activation in human cells. Proc Natl Acad Sci U S A *103*, 17337-17342.

Li, S., Wang, L., Fu, B., Berman, M.A., Diallo, A., and Dorf, M.E. (2014). TRIM65 regulates microRNA activity by ubiquitination of TNRC6. Proc Natl Acad Sci U S A *111*, 6970-6975.

Li, Y., Han, J., Zhang, Y., Cao, F., Liu, Z., Li, S., Wu, J., Hu, C., Wang, Y., Shuai, J.*, et al.* (2016). Structural basis for activity regulation of MLL family methyltransferases. Nature *530*, 447-452.

Li, Y., Lin, H., Deng, C., Yang, P., and Zhang, X. (2008). Highly selective and rapid enrichment of phosphorylated peptides using gallium oxide-coated magnetic microspheres for MALDI-TOF-MS and nano-LC-ESI-MS/MS/MS analysis. Proteomics *8*, 238-249.

Lian, S.L., Li, S., Abadal, G.X., Pauley, B.A., Fritzler, M.J., and Chan, E.K. (2009). The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. Rna *15*, 804-813.

Lima, W.F., Prakash, T.P., Murray, H.M., Kinberger, G.A., Li, W., Chappell, A.E., Li, C.S., Murray, S.F., Gaus, H., Seth, P.P.*, et al.* (2012). Single-stranded siRNAs activate RNAi in animals. Cell *150*, 883-894.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. Nat Struct Mol Biol *11*, 576-577.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science (New York, NY) *305*, 1437-1441.

Liu, J., Hu, J., and Corey, D.R. (2012). Expanding the action of duplex RNAs into the nucleus: redirecting alternative splicing. Nucleic Acids Res *40*, 1240-1250.

Liu, J., Hu, J., Hicks, J.A., Prakash, T.P., and Corey, D.R. (2015). Modulation of Splicing by Single-Stranded Silencing RNAs. Nucleic acid therapeutics.

Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., 3rd, Parker, R., and Hannon, G.J. (2005). A role for the P-body component GW182 in microRNA function. Nature cell biology *7*, 1261-1266.

Lopez-Orozco, J., Pare, J.M., Holme, A.L., Chaulk, S.G., Fahlman, R.P., and Hobman, T.C. (2015). Functional analyses of phosphorylation events in human Argonaute 2. Rna *21*, 2030-2038.

Lopez-Saavedra, A., Gomez-Cabello, D., Dominguez-Sanchez, M.S., Mejias-Navarro, F., Fernandez-Avila, M.J., Dinant, C., Martinez-Macias, M.I., Bartek, J., and Huertas, P. (2016). A genome-wide screening uncovers the role of CCAR2 as an antagonist of DNA end resection. Nature communications *7*, 12364.

Lv, J., Liu, H., Wang, Q., Tang, Z., Hou, L., and Zhang, B. (2003). Molecular cloning of a novel human gene encoding histone acetyltransferase-like protein involved in transcriptional activation of hTERT. Biochemical and biophysical research communications *311*, 506-513.

Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. Nature *434*, 666-670.

Magni, M., Ruscica, V., Buscemi, G., Kim, J.E., Nachimuthu, B.T., Fontanella, E., Delia, D., and Zannini, L. (2014). Chk2 and REGgamma-dependent DBC1 regulation in DNA damage induced apoptosis. Nucleic Acids Res *42*, 13150-13160.

Magni, M., Ruscica, V., Restelli, M., Fontanella, E., Buscemi, G., and Zannini, L. (2015). CCAR2/DBC1 is required for Chk2-dependent KAP1 phosphorylation and repair of DNA damage. Oncotarget *6*, 17817-17831.

Maillet, L., Tu, C., Hong, Y.K., Shuster, E.O., and Collart, M.A. (2000). The essential function of Not1 lies within the Ccr4-Not complex. Journal of molecular biology *303*, 131-143.

Mann, M., and Jensen, O.N. (2003). Proteomic analysis of post-translational modifications. Nature biotechnology *21*, 255-261.

Mann, M., and Wilm, M. (1994). Error-tolerant identification of peptides in sequence databases by peptide sequence tags. Analytical chemistry *66*, 4390-4399.

