An investigation of the RNA induced silencing complex

and its therapeutic implications

by

Vishalakshi Krishnan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2013

Doctoral Committee:

Professor Nils G. Walter, Chair
Assistant Professor Julie Suzanne Biteen
Professor Carol A. Fierke
Assistant Professor Aaron C. Goldstrohm

2013

To Mom, Dad, Alamu and Sethu

Acknowledgements

I would like to extend my immense gratitude to all the people who have made this dissertation possible. First and foremost, I would like to express my heartfelt gratitude to my PhD advisor, Dr. Nils G. Walter, for accepting me into his group and guiding me through my graduate studies. Nils has been an awesome mentor and is one of the most organized and hardworking person I have seen in my life. His unconditional support, feedback and kindness has helped me tremendously not only in research but also personally, considering this is my first experience in a foreign country. I am very thankful and indebted to you Nils for this wonderful experience which I have greatly benefited from and I wish you only the best for the rest of your life.

I would like to extend my thanks to my wonderful committee members who have unfailingly supported me and guided me through my research. You were an awesome committee, who always had my back and guided me with your useful suggestions and comments to drive my research in the right direction. I am very grateful to you for all your support and I seriously couldn't have asked for a better thesis committee.

My lab members- both past and present, the most awesome people I have met in my life and would hate to leave behind, you all rock. I am very very grateful to all of you for making my stay in the Walter Lab unbelievably pleasant, enjoyable, fun, exciting and insightful both scientifically and otherwise. As members of the lab we are pretty much married to the lab and this marriage wouldn't have lasted this long without YOU. I love you all, and I will value you for the rest of my life.

My friends in Ann Arbor, other parts of the country and from back home, you all have been the greatest pillar of support and I seriously couldn't have

asked for more. You all (you know who you are :)) are my extended family and I couldn't have reached this point without you. I am extremely thankful to you all for all your support, for listening to me and putting up with me when I was happy, sad, angry, disappointed, wild and crazy. Without you'll I don't think I would have lasted this long in a foreign land, miles and miles away from my family. I extremely grateful for our friendship. You all simply are the best.

Last but not the least, I would like to thank my ever supportive family- My dad, mom and sister for their unfailing prayers, encouragement, motivation, support, and faith in me. You are the most divine people in the world and I love you so very much that no words can express it. I am very thankful to you for mustering up the courage to send me miles and miles away from home to pursue my graduate studies. I love you for providing me with only the best things in the world. I would like to thank my in-laws for their continued support and encouragement during my graduate studies. Finally, my beloved friend, mentor, colleague, husband and my better half, Sethu Pitchiaya who has been beside me every waking moment in my life for the last 10 years- I am extremely thankful to you for being there for me and with me. I am very certain that I couldn't have done this without you in my life. I am very thankful and indebted to you for supporting me through my graduate studies and being a great life partner. I look forward to many more years of pure awesomeness with you. Love you lots and thank you for everything.

Table of Contents

List of Figures

List of Appendices

Abstract

An investigation of the RNA induced silencing complex

and its therapeutic implications

By

Vishalakshi Krishnan

Chair: Nils G. Walter

Our understanding of the central dogma of biology has changed greatly since the discovery of non-protein coding RNAs that play key regulatory roles in protein expression. The research presented in this thesis is an investigation of a highly conserved type of regulatory RNA, small interfering RNA (siRNA), which regulates gene expression at the post-transcriptional level through the RNA interference (RNAi) pathway. RNAi is a conserved gene regulatory mechanism employed by most eukaryotes as a key component of their innate immune response to viruses and retrotransposons, wherein exogenous long dsRNAs are cleaved (diced) into ~21-28nt long siRNA duplexes with characteristic 5′-phosphate and 3′-dinucleotide termini that are subsequently loaded into the effector RNA Induced Silencing Complex (RISC). The mature duplex siRNA is then unwound and activated in a strand-specific and ATP-dependent manner. Upon recognition of a target mRNA, the siRNA guides RISC to initiate sequence

x

specific mRNA cleavage (slicing) by the Argonaut (Ago) protein, thus inhibiting the production of a specific protein.

As a countermeasure to RNAi, many viruses have evolved viral RNA silencing suppressors (RSS) that tightly, and presumably quantitatively, bind siRNAs to thwart RNAi-mediated degradation. Viral RSS proteins also act across kingdoms as potential immunosuppressors in gene therapeutic applications. Here we report fluorescence quenching and electrophoretic mobility shift assays that probe siRNA binding by the dimeric RSS p19 from Carnation Italian Ringspot Virus, as well as by human Dicer and RISC assembly complexes. We find that the siRNA:p19 interaction is readily reversible and p19 efficiently competes with recombinant Dicer and inhibits the formation of RISC-related assembly complexes found in human cell extract. Computational modeling based on these results provides evidence for the transient formation of a ternary complex between siRNA, human Dicer, and p19. An expanded model of RNA silencing indicates that multiple turnover by reversible binding of siRNAs potentiates the efficiency of the RSS protein.

siRNAs have yielded powerful insights into biological function through reverse genetic approaches and have given birth to a new field of gene therapeutics. A number of questions are outstanding concerning the potency of siRNAs, necessitating an understanding of how short double-stranded RNAs are processed by the cell. Recent work suggests that unmodified siRNAs are protected in the intracellular environment, although the mechanism of protection still remains unclear. We have developed a set of doubly-fluorophore labeled RNAs (more precisely, RNA/DNA chimeras) to probe in real-time the stability of siRNAs and related molecules by fluorescence resonance energy transfer (FRET). We found that these RNA probes are substrates for relevant cellular degradative processes, including the RNase H1 mediated degradation of an DNA/RNA hybrid and Dicer-mediated cleavage of a 24-nucleotide (per strand) double-stranded RNA. In addition, we found that 21- and 24-nucleotide double-stranded RNAs are relatively protected in human cytosolic cell extract, but less

so in blood serum, whereas an 18-nucleotide double-stranded RNA is less protected in both fluids. These results suggest that RNAi effector RNAs are specifically protected in the cellular environment and may provide an explanation for recent results showing that unmodified siRNAs in cells persist intact for extended periods of time.

Despite the vast application of RNAi in functional genomics and therapeutics, there still are also several outstanding questions about the composition of siRISC. Although it is known that the minimal RISC is composed of Dcr, TRBP and Ago2, the roles played by other Ago2 and RISC associated proteins are still under debate. Much of the complexity associated with assigning specific roles to the pathways components arises from the inherent complexity of the pathway itself and from the inability to develop a comprehensive system of study for RNAi. Our results demonstrate that the two systems we and others have employed, i.e., the HeLa whole cell extract and the HeLa S100 cytoplasmic extract, utilize siRNA duplexes versus single-stranded antisense siRNA with differential efficiency in cleaving a target mRNA, despite their uniform ability to perform slicing. We have shown that the ability of HeLa cytoplasmic extract to utilize siRNA duplexes better than single-stranded antisense siRNA may be due to elevated concentrations of two RNA helicases, MOV10 and RHA. Additionally, we have shown that ATP has an inhibitory effect on the target cleavage mediated by HeLa cytoplasmic extract. Our results suggest a potential role for MOV10 and RHA in siRNA duplex loading of RISC. We also hypothesize that ATP may be mediating some yet-to-be-identified allosteric and/or structural change in RISC that has an inhibitory effect on its target mRNA cleavage activity. Taken together, the work presented in this thesis has furthered our understanding of how siRNAs affect functional genomics and therapeutic applications.

Chapter 1

RNA interference: The nuts and bolts

1.1 Introduction

The advances in molecular biology since the early 1900s have been staggering and have enticed scientists from all over the world to work towards understanding life at its most fundamental level. The breadth of research conducted thus far in this field has gone to prove the existence of ever more complex molecular processes in increasingly complex organisms.

The beginning of the golden era of molecular biology was marked by the demonstration of a linear relationship between genes (DNA) and proteins by George Beadle and Edward Tatum in 1940[1], followed by the discovery of DNA as the genetic material by the famous Avery-MacLeod-McCarty transformation experiment in 1944[2], and the Hershey-Chase hereditary experiment in 1952[3]. The seminal work by Rosalind Franklin influenced the work of James D. Watson and Francis Crick, leading to the recognition of DNA as a double helix in 1953[4]. This discovery, followed by the influential presentation by Crick in 1957[5] that the Central Dogma of Biology is a sequential relationship between DNA, RNA and protein, had a major impact on the advancement of molecular biology.

The Central Dogma of Biology describes the flow of genetic information to be sequential residue- by-residue and unidirectional, i.e., the genetic information stored in the double-stranded DNA molecule is translated via single-stranded ribonucleic acid (RNA, specifically messenger RNA, mRNA) molecules with the help of ribosomes to proteins[6]. It was believed that the main role of RNAs within a cell was to code for proteins and hence most research done by molecular biologists was focused on understanding the structure, function and interactions between the two macromolecules: Deoxyribonucleic acids (DNA) and Proteins.

Ever since, our understanding of this dogma and the role of RNA within the cell have evolved significantly, owing to the seminal discovery of ribozymes (RNAs with catalytic properties) by Thomas Cech and Sidney Altman in the 1980s[7, 8] as well as the discovery, made during the Human Genome Project in 2000[9, 10], that approximately 98.5% of the human genome does not code for protein[11]. Additionally, it has been noted that the number of non-protein coding RNAs, but not as much the number of protein coding mRNAs, increases with increasing complexity of the organism (**Fig. 1.1**) [12, 13]. Since then a number of non-coding RNA classes have been discovered that play either a structural, catalytic or regulatory role, such as transfer RNA (tRNA), ribosomal RNA (rRNA), short interfering RNA (siRNA), microRNA (miRNA), Piwi-interacting RNA (piRNA), small nuclear RNA (snRNA), small nucleolar (snoRNA) and others[11]. These non-coding RNAs have critical impact on several diseases such as cancer, diabetes, Alzheimer's, autism etc[14]. Despite the fact that non-coding RNAs have been studied extensively over the last decade, as evident from the exponential increase in the number of PubMed entries for research articles on non coding RNAs, (**Fig. 1.2**) the mechanism of action of most of them still needs to be brought to light. My research in the Walter Lab at the Department of Chemistry, University of Michigan, has focused on monitoring key steps in the mechanism of action of one such type of regulatory non-coding RNA called small interfering RNA (siRNA).

1.2 Non coding RNAs: The face of the current era of Molecular Biology

Non-coding RNAs (ncRNAs) are ribonucleic acids that are transcribed from DNA and predominantly do not code for proteins. While the initial assumption was that DNA that was transcribed into RNA is always translated into protein and DNA not coding for protein was considered evolutionary 'junk', recent findings suggest that the majority of the human genome is in fact transcribed into non-protein coding RNAs or ncRNAs that contain information or have a function[12]. ncRNAs can be either transcribed as their own genes (RNA genes), or can be derived from mRNA introns, appended to an mRNA as untranslated regions (UTRs), alternatively spliced and/or processed into smaller products[11].

ncRNAs predominantly are single-stranded stretches of nucleotides that may contain self-complementary sequences enabling them to form highly complex structures[15]. Like proteins, ncRNAs can assume specific secondary structures such as short helices, hairpin loops, bulges and internal loops[16]. These secondary and tertiary structures are stabilized by metal ions such as Mg^{2+} [17] that help ncRNAs achieve functional forms similar to proteins[18, 19]. Over a decade, ncRNAs have been found to be involved in a plethora of cellular processes that include protein synthesis (transfer RNA, ribosomal RNA), gene regulation (short interfering RNA, microRNA), RNA processing (ribozymes), to name a few[19, 20]. These ncRNAs have been grouped into diverse classes of RNA molecules with structural, catalytic, or regulatory functions, ranging in size from alternatively spliced transcripts of many kilo-bases (kb) in length to 22-nucleotide small RNAs[11]. Additionally, several other ncRNAs exist whose functions within the cell still remain unknown.

Structural ncRNAs that have been known for a while now include transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) that play a role in mRNA translation, small nuclear RNAs (snRNAs) or spliceosomal uRNAs that are involved in splicing, and the common small nucleolar RNAs (snoRNAs) that are involved in the modification of rRNAs. These structural ncRNAs play an important
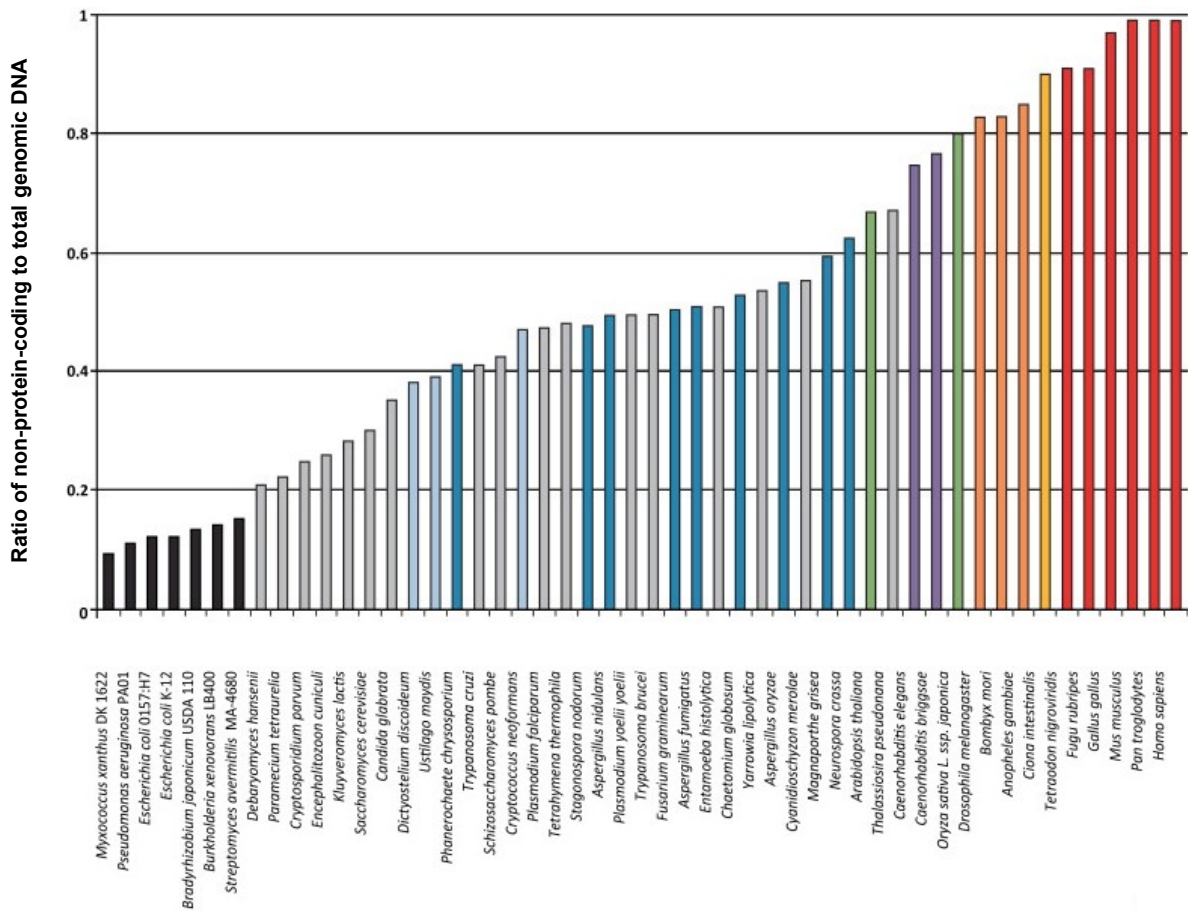
Figure 1.1. The fraction of non-protein-coding DNA increases with the increasing complexity of the organism. Plot illustrating the percent ncDNA per sequenced genome across phyla. The prokaryote and bacterial genomes are depicted in black. Single-celled organisms are shown in gray, organisms known to be both single and multi-cellular depending on lifecycle are light blue, basal multi-cellular organisms are blue, plants are green, nematodes are purple, arthropods are orange, chordates are yellow, and vertebrates are red. Reprinted with permission from John Wiley and Sons[13].
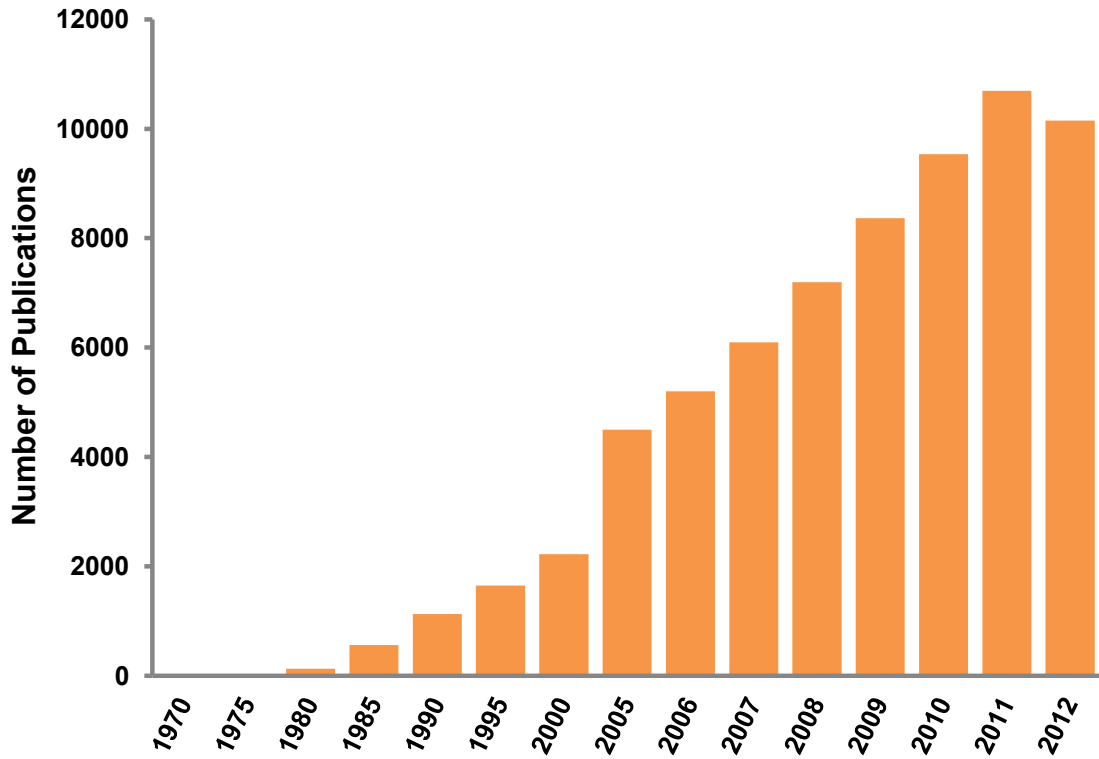
Figure 1.2. Increase in amount of research on non-coding RNAs. Graph depicting the exponential increase in the number of publications on non-coding RNAs over the last few decades.

role in catalytic process in addition to sequence specific recognition of RNA substrates[21, 22, 23, 24, 25]. Two major rRNAs constitute about 60% of the total weight of ribosome, which is composed of a large and a small subunit made of proteins and RNAs. The rRNAs have a structural role, acting as a scaffold defining the positions of the ribosomal proteins and aiding in the binding of the two ribosomal subunits. The 40S small ribosomal subunit contains one 18S rRNA and 33 proteins, whereas the three rRNAs (5S, 5.8S and 28S) that comprise the catalytic core[26, 27, 28, 29] are present, along with 46 other proteins, in the 60S large subunit of the ribosome, which acts as a ribozyme and catalyzes peptide bond formation[30, 31]. The mRNA to be translated is sandwiched between the two subunits of the ribosome, where the small subunit binds the mRNA and monitors the interaction of the tRNA anticodon, whereas the mRNA codon and the large subunit houses the three binding sites, A, P and E, for tRNAs[32]. During translation, the A site is occupied by an aminoacyl-tRNA (a tRNA that contains a complementary anticodon on one end and is charged with an amino acid on the other end), which attacks the ester linkage of the peptidyl-tRNA present within the P site to form a new peptide bond by peptidyl transfer. Subsequently, the attacked tRNA, now uncharged, exits the ribosome through the E site. Based on a similar principle of ordered assembly, snRNAs (commonly referred to as U-RNAs owing to their high uridine content) associate with several proteins to form the spliceosome[33], which is the complex responsible for removal of introns from pre-mRNA within the nucleus, through the post-transcriptional process called splicing. Each snRNA (U1, U2, U4, U5 and U6) associates with a set of specific protein and is referred to as small nuclear ribonucleoproteins (snRNP) complex (often pronounced "snurps")[34]. In addition to being involved in the assembly of spliceosome, U1 snRNP is also responsible for initiating spliceosomal activity by base pairing with the pre-mRNA, which has also been shown to protect pre-mRNA from polyadenylation as well as premature cleavage[35]. Besides splicing, U1 snRNA has also been implicated in regulation of transcription initiation by RNA polymerase II through interaction with the transcription initiation factor TFIIH and possibly in cell cycle regulation[36, 37].

Several other ncRNAs, for example, ribozymes and riboswitches exist that have functional roles in various cellular processes. Ribozymes or RNA enzymes are small ncRNA molecules with the ability to catalyze a specific biochemical reaction, similar to protein enzymes. The first ribozymes discovered in 1980s by Thomas R. Cech and Sidney Altman were the self-splicing Group I intron of an RNA transcript in Tetrahymena thermophila[38] and the RNA component of the bacterial enzyme RNase P that is involved in the maturation of pre-tRNAs[39], respectively. Since then, ribozymes have been found to be involved in a variety of cellular processes including protein synthesis as part of the large subunit rRNA, and in RNA processing such as RNA splicing, viral replication and tRNA biosynthesis. Examples of small ribozymes include the hammerhead ribozyme, the hepatitis delta virus (HDV) ribozyme, the Varkud satellite (VS) ribozyme and the *glmS* ribozyme[130]. The group of hammerhead ribozymes are ubiquitous catalytic RNAs[40, 41] that catalyze site-specific reversible self-cleavage reactions within RNA molecules[42]. They occur in the intronic regions of a few specific genes and point to a conserved biological role during pre-mRNA biosynthesis[43]. Through a similar mechanism of self-cleavage within the RNA molecule, the HDV ribozyme achieves monomerization of longer RNA transcripts synthesized during the rolling circle replication of the viral HDV genome; it is thought to be the only catalytic RNA required for viability of a human pathogen[44].

Riboswitches, regulatory segments mostly found upstream of the coding sequence in mRNA transcripts, bind small molecules (ligands) that control the expression of downstream genes, thereby facilitating the self-regulation of mRNA abundance. Riboswitches adopt various mechanisms to mediate gene regulation[45], which includes ligand induced formation of terminator hairpins that lead to transcription termination (e.g., PreQ$_1$-I riboswitch in *Bacillus subtilis*)[46], ligand mediated structural changes that inhibit translation by sequestering the ribosome-binding site (e.g., PreQ1-II riboswitch found in certain species of *Streptococcus*)[47], ligand mediated alternate structure formation of pre-mRNA that affects splicing, thereby affecting the expression of downstream genes (e.g., thiamin pyrophosphate riboswitch)[48], and ligand mediated ribozyme-like self-

cleavage (e.g., by the *glmS* riboswitch)[49] of mRNAs. The glucosamine-6-phosphate dependent *glmS* riboswitch is an ncRNA that is both a ribozyme, catalyzing self-cleavage, and a riboswitch regulating genes in response to binding glucosamine-6-phosphate (GlcN6P)[50]. The *glmS* ribozyme is present in the 5' untranslated region (UTR) of the mRNA that codes for the *glmS* enzyme, which is responsible for GlcN6P biosynthesis in bacteria[51]. The ribozyme binds GlcN6P, which causes self-cleavage of the mRNA backbone leading to its degradation and, consequently, reduced synthesis of the *glmS* enzyme encoded by the mRNA. The ribozyme thus employs a negative feedback loop to regulate GlcN6P biosynthesis when it senses an already high level within the cell[52, 53]. Other examples of riboswitches include the flavin mononucleotide (FMN) riboswitch, the cobalamin riboswitch, the lysine riboswitch and the S-adenosyl methionine riboswitch that bind their respective metabolite to regulate its biosynthesis and transport, and the glycine riboswitch, the glutamine riboswitch and the purine riboswitch that regulate genes involved in the metabolism and transport of their respective metabolite binding partners[45, 54, 55].

The expression of most genes is regulated by *trans*-acting ncRNAs, i.e., regulatory RNAs that function by base-pairing with complementary sequences in other RNAs and DNA, to form RNA:RNA or RNA:DNA duplexes, or *cis*-acting ncRNAs that form secondary structures upon binding of ligands and are recognized, and acted upon, by regulatory complexes. The 5`UTRs, 3`UTRs, and intronic regions in pre-mRNAs contain *cis*-acting regulatory ncRNAs, such as riboswitches, internal ribosome entry sites (IRES) and RNA leader sequences, that receive *trans*-acting signals by forming secondary structures that recruit regulatory proteins[56, 57, 58]. The UTRs also act as sensors of *trans*-acting regulatory ncRNAs, like miRNAs that act through partial sequence complementary to one or more messenger RNA (mRNA) molecules, generally in 3' UTRs[59, 60]. Interestingly, it was noted that the average length of UTRs in mRNAs and the fraction of alternatively spliced genes increases with developmental complexity in animals, indicating the existence of increased and sophisticated mRNA regulation in the higher organisms[61, 62]. Another group of

regulatory ncRNAs includes siRNAs that are similar to miRNAs but use the RNA interference (RNAi) cellular pathway to control gene expression through a different, yet related mechanism[63]. siRNAs, miRNAs and RNAi are discussed in more detail in the subsequent sections.

In summary, the discovery of functional ncRNAs not only rectified the fundamental misconception of genetic programming in higher organisms but also brought to light the existence of an enormous repertoire of ncRNAs and a sophisticated RNA-based regulatory system. Recently a number of short and long ncRNAs have been discovered, which constitute another previously hidden layer of molecular genetic signals, helping us gain a better understanding of the complexities involved in human transcriptome expression and the tremendous diversity of characteristics among mammals and other complex organisms. The work presented in this thesis will outline how studies of the interaction of a class of ncRNAs (siRNA) with either viral RNA silencing suppressor (RSS) or RNA induced silencing complex (RISC) proteins, and of chemical modification of ncRNAs provide important insight into the function and applications of this important class of ncRNAs and the underlying RNAi pathway mechanisms.

1.3 History and mechanism of RNA interference

RNA interference (RNAi) is a fundamentally important regulatory mechanism in which small RNAs enable specific regulation of gene expression at the RNA level. Found in fungi, plants, and animals, RNAi is thought to have evolved to combat viral infections and has had a tremendous impact on fundamental and applied research. RNAi is currently one of the most promising new approaches for disease therapy and RNA silencing by miRNAs has been shown to play a fundamental role in chromatin remodeling, genome rearrangement, developmental timing, brain morphogenesis, and stem cell maintenance[64, 65, 66, 67, 68, 69, 70].

RNAi was first observed in transgenic petunia plants[71] as a phenomenon wherein small ncRNA fragments act as triggers to either repress or destroy complementary mRNA targets, which over time was collectively called post-transcriptional gene silencing (PTGS). In 1998, Craig C. Mello and Andrew Fire were the first to identify double-stranded RNA as the causative agent for this phenomenon and termed it RNAi[72]. Since then a lot of research has been performed to elucidate the RNAi mechanism. It has been found that two types of small RNA molecules – exogenous siRNA and endogenous miRNA - are central to a pair of closely related pathways, RNAi and RNA silencing, respectively.

miRNAs involved in RNA silencing are *trans*-acting regulatory ncRNAs that are expressed endogenously from genomic DNA and processed into ~20-30nt long ncRNAs that bind to short complementary seed sequences (6-8 nt) in the 3'UTRs of target mRNAs, thereby regulating mRNA gene expression at the post-transcriptional level[73]. It has been predicted computationally that over 50% of human genes are controlled by ~1,500 different miRNAs through an adaptive, complex, regulatory network where a single miRNA controls many different (>100) mRNAs and a single mRNA often has multiple binding sites for the same or different miRNAs[74, 75]. miRNA biosynthesis is tightly regulated[74] wherein primary transcripts (pri-miRNAs) are transcribed by RNA polymerase II (Pol II) or Pol III containing either a single miRNA or a cluster of miRNAs with one or several stem-loop structures containing imperfectly base-paired stems[76, 77]. These pri-miRNAs are then cleaved into precursor-miRNAs (pre-miRNA), which are ~70 nt long and contain a single stem-loop with a 2-nt overhang on their 3'-end, by a protein complex containing the RNase III-type endonuclease Drosha and the microprocessor DGCR8. These pre-miRNAs are then transported by Exportin-5 (Xpo-5) via the nuclear pore complex (NPC) to the cytoplasm, where they are cleaved by a Dicer containing complex into ~22 nt long mature miRNA duplexes with 2-nt overhangs on the 3'-end of each strand. As a part of the miRNA-induced silencing complex (miRISC), one strand of the mature miRNA duplex, i.e., the guide strand, base pairs with the partially complementary target mRNA to mediate its repression[78] and the passenger strand is cleaved. In cases

where the miRNA is perfectly complementary to the mRNA target, there is complete knockdown of gene expression resulting from the endonucleolytic cleavage of the target mRNA, as expected for siRNAs[79] (**Fig. 1.3**).

By contrast, exogenous long double-stranded RNAs (dsRNAs) that result from RNA virus replication, convergent transcription of cellular genes, mobile genetic elements, self-annealing transcripts, or experimental transfection, supply substrates for the siRNA mediated RNAi pathway[80]. siRNAs are another class of *trans*-acting regulatory ncRNAs that bind the target mRNA with complete complementarity, leading to its endonucleolytic cleavage, thereby regulating its gene expression at the post-transcriptional level. The exogenous long dsRNAs are cleaved into siRNA duplexes, ~21-28 base pairs (bp) in length, with characteristic 5′-Phosphate modification and 3′-dinucleotide overhang, by a complex that contains an endogenous RNase III-type enzyme called Dicer (Dcr)[81, 82, 83, 84]. These siRNA duplexes are thought to be subsequently loaded onto an effector RNA Induced Silencing Complex[85] (RISC) that contains Argonaute (Ago), an RNase H-type protein. The siRNA duplexes are then unwound and activated in a strand-specific and ATP-dependent manner within RISC[86]. In humans, it has been proposed that the siRNA duplex is loaded first into the RISC-loading ternary complex, consisting of Argonaute, Dicer and TAR-RNA binding protein (TRBP)[87] that then further matures. Upon engagement of a target mRNA, the unwound single-stranded siRNA (namely the guide strand) guides siRISC to initiate sequence specific mRNA (slicing) and degradation[88], thus stopping the production of a specific protein (**Fig. 1.3**).

During strand unwinding and strand selection of the siRNA duplex, the RNA strand with the thermodynamically less stable 5′-end is preferentially loaded into RISC as the guide strand[89, 90] and the passenger strand is cleaved. Passenger strands are not always byproducts of siRNA biogenesis; they can also be loaded into siRISC to function as siRNAs[91, 92]. In *Drosophila,* strand selection is thought to be mediated by R2D2, which binds to the more stable end of siRNAs as they are loaded into RISC[93]. R2D2 is a *Drosophila* protein that has
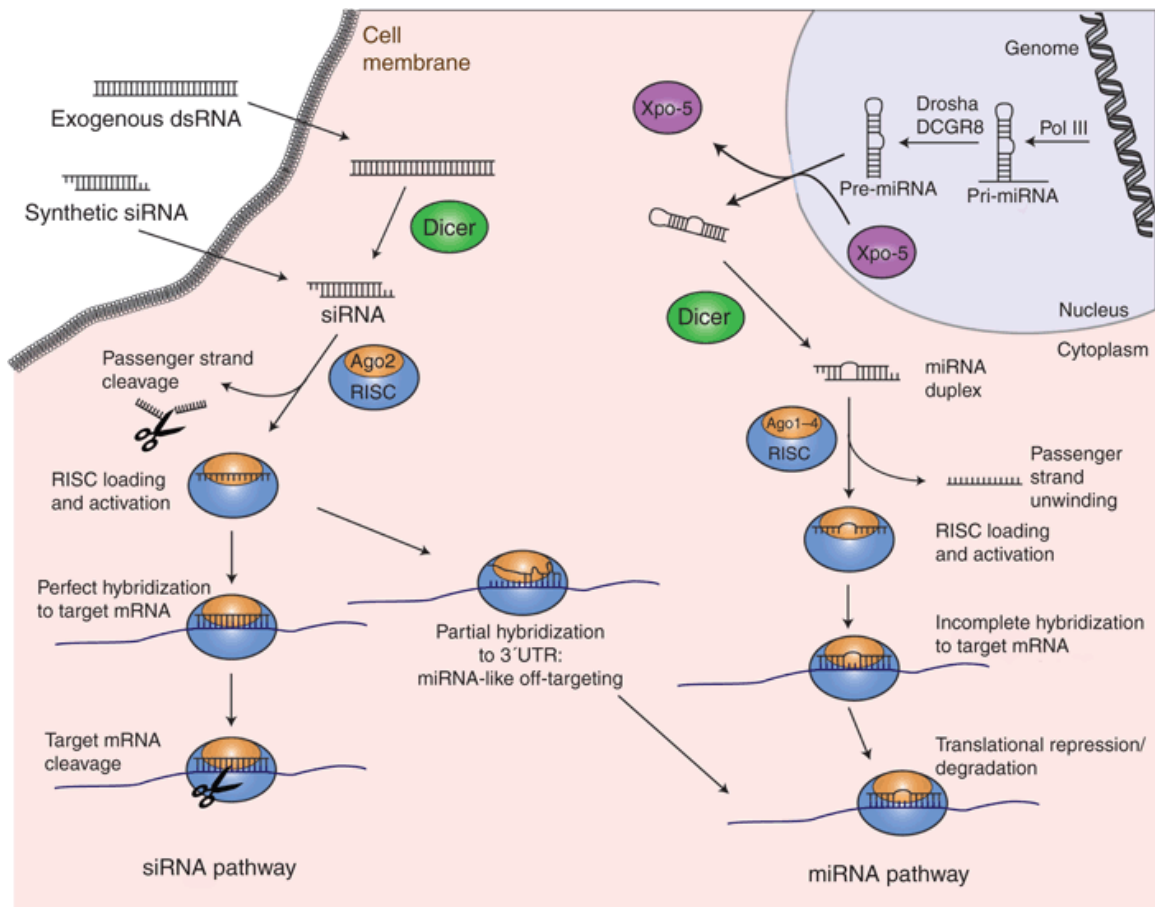
Figure 1.3. A graphical representation of the siRNA and miRNA pathway.
The biogenesis of siRNA and miRNAs and their mechanism of regulation of gene
expression. Reprinted with permission from Nature Publishing Group[94].

two dsRNA-binding domains (R2) and is associated with DCR-2 (D2). It has been shown that in addition to thermodynamic parameters, the processing polarity (i.e., the end from which the long dsRNA substrate is cleaved) of dsRNA by Dicer can also play a role in guide strand selection[95]. In the absence of thermodynamic instability, the RNA strand that has its 3'-end bound to Dicer has been shown to be preferentially incorporated as the guide strand[95], thus making it equally likely for either strand of the siRNA duplex to affect RNAi in this context.

Despite our expanding knowledge in this field, the composition of the effector RISC still remains under debate. Although active RISCs of varying sizes ranging from ~160-500 kDa have all been shown to cleave mRNA targets [86, 88, 96], it is still unknown if these complexes represent the complete active RISC or if they just contain the core proteins responsible for the RNAi activity. *In vitro* experiments with *Drosophila* embryo extracts resulted in a model for RISC assembly in which complexes were seen to form sequentially, containing Dicer both at the beginning and end, and ultimately causing target mRNA cleavage[97]. These findings suggest an extended role for Dicer beyond the initiation phase of RNAi, at least in the fruit fly. Human Dicer (hDcr) in complex with TRBP has been shown to interact with Ago proteins[98, 99], where Dicer binds to the PIWI-box of Ago proteins through one of its RNase III domains[100]. Ago has been shown to be present at the heart of silencing effector complexes with single-stranded siRNA/miRNA[101, 102] and in pre-miRNA processing complexes that transfer the miRNA into a target-mRNA cleaving complex[103]. Several proteins have been shown to be associated with Ago proteins either directly or through adaptor proteins that mediate strand selection, strand unwinding, passenger strand removal and repression either by stimulating mRNA decay or translational repression[99, 103, 104]. The work presented in Chapter 4 of this thesis will focus on identifying the functional role of a couple of previously identified Ago associated proteins in the RNAi pathway.

1.4 Viral suppressors of RNA silencing

RNAi is a conserved sequence-specific gene regulation mechanism that also functions as an adaptive inducible antiviral defense against invasive foreign nucleic acids[105]. In addition to exogenous virus derived siRNAs, it has been demonstrated that mammalian cells have endogenous cellular miRNAs that can sequester and degrade viral mRNAs through the related RNA silencing pathway[106]. Several endogenous anti-viral miRNAs exist, including miR-28, miR-125b, miR-150, miR-223, miR-382 and miRNA-29a, that target different regions of the human immunodeficiency virus-1 (HIV-1) mRNAs[107]; miR-199a and miR-32 restrict Hepatitis C virus (HCV) and primate foamy virus type 1 (PFV-1) replication, respectively[108, 109]; and miR-100 and miR-101 function against human cytomegalovirus[110].

Pathogenic viruses have developed, in turn, RNA silencing suppressor (RSS) mechanisms to counteract the host anti-viral defense. This phenomenon was first observed in plant viruses such as potato virus and cucumber mosaic virus[111]. In addition, it has been predicted that most mammalian viruses, i.e., HIV, influenza A virus, HCV, Vaccinia, Ebola virus and PFV-1, that elicit host anti-viral defenses contain viral RSS proteins that harbor dsRNA binding domains and GW/WG motifs in their sequences[112, 113]. RSS proteins of plant viruses have been well characterized and shown to target most steps of the RNAi pathway, including viral RNA recognition, dicing, RISC assembly, RNA targeting and amplification[114](**Fig. 1.4**). Several viral RSS proteins have been identified that inhibit RNAi, including the p14 protein of Pothos latent aureusvirus and p38 of Turnip crinkle virus, which bind dsRNAs size-independently and inhibit their processing to siRNAs[115]; the p19 protein of Tombusviruses, the 2b protein of Tomato aspermy cucumovirus and the B2 protein of Flock house virus, which all sequester siRNAs by binding double-stranded siRNA with high affinity in a sequence non-specific, but size-specific manner[116]; and the 2b protein of Cucumber mosaic virus (CMV) that interacts with Ago to inhibit slicing activity and prevent the spread of the long-range silencing signal[117]. Thus, the viral RSS

proteins often target one or more conserved core components of the RNAi pathway such as siRNAs, Ago and Dcr proteins to suppress the pathway in host cells.

Understanding the activity of viral RSS proteins has important implications both in basic research that employs siRNA/miRNA to study viral infection and pathogenesis and in siRNA/miRNA/shRNA based gene therapy approaches. Deciphering the mechanism of RSS proteins will make them potential targets for future therapies and drug development. To better understand the mechanism of sequestration of siRNA, the research presented in Chapter 2 of this thesis focuses on understanding the siRNA binding and dissociation kinetics by one such viral RSS protein, p19, found in Carnation Italian Ringspot Virus (CIRV) and other tombusviruses. p19 is a 19 KDa protein that functions as a homodimer and binds siRNA duplex in a positively charged surface cleft. It has a pair of tryptophan residues that stack on both terminal base pairs of an siRNA and facilitates a sequence-independent, caliper-like size selection for siRNA duplexes[118]. p19 also has been shown to inhibit siRNA mediated target mRNA cleavage and RISC assembly in Drosophila embryo extracts[116].

1.5 RNAi therapeutics

RNAi based therapeutics has already had a transformational effect on modern medicine and offers significant potential for treatment against potentially any target disease. A lot of research is currently being performed to validate the use of siRNAs as therapeutic agents. Synthetic siRNAs or short hairpin RNAs (shRNAs) leverage the host RNAi machinery to reduce expression of any pathological molecular RNA target with high selectivity, which is impractical to achieve with pharmaceutical approaches involving small molecules. The major challenges that concern the use of siRNA duplexes for therapeutic applications are their delivery, stability, specificity and potency[119]. siRNAs are extremely hydrophilic macromolecules with exposed sugar-phosphate backbone and are
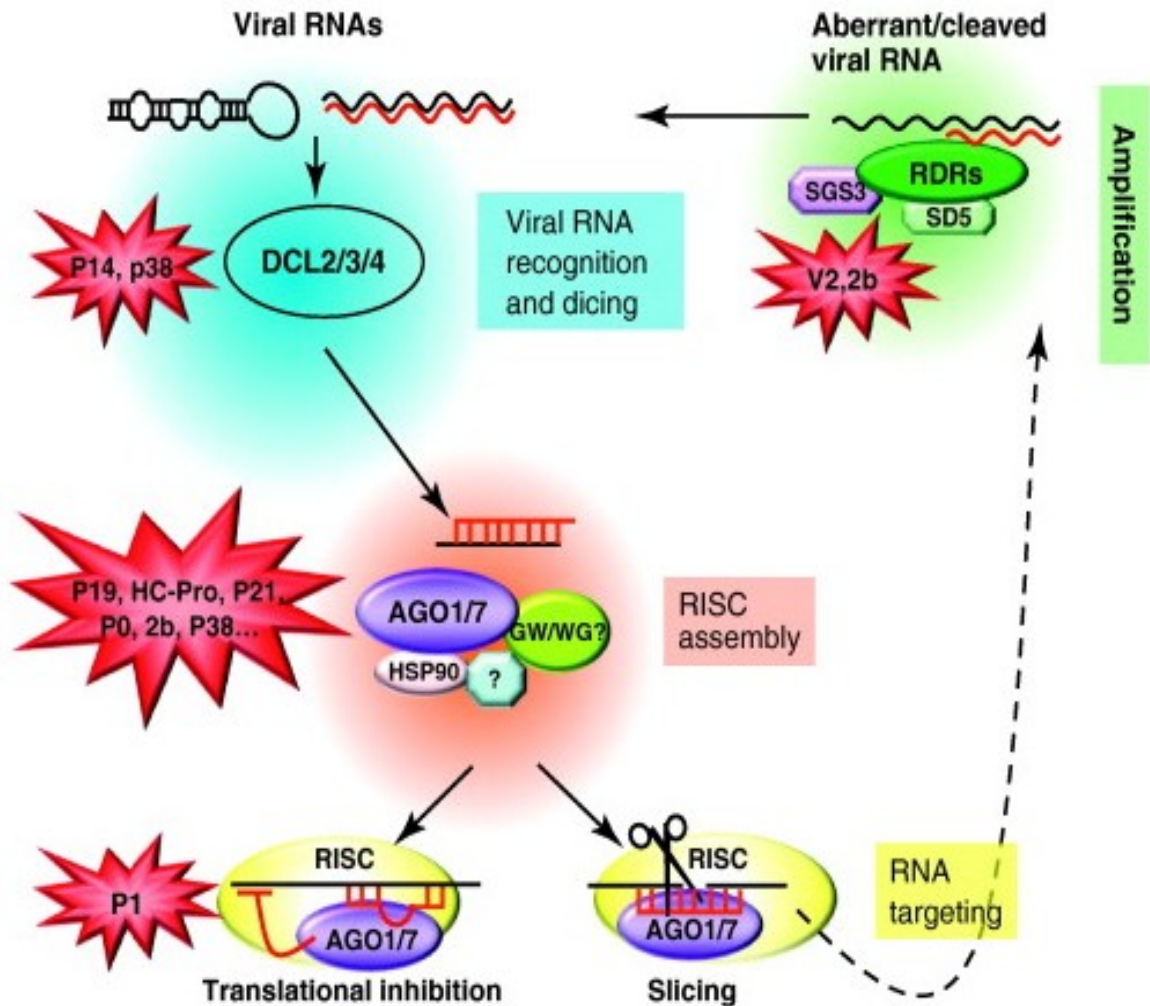
Figure 1.4. A model illustrating the steps of the RNA silencing pathway and their suppression by various RSS proteins. The Viral-silencing suppressor proteins can inhibit the RNAi at multiple points in the pathway, thereby preventing the anti-viral host response. The model represents the various steps at which the RSS proteins (i.e. P14, P38, V2, 2b, P19, HC-Pro, P21, P0 and P1) interact with the silencing pathways. Reprinted with permission from Elsevier[114].

highly susceptible to degradation by serum nucleases[120]. However, it has been demonstrated by several research groups that introducing specific chemical modifications to synthetic siRNAs renders them both stable and of increased *in vitro* potency, depending on the environment of the siRNA[121, 122, 123]. Such modifications include a 3′-phosphorothioate backbone linkage for exonuclease resistance and 2′-O-methyl, 2'-O-methoxyethyl and 2′-fluoro modifications (2′-OMe, 2′-O-MOE and 2′-F, respectively) for endonuclease resistance[123]. It was also observed that modifications at the 5'-end of the guide strand was less well tolerated than those at the 3'-end; that 2′-F modification on the sugar is well tolerated on the guide strand compared to 2′-O-MOE modification; and that the incorporation of 2′-O-Me or 2′-O-MOE in the passenger strand does not affect the activity of the siRNA[124]. Another modification that was tolerated on both strands in moderation was a 4′-thioribose modification[125], which adopts conformations very similar to the C3′-endo pucker observed for unmodified sugars and increases the potency of siRNAs when used both individually and in combination with 2′-O-Me and 2′-O-MOE modifications[126]. Off-target effects due to partial homology with the siRNAs were shown to be considerably reduced by introducing a 2′-O-Me modification at a single nucleotide position (position 2) of the guide strand[127]. Although chemical modifications have been shown to clearly stabilize siRNAs and improve their potency, the work presented in Chapter 3 of this thesis along with a few pieces of evidence in the literature suggest that siRNAs are relatively protected within the cell and chemical modification may not be imperative after all[128, 129]. This protection of siRNA may be due to binding of the siRNA by RISC or other RNAi components, which keeps it largely hidden from nucleolytic decay either as the base-paired siRNA duplex or as the single guide strand processed into active RISC.