Mao, C.D., and Byers, S.W. (2011). Cell-context dependent TCF/LEF expression and function: alternative tales of repression, de-repression and activation potentials. Critical reviews in eukaryotic gene expression *21*, 207-236.

Mao, Q., Li, Y., Zheng, X., Yang, K., Shen, H., Qin, J., Bai, Y., Kong, D., Jia, X., and Xie, L. (2008). Up-regulation of E-cadherin by small activating RNA inhibits cell invasion and migration in 5637 human bladder cancer cells. Biochemical and biophysical research communications *375*, 566-570.

Mao, Q., Zheng, X., Yang, K., Qin, J., Bai, Y., Jia, X., Li, Y., and Xie, L. (2010). Suppression of migration and invasion of PC3 prostate cancer cell line via activating E-cadherin expression by small activating RNA. Cancer investigation *28*, 1013-1018.

Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., and Filipowicz, W. (2014). Structural and biochemical insights to the role of the CCR4-NOT complex and DDX6 ATPase in microRNA repression. Mol Cell *54*, 751-765.

Matsui, M., Chu, Y., Zhang, H., Gagnon, K.T., Shaikh, S., Kuchimanchi, S., Manoharan, M., Corey, D.R., and Janowski, B.A. (2013). Promoter RNA links transcriptional regulation of inflammatory pathway genes. Nucleic Acids Res *41*, 10086-10109.

Matsui, M., and Corey, D.R. (2016). Non-coding RNAs as drug targets. Nature reviews Drug discovery.

Matsui, M., Li, L., Janowski, B.A., and Corey, D.R. (2015). Reduced Expression of Argonaute 1, Argonaute 2, and TRBP Changes Levels and Intracellular Distribution of RNAi Factors. Scientific reports *5*, 12855.

Matsui, M., Sakurai, F., Elbashir, S., Foster, D.J., Manoharan, M., and Corey, D.R. (2010). Activation of LDL receptor expression by small RNAs complementary to a noncoding transcript that overlaps the LDLR promoter. Chemistry & biology *17*, 1344-1355.

McCue, A.D., Panda, K., Nuthikattu, S., Choudury, S.G., Thomas, E.N., and Slotkin, R.K. (2015). ARGONAUTE 6 bridges transposable element mRNA-derived siRNAs to the establishment of DNA methylation. The EMBO journal *34*, 20-35.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. Nat Rev Genet *14*, 447-459.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell *15*, 185-197.

Meister, G., Landthaler, M., Peters, L., Chen, P.Y., Urlaub, H., Luhrmann, R., and Tuschl, T. (2005). Identification of novel argonaute-associated proteins. Current biology : CB *15*, 2149-2155.

Mette, M.F., Aufsatz, W., van der Winden, J., Matzke, M.A., and Matzke, A.J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. The EMBO journal *19*, 5194-5201.

Moretti, F., Kaiser, C., Zdanowicz-Specht, A., and Hentze, M.W. (2012). PABP and the poly(A) tail augment microRNA repression by facilitated miRISC binding. Nat Struct Mol Biol *19*, 603-608.

Morris, K.V., Chan, S.W., Jacobsen, S.E., and Looney, D.J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. Science (New York, NY) *305*, 1289-1292.

Mulder, K.W., Brenkman, A.B., Inagaki, A., van den Broek, N.J., and Timmers, H.T. (2007). Regulation of histone H3K4 tri-methylation and PAF complex recruitment by the Ccr4-Not complex. Nucleic Acids Res *35*, 2428-2439.

Naftelberg, S., Schor, I.E., Ast, G., and Kornblihtt, A.R. (2015). Regulation of alternative splicing through coupling with transcription and chromatin structure. Annual review of biochemistry *84*, 165-198.

Nishi, K., Nishi, A., Nagasawa, T., and Ui-Tei, K. (2013). Human TNRC6A is an Argonaute-navigator protein for microRNA-mediated gene silencing in the nucleus. RNA *19*, 17-35.

Nishi, K., Takahashi, T., Suzawa, M., Miyakawa, T., Nagasawa, T., Ming, Y., Tanokura, M., and Ui-Tei, K. (2015). Control of the localization and function of a miRNA silencing component TNRC6A by Argonaute protein. Nucleic Acids Res *43*, 9856-9873.

Oeffinger, M. (2012). Two steps forward--one step back: advances in affinity purification mass spectrometry of macromolecular complexes. Proteomics *12*, 1591-1608.