1.6 Conclusion

RNAi and for that matter siRNAs have found tremendous application in various areas of biomedicine, ranging from research based applications for gene

knock down, functional genomics and genome-scale RNAi screens to biotechnology and medical application as therapeutics. This extensive application of siRNAs in several arenas only bolsters the need for a better understanding of the RNAi pathway and its components. The research presented in this thesis sheds light on several aspects of the pathway that have significant therapeutic implications.

1.7 References

1.   Beadle, G. W., Tatum, E. L. (1941). Genetic Control of Biochemical Reactions in Neurospora. *Proceedings of the National Academy of Sciences* **27** (11), 499–506.

2.   Avery, O. T., MacLeod, C. M., McCarty, M. (1944). Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Deoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III. *Journal of Experimental Medicine* **79** (2), 137–158.

3.   Hershey, A.D. and Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteriophage. *Journal of General Physiology* **36** (1), 39-56.

4.   Watson, J. D. and Crick, F. H. C. (1953). A Structure for Deoxyribose Nucleic Acid. *Nature* **171** (4356), 737-738.

5.   Crick, F.H.C. (1958). On Protein Synthesis. *The Symposia of the Society for Experimental Biology XII* 139-163.

6.   Crick, F. H. C. (1970). Central dogma of molecular biology. *Nature* **227** (5258), 561–563.

7.   Altman, S. et al. (1989). Catalysis by the RNA subunit of RNase P--a minireview. *Gene* **82** (1), 63-64.

8.   Cech, T. R. (1990). Self-splicing of group I introns. *Annual Reviews of Biochemistry* **59,** 543-568.

9.   Lander, E. S., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

10. Venter, J. C., et al. (2001). The sequence of the human genome. *Science* **291**, 1304-51.

11. Mattick, J. S., and Makunin, I. V. (2006). Non-coding RNA. *Human Molecular Genetics* **15** (1), R17-29.

12. Mattick, J.S. (2004). The hidden genetic program of complex organisms. *Scientific American* **291** (4), 60-67.

13. Taft, R. J., Pheasant, M. and Mattick, J. S. (2007). The relationship between non-protein-coding DNA and eukaryotic complexity. *Bioessays* **29** (3), 288–299.

14. Robinson, R. (2004). RNAi Therapeutics: How Likely, How Soon?. *PLoS Biology* **2**(1), e28.

15. Tinoco, I. and Bustamante, C. (1999). How RNA folds. *Journal of Molecular Biology* **293** (2), 271–281.

16. Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M. and Turner, D. H. (2004). Incorporating chemical modification constraints

into a dynamic programming algorithm for prediction of RNA secondary structure. *Proceedings of the National Academy of Sciences* **101** (19), 7287–7292.

17. Tan, Z. J. and Chen, S. J. (2008). Salt dependence of nucleic acid hairpin stability. *Biophysical Journal* **95** (2), 738–752.

18. Higgs, P. G. (2000). RNA secondary structure: physical and computational aspects. *Quarterly Reviews of Biophysics* **33** (3), 199–253.

19. Nissen, P., Hansen, J., Ban, N., Moore, P. B. and Steitz, T. A. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science **289** (5481), 920–930.

20. Rossi, J. J. (2004). Ribozyme diagnostics comes of age. *Chemistry & Biology* **11** (7), 894–95.

21. Gesteland, R. F., Cech, T. R. and Atkins, J. F. (2006). The RNA World, 3rd edition. *Cold Spring Harbor Laboratory Press*.

22. Steitz, T. A. and Moore, P. B. (2003). RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends in Biochemical Sciences* **28**, 411–418

23. Noller, H. F. (2005). RNA structure: reading the ribosome. *Science* **309**, 1508–1514.

24. Nilsen, T. W. (2003). The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* **25**, 1147–1149.

25. Butcher, S. E. and Brow, D. A. (2005). Towards understanding the catalytic core structure of the spliceosome. *Biochemical Society Transactions* **33**, 447–449.

26. Ramakrishnan, V. (2011). The eukaryotic Ribosome. *Science* **331** (6018), 681-682.

27. Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G. and Yusupov, M. (2011). The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* **334** (6062), 1524–1529.

28. Rabl, J., Leibundgut, M., Ataide, S. F., Haag, A. and Ban, N. (2010). Crystal Structure of the Eukaryotic 40S Ribosomal Subunit in Complex with Initiation Factor 1. *Science* **331** (6018), 730–736.

29. Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S. and Ban N. (November 2011). "Crystal Structure of the Eukaryotic 60S Ribosomal Subunit in Complex with Initiation Factor 6". *Science* **334** (6058): 941–948.

30. Cech, T. (2000). Structural biology. The ribosome is a ribozyme. *Science* **289** (5481), 878–879.

31. Rodnina, M. V., Beringer, M. and Wintermeyer, W. (2007). How ribosomes make peptide bonds. *Trends in Biochemical Sciences* **32** (1), 20–26.

32. Selmer, M., Dunham, C. M., Murphy, F. V., Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R. and Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313** (5795), 1935–1942.

33. Wahl, M.C., Will, C.L. and Luhrmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* **136**(4), 701-18.

34. Matera, A. G., Terns, R. M. and Terns, M. P. (2007) Non-coding RNAs: lessons from the small nuclear and small nucleolar RNA. *Nature Reviews Molecular Cell Biology* **8**, 209-220.

35. Kaida, D., Berg, M. G., Younis, I., Kasim, M., Singh, L. N., Wan, L. and Dreyfuss, G. (2010). U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–66.

36. Kwek, K.Y., Murphy, S., Furger, A., Thomas, B., O'Gorman, W., Kimura, H., Proudfoot, N.J. and Akoulitchev, A. (2002). U1 snRNA associates with TFIIH and regulates transcriptional initiation. *Nature Structural Biology* **9**, 800–805.

37. O'Gorman, W., Thomas, B., Kwek, K.Y., Furger, A. and Akoulitchev, A. (2005). Analysis of U1 snRNA interaction with cyclin H. *Journal of Biological Chemistry* **280**, 36920–36925.

38. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31** (1) 147–157.

39. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **35** (3 Pt 2), 849–857.

40. Perreault, J., Weinberg, Z., Roth, A., Popescu, O., Chartrand, P., Ferbeyre, G. and Breaker, R. R. (2011). Identification of Hammerhead Ribozymes in All Domains of Life Reveals Novel Structural Variations. *PLoS Computational Biology* **7** (5), e1002031.

41. Hammann, C., Luptak, A., Perreault, J. and de la Peña, M. (2012). The ubiquitous hammerhead ribozyme. *RNA* **18** (5), 871–885.

42. Forster, A. C. and Symons, R. H. (1987). Self-cleavage of plus and minus RNAs of a virusoid and a structural model for the active sites. *Cell* **49** (2), 211–220.

43. García-Robles, I., Sánchez-Navarro, J. and de la Peña, M. (2012). Intronic hammerhead ribozymes in mRNA biogenesis. *Biological Chemistry* **393** (11), 1317–1326.

44. Kuo, M. Y., Sharmeen, L., Dinter-Gottlieb, G. and Taylor, J. (1988). Characterization of self-cleaving RNA sequences on the genome and

antigenome of human hepatitis delta virus. *Journal of Virology* **62** (12), 4439–4444.

45. Tucker, B. J. and Breaker, R. R. (2005). Riboswitches as versatile gene control elements. *Current Opinions in Structural Biology* **15** (3), 342–348.

46. Roth, A., Winkler, W. C., Regulski, E. E., Lee, B. W., Lim, J., Jona, I., Barrick, J. E., Ritwik, A., Kim, J. N., Welz, R., Iwata-Reuyl, D. and Breaker, R. R. (2007). A riboswitch selective for the queuosine precursor preQ1 contains an unusually small aptamer domain. *Nature Structure and Molecular Biology* **14** (4), 308–317.

47. Meyer, M. M., Roth, A., Chervin, S. M., Garcia, G. A. and Breaker, R. R. (2008). Confirmation of a second natural preQ1 aptamer class in Streptococcaceae bacteria. *RNA* **14** (4), 685.

48. Cheah, M. T., Wachter, A., Sudarsan, N. and Breaker, R. R. (2007). Control of alternative RNA splicing and gene expression by eukaryotic riboswitches. *Nature* **447** (7143), 497–500.

49. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. and Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428** (6980), 281–286.

50. Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. and Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428** (6980), 281-286.

51. Barrick, J. E., Corbino, K. A., Winkler, W. C., Nahvi, A., Mandal, M., Collins, J., Lee, M., Roth, A., Sudarsan, N., Jona, I., Wickiser, J. K. and Breaker R. R. (2004). New RNA motifs suggest an expanded scope for riboswitches in bacterial genetic control. *Proceedings of the National Academy of Sciences U.S.A.* **101** (17), 6421–6426.

52. Collins, J. A., Irnov, I., Baker, S. and Winkler, W. C. (2007). Mechanism of mRNA destabilization by the glmS ribozyme. *Genes and Development* **21** (24), 3356–3368.

53. Tinsley, R.A., Furchak, J.R. and N.G. Walter. 2007. Trans-acting glmS catalytic riboswitch: locked and loaded. *RNA* **13**(4), 468-477.

54. Barrick, J. E. and Breaker, R. R. (2007). The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biology* **8** (11), R239.

55. Mandal, M. and Breaker, R. R. (2004). Gene regulation by riboswitches. *Nature Reviews Molecular and Cell Biology* **5,** 451-463.

56. Kuersten, S. and Goodwin, E.B. (2003). The power of the 30 UTR: translational control and development. *Nature Reviews Genetics* **4**, 626–637.

57. Sudarsan, N., Barrick, J.E. and Breaker, R.R. (2003). Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA* **9**, 644–647.

58.	Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. *Science* **309**, 1514–1518.

59.	Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K. and Slack, F.J. (2004). The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes and Development* **18**, 132–137.

60.	Farh, K.K., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B. and Bartel, D.P. (2005). The widespread impact of mammalian microRNAs on mRNA repression and evolution. *Science* **310**, 1817–1821.

61.	Frith, M.C., Pheasant, M. and Mattick, J.S. (2005). The amazing complexity of the human transcriptome. *European Journal of Human Genetics* **13**(8), 894-7.

62.	Matlin, A.J., Clark, F. and Smith, C.W. Understanding alternative splicing: towards a cellular code. (2005). *Nature Reviews Molecular and Cell Biology* **6**(5), 386-398.

63.	Mattick, J. S. and Makunin, I. V. (2005). Small regulatory RNAs in mammals. *Human Molecular Genetics* **14**, R121–R132.

64.	Castanotto, D. and Rossi, J. J. (2009). The promises and pitfalls of RNA-interference-based therapeutics. *Nature* **457** (7228), 426-433.

65.	Maine, E. M. (2008). Studying gene function in Caenorhabditis elegans using RNA-mediated interference. *Briefings in Functional Genomics and Proteomics* **7**, 184-194.

66.	Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. S., Martienssen, R. A. (2002). Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. *Science* **297** (5588), 1833-1837.

67.	Mochizuki, K., Fine, N. A., Fujisawa, T., Gorovsky, M. A. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. *Cell* **110** (6), 689-699.

68.	Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D. L., Fire, A., Ruvkun, G., 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** (1), 23-34.

69.	Giraldez, A. J., Cinalli, R. M., Glasner, M. E., Enright, A. J., Thomson, J. M., Baskerville, S., Hammond, S. M., Bartel, D. P. and Schier, A. F. (2005). MicroRNAs regulate brain morphogenesis in zebrafish. *Science* **308** (5723), 833-838.

70.	Hatfield, S. D., Shcherbata, H. R., Fischer, K. A., Nakahara, K., Carthew, R. W. and Ruohola-Baker, H. (2005). Stem cell division is regulated by the microRNA pathway. *Nature* **435** (7044), 974-978.

71. Napoli, C., Lemieux, C. and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in *trans*. *Plant Cell* **2** (4), 279–289.

72. Fire, A., Xu, S., Montgomery, M., Kostas, S., Driver, S. and Mello, C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391** (6669), 806–811.

73. Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* **136**(2), 215-233.

74. Krol, J., Loedige, I. and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews Genetics* **11**(9), 597-610.

75. Farh, K. K., Grimson, A., Jan, C., Lewis, B. P., Johnston, W. K., Lim, L. P., Burge, C. B. and Bartel, D. P. (2005). The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* **310**(5755), 1817-1821.

76. Borchert, G.M., Lanier, W. and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. *Nature Structure and Molecular Biology* **13**(12), 1097-1101.

77. Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**(2), 281-297.

78. Chekulaeva, M. and Filipowicz, W. (2009). Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Current Opinions in Cell Biology* **21**(3), 452-460.

79. Doench, J.G., Petersen, C.P. and Sharp, P.A. (2003). siRNAs can function as miRNAs. *Genes and Development* **17**(4), 438-442.

80. Jinek, M. and Doudna, J. A. (2009). A three-dimensional view of the molecular machinery of RNA interference. *Nature* **457,** 405-412.

81. Hamilton, A. J. and Baulcombe, D. C. (1999). A species of small anti-sense RNA in posttranscriptional gene silencing in plants. *Science* **286**, 950–952.

82. 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.

83. Zamore, P. D., Tuschl, T., Sharp, P. A. and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**, 25–33.

84. Macrae, I. J., Zhou, K. and Doudna, J. A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. *Nature Structure and Molecular Biology* **14**, 934–940 .

85. Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293–296.

86. Nykänen, A., Haley, B. and Zamore, P. D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107**, 309–321.

87. Gregory, R. I., Chendrimada, T. P., Cooch, N. and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631.

88. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. and Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563–574.

89. Khvorova, A., Reynolds, A. and Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**, 209.

90. Schwarz, D. S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199.

91. Tomari, Y., Matranga, C., Haley, B., Martinez, N. and Zamore, P. D. (2004a). A protein sensor for siRNA asymmetry. *Science* 306, 1377.

92. Rose, S. D., Kim, D. H., Amarzguioui, M., Heidel, J. D., Collingwood, M. A., Davis, M. E., Rossi, J. J. and Behlke, M. A. (2005). Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Research* **33**, 4140.

93. Tomari, Y., Matranga, C., Haley, B., Martinez, N. and Zamore, P. D. (2004a). A protein sensor for siRNA asymmetry. *Science* **306**, 1377.

94. Kanasty, R. L., Whitehead, K. A., Vegas, A. J. and Anderson, D. G. (2012). Action and reaction: the biological response to siRNA and its delivery vehicles. *Molecular Therapy* **20**(3), 513-524.

95. Rose, S. D., Kim, D. H., Amarzguioui, M., Heidel, J. D., Collingwood, M. A., Davis, M. E., Rossi, J. J. and Behlke, M. A. (2005). Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Research* **33**, 4140-4156.

96. 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* **293**, 1146–1150.

97. Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W. and Sontheimer, E. J. (2004). A Dicer-2-dependent 80S complex cleaves targeted mRNAs during RNAi in Drosophila. *Cell* **117**, 83–94.

98. Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K. and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740–744.

99. Haase, A.D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A. and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Report* **6**, 961–967.

100. 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 Report* **5**, 189–194.

101. Meister, G. and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343–349.

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

103. Meister, G., Landthaler, M., Peters, L., Chen, P. Y., Urlaub, H., Lührmann, R. and Tuschl T. (2005). Identification of novel argonaute-associated proteins. *Current Biology* **15**(23), 2149-2155.

104. Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., 3rd, Parker, R. and Hannon, G.J. (2005a). A role for the P-body component GW182 in microRNA function. *Nature Cell Biology* **7**, 1161–1166.

105. Ding, S.W. (2010). RNA-based antiviral immunity. *Nature Reviews Immunology* **10**, 632–644.

106. Mehlaa, S. B., Vakhariaa, J., Mehlaa, R., Abrehaa, M., Kanwarb, J. R., Tikooc, A. and Chauhana, A. Viral RNA silencing suppressors (RSS): Novel strategy of viruses to ablate the host RNA interference (RNAi) defense system. (2011). *Virus Research* **155**, 1-9.

107. Huang, J., Wang, F., Argyris, E., Chen, K., Liang, Z., Tian, H., Huang, W., Squires, K., Verlinghieri, G., Zhang, H., 2007. Cellular microRNAs contribute to HIV-1 latency in resting primary CD4(+) T lymphocytes. *Nature Medicine* **13**, 1241–1247.

108. Murakami, Y., Aly, H.H., Tajima, A., Inoue, I., Shimotohno, K., 2009. Regulation of the hepatitis C virus genome replication by miR-199a. *Journal of Hepatology* **50** (3), 453–460.

109. Lecellier, C.H., Dunoyer, P., Arar, K., Lehmann-Che, J., Eyquem, S., Himber, C., Saib, A., Voinnet, O., 2005. A cellular microRNA mediates antiviral defense in human cells. *Science* **308**, 557–560.

110. Wang, F.Z., Weber, F., Croce, C., Liu, C.G., Liao, X., Pellett, P.E., 2008. Human cytomegalovirus infection alters the expression of cellular microRNA species that affect its replication. *Journal of Virology* **82**, 9065–9074.

111. Brigneti, G., Voinnet, O., Li, W.X., Ji, L.H., Ding, S.W. and Baulcombe, D.C. (1998). Viral pathogenicity determinants are suppressors of transgene silencing in Nicotiana benthamiana. *EMBO Journal* **17**, 6739–6746.

112. Campbell, G.R. and Loret, E.P. (2009). What does the structure–function relationship of the HIV-1 Tat protein teach us about developing an AIDS vaccine? *Retrovirology* **6**, 50.

113. Azevedo, J., Garcia, D., Pontier, D., Ohnesorge, S., Yu, A., Garcia, S., Braun, L., Bergdoll, M., Hakimi, M.A., Lagrange, T., Voinnet, O., 2010. Argonaute quenching and global changes in Dicer homeostasis caused by a pathogen-encoded GW repeat protein. *Genes and Development* **24** (9), 904–915.

114. Burgyan, J. and Havelda, Z. (2011). Viral suppressors of RNA silencing. *Trends in Plant Science* **16** (5), 265-272.

115. Merai, Z., Kerényi, Z., Kertész, S., Magna, M., Lakatos, L. and Silhavy, D. (2006). Double-stranded RNA binding may be a general plant RNA viral strategy to suppress RNA silencing. *Journal of Virology* **80**, 5747– 5756.

116. Lakatos, L. Csorba, T., Pantaleo, V., Chapman, E. J., Carrington, J. C., Liu, Y. P., Dolja, V. V., Calvino, L. F., López-Moya, J. J. and Burgyán, J. (2006). Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO Journal* **25**, 2768–2780.

117. Guo, H.S. and Ding, S.W. (2002) A viral protein inhibits the long range signaling activity of the gene silencing signal. *EMBO Journal* **21**, 398–407.

118. Vargason, J., Szittya, G., Burgyán, J. and Hall, T. M. (2003). Size selective recognition of siRNA by an RNA silencing suppressor. *Cell* **115**, 799–811.

119. Bumcrot, D., Manoharan, M., Koteliansky, V. and Sah, D. W. Y. (2006). RNAi therapeutics: a potential new class of pharmaceutical drugs. *Nature Chemical Biology* **2** (12), 711-719.

120. Hoerter, J.A.H. and Walter, N.G. (2007). Chemical modification resolves the asymmetry of siRNA strand degradation in human blood serum. *RNA* **13**, 1887-1893.

121. Yu, D., Pendergraff, H., Liu, J., Corey, D. R, et al. (2012). Single-stranded RNAs use RNAi to potently and allele-selectively inhibit mutant Huntingtin expression. *Cell* **150** (5), 895-908.

122. Lima, W. F., Prakash, T. P., Crooke, S. T., et al. (2012). Single-stranded siRNAs activate RNAi in animals. *Cell* **150** (5), 883-894.

123. Choung, S., Kim, Y.J., Kim, S., Park, H.O. and Choi, Y.C. (2006). Chemical modification of siRNAs to improve serum stability without loss of efficacy. *Biochemical and Biophysical Research Communications* **342**, 919–927.

124. Prakash, T.P. et al. (2005). Positional effect of chemical modifications on short interference RNA activity in mammalian cells. *Journal of Medicinal Chemistry* **48**, 4247–4253.

125. Hoshika, S., Minakawa, N. and Matsuda, A. (2004).Synthesis and physical and physiological properties of 4'-thioRNA: application to post-modification of RNA aptamer toward NF-κB. *Nucleic Acids Research* **32**, 3815–3825.

126. Dande, P. et al. (2006).Improving RNA interference in mammalian cells by 4'-thio-modified small interfering RNA (siRNA): effect on siRNA activity and nuclease stability when used in combination with 2'-O-alkyl modifications. *Journal of Medicinal Chemistry* **49**, 1624–1634.

127. Jackson, A.L. et al. (2006). Position-specific chemical modification of siRNAs reduces "off-target" transcript silencing. *RNA* **12**, 1197–1205.

128. Bartlett, D. W. and Davis, M. E. (2007). Effect of siRNA nuclease stability on the in vitro and in vivo kinetics of siRNA-mediated gene silencing. *Biotechnology and Bioengineering* **97**, 909-921.

129. Berezhna, S. Y., Supekova, L., Supek, F., Schultz, P. G. and Deniz, A. A. (2006). siRNA in human cells selectively localizes to target RNA sites. *Proceedings of the National Academy of Sciences* **103**, 7682-7687.

130. Doherty, E.A. and J.A. Doudna. (2001). Ribozyme structures and mechanisms. *Annual Reviews Biophysics Biomolecular Structure* **30**, 457-475.

Chapter 2

Viral RNAi suppressor reversibly binds siRNA to outcompete Dicer and RISC via multiple-turnover[1, 2]

2.1 Introduction

Over the past decade, RNA interference (RNAi) has become recognized as a benchmark laboratory and potential clinical tool to control gene expression, as well as a highly conserved eukaryotic immune response against viral pathogens or transposons.[1-4] During viral infection, plant and invertebrate cells are known to activate RNAi where viral RNA is specifically recognized by the cytoplasmic RNase III-type endonuclease Dicer and degraded by an RNA-induced silencing complex (RISC). Dicer cleaves exposed stretches of double-stranded RNA into small interfering RNAs (siRNAs) 21-24 nucleotides in length (depending on species). Dicer then delivers the double-stranded siRNA via the RISC-loading complex[5] to RISC,[6,7] where the passenger strand of the duplex is cleaved and released while the guide strand is retained to guide specific cleavage of complementary viral RNA genomes or transcripts.[8,9] Dicer is thus thought to play a critical dual role in RNAi by both generating virus-specific siRNA duplexes and assembling them with the necessary protein

components into active RISC.[10] In mammalian including human cells, artificial siRNAs have been shown to confer antiviral immunity,[11] but only very recently evidence has emerged to suggest that small RNAs including siRNAs are also directly generated from infecting viruses.[12]

To subvert the host RNAi response many viruses have evolved small proteins that act as RNA silencing suppressors (RSSs).[13,14] One of the best characterized RSSs is exemplified in the plant pathogen Carnation Italian Ringspot Virus (CIRV), which encodes a 19 kDa protein termed p19, an important pathogenicity factor that is highly conserved among all tombusviruses.[14] Crystal structures[15,16] and recent molecular dynamics simulations[17] show that homodimers of p19 bind a single siRNA duplex in a positively charged surface cleft (**Fig. 2.1a**). Pairs of tryptophan residues stack on both terminal base pairs to establish a sequence-independent, caliper-like size selection for siRNA duplexes with 19 base pairs and 2-nucleotide long 3′-overhangs. Sequestration and binding of Dicer-generated siRNAs by p19 has been observed to correlate with reduced viral RNA degradation, systemic symptom spread, and sustainability of the viral phenotype after infection.[18-21] Previous electrophoretic mobility shift assays (EMSAs) have suggested that the affinity of dimeric p19 for siRNA is high with an apparent equilibrium dissociation constant of K_P = 0.17 ± 0.02 nM.[15] Such tight binding is consistent with a widely accepted sequestration model wherein the p19 dimer near-quantitatively binds siRNA duplexes to suppress RNAi.[14,18] Binding of double-stranded RNA is considered to be the most common RSS mechanism,[22,23] yet measurements of siRNA binding and dissociation kinetics to better understand sequestration have not been reported to date for any RSS protein.

Here we develop solution-based fluorescence quenching assays that observe siRNA:p19 (SP) binding to be readily reversible with not only a rapid binding rate constant ($(1.69 \pm 0.07) \times 10^8$ $M^{-1}s^{-1}$), but also a dissociation rate constant (k_{off} = 0.062 ± 0.002 s^{-1}) markedly faster than those of slow-release nucleic acid binding proteins.[24,25] Our results are consistent, within error, with the high apparent affinity measured in EMSAs, yet expand on the sequestration

model. If binding were completely irreversible, it would require the virus to generate sufficient quantities of p19 dimer to quantitatively bind all small RNAs produced enzymatically by either Dicer cleavage or the amplification effected in a plant cell through RNA-dependent RNA polymerase.[2] By contrast, facile reversibility of SP complex formation opens the possibility that multiple-turnover may potentiate p19's efficacy as RSS. In particular, if p19 were to disrupt the siRNA:Dicer (SD) interaction via multiple-turnover, it would potently interfere with the role of Dicer as obligatory RISC assembly factor.

Consistent with such a mechanism of p19 action, we observe the efficient competition of p19 for siRNA binding with active recombinant human Dicer in the absence of other known RISC components, as well as the prominent disruption of RISC assembly complexes observed in human HeLa cell extract containing the three core RISC-loading complex components Dicer, TRBP, and Ago2.[5] Mathematical modeling demonstrates that the Dicer competition data are only accurately reproduced assuming the formation of a transient ternary complex, siRNA:Dicer:p19 (SDP). Taken together, a multiple-turnover "catch and release" mechanism emerges wherein p19 repeatedly promotes dissociation of SD complexes to efficiently suppress Dicer-mediated RISC assembly.

2.2 Materials and Methods

Sources of p19, Dicer, HeLa cell extract, and RNA. Purified p19 protein was prepared as previously described,[15] and its dimer concentration was measured by Bradford assay. Recombinant human Dicer was prepared by Genlantis using the Baculovirus system with a commercially available protocol from Invitrogen. The His-tagged Dicer was purified on a Ni^{2+} column and the His tag removed by Tev protease cleavage, which should remove any contaminating factors that do not remain bound to the recombinant Dicer. S100 cytosolic HeLa cell extract was obtained from Jena Biosciences and the integrity of Dicer in both samples was tested by Western blot analysis (**Fig. 2.2a**). Additionally, we demonstrated biological activity of the recombinant Dicer through cleavage assays using long double-stranded RNA substrates (**Fig. 2.2b**). An estimate of Dicer concentration

in the partially purified preparation was derived from Bradford assays to measure the total protein concentration in combination with the 5% purity quoted by Genlantis and verified by Coomassie staining of a denaturing SDS-PAGE gel (**Fig. 2.2c**). This estimate is conservative as higher relative purities would only imply that even higher concentrations of Dicer are needed to efficiently compete with p19 for siRNA binding. All RNAs were synthesized by the HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory at the Yale University School of Medicine, then deprotected and purified as previously described.[27,56] siRNA sequences were derived from the firefly luciferase gene[26] and were as follows: Sense strand 5´-P-CGU ACG CGG AAU ACU UCG AAA-3´; antisense strand 5´-P-UCG AAG (Fluorescein-dT)AU UCC GCG (Amino C6-dT)AC GUG-3´, where P denotes a 5´-phosphate, and Fluorescein-dT and Amino C6-dT denote the phosphoramidites of the same name from Glen Research; the antisense strand was also synthesized with only one or no modification wherein U replaced the modified dT. Where needed, tetramethylrhodamine was coupled to the Amino C6 linker and the RNA purified as described.[56] miRNA sequences were derived from the human wild-type let-7a miRNA and were as follows: Let-7a antisense strand 5´-P-UGA GGU AGU AGG UUG UAU AGU U-3´; Let-7a sense strand 5´-P-CUA UAC AAU CUA CUG UCU UUC C-3´.

Steady-state fluorescence spectroscopy. Standard fluorescence measurements were performed on an Aminco Bowman spectrofluorometer (Thermo Scientific) at 37 °C. siRNA duplexes were formed by mixing an excess of 100 nM of unlabeled sense strand with 50 nM labeled sense strand in near-physiologic standard buffer (50 mM Tris-Acetate, pH 7.4, 80 mM KCl, 20 mM NaCl, 1 mM MgCl$_2$, 1 mM DTT, 0.02% (v/v) Tween 20), heat annealing for 2 min at 70 °C, and cooling to room temperature over 5 min. Fluorescein/tetramethylrhodamine doubly-labeled siRNAs were excited at 490 nm (4 nm slit width) and emission alternately detected at 520 nm (for fluorescein, F) and 585 nm (8 nm slit width, for tetramethylrhodamine, T) to calculate the T/F ratio; F or T singly-labeled siRNAs were excited/detected at 490/520 nm or 565/585nm, respectively. In a typical experiment, p19 was added to 50 nM

labeled siRNA and mixed manually to a final concentration of 500 nM unless otherwise indicated. For a chase experiment, unmodified siRNA duplex (formed through annealing the unlabeled composite sense and antisense RNA strands in a 2:1 ratio in standard buffer) was added to a final concentration of 3 μM. Saturation by the chase was ensured by triplicate trials using 500 nM, 750 nM, 1.5 μM, and 3 μM (**Appendix A, Fig. A.1**).

To observe fast SP complex formation, siRNA samples were prepared as above in standard buffer to a final concentration of 40 nM. In standard buffer p19 was added to varying (at least 8-fold) final excess concentrations (0.32 – 1.76 μM) at 37 $^\circ$C in a KinTek PMA23B stopped-flow spectrofluorometer while exciting at 520 nm and detecting emission at 585 nm. All time course were fit with single-exponential increase or decrease functions in Origin (OriginLab) as described.[57]

Electrophoretic mobility shift assays. For radioactive EMSAs the non-phosphorylated antisense strand was first 5′-^{32}P labeled with T4 polynucleotide kinase (PNK) and γ-^{32}P-ATP at an RNA concentration of 800 nM. PNK was inactivated by heating to 90 $^\circ$C for 10 min, then the 5′ phosphorylated complementary strand was added in 2-fold molar excess and the reaction slowly cooled to room temperature to anneal the two strands. The siRNA duplex was further purified by nondenaturing, 20% (w/v) polyacrylamide gel electrophoresis. The duplex siRNA was cut out, eluted overnight into 1 mM EDTA at 4 $^\circ$C, ethanol precipitated, and dissolved in RNase-free water. Scintillation counting was done on a Beckman LS6500 Multipurpose Scintillation Counter. EMSAs were performed as described previously[15] on 10-cm non-denaturing 12% (w/v) polyacrylamide gels in 0.5x TBE (44.5 mM Tris-Borate, 1 mM EDTA) and run at 500 V and 4 $^\circ$C for 2 h. Dicer and p19 at varying final concentrations were added to radiolabeled siRNA duplex (50,000 cpm, < 400 pM of 5′-^{32}P labeled duplex) in 10 μL standard buffer and incubated for 0.5 hrs. Next, 10 μL non-denaturing loading buffer (10% glycerol, 0.5x TBE, 0.025% bromophenol blue, and 0.025% xylene cyanol) was added and each sample loaded into a well of the EMSA gel. Competition experiments were performed by preincubating < 400 pM siRNA

duplex (prepared as above) in standard buffer either with 140 nM Dicer and increasing concentrations of p19, or with 0.17 nM or 2.5 nM p19 and increasing concentrations of Dicer. Gel chase experiments were done in standard buffer with 70 and 10 nM of Dicer or p19 respectively to < 400pM 5′-^{32}P labeled siRNA and addition of 750 nM unlabeled siRNA chase for 0, 10, 30, 60 and 120 min. Gels were wrapped in plastic wrap and exposed to phosphor screens, which were scanned on a Storm 840 PhosphorImager and quantified using Image Quant software (Molecular Dynamics). The relative fraction f of relevant protein complexes was fit with the following hyperbolic binding isotherm to extract saturation point f_{max} and half-titration point $K_{D,app}$:

$$f = f_{max} \frac{[protein]}{[protein] + K_{D,app}}$$

RISC assembly complex formation. To observe RISC assembly complexes 15, 50, or 75% (v/v) cytosolic HeLa cell extract was incubated with < 400 pM siRNA duplex (prepared as above) in a final volume of 10 µl HeLa buffer [8% (v/v) Lysis Buffer (23.7 mM HEPES, pH 7.5, 79 mM potassium acetate, 1.58 mM magnesium acetate, 5 mM DTT, and 1 mg/ml Perabloc SC (Roche)), supplemented with 20% (v/v) 5x RNAi mix (125 mM creatine phosphate, 5 mM ATP, 25 mM DTT), 2% (w/v) creatine kinase, and 20% (v/v) RNA guard RNase Inhibitor (GE Bioscience)] for 2 h at 4 ^{o}C, unless otherwise specified. For competition experiments, p19 was added in increasing concentrations to the HeLa cell extract/siRNA mixture either immediately or after a 2 h preincubation. Time courses were stopped by direct loading onto a running nondenaturing 4% (w/v) polyacrylamide gel in 0.5x TBE (44.5 mM Tris-Borate, 1 mM EDTA), run at ~50 mW/cm^{2} and 4 ^{o}C. For comparison with a previous assay, 50 and 75% (v/v) HeLa cell extract was incubated with siRNA as described in[28] for 1 h at 4 ^{o}C to 37 ^{o}C and analyzed by EMSA, yielding similar results (**Appendix A, Fig. A.3**). 15% (v/v) extract was preferred for most of our work due to the better pH control afforded by this lower concentration and to conserve material. For antibody supershift EMSA, cytosolic HeLa cell extract (24 µg total protein) was pre-

incubated with 0.8 µg or 1.6 µg of anti-Dicer rabbit polyclonal antibody (Santa Cruz Biotechnology, Inc.) and 0.8 µg or 1.6 µg of unrelated control antibodies, either anti-DsRed rabbit polyclonal antibody (Clontech Laboratories, Inc.) or anti-β actin mouse monoclonal antibody (Sigma-Aldrich, Inc.), for 2 h at 4 °C prior to the RISC assembly assay. The final volume was 60 µl per lane. Nondenaturing 4% (w/v) polyacrylamide gel electrophoresis was used to resolve these complexes as described above. Time courses and titration curves were fit as above.

Time-resolved FRET. Time-resolved donor (fluorescein, F) fluorescence decays were collected in the presence and absence of acceptor (tetramethylrhodamine, T) using time-correlated single-photon counting, as previously described.[58,59] As before, the donor decay in the presence of acceptor was modeled as a weighted average of the decays for each donor-acceptor distance in the ensemble of molecules assuming a three-dimensional Gaussian distance distribution with a Förster distance of 54 Å between F and T.[58,59]

Western Blots for Dicer Detection. For denaturing Western blots, the cytosolic HeLa cell extract, the recombinant Dicer enzyme preparation, and recombinant TRBP (Abcam) and Ago2 (Abnova) proteins were run on a NuPAGE Novex 4-12% gradient Bis-Tris Gel in the presence of a molecular weight marker (ECL Plex Rainbow marker). The gel was then electroblotted in Tris-Glycine buffer (25 mM Tris, 192 mM Glycine, and 20% (v/v) Methanol) onto a Polyvinylidene Fluoride (PVDF) membrane (Immobilon-P Membrane, Millipore) over 75 min at 300 mA, using a Bio-Rad Trans-Blot SD semi-dry transfer cell per the manufacturer's instructions. After transfer the membrane was probed with either rabbit anti-Dicer (Santa Cruz Biotechnology), mouse anti-TRBP or mouse anti-Ago2 (both Abnova) followed by goat anti-Rabbit-Cy5 and goat anti-mouse-Cy3 (both Amersham), respectively. The probed blots in **Fig. 2.2a,d,e** were visualized on a Typhoon 9410 Variable Mode Imager (GE Healthcare Life Sciences).

For non-denaturing Westerns blots, RISC assembly complexes were formed as described above but reactions were scaled up by 6-fold. Gels were

soaked in 0.1% (w/v) SDS for 15 min and then electroblotted in Tris-Glycine transfer buffer (25 mM Tris, 25 mM Glycine) onto a PVDF membrane (Immobilon-P Membrane, Millipore) over 75 min at 300 mA, using a Bio-Rad Trans-Blot SD semi-dry transfer cell per the manufacturer's instructions. After transfer, the proteins were fixed to the membrane by incubating in 5-10% (v/v) acetic acid, rinsing with deionized water, and air drying. The membrane was probed with a rabbit primary antibody against Dicer (Santa Cruz Biotechnology), followed by a goat anti-rabbit secondary antibody, conjugated with horseradish peroxidase (Zymed, Invitrogen). The blot was developed using a peroxide/enhancer solution (ECL-Plus detection, Amersham), and visualized on a Typhoon 9410 Variable Mode Imager (**Fig. 2.6**).

Modeling p19 in the context of RNAi. For the model described schematically in (**Fig. 2.7a**), changes in viral protein expression levels were monitored at steady-state in the absence and presence of the RSS p19. Changes in free p19 protein over time depend on terms representing the natural rates of production (r_x r_t / r_{dm}), degradation (r_{dp} [p19]), the extent of siRNA binding (k_{on} [siRNA] [P]), and the siRNA:p19 complex dissociation (k_{off} [SP]). Production of p19 is governed by its transcription rate (rx), its messenger RNA degradation rate (r_{dm}), and its translation rate (r_t), which are assumed to have reached equilibrium on a timescale faster than that of repression. The p19 protein is degraded by the first-order rate constant (r_{dp}), leading to:

$$p19'[\tau] = \frac{r_x r_t}{r_{dm}} - r_{dp}[p19] - k_{on}[siRNA][p19] + k_{off}[SP]$$

[1]

The amount of usable siRNA in the system, i.e., the fraction able to associate with RISC, is reduced by the amount of siRNA bound to p19. Production of siRNA is dependent on the catalytic rate constant of Dicer (k_{catD}), and siRNA is degraded with a first order rate constant (r_{ds}). The parameter *n* is the fraction of siRNA that reenters the RNAi pathway after p19 dissociation, yielding:

$$siRNA'[\tau] = k_{catD}[Dicer] - r_{ds}[siRNA] - k_{on}[siRNA][p19] + n * k_{off}[SP]$$

[2]

The time evolution of the viral protein of interest that serves as a reporter of RNAi suppression and p19 enhancement is described as:

$$viral'[\tau] = \frac{r_x r_t}{\frac{k_{catR}}{K_{DR}}[RISC][siRNA] + r_{dm}} - r_{dp}[viral]$$

[3]

where (k_{catR}) is the catalytic rate constant for RISC-mediated cleavage of mRNA. The additional term in the denominator of Equation 3 is a quasi-steady state assumption used in previous work[33,60] to describe RISC mediated degradation of viral messenger RNA (mRNA). It assumes that the active concentration of siRNA-loaded RISC is approaching its steady-state value with rate constant k_{catR}. This buildup of siRNA-loaded RISC contributes to the rate of degradation of the viral mRNA, as described by the concentrations of RISC and siRNA, as well as the dissociation constant K_{DR}. The proteins not silenced by RNAi are represented at steady-state by unitless constants:

$$[Dicer] = \frac{r_x r_t}{r_{dp} r_{dm}} \quad [4] \qquad [RISC] = \frac{r_x r_t}{r_{dp} r_{dm}} \quad [5]$$

Equations 1-3 are rescaled as described by Marshall[33] and solved at steady-state to yield:

$$P = \frac{1}{1 - Aa\frac{k_{off}}{kcat_D} + (Av\beta)S} \quad [6] \qquad S = \frac{1 + \frac{k_{off}}{kcat_D}naP}{\beta + \beta vP} \quad [7] \qquad V = \frac{1}{1 + \beta yS} \quad [8]$$

where y and β are the parameters previously described as the efficiency of the RNAi pathway and the RNAi settling time, respectively.[33] The parameters a, v, and A are the fraction of siRNA-bound relative to total p19, the efficiency of p19, and the efficiency of Dicer in a particular cell line, respectively. The level of the viral protein at steady-state is dependent mostly on S. For example, in the absence of RNAi (i.e., in the absence of Dicer) $S = 0$, so that $V = 1$, that is, the viral protein level is not suppressed. Substituting Equation 6 into Equation 7 yields two solutions for S, where we choose S_2 as the more biologically relevant one:

$$S_1 = \frac{1 - aA\frac{k_{off}}{kcat_D} + v - Av + \sqrt{\left(-4A(-1 + a\frac{k_{off}}{kcat_D}(A - n))\frac{v}{\beta^2} + (-1 + aA\frac{k_{off}}{kcat_D} + (-1 + A)v)^2\right)}}{2Av\beta}$$

$$S_2 = \frac{1 - aA\frac{k_{off}}{kcat_D} + v - Av - \sqrt{\left(-4A(-1 + a\frac{k_{off}}{kcat_D}(A - n))\frac{v}{\beta^2} + (-1 + aA\frac{k_{off}}{kcat_D} + (-1 + A)v)^2\right)}}{2Av\beta}$$

[9]

Substituting S_2 into Equation 8 yields:

$$V_{Rev} = \frac{1}{1 + \dfrac{y(1 - aA\frac{k_{off}}{kcat_D} + v - Av - \sqrt{\left(-4A(-1 + a\frac{k_{off}}{kcat_D}(A-n))\frac{v}{\beta^2} + (-1 + aA\frac{k_{off}}{kcat_D} + (-1+A)v)^2\right)}}{2Av}}}$$

[10]

If p19 binding is assumed to be irreversible (i.e., $k_{off} = 0$), then V solves to:

$$V_{Irr} = \frac{2Av}{\left(-1 - v + \sqrt{1 + 2(1 + A)v + (-1 + A)^2v^2}\right)y + Av(2 + y)}$$

[11]

In the absence of p19 (i.e., $P = 0$), based on Equations 7 and 8, V reduces to the fully suppressed protein level:

$$V_o = \frac{1}{1+y} \qquad \text{[12]}$$

The enhancement of viral protein production due to either irreversible (ϕ_{Irr}) or reversible p19 binding (ϕ_{Rev}) is then found to be (see also **Fig. 2.7b**):

$$\phi_{Irr} = \frac{V_{Irr}}{V_o} \qquad \text{[13]} \qquad Rev = \frac{V_{Rev}}{V_o} \qquad \text{[14]}$$

2.3 Results

Fluorescence assays to observe siRNA:p19 complex formation. We site-specifically labeled the antisense (guide) strand of an established luciferase-targeting siRNA duplex[26,27] with either fluorescein or tetramethylrhodamine as fluorescence quenching probes in positions not expected to interfere with p19 binding (**Fig. 2.1a**). We found that only the tetramethylrhodamine label was efficiently quenched upon p19 addition (**Fig. 2.1b**). Addition of buffer or other proteins such as recombinant Dicer caused no such quenching (**Appendix A, Fig. A.1**). To observe a robust quenching signal upon p19 addition, we proceeded to label with both fluorophores together, allowing us to monitor a normalized tetramethylrhodamine:fluorescein (T/F) fluorescence ratio that is insensitive to dilution effects and other perturbations of an otherwise relative fluorescence signal. Next, we tested by EMSA p19 binding of this doubly-labeled siRNA against that of the unlabeled siRNA and detected no interference from the fluorophores (K_P = 0.17 ± 0.03 nM in both cases, consistent with earlier EMSA studies[15] (**Fig. 2.1c**). Time-resolved fluorescence resonance energy transfer
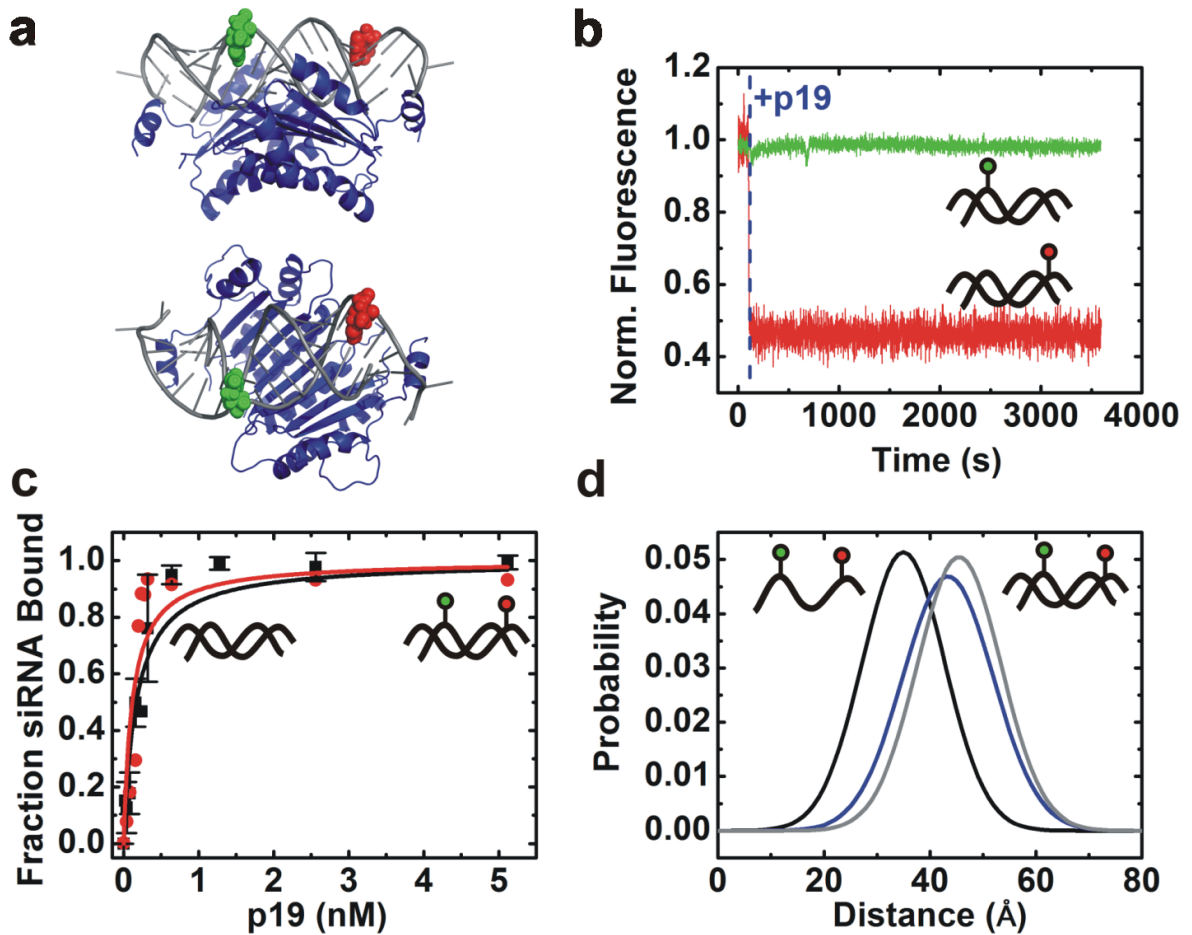
Figure 2.1. Fluorescence assays to detect siRNA:p19 complex formation.
(a) Crystal structure of p19 (blue) bound to an siRNA (grey). The nucleotides to which the fluorophores fluorescein (F) and tetramethylrhodamine (T) are attached are highlighted in green and red, respectively. **(b)** The T-only labeled siRNA (red) is quenched upon p19 addition, while the F-only labeled siRNA (green) is unaffected by p19. **(c)** EMSA analysis shows that the doubly-labeled ft-siRNA (red) binds to p19 with the same affinity as the unlabeled siRNA (black). **(d)** Time-resolved FRET measures the F:distances for the single-stranded, protein-free double-stranded, and p19-bound double-stranded siRNA 35 Å, 43.4 Å, and 44.5 Å, respectively. Reprinted with permission from Elsevier.
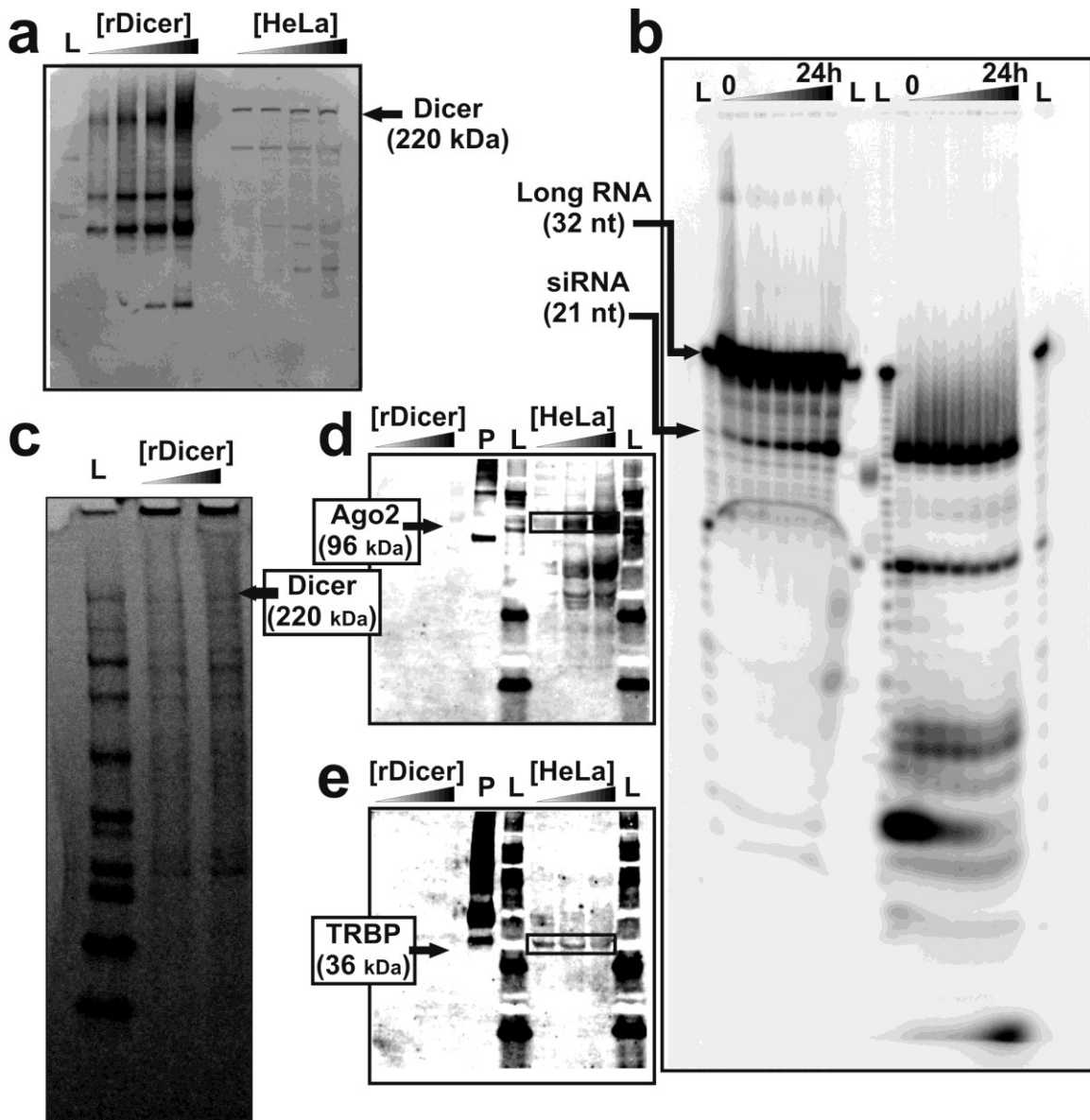
Figure 2.2. Characterization of recombinant (r)Dicer and HeLa cell extract.
(a) Denaturing Western blot against Dicer showing the presence of full length Dicer p rotein (218 kDa) in both extracts. Lane 1 (from left, marked "L"): Full-Range Rainbow Molecular Weight Marker (RPN800E). Lanes 2-5: Recombinant Dicer (Genlantis) – 3, 6, 9, and 12 μL, respectively. Lanes 6-9: Cytosolic HeLa cell extract – 3, 6, 9, and 12 μL, respectively. The multiple bands seen on the Western blot may result from degradation of Dicer or cross-reactivity of the antibody. **(b)** Cleavage assay measuring the activity of recombinant Dicer. Recombinant Dicer was incubated with radio-labeled double-stranded 32 nucleotide RNA for 1, 3, 5, 7, 9, 12, and 24 h in the Genlantis supplied buffer (lanes 2-8 from left to right) and in the near-physiologic standard buffer (lanes 12-18) used for the Dicer assays (50 mM Tris-Acetate, pH 7.4, 80 mM KCl, 20 mM

41

NaCl, 1 mM MgCl$_2$, 1 mM DTT, 0.02% (v/v) Tween 20). In the company-supplied buffer, cleavage was observed over the recommended incubation time of 12 h. In the near-physiologic buffer, all RNA was 100% cleaved into mature siRNA as early as 1 h after incubation. **(c)** Coomassie stained 4-12% SDS-PAGE gel. Lane 1 (from left, marked "L"): Full-Range Rainbow Molecular Weight Marker (RPN800E). Lane 2: Recombinant Dicer (Genlantis) – 5 µL. Lane 3: Recombinant Dicer (Genlantis) – 10 µL. Fraction of full length Dicer was calculated as 5 ± 2% **(d)** Western blot of recombinant Dicer and HeLa cell extract against Ago2. **(e)** Western blot of recombinant Dicer and HeLa cell extract against TRBP. Reprinted with permission from Elsevier.

(trFRET) between the two fluorophores revealed only a small increase in fluorophore distance distribution upon p19 addition (with the mean distance increasing from 43.4 to 44.5 Å; **Fig. 2.1d**), indicating that formation of the SP complex leads to only a minor deformation of the siRNA duplex, as expected from X-ray crystallography.[15] We conclude that the doubly-labeled siRNA duplex is a suitable reporter for formation of the siRNA:p19 interaction in solution.

p19 binds siRNA rapidly and reversibly. Addition of increasing excess concentrations of p19 to the doubly-labeled siRNA led to an increase in the pseudo-first order rate constant of tetramethylrhodamine quenching as monitored by the T/F ratio, yielding a bimolecular binding rate constant of k_{on} = (1.69 ± 0.07) × 10^8 $M^{-1}s^{-1}$ (**Fig. 2.3a,b**). Next, dissociation of the preformed SP complex was monitored upon addition of a large excess of unlabeled siRNA (chase), yielding a dissociation rate constant of k_{off} = 0.062 ± 0.002 s^{-1} (**Fig. 2.3c and Appendix A, Fig. A.1**). Together with the rapid k_{on}, a solution-based dissociation equilibrium constant of k_{off}/k_{on} = $K_{P,sol}$ = 0.37 ± 0.08 nM is obtained, within 2-fold of the EMSA predicted value (0.17 nM), attesting to a high-affinity, yet also notably reversible siRNA:p19 binding mode.

p19 efficiently competes with recombinant human Dicer for siRNA binding. The critical dual role of Dicer in RNAi as both the enzyme that produces siRNAs and as a RISC assembly factor[10] suggests the siRNA:Dicer (SD) interaction as a particularly vulnerable step during RNA silencing. Previous studies in embryo extracts of the fruitfly *Drosophila* have shown that p19 suppresses RNA silencing by competing with the Dcr2-R2D2 complex.[18,22] To determine whether p19 also competes with human Dicer for siRNA, we employed EMSA using radio-labeled siRNA duplexes and an active recombinant human Dicer preparation, devoid of other known RISC proteins (**Fig. 2.2**) (available as a relatively crude purification of heterologously overexpressed Dicer from Genlantis, see Materials and Methods). We measured the apparent equilibrium dissociation constant of the human SD complex as K_D = 3.7 ± 0.4 nM (**Fig. 2.4a and Appendix A, Fig. A.2a**). The presence of a low concentration of p19 (0.17 nM), equivalent to the K_P of the SP complex, weakened this interaction by >25-fold to
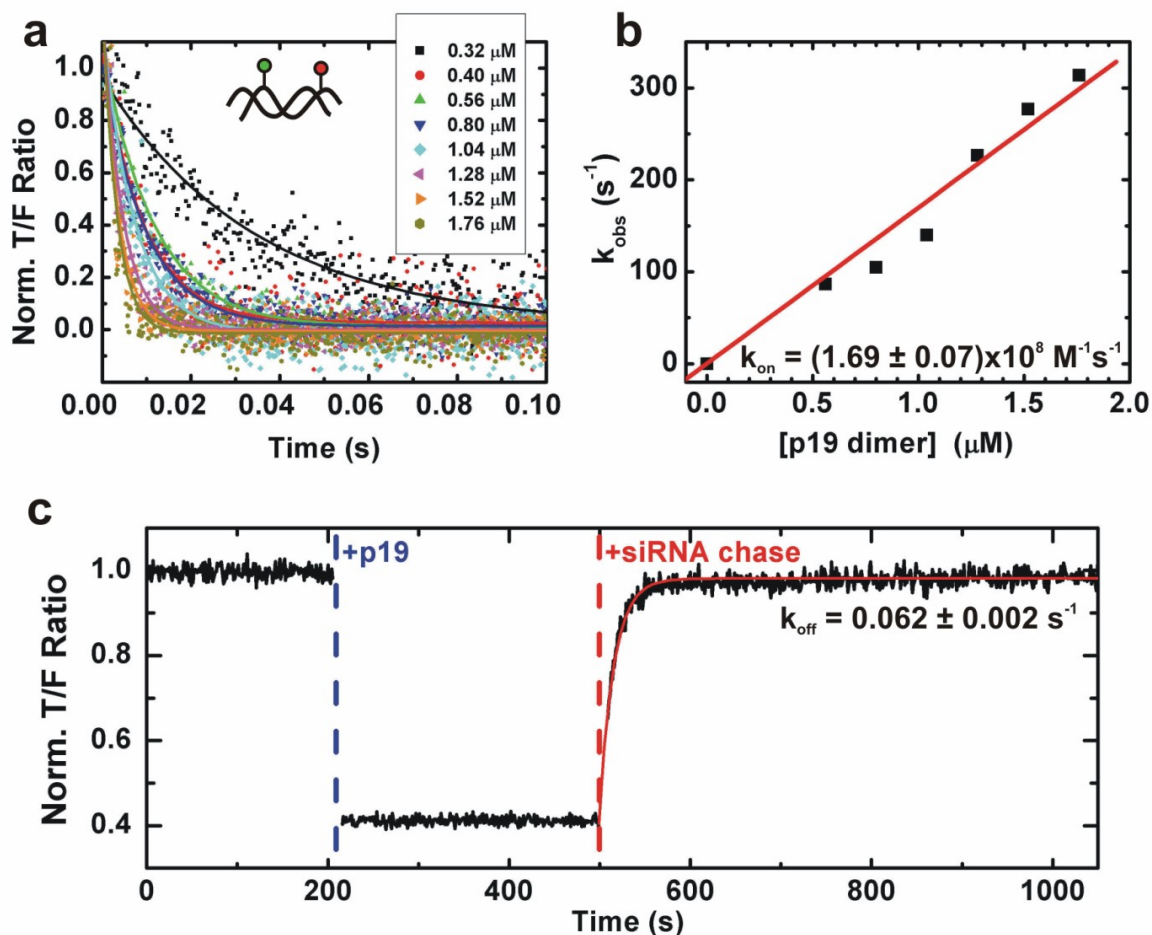
Figure 2.3. Kinetics of formation and dissociation of the siRNA:p19 complex. (a) Pseudo-first order kinetics of the binding of doubly-labeled siRNA (50 nM) to excess p19, measured through the T/F ratio during stopped-flow mixing. Scatter points denote raw data; lines denote single-exponential fits to extract rate constants. **(b)** Plot of the observed pseudo-first rate constants from panel A over the p19 concentration. The linear fit yields the indicated bimolecular binding rate constant, k_{on}. **(c)** Fluorescence detection of the binding of doubly-labeled siRNA (50 nM) to p19 (500 nM) and of complex dissociation upon addition of an unlabeled siRNA chase (3 μM). A single-exponential fit (red line) yields the indicated dissociation rate constant, k_{off}. Reprinted with permission from Elsevier.

44
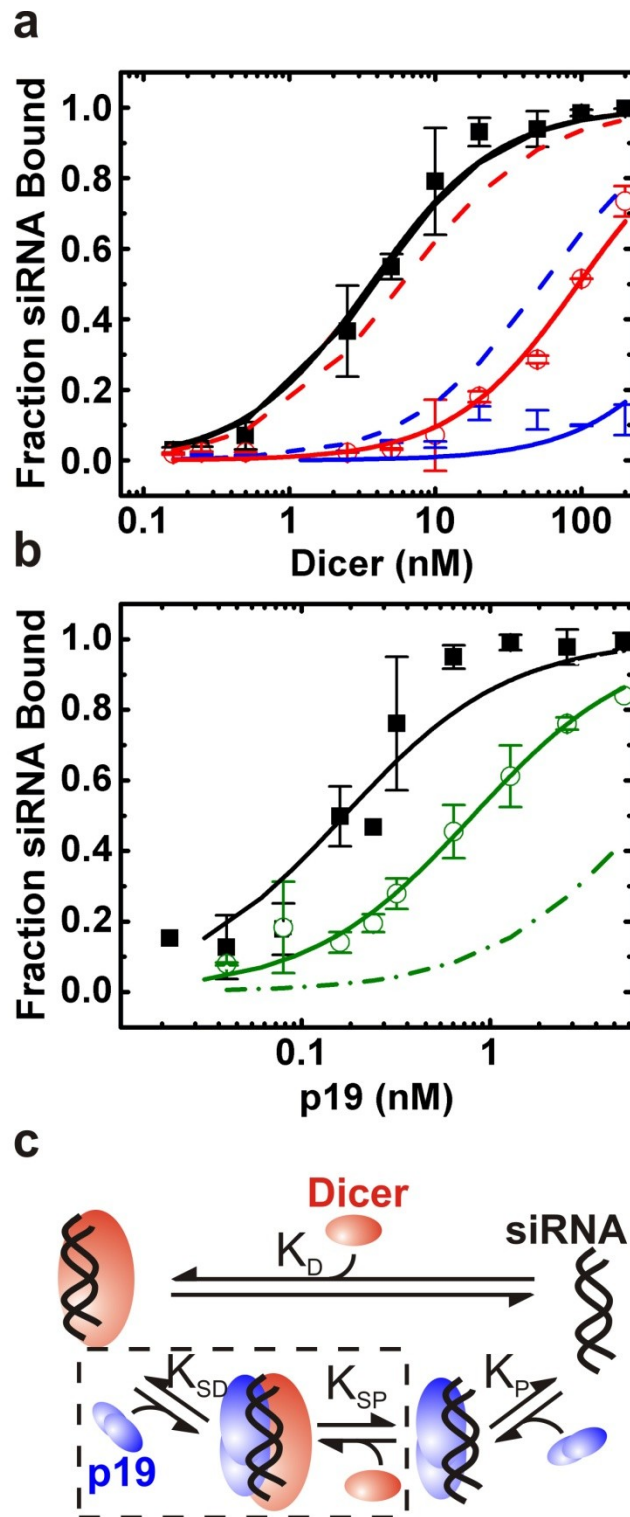
Figure 2.4. Competition of p19 with human Dicer for siRNA binding as detected by EMSA. (a) The fraction of siRNA bound to Dicer over increasing Dicer concentrations in the absence of p19 (black squares), and presence of

0.17 nM (open red circles) or 2.5 nM p19 (blue dots). Solid lines, corresponding fits of hyperbolic binding isotherms. Dashed lines indicate the predicted curves based on a dissociative binding mechanism (siRNA has to first dissociate from Dicer to bind p19). **(b)** Fraction of siRNA bound to p19 over increasing p19 concentrations in the absence (black squares) or presence of a saturating Dicer concentration of 140 nM (open green circles). Solid lines, corresponding fits of hyperbolic binding isotherms. The dashed-dotted line indicates the predicted curve based on a dissociative binding mechanism. **(c)** Schematic of the two competition models between Dicer and p19 for siRNA; the dashed box highlights the part unique to the ternary complex formation model. Reprinted with permission from Elsevier.

~100 nM (**Fig. 2.4a and Appendix A, Fig. A.2b**). The presence of a saturating concentration of p19 (2.5 nM) weakened the interaction still further by >270-fold to an estimated lower limit of ~1.0 μM (**Fig. 2.4a and Appendix A, Fig. A.2c**).These findings led us to hypothesize that p19 can function to inhibit RNAi in humans by interrupting the SD complex before Dicer can hand off siRNA to RISC. Reciprocal competition experiments showed that the presence of even saturating concentrations (140 nM) of Dicer lowered the apparent affinity of p19 for siRNA only marginally (~4-fold from 0.17 ± 0.03 nM to 0.81 ± 0.06 nM; **Fig. 2.4b and Appendix A, Fig. A.2d,e**), further supporting the hypothesis that strong competition with the SD complex is an important mode of action for p19 inhibition of RNAi in human cells.

Human siRNA-containing complexes formed in cytosolic cell extract are vulnerable to p19 challenge. Previous studies with *Drosophila* extracts have shown that p19 inhibits RISC assembly while disrupting preassembled RISC to a lesser extent.[22] Our result that p19 disrupts the complex of recombinant human Dicer with siRNA suggests that p19 may also be effective in suppressing RISC assembly in mammals. To test this hypothesis and probe p19's effect on the formation of human RISC-related complexes, we used a modified *in vitro* RISC assembly assay.[28,29] Specifically, we incubated radio-labeled siRNA duplex with 15% (v/v) cytosolic HeLa cell extract, which we showed to contain the three core RISC-loading complex components Dicer, TRBP, and Ago2 (**Figs. 2.2 and 2.6**), and analyzed the products by EMSA and autoradiography (**Fig. 2.5a,b**). Three distinct siRNA-containing complexes were observed that correspond to those previously observed, including a slowly-migrating complex D that is thought to contain Dicer, as well as two faster migrating bands termed C1 and C2[18,30] (**Fig. 2.5b**). The normalized fraction of complex C2 decreases over a 2-h period, whereas the complex C1 and D fractions increase (with estimated rate constants of 2×10^{-3} s^{-1}, 0.8×10^{-3} s^{-1}, and 0.4×10^{-3} s^{-1}, respectively, **Fig. 2.5a**). To test for the presence of Dicer we performed supershift assays with anti-Dicer polyclonal antibody and found that only complex D specifically supershifts with this antibody (**Fig. 2.5c**). Similarly, Dicer is detected by Western blot in
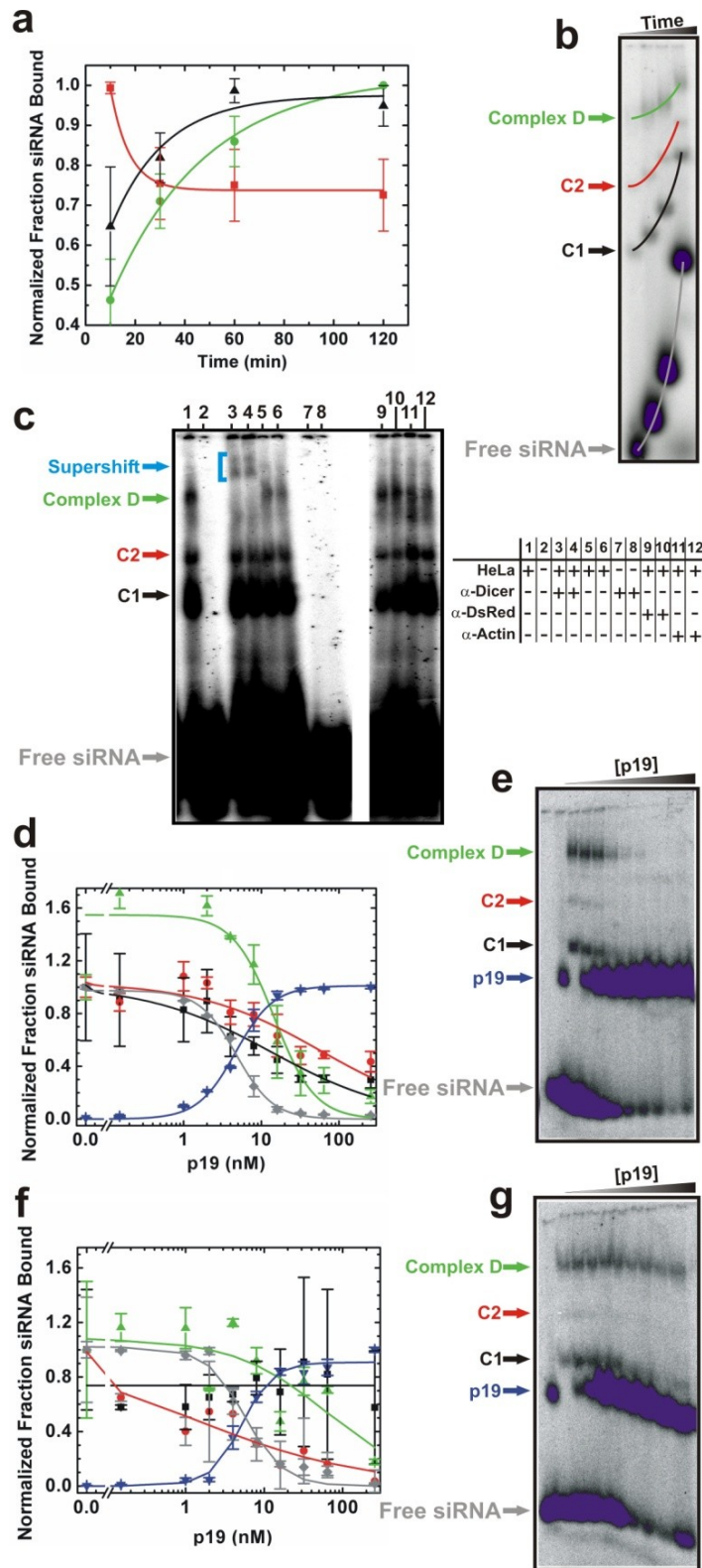
Figure 2.5. Competition of p19 with human siRNA-containing complexes found in cytosolic HeLa cell extract. (a) Normalized fraction of siRNA bound to complexes D (green), C2 (red), and C1 (black), where the highest point is

48

normalized to unity in each complex over time, as derived from the gel in panel B. The relative abundances of C1 and complex D increase over time with estimated rate constants of 8×10^{-4} s^{-1} and 4×10^{-4} s^{-1}, respectively, whereas C2 decreases in relative abundance with a rate constant of 2×10^{-3} s^{-1}. **(b)** Formation of complexes D, C2, and C1 after 10, 30, 60 and 120 min incubation of siRNA in cytosolic HeLa cell extract. Samples were loaded onto a running 4% native polyacrylamide gel, leading to the indicated differences in migration. **(c)** Anti-Dicer antibody-based supershift of siRNA-containing complexes. Dicer containing complex D shifts in the presence of anti-Dicer antibody (lane 3: 0.8 µg; lane 4: 1.6 µg). Lanes 1, 5 and 6 indicate the mobility of the complexes in the absence of any antibody. Lanes 9-12 show negative controls where 0.8 µg or 1.6 µg of anti-actin (lanes 9, 10) or anti-DsRed (lanes 11, 12) were added. **(d)** Quantification of the normalized fraction of siRNA bound in each complex, as derived from gels as in panel E. Fractions were normalized to the zero time point. **(e)** Formation of complexes after 2.5 h incubation of siRNA in cytosolic HeLa cell extract, and increasing concentrations of p19. **(f)** Quantification of the normalized fraction of siRNA bound in each complex, as derived from gels as in panel E, normalized as in D. **(g)** Formation of complexes D, C1, and C2 after 2 h incubation of siRNA in cytosolic HeLa cell extract, followed by the addition of increasing concentrations of p19 and further incubation for 30 min. Reprinted with permission from Elsevier.
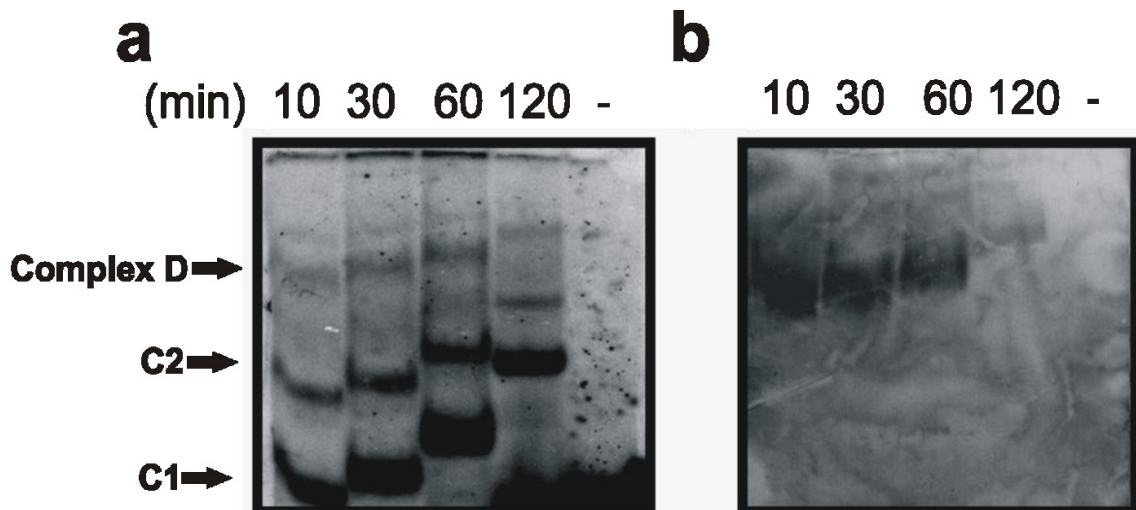
Figure 2.6. Non-denaturing western blot detection of Dicer. **(a)** A scaled-up version of the non-denaturing time course gel in (Figure 4B), imaged for radio-labeled siRNA, yields sufficient complexes for Western blot detection of Dicer. **(b)** Western blot of a similar non-denaturing gel to (a) probed with a rabbit primary antibody against Dicer, followed by a goat anti-rabbit secondary antibody, conjugated with horseradish peroxidase. Samples were loaded onto a running gel, leading to the observed differences in migration. Reprinted with permission from Elsevier.

complex D only (**Fig. 2.6**). Additionally, preliminary mass spectrometry analysis of the three siRNA containing complexes was performed to identify the proteins associated with these complexes and with the siRNA (**Appendix C**). Notably, adding increasing p19 concentrations to the cell extract prior to siRNA addition increasingly impairs the formation of all three siRNA-containing complexes, concomitant with an accumulation of the competing SP complex (**Figs. 2.5d,e**). When p19 is added to the RISC assembly assay after 2 h of pre-incubation with siRNA, the formation of particularly the Dicer-containing complex D is inhibited to a lesser extent (**Figs. 2.5f,g**). Similar results were obtained at higher HeLa cell extract concentrations and temperatures (**Appendix A, Fig A.3**).

Modeling supports formation of a transient ternary siRNA:Dicer:p19 complex. The mechanism of competition between p19 and human Dicer can be determined from mathematically modeling the two possible modes of competition for a two-protein-single-substrate system. First, competition may occur through a dissociative binding mode, where the higher affinity protein shifts the relative equilibrium of the system by binding all available substrate as it dissociates from the competitor, thus establishing a new equilibrium position (**Fig. 2.4c**). In the case of p19 and Dicer, siRNAs are pre-associated with Dicer at the onset, due to their cleavage from longer RNAs, making such dissociative binding by p19 relatively fast only if spontaneous dissociation from Dicer is rapid. Second, competition may occur through formation of a ternary complex intermediate involving siRNA, Dicer, and p19 (**Fig. 2.4c**). Using the previously determined gel-based apparent dissociation equilibrium constants for p19 (K_P = 0.17 ± 0.02 nM) and Dicer (K_D = 3.7 ± 0.4 nM), we can discriminate between these two possible modes of competition by calculating the resulting expected siRNA:Dicer and siRNA:p19 complex fractions and comparing them to experiments for a given set of equilibrium conditions.[31]

Assumption of the dissociative binding mechanism shows poor correlation with experiment, as evident from the facts that changes upon adding Dicer are consistently overestimated (**Fig. 2.4a**), while changes upon adding p19 are consistently underestimated (**Fig. 2.4b**). This poor correlation suggests that

ternary complex formation is needed to explain the observed efficient disruption of the siRNA:Dicer complex in the presence of p19. The inclusion of ternary complex formation into the mechanism creates a thermodynamic cycle that requires the ratio of K_D/K_{SP} to equal that of K_P/K_{SD}.[32] Rearranging and using the values of K_D and K_P we find that $K_{SP}/K_{SD} = K_D/K_P = 21.6$ and, therefore, $K_{SP} = 21.6 \, K_{SD}$. This result suggests that the dissociation of the ternary complex into the binary siRNA:p19 complex is >20-fold more likely than dissociation into the siRNA:Dicer complex, indicating that siRNAs can indeed be effectively shuttled toward the p19 bound state by assuming transient formation of a ternary siRNA:Dicer:p19 complex. The resulting model correlates well with experiment and explains both the weak effect that Dicer addition has on the siRNA:p19 complex and the large effect that p19 addition has on the siRNA:Dicer complex (**Figs. 2.4a,b**). A similar result is observed when p19 is placed in the context of the RNAi pathway, as follows.

Modeling illustrates the benefits of reversible siRNA binding by p19. To further interpret and unify our results, we expanded a previously published, simplified steady-state kinetic model that encapsulates key features of the RNAi pathway, including RNA cleavage by Dicer, siRNA incorporation into RISC, and the resulting RISC-mediated RNA degradation.[33] We included the experimentally observed competition of p19 with the siRNA:Dicer interaction and p19's reversible interaction with siRNA. From this model the possibility of a multiple-turnover "catch and release" mechanism emerges wherein p19 is effectively recycled after successfully competing with Dicer and dissociating again from the siRNA (**Fig. 2.7a**). We next sought to establish what impact, if any, such a mechanism has on the ability of p19 to suppress RNAi (for details of the model and its analysis see **Appendix A**).

Our model suggests that the intracellular steady-state load of viral proteins (and thus viral viability) is specified by three global parameters (**Fig. 2.7a**); the efficiencies of Dicer (*A*), RNAi in the specific host cell type (*y*), and p19 (*v*). These three global parameters are each composed of detailed mechanistic parameters (**Table 2.1** and **Appendix A**). In the case of irreversible binding of
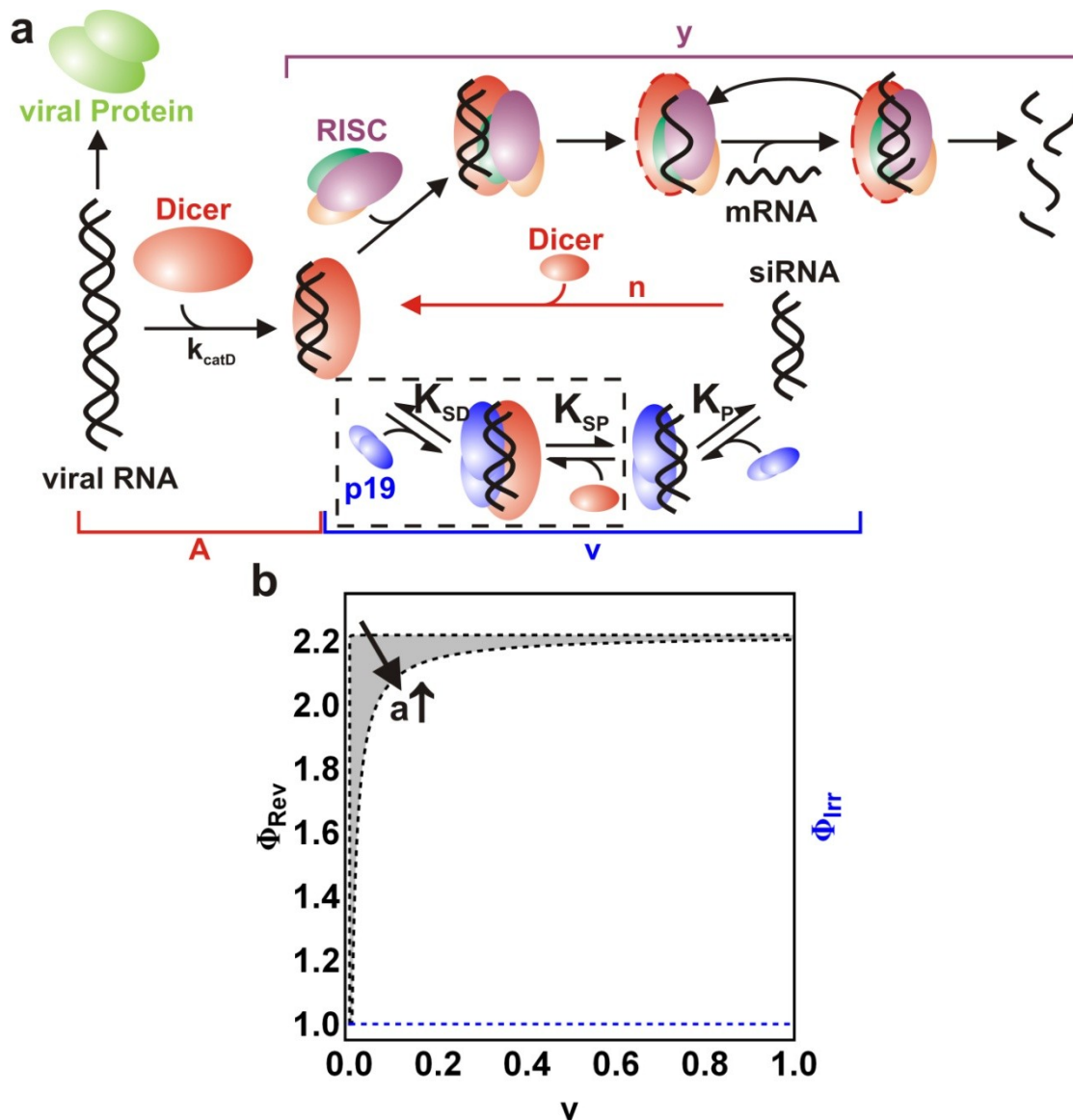
Figure 2.7. Steady-state kinetic modeling of p19 action in RNAi suppression. (a) Schematic of our expanded "catch and release" mathematical model. *A*, *y*, and *v* describe the efficiencies of Dicer, RNAi in the cell, and p19, respectively. *n* is the fraction of siRNAs that successfully reenters the RNAi pathway after release from p19. The mature human RISC (with single stranded siRNA bound) may or may not contain Dicer (dashed outline). **(b)** Enhancement of viral protein load over unsuppressed protein level, assuming reversible and irreversible p19 binding (ϕ_{Rev} and ϕ_{Irr}, respectively). The fraction of siRNA:p19 complex over free p19, *a*, will increase as the efficiency of p19 (*v*) increases. The enhancement in the level of viral protein due to reversible p19 binding will therefore fall within the range (grey area) determined for *a* ranging from 0.1 (upper bound) to 200 (lower bound). Reprinted with permission from Elsevier.

p19 to an siRNA (single turnover), the relative enhancement of the viral protein load in the presence over the absence of p19, ϕ_{Irr}, is then described by:

$$\phi_{Irr} = \frac{2Av(1+y)}{\left(-1-v+\sqrt{1+2(1+A)v+(-1+A)^2v^2}\right)y+Av(2+y)}$$

leading to a negligible dependence on the p19 suppression efficiency v (**Fig. 2.7b**) when using a realistic parameter space (**Table 2.1** and Materials and Methods).