Ohrt, T., Muetze, J., Svoboda, P., and Schwille, P. (2012). Intracellular localization and routing of miRNA and RNAi pathway components. Current topics in medicinal chemistry *12*, 79-88.

Ohrt, T., Mutze, J., Staroske, W., Weinmann, L., Hock, J., Crell, K., Meister, G., and Schwille, P. (2008). Fluorescence correlation spectroscopy and fluorescence cross-correlation spectroscopy reveal the cytoplasmic origination of loaded nuclear RISC in vivo in human cells. Nucleic Acids Res *36*, 6439-6449.

Parker, J.S., Roe, S.M., and Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. The EMBO journal *23*, 4727-4737.

Peng, W., Togawa, C., Zhang, K., and Kurdistani, S.K. (2008). Regulators of cellular levels of histone acetylation in Saccharomyces cerevisiae. Genetics *179*, 277-289.

Pfaff, J., Hennig, J., Herzog, F., Aebersold, R., Sattler, M., Niessing, D., and Meister, G. (2013). Structural features of Argonaute-GW182 protein interactions. Proc Natl Acad Sci U S A *110*, E3770-3779.

Pfaff, J., and Meister, G. (2013). Argonaute and GW182 proteins: an effective alliance in gene silencing. Biochemical Society transactions *41*, 855-860.

Phelan, M.L., Sif, S., Narlikar, G.J., and Kingston, R.E. (1999). Reconstitution of a core chromatin remodeling complex from SWI/SNF subunits. Mol Cell *3*, 247-253.

Piao, X., Zhang, X., Wu, L., and Belasco, J.G. (2010). CCR4-NOT deadenylates mRNA associated with RNA-induced silencing complexes in human cells. Molecular and cellular biology *30*, 1486-1494.

Place, R.F., Li, L.C., Pookot, D., Noonan, E.J., and Dahiya, R. (2008). MicroRNA-373 induces expression of genes with complementary promoter sequences. Proc Natl Acad Sci U S A *105*, 1608-1613.

Portnoy, V., Lin, S.H., Li, K.H., Burlingame, A., Hu, Z.H., Li, H., and Li, L.C. (2016). saRNA-guided Ago2 targets the RITA complex to promoters to stimulate transcription. Cell research *26*, 320-335.

Qi, H.H., Ongusaha, P.P., Myllyharju, J., Cheng, D., Pakkanen, O., Shi, Y., Lee, S.W., Peng, J., and Shi, Y. (2008). Prolyl 4-hydroxylation regulates Argonaute 2 stability. Nature *455*, 421-424.

Rand, T.A., Ginalski, K., Grishin, N.V., and Wang, X. (2004). Biochemical identification of Argonaute 2 as the sole protein required for RNA-induced silencing complex activity. Proc Natl Acad Sci U S A *101*, 14385-14389.

Rappsilber, J. (2011). The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. Journal of structural biology *173*, 530-540.

Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., and Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. Rna *11*, 1640-1647.

Reinhart, B.J., and Bartel, D.P. (2002). Small RNAs correspond to centromere heterochromatic repeats. Science (New York, NY) *297*, 1831.

Robb, G.B., Brown, K.M., Khurana, J., and Rana, T.M. (2005). Specific and potent RNAi in the nucleus of human cells. Nat Struct Mol Biol *12*, 133-137.

Robb, G.B., and Rana, T.M. (2007). RNA helicase A interacts with RISC in human cells and functions in RISC loading. Mol Cell *26*, 523-537.

Rogers, K., and Chen, X. (2013). Biogenesis, turnover, and mode of action of plant microRNAs. The Plant cell *25*, 2383-2399.

Ruda, V.M., Chandwani, R., Sehgal, A., Bogorad, R.L., Akinc, A., Charisse, K., Tarakhovsky, A., Novobrantseva, T.I., and Koteliansky, V. (2014). The roles of individual mammalian argonautes in RNA interference in vivo. PLoS One *9*, e101749.

Rudel, S., Flatley, A., Weinmann, L., Kremmer, E., and Meister, G. (2008). A multifunctional human Argonaute2-specific monoclonal antibody. RNA *14*, 1244-1253.