In the case of a reversibly formed SP complex, as observed in our experimental work, two additional model parameters come into play (**Fig. 2.7a**), the fraction of siRNAs that successfully reenters the RNAi pathway after release from p19 (n) and the fraction of the total p19 that is bound to siRNA (a). Values of n close to zero lead to recycling of p19 protein without reincorporation of the released siRNA into RISC, whereas values of n close to unity effectively reduce the free concentration of p19 without changing the fate of the bound siRNA as an eventual guide for RISC. Previous studies have shown that Dicer substrates at sub-nanomolar concentrations are up to 100-fold more effective in gene knockdown than mature siRNAs,[6,34] suggesting a low probability of n = 0.01 for reincorporation of free siRNA into the RNAi pathway especially at the low siRNA concentrations expected during early stages of viral infection. Previous experimental studies in *Drosophila* embryo lysate have shown that p19 is effective in preventing low nanomolar concentrations of siRNAs from incorporation into RISC but less effective in disrupting mature RISC,[18,22] a feature also evident in our data (compare **Fig. 2.5e,g**), suggesting that indeed early stages of viral infection, where n is small, may be most effective for RNAi suppression by p19. The relative enhancement of viral protein load in the presence over the absence of p19 in our reversible binding model, ϕ_{Rev}, is then described by:

$$\phi_{Rev} = \frac{2Av(1+y)}{2Av+y(1-aA\frac{k_{off}}{k_{cat_D}}+v-Av-\sqrt{\left(-4A(-1+a\frac{k_{off}}{k_{cat_D}}(A-n))\frac{v}{\beta^2}+(-1+aA\frac{k_{off}}{k_{cat_D}}+(-1+A)v)^2\right)}}$$

where k_{off} is the dissociation constant of the SP complex and k_{cat_D} is the cleavage rate constant of Dicer. This model predicts a hyperbolic dependence of

ϕ_{Rev} on the p19 efficiency v, with little dependence on a (**Fig. 2.7b**), which can be estimated from our experimental data in **Fig. 2.4a,b** to range between 0.1-200. Reversible siRNA binding by p19 is thus found to enhance viral protein production by up to ~2.2-fold over that observed for the irreversible binding model (**Fig. 2.7b**). Notably, even a value of n = 0.1 does not diminish this effect much due to a close-to-linear decrease of ϕ_{Rev} toward unity with an increase of n toward unity (data not shown).

2.4 Discussion

The small, positive-sense, single-stranded RNA genome of tombusviruses encodes p19, an RSS thought critical for symptom spread and cytotoxicity of this class of plant pathogens[20,21,35,36] through its ability to bind and thus sequester Dicer-generated siRNAs.[14,18] The proposed use of p19 in therapeutic and research applications including mammalian gene therapy,[37] vaccine production,[38,39] and as a human retroviral protein homolog[40] motivates the study of p19 outside its natural host environment, across kingdom boundaries. We here have investigated binding of a generic siRNA by CIRV p19 through fluorescence quenching assays in combination with radioactive EMSAs. Noticeable dissociation of an siRNA from p19 (k_{off} = 0.062 ± 0.002 s^{-1}) is counterbalanced by fast association to yield a solution based dissociation equilibrium constant (K_D = 0.37 ± 0.08 nM; **Fig. 2.3**), within 2.2-fold of the high affinity previously deduced in EMSAs.[15] The dissociation of siRNA from p19 is nearly three orders-of-magnitude faster than that observed for many high-affinity nucleic acid-protein complexes with dissociation rate constants on the order of 10^{-4} s^{-1}.[24,25,41,42] Strikingly, the rate and equilibrium constants for the siRNA:p19 complex resemble those of HIV-1 protein Rev for interaction with Rev Response Element (RRE) RNA (k_{on} = 5.3 x10^8 M^{-1} s^{-1}, k_{off} = 0.14 s^{-1}), perhaps due to the similarly (primarily) non-sequence specific binding of a small protein to a (partially) base paired RNA duplex.[43]

The presence of even low concentrations of p19 dramatically weakens the interaction of recombinant human Dicer with its siRNA product (25- to >270-fold,

depending on the p19 concentration, **Fig. 2.4a**), in the absence of other detectable RISC components. Conversely, the presence of saturating concentrations of Dicer only slightly lowers (~4-fold) the affinity of p19 for siRNA (**Fig. 2.4b**), supporting the hypothesis that strong competition with the SD complex is an important mode of p19 action *in vivo*. Accordingly, assembly assays show that p19 also strongly disrupts RISC-related human siRNA-protein complexes observed in HeLa cell extract that we show to contain all core RISC-loading complex components (Dicer, Ago2, and TRBP;[5] **Fig. 2.2d,e**), suggesting that p19 is able to disrupt the SD complex even in the context of the full complement of RISC components (**Fig. 2.5**). Our experimental results inform mathematical modeling of the competition between Dicer and p19 for siRNA, as well as of the steady-state interactions between siRNA, Dicer, p19, RISC, and viral (m)RNA. We determine that the minimal model necessary to achieve the experimentally observed level of competition of p19 with Dicer for siRNA requires the postulation of a transient ternary complex that favors dissociation into the siRNA:p19 complex and free Dicer (**Fig. 2.4**). In addition, an expanded steady-state kinetic model predicts a significant enhancement in viral protein load when the siRNA:p19 interaction is assumed to be reversible, as observed experimentally. Our results thus invoke a multiple-turnover "catch and release" mechanism wherein p19, unlike the released siRNA,[6,34] is effectively recycled after successfully competing with Dicer (**Fig. 2.7**). Notably, if the dissociation rate constant of the siRNA:p19 complex were lower and thus closer to that of a slow-release protein, p19 would be more efficient in competing with Dicer for siRNA, but also less efficient in releasing the siRNA to achieve multiple-turnover. It is tempting to speculate that evolution in this way has yielded a p19 with siRNA binding properties optimized for RNAi suppression.

As any mathematical model, ours relies on some assumptions, including an experimentally derived estimate for a low probability *n* of 1% for reincorporation of siRNA released from p19 into the RNAi pathway.[6,34] The dependence of the predicted RNAi suppression has a relatively shallow dependence on *n*, and for any *n* < 1 predicts an enhancement of suppression

through the experimentally observed reversibility of the siRNA:p19 interaction. Strikingly, the estimated 2.2-fold enhancement in RNAi suppression through reversible siRNA:p19 complex formation is similar in magnitude to the typical RNAi effect that microRNAs have on the expression of cellular genes,[9,44] suggesting that even a modest change in gene expression level as predicted here can have profound biological consequences, in our case on the likelihood of survival of a virus in its host.

Several implications for the struggle between the virus and the cellular RNAi machinery arise from our work. First, in gene therapeutic applications p19 will be required to be effective outside of its natural host environment, where a study of the protein's potential interaction partners so far has been lacking. Here, we observe p19 in the context of human cellular components and find evidence for ternary complex formation between p19, an siRNA and human Dicer. Our experimental data suggest that the observed efficient competition of p19 with the siRNA:Dicer complex (**Fig. 2.4**) and the resulting Dicer-containing RISC assembly complexes (**Fig. 2.5**) targets a particularly critical step on the path to an activated RISC. Recent 3D reconstructions from negatively-stained electron microscopy images of human Dicer complexes depict it as a fist-like shape with a central flat surface that is thought to bind the siRNA.[4,45,46] This architecture along with the known binding mode of p19 raises the possibility that p19 could bind the opposite face of the siRNA in a ternary complex with Dicer. Notably, crystal structures of several other viral suppressors also appear to bind only one face of the siRNA, leaving the opposite face open to potentially accommodate proteins such as Dicer.[47] While we did not observe formation of a super-shifted band as indicative of such a ternary complex in our EMSAs with siRNA, p19, and human Dicer (**Figs. 2.5d-g**), the complex may simply be too short-lived or our EMSA too low in resolution for the complex to be detected.

Second, the previously proposed simple sequestration model would require stoichiometric p19 concentrations relative to all siRNA duplexes, native to CIRV or not, that are generated by Dicer, as well as all other small duplex RNAs bound by p19, including defective interfering RNAs,[20] the siRNA amplification

products generated in plant and other cells by RNA-dependent RNA polymerase activity,[2,48] and perhaps miRNA duplexes.[19,49,50] We independently confirmed the ability of p19 to bind a human miRNA using the conserved, ubiquitous, and well-studied let-7a miRNA, which has been implicated as a tumor suppressant.[51] Remarkably, we observed that p19 binds the partially mismatched, double-stranded 22-nucleotide let-7a miRNA and our 21-nucleotide siRNA with equal affinities of 0.16 ± 0.02 nM and 0.17 ± 0.03 nM, respectively (**Appendix A, Fig. A.6**). Similar to our observations, a recent study used an affinity-capture approach to show that p19 from CIRV retains high affinity for let-7 miRNA, as well as a fluorophore labeled siRNA.[50] These observations further underscore that p19 may face a large background of competitor RNAs while targeting its siRNA substrate, although the specific base pairing patterns of some plant miRNAs may disfavor their binding to p19.[52] Our finding that formation of the high-affinity siRNA:p19 complex is readily reversible (**Fig. 2.3**) while still interfering with siRNA binding by Dicer and RISC-related assembly complexes *in vitro* (**Figs. 2.4 and 2.5**) raises the possibility that sub-stoichiometric p19 concentrations may suffice *in vivo* to prevent siRNAs by multiple-turnover from assembling into RISC. In this fashion, the virus could keep pace with the amplified host response during systemic RNAi self-defense of the plant host against CIRV.[53]

Third, while our results have an immediate impact on our understanding of the physicochemical properties and biological function of p19, a more practical consideration also arises from our work. The future of RNAi based strategies may well be dependent on the ability to control and predict the relative intracellular expression levels of RISC components and potential therapeutics.[37,54] Short hairpin RNAs (shRNA), for example, while shown *in vivo* to efficiently induce gene silencing, in long-term investigations have been found to cause dose-dependent liver injury and ultimately death in adult mice when sustained at high levels,[54] and to cause dysregulation of microRNAs and increased tumorigenesis when marginally dosed.[55] Co-expression with p19 could help ameliorate such toxic effects of shRNAs. In addition, CIRV p19 and other

RSS's have been shown to aid in the transfection and difficult over-expression of alpha-, adeno-, and lentiviruses used in gene therapy and in plant based vaccines, through their ability to suppress the intracellular RNAi immune response.[37] For such immunosuppression to be effective, dosing of the RSS is important, analogous to the use of immunosuppressants after organ transplant surgery. In fact, our observation that at least some miRNAs tightly bind to p19 cautions that adverse developmental effects due to a sequestration of essential cellular RNAs has to be carefully avoided. Understanding and modeling relevant biophysical properties of RNAi processes in the presence of RSSs, as advanced here, is thus anticipated to become critical for a broader and safer use of RNAi therapeutics in human disease control.

2.5 Acknowledgments

Name	Description (units)	Determined From	Value
kcatD	Catalytic Rate of Dicer (nM / h)	Literature[61]	8.143×10^6
rdp	Protein Degradation Rate Constant (h^{-1})	Literature[62]	0.35
rds	siRNA Degradation Rate Constant (h^{-1})	Literature[62]	0.029
koff	p19 Dissociation Rate Constant (h^{-1})	Determined Experimentally	223.2
kon	p19 Association Rate Constant ($nM^{-1} h^{-1}$)	Determined Experimentally	608.4
n	Fraction of siRNA that reenter the RNAi pathway	Literature[63]	0.01
a	Fraction of bound verses unbound p19	Estimated from experimental data	1 - 100
Cellular Volume	Intracellular volume (L)	Literature[62]	1.4×10^{-12}
Grouped Parameters			
y	RNAi Efficacy Parameter specific to HEK293 cells	Literature[33]	1.22
A	Dicer Efficacy Parameter	kcatD / rdp	2.33×10^7
β	RNAi Settling Time	rds / rdp[33]	8.29×10^{-2}
v	p19 Efficacy Parameter	Range	0 - 1

Table 2.1. Model Parameters

2.6 References

1. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-881.
2. Carthew, R. W. & Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136**, 642-655.
3. Castanotto, D. & Rossi, J. J. (2009). The promises and pitfalls of RNA-interference-based therapeutics. *Nature* **457**, 426-433.
4. Sashital, D. G. & Doudna, J. A. (2010). Structural insights into RNA interference. *Curr Opin Struct Biol* **20**, 90-97.
5. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. & Doudna, J. A. (2008). In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* **105**, 512-517.
6. Kim, D. H., Behlke, M. A., Rose, S. D., Chang, M. S., Choi, S. & Rossi, J. J. (2005). Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat Biotechnol* **23**, 222-226.
7. Collingwood, M. A., Rose, S. D., Huang, L., Hillier, C., Amarzguioui, M., Wiiger, M. T. *et al.* (2008). Chemical modification patterns compatible with high potency dicer-substrate small interfering RNAs. *Oligonucleotides* **18**, 187-200.
8. Ji, X. (2008). The mechanism of RNase III action: how dicer dices. *Curr Top Microbiol Immunol* **320**, 99-116.
9. Siomi, H. & Siomi, M. C. (2009). On the road to reading the RNA-interference code. *Nature* **457**, 396-404.
10. Jaskiewicz, L. & Filipowicz, W. (2008). Role of Dicer in posttranscriptional RNA silencing. *Curr Top Microbiol Immunol* **320**, 77-97.
11. Gitlin, L. & Andino, R. (2003). Nucleic acid-based immune system: the antiviral potential of mammalian RNA silencing. *J Virol* **77**, 7159-7165.
12. Parameswaran, P., Sklan, E., Wilkins, C., Burgon, T., Samuel, M. A., Lu, R. *et al.* (2010). Six RNA viruses and forty-one hosts: viral small RNAs and modulation of small RNA repertoires in vertebrate and invertebrate systems. *PLoS Pathog* **6**, e1000764.
13. Qu, F. & Morris, T. J. (2005). Suppressors of RNA silencing encoded by plant viruses and their role in viral infections. *FEBS Lett* **579**, 5958-5964.
14. Scholthof, H. B. (2006). The Tombusvirus-encoded P19: from irrelevance to elegance. *Nat Rev Microbiol* **4**, 405-411.
15. Vargason, J. M., Szittya, G., Burgyan, J. & Hall, T. M. (2003). Size selective recognition of siRNA by an RNA silencing suppressor. *Cell* **115**, 799-811.
16. Ye, K., Malinina, L. & Patel, D. J. (2003). Recognition of siRNA by a viral suppressor of RNA silencing. *Nature* **426**, 874-878.
17. Xia, Z., Zhu, Z., Zhu, J. & Zhou, R. (2009). Recognition mechanism of siRNA by viral p19 suppressor of RNA silencing: a molecular dynamics study. *Biophys J* **96**, 1761-1769.
18. Lakatos, L., Szittya, G., Silhavy, D. & Burgyan, J. (2004). Molecular mechanism of RNA silencing suppression mediated by p19 protein of tombusviruses. *EMBO J* **23**, 876-884.

19. Dunoyer, P., Lecellier, C. H., Parizotto, E. A., Himber, C. & Voinnet, O. (2004). Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. *Plant Cell* **16**, 1235-1250.

20. Havelda, Z., Hornyik, C., Valoczi, A. & Burgyan, J. (2005). Defective Interfering RNA Hinders the Activity of a Tombusvirus-Encoded Posttranscriptional Gene Silencing Suppressor. *J Virol* **79**, 450-457.

21. Omarov, R., Sparks, K., Smith, L., Zindovic, J. & Scholthof, H. B. (2006). Biological relevance of a stable biochemical interaction between the tombusvirus-encoded P19 and short interfering RNAs. *J Virol* **80**, 3000-3008.

22. Lakatos, L., Csorba, T., Pantaleo, V., Chapman, E. J., Carrington, J. C., Liu, Y. P. *et al.* (2006). Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *EMBO J* **25**, 2768-2780.

23. Merai, Z., Kerenyi, Z., Kertesz, S., Magna, M., Lakatos, L. & Silhavy, D. (2006). Double-stranded RNA binding may be a general plant RNA viral strategy to suppress RNA silencing. *J Virol* **80**, 5747-5756.

24. Mitsis, P. G. & Wensink, P. C. (1989). Identification of yolk protein factor 1, a sequence-specific DNA-binding protein from Drosophila melanogaster. *J Biol Chem* **264**, 5188-5194.

25. Ramsden, J. J. & Dreier, J. (1996). Kinetics of the interaction between DNA and the type IC restriction enzyme EcoR124II. *Biochemistry* **35**, 3746-3753.

26. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. & Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563-574.

27. Hoerter, J. A. & Walter, N. G. (2007). Chemical modification resolves the asymmetry of siRNA strand degradation in human blood serum. *RNA* **13**, 1887-1893.

28. Pellino, J. L., Jaskiewicz, L., Filipowicz, W. & Sontheimer, E. J. (2005). ATP modulates siRNA interactions with an endogenous human Dicer complex. *RNA* **11**, 1719-1724.

29. Pham, J. W. & Sontheimer, E. J. (2005). Separation of Drosophila RNA silencing complexes by native gel electrophoresis. *Methods Mol Biol* **309**, 11-16.

30. Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W. & Sontheimer, E. J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. *Cell* **117**, 83-94.

31. Sigurskjold, B. W. (2000). Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal Biochem* **277**, 260-266.

32. Wreggett, K. A. & De Lean, A. (1984). The ternary complex model. Its properties and application to ligand interactions with the D2-dopamine receptor of the anterior pituitary gland. *Mol Pharmacol* **26**, 214-227.

33. Marshall, W. F. (2008). Modeling recursive RNA interference. *PLoS Comput Biol* **4**, e1000183.

34. Hefner, E., Clark, K., Whitman, C., Behlke, M. A., Rose, S. D., Peek, A. S. *et al.* (2008). Increased potency and longevity of gene silencing using validated Dicer substrates. *J Biomol Tech* **19**, 231-237.

35. Chu, M., Desvoyes, B., Turina, M., Noad, R. & Scholthof, H. B. (2000). Genetic dissection of tomato bushy stunt virus p19-protein mediated host-dependent symptom induction and systemic invasion. *Virology* **266**, 79-87.

36. Silhavy, D., Molnar, A., Lucioli, A., Szittya, G., Hornyik, C., Tavazza, M. *et al.* (2002). A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *EMBO J.* **21**, 3070-3080.

37. de Vries, W., Haasnoot, J., van der Velden, J., van Montfort, T., Zorgdrager, F., Paxton, W. *et al.* (2008). Increased virus replication in mammalian cells by blocking intracellular innate defense responses. *Gene Ther* **15**, 545-552.

38. Lombardi, R., Circelli, P., Villani, M. E., Buriani, G., Nardi, L., Coppola, V. *et al.* (2009). High-level HIV-1 Nef transient expression in Nicotiana benthamiana using the P19 gene silencing suppressor protein of Artichoke Mottled Crinckle Virus. *BMC Biotechnol* **9**, 96.

39. Zheng, N., Xia, R., Yang, C., Yin, B., Li, Y., Duan, C. *et al.* (2009). Boosted expression of the SARS-CoV nucleocapsid protein in tobacco and its immunogenicity in mice. *Vaccine* **27**, 5001-5007.

40. Qian, S., Zhong, X., Yu, L., Ding, B., de Haan, P. & Boris-Lawrie, K. (2009). HIV-1 Tat RNA silencing suppressor activity is conserved across kingdoms and counteracts translational repression of HIV-1. *Proc Natl Acad Sci U S A* **106**, 605-610.

41. Mougel, M., Ehresmann, B. & Ehresmann, C. (1986). Binding of Escherichia coli ribosomal protein S8 to 16S rRNA: kinetic and thermodynamic characterization. *Biochemistry* **25**, 2756-2765.

42. Seimiya, M. & Kurosawa, Y. (1996). Kinetics of binding of Antp homeodomain to DNA analyzed by measurements of surface plasmon resonance. *FEBS Lett* **398**, 279-284.

43. Pond, S. J., Ridgeway, W. K., Robertson, R., Wang, J. & Millar, D. P. (2009). HIV-1 Rev protein assembles on viral RNA one molecule at a time. *Proc Natl Acad Sci USA* **106**, 1404-1408.

44. Krol, J., Loedige, I. & Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* **11**, 597-610.

45. Lau, P. W., Potter, C. S., Carragher, B. & MacRae, I. J. (2009). Structure of the human Dicer-TRBP complex by electron microscopy. *Structure* **17**, 1326-1332.

46. Wang, H. W., Noland, C., Siridechadilok, B., Taylor, D. W., Ma, E., Felderer, K. *et al.* (2009). Structural insights into RNA processing by the human RISC-loading complex. *Nat Struct Mol Biol* **16**, 1148-1153.

47. Yang, J. & Yuan, Y. A. (2009). A structural perspective of the protein-RNA interactions involved in virus-induced RNA silencing and its suppression. *Biochim Biophys Acta* **1789**, 642-652.

48. Lipardi, C. & Paterson, B. M. (2009). Identification of an RNA-dependent RNA polymerase in Drosophila involved in RNAi and transposon suppression. *Proc. Natl. Acad. Sci. USA* **106**, 15645-15650.

49. Chapman, E. J., Prokhnevsky, A. I., Gopinath, K., Dolja, V. V. & Carrington, J. C. (2004). Viral RNA silencing suppressors inhibit the microRNA pathway at an intermediate step. *Genes Dev* **18**, 1179-1186.

50. Jin, J., Cid, M., Poole, C. B. & McReynolds, L. A. (2010). Protein mediated miRNA detection and siRNA enrichment using p19. *Biotechniques* **48**, xvii-xxiii.

51. Boyerinas, B., Park, S. M., Hau, A., Murmann, A. E. & Peter, M. E. (2011). The role of let-7 in cell differentiation and cancer. *Endocr Relat Cancer* **17**, F19-36.

52. Varallyay, E., Valoczi, A., Agyi, A., Burgyan, J. & Havelda, Z. (2010). Plant virus-mediated induction of miR168 is associated with repression of ARGONAUTE1 accumulation. *EMBO J* **29**, 3507-3519.

53. Baulcombe, D. (2004). RNA silencing in plants. *Nature* **431**, 356-363

54. Grimm, D., Streetz, K. L., Jopling, C. L., Storm, T. A., Pandey, K., Davis, C. R. *et al.* (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* **441**, 537-541.

55. Beer, S., Bellovin, D. I., Lee, J. S., Komatsubara, K., Wang, L. S., Koh, H. *et al.* (2010). Low-level shRNA Cytotoxicity Can Contribute to MYC-induced Hepatocellular Carcinoma in Adult Mice. *Mol Ther* **18**, 161-170.

56. Walter, N. G. (2002). Probing RNA structural dynamics and function by fluorescence resonance energy transfer. In *Curr. Protocols Nucleic Acid Chem*, pp. 11.10.11-11.10.23.

57. Gondert, M. E., Tinsley, R. A., Rueda, D. & Walter, N. G. (2006). The catalytic core structure of the trans-acting HDV ribozyme is subtly influenced by sequence variation outside the core. *Biochemistry* **45**, 7563-7573.

58. Rueda, D., Wick, K., McDowell, S. E. & Walter, N. G. (2003). Diffusely bound Mg2+ ions slightly reorient stems I and II of the hammerhead ribozyme to increase the probability of formation of the catalytic core. *Biochemistry* **42**, 9924-9936.

59. Gerczei, T., Shah, B. N., Manzo, A. J., Walter, N. G. & Correll, C. C. (2009). RNA chaperones stimulate formation and yield of the U3 snoRNA-Pre-rRNA duplexes needed for eukaryotic ribosome biogenesis. *J Mol Biol* **390**, 991-1006.

60. Groenenboom, M. A., Maree, A. F. & Hogeweg, P. (2005). The RNA silencing pathway: the bits and pieces that matter. *PLoS Comput Biol* **1**, 155-165.

61. Ma, E., MacRae, I. J., Kirsch, J. F. & Doudna, J. A. (2008). Autoinhibition of human dicer by its internal helicase domain. *J Mol Biol* **380**, 237-243.

62. Bartlett, D. W. & Davis, M. E. (2006). Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging. *Nucleic Acids Res* **34**, 322-333.

63. Rose, S. D., Kim, D. H., Amarzguioui, M., Heidel, J. D., Collingwood, M. A., Davis, M. E. *et al.* (2005). Functional polarity is introduced by Dicer processing of short substrate RNAs. *Nucleic Acids Res* **33**, 4140-4156.

Chapter 3

siRNA-like double-stranded RNAs are specifically protected against degradation in human cell extract[3, 4]

3.1 Introduction

RNA interference (RNAi), first described in the nematode *C. elegans*[1], is a set of conserved post-transcriptional gene silencing pathways in eukaryotes. Following the first demonstration that RNAi is functional in human cells and receptive to using synthetic small interfering (si)RNA effector molecules[2], significant progress has been made in harnessing the RNAi pathway for functional genomics and for therapies targeting previously "undruggable" genetic targets[3, 4, 5, 6]. Inside the cell, double-stranded siRNAs specifically interact with a number of proteins including the endoribonuclease Dicer that cleaves them from long precursor RNAs and steers incorporation of this product into a multiprotein complex referred to as the RNA-induced silencing complex (RISC). One strand of the siRNA duplex is retained in RISC and acts as a template for the sequence-specific identification and site-specific cleavage of messenger (m)RNAs and concomitant reduction in gene expression.

The inherent sensitivity of RNA molecules both inside and outside the cell towards degradation by ribonucleases motivated an immediate focus on

---

development of chemically modified siRNAs that can withstand nucleolytic extracellular environments and reach the cellular milieu intact[3]. Investigationinto the determinants of siRNA stability have focused on extracellular stability, intracellular stability, and potency of silencing. First, to mimic the extracellular environment several groups have used blood serum and purified nucleases to determine the sensitivity of siRNAs to degradation. Chemical modifications were found to impart a considerable degree of nuclease resistance to siRNAs while retaining the ability to induce silencing through the RNAi pathway[7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18]. Furthermore, chemical modifications were found to be absolutely essential for systemic delivery of siRNA, both to mediate binding to serum proteins to increase circulating half-life and to block sites vulnerable to nuclease attack[14,15]. Only few exceptions to this tendency towards RNA-hostile extracellular environments exist[19].

The intracellular milieu, which is the functional environment of the siRNA, is the second environment to which the siRNA is exposed. By monitoring the expression of a target gene, several groups were able to assess the potency and duration of siRNA effects, including a number of studies that focused on determining whether chemically modified siRNAs are more potent than unmodified siRNAs[7, 8, 9, 10, 11, 13, 16]. A recent study has shown, however, that enhanced intracellular nucleolytic stability is not necessarily correlated with increased duration of the silencing effect[8]. In fact, the authors found that silencing in non-dividing cells persisted for up to one month from a single dose of an unmodified siRNA, suggesting that siRNAs may be quite stable inside the cell.

Several groups have addressed questions of intracellular siRNA stability and localization by introducing fluorophore modified siRNAs into live cells and using various microscopy techniques[20]. siRNAs are actively exported from the nucleus[21], except in cases where the RNA target is located in the nucleus[22]. Fluorescence fluctuation spectroscopy has been utilized in a separate study to assess the integrity of labeled intracellular RNAs, revealing that doubly-labeled RNA suitable for fluorescence resonance energy transfer (FRET) measurement between the fluorophores is relatively unstable in single-stranded form compared

to the corresponding siRNA duplex[23]. Intracellular FRET imaging of dsRNAs has also been employed to show that intact siRNA duplexes accumulate in cellular foci identified as P-bodies[24, 25]. We have used FRET labeled single-stranded RNAs to show that secondary structure in general attenuates degradation in human cell extracts[26].

The precise origin of the apparent intracellular protection of siRNAs is unclear. Long duration of silencing in non-dividing cells may be indicative of the stability of the single-stranded guide strand incorporated into RISC. Prior to RISC maturation, it can be assumed that some stability is derived from the decreased nuclease sensitivity of double-stranded RNA compared to single-stranded RNA [23; 26]. Previous intracellular fluorescence studies[21, 22, 23, 24, 25] have focused, however, only on canonical 21-22 nucleotide (nt, per strand) siRNAs, while the observation that the RNAi pathway can utilize 21- to 27-nt double-stranded RNAs[27] suggests that intracellular double-stranded RNAs are discriminated based on length.

To better evaluate the potential of siRNAs in gene therapeutics, a real-time characterization of their degradation kinetics under extracellular and intracellular conditions is necessary. Such an assay should include rapid and precise assessment of RNA stability as a function of bodily fluid, base paring partner, and RNA size and should eventually be amenable to high-throughput screening for optimizing siRNA drugs. To this end, we here have developed a generalizable series of RNAs modified for FRET measurements, where two uridines in positions 7 and 16 of the 21-nt guide strand were replaced with amino-modified deoxythymidines to introduce two fluorophores that undergo distance dependent FRET (**Fig. 3.1**). We varied the size of the labeled strand to generate double-stranded (ds)RNAs ranging from 18-24 nt per strand (while keeping the fluorophore distance constant, **Fig. 3.1**). We find that our dsRNA probes interact with enzymes in RNAi-active cytosolic extract from human HeLa cells, including RNase H1 (21-nt DNA/RNA hybrid) and Dicer (24-nt dsRNA). The DNA/RNA hybrid degradation by RNase H1 is quite efficient, with observed rate constants exceeding those for the single-stranded RNA, which is not expected to

be an RNase H1 substrate. Furthermore, our experimental design allows us to correlate the previously studied potency/size relationships for siRNAs[27, 28] with a real-time readout of RNA stability. Our results demonstrate that 21- and 24-nt dsRNAs are protected in cell extract relative to the 18-nt dsRNA, potentially contributing to the longevity of intracellular RNAi effects.

3.2 Materials and Methods

Nucleic acid synthesis and labeling. All RNAs were synthesized by the HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory at the Yale University School of Medicine, deprotected, TMR labeled (as necessary), and purified as previously described[39, 40, 41]. More specifically, the guide strand was labeled during synthesis with Fluorescein dT and Amino-Modifier C6 dT in the positions indicated in **Fig. 3.1**; the amino-modifier was subsequently coupled to TMR succinimidyl ester (Invitrogen), the RNA purified away from free dye by C18-reversed phase HPLC, and its purify checked by denaturing polyacrylamide gel electrophoresis[39, 40, 41]. Sequences are derived from the firefly luciferase gene and have been shown to be effective in RNAi-mediated target cleavage[12, 28]. All RNAs were purchased with a 5' phosphate group, except for a second fluorophore-labeled 21-nt guide strand with a 5' OH used for radiolabeling. The DNA oligonucleotides used in our studies (Sequence = VI-2, V-2, and S1-DNA)[30] were synthesized by Invitrogen and had the following sequences: VI-2, CGGTCGCTCCGTGTGGCTTGGGTTGGGTGTGGCAGTGAC; V-2, GCTGGTCTCTGCGGGTTGTTGCGCCGCGGCACCCTTGGCA; S1-DNA, CGTACGCGGAATACTTTGAAA.

Steady-state FRET assays of RNA degradation. A final concentration of 50 nM FL and TMR doubly-labeled RNAs was annealed either alone (single-stranded RNA) or with 100 nM (final concentration, in excess for saturation) of an appropriate DNA or RNA complement by heating to 70º C for 2 min and cooling to room temperature over 10 min in a standard buffer of 50 mM Tris-HOAc, pH 7.4, 80 mM KCl, 20 mM NaCl, and 1 mM MgCl$_2$ (chosen to mimic physiologic conditions). After the 10-min cooling period at room temperature, RNA solutions

were incubated in a circulating water bath at 37° C for 5 min before the samples were added to a fluorometer cuvette, also controlled at 37° C. After the initiation of collection of donor (F_{520}) and acceptor fluorescence emission signal (F_{585}) as described previously[41], a baseline measurement was recorded for approximately 50-100 seconds before a fraction of either HeLa S100 cytoplasmic cell extract (a gift from Danny Reinberg, Department of Biochemistry, Rutgers University, prepared and dialyzed following published protocols[42]), fetal bovine serum (Gibco), or purified Dicer (Stratagene, 2 Units) was added and the resulting signal evolution recorded. The FRET ratio Q ($F_{585}/(F_{585}+F_{520})$) was plotted over time t and fit where appropriate with either the single- or double-exponential form of the following expression to extract the observed rate constants k_{obs} (please note that in case of a decreasing data curve the pre-exponential factor A is positive, otherwise negative):

$$y = y_0 + A_1 e^{-k_{obs,1}t} + A_2 e^{-k_{obs,2}t}$$ (Eq. 1)

These assays were carried out over a range of 0-30% (v/v) HeLa cell extract and serum volume additions, equivalent to 0-2.5 mg/ml total HeLa protein (concentration measured by Bradford assay) and 0-11.1 mg/ml total serum protein (concentration provided by Gibco). The observed rate constants were plotted as a function of total protein content and fit with the following hyperbolic binding equation that does not assume cooperativity:

$$y = \frac{v_{max}[protein]}{protein_{1/2} + [protein]}$$ (Eq. 2)

In this equation, v_{max} is the maximal rate constant at saturation, and $protein_{1/2}$ is the half-titration point.

RNase H1 inhibition assays. Fluorometer assays to monitor RNase H1 inhibition by DNA aptamers VI-2 and V-2[30] were conducted identically to the fluorometer assays described above with the following exceptions. The HeLa cell extract used in these assays was pre-incubated with a defined concentration of the pre-annealed DNA aptamer or the negative control DNA S1 for 30 min at 37° C before mixing with the DNA/RNA hybrid in the fluorometer. In addition, the

DNA/RNA hybrid was prepared in the presence of the respective aptamer or control DNA at a concentration equal to that in the HeLa cell extract. Again, the FRET Ratio Q for each trace was extracted and fit with a single-exponential expression as above, and the observed rate constants were plotted as a function of aptamer concentration and fit with the following hyperbolic binding equation below where no cooperativity was assumed:

$$y = v_0 - \frac{(v_0 - v_{\min})[aptamer]}{aptamer_{1/2} + [aptamer]} \qquad \text{(Eq. 3)}$$

Here, v_{min} is the minimal rate constant reached at saturation and $aptamer_{1/2}$ is the half-titration point.

FRET-monitored gel electrophoretic analysis of RNA fragments. Denaturing gel electrophoresis was paired with FRET detection of the RNA to gain qualitative insight into the processing of the labeled RNAs by Dicer, cell extract, and serum. The doubly-labeled RNA strand was prepared at 500 nM final concentration in standard buffer, the complementary RNA or DNA as necessary added to 1 µM concentration and annealed as described above. After cooling, Dicer, cell extract, or serum was added as indicated and the samples were incubated at 37° C until the reactions were stopped by mixing with a final concentration of 10% (v/v) Contrad 70 at pH 9.3 (Decon Labs) as previously described[26]. For the samples of the FRET gel in **Fig. 3.2C**, serum was added to a final total protein concentration of 11.1 mg/ml for the DNA/RNA hybrid and dsRNA, but only 0.93 mg/ml for the ssRNA. Time points for these reactions were as follows: ssRNA, 1, 5, and 10 min; DNA/RNA hybrid and dsRNA, 0.5, 1 and 2.5 h. For the samples of the FRET gel in **Fig. 3.2G**, HeLa cell extract was added throughout to a final total protein concentration of 2.49 mg/ml and time points were taken at 5, 10, and 20 min for the ssRNA and the DNA/RNA hybrid, while time points of 5, 20, and 60 min were taken for the dsRNA. The samples for the FRET gel in **Fig. 3.4B** were prepared by exposing each of the 18, 21, and 24-nt double-stranded RNAs to Dicer enzyme (1 unit), HeLa cell extract (2.49 mg/ml), or buffer alone for 200 min (3.3 h). The buffer for the HeLa and buffer only reactions was composed of our standard buffer supplemented with 10 mM DTT,

and the Dicer reactions contained the buffer supplied and recommended by the manufacturer whose final composition was 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, and 2.5 mM $MgCl_2$. The T1 and OH lanes in the gels refer to reference ladders using RNase T1 and alkali (OH⁻) digestion, respectively, prepared as described[12] except that no carrier tRNA was added to the RNase T1 digest. For all reactions, 10 pmol/lane of labeled RNA was mixed with loading buffer for a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue, and 40% (v/v) formamide, and the products were separated on a denaturing, 8 M urea, 20% (w/v) polyacrylamide gel. Fluorescence detection of the RNA was accomplished using a FluorImager SI fluorescence scanner with ImageQuant software (Molecular Dynamics) as previously described[43, 44].

Radioactivity-monitored gel electrophoretic analysis of RNA fragments and RISC-related RNA:protein complexes. The 21-nt non-phosphorylated, doubly-fluorophore labeled RNA was 5'-^{32}P labeled with T4 polynucleotide kinase (PNK) and γ-^{32}P-ATP, mixed with loading buffer to give a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue, and 40% (v/v) formamide, and gel purified on a denaturing, 8 M urea, 20% (w/v) polyacrylamide gel. The radiolabeled full-length RNA was excised from the gel, diffusion eluted into 1 mM EDTA overnight, and ethanol precipitated. The dried RNA was dissolved in water.

5'-^{32}P labeled RNA (50,000 cpm/lane) was mixed with 50 nM non-radiolabeled, but 5' phosphorylated, doubly-fluorophore labeled RNA and 100 nM of the complementary 21-nt RNA strand, and annealed as described above in standard buffer. Degradation was initiated by the addition of a final total protein concentration of 2.49 mg/ml cell extract or 11.1 mg/ml serum at 37°C. For the analysis in **Fig. 3.5D,** the assays were carried out over a range of 0-30% (v/v) HeLa cell extract. Time points were taken after 2, 5, 10, 20, 40, and 80 min for the cell extract sample and 5, 10, 20, 40, 80, 120, and 180 min for the serum sample. For the analysis in **Fig. 3.5D**, time points were instead taken 2, 5, 10, 20 and 40 min after the addition of cell extract. All time points were quenched by mixing with a final concentration of 10% (v/v) Contrad 70 at pH 9.3 (Decon Labs)[26]. Loading buffer was added to a final concentration of 1×TBE, 0.025%

(w/v) bromophenol blue, 0.025% (w/v) xylene cyanol, and 40% (v/v) formamide, and the degradation products separated on a denaturing, 8 M urea, 20% (w/v) polyacrylamide gel. Individual bands were identified by comparison with sequencing ladders from partial digestion with G-specific RNase T1 and alkali as described[45], quantified and normalized to the sum of all bands in a lane using a PhosphorImager Storm 840 with Image Quant software (Molecular Dynamics). Any error bars are derived from three independent determinations.

For the radioactive dsRNA degradation data in **Fig. 3.5D** the normalized ratio of intact dsRNA over the sum of all bands in a lane was plotted over time t and fitted with a simple linear regression line to estimate the observed rate constant k_{obs}:

$$y = y_0 - k_{obs}.t \qquad\qquad\qquad\qquad \text{(Eq. 4)}$$

These assays were carried out over a range of 0-30% (v/v) added HeLa cell extract, equivalent to 0-2.5 mg/ml total HeLa protein (concentration measured by Bradford assay). The observed rate constants were plotted as a function of total protein content and fitted as above with hyperbolic binding equation 2.

The gel electrophoretic mobility shift assays of RISC-related RNA:protein complexes were initiated by the addition of a final total protein concentration of 2.49 mg/ml cell extract to the annealed 5'-^{32}P labeled RNA in standard buffer at 4°C. After 2 h incubation the reaction was stopped by mixing with non-denaturing loading buffer to a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue and 0.025% (w/v) xylene cyanol and the RNA-associated complexes were separated by non-denaturing 4% (w/v) polyacrylamide gel electrophoresis at 4°C, quantified and normalized to the sum of all bands in a lane using a Typhoon 9410 Variable Mode Imager with Image Quant software (GE Healthcare), essentially as described[35]. Three independent analyses gave similar results.

3.3 Results

We here have developed a set of fluorescein (FL) and tetramethylrhodamine (TMR) modified RNAs (or, more precisely, RNA-DNA chimeras with two 2'-deoxy modifications necessary for fluorophore labeling; **Fig. 3.1**) for use in studies of intracellular and extracellular RNA stability by real-time FRET. To specifically address the question of how double-stranded RNAs are recognized as substrates by intra- and extracellular ribonucleases, we developed three dsRNAs (**Fig. 3.1**) of 18, 21, and 24 nt length that bracket the canonical human siRNA length and are derived from an siRNA previously developed and validated against the *pp-luc* luciferase gene[12, 28]. We fluorophore labeled the strand that dominantly acts as guide for the RNAi machinery and is antisense to the luciferase mRNA[12, 28]. We first sought to compare the degradation of these dsRNAs with those of the corresponding, relatively unstructured single-stranded RNA (ssRNA, **Fig. 3.1**) and its hybrid with a fully complementary DNA under our standard conditions of near-physiologic buffer (50 mM Tris-HOAc, pH 7.4, 80 mM KCl, 20 mM NaCl, and 1 mM $MgCl_2$) at 37° C.

Double-stranded RNA is stabilized in both cell extract and serum. To establish the relative degradation of a 21-nt dsRNA compared to a DNA/RNA hybrid and ssRNA, we applied our steady-state FRET assay to compare the degradation of these complexes upon the addition of 30% (v/v) of either fetal bovine serum (FBS) or HeLa S100 cytosolic cell extract, equivalent to 11.1 mg/ml total serum and 2.49 mg/ml total cellular protein, respectively. The blood serum contains a significant fraction of RNase A-like enzymes[29], whereas the cell extract contains the protein components of the RNAi pathway and was previously tested as RNAi-active[12].

In general, a decrease in FRET ratio should be indicative of degradation of the labeled strand, as cleavage between the fluorophore labels is expected to lead to a decrease in FRET efficiency. Surprisingly, when the dsRNAs were incubated in serum, we observed two fast phases representing increases in
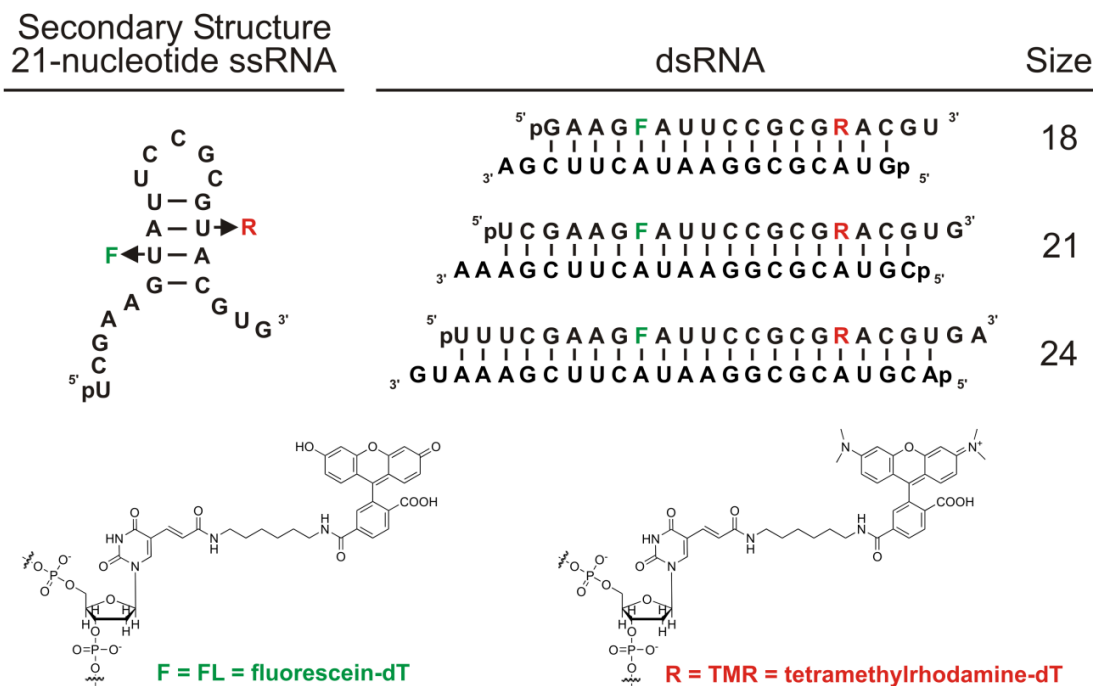
Figure 3.1. Design of the single- and double-stranded RNAs used in this study. The sequences for the RNAs are derived from the *pp-luc* firefly luciferase gene and the corresponding siRNA previously reported[12,28]. The double-stranded RNAs consist of an identical central sequence and internal FL and TMR modifications, and the flanking sequences are derived from the luciferase gene. All dsRNAs are fully Watson-Crick base paired and contain the 2-nt 3' overhangs characteristic of siRNAs; "p" indicates 5' phosphate. Reprinted with permission from PLOS.
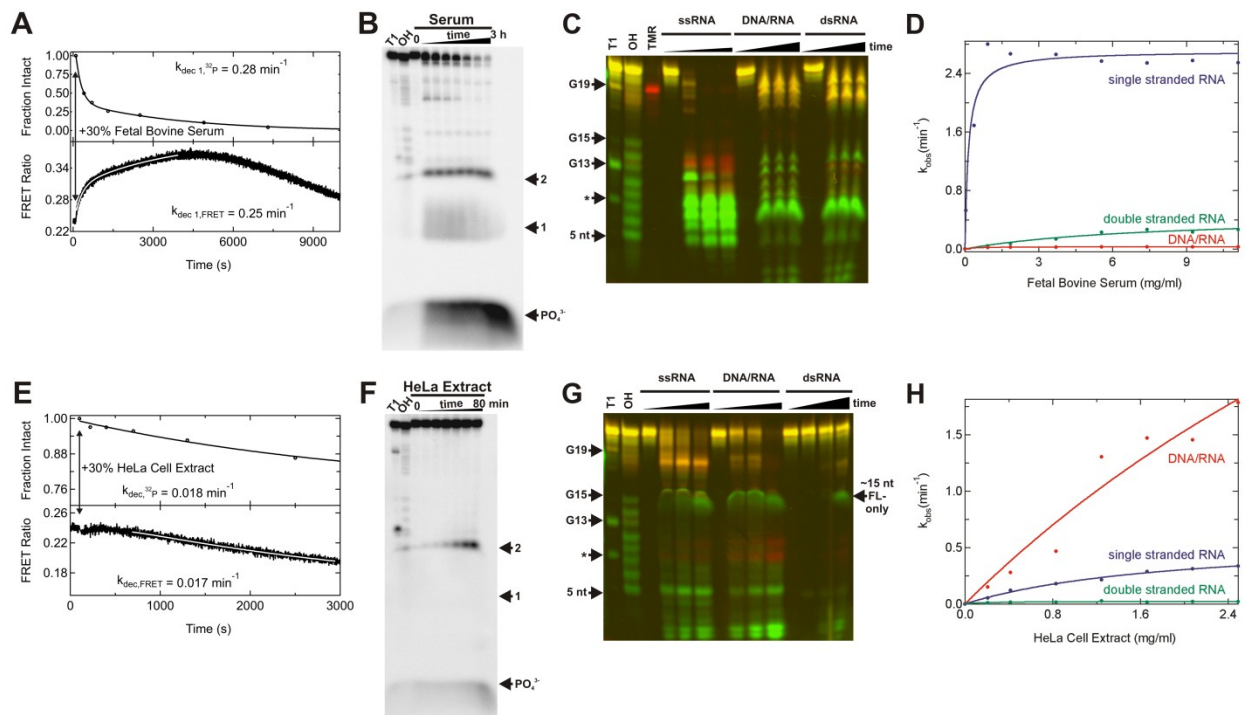
Figure 3.2. Monitoring the degradation of ssRNAs and dsRNAs in fetal bovine serum and HeLa S100 cytosolic cell extract. (A) Degradation time course of the 21-nt dsRNA in 30% (v/v) serum, as monitored either in real-time by steady-state FRET (bottom) or discontinuously by radioactive labeling and denaturing gel electrophoretic analysis of time points (top). Solid lines represent double-exponential fits to both types of degradation data. In both cases the faster of the two rate constants are reported here and elsewhere and demonstrate good agreement between the FRET and radioactive assays. **(B)** Denaturing radioactive gel electrophoretic analysis corresponding to the data in panel A (top). RNase T1 ("T1") and alkali ("OH") ladders were run alongside varying time points of RNA degradation to identify the products. Arrows denote bands corresponding to positions 1, 2 (from the 5' end), and PO_4^{3-}. **(C)** Denaturing FRET gel revealing the serum degradation patterns of the indicated RNAs as a function of time, alongside RNase T1 ("T1", leading to cleavage 3' of G as indicated by arrows) and alkali ("OH") ladders as well as a TMR-only labeled color calibration standard. The * denotes a double cut at both G13 and G3. **(D)** Observed degradation rate constants for the various RNAs as a function of total serum protein concentration, quantified from FRET time traces as exemplified in panel A (bottom). **(E)** Degradation time course of the 21-nt dsRNA in 30% (v/v) cell extract, as monitored either in real-time by steady-state FRET (bottom) or discontinuously by radioactive labeling and denaturing gel electrophoretic analysis of time points (top). Solid lines represent single-exponential fits to both types of degradation data with the respective rate constants shown, demonstrating good agreement between the FRET and radioactive assays. **(F)** Denaturing radioactive gel electrophoretic analysis corresponding to the data in panel E (top). RNase T1 ("T1") and alkali ("OH") ladders were run alongside

varying time points of RNA degradation to identify the products. Arrows denote bands corresponding to positions 1, 2 (from the 5' end), and PO_4^{3-}. **(G)** Denaturing FRET gel revealing the cell extract degradation patterns of the indicated RNAs as a function of time, alongside RNase T1 ("T1", leading to cleavage 3' of G as indicated by arrows) and alkali ("OH") ladders. The * denotes a double cut at both G13 and G3. **(H)** Observed degradation rate constants for the various RNAs as a function of total cell extract protein concentration, quantified from FRET time traces as exemplified in panel E (bottom). Further details of the experiments are found in Materials and Methods. Reprinted with permission from PLOS.

FRET ratio, followed by a slow decrease (**Fig. 3.2A**). Assays using purified RNase A (the major type of ribonuclease present in serum) qualitatively reproduced this shape of the FRET ratio time trace (**Appendix B, Fig. B.1**). In cell extract, the fast phase(s) decreased significantly (**Fig. 3.2E**). To assess which phase and corresponding rate constant from the FRET traces reflect RNA degradation, we used 5'-^{32}P radiolabeling of the 21-nt fluorophore-modified guide strand and denaturing gel electrophoretic analysis[12] to determine the degradation rate constant independent of the real-time output of our FRET trace (**Fig. 3.2B,F**). Gels were quantified for the fraction of the guide strand that was intact, therefore quantifying only primary cleavage events. The resulting data were plotted and fit with exponential functions (double-exponential increase for serum and single-exponential decrease for HeLa cell extract). Comparison with these data unambiguously identifies the relevant degradation rate constants from the FRET time traces as the dominant rise phases in serum and the dominant decrease phase in HeLa cell extract (**Fig. 3.2A,E**). We note that the curvatures and rates of the radioactively monitored decay time courses are well reflected in the real-time FRET time traces (despite the fact that the FRET trace in serum is inverted), attesting to the validity of using FRET to monitor the integrity of the guide strand over time. The increase in FRET particularly upon degradation in serum may be due to, for example, transient formation of doubly-labeled, (partially) single-stranded RNAs with shorter fluorophore distances.

We next qualitatively showed by denaturing FRET gel analysis that serum manifests more efficient cleavage of the ssRNA compared to either the DNA/RNA or the dsRNA duplex (**Fig. 3.2C**). While the cleavage patterns of the DNA/RNA hybrid and dsRNA are quite similar, the cleavage pattern of the ssRNA is notably different. These differences include accumulation of a slight TMR-only band in the ssRNA sample (note the similar color to the reference TMR-only control lane), which indicates selective removal of the FL donor fluorophore. There is also significant accumulation of intermediate size (5-13 nt) products that are either singly TMR or FL labeled. By contrast, products smaller than 5 nt are almost absent from the ssRNA lanes, suggesting primarily

endonucleolytic, rather than 5' exonucleolytic, decay (and/or very rapid degradation of smaller fragments). Some accumulation of short FL-only labeled products is seen upon degradation of the DNA/RNA hybrid and dsRNA, suggesting limited 5' exonuclease activity.

The degradation patterns in HeLa cell extract are markedly different (**Fig. 3.2G**). Most notably, the DNA/RNA hybrid is more completely degraded in HeLa cell extract, whereas the ssRNA is more completely degraded in serum (compare with **Fig. 3.2C**). Another difference between serum and cell extract is the significant accumulation of short (<5 nt) products in the HeLa lanes, suggesting that these samples contain substantial detectable 5' and 3' exonucleolytic activities. The gel also shows that degradation of ssRNA by HeLa cell extract generates long (15-20 nt) exonucleolytic products that are not found for either the DNA/RNA hybrid or dsRNA (**Fig. 3.2G**). There is also a FL-only labeled, somewhat diffuse band found in all samples incubated with cell extract that co-migrates with an ~15-nt RNA fragment (**Fig. 3.2G**). Given that the two fluorophores are attached only 8 nt apart (**Fig. 3.1**) this band may be evidence of a tight interaction of a component of the cell extract with a shorter, FL-only labeled fragment, leading to a band shift in the denaturing gel.

Applying our real-time FRET degradation assay to the three fluorophore labeled RNA variants allowed us to determine their decay rate constants over a range of total protein concentrations from added serum or cell extract (**Figs. 3.2 D,H**). Consistent with our qualitative observations by gel electrophoresis, the fastest observed rate constants are found for the decay of ssRNA in serum (maximum velocity at saturation, v_{max}= 2.7 min^{-1}; half-titration point FBS$_{1/2}$= 0.14 mg/ml; **Fig. 3.2D**). Degradation rate constants are significantly slower upon formation of secondary structure, as the dsRNA (v_{max}= 0.47 min^{-1}; FBS$_{1/2}$= 7.7 mg/ml) and particularly the DNA/RNA hybrid (v_{max}= 0.03 min^{-1}; FBS$_{1/2}$= 0.46 mg/ml) are less efficiently degraded.

The relative decay kinetics in HeLa cell extract (**Fig. 3.2H**) are significantly different from those in serum (compare with **Fig. 3.2D**). Here the DNA/RNA hybrid is most efficiently degraded (v_{max}= 7.4 min^{-1}; HeLa$_{1/2}$= 7.6 mg/ml; **Fig.**

3.2H), while the ssRNA is relatively protected (v_{max}= 0.60 min^{-1}; HeLa$_{1/2}$= 1.9 mg/ml). We find that the secondary structured dsRNA is most strongly protected in cell extract, leading to only a slow decay (v_{max}= 0.02 min^{-1}; HeLa$_{1/2}$= 0.18 mg/ml). In both cell extract and blood serum, the formation of dsRNA thus protects ssRNA against nucleolytic degradation.

DNA/RNA hybrid is degraded by RNase H1 in HeLa cell extract. One of the most prominent features of our decay data is the fast degradation of the DNA/RNA hybrid in HeLa S100 cytosolic cell extract, contrasting with its strong protection in blood serum. We hypothesized that the efficient degradation of the DNA/RNA hybrid in cell extract results from the activity of human RNase H enzymes. RNase H activity degrades the RNA in a DNA/RNA hybrid and is important in nucleic acid processing in the cell, including in DNA replication. Two DNA aptamers have previously been in vitro selected for catalytic inhibition of human RNase H1[30], allowing us to specifically test for RNase H1 degradation of the fluorophore-labeled DNA/RNA hybrid. **Fig. 3.3A** shows the time evolution of the FRET ratio in several DNA/RNA hybrid degradation assays. In the absence of any added DNA or the presence of a negative control DNA, degradation is rapid. When the inhibitory RNase H1 aptamers VI-2 or V-2 were added, we observed a marked reduction in the observed decay rate constant, indicating specific inhibition of RNase H1 in the degradation of the fluorophore-labeled DNA/RNA hybrid. These experiments were repeated over a range of aptamer concentrations, and a dose dependent decrease in rate constant was observed for both aptamers (**Fig. 3.3B**; VI-2: initial rate constant v_0= 1.35 min^{-1}, minimal rate constant at saturation v_{min}= 0.75 min^{-1}, half-titration point VI-2$_{1/2}$= 40 nM; V-2: v_0= 1.34 min^{-1}, v_{min}= 0.4 min^{-1}, V-2$_{1/2}$= 90 nM). These observations indicate that human RNAse H1 is (one of) the enzyme(s) responsible for degrading a DNA/RNA hybrid in HeLa cell extract.

FRET labeled 24-nucleotide dsRNA is a Dicer substrate. Interaction with Dicer is one of the earliest and most crucial steps in RNAi, wherein larger dsRNAs are cleaved into 21-nt fragments that Dicer then loads into the RISC complex[31, 32, 33, 34]. To test whether our 24-nt dsRNA is an authentic substrate for
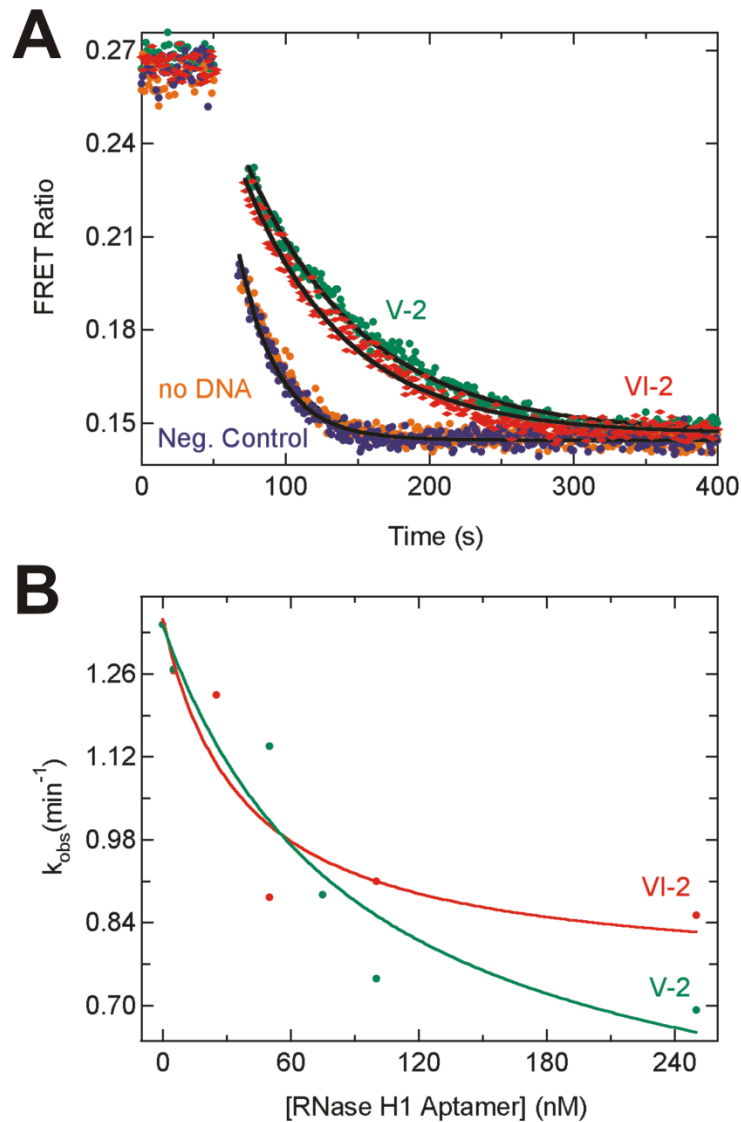
Figure 3.3. Inhibition of degradation of the doubly-labeled DNA/RNA hybrid in 30% (v/v) HeLa S100 cytosolic cell extract upon addition of DNA aptamers. (A) Steady-state FRET assays showing attenuation of degradation by RNase H1-specific DNA aptamers VI-2 (red) and V-2 (green) at 500 nM concentration. No such attenuation is observed in the absence of inhibitor (orange) or when 500 nM of the negative control S1-DNA is included (blue). Solid black lines represent single-exponential fits to the data to extract rate constants. **(B)** Rate constants of DNA/RNA degradation as a function of VI-2 or V-2 aptamer concentration. Solid lines are fits with a hyperbolic binding equation, revealing the nanomolar binding constants for the DNA aptamers reported in the text. Reprinted with permission from PLOS.

Dicer, we employed both FRET assays and gel based analysis. First, we collected steady-state FRET time traces for our 18-, 21-, and 24-nt dsRNAs upon addition of purified human Dicer (**Fig. 3.4A**). While the traces for the 18- and 21-nt dsRNAs hold steady over more than 30 min incubation time, the trace of the 24-nt dsRNA significantly increases, suggestive of an interaction with Dicer (**Fig. 3.4A**). Next, we incubated the three dsRNAs with either human Dicer or HeLa cell extract for 200 min and analyzed the degradation products on a denaturing FRET gel (**Fig. 3.4B**). As expected, the 24-nt dsRNA is processed by Dicer to the mature 21-nt siRNA. In HeLa cell extract, the 18-nt dsRNA exhibits a slightly stronger degradation band at ~7 nt length than either the 21- or 24-nt dsRNAs, but with that exception the lanes are quite similar. We only observed a weak 21-nt Dicer-derived cleavage band in the 24-nt dsRNA lane, suggesting that processing by Dicer under these conditions in cell extract is slow, perhaps because the incorporation into Dicer containing protein complexes is modest for our fluorophore-labeled dsRNAs (see also **Fig. 3.5E** and corresponding discussion below). Overall, degradation is slow over the 200 min incubation and seems common to all dsRNAs in HeLa cell extract, generating the diffuse, FL-only band that co-migrates with an ~15 nt RNA fragment and is also observed in **Fig. 3.2G**.

21- and 24-nucleotide dsRNAs are protected in cell extract. Next, we employed our steady-state FRET assay to the 18-, 21-, and 24-nt dsRNAs and extracted the rate constants from the time traces that were identified in **Fig. 3.2** as reflective of the RNA degradation kinetics (the rising phase in fetal bovine serum and the decreasing phase in HeLa extract, **Fig. 3.5**). The results from serum (**Fig. 3.5B**) show that the degradation rate constants progressively decrease with length of the dsRNA (18-nt dsRNA: v_{max}= 1.2 min^{-1}, Serum$_{1/2}$= 4.5 mg/ml; 21-nt dsRNA: v_{max}= 0.47 min^{-1}, Serum$_{1/2}$= 7.7 mg/ml; 24-nt dsRNA: v_{max}= 0.09 min^{-1}, Serum$_{1/2}$= 1.7 mg/ml). This trend correlates with the increasing distance of the fluorophores from the ends as well as the higher thermodynamic stability of the longer duplexes (**Fig. 3.1**). By comparison, the decay rate constants observed in cell extract appear to be an order of magnitude slower in
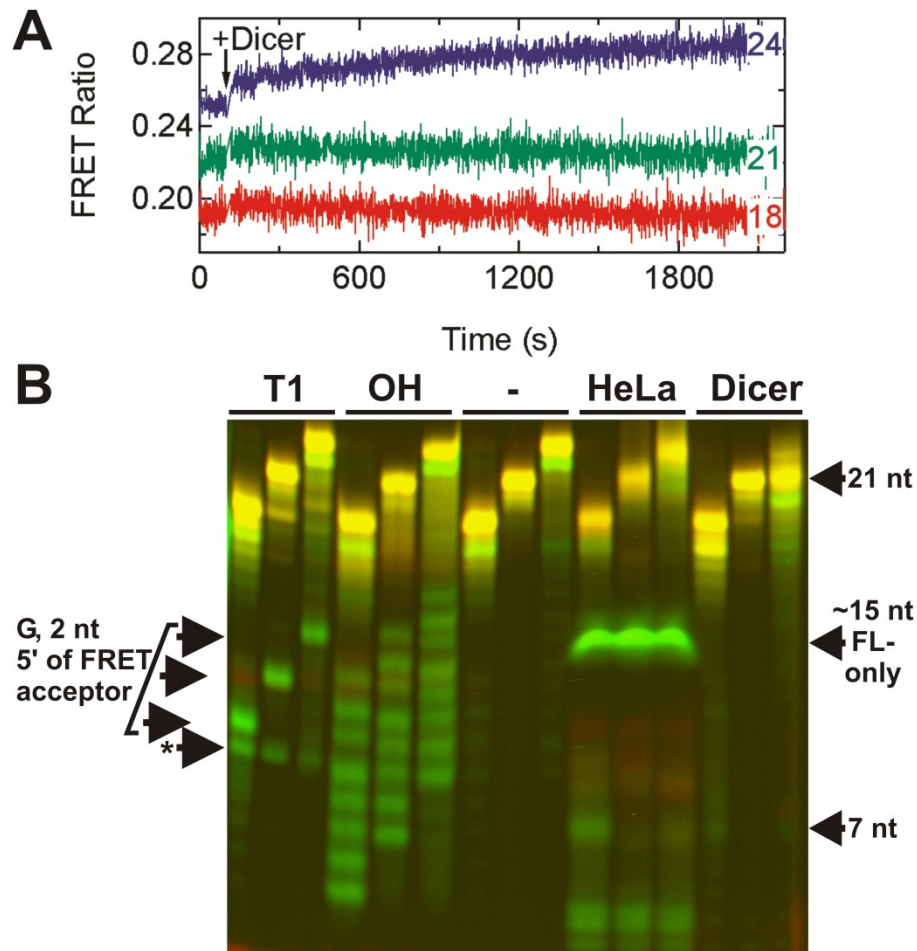
Figure 3.4. Processing of doubly-labeled dsRNAs of varying length by HeLa S100 cytosolic cell extract and purified human Dicer. (A) Steady-state FRET time traces of 18-, 21-, and 24-nt dsRNAs as indicated upon Dicer addition; only the 24-nt dsRNA time trace slightly increases. (***B***) Denaturing FRET gel revealing the degradation patterns of the 18-, 21-, and 24-nt dsRNAs (from left to right in each triplet of bands) as a function of time, alongside RNase T1 ("T1", leading to cleavage 3' of G as indicated by arrows on the left) and alkali ("OH") ladders. "-" indicates incubation in buffer only, while the last two triplets of bands show samples incubated in HeLa cell extract and purified Dicer enzyme, respectively. Arrows to the right mark specific degradation products. Reprinted with permission from PLOS.
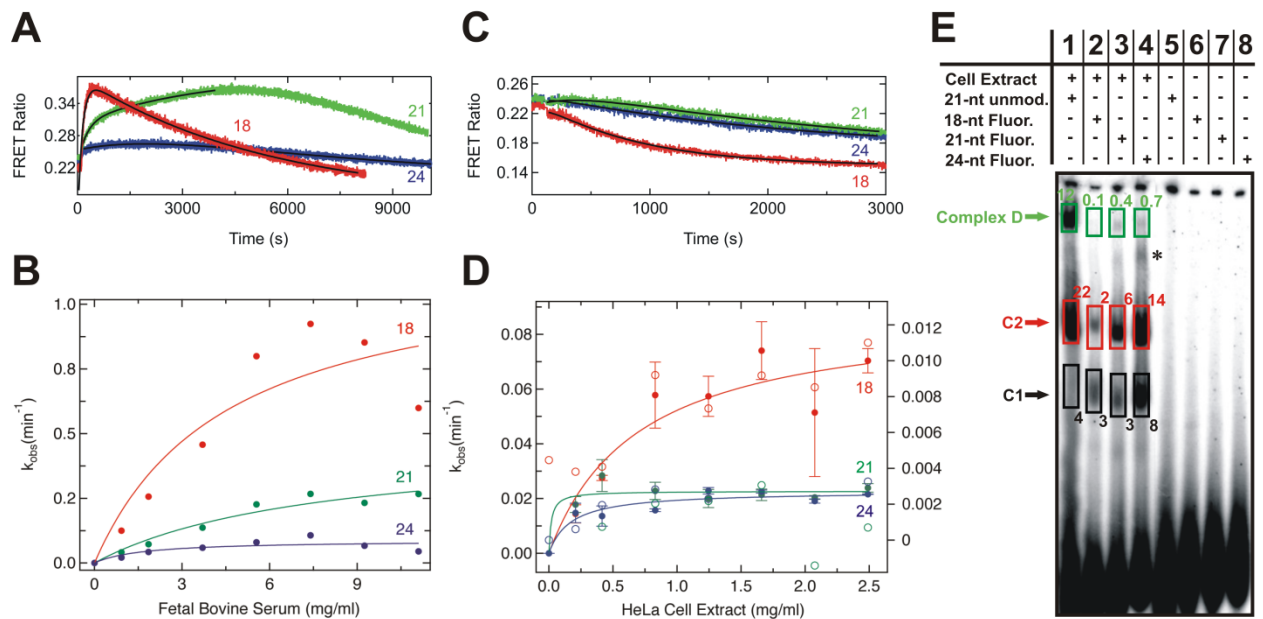
Figure 3.5. Degradation kinetics of 18-, 21-, and 24-nt dsRNAs in fetal bovine serum and HeLa S100 cytosolic cell extract. (A) Steady-state FRET time traces upon degradation of 18-, 21-, and 24-nt dsRNAs (red, green, and blue respectively) in serum. The solid black lines represent double-exponential fits to the data, the fast rate constants of which are reported in panel B. **(B)** The resulting rate constants of degradation of the 18-, 21-, and 24-nt dsRNAs in serum plotted as a function of total serum protein concentration. Solid lines indicate fits with a hyperbolic binding equation. **(C)** Steady-state FRET time traces upon degradation of 18-, 21-, and 24-nt dsRNAs (red, green, and blue respectively) in cell extract. The solid black lines represent single-exponential fits to the data, the rate constant of which are reported in panel D. **(D)** The rate constants of degradation, derived from steady-state FRET (solid circles, left y-axis) and radioactive assays (open circles, right y-axis) of the 18-, 21-, and 24-nt dsRNAs (red, green, and blue respectively) in cell extract and plotted as a function of total cell extract protein concentration. Solid lines indicate fits of the FRET-derived data with a hyperbolic binding equation. **(E)** Formation of complexes D, C2 and C1 after 2 h of incubation with radiolabeled 21-nt dsRNA with no modification (lane 1), 18-nt dsRNA (lane 2), 21-nt dsRNA (lane 3) and 24-nt dsRNA (lane 4) in HeLa cytosolic extract. Lanes 5-8 represent the respective negative controls in the absence of cell extract with the dsRNAs in the same order from left to right. Normalized fractions in percent are indicated for each boxed band; * indicates an unknown complex. Reprinted with permission from PLOS.

the FRET-based assays (filled circles in **Fig. 3.5D**, left y-axis) and even more so using 5' ^{32}P-labeled, fluorophore-labeled dsRNAs (open circles in **Fig. 3.5D**, right y-axis), measured in a less accurate, discontinuous radioactive assay where aliquots were taken periodically and decay products analyzed by denaturing gel electrophoresis as in **Fig. 3.2**. Steady-state FRET assays revealed that the 18-nt dsRNA is the most rapidly degraded of the three RNAs (v_{max}= 0.09 min^{-1}, HeLa$_{1/2}$= 0.70 mg/ml). In contrast to the serum result, the 21- and 24-nt dsRNAs exhibit the same low v_{max} values, signifying that they are degraded at about the same rate in cell extract (21-nt dsRNA: v_{max}= 0.02 min^{-1}, HeLa$_{1/2}$= 0.02 mg/ml; 24-nt dsRNA: v_{max}= 0.02 min^{-1}, HeLa$_{1/2}$= 0.17 mg/ml). The same trend was observed using the 5' ^{32}P-radiolabeled dsRNAs (18-nt dsRNA: v_{max}= 0.011 min^{-1}, HeLa$_{1/2}$= 0.68 mg/ml; 21-nt dsRNA: v_{max}= 0.0031 min^{-1}, HeLa$_{1/2}$= 0.3 mg/ml; 24-nt dsRNA: v_{max}= 0.0032 min^{-1}, HeLa$_{1/2}$= 0.29 mg/ml). These results suggest that the size dependence of the protection of dsRNA from decay in cell extract is distinct from that in serum. It should be noted that we also observed that the deoxythymidine and fluorophore modifications offered a general, modest protection to the RNAs, as may be expected (data not shown).

Longer fluorophore labeled RNAs are more efficiently assembled into RISC-related complexes. To investigate the ability of the fluorophore labeled 18-, 21-, and 24-nt dsRNAs to form protein complexes upon incubation with HeLa cell extract we performed gel electrophoretic mobility shift assays under non-denaturing conditions to identify such complexes based on a modified RISC assembly assay (**Fig. 3.5E**)[35]. Characteristic RISC-related RNA:protein complexes D (Retention factor, Rf: 0.2), C2 (0.4) and C3 (0.5) were observed, similar to those found previously using the unmodified 21-nt dsRNA[35]. Complex D in particular is known to contain Dicer[35, 36]. We observed that the fraction of the fluorophore-labeled ^{32}P-labeled RNAs incorporated into Complex D is less than that of ^{32}P-labeled 21-nt dsRNA devoid of any modification, consistent with the slow processing of 24-nt dsRNA by Dicer (**Fig. 3.4B**), whereas complexes C1 and C2 are more highly populated (**Fig. 3.5E**). Importantly, the fraction of 18-mer dsRNA incorporated into any of these complexes is significantly less than those

of either the 21-mer or 24-mer dsRNA, which may explain the comparatively lesser protection of the 18-mer relative to the 21-mer and 24-mer dsRNAs in HeLa cell extract.

3.4 Discussion

Potent and specific induction of mRNA degradation by siRNAs is a key determinant for harnessing RNAi in reverse genetic studies and therapeutic applications. It remains unclear, however, just how stable siRNAs are inside the cell. We have previously shown the decreased nuclease sensitivity of dsRNA compared to ssRNA in human HeLa cytosolic cell extracts[26], and recent work has employed FRET fluctuation spectroscopy to demonstrate this trend inside cells[23]. We here have established real-time assays to measure degradation profiles and kinetics for small, RNAi-related RNAs in both blood serum and RNAi-active cytosolic cell extract. We show that FRET-labeled RNAs interact with relevant cellular enzymes including RNase H1 and Dicer. A DNA/RNA hybrid is strongly protected in serum, but rapidly degraded by RNase H1 in cell extract. 21- and 24-nt double-stranded RNAs are equivalently protected in HeLa cell extract, but degraded considerably faster in blood serum with their relative decay rate constants correlating with their relative lengths. Given that the 24-nt dsRNA is a substrate of human Dicer and the 21-nt dsRNA its product, these results suggest that RNAs that are involved in RNAi are specifically protected in the cellular milieu. Our observations may thus explain the relative longevity of intracellular RNAi effects induced even by chemically unmodified siRNAs[8].

As expected, dsRNAs are more stable against nucleases than the corresponding ssRNA in both HeLa cell extract and blood serum (**Fig. 3.2**). These results are consistent with the relative stiffness of the RNA duplex, stabilizing the phosphodiester and ribose torsion angles in conformations less amenable to the in-line attack conformation needed by ribonucleases such as the RNase A-like enzymes dominant in serum[29, 37]. Hybridization to a DNA further stabilizes the RNA component in blood serum, perhaps due to perturbation of the A-form helical structure or lack of 2'-OH functionality on an entire strand of the

duplex. By contrast, in cell extract the DNA/RNA hybrid is rapidly degraded by RNase H1, as shown by enzyme-specific inhibition by two *in vitro* selected DNA aptamers[30] (**Fig. 3.3**). In human cells, a second, disparate, multi-subunit enzyme, termed RNase H2, has been characterized that degrades DNA/RNA hybrids[38]. This may be the enzyme that causes the residual DNA/RNA decay upon maximal inhibition of RNase H1 (**Fig. 3.3**).

The size range of the dsRNAs we tested for their stability in serum and cell extract was chosen to bracket the specific 21-nt size of siRNAs in human cells. dsRNAs larger than ~21 nt are cleaved by human Dicer into siRNAs during the initial step of RNAi, as observed here for the 24-nt dsRNA, followed by incorporation into the RISC-loading complex (RLC), comprising the proteins Dicer, Ago2, and TRBP[33, 34]. Since cellular delivery of 21-nt siRNAs induces RNAi, albeit not as effectively as does delivery of Dicer substrates[27], both types of dsRNAs must be interacting with the components of the RNAi machinery, in contrast to shorter dsRNAs. The similar protection of 21- and 24-nt dsRNAs in our assays relative to the corresponding 18-nt dsRNA (**Fig. 3.5**) is consistent with these findings. In addition, the fraction of 18-mer dsRNA incorporated into RISC-related protein complexes is significantly less than those of either the 21-mer or 24-mer dsRNAs (**Fig. 3.5E**). This result suggests that the comparatively faster degradation rate of the 18-mer versus the 21-mer and 24-mer dsRNAs may be linked to the existence of protective RNA:protein complexes in HeLa cell extract preferably for the longer dsRNAs. By contrast, the efficiency of dsRNA degradation in blood serum monotonically decreases with increasing RNA size, that is, with distance of the FRET fluorophores from the ends and overall thermodynamic stability of the duplex (**Fig. 3.5**). These observations are in line with the known properties of RNase A, where the limiting step of enzyme action on dsRNA substrates is invasion of the duplex to access an in-line attack conformation and cleave the RNA backbone[37].

In summary, our results are consistent with established steps in the RNAi pathway where dsRNAs of ~21 nucleotides per strand are bound by the RLC or other RNAi components and remain largely hidden from nucleolytic decay either

as the base-paired siRNA duplex or as the single guide strand processed into active RISC. The real-time *in vitro* FRET assays established here open up the possibility for future high-throughput screening of the nucleolytic stability of additional chemically modified RNAi effector molecules.

3.5 Acknowledgments

3.6 References

1. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-11.
2. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. & Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-8.
3. Rana, T. M. (2007). Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* **8**, 23-36.
4. Siomi, H. & Siomi, M. C. (2009). On the road to reading the RNA-interference code. *Nature* **457**, 396-404.
5. Castanotto, D. & Rossi, J. J. (2009). The promises and pitfalls of RNA-interference-based therapeutics. *Nature* **457**, 426-33.
6. Lee, S. K. & Kumar, P. (2009). Conditional RNAi: towards a silent gene therapy. *Adv Drug Deliv Rev* **61**, 650-64.
7. Amarzguioui, M., Holen, T., Babaie, E. & Prydz, H. (2003). Tolerance for mutations and chemical modifications in a siRNA. *Nucleic Acids Res* **31**, 589-95.
8. Bartlett, D. W. & Davis, M. E. (2007). Effect of siRNA nuclease stability on the in vitro and in vivo kinetics of siRNA-mediated gene silencing. *Biotechnol Bioeng* **97**, 909-21.
9. Braasch, D. A., Jensen, S., Liu, Y., Kaur, K., Arar, K., White, M. A. & Corey, D. R. (2003). RNA interference in mammalian cells by chemically-modified RNA. *Biochemistry* **42**, 7967-75.
10. Chiu, Y. L. & Rana, T. M. (2003). siRNA function in RNAi: a chemical modification analysis. *RNA* **9**, 1034-48.
11. Czauderna, F., Fechtner, M., Dames, S., Aygun, H., Klippel, A., Pronk, G. J., Giese, K. & Kaufmann, J. (2003). Structural variations and stabilising modifications of synthetic siRNAs in mammalian cells. *Nucleic Acids Res* **31**, 2705-16.
12. Hoerter, J. A. & Walter, N. G. (2007). Chemical modification resolves the asymmetry of siRNA strand degradation in human blood serum. *RNA* **13**, 1887-1893.
13. Layzer, J. M., McCaffrey, A. P., Tanner, A. K., Huang, Z., Kay, M. A. & Sullenger, B. A. (2004). In vivo activity of nuclease-resistant siRNAs. *RNA* **10**, 766-71.
14. Soutschek, J., Akinc, A., Bramlage, B., Charisse, K., Constien, R., Donoghue, M., Elbashir, S., Geick, A., Hadwiger, P., Harborth, J., John, M., Kesavan, V., Lavine, G., Pandey, R. K., Racie, T., Rajeev, K. G., Rohl, I., Toudjarska, I., Wang, G., Wuschko, S., Bumcrot, D., Koteliansky, V., Limmer, S., Manoharan, M. & Vornlocher, H. P. (2004). Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **432**, 173-8.
15. Zimmermann, T. S., Lee, A. C., Akinc, A., Bramlage, B., Bumcrot, D., Fedoruk, M. N., Harborth, J., Heyes, J. A., Jeffs, L. B., John, M., Judge, A.

D., Lam, K., McClintock, K., Nechev, L. V., Palmer, L. R., Racie, T., Rohl, I., Seiffert, S., Shanmugam, S., Sood, V., Soutschek, J., Toudjarska, I., Wheat, A. J., Yaworski, E., Zedalis, W., Koteliansky, V., Manoharan, M., Vornlocher, H. P. & MacLachlan, I. (2006). RNAi-mediated gene silencing in non-human primates. *Nature* **441**, 111-4.

16. Potenza, N., Moggio, L., Milano, G., Salvatore, V., Di Blasio, B., Russo, A. & Messere, A. (2008). RNA Interference in Mammalia Cells by RNA-3'-PNA Chimeras. *Int J Mol Sci* **9**, 299-315.

17. Behlke, M. A. (2008). Chemical modification of siRNAs for in vivo use. *Oligonucleotides* **18**, 305-19.

18. Watts, J. K., Deleavey, G. F. & Damha, M. J. (2008). Chemically modified siRNA: tools and applications. *Drug Discov Today* **13**, 842-55.

19. de Fougerolles, A., Vornlocher, H. P., Maraganore, J. & Lieberman, J. (2007). Interfering with disease: a progress report on siRNA-based therapeutics. *Nat Rev Drug Discov* **6**, 443-53.

20. Ohrt, T. & Schwille, P. (2008). siRNA modifications and sub-cellular localization: a question of intracellular transport? *Curr Pharm Des* **14**, 3674-85.

21. Ohrt, T., Merkle, D., Birkenfeld, K., Echeverri, C. J. & Schwille, P. (2006). In situ fluorescence analysis demonstrates active siRNA exclusion from the nucleus by Exportin 5. *Nucleic Acids Res* **34**, 1369-80.

22. Berezhna, S. Y., Supekova, L., Supek, F., Schultz, P. G. & Deniz, A. A. (2006). siRNA in human cells selectively localizes to target RNA sites. *Proc Natl Acad Sci USA* **103**, 7682-7.

23. Raemdonck, K., Remaut, K., Lucas, B., Sanders, N. N., Demeester, J. & De Smedt, S. C. (2006). In situ analysis of single-stranded and duplex siRNA integrity in living cells. *Biochemistry* **45**, 10614-23.

24. Jagannath, A. & Wood, M. J. (2009). Localization of double-stranded small interfering RNA to cytoplasmic processing bodies is Ago2 dependent and results in up-regulation of GW182 and Argonaute-2. *Mol Biol Cell* **20**, 521-9.

25. Jarve, A., Muller, J., Kim, I. H., Rohr, K., MacLean, C., Fricker, G., Massing, U., Eberle, F., Dalpke, A., Fischer, R., Trendelenburg, M. F. & Helm, M. (2007). Surveillance of siRNA integrity by FRET imaging. *Nucleic Acids Res* **35**, e124.

26. Uhler, S. A., Cai, D., Man, Y., Figge, C. & Walter, N. G. (2003). RNA degradation in cell extracts: real-time monitoring by fluorescence resonance energy transfer. *J Am Chem Soc* **125**, 14230-1.

27. Kim, D. H., Behlke, M. A., Rose, S. D., Chang, M. S., Choi, S. & Rossi, J. J. (2005). Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat Biotechnol* **23**, 222-6.

28. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. & Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563-74.

29. Schein, C. H. (1997). From housekeeper to microsurgeon: the diagnostic and therapeutic potential of ribonucleases. *Nat Biotechnol* **15**, 529-36.

30. Pileur, F., Andreola, M. L., Dausse, E., Michel, J., Moreau, S., Yamada, H., Gaidamakov, S. A., Crouch, R. J., Toulme, J. J. & Cazenave, C. (2003). Selective inhibitory DNA aptamers of the human RNase H1. *Nucleic Acids Res* **31**, 5776-88.

31. Pham, J. W., Pellino, J. L., Lee, Y. S., Carthew, R. W. & Sontheimer, E. J. (2004). A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in Drosophila. *Cell* **117**, 83-94.

32. Tomari, Y., Matranga, C., Haley, B., Martinez, N. & Zamore, P. D. (2004). A protein sensor for siRNA asymmetry. *Science* **306**, 1377-80.

33. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. & Doudna, J. A. (2008). In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci USA* **105**, 512-7.

34. Wang, H. W., Noland, C., Siridechadilok, B., Taylor, D. W., Ma, E., Felderer, K., Doudna, J. A. & Nogales, E. (2009). Structural insights into RNA processing by the human RISC-loading complex. *Nat Struct Mol Biol* **16**, 1148-53.

35. Rawlings, R. A., Krishnan, V. & Walter, N. G. (2011). Viral RNAi Suppressor Reversibly Binds siRNA to Outcompete Dicer and RISC via Multiple Turnover. *J Mol Biol* **in press**.

36. Pellino, J. L., Jaskiewicz, L., Filipowicz, W. & Sontheimer, E. J. (2005). ATP modulates siRNA interactions with an endogenous human Dicer complex. *Rna* **11**, 1719-24.

37. Raines, R. T. (1998). Ribonuclease A. *Chem Rev* **98**, 1045-1066.

38. Cerritelli, S. M. & Crouch, R. J. (2009). Ribonuclease H: the enzymes in eukaryotes. *FEBS J* **276**, 1494-505.

39. Walter, N. G. & Burke, J. M. (2000). Fluorescence assays to study structure, dynamics, and function of RNA and RNA-ligand complexes. *Methods Enzymol* **317**, 409-40.

40. Walter, N. G. (2002). Probing RNA structural dynamics and function by fluorescence resonance energy transfer (FRET). *Curr Protoc Nucleic Acid Chem* **11.10**, 11.10.1-11.10.23.

41. Hoerter, J. A., Lambert, M. N., Pereira, M. J. & Walter, N. G. (2004). Dynamics inherent in helix 27 from Escherichia coli 16S ribosomal RNA. *Biochemistry* **43**, 14624-36.

42. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11**, 1475-89.

43. Ramirez-Carrozzi, V. & Kerppola, T. (2001). Gel-based fluorescence resonance energy transfer (gelFRET) analysis of nucleoprotein complex architecture. *Methods* **25**, 31-43.

44. Pereira, M. J., Harris, D. A., Rueda, D. & Walter, N. G. (2002). Reaction pathway of the trans-acting hepatitis delta virus ribozyme: a conformational change accompanies catalysis. *Biochemistry* **41**, 730-40.

45. Harris, D. A., Tinsley, R. A. & Walter, N. G. (2004). Terbium-mediated footprinting probes a catalytic conformational switch in the antigenomic hepatitis delta virus ribozyme. *J Mol Biol* **341**, 389-403.

Chapter 4

Compositional discrepancies in cell extracts facilitate different modes of RISC loading and activation[5]

4.1 Introduction

RNA interference (RNAi) and RNA silencing are a pair of key intracellular gene regulatory mechanism guided by small exogenous and endogenous ncRNAs known as short interfering RNAs (siRNAs) and microRNAs (miRNAs)[1, 2, 3], respectively. Exogenous siRNAs or long double-stranded RNAs (dsRNAs) are interpreted as invading gene elements by cells and trigger the cellular defense mechanism called RNAi[4]. Within the cell, these long dsRNAs are processed into mature siRNA duplexes by complexes containing the endoribonuclease Dicer (Dcr). These siRNA duplexes are subsequently loaded into an effector RNA Induced Silencing Complex[5] (RISC) that contain RNase-H type Argonaute (Ago) proteins[6]; then, the two strands are unwound, and RISC is guided by the one retained RNA strand to cleave a complementary target RNAs in a sequence dependent manner[2, 3].