Rudel, S., Wang, Y., Lenobel, R., Korner, R., Hsiao, H.H., Urlaub, H., Patel, D., and Meister, G. (2011). Phosphorylation of human Argonaute proteins affects small RNA binding. Nucleic Acids Res *39*, 2330-2343.

Ruthenburg, A.J., Wang, W., Graybosch, D.M., Li, H., Allis, C.D., Patel, D.J., and Verdine, G.L. (2006). Histone H3 recognition and presentation by the WDR5 module of the MLL1 complex. Nat Struct Mol Biol *13*, 704-712.

Sahin, U., Lapaquette, P., Andrieux, A., Faure, G., and Dejean, A. (2014). Sumoylation of human argonaute 2 at lysine-402 regulates its stability. PloS one *9*, e102957.

Schirle, N.T., and MacRae, I.J. (2012). The crystal structure of human Argonaute2. Science (New York, NY) *336*, 1037-1040.

Schlereth, K., Heyl, C., Krampitz, A.M., Mernberger, M., Finkernagel, F., Scharfe, M., Jarek, M., Leich, E., Rosenwald, A., and Stiewe, T. (2013). Characterization of the p53 cistrome--DNA binding cooperativity dissects p53's tumor suppressor functions. PLoS genetics *9*, e1003726.

Schraivogel, D., and Meister, G. (2014). Import routes and nuclear functions of Argonaute and other small RNA-silencing proteins. Trends in biochemical sciences *39*, 420-431.

Schraivogel, D., Schindler, S.G., Danner, J., Kremmer, E., Pfaff, J., Hannus, S., Depping, R., and Meister, G. (2015a). Importin-beta facilitates nuclear import of human GW proteins and balances cytoplasmic gene silencing protein levels. Nucleic Acids Res *43*, 7447-7461.

Schraivogel, D., Schindler, S.G., Danner, J., Kremmer, E., Pfaff, J., Hannus, S., Depping, R., and Meister, G. (2015b). Importin-beta facilitates nuclear import of human GW proteins and balances cytoplasmic gene silencing protein levels. Nucleic acids research.

Schurmann, N., Trabuco, L.G., Bender, C., Russell, R.B., and Grimm, D. (2013). Molecular dissection of human Argonaute proteins by DNA shuffling. Nat Struct Mol Biol *20*, 818-826.

Schwartz, J.C., Younger, S.T., Nguyen, N.B., Hardy, D.B., Monia, B.P., Corey, D.R., and Janowski, B.A. (2008). Antisense transcripts are targets for activating small RNAs. Nat Struct Mol Biol *15*, 842-848.

Sharma, K., D'Souza, R.C., Tyanova, S., Schaab, C., Wisniewski, J.R., Cox, J., and Mann, M. (2014). Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. Cell reports *8*, 1583-1594.

Shen, J., Xia, W., Khotskaya, Y.B., Huo, L., Nakanishi, K., Lim, S.O., Du, Y., Wang, Y., Chang, W.C., Chen, C.H.*, et al.* (2013). EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. Nature *497*, 383-387.

Shi, Y., Lee, J.S., and Galvin, K.M. (1997). Everything you have ever wanted to know about Yin Yang 1. Biochimica et biophysica acta *1332*, F49-66.

Shimada, Y., Mohn, F., and Buhler, M. (2016). The RNA-induced transcriptional silencing complex targets chromatin exclusively via interacting with nascent transcripts. Genes & development *30*, 2571-2580.

Simon, E.S., Young, M., Chan, A., Bao, Z.Q., and Andrews, P.C. (2008). Improved enrichment strategies for phosphorylated peptides on titanium dioxide using methyl esterification and pH gradient elution. Analytical biochemistry *377*, 234-242.

Skourti-Stathaki, K., Kamieniarz-Gdula, K., and Proudfoot, N.J. (2014). R-loops induce repressive chromatin marks over mammalian gene terminators. Nature *516*, 436-439.

Smits, A.H., Jansen, P.W., Poser, I., Hyman, A.A., and Vermeulen, M. (2013). Stoichiometry of chromatin-associated protein complexes revealed by label-free quantitative mass spectrometry-based proteomics. Nucleic Acids Res *41*, e28.

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. Science (New York, NY) *305*, 1434-1437.

Stalder, L., Heusermann, W., Sokol, L., Trojer, D., Wirz, J., Hean, J., Fritzsche, A., Aeschimann, F., Pfanzagl, V., Basselet, P.*, et al.* (2013). The rough endoplasmatic reticulum is a central nucleation site of siRNA-mediated RNA silencing. The EMBO journal *32*, 1115-1127.

Stonestrom, A.J., Hsu, S.C., Jahn, K.S., Huang, P., Keller, C.A., Giardine, B.M., Kadauke, S., Campbell, A.E., Evans, P., Hardison, R.C.*, et al.* (2015). Functions of BET proteins in erythroid gene expression. Blood *125*, 2825-2834.

Sunyaev, S., Liska, A.J., Golod, A., Shevchenko, A., and Shevchenko, A. (2003). MultiTag: multiple error-tolerant sequence tag search for the sequence-similarity identification of proteins by mass spectrometry. Analytical chemistry *75*, 1307-1315.

Suzawa, M., Noguchi, K., Nishi, K., Kozuka-Hata, H., Oyama, M., and Ui-Tei, K. (2017). Comprehensive Identification of Nuclear and Cytoplasmic TNRC6A-Associating Proteins. Journal of molecular biology.

Suzuki, K., Shijuuku, T., Fukamachi, T., Zaunders, J., Guillemin, G., Cooper, D., and Kelleher, A. (2005). Prolonged transcriptional silencing and CpG methylation induced by siRNAs targeted to the HIV-1 promoter region. Journal of RNAi and gene silencing : an international journal of RNA and gene targeting research *1*, 66-78.

Swarts, D.C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R.F., Koonin, E.V., Patel, D.J., and van der Oost, J. (2014). The evolutionary journey of Argonaute proteins. Nat Struct Mol Biol *21*, 743-753.

Tahbaz, N., Kolb, F.A., Zhang, H., Jaronczyk, K., Filipowicz, W., and Hobman, T.C. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. EMBO reports *5*, 189-194.

Taira, K. (2006). Induction of DNA methylation and gene silencing by short interfering RNAs in human cells. Nature *441*, 1176.

Takahashi, T., Zenno, S., Ishibashi, O., Takizawa, T., Saigo, K., and Ui-Tei, K. (2014). Interactions between the non-seed region of siRNA and RNA-binding RLC/RISC proteins, Ago and TRBP, in mammalian cells. Nucleic Acids Res *42*, 5256-5269.

Takimoto, K., Wakiyama, M., and Yokoyama, S. (2009). Mammalian GW182 contains multiple Argonaute-binding sites and functions in microRNA-mediated translational repression. RNA *15*, 1078-1089.

Till, S., Lejeune, E., Thermann, R., Bortfeld, M., Hothorn, M., Enderle, D., Heinrich, C., Hentze, M.W., and Ladurner, A.G. (2007). A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. Nat Struct Mol Biol *14*, 897-903.

Ting, A.H., Schuebel, K.E., Herman, J.G., and Baylin, S.B. (2005). Short double-stranded RNA induces transcriptional gene silencing in human cancer cells in the absence of DNA methylation. Nature genetics *37*, 906-910.

Tolia, N.H., and Joshua-Tor, L. (2007). Slicer and the argonautes. Nature chemical biology *3*, 36-43.

Torchy, M.P., Hamiche, A., and Klaholz, B.P. (2015). Structure and function insights into the NuRD chromatin remodeling complex. Cellular and molecular life sciences : CMLS *72*, 2491-2507.

Torres, M.P., Thapar, R., Marzluff, W.F., and Borchers, C.H. (2005). Phosphatase-directed phosphorylation-site determination: a synthesis of methods for the detection and identification of phosphopeptides. J Proteome Res *4*, 1628-1635.

Trapp, J., Armengaud, J., Salvador, A., Chaumot, A., and Geffard, O. (2014). Next-generation proteomics: toward customized biomarkers for environmental biomonitoring. Environmental science & technology *48*, 13560-13572.

Trudgian, D.C., and Mirzaei, H. (2012). Cloud CPFP: a shotgun proteomics data analysis pipeline using cloud and high performance computing. J Proteome Res *11*, 6282-6290.