Although it has been demonstrated that a minimal human RISC, containing just Dcr, TRBP (TAR RNA binding protein) and Ago2, mediates

91

siRNA-guided RNA cleavage[7, 8], several other Ago associated proteins have been shown to enhance RISC activities[9]. These proteins include Heat-shock protein 90 (Hsp90), GW182, RNA Helicase A (RHA) and Moloney leukemia virus 10 (MOV10) protein. Hsp90 is a 90 kDa ubiquitous molecular chaperone that regulates the function or turnover of a specific subset of signaling proteins, cell cycle regulators, and apoptotic factors[10]. Hsp90 is required for Ago2 to bind the siRNA duplex; it has also been shown to be responsible for the transfer of siRNA duplexes from the RISC loading complex (RLC) to RISC and the unwinding of siRNA duplex to single-stranded siRNAs in *Drosophila*, possibly by mediating conformational rearrangement of the Ago2 protein[11]. In mammals, it was shown that Hsp90 binds the N-terminus of Ago2 and is needed for interaction between Ago2 and Dcr[12], suggesting a potential role for Hsp90 in converting RLC into RISC. Additionally, Hsp90 activity has been shown to be essential for the stability of non-RNA bound Ago2 proteins, Processing bodies (P-bodies) and GW182 (P-Body marker)[13], as well as for the efficient targeting of Ago2 to P-bodies and stress granules[14]. Besides its role in siRNA loading and the efficient utilization of exogenous siRNA, Hsp90 also plays a significant role in miRNA loading[13]. It has been shown to affect the function of Ago2, and therefore RISC activity in both siRNA and miRNA mediated sequence-specific cleavage and gene regulation, respectively[14]. GW182, a 182-kDa core P-body component that is imperative for P-body integrity, is another protein that interacts with Ago and is required for both miRNA mediated translational repression of mRNAs[15] and siRNA mediated mRNA decay[16]. GW182 interacts with the miRNA bound Ago2 through the N-terminal GW repeats[17] and the CCR4-NOT deadenylase complex through C-terminal-conserved W motifs[18], thereby linking the miRNA pathway to mRNA degradation. Recent studies have proposed a new role for GW182 in regulating stability of Ago-bound miRNAs where the interaction of the N-terminal region of GW182 with the miRNA-Ago complex inhibits the activity of an 3′–5′ exoribonuclease complex responsible for miRNA degradation[19].

Ago has been observed to interact with a few helicases such as MOV10[20] and RHA[21] whose role in the RNA silencing pathway is still under investigation.

MOV10 is a 130-kDa DExD-box helicase which, similar to Ago, was observed to localize to P-bodies[20]. MOV10 is an ortholog of Armitage, found in *Drosophila melanogaster*, which was suggested to be involved in the incorporation of single stranded siRNA into RISC[22]. The MOV10 protein is required for small RNA guided mRNA cleavage and has been shown to interact with mature-miRNA containing protein complexes functioning downstream of the Dicer cleavage step observed in *Drosophila*[20]. PAR-CLIP data, which capture transient interactions of RNAs with partner proteins, indicate that MOV10 mostly localizes to the 3' untranslated region of the mRNA and significantly overlaps with Ago2 binding sequences, suggesting that MOV10 and AGO2 bind to the same positions on the RNA[23]. It was also demonstrated that the interaction between MOV10 and Ago2 is RNA dependent,[23] however, the exact function of MOV10 in the RNA silencing pathway is still unclear. RHA is a 142 kDa DEAH-box protein that was found to interact specifically with active RISC[21]. RHA, similar to maleless, its ortholog in *Drosophila*[24], is hypothesized to play an important role in transcriptional regulation of genes. RHA has been shown to be required for the recruitment of siRNAs into Ago2 (RISC loading) containing complexes and formation of active RISC, where it may function in RISC loading by inducing the remodeling of the RLC[21]. Although it was shown to have the ability to unwind RNA duplexes[21], the specific protein(s) responsible for passenger strand removal remain a mystery.

Despite our expanding knowledge in this field, the full composition of RISC remains unknown. Although active RISCs of varying sizes ranging from ~160-500 kDa have all been shown to cleave mRNA targets[7, 25, 26], it is still unknown if these complexes represent the complete RISC or if they just contain the core proteins responsible for the RNAi activity. Much of this complexity we believe arises from the lack of a standard system of study for understanding the RNAi pathway. The research present here describes the systematic differences observed in the RNAi pathway when using two different preparations of HeLa cell extract, i.e., S100 cytoplasmic extract[27] and translationally active whole cell extract, both of which have been used interchangeably in the field as a model system to probe RNAi. Additionally, we found that MOV10 is required for the

RNA silencing machinery to utilize siRNA duplexes and mediate target mRNA cleavage.

4.2 Materials and Methods

Nucleic acid synthesis. siRNAs were synthesized by the HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory at the Yale University School of Medicine, then deprotected and purified as previously described.[28, 29] siRNA sequences were derived from the firefly luciferase gene[30] and were as follows: sense strand (21-mer) 5´-P-CGU ACG CGG AAU ACU UCG AAA-3´; antisense strand (21-mer) 5´-P-UCG AAG UAU UCC GCG UAC GUG-OH-3´; sense strand (32-mer) 5'-P- CAU CAC GUA CGC GGA AUA CUU CGA AAU GUC CG-OH-3'; antisense strand (32-mer) 5'-P- GAC AUU UCG AAG UAU UCC GCG UAC GUG AUG UU-OH-3', where P denotes a 5´-phosphate, and OH denotes a 3' hydroxyl. Both 5' phosphorylated and unphosphorylated versions of the 21-mer were purchased.

P^{32} labeling of 21-mer and 32-mer RNAs. The non-phosphorylated, antisense strand was 5'-^{32}P labeled with T4 polynucleotide kinase (PNK- NEB) and γ-^{32}P-ATP, mixed with loading buffer to give a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue, and 40% (v/v) formamide, and gel purified on a denaturing, 8 M urea, 20% (w/v) polyacrylamide gel. The radiolabeled full-length RNA was excised from the gel, diffusion eluted into 1 mM EDTA overnight, and ethanol precipitated. The dried RNA was dissolved in water. For radiolabled siRNA duplexes, the non-phosphorylated antisense strand was first 5´-^{32}P labeled with T4 polynucleotide kinase (PNK- NEB) and γ-^{32}P-ATP. PNK was inactivated by heating to 65 °C for 20 min, then the 5´ phosphorylated complementary strand was added in 1.5-fold molar excess and the reaction was heated to 90 °C for 3 min and slowly cooled to room temperature (or incubated 37 °C for 1 hr) to anneal the two strands. The siRNA duplex was mixed with loading buffer to give a final concentration of 10% glycerol, 1×TBE and 0.025%

(w/v) bromophenol blue and further purified by nondenaturing, 20% (w/v) polyacrylamide gel electrophoresis. The duplex siRNA was cut out, eluted overnight into 1 mM EDTA at 4 °C, ethanol precipitated, and dissolved in RNase-free water. Scintillation counting was done on a Beckman LS6500 Multipurpose Scintillation Counter.

**Target mRNA synthesis and labeling with P³². **The target luciferase mRNA was synthesized as described[31], but with the following modifications. PCR amplification of the pGEM-luc plasmid (Promega) using the Taq 2X Master Mix (NEB) and the following primers, Forward: GGCATAAAGAATTGAAGA and Reverse: GCGTAATACGACTCACTATAGGAACAATTGCTTTTACAG, resulted in a 181 bp PCR product containing an upstream T7 RNA polymerase promoter sequence for transcription. The PCR product was then *in vitro* transcribed using transcription reaction mixture containing 40 mM Tris-HCl (pH 7.5), 15 mM MgCl2, 5 mM dithiothreitol (DTT), 2 mM spermidine, NTPs (4 mM each), 5 units/mL inorganic pyrophosphatase, and 0.1 mg/mL T7 RNA polymerase [purified in the His-tagged form from an overexpressing strain as described previously[31]] and incubated at 37 °C for 4 hrs. The 161 nt long full length transcript was mixed with loading buffer to give a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue, and 40% (v/v) formamide, and extracted from a denaturing, 8 M urea, 20% (w/v) polyacrylamide gel electrophoresis by UV shadowing, diffusion elution of small gel slices into 1 mM EDTA, and ethanol precipitation. The dried RNA was dissolved in water and stored at -20 °C until further use. The purified target luciferase mRNA was 5'-cap P³² labeled using α-P³²-GTP and guanylyltransferase enzyme (ScriptCap Capping enzyme from Epicenter) as recommended by the supplier. Cap labeled RNA was mixed with loading buffer to give a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue, and 40% (v/v) formamide, and purified on a denaturing, 8 M urea, 15% (w/v) polyacrylamide gel. The RNA was excised, eluted into 1 mM EDTA at 4°C overnight, ethanol precipitated, and dissolved in water.

Cell Extracts. S100 cytosolic HeLa cell extract was obtained from Jena Biosciences, whereas translationally active HeLa whole cell lysate was obtained from ThermoScientific. Although the steps involved in the preparation of these extracts are different, in the final step of their preparation the extracts are dialyzed against buffers containing similar ionic strength, thus reducing the risk of ionic strength causing disparities in our comparative analysis. The S100 cytosolic HeLa cell extract contains 20 mM Tris-HCl, pH 7.9, 100 mM KCl, 0.2 mM EDTA, 1 mM DTT, 20% glycerol and the HeLa whole cell lysate contains 20 mM Hepes–KOH pH 7.5, 135 mM potassium acetate, 30 mM KCl, and 1.65 mM magnesium acetate. The ionic concentrations of these extract when used in the activity assays are comparable upon addition of the respective reaction buffers mentioned below.

Western Blot Analysis. For denaturing Western blots, the S100 cytosolic HeLa cell extract and the HeLa whole cell lysate were run on a NuPAGE Novex 4-12% gradient Bis-Tris Gel in the presence of a molecular weight marker (ECL Plex Rainbow marker). The gel was then electroblotted in Tris-Glycine buffer (25 mM Tris, 192 mM Glycine, and 20% (v/v) Methanol) onto a Polyvinylidene Fluoride (PVDF) membrane (Immobilon-P Membrane, Millipore) over 75 min at 300 mA, using a Bio-Rad Trans-Blot SD semi-dry transfer cell per the manufacturer's instructions. After transfer the membrane was probed with one of these antibodies: rabbit anti-Dicer (Santa Cruz Biotechnology), mouse anti-TRBP (Abnova), mouse anti-Ago2 (Abnova), mouse anti- GW182 (Santa Cruz Biotechnology), rabbit anti-MOV10 (Proteintech), rabbit anti-RHA (Vaxron), and mouse anti-Hsp90 (Enzo life sciences). Mouse anti-beta actin or rabbit anti-beta actin was used as a loading control for each of these westerns. The incubation with the primary antibody was followed by goat anti-Rabbit-Cy5 and goat anti-mouse-Cy3 (both Amersham), respectively. The probed blots were visualized on a Typhoon 9410 Variable Mode Imager (GE Healthcare Life Sciences) and quantified using Image Quant software (Molecular Dynamics).

Dicing assay. Radiolabeled 32 nt long double stranded substrates(40000cpm, ~300pM) was incubated with 40% (v/v) of either HeLa S100 extract or HeLa whole cell lysate, 25 mM creatine phosphate, 5 mM DTT, 10 mM HEPES, 150 mM potassium acetate, 0.3 mg/mL Pefabloc SC, 20 ug/mL creatine kinase, 0.8 U/uL RNasin and 10% (v/v) glycerol in a final volume of 10 uL. The reaction also contains 1mM ATP and 3 mM Magnesium Acetate unless otherwise stated. The entire reaction was incubated at 37 °C for 1 hr. The reaction was stopped by heating to 65 °C for 15 min and mixed with loading buffer to a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue, and 40% (v/v) formamide, and analyzed on a denaturing, 8 M urea, 20% (w/v) polyacrylamide gel. Gels were then wrapped in plastic wrap and exposed to phosphor screens, which were scanned on a Typhoon 9410 Variable Mode Imager (GE Healthcare Life Sciences) and quantified using Image Quant software (Molecular Dynamics).

Slicing assay and ATP depletion. For the slicing assays, 100 nM non radiolabeled siRNA duplex (21-mer) was preincubated with 40% (v/v) of either HeLa S100 extract or HeLa whole cell lysate, 25 mM creatine phosphate, 5 mM DTT, 10 mM HEPES, 150 mM potassium acetate, 0.3 mg/mL Pefabloc SC, 20 ug/mL creatine kinase, 0.8 U/uL RNasin and 10% (v/v) glycerol in a final volume of 10 uL at 4 °C for 15 min. The reaction also contains 1mM ATP and 3 mM Magnesium Acetate unless otherwise stated. The reaction was then warmed to 37 °C for 15 min, 40000cpm (~300pM) of radiolabeled target luciferase mRNA was added and the reaction was further incubated at 37 °C for 30 min. A time course of the slicing assay was performed to obtain the time at which target mRNA cleavage was maximal. The reaction was stopped by heating to 65 °C for 15 min and mixed with loading buffer to a final concentration of 1×TBE, 0.025% (w/v) bromophenol blue, and 40% (v/v) formamide, and analyzed on a denaturing, 8 M urea, 20% (w/v) polyacrylamide gel. Gels were then wrapped in plastic wrap and exposed to phosphor screens, which were scanned on a Typhoon 9410 Variable Mode Imager (GE Healthcare Life Sciences) and quantified using Image Quant software (Molecular Dynamics). For ATP depletion experiments, ATP, creatine phosphate, and creatine kinase were omitted and

ATP in extract was depleted by incubating the extract with 25 mM glucose and 0.5 units per µl hexokinase for 30 min at 15 °C. Lysis buffer instead of hexokinase was used for mock-treated samples. We substituted ATP with EDTA for the –ATP condition to normalize the free Mg^{2+} concentration and added equimolar concentration of Mg^{2+} for the excess ATP conditions.

4.3 Results

HeLa S100 cytosolic extract and HeLa whole cell lysate are viable for studying RNAi. We first confirmed that S100 cytosolic HeLa extract and HeLa whole cell lysate, as previously described[30, 37], contain Dcr, TRBP and Ago2 (**Fig. 4.1a**), the three protein components required for formation of minimal RISC. The process of cleaving (dicing) long RNA duplexes into mature siRNA/miRNA[32, 33, 34] and the cleavage (slicing) of target mRNA by RISC containing guide strand (the single-stranded guide also known as anti-sense strand)[35, 36] are key established and characteristic steps of the RNAi pathway. We therefore showed that both S100 cytosolic HeLa extract and HeLa whole cell lysate can perform both dicing and slicing (**Fig. 4.1b, c**), by employing a segment of luciferase mRNA that has been used in previous studies[30], thereby confirming that the core RISC components are functional. We also showed that the HeLa S100 cytosolic extract and HeLa whole cell lysate cleave a long dsRNA dicer substrate to the same extent to form a mature 21-nt long siRNA duplex, suggesting no difference in the dicing activity of both the extracts (**Fig. 4.2**).

Single-stranded siRNAs are incorporated into RISC more efficiently than siRNA duplexes in HeLa S100 cytosolic extract, but the opposite is true for HeLa whole cell lysate. S100 cytosolic HeLa extract has been previously used to demonstrate the ability of single-stranded antisense siRNAs to guide target mRNA cleavage[30]. When we use similar conditions in our slicing assays with HeLa S100 cytosolic extract, we observed that target mRNA cleavage achieved with single stranded antisense guide strand was approximately 5- to 10-fold more

efficient than that obtained using siRNA duplexes (**Fig. 4.3b**). This is contrary to a report that single-stranded siRNAs are less effective and require 10- to 100-fold higher concentrations than siRNA duplexes[27]. Our results are in concordance with previous work, however, which reported that rabbit reticulocyte lysate also lacks the ability to load siRNA/miRNA duplexes into RISC/miRNP unless the si/miRNA is preheated and preannealed to the mRNA before addition to extract[37]. In contrast, slicing assays performed with HeLa whole cell lysate showed that the extent of target mRNA cleavage obtained using the duplex siRNA was ~10-fold greater than that seen with single-stranded antisense siRNA (**Fig. 4.3a**), supporting the notion that siRNA duplexes are required for formation of mature and active RISC[32, 35, 36]. We also observed that supplementing the slicing assays performed with HeLa S100 cytosolic extract with increasing amounts of HeLa whole cell lysate enhanced the ability of the HeLa S100 cytosolic extract to utilize siRNA duplexes more efficiently to cleave complementary target mRNA (**Fig 4.4**). Thus, the inability of the HeLa S100 cytosolic extract to utilize siRNA duplex may be due to the lack of potential RISC loading factor(s).

MOV10 and RHA may have a potential role in loading siRNA duplexes into Ago. To address the question of why siRNA duplexes are utilized better in HeLa whole cell lysate compared to HeLa S100 cytosolic extract, we analyzed the protein of both extracts. In particular, we performed quantitative Western blot analyses of both cell extracts for known Ago2 interacting proteins that include Hsp90, GW182, RHA, MOV10, Dcr and TRBP with beta-actin as loading control (**Fig. 4.5**). Whereas the amount of Ago2, GW182 and TRBP present in both extracts are comparable, Dcr, RHA and MOV10 are significantly enriched in HeLa whole cell lysate compared to HeLa S100 cytosolic extract. More specifically, Dcr and RHA are present in 2- to 4-fold excess and MOV10 in >10-fold excess in HeLa whole cell lysate compared to S100 cell extract. It is noteworthy that MOV10 and RHA are both helicases that have been previously shown to be required for formation of mature RISC, and have been hypothesized
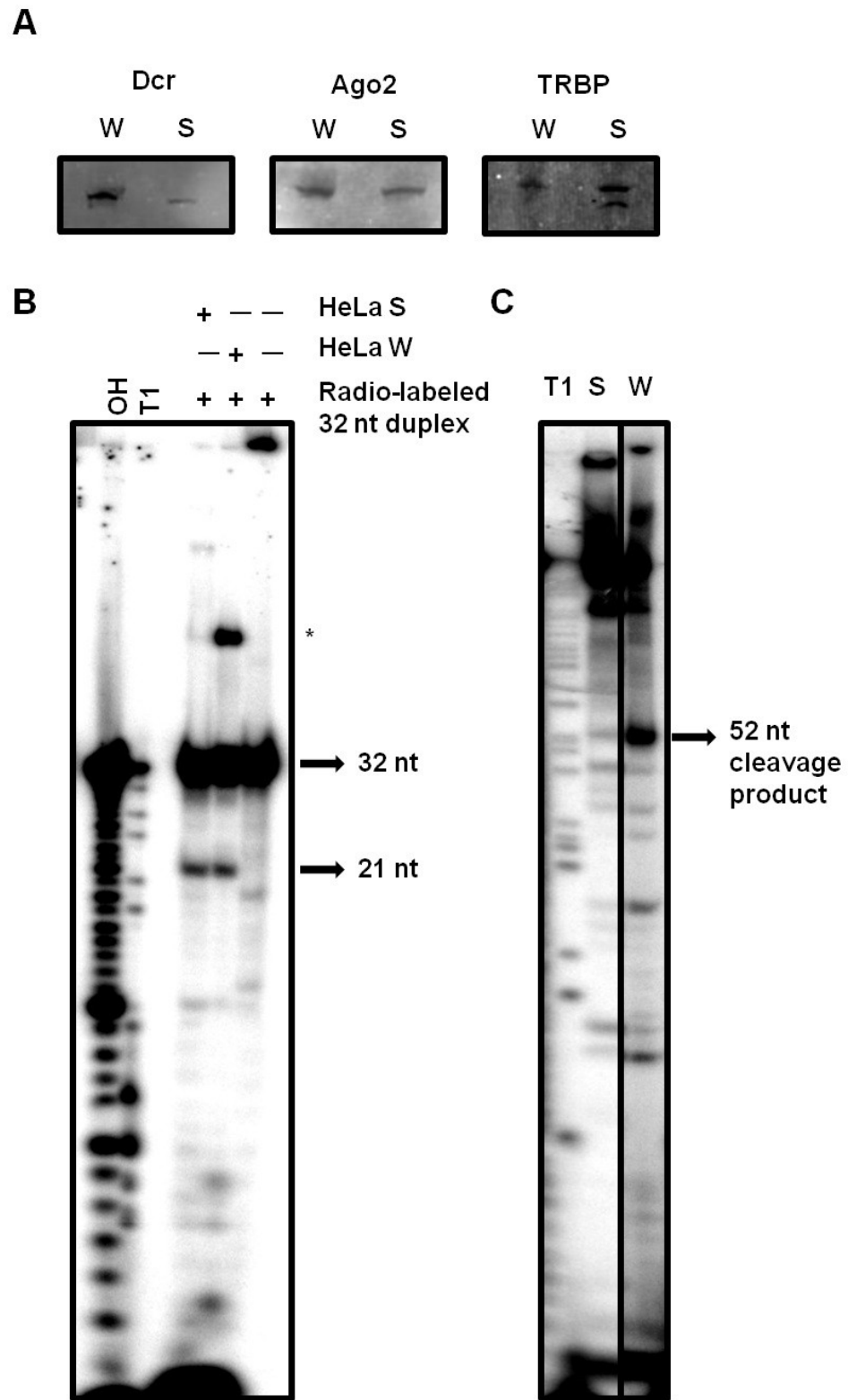
Figure 4.1. HeLa whole cell lysate and HeLa S100 cytoplasmic extract are sutiable systems for studying RNAi. (a) Denaturing Western blot analyses of

extracts shows the presence of Dcr, Ago and TRBP in both HeLa whole cell lysate(W) and HeLa S100 cytoplasmic extract(S). **(b)** Dicing assay showing the ability of both W and S to cleave a 32 nt long radiolabeled RNA duplex to a 21 nt long mature siRNA duplex. The hyrodxyl (OH) ladder and the RNaseT1 (T1) ladder were prepared as described[48]. Lanes 3 and 4 show slicing activity with S and W respectively. Lane 5 is a control with no extract added to the reaction. **(b)** Slicing assay showing the ability of both W and S to cleave a 161 nt long radiolabeled target mRNA to a 52 nt long cleaved product. The RNaseT1 (T1) ladder were prepared as described[48]. Lanes 2 and 3 show slicing activity with S and W respectively.
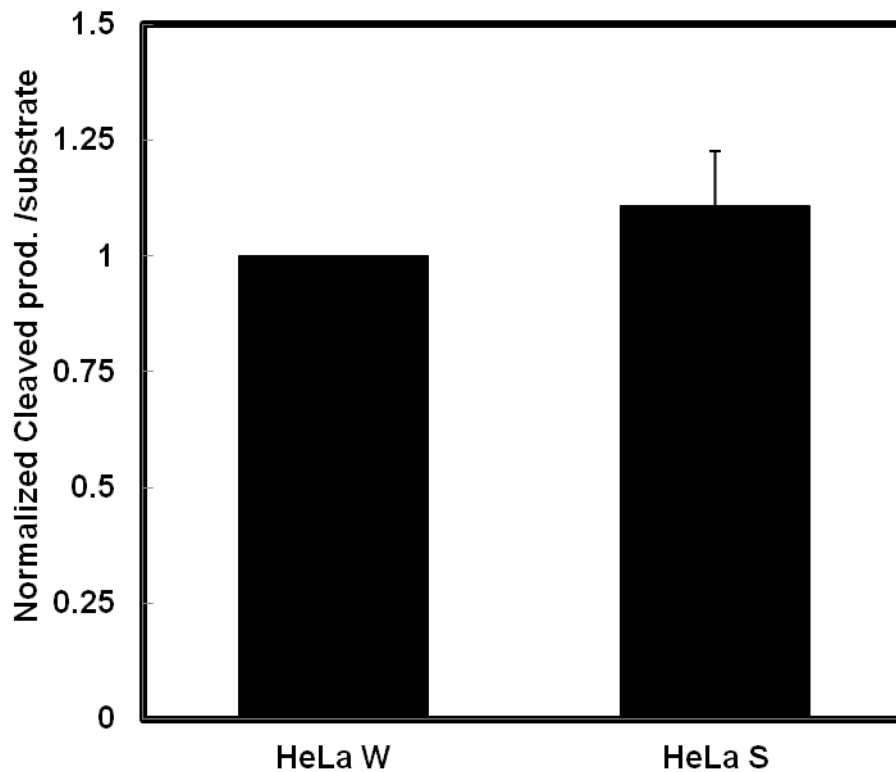
Figure 4.2. HeLa whole cell lysate and HeLa S100 cytoplasmic extract exhibit similar dicing activity. A graphical representation of the normalized and averaged fraction of cleaved 21-nt long RNA duplex over full length dsRNA substrate for dicing activity by HeLa whole cell lysate (HeLa W) and HeLa S100 cytoplasmic extract (HeLa S).
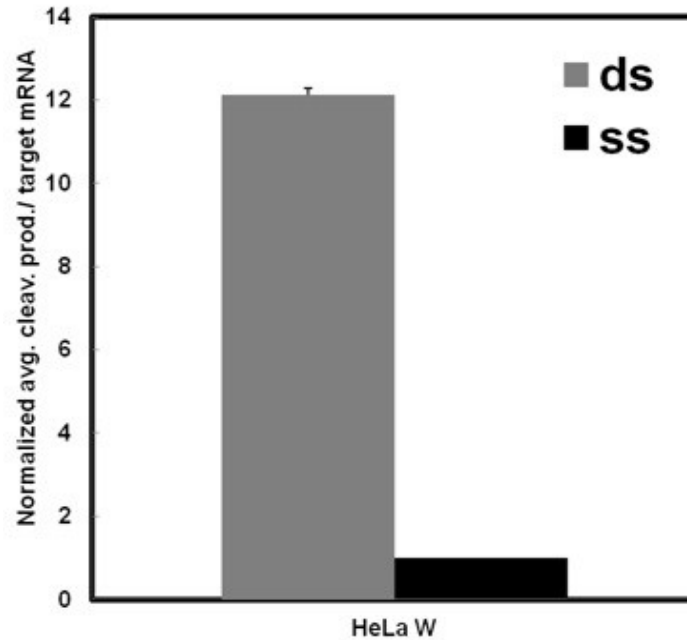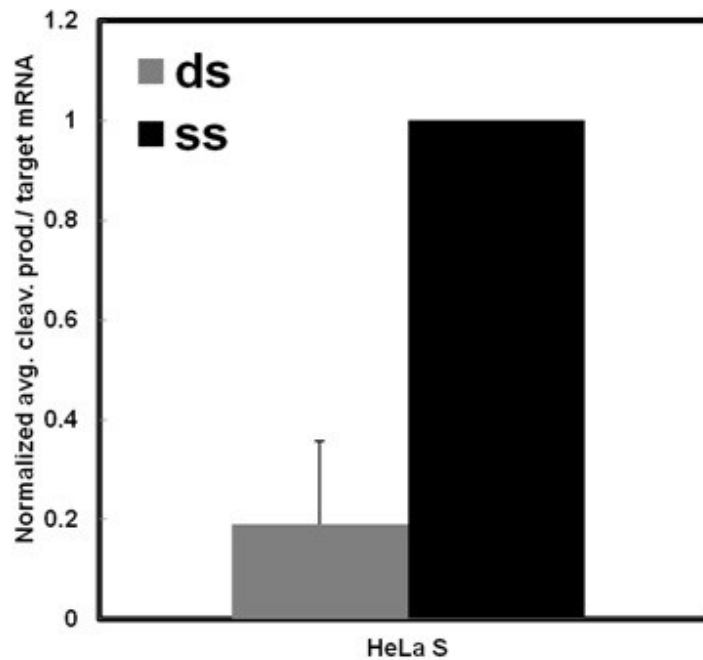
Figure 4.3. HeLa whole cell lysate and HeLa S100 cytoplasmic extract utilize single stranded antisense siRNA and siRNA duplexes to variable extents to mediate target mRNA cleavage. A graphical representation of the normalized and averaged fraction of cleaved target mRNA product over full length target mRNA for slicing activity by **(A)** HeLa whole cell lysate (HeLa W) and **(B)** HeLa S100 cytoplasmic extract (HeLa S). The gray and black bars indicate siRNA duplex (ds) and single stranded antisense siRNA (ss) mediated target mRNA cleavage respectively.
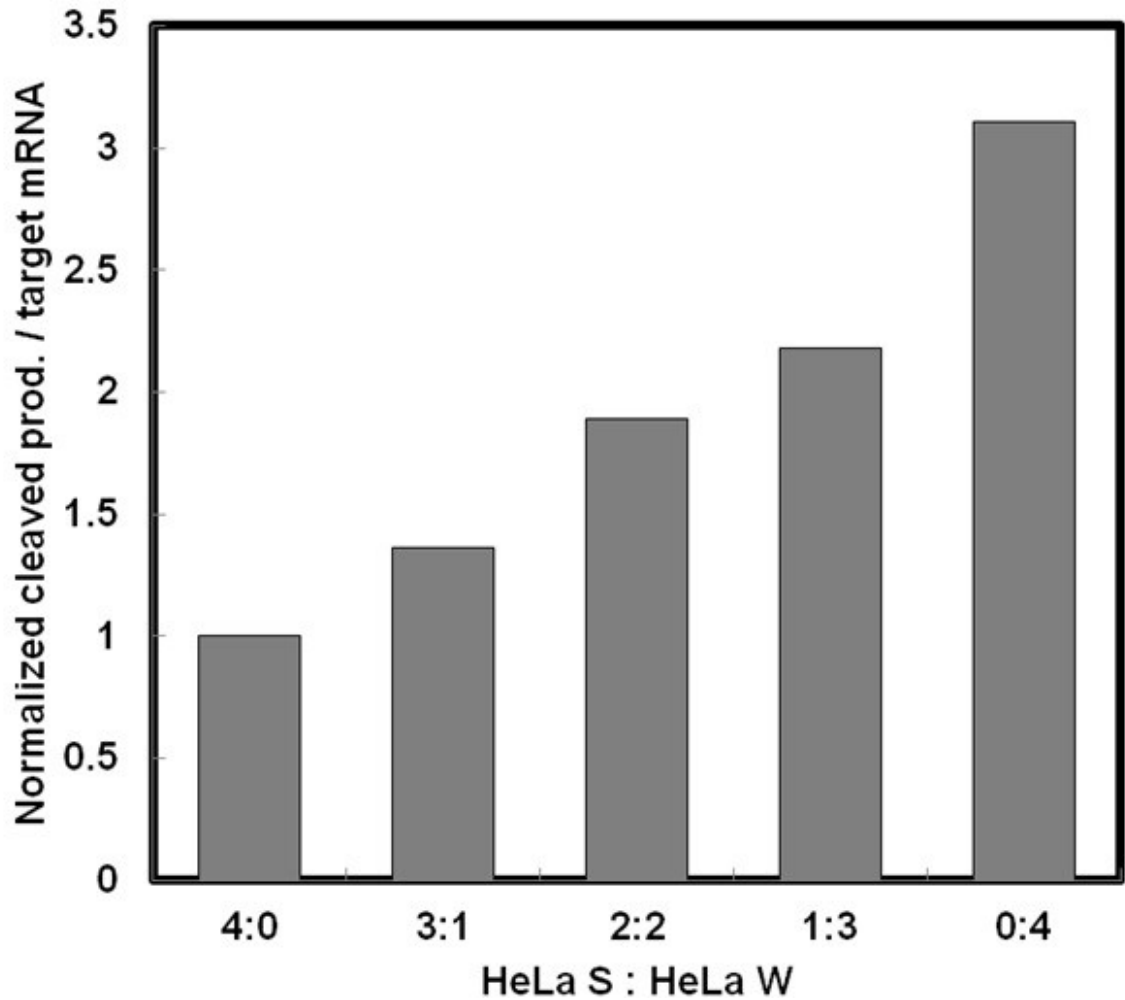
Figure 4.4. Addition of HeLa whole cell lysate enhances ds-siRNA mediated silencing activity in HeLa S100 cytoplasmic. A graphical representation of the normalized and averaged fraction of cleaved target mRNA product over full length target mRNA for slicing activity mediated by different ratios of HeLa whole cell lysate (HeLa W): HeLa S100 cytoplasmic extract (HeLa S) present in cleavage reactions.
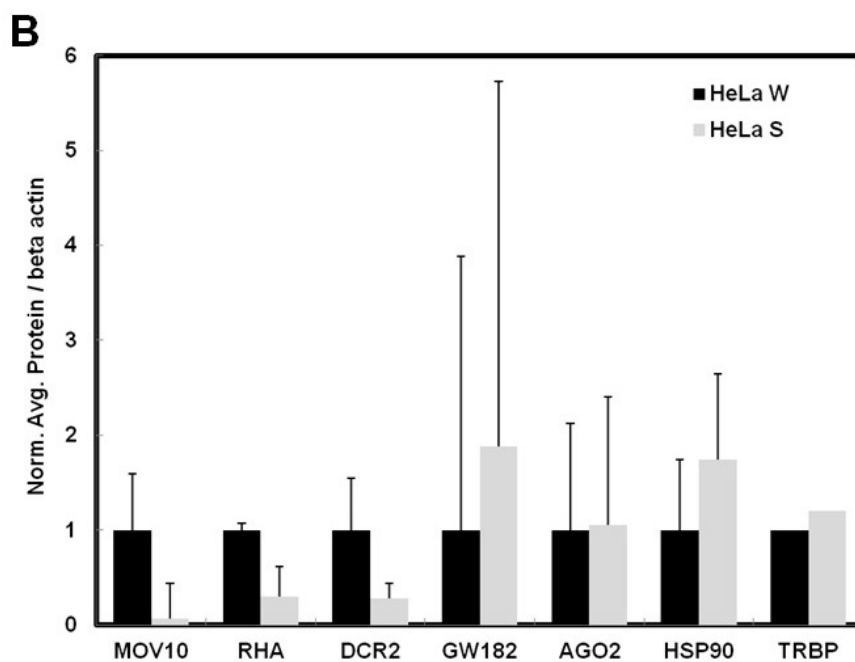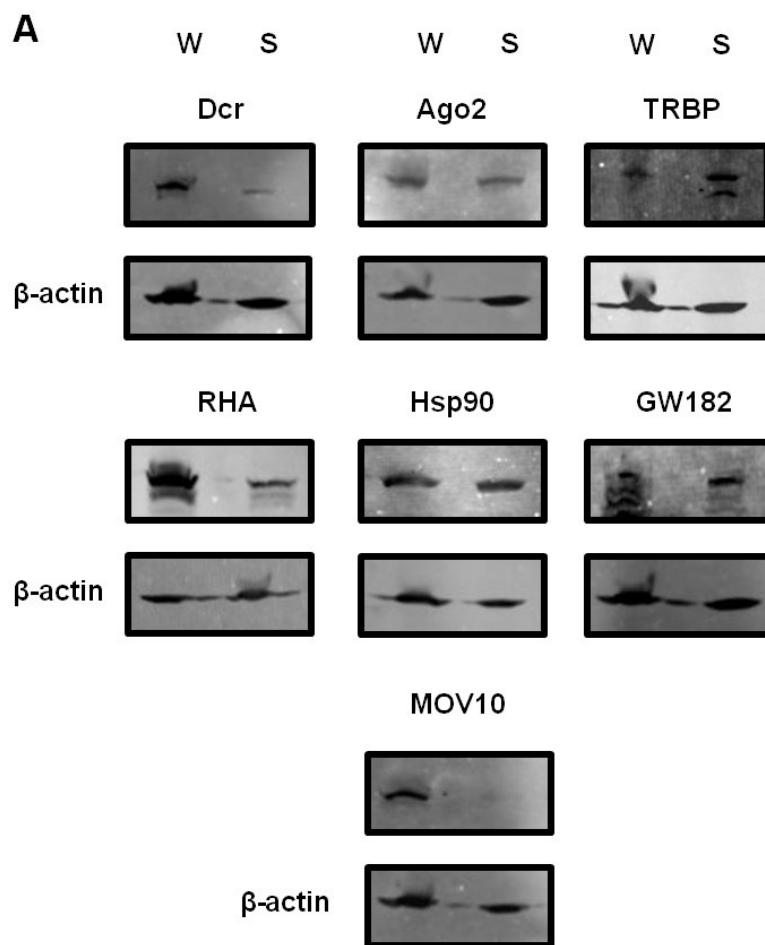
Figure 4.5. Quantitative Western blot analyses of Ago associated proteins in HeLa whole cell lysate and HeLa S100 cytoplasmic. (a) Denaturing Western blot analyses of extracts using antibodies against Dcr, Ago, TRBP, RHA, Hsp90, GW182 and MOV10 in both HeLa whole cell lysate(W) and HeLa S100 cytoplasmic extract(S). Beta- Actin was used as a loading control. **(b).** Graphical representation of the normalized and averaged ratio of protein of interest over beta-actin. The black and gray bars represent HeLa whole cell lysate(HeLa W) and HeLa S100(HeLa S) cytoplasmic extract respectively.

to have a potential role in strand unwinding and loading of siRNA duplex into Ago2[20, 21].

ATP has an inhibitory effect on target mRNA cleavage mediated by single-stranded antisense siRNA in HeLa S100 cytosolic extract. It has been previously reported that ATP drives RISC assembly and is required for efficient loading of small-RNA duplexes into AGO proteins but not for subsequent unwinding of siRNAs[37]. To investigate if addition of ATP had an effect on the ability of HeLa S100 cytosolic extract to utilize siRNA duplexes for RNAi, we performed slicing assays with extracts supplemented with ATP (**Fig. 4.6a**). Addition of ATP has an inhibitory effect on target mRNA cleavage mediated both by single-stranded antisense siRNA and siRNA duplex (**Fig. 4.6a**). Furthermore, completely depleting ATP from the extract led to enhanced cleavage activity (**Fig. 4.6b**). This is contrary to previous reports that suggested that ATP-dependent RISC loading and helicase activities are required for subsequent ATP-independent target mRNA cleavage[39, 40].

3.4 Discussion

The potential of the RNAi pathway to utilize exogenous genetic material to activate sequence specific regulation of gene expression has been exploited extensively in functional genomics and therapeutic applications. Although RNAi has found several applications, the mechanistic details of the pathway, composition of functional complexes and significance of protein role-players remain unclear. Enormous effort has focused on *Drosophila, C. elegans* and mammalian systems both *in vivo* (in live organisms and cultured cells) and *in vitro* (using cell extracts and purified proteins), to gain a better understanding of the pathway. In this study, we have employed seemingly well established preparations of HeLa cell extracts, in particular HeLa whole cell lysate and HeLa S100 cytosolic extract[41], which have been previously used by several research groups as model systems to study RNAi[30, 37] (**Fig. 4.1**).
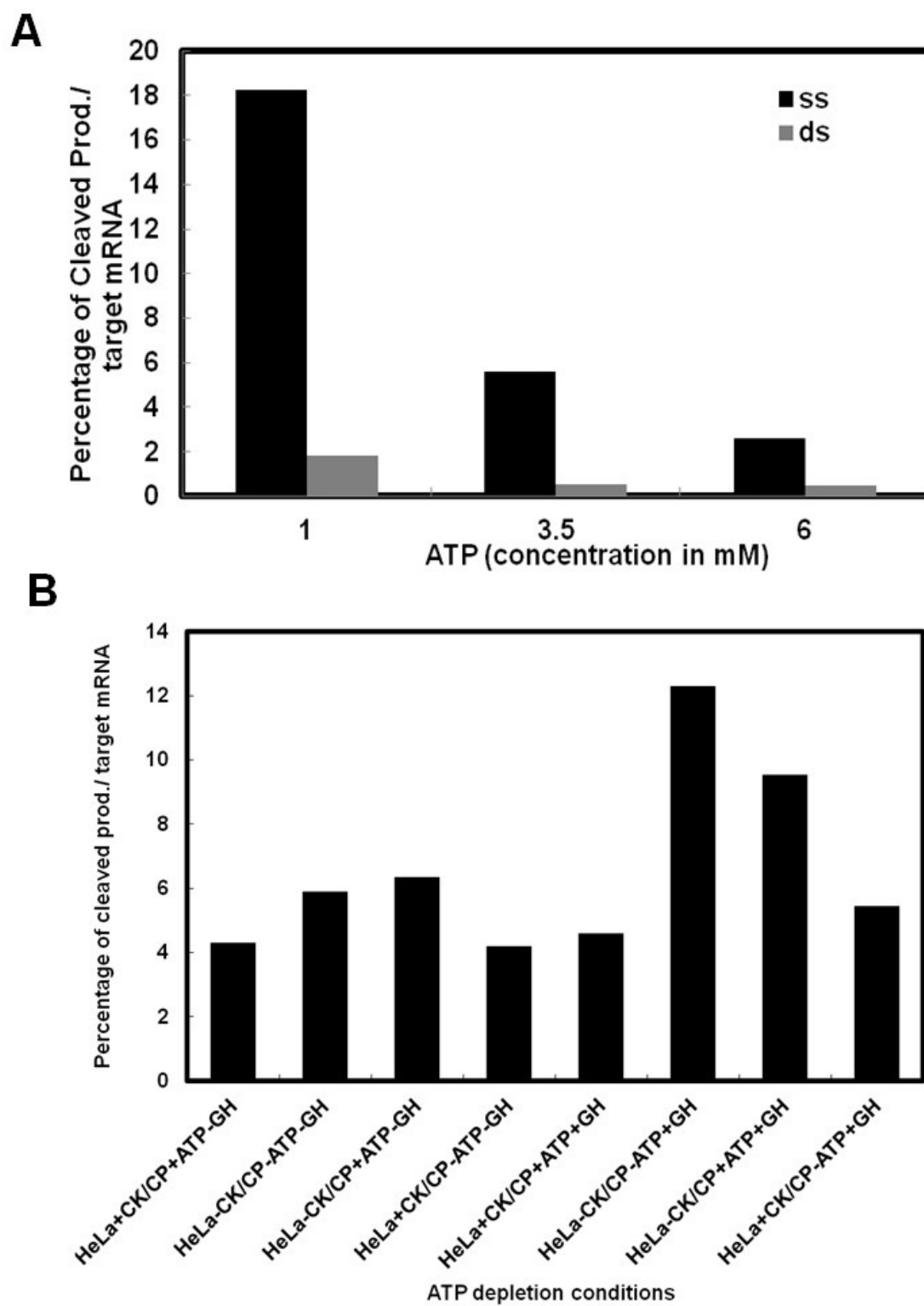
Figure 4.6. Effect of ATP on single stranded antisense siRNA and siRNA duplexes to variable extents to mediate target mRNA cleavage in HeLa

S100 cytoplasmic. Graphical representation of the normalized and averaged fraction of cleaved target mRNA product over full length target mRNA for slicing activity HeLa S100 cytoplasmic extract. **(a)** The slicing assay was performed using 1 mM, 3.5 mM or 6 mM ATP as shown in the graph. The gray and black bars indicate siRNA duplex (ds) and single stranded antisense siRNA (ss) mediated target mRNA cleavage respectively. **(b)** The slicing assay was performed under the indicated ATP depletion conditions. ATP regeneration system: CK- Creatine Kinase, CP- Creatine Phosphate; G- glucose; H- Hexokinase. HeLa S100 cytosolic cell extracts were treated with glucose and Hexokinase to deplete any residual ATP present in the extract.
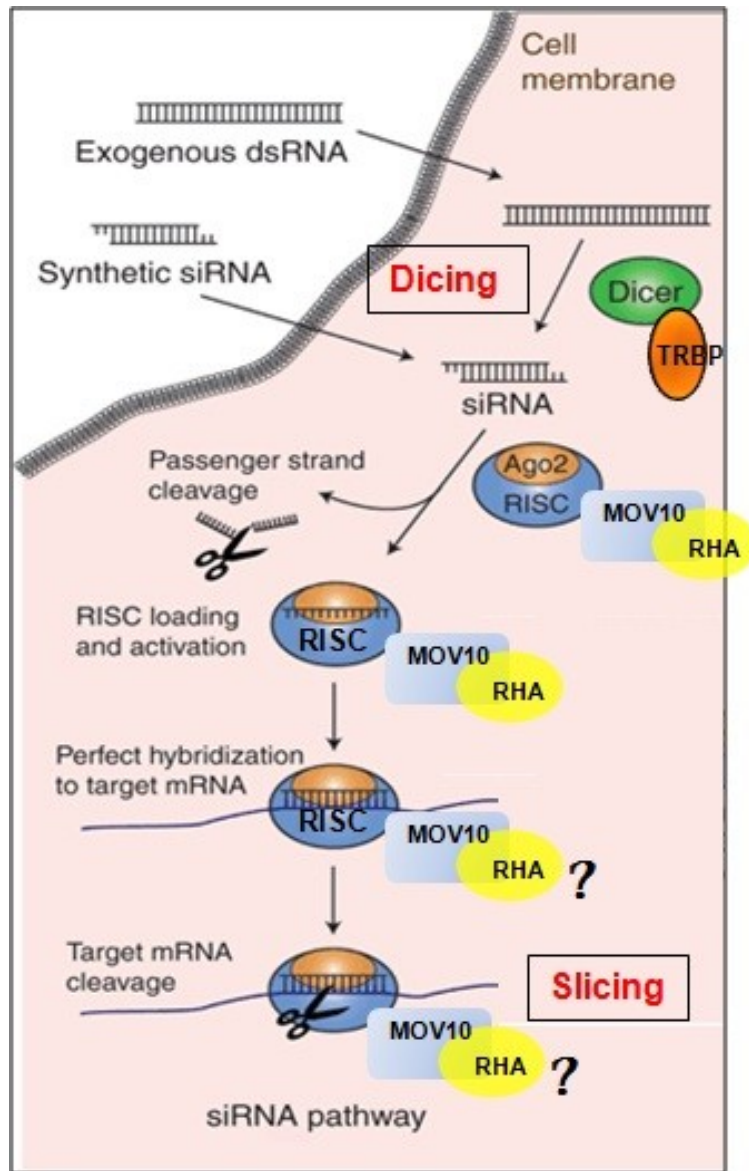
Figure 4.7. MOV10/RHA have a potential role in RISC loading of siRNA duplexes and strand unwinding

Upon performing slicing assays, we found that HeLa whole cell lysate utilizes siRNA duplexes more efficiently than single-stranded antisense siRNAs, similar to intracellular observations[32]. However, the converse was true for HeLa S100 cytosolic extract (**Fig. 4.3**), i.e., single-stranded antisense siRNAs elicited RNAi more robustly than siRNA duplexes. This observation is contrary to previous reports that have shown the S100 cytoplasmic extracts to bring about efficient mRNA cleavage with siRNA duplexes compared to single-stranded siRNAs[30]. One potential rationale for such contrasting results lies in the methods of extract preparation, where the S100 cytoplasmic extract lacks known RISC associated factors that localize to the nucleus[42, 43, 44] and potential high molecular weight RISC associated proteins and complexes[45] that may be removed during centrifugation steps used for enrichment of the cytoplasmic fraction. We also observed that supplementing HeLa S100 cytosolic extract with HeLa whole cell lysate in slicing assays increased the ability of the S100 extract to utilize siRNA duplex much more efficiently as shown in **Fig 4.4**. On the contrary dicing assays performed with these extracts showed that both the extracts can cleave a Dicer substrate with the same efficiency (**Fig 4.2**). The quantitative Western blot analyses of RNAi related and Ago interacting proteins presented in **Fig. 4.3**, together with previous reports[44, 43, 42, 32,21] that these proteins localize to the nucleus, strongly support the above hypothesis. More specifically, MOV10, RHA and Dcr, which have been found to localize to the cell nucleus to significant extent[44, 43, 42, 32, 21], were significantly more abundant in whole cell lysate than S100 cytosolic extract. This strongly suggests that the poor efficiency of siRNA duplexes to elicit RNAi in S100 extract is due to the loss of these nuclear localized RISC loading factors. Previous reports have also shown that the high-molecular weight (~500 kDa) RLC complex containing Ago, Dcr and TRBP is formed prior to the encounter with pre-miRNAs or double-stranded siRNAs in humans[46], thus increasing the likelihood of losing such complexes during the preparation of S100 cytoplasmic extracts. Additionally, MOV10 was found to be associated with ~2 MDa TRBP-containing complex that mediate miRNA silencing[45]. Put together, the results presented here suggest that a systematic

difference in extract preparation and abundance of the RNA helicases MOV10 and RHA as potentially important cofactors involved in loading double-stranded siRNA duplexes into RISC may critically modulate RNAi efficiency. These helicases have been previously shown to associate with Ago and are required for formation of mature and active RISC, however, their exact roles in RISC still remain unknown[20, 21]. Additionally the fact that both these extracts have similar dicing activity suggests a role for MOV10 and RHA after the dicing step in the RNAi pathway as depicted in our model (**Fig 4.7**).