Trudgian, D.C., Ridlova, G., Fischer, R., Mackeen, M.M., Ternette, N., Acuto, O., Kessler, B.M., and Thomas, B. (2011a). Comparative evaluation of label-free SINQ normalized spectral index quantitation in the central proteomics facilities pipeline. Proteomics *11*, 2790-2797.

Trudgian, D.C., Ridlova, G., Fischer, R., Mackeen, M.M., Ternette, N., Acuto, O., Kessler, B.M., and Thomas, B. (2011b). Comparative evaluation of label-free SINQ normalized spectral index quantitation in the central proteomics facilities pipeline. Proteomics *11*, 2790-2797.

Trudgian, D.C., Thomas, B., McGowan, S.J., Kessler, B.M., Salek, M., and Acuto, O. (2010). CPFP: a central proteomics facilities pipeline. Bioinformatics *26*, 1131-1132.

Turunen, M.P., Lehtola, T., Heinonen, S.E., Assefa, G.S., Korpisalo, P., Girnary, R., Glass, C.K., Vaisanen, S., and Yla-Herttuala, S. (2009). Efficient regulation of VEGF expression by promoter-targeted lentiviral shRNAs based on epigenetic mechanism: a novel example of epigenetherapy. Circulation research *105*, 604-609.

Ui-Tei, K., Nishi, K., Takahashi, T., and Nagasawa, T. (2012). Thermodynamic Control of Small RNA-Mediated Gene Silencing. Frontiers in genetics *3*, 101.

Underwood, C.J., and Martienssen, R.A. (2015). Argonautes team up to silence transposable elements in Arabidopsis. The EMBO journal *34*, 579-580.

van Deutekom, J.C., Janson, A.A., Ginjaar, I.B., Frankhuizen, W.S., Aartsma-Rus, A., Bremmer-Bout, M., den Dunnen, J.T., Koop, K., van der Kooi, A.J., Goemans, N.M.*, et al.* (2007). Local dystrophin restoration with antisense oligonucleotide PRO051. The New England journal of medicine *357*, 2677-2686.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science (New York, NY) *303*, 672-676.

Vickers, T.A., Koo, S., Bennett, C.F., Crooke, S.T., Dean, N.M., and Baker, B.F. (2003). Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. J Biol Chem *278*, 7108-7118.

Voit, T., Topaloglu, H., Straub, V., Muntoni, F., Deconinck, N., Campion, G., De Kimpe, S.J., Eagle, M., Guglieri, M., Hood, S.*, et al.* (2014). Safety and efficacy of drisapersen for the treatment of Duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. The Lancet Neurology *13*, 987-996.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I., and Martienssen, R.A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science (New York, NY) *297*, 1833-1837.

Wakiyama, M., Takimoto, K., Ohara, O., and Yokoyama, S. (2007). Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. Genes & development *21*, 1857-1862.

Wassenegger, M., Heimes, S., Riedel, L., and Sanger, H.L. (1994). RNA-directed de novo methylation of genomic sequences in plants. Cell *76*, 567-576.

Watts, J.K., and Corey, D.R. (2012). Silencing disease genes in the laboratory and the clinic. The Journal of pathology *226*, 365-379.

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.G., and Qi, Y. (2012). A role for small RNAs in DNA double-strand break repair. Cell *149*, 101-112.

Wei, Y., Li, L., Wang, D., Zhang, C.Y., and Zen, K. (2014). Importin 8 regulates the transport of mature microRNAs into the cell nucleus. J Biol Chem *289*, 10270-10275.

Weinberg, M.S., and Morris, K.V. (2016). Transcriptional gene silencing in humans. Nucleic Acids Res *44*, 6505-6517.

Weinmann, L., Hock, J., Ivacevic, T., Ohrt, T., Mutze, J., Schwille, P., Kremmer, E., Benes, V., Urlaub, H., and Meister, G. (2009). Importin 8 is a gene silencing factor that targets argonaute proteins to distinct mRNAs. Cell *136*, 496-507.

Weth, O., and Renkawitz, R. (2011). CTCF function is modulated by neighboring DNA binding factors. Biochemistry and cell biology = Biochimie et biologie cellulaire *89*, 459-468.

Wierzbicki, A.T., Haag, J.R., and Pikaard, C.S. (2008). Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell *135*, 635-648.

Wilson, R.C., and Doudna, J.A. (2013). Molecular mechanisms of RNA interference. Annual review of biophysics *42*, 217-239.