Another interesting observation from our results was the inhibitory effect of ATP on siRNA duplex and single-stranded antisense siRNA mediated target mRNA cleavage (**Fig 4. 6**). Although this observation needs further investigation, similar observations using Rabbit Reticulocyte extract suggested that these extracts lack miRNP or RISC loading activity and may employ an alternative mechanism to utilize the exogenous single-stranded antisense siRNA to mediate target mRNA cleavage[38]. Alternatively, as observed in the case of Dicer, where the deletion of the ATP-dependent helicase domain results in increased cleavage activity by the enzyme[47], ATP may be mediating some yet-to-be-identified allosteric and/or structural changes in RISC that have an inhibitory effect on its target mRNA cleavage activity.

In summary, our results suggest a potentially critical role for MOV10 and RHA in RISC loading of mature siRNA duplexes. Although the use of different systems to study RNAi shed light on various mechanisms in RNAi and the roles for new Ago or RISC associated proteins, it also does add a layer of complexity to interpreting the pathway mechanism. Follow up studies on the inhibitory effect of ATP on single-stranded antisense siRNA mediated target mRNA cleavage may lead to the discovery of a yet-to-be-identified mechanistic detail of target degradation.

4.5 Acknowledgments

4.6 References

1. Bartel, D.P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297.
2. Dykxhoorn, D.M., Novina, C.D. and Sharp, P.A. (2003). Killing the messenger: Short RNAs that silence gene expression. *Nature Reviews Molecular and Cell Biology* **4**, 457–467.
3. Meister, G. and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343–349.
4. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806–811.
5. Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293–296.
6. Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G. and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell* **15**, 185–197.
7. Gregory, R. I., Chendrimada, T. P., Cooch, N. and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631.
8. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. and Doudna, J. A. (2005). In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences* **105**(2): 512–517.
9. Peters L, Meister G. (2007). Argonaute proteins: mediators of RNA silencing. *Molecular Cell* **26**(5),611-623.
10. Picard, D. (2002). Heat-shock protein 90, a chaperone for folding and regulation. *Cell and Molecular Life Sciences* **59**, 1640–1648.
11. Miyoshi, T., Takeuchi, A., Siomi, H. and Siomi, M. C. (2010). A direct role for Hsp90 in pre-RISC formation in Drosophila. *Nature Structure and Molecular Biology* **17**(8),1024-1026.
12. 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 Report* **5**, 189–194.
13. Johnston, M., Geoffroy, M. C., Sobala, A., Hay, R. and Hutvagner, G. (2010). HSP90 protein stabilizes unloaded argonaute complexes and microscopic P-bodies in human cells. *Molecular Biology of the Cell* **21**(9):1462-1469.
14. Pare, J. M., Tahbaz, N., Lopez-Orozco, J., LaPointe, P., Lasko, P. and Hobman, T. C. (2009). Hsp90 regulates the function of argonaute 2 and its recruitment to stress granules and P-bodies. Mol. Biol. *Cell* **20**, 3273–3284.
15. Liu J, Rivas F.V, Wohlschlegel J., Yates J.R., Parker R. and Hannon G.J. (2005). A role for the P-body component GW182 in microRNA function. *Nature Cell Biology* **7**, 1261–1266.

16. 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,**1167–1174.

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

18. 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. *Nature Structure and Molecular Biology* **18,** 1218–1226.

19. Yao, B., La, L. B., Chen, Y. C., Chang, L. J. and Chan, E. K.(2012). Defining a new role of GW182 in maintaining miRNA stability. *EMBO Report* **13**(12),1102-1108.

20. 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* **15**, 2149–2155.

21. Robb, G. B. and Rana, T. M. (2007). RNA helicase A interacts with RISC in human cells and functions in RISC loading. *Molecular Cell* **26**(4),523-537.

22. Tomari, Y., Du, T., Haley, B., Schwarz, D.S., Bennett, R., Cook, H.A., Koppetsch, B.S., Theurkauf, W.E. and Zamore, P.D. (2004). RISC assembly defects in the Drosophila RNAi mutant armitage. *Cell* **116**, 831–841.

23. Sievers, C., Schlumpf, T., Sawarkar, R., Comoglio, F. and Paro R. (2012). Mixture models and wavelet transforms reveal high confidence RNA-protein interaction sites in MOV10 PAR-CLIP data. *Nucleic Acids Research* **40**(20), e160.

24. Lucchesi, J.C. (1998). Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation. *Current Opinions in Genetics and Development* **8**, 179–184.

25. Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293–296.

26. 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* **293**, 1146–1150.

27. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. & Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563-74.

28. Hoerter, J. A. & Walter, N. G. (2007). Chemical modification resolves the asymmetry of siRNA strand degradation in human blood serum. *RNA* **13**, 1887-1893.

29. Walter, N. G. (2002). Probing RNA structural dynamics and function by fluorescence resonance energy transfer. In *Curr. Protocols Nucleic Acid Chem*, pp. 11.10.11-11.10.23.

30. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. & Tuschl, T. (2002). Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563-574.

31. He, B., Rong, M., Lyakhov, D., Gartenstein, H., Diaz, G., Castagna, R., McAllister, W. T. and Durbin, R. K. (1997). Rapid mutagenesis and purification of phage RNA polymerases. *Protein Expression and Purification* **9**, 142-151.

32. 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**(6836):494-498.

33. Tuschl, T. (2001). RNA interference and small interfering RNAs. *Chembiochem: a European journal of chemical biology* **2**(4):239-245.

34. Elbashir, S. M., Lendeckel, W. and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes and Development* **15**(2):188-200.

35. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D.P. and Sharp, P. A. (1999). Targeted mRNA degradation by double-stranded RNA in vitro. *Genes and Development* **13**(24), 3191-3197.

36. Zamore, P. D., Tuschl, T., Sharp, P. A. and Bartel, D. P. (2000). RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* **101**(1), 25-33.

37. Wang, B., Love, T. M., Call, M. E., Doench, J. G. and Novina, C. D. (2006). Recapitulation of short RNA-directed translational gene silencing in vitro. *Molecular Cell* **22**(4),553-560.

38. Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q. and Tomari Y. (2010). ATP-dependent human RISC assembly pathways. *Nature Structure and Molecular Biology* 17(1),17-23.

39. Nykänen, A., Haley, B. and Zamore, P.D. (2001). ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107**, 309–321.

40. Schwarz, D. S., Hutvágner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P. D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**, 199–208.

41. Mayeda, A. and Krainer, A. R. Preparation of Hela Cell Nuclear and Cytosolic S100 Extracts for In Vitro Splicing. *Methods in Molecular Biology* **118**(RNA-Protein Interaction Protocols)**,** 309-314.

42. Tang, H., McDonald, D., Middlesworth, T., Hope, T.J. and Wong- Staal, F. (1999). The carboxyl terminus of RNA helicase A contains a bidirectional nuclear transport domain. *Molecular and Cellular Biology* **19**, 3540– 3550.

43. El Messaoudi-Aubert, S., Nicholls, J., Maertens, G. N., Brookes, S., Bernstein, E. and Peters, G. (2010). Role for the MOV10 RNA helicase in Polycomb-mediated repression of the INK4a tumor suppressor. *Nature Structure and Molecular Biology* **17**(7),862-868.

44. Ando, Y., Tomaru, Y., Morinaga, A., *et al.* (2011). Nuclear pore complex protein mediated nuclear localization of dicer protein in human cells. *PLoS One* **6**(8):e23385.

45. Chendrimada, T. P., Finn, K. J., Ji, X., Baillat, D., Gregory, R. I., Liebhaber, S. A., Pasquinelli, A. E. and Shiekhattar, R. (2007). MicroRNA silencing through RISC recruitment of eIF6. *Nature* **447**(7146):823-828.

46. Maniataki, E. and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes and Development* **19**, 2979–2990.

47. Ma, E., MacRae, I. J., Kirsch, J. F. and Doudna, J. A. (2008). Autoinhibition of human dicer by its internal helicase domain. *Journal of Molecular Biology* **380**(1):237-243.

48. Ziehler, W. A. and Engelke, D. R. (2006). Probing RNA structure with chemical reagents and enzymes. Current Protocols in Nucleic Acid Chemistry, 6.1.1-6.1.21.

Chapter 5

Summary and Future Directions

RNAs and more specifically non-coding RNAs (ncRNAs) have gained immense popularity over the last decade, since the discovery that RNA plays key structural, functional and regulatory roles within the cell[1]. All of the research presented in this thesis has focused on understanding the role played by one such regulatory ncRNA called short interfering RNA (siRNA). siRNAs are mostly exogenous in origin and can be generated from long double-stranded (ds)RNA derived from trespassing RNA viruses, mobile genetic elements, self annealing transcripts or experimental transfection[2]. RNase III enzyme Dicer (Dcr) cleaves this long ds-RNA into ~21-28 nt siRNAs that are the functional triggers for the RNA interference (RNAi) pathway[3]. These mature siRNAs are then incorporated into a RNA Induced Silencing Complex (RISC) that mediates the sequence specific cleavage of a target mRNA with perfectly complementary sequence[4, 5, 6]. Synthetic mature siRNAs are also incorporated efficiently into RISC and can equally function as triggers for the RNAi pathway[7]. The inherent ability of the RNAi pathway to utilize synthetic siRNAs as triggers to mediate sequence specific target mRNA cleavage has been exploited in functional genomics and therapeutic applications.

The aspects of RNAi that have been studied as part of the work reported in this thesis include the investigation of: (1) the interactions of p19, a viral silencing suppressor protein, with siRNAs and RNAi pathway components; (2) the chemical stability of siRNA-like small RNAs in cell-like and blood-like environments; and (3) the process of RISC loading and activation in model cell

extracts that have demonstrated differential capacity for utilizing single-stranded versus double-stranded siRNAs.

5.1 Kinetic characterization of RNA silencing suppressor protein, p19

The RNAi pathway is believed to have evolved to combat viral infections and to protect host cells from mobile genetic elements by inhibiting the replication of virus-derived transposons. As a countermeasure, viruses are thought to have evolved, in parallel, RNAi silencing suppressor (RSS) proteins[8, 9, 10] to circumvent the RNAi mechanism. p19 is one such RSS protein, which has been shown to function as a homodimer in tombusviruses[11]. p19 binds mature siRNA duplexes, products of Dcr, in a caliper-like size specific, but sequence non-specific manner, thus making it unavailable for incorporation into mature RISC, thus inhibiting RNAi mediated virus RNA cleavage[12, 13, 14].

The work presented in Chapter 2 of this thesis characterizes the interaction between p19 and components of the RNAi pathways using biochemical assays and computational modeling. A solution-based fluorescence quenching assay performed with p19 showed that it binds reversibly to siRNAs with fast association ($(1.69 \pm 0.07) \times 10^8$ $M^{-1}s^{-1}$) and dissociation (0.062 ± 0.002 s^{-1}) rate constants (**Fig 2.3**). It was observed that p19 binds to siRNA with a significantly higher affinity ($K_{P,sol} = 0.37 \pm 0.08$ nM) than Dcr ($K_D = 3.7 \pm 0.4$ nM) (**Fig 2.4a**). Competition of p19 with human Dicer for siRNA binding detected by EMSA showed that p19 efficiently competes with human Dcr, thereby disrupting the siRNA:Dcr complex (**Fig 2.4a**). This led to the hypothesis that p19 may be able to function to inhibit RNAi in humans by inhibiting the process by which Dicer hands off siRNA to RISC. Conversely, Dcr was unable to disrupt the p19:siRNA complex even when present at saturating concentrations, owing to the high affinity of p19 for siRNAs (**Fig 2.4b**). Additionally, from competition of p19 with human siRNA-containing complexes found in cytosolic HeLa cell extract we observed that adding p19 to the cell extract prior to siRNA addition increasingly impairs the formation of siRNA-containing complexes, concomitant with an accumulation of the competing siRNA:p19 complex (**Fig 2.5d,e**).

However, when p19 was added to fully mature siRNA containing complexes, the formation of particularly the Dicer-containing complex D was inhibited to a lesser extent (**Fig 2.5f,g; Fig 2.6**), suggesting that a tertiary interaction with the siRNA:Dcr complex is required for the interference of p19 with the RNAi pathway.

Computational modeling informed by our experimental results and reports from the literature further support the formation of a ternary complex that favors dissociation into the siRNA:p19 complex and free Dicer (**Fig 2.4c**). Additionally, it predicts a significant enhancement in viral protein load when the siRNA:p19 interaction is assumed to be reversible, as observed experimentally, thus invoking a multiple-turnover "catch and release" mechanism wherein p19, unlike the released siRNA, is effectively recycled after successfully competing with Dicer (**Fig. 2.7**).

RSS proteins play a key role in sustaining viral infection by combating the host anti-viral response posed by the RNAi pathway. RSS proteins may thus serve as ideal drug targets for shunting viral inhibitory effects. Thus, investigating RSS proteins as drug targets would facilitate restoring the host anti-viral mechanism during viral infection. Additionally, the techniques presented in this chapter could be employed likewise to systematically characterize other RSS proteins both biochemically and biophysically, to understand better their interaction with the RNAi pathway components. Delineating their interaction kinetics and mechanism of action would further their potential as viable drug targets. For instance, competing non-specific RNAs of the appropriate size could be dosed as "bait" for p19 binding, thus suppressing p19's inhibitory effect on RISC and ensuring robust antiviral response. It would also be interesting to investigate the role of Viral RSS proteins as potential immunosuppressors in viral gene therapy applications such as adenoviral gene therapy to suppress the host immune response against these viruses, very similar to the use of immunosuppressants in organ transplants.

5.2 Stability of siRNAs and their implication in RNAi therapeutics

The ability of synthetic siRNAs to function as effector molecules of the RNAi pathway has significant implications in RNAi therapeutics[15, 16]. The major challenges that concern the use of siRNA duplexes for therapeutic applications are their delivery, stability, specificity and potency[17]. Much research has gone into modifying the pharmacokinetic properties of siRNAs to enhance their utility as therapeutic agents. The primary problem posed by nuclease sensitive RNA molecules is their low stability when exposed to the nuclease-rich cellular environment. Although chemical modifications have been shown to provide significant stability to siRNA in the cellular environment[15, 18], questions regarding their utility within the cell persist.

The work presented in Chapter 3 of this thesis establishes a real-time assay to measure degradation profiles and kinetics for small, RNAi-related RNAs in both blood serum and RNAi-active cytosolic cell extract. We observed that fluorophore labeled small siRNA-like single-stranded (ss)RNAs, dsRNAs and DNA/RNA hybrids interact with several cellular enzymes, including RNase H1 and Dcr. It was observed that dsRNAs were more stable in both blood serum and RNAi-active cytosolic cell extract compared to ssRNA, although its hybridization to a DNA further stabilized the RNA component in blood serum (**Fig 3.2**). Additionally, the DNA/RNA hybrid that was strongly protected in serum, but rapidly degraded by RNase H1 in cell extract, which was confirmed by enzyme-specific inhibition of RHase H1 by two *in vitro* selected DNA aptamers (**Fig 3.2, Fig 3.3**).

We also observed a size specific protection of dsRNAs in both blood serum and RNAi-active cytosolic cell extract where 21- and 24-nt dsRNAs, but not 18-nt dsRNA, were equally protected in S100 cytosolic HeLa cell extract. However, in blood serum the relative decay rate constants correlated with their relative lengths, i.e., the rate of degradation decreased with increasing size of the dsRNA (**Fig 3.5**). We have shown that the 21- and 24-nt dsRNAs were incorporated in RISC related complexes (as seen in Chapter 2) to a greater extent than 18-nt dsRNAs, contributing to the faster decay rate of the 18-mer in

HeLa cell extract (**Fig 3.5**). The 24-nt dsRNA is a substrate of human Dicer (**Fig 3.4)** and the 21-nt dsRNA Dicer's product; coincidentally, these RNAs were protected specifically in the cellular milieu. Our results are consistent with established steps in the RNAi pathway where 21-nt dsRNAs are bound by the RISC loading complex (RLC) or other RNAi components and remain largely protected from nucleolytic decay, thus explaining the relative longevity of intracellular RNAi effects induced even by chemically unmodified siRNAs.

Although it has been hypothesized that the RLC or other RNAi associated proteins are responsible for binding to and protecting siRNA-like small RNAs upon entry into the cellular environment, the precise mechanism of this protection is still unclear. Future studies should focus on investigating what proteins or mechanisms are involved in rendering such protection to unmodified siRNA-like small RNAs. In addition to their localization and activity within the cell, there have been reports of miRNA in the bloodstream in a highly stable, extracellular form[19]. However, the mechanism underlying their remarkable stability in the RNase-rich environment of blood is still under debate. There has been evidence that these miRNAs are encapsulated in vesicles in the plasma, but it has also been shown that a significant portion of these miRNAs are associated with Ago2 complexes that may be responsible for the stability of plasma miRNAs[19]. Probing the mechanism behind the stability of these circulating miRNAs present in blood might give some insights into methods for delivery and improving the stability of exogenous siRNA introduced into the plasma for therapeutic purposes and can possibly function as signaling molecules as described [20].

5.3 Differential modes of RISC loading and activation facilitated by cell extracts

Despite the wide application of RNAi in basic research and therapeutics it remains a poorly understood mechanism. Although Dcr, Ago and TRBP have been identified as core components of a siRNA-programmed RISC that mediates sequence specific cleavage of mRNA target[21, 22], there have been conflicting ideas in the field about the potential role for other proteins that associate with

Ago2 or RISC[23]. Much of this complexity may arise from the lack of a standard system of study for understanding the RNAi pathway.

The work presented in Chapter 4 of this thesis describes the systematic differences observed in the RNAi pathway when using two different preparations of HeLa cell extract, i.e., S100 cytosolic cell extract and translationally active whole cell extract. We observed that despite demonstrating the ability to perform dicing and slicing activity (**Fig 4.1**), the extracts possess an inherent difference in their ability to utilize double- versus single-stranded siRNAs to mediate target mRNA cleavage. The HeLa whole cell extract accommodated siRNA duplexes more efficiently than single-stranded siRNAs to mediate target cleavage, as observed in previous reports. By contrast, the opposite was true for HeLa S100 cytosolic extract, which lacks any nuclear components that might be involved in the RNAi pathway (**Fig 4.2**). Quantitative Western blot analysis of both extracts for Ago-associated proteins indicated that the HeLa S100 cytoplasmic extract had much lower concentrations of three proteins, MOV10, RHA and Dcr, which have been shown to be present in both the nuclear and cytosolic compartments of the cell and are required for mature RISC formation and activity (**Fig 4.3**). Additionally, we noticed that ATP in HeLa S100 cytosolic extract has an inhibitory effect on both siRNA duplex and single-stranded guide strand mediated target mRNA cleavage (**Fig 4.4**). Taken together, our results suggest a potential role for MOV10 and RHA in siRNA duplex loading of RISC. We also hypothesize that ATP may be mediating some yet-to-be-identified allosteric and/or structural change in RISC that has an inhibitory effect on its target mRNA cleavage activity.

Although MOV10 and RHA have been implicated in RISC loading and active RISC formation, the first step in confirming this observation would be to supplement HeLa S100 cytoplasmic extract with different molecular weight fractions of HeLa whole cell extract or purified recombinant MOV 10 or RHA and ask whether functional complementation is observed upon siRNA duplex loading into RISC. A reconstitution experiment with purified recombinant proteins of RISC would verify the implied role of MOV10 and RHA in RISC loading and active RISC formation. Additionally, follow up studies on the inhibitory effect of

ATP on single-stranded antisense siRNA mediated target mRNA cleavage would help in unraveling the mechanistic details of target degradation. Many conflicting ideas exist related to the role of RISC associated proteins, such as the presence of Dicer in the mature effector RISC, in the RNAi pathway, which could be attributed to the fact that most results reported thus far are from unsynchronized bulk assays. In order to gain deeper mechanistic insights into the role of siRNA-containing complexes in RISC loading, unwinding of the siRNA duplexes, binding to the target mRNA and cleavage, in real-time, single-molecule fluorescence microscopy tools can be employed. Observing these processes at the single-molecule level will shed light on the heterogeneity associated with these RNAi processes and will help delineate their kinetics.

5.4 References

1. Mattick, J. S., and Makunin, I. V. (2006). Non-coding RNA. *Human Molecular Genetics* **15** (1), R17-29.

2. Kanasty, R. L., Whitehead, K. A., Vegas, A. J. and Anderson, D. G. (2012). Action and reaction: the biological response to siRNA and its delivery vehicles. *Molecular Therapy* **20**(3), 513-524.

3. 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**(6836):494-498.

4. Hammond, S. M., Bernstein, E., Beach, D. and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293–296.

5. Elbashir, S. M., Lendeckel, W. and Tuschl, T. (2001). RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes and Development* **15**(2):188-200.

6. Gregory, R. I., Chendrimada, T. P., Cooch, N. and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631.

7. Kim, D. H., Behlke, M. A., Rose, S. D., Chang, M. S., Choi, S. & Rossi, J. J. (2005). Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nature Biotechnology* **23**, 222-226.

8. Qu, F. and Morris. T. J. (2005). Suppressors of RNA silencing encoded by plant viruses and their role in viral infections. *FEBS Letters* **579**(26),5958-5964.

9. Roth, B. M., Pruss, G. J., and Vance, V. B. (2004). Plant viral suppressors of RNA silencing. *Virus Research* **102**(1),97-108.

10. Voinnet, O. (2001). RNA silencing as a plant immune system against viruses. *Trends in Genetics* **17**(8),449-459.

11. Scholthof, H. (2006) The Tombusvirus-encoded P19: from irrelevance to elegance. *Nature Reviews Microbiology* **4**, 405-411.

12. Vargason, J. M., Szittya, G., Burgyan, J. and Hall, T. M. (2003). Size selective recognition of siRNA by an RNA silencing suppressor. *Cell* **115,** 799-811.

13. Ye, K., Malinina, L. and Patel, D. J. (2003). Recognition of siRNA by a viral suppressor of RNA silencing. *Nature* **426**, 874-878.

14. Xia, Z., Zhu, Z., Zhu, J. and Zhou, R. (2009). Recognition mechanism of siRNA by viral p19 suppressor of RNA silencing: a molecular dynamics study. *Biophysical Journal* **96**(5), 1761-1769.

15. Rana, T. M. (2007). Illuminating the silence: understanding the structure and function of small RNAs. *Nature Reviews Molecular and Cell Biology* **8**, 23-36.

16. de Fougerolles, A., Manoharan, M., Meyers, R., and Vornlocher, H. P. (2005).RNA interference in vivo: toward synthetic small inhibitory RNA-based therapeutics. *Methods in Enzymology* **392**, 278-96.

17. de Fougerolles, A., Vornlocher, H. P., Maraganore, J., and Lieberman, J. (2007). Interfering with disease: a progress report on siRNA-based therapeutics. *Nature Reviews Drug Discovery* **6**, 443-53.

18. Hoerter, J. A., and Walter, N. G. (2007). Chemical modification resolves the asymmetry of siRNA strand degradation in human blood serum. *RNA* **13**, 1887-93.

19. Arroyoa, J. D., Chevilleta, J. R., Kroh, E. M., *et al.* (2011). Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proceedings of the National Academy of Sciences* Early Edition, 1-6.

20. Lehmann, S. M., Krüger, C., Park, B., *et al*. (2012). An unconventional role for miRNA: let-7 activates Toll-like receptor 7 and causes neurodegeneration. *Nature Neuroscience* **15**(6), 827-835.

21. Gregory, R. I., Chendrimada, T. P., Cooch, N. and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**, 631.

22. MacRae, I. J., Ma, E., Zhou, M., Robinson, C. V. and Doudna, J. A. (2005). In vitro reconstitution of the human RISC-loading complex. *Proceedings of the National Academy of Sciences* **105**(2): 512–517.

23. Peters L, Meister G. (2007). Argonaute proteins: mediators of RNA silencing. *Molecular Cell* **26**(5),611-623.

Viral RNAi suppressor reversibly binds siRNA to outcompete Dicer and RISC via multiple-turnover
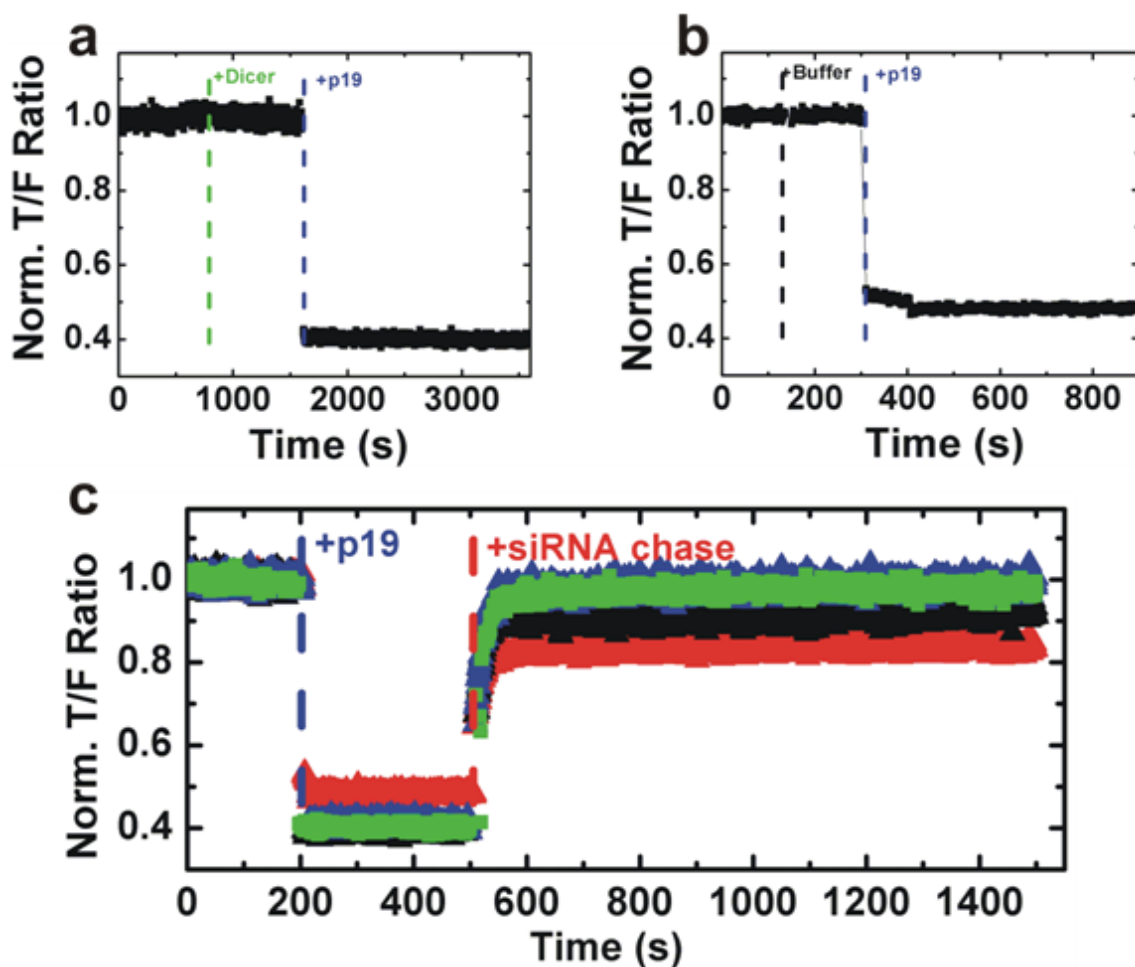
Figure A.1. Concentration dependence of siRNA:p19 dissociation kinetics. (a, b) Doubly-labeled siRNA (50 nM) after addition of Dicer (5 nM green dotted line) or Buffer (black dotted line) respectively, followed by addition of p19 (500

nM, blue dotted line). **(c)** Doubly-labeled siRNA (50 nM) was bound by p19 (500 nM, blue dotted line) after which excess unlabeled siRNA was added at concentrations of 500 nM (red), 750 nM (black), 1.5 µM (green), 3 µM (blue), respectively. Reprinted with permission from Elsevier.
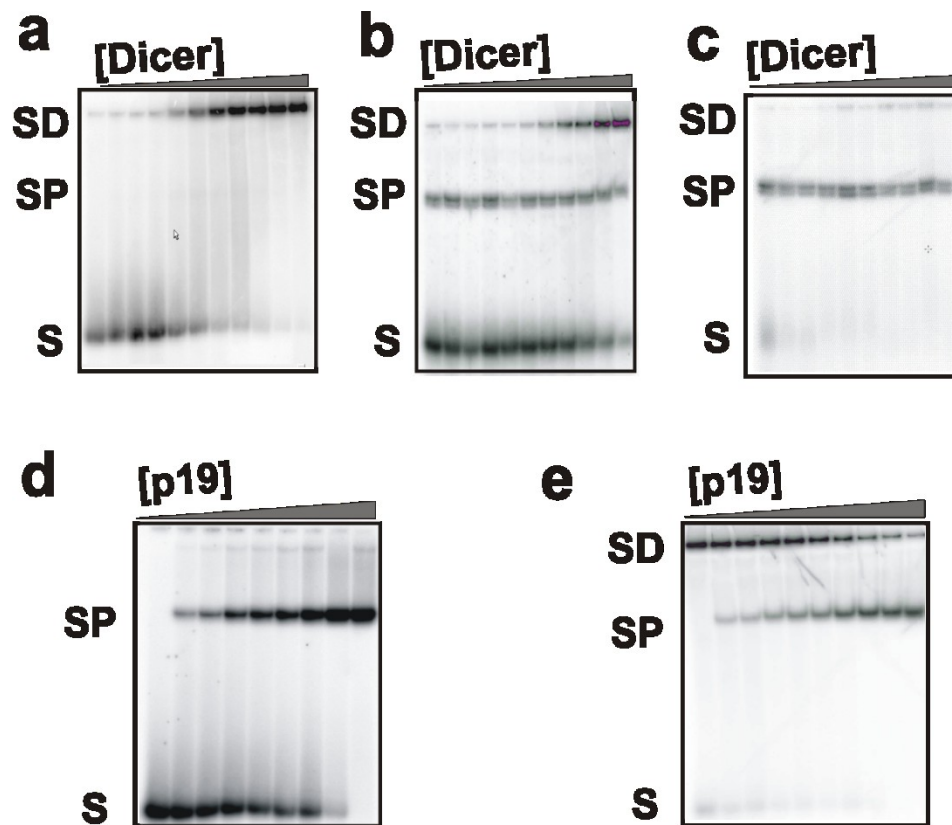
Figure A.2. Competition between p19 and human Dicer for siRNA binding as shown by EMSAs. siRNA (S) labeled with [γ–^{32}P] ATP are bound to **(a)** Recombinant human Dicer (D) (Genlantis Inc.) where Dicer concentrations are 0, 0.16, 0.25, 0.5, 2.5, 5, 10, 16, 50 ,100, 200 nM, respectively. Samples were incubated for 30 min before loading onto a 12% nondenaturing polyacrylamide gel. **(b)** siRNA pre-incubated with p19 at a concentration of 0.16 nM for 30 min followed by Dicer addition as described above. **(c)** siRNA pre-incubated with p19 at a concentration of 2.56 nM for 30 min followed by Dicer addition as described above. **(d)** p19 only (P) where p19 concentrations are 0, 0.02, 0.04, 0.08, 0.16, 0.32, 0.64, 1.28, 2.56, 5.12 nM, respectively. Samples were incubated for 30 min before loading onto a 12% nondenaturing polyacrylamide gel. **(e)** siRNA pre-incubated with Dicer at a concentration of 140 nM for 30 min followed by p19 addition as described in panel d. Reprinted with permission from Elsevier.
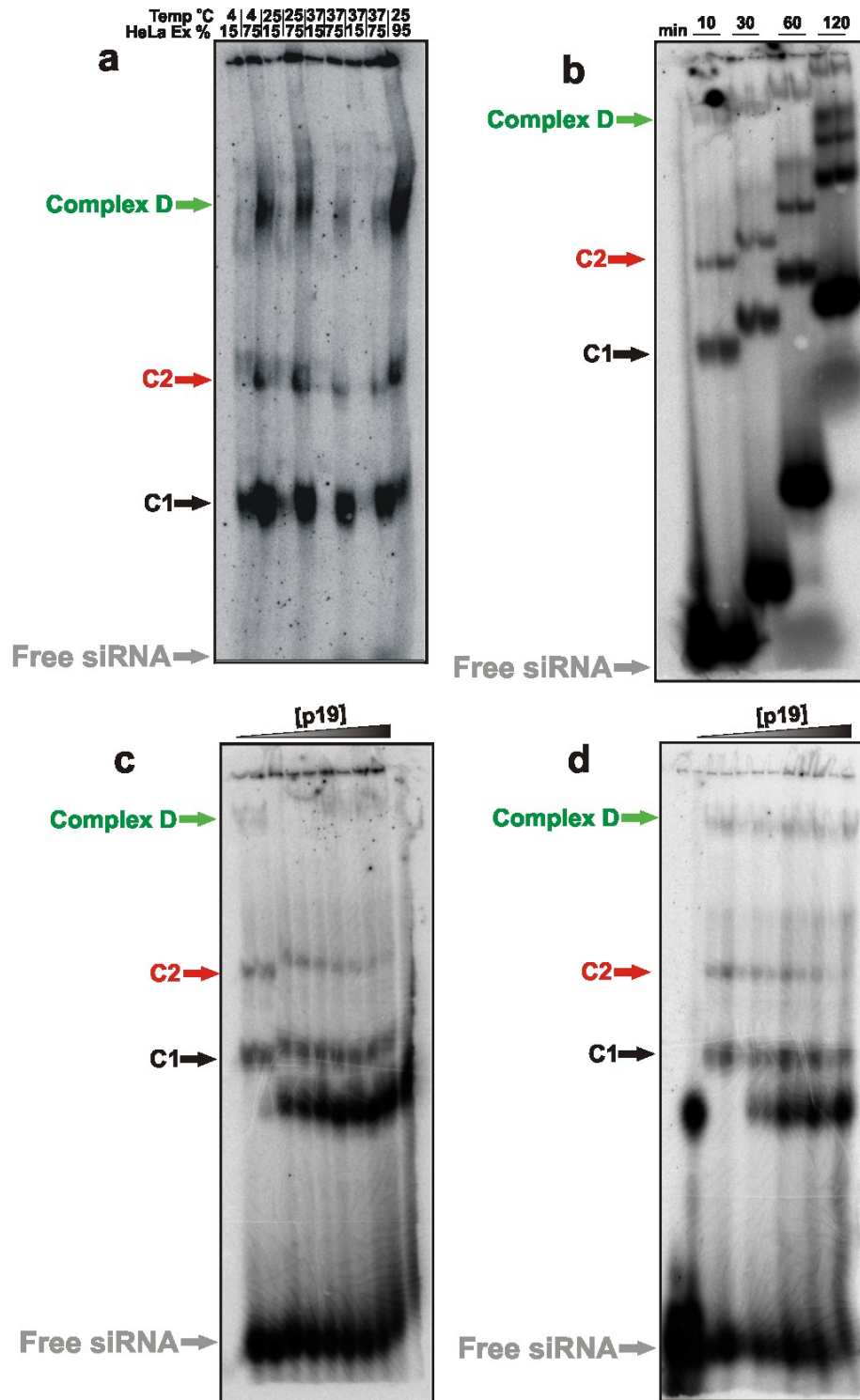
Figure A.3. Comparison of siRNA complex formation for different HeLa cell extract concentrations and temperatures. (a) (lanes 1-2) 15% or 75% (v/v) HeLa extract conditions incubated with siRNA for 1 h at 4 °C, (lanes 3-4) incubated at 25 °C, (lanes 5-8) incubated at 37 °C, and (lane 9) 93% (v/v) HeLa

cell extract in the absence of an ATP generating system and in supplied buffer (20 mM HEPES-Na pH 7.9, 42 mM ammonium sulfate, 0.2 mM EDTA, 0.5 mM DTT and 20% glycerol). The effect of incubation temperature and HeLa extract concentration on complex formation was limited. The same migration patterns are seen under all conditions. **(b)** Formation of complexes D, C2, and C1 after 10, 30, 60 and 120 min incubation of siRNA in 50% cytosolic HeLa cell extract. Samples were loaded onto a running 4% native polyacrylamide gel, leading to the indicated differences in migration. **(c)** Formation of complexes after 2.5 h incubation of siRNA in 50% (v/v) cytosolic HeLa cell extract, and increasing concentrations of p19. **(d)** Formation of complexes after 2 h incubation of siRNA in 50% (v/v) cytosolic HeLa cell extract, followed by the addition of increasing concentrations of p19 and further incubation for 30 min. Reprinted with permission from Elsevier.
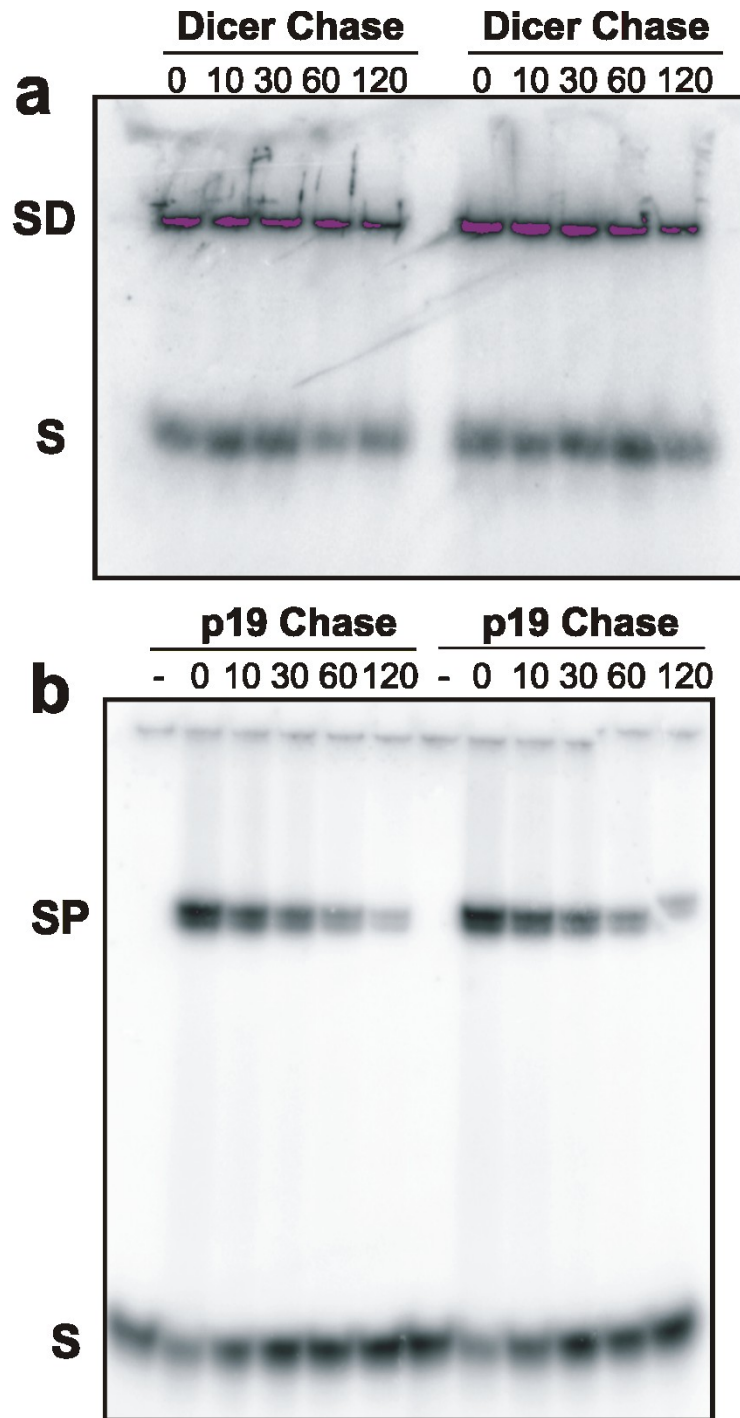
Figure A.4. Gel-based dissociation chase assays. Duplicate gel chase experiments were done in standard buffer with **(a)** 70 nM recombinant Dicer or **(b)** 10 nM of p19 added to 5′-^{32}P labeled siRNA before addition of 750 nM unlabeled siRNA chase at 0, 10, 30, 60 and 120 min. Reprinted with permission from Elsevier.
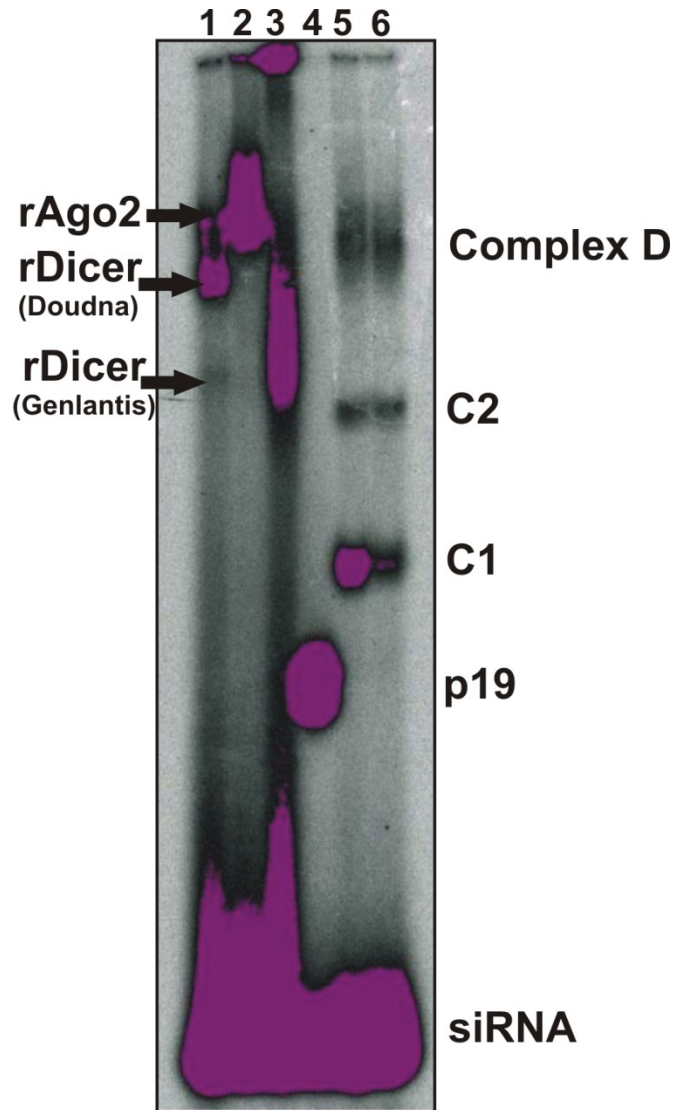
Figure A.5. Size comparison of protein-siRNA binding complexes. $[\gamma-^{32}P]$ ATP labeled siRNA incubated with: purified rDicer provided from the J. Doudna lab (lane 1); purified recombinant Ago2 protein provided from the J. Doudna lab (lane 2); rDicer (Genlantis Inc.; lane 3); purified p19 (lane 4); 15% (v/v) HeLa cell extract (lanes 5, 6).   All samples were incubated for 30 min before loading. Nondenaturing 4% (w/v) polyacrylamide gel electrophoresis was used to resolve these complexes as described above. Although Dicer is a larger protein (MW 218 kDa) compared to Ago2 (MW 97 kDa), it has an isoelectric point of 5.47 compared to 9.39 for Ago2. This difference contributes to Ago2 running as an apparently slower migrating complex in buffer at pH 7.4. Reprinted with permission from Elsevier.
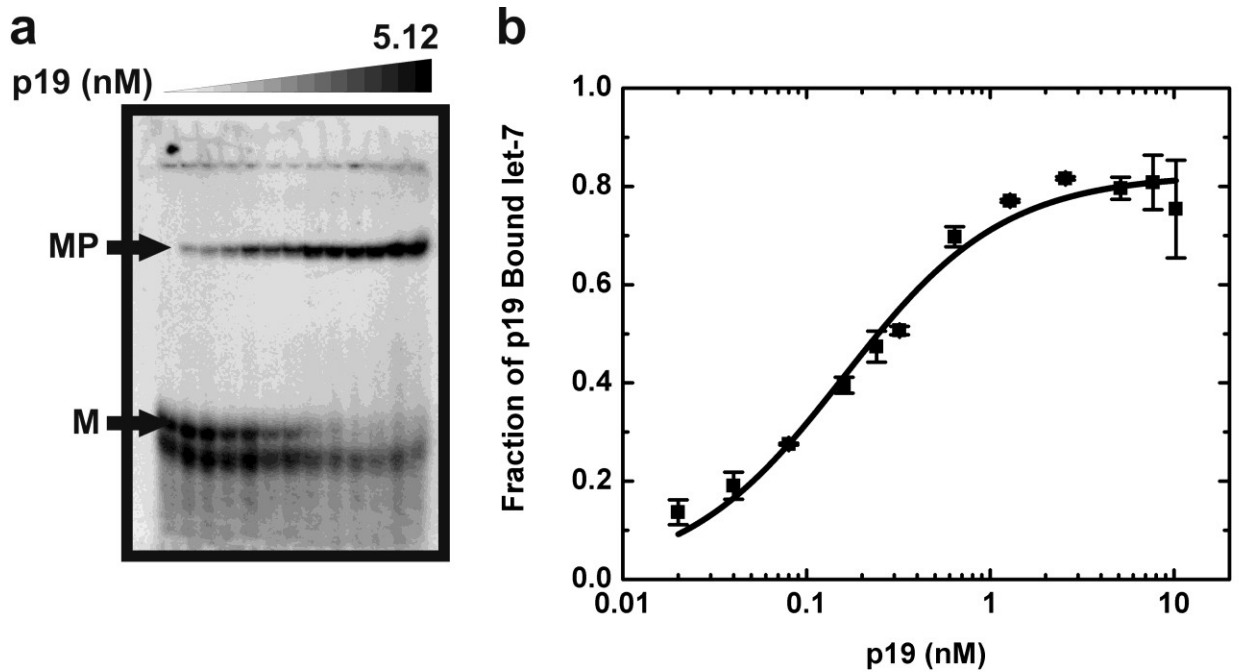
Figure A.6. Complex formation between p19 and miRNA let7a as shown by EMSAs. miRNA (M) labeled with $[\gamma\text{-}^{32}P]$ ATP is bound to p19 (P) to form the complex (MP). **(a)** Representative EMSA. **(b)** Analysis of binding by EMSA yields a p19-miRNA affinity of 0.16 ± 0.02 nM. Reprinted with permission from Elsevier.

Appendix B

siRNA-like double-stranded RNAs are specifically protected against degradation in human cell extract
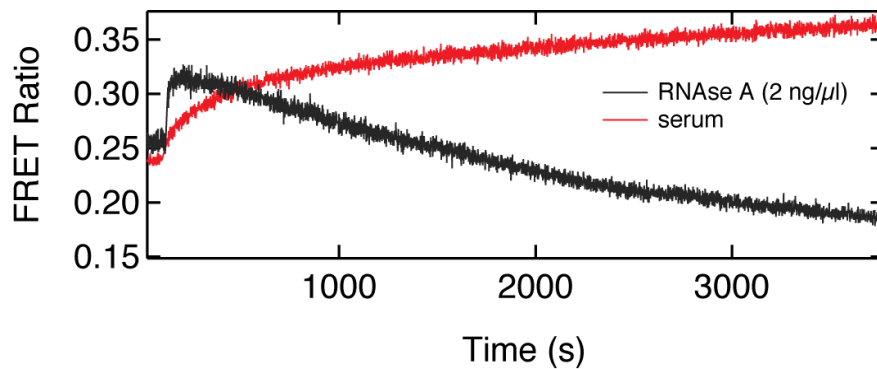
Figure B.1. FRET ratio time trace upon incubating doubly fluorophore labeled 21-nt dsRNA with either serum (red) or 2 ng/ml of purified RNase A (black). The real-time FRET traces for each of these conditions reveal that both serum and the purified nuclease manifest a significant initial increase in FRET ratio. Reprinted with permission from PLOS.

Appendix C

Mass spectrometry analysis of siRNA associated protein complexes obtained from cytosolic HeLa cell extract

C.1 Introduction

Three distinct siRNA-containing complexes, complex D, C1 and C2, were observed during the RISC assembly complex formation assay using cytosolic HeLa cell extract (**Fig 2.5b**). We confirmed the presence of Dicer in complex D using both native western blot analysis and supershift assays (**Fig 2.6 and 2.5c**). To further identify the proteins associated with these siRNA bound complexes we performed Mass spectrometry analysis.

C.2 Materials and Method

The siRNA containing complexes, D, C1 and C2 were formed as described in Chapter 2 using unlabeled siRNAs and resolved on a non-denaturing gel (**Fig 2.5b**). The portion of the gel containing these complexes was excised and the proteins were eluted by passive elution into a buffer containing 100mM sodium acetate; 0.1% SDS; 100mM DTT. The eluted samples were run on a 10% NuPAGE Novex Bis-Tris Gel for about 1.5-2 cm into the gel, stained with Coomassie Brilliant Blue R-250 and destained before submitting it to NextGen Sciences for mass spectrometry analysis (**Fig C.1**). Each sample lane was excised into 10 equal segments using a grid. Each segment was digested using trypsin as described below. Gel digests were analyzed using LC/MS/MS with a
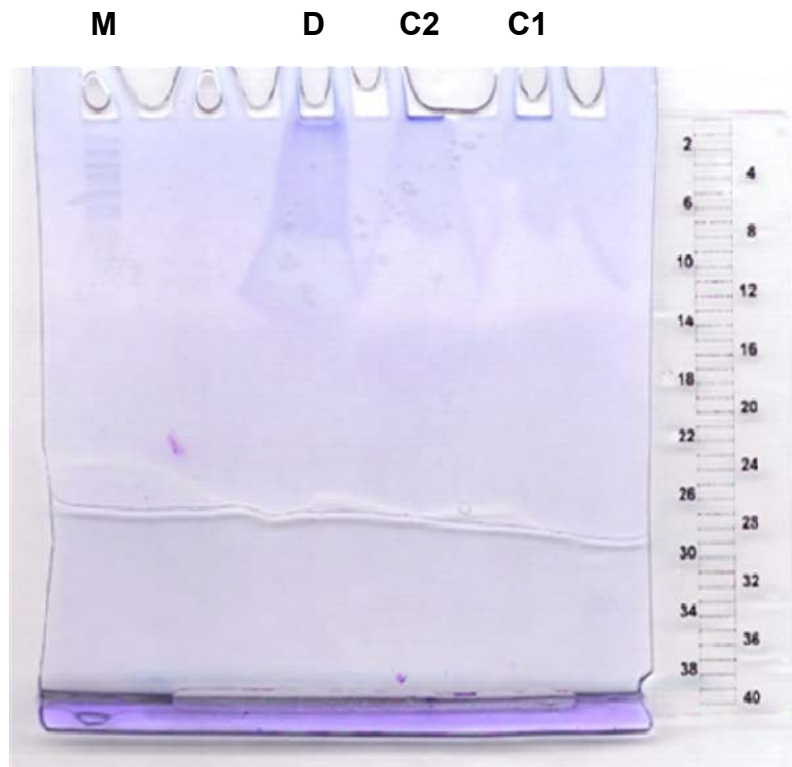
Figure C.1. Denaturing electrophoresis of complexes D, C2 and C1. Denaturing electrophoresis of siRNA associated complexes on a 10% NuPAGE Novex Bis-Tris Gel for about 1.5-2 cm into the gel, stained with Coomassie Brilliant Blue R-250 and destained. Lane M indicates the protein size marker.

1 h gradient on a LTQ Orbitrap XL mass spectrometer. Product ion data were searched against the concatenated forward and reverse IPI Human v3.46 database using the Mascot search engine. The database was appended with commonly observed background proteins (cRAP) to prevent false assignment of peptides derived from those proteins. Mascot output files were parsed into the Scaffold program for collation into non-redundant lists per lane and filtering to assess false discovery rates and allow only correct protein identifications. Spectral counts per protein were output to the Excel summary, these are a semi-quantitative measure of abundance across samples. Spectral count reflects the number of matched peptides and the number of times those peptides were observed.

In-gel digestion

Samples were subjected to proteolytic digestion on a ProGest (Genomic Solutions) workstation where they were reduced with DTT at 60 ºC, allowed to cool to room temperature, alkylated with iodoacetamide and incubated at 37 ºC for 4h in the presence trypsin. Formic acid was added to stop the reaction and the supernatant was analyzed directly.

LC/MS/MS

Samples were analyzed by nano LC/MS/MS on a ThermoFisher LTQ or LTQ Orbitrap XL. 30 µl of hydrolysate was loaded onto a 5 mm, 75 µm ID C12 (Jupiter Proteo, Phenomenex) vented column at a flow-rate of 10µL/min. Gradient elution was over a 15 cm, 75 µm ID C12 column at 300 nL/min. A 1h gradient was employed. The mass spectrometer was operated in data-dependent mode, the six most abundant ions were selected for MS/MS. The Orbitrap MS scan was performed at 60,000 FWHM resolution. MS/MS data were searched using a local copy of Mascot. The parameters for all LC/MS/MS searches were as follows: Type of search : MS/MS Ion Search, Enzyme : Trypsin, Fixed modifications : Carbamidomethyl (C), Variable modifications : Oxidation (M), Acetyl (N-term), Pyro-glu (N-term Q), Mass values : Monoisotopic, Protein Mass : Unrestricted, Peptide Mass Tolerance : ± 5 ppm (Orbitrap); ± 2.0 Da (LTQ), Fragment Mass Tolerance: ± 0.5 Da (LTQ), Max Missed Cleavages : 2

SCAFFOLD

Samples were processed in the Scaffold software using DAT files generated by Mascot. Parameters for LTQ data require a minimum of 3 peptides matching per protein with minimum probabilities of 90% at the protein level and 50-80% at the corresponding peptide level.

C.3 Results

A combined total of 361 proteins were detected (with 2 or more unique peptides/protein) in these experiments (**Table C.1, Table C.2**).

C.4 Conclusions

The mass spectrometry analysis identified 361 proteins from the samples obtained by eluting the siRNA associated complexes from a non-denaturing polyacrylamide gel. We identified few proteins which have been shown to be involved with the RNAi pathway, eg: Heat shock protein 90 (Hsp90), however further investigation is required to confirm the role of these proteins in siRNA associated complexes and interaction if any with the RSS protein, p19.

Parameters	D	C2	C1
Total no. of proteins	105	116	210
Total Spectral Count	6181	5668	5090
Total Unique Peptides	805	976	1447
Mean Spectral Count	58.9	48.9	24.2
Median Spectral Count	25.5	17	9.5
Mean Unique Peptides	7.7	8.4	6.9
Median Unique Peptides	5	5	4
No. of reverse hits	0	0	0
% FDR	0	0	0

Table C.1. Summary of proteins and peptides identified.

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
KRT1 Keratin, type II cytoskeletal 1	IPI00220327 (+1)	66 kDa	484	283	291
KRT10 Keratin, type I cytoskeletal 10	IPI00009865	60 kDa	352	243	236
sp_TRYP_PIG	IPIsp_TRYP_PIG	24 kDa	220	265	233
KRT2 Keratin, type II cytoskeletal 2 epidermal	IPI00021304 (+1)	66 kDa	355	187	207
HSPA8 Isoform 1 of Heat shock cognate 71 kDa protein	IPI00003865	71 kDa	0	64	196
VCL Isoform 1 of Vinculin	IPI00291175 (+1)	117 kDa	0	0	141
KRT9 Keratin, type I cytoskeletal 9	IPI00019359 (+1)	62 kDa	264	133	137
HSPA1B;HSPA1A Heat shock 70 kDa protein 1	IPI00304925	70 kDa	0	0	137
KRT5 Keratin, type II cytoskeletal 5	IPI00009867	62 kDa	227	117	114
TPI1 Isoform 1 of Triosephosphate isomerase	IPI00465028	31 kDa	0	0	114
LDHB L-lactate dehydrogenase B chain	IPI00219217	37 kDa	0	0	113
KRT6B Keratin, type II cytoskeletal 6B	IPI00293665	60 kDa	170	100	100
KRT14 Keratin, type I cytoskeletal 14	IPI00384444	52 kDa	155	84	92
NME2 Nucleoside diphosphate kinase	IPI00604590 (+1)	33 kDa	0	16	83
KRT16 Keratin, type I cytoskeletal 16	IPI00217963	51 kDa	118	0	80
LDHA Isoform 1 of L-lactate dehydrogenase A chain	IPI00217966	37 kDa	0	0	80
PHGDH D-3-phosphoglycerate dehydrogenase	IPI00011200	57 kDa	0	0	76
ACTG1 Actin, cytoplasmic 2	IPI00021440	42 kDa	51	486	75
IQGAP1 Ras GTPase-activating-like protein IQGAP1	IPI00009342	189 kDa	0	0	72
S100A4 Protein S100-A4	IPI00032313	12 kDa	0	0	56
M6PRBP1 Isoform B of Mannose-6-phosphate receptor-binding protein 1	IPI00303882	47 kDa	0	0	56
EEF1A1 Elongation factor 1-alpha 1	IPI00396485 (+1)	50 kDa	14	14	49
PRDX2 Peroxiredoxin-2	IPI00027350	22 kDa	0	0	49

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
NP CDNA FLJ25678 fis, clone TST04067, highly similar to PURINE NUCLEOSIDE PHOSPHORYLASE	IPI00017672 (+1)	33 kDa	0	0	49
TALDO1 Transaldolase	IPI00744692	38 kDa	0	0	48
ATIC Bifunctional purine biosynthesis protein PURH	IPI00289499	65 kDa	0	0	48
STIP1 Stress-induced-phosphoprotein 1	IPI00013894 (+1)	63 kDa	0	0	46
TUBB Tubulin beta chain	IPI00011654 (+1)	50 kDa	139	46	44
VCP Transitional endoplasmic reticulum ATPase	IPI00022774	89 kDa	0	0	44
MDH1 Malate dehydrogenase, cytoplasmic	IPI00291005	36 kDa	0	0	43
TUBB2C Tubulin beta-2C chain	IPI00007752	50 kDa	122	0	41
PFN1 Profilin-1	IPI00216691	15 kDa	0	0	40
KYNU Kynureninase	IPI00003818	52 kDa	0	0	40
G6PD Isoform Long of Glucose-6-phosphate 1-dehydrogenase	IPI00216008 (+1)	64 kDa	0	0	40
PRDX6 Peroxiredoxin-6	IPI00220301	25 kDa	0	0	39
HPRT1 Hypoxanthine-guanine phosphoribosyltransferase	IPI00218493	25 kDa	0	0	38
sp_ALBU_BOVIN	IPIsp_ALBU_BOVIN	69 kDa	67	37	37
GOT1 Aspartate aminotransferase, cytoplasmic	IPI00219029	46 kDa	0	0	36
ACTN1 Alpha-actinin-1	IPI00013508 (+1)	103 kDa	0	0	35
SERPINB1 Leukocyte elastase inhibitor	IPI00027444	43 kDa	0	0	34
GAPDH Glyceraldehyde-3-phosphate dehydrogenase	IPI00219018	36 kDa	9	0	33
SRI Sorcin	IPI00027175 (+1)	22 kDa	0	0	33
GSTO1 Glutathione transferase omega-1	IPI00019755	28 kDa	0	0	33
ALDH9A1 4-trimethylaminobutyraldehyde dehydrogenase	IPI00873817	54 kDa	0	0	33
TUBA1B Tubulin alpha-1B chain	IPI00387144 (+1)	50 kDa	84	23	31
EIF4A1 Eukaryotic initiation factor 4A-I	IPI00025491 (+1)	46 kDa	0	136	31
sp_CAS1_BOVIN	IPIsp_CAS1_BOVIN	25 kDa	38	32	30

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
MCM2 DNA replication licensing factor MCM2	IPI00184330	102 kDa	0	4	29
PSMB6 Proteasome subunit beta type-6 precursor	IPI00000811 (+1)	25 kDa	0	0	28
IMPDH2 Inosine-5'-monophosphate dehydrogenase 2	IPI00291510	56 kDa	0	0	28
KRT77 Keratin 77	IPI00376379	62 kDa	0	0	27
CAND1 Isoform 1 of Cullin-associated NEDD8-dissociated protein 1	IPI00100160 (+1)	136 kDa	0	0	26
MCM7 Isoform 1 of DNA replication licensing factor MCM7	IPI00299904	81 kDa	0	0	23
CCT2 T-complex protein 1 subunit beta	IPI00297779	57 kDa	0	0	22
ACTN4 Alpha-actinin-4	IPI00013808	105 kDa	0	0	22
HSP90AB1 Heat shock protein HSP 90-beta	IPI00414676	83 kDa	76	450	21
RUVBL2 RuvB-like 2	IPI00009104	51 kDa	0	0	20
HSP90AA1 heat shock protein 90kDa alpha (cytosolic), class A member 1 isoform 1	IPI00382470 (+1)	98 kDa	0	450	19
STMN1 Stathmin	IPI00479997	17 kDa	0	0	19
PSMA6 Proteasome subunit alpha type-6	IPI00029623	27 kDa	0	0	19
CAP1 Adenylyl cyclase-associated protein 1	IPI00008274 (+2)	52 kDa	0	0	19
EEF2 Elongation factor 2	IPI00186290	95 kDa	0	0	18
EEF1G Elongation factor 1-gamma	IPI00000875 (+1)	50 kDa	0	0	18
CAPG Macrophage-capping protein	IPI00027341 (+1)	39 kDa	0	0	18
GARS 85 kDa protein	IPI00465260 (+1)	85 kDa	0	0	18
TARS Threonyl-tRNA synthetase, cytoplasmic	IPI00329633	83 kDa	0	0	18
RNH1 Ribonuclease inhibitor	IPI00550069	50 kDa	675	60	17
HSPB1 Heat shock protein beta-1	IPI00025512	23 kDa	0	0	17
DDT D-dopachrome decarboxylase	IPI00293867	13 kDa	0	0	17
PSME2 Uncharacterized protein PSME2	IPI00384051 (+1)	29 kDa	0	0	17
MSN Moesin	IPI00219365 (+1)	68 kDa	0	26	16

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
ACAT2 Acetyl-CoA acetyltransferase, cytosolic	IPI00291419	41 kDa	0	0	16
ARF1 ADP-ribosylation factor 1	IPI00215914 (+1)	21 kDa	0	0	16
MCM3 DNA replication licensing factor MCM3	IPI00013214	91 kDa	0	3	15
PGD 6-phosphogluconate dehydrogenase, decarboxylating	IPI00219525 (+1)	53 kDa	0	0	15
DUT Isoform DUT-M of Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial precursor	IPI00013679 (+3)	27 kDa	0	0	14
GNPDA1 Glucosamine-6-phosphate isomerase 1	IPI00009305	33 kDa	0	0	14
TXNRD1 Isoform 5 of Thioredoxin reductase 1, cytoplasmic	IPI00554786 (+4)	55 kDa	0	0	14
RPS27A;UBC;UBB ubiquitin and ribosomal protein S27a precursor	IPI00179330 (+22)	18 kDa	11	7	13
DCD Dermcidin precursor	IPI00027547 (+1)	11 kDa	0	12	13
ACP1 Isoform 1 of Low molecular weight phosphotyrosine protein phosphatase	IPI00219861 (+2)	18 kDa	0	0	13
PSMA3 Isoform 2 of Proteasome subunit alpha type-3	IPI00171199 (+1)	28 kDa	0	0	13
USO1 Putative uncharacterized protein DKFZp451D234	IPI00031583	109 kDa	0	0	13
TRIM28 Isoform 1 of Transcription intermediary factor 1-beta	IPI00438229 (+1)	89 kDa	0	0	13
USP9X ubiquitin specific protease 9, X-linked isoform 4	IPI00003964 (+1)	290 kDa	0	0	13
ALB Isoform 1 of Serum albumin precursor	IPI00745872 (+1)	69 kDa	27	17	12
sp_CASK_BOVIN	IPIsp_CASK_BOVIN	21 kDa	9	10	12
#NAME?	IPI00453476 (+1)	29 kDa	0	2	12
MCM6 DNA replication licensing factor MCM6	IPI00031517	93 kDa	0	0	12
PSMA1 Isoform Short of Proteasome subunit alpha type-1	IPI00016832 (+2)	30 kDa	0	0	12
RUVBL1 Isoform 1 of RuvB-like 1	IPI00021187 (+1)	50 kDa	0	0	12
PSME3 Isoform 1 of Proteasome activator complex subunit 3	IPI00030243 (+1)	30 kDa	0	0	11
TMSB10 Thymosin beta-10	IPI00220827	5 kDa	0	0	11
ADSS Adenylosuccinate synthetase isozyme 2	IPI00026833	50 kDa	0	0	11

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
MTHFD1 C-1-tetrahydrofolate synthase, cytoplasmic	IPI00218342 (+1)	102 kDa	0	0	11
GART Isoform Long of Trifunctional purine biosynthetic protein adenosine-3	IPI00025273	108 kDa	0	0	11
KIAA0664 Putative eukaryotic translation initiation factor 3 subunit	IPI00024425	147 kDa	0	0	11
PDCD6 Programmed cell death protein 6	IPI00025277	22 kDa	0	0	10
ADSL Isoform 1 of Adenylosuccinate lyase	IPI00026904 (+1)	55 kDa	0	0	10
- 22 kDa protein	IPI00219910 (+1)	22 kDa	0	0	10
CBR1 Carbonyl reductase [NADPH] 1	IPI00295386	30 kDa	0	0	10
PRMT5 Protein arginine N-methyltransferase 5	IPI00441473	73 kDa	0	0	10
PSME1 Proteasome activator complex subunit 1	IPI00479722	29 kDa	0	0	10
MYL6;MYL6B Isoform Non-muscle of Myosin light polypeptide 6	IPI00335168 (+2)	17 kDa	0	0	10
ARF4 ADP-ribosylation factor 4	IPI00215918	21 kDa	0	0	10
CAPZA1 F-actin-capping protein subunit alpha-1	IPI00005969 (+1)	33 kDa	0	0	9
TPD52L2 Isoform 2 of Tumor protein D54	IPI00221178 (+6)	20 kDa	0	0	9
CAPZB Isoform 1 of F-actin-capping protein subunit beta	IPI00026185 (+2)	31 kDa	0	0	9
TPD52 tumor protein D52 isoform 2	IPI00619951 (+4)	22 kDa	0	0	9
COTL1 Coactosin-like protein	IPI00017704	16 kDa	0	0	9
AHCY Adenosylhomocysteinase	IPI00012007	48 kDa	0	0	9
VPS35 Vacuolar protein sorting-associated protein 35	IPI00018931	92 kDa	0	0	9
WARS Tryptophanyl-tRNA synthetase, cytoplasmic	IPI00295400 (+1)	53 kDa	0	0	9
MCM5 DNA replication licensing factor MCM5	IPI00018350 (+1)	82 kDa	0	0	9
PIP Prolactin-inducible protein precursor	IPI00022974	17 kDa	17	8	8
sp_CASB_BOVIN	IPIsp_CASB_BOVIN	25 kDa	0	0	8
PSMD5 26S proteasome non-ATPase regulatory subunit 5	IPI00002134	56 kDa	0	0	8
LAP3 Isoform 1 of Cytosol aminopeptidase	IPI00419237 (+1)	56 kDa	0	0	8

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
IDH1 Isocitrate dehydrogenase [NADP] cytoplasmic	IPI00027223	47 kDa	0	0	8
DPYSL2 Dihydropyrimidinase-related protein 2	IPI00257508	62 kDa	0	0	8
PSMA7 Isoform 1 of Proteasome subunit alpha type-7	IPI00024175	28 kDa	0	0	8
PSMB5 29 kDa protein	IPI00375704 (+1)	29 kDa	0	0	8
MCM4 DNA replication licensing factor MCM4	IPI00018349	97 kDa	0	0	8
LYZ Lysozyme C precursor	IPI00019038 (+1)	17 kDa	23	0	7
PSMA5 Proteasome subunit alpha type-5	IPI00291922	26 kDa	0	0	7
RPLP0 60S acidic ribosomal protein P0	IPI00008530 (+2)	34 kDa	0	0	7
PSMD11 Proteasome 26S non-ATPase subunit 11 variant (Fragment)	IPI00105598	48 kDa	0	0	7
FH Isoform Mitochondrial of Fumarate hydratase, mitochondrial precursor	IPI00296053 (+1)	55 kDa	0	0	7
PARK7 Protein DJ-1	IPI00298547	20 kDa	0	0	7
STAT3 Isoform Del-701 of Signal transducer and activator of transcription 3	IPI00306436 (+2)	88 kDa	0	0	7
CKM Creatine kinase M-type	IPI00027487 (+1)	43 kDa	0	0	7
ESD S-formylglutathione hydrolase	IPI00411706 (+1)	31 kDa	0	0	7
NASP Isoform 1 of Nuclear autoantigenic sperm protein	IPI00179953 (+1)	85 kDa	71	0	6
sp_CAS2_BOVIN	IPIsp_CAS2_BOVIN	26 kDa	11	9	6
PGM1 Isoform 1 of Phosphoglucomutase-1	IPI00219526 (+1)	61 kDa	0	0	6
COPS4 COP9 signalosome complex subunit 4	IPI00171844	46 kDa	0	0	6
PSMA4 Proteasome subunit alpha type-4	IPI00299155 (+1)	29 kDa	0	0	6
MAPK1 Mitogen-activated protein kinase 1	IPI00003479	41 kDa	0	0	6
HINT1 Histidine triad nucleotide-binding protein 1	IPI00239077	14 kDa	0	0	6
XPO7 Exportin-7	IPI00302458 (+2)	124 kDa	0	0	6
ADK Isoform Short of Adenosine kinase	IPI00234368 (+1)	39 kDa	0	0	6

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
PSMA2 Proteasome subunit alpha type-2	IPI00219622	26 kDa	0	0	5
AKR1B1 Aldose reductase	IPI00413641 (+1)	36 kDa	0	0	5
NUP93 Nuclear pore complex protein Nup93	IPI00397904	93 kDa	0	0	5
CPNE3 Copine-3	IPI00024403	60 kDa	0	0	5
CMBL Carboxymethylenebutenolidase homolog	IPI00383046	28 kDa	0	0	5
MAP2K1 Dual specificity mitogen-activated protein kinase kinase 1	IPI00219604	43 kDa	0	0	5
FAM49B Protein FAM49B	IPI00303318 (+1)	37 kDa	0	0	5
MAPK3 Mitogen-activated protein kinase 3	IPI00018195 (+1)	43 kDa	0	0	5
FLG2 Filaggrin-2	IPI00397801	248 kDa	7	0	4
KPRP Keratinocyte proline-rich protein	IPI00514908	64 kDa	0	0	4
PGM2 Phosphoglucomutase-2	IPI00550364	68 kDa	0	0	4
IMPA2 Isoform 1 of Inositol monophosphatase 2	IPI00023635 (+1)	31 kDa	0	0	4
CARM1 Isoform 1 of Histone-arginine methyltransferase CARM1	IPI00412880 (+2)	63 kDa	0	0	4
KIAA1598 Isoform 3 of Shootin-1	IPI00448751 (+4)	74 kDa	0	0	4
PCBP2 poly(rC)-binding protein 2 isoform b	IPI00012066 (+3)	38 kDa	0	0	4
PFKM cDNA FLJ44241 fis, clone THYMU3008436, highly similar to 6-phosphofructokinase, muscle type	IPI00465179 (+1)	93 kDa	0	0	4
GPS1 G protein pathway suppressor 1 isoform 2	IPI00156282 (+1)	56 kDa	0	0	4
GEMIN5 Gem-associated protein 5	IPI00291783	169 kDa	0	0	4
DSP Isoform DPI of Desmoplakin	IPI00013933 (+1)	332 kDa	3	0	3
KPNB1 Importin subunit beta-1	IPI00001639	97 kDa	0	93	3
STAT5A Signal transducer and activator of transcription 5A	IPI00030783 (+1)	91 kDa	0	0	3
CUL4A Isoform 1 of Cullin-4A	IPI00419273	88 kDa	0	0	3
ME1 NADP-dependent malic enzyme	IPI00008215	64 kDa	0	0	3
COPS5 COP9 signalosome complex subunit 5	IPI00009958 (+1)	38 kDa	0	0	3

147

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
NAMPT Isoform 1 of Nicotinamide phosphoribosyltransferase	IPI00018873 (+1)	56 kDa	0	0	3
ACOT7 Isoform 1 of Cytosolic acyl coenzyme A thioester hydrolase	IPI00010415 (+6)	42 kDa	0	0	3
ACO1 Cytoplasmic aconitate hydratase	IPI00008485	98 kDa	0	0	3
EFHD2 EF-hand domain-containing protein D2	IPI00060181	27 kDa	0	0	3
NUDC Nuclear migration protein nudC	IPI00550746 (+1)	38 kDa	0	0	3
WDR1 Isoform 2 of WD repeat-containing protein 1	IPI00216256 (+3)	58 kDa	0	0	3
PSMD2 26S proteasome non-ATPase regulatory subunit 2	IPI00012268	100 kDa	0	0	3
COASY Bifunctional coenzyme A synthase	IPI00184821	62 kDa	0	0	3
SPAG9 145 kDa protein	IPI00218097 (+6)	145 kDa	0	0	3
PGP Phosphoglycolate phosphatase	IPI00177008	34 kDa	0	0	3
NAP1L1 Nucleosome assembly protein 1-like 1	IPI00023860 (+1)	45 kDa	30	0	2
AARS Alanyl-tRNA synthetase, cytoplasmic	IPI00027442	107 kDa	0	48	2
C1orf68;LOC100129271 Skin-specific protein 32	IPI00023078	26 kDa	0	0	2
SLK Isoform 1 of STE20-like serine/threonine-protein kinase	IPI00022827 (+1)	143 kDa	0	0	2
EEF1B2 Elongation factor 1-beta	IPI00178440 (+1)	25 kDa	0	0	2
MGEA5 Isoform 2 of Bifunctional protein NCOAT	IPI00181391 (+2)	95 kDa	0	0	2
HSPA4L Heat shock 70 kDa protein 4L	IPI00295485 (+1)	94 kDa	0	0	2
PSMB2 Proteasome subunit beta type-2	IPI00028006 (+1)	23 kDa	0	0	2
NIT2 Nitrilase homolog 2	IPI00549467 (+1)	31 kDa	0	0	2
NCDN Isoform 1 of Neurochondrin	IPI00549543 (+2)	79 kDa	0	0	2
COPS2 Isoform 2 of COP9 signalosome complex subunit 2	IPI00018813 (+1)	52 kDa	0	0	2
PSMB4 Proteasome subunit beta type-4 precursor	IPI00555956	29 kDa	0	0	2
UMPS Isoform 1 of Uridine 5'-monophosphate synthase	IPI00003923	52 kDa	0	0	2
RABGAP1 Isoform 1 of Rab GTPase-activating protein 1	IPI00016702 (+1)	122 kDa	0	0	2

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
ACLY ATP-citrate synthase	IPI00021290 (+1)	121 kDa	0	0	2
PRKAR1A cAMP-dependent protein kinase type I-alpha regulatory subunit	IPI00021831	43 kDa	0	0	2
BLVRA Biliverdin reductase A precursor	IPI00294158	33 kDa	0	0	2
NANS Sialic acid synthase	IPI00147874	40 kDa	0	0	2
AHCYL2 Putative adenosylhomocysteinase 3	IPI00101645 (+5)	67 kDa	0	0	2
ATP6V1B2 Vacuolar ATP synthase subunit B, brain isoform	IPI00007812 (+1)	57 kDa	0	0	2
GALK1 Galactokinase	IPI00019383 (+2)	42 kDa	0	0	2
CPNE1 Copine-1	IPI00018452 (+8)	59 kDa	0	0	2
ANKFY1 ankyrin repeat and FYVE domain containing 1 isoform 2	IPI00375301	55 kDa	0	0	2
SPRR2G Small proline-rich protein 2G	IPI00000849 (+4)	8 kDa	0	0	2
STAT1 Isoform Alpha of Signal transducer and activator of transcription 1-alpha/beta	IPI00030781 (+2)	87 kDa	0	0	2
PSMC3 26S protease regulatory subunit 6A	IPI00018398	49 kDa	0	0	2
USP14 Ubiquitin carboxyl-terminal hydrolase 14	IPI00219913 (+3)	56 kDa	0	0	2
RELA Isoform 2 of Transcription factor p65	IPI00219084 (+4)	59 kDa	0	0	2
IDI1 14 kDa protein	IPI00645307	14 kDa	0	0	2
APOA1 Apolipoprotein A1	IPI00853525 (+1)	28 kDa	0	0	2
PDIA3 Protein disulfide-isomerase A3 precursor	IPI00025252	57 kDa	0	0	2
GSR Isoform Mitochondrial of Glutathione reductase, mitochondrial precursor	IPI00016862 (+1)	56 kDa	0	0	2
PITPNA Phosphatidylinositol transfer protein alpha isoform	IPI00216048	32 kDa	0	0	2
GRHPR Glyoxylate reductase/hydroxypyruvate reductase	IPI00037448 (+2)	36 kDa	0	0	2
PYGL Glycogen phosphorylase, liver form	IPI00783313	97 kDa	0	0	2
ANXA5 Annexin A5	IPI00329801 (+2)	36 kDa	151	0	0
NSUN2 tRNA	IPI00306369	86 kDa	111	0	0

149

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
KRT17 Keratin, type I cytoskeletal 17	IPI00450768	48 kDa	103	0	0
PLA2G4A Cytosolic phospholipase A2	IPI00026108	85 kDa	91	0	0
CLIC1 Chloride intracellular channel protein 1	IPI00010896	27 kDa	79	17	0
NCL Isoform 1 of Nucleolin	IPI00604620 (+1)	77 kDa	77	0	0
GDA Guanine deaminase	IPI00465184 (+1)	51 kDa	72	0	0
KRT73 Isoform 1 of Keratin, type II cytoskeletal 73	IPI00174775	59 kDa	72	0	0
NRD1 nardilysin (N-arginine dibasic convertase) isoform a	IPI00243221 (+2)	139 kDa	67	0	0
TPT1 Translationally-controlled tumor protein	IPI00550900	20 kDa	65	0	0
sp_K1M1_SHEEP	IPIsp_K1M1_SHEEP	47 kDa	64	0	0
KRT13 Isoform 1 of Keratin, type I cytoskeletal 13	IPI00009866	50 kDa	63	0	0
BBS1;DPP3 Isoform 1 of Dipeptidyl-peptidase 3	IPI00020672	83 kDa	62	0	0
APEH Acylamino-acid-releasing enzyme	IPI00337741	81 kDa	54	0	0
HAT1 Histone acetyltransferase type B catalytic subunit	IPI00024719	50 kDa	49	0	0
USP5 Isoform Long of Ubiquitin carboxyl-terminal hydrolase 5	IPI00024664 (+1)	96 kDa	49	0	0
KRT71 Keratin, type II cytoskeletal 71	IPI00061200	57 kDa	48	0	0
KRT83 Keratin type II cuticular Hb3	IPI00297795 (+2)	54 kDa	44	0	0
CAPN2 Calpain-2 catalytic subunit precursor	IPI00289758 (+1)	80 kDa	43	0	0
TXN Thioredoxin	IPI00216298 (+1)	12 kDa	42	0	0
FDPS Farnesyl diphosphate synthase	IPI00101405	48 kDa	40	0	0
GLO1 Lactoylglutathione lyase	IPI00220766	21 kDa	38	0	0
CAPNS1 Calpain small subunit 1	IPI00025084	28 kDa	37	0	0
CKB Creatine kinase B-type	IPI00022977	43 kDa	36	0	0
YWHAB Isoform Short of 14-3-3 protein beta/alpha	IPI00759832	28 kDa	34	0	0
KRT4 keratin 4	IPI00290078	64 kDa	33	0	0

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
LOC389842 similar to RanBP1	IPI00399212 (+3)	23 kDa	32	0	0
PPP2R1A Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	IPI00554737	65 kDa	31	13	0
RBBP7 Histone-binding protein RBBP7	IPI00395865	48 kDa	30	0	0
SET Isoform 1 of Protein SET	IPI00072377 (+4)	33 kDa	28	0	0
NACA Nascent polypeptide-associated complex subunit alpha	IPI00023748 (+2)	23 kDa	27	3	0
IPO7 Importin-7	IPI00007402	120 kDa	26	51	0
NUDT5 ADP-sugar pyrophosphatase	IPI00296913 (+1)	24 kDa	25	0	0
KRT27 Keratin, type I cytoskeletal 27	IPI00328103	50 kDa	25	0	0
PBK Lymphokine-activated killer T-cell-originated protein kinase	IPI00306708	36 kDa	23	0	0
PUS7 Pseudouridylate synthase 7 homolog	IPI00044761	75 kDa	23	0	0
C6orf108 c-Myc-responsive protein Rcl	IPI00007926	19 kDa	22	0	0
LACRT Extracellular glycoprotein lacritin precursor	IPI00020487	14 kDa	22	0	0
S100A8 Protein S100-A8	IPI00007047	11 kDa	20	0	0
GLRX3 Glutaredoxin-3	IPI00008552	37 kDa	19	0	0
RAN GTP-binding nuclear protein Ran	IPI00643041 (+1)	24 kDa	18	10	0
SMS Isoform 1 of Spermine synthase	IPI00005102	41 kDa	18	0	0
YWHAQ 14-3-3 protein theta	IPI00018146	28 kDa	18	0	0
SSB Lupus La protein	IPI00009032	47 kDa