Wu, L., Mao, L., and Qi, Y. (2012). Roles of dicer-like and argonaute proteins in TAS-derived small interfering RNA-triggered DNA methylation. Plant physiology *160*, 990-999.

Xu, X., and Stern, D.F. (2003). NFBD1/MDC1 regulates ionizing radiation-induced focus formation by DNA checkpoint signaling and repair factors. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *17*, 1842-1848.

Yang, K., Zheng, X.Y., Qin, J., Wang, Y.B., Bai, Y., Mao, Q.Q., Wan, Q., Wu, Z.M., and Xie, L.P. (2008). Up-regulation of p21WAF1/Cip1 by saRNA induces G1-phase arrest and apoptosis in T24 human bladder cancer cells. Cancer letters *265*, 206-214.

Yang, M., Haase, A.D., Huang, F.K., Coulis, G., Rivera, K.D., Dickinson, B.C., Chang, C.J., Pappin, D.J., Neubert, T.A., Hannon, G.J.*, et al.* (2014). Dephosphorylation of tyrosine 393 in argonaute 2 by protein tyrosine phosphatase 1B regulates gene silencing in oncogenic RAS-induced senescence. Molecular cell *55*, 782-790.

Yang, W.Q., and Zhang, Y. (2012). RNAi-mediated gene silencing in cancer therapy. Expert opinion on biological therapy *12*, 1495-1504.

Yao, B., Li, S., Jung, H.M., Lian, S.L., Abadal, G.X., Han, F., Fritzler, M.J., and Chan, E.K. (2011). Divergent GW182 functional domains in the regulation of translational silencing. Nucleic Acids Res *39*, 2534-2547.

Yoshida, H., Hirano, K., Sato, T., Mitsuda, N., Nomoto, M., Maeo, K., Koketsu, E., Mitani, R., Kawamura, M., Ishiguro, S.*, et al.* (2014). DELLA protein functions as a transcriptional activator

through the DNA binding of the indeterminate domain family proteins. Proc Natl Acad Sci U S A *111*, 7861-7866.

Younger, S.T., and Corey, D.R. (2011). Transcriptional gene silencing in mammalian cells by miRNA mimics that target gene promoters. Nucleic Acids Res *39*, 5682-5691.

Yu, F., Li, N., and Yu, W. (2016). PIPI: PTM-Invariant Peptide Identification Using Coding Method. Journal of Proteome Research *15*, 4423-4435.

Zekri, L., Huntzinger, E., Heimstadt, S., and Izaurralde, E. (2009). The silencing domain of GW182 interacts with PABPC1 to promote translational repression and degradation of microRNA targets and is required for target release. Molecular and cellular biology *29*, 6220-6231.

Zeng, Y., and Cullen, B.R. (2002). RNA interference in human cells is restricted to the cytoplasm. Rna *8*, 855-860.

Zeng, Y., Sankala, H., Zhang, X., and Graves, P.R. (2008). Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies. The Biochemical journal *413*, 429-436.

Zhang, X., Li, H., Burnett, J.C., and Rossi, J.J. (2014). The role of antisense long noncoding RNA in small RNA-triggered gene activation. Rna *20*, 1916-1928.

Zhang, Y., Fonslow, B.R., Shan, B., Baek, M.C., and Yates, J.R., 3rd (2013). Protein analysis by shotgun/bottom-up proteomics. Chemical reviews *113*, 2343-2394.

Zhu, Y., Rowley, M.J., Bohmdorfer, G., and Wierzbicki, A.T. (2013). A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. Mol Cell *49*, 298-309.

Zilberman, D., Cao, X., and Jacobsen, S.E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science (New York, NY) *299*, 716-719.

Zipprich, J.T., Bhattacharyya, S., Mathys, H., and Filipowicz, W. (2009). Importance of the C-terminal domain of the human GW182 protein TNRC6C for translational repression. Rna *15*, 781-793.

Zou, L.H., Shang, Z.F., Tan, W., Liu, X.D., Xu, Q.Z., Song, M., Wang, Y., Guan, H., Zhang, S.M., Yu, L.*, et al.* (2015). TNKS1BP1 functions in DNA double-strand break repair though facilitating DNA-PKcs autophosphorylation dependent on PARP-1. Oncotarget *6*, 7011-7022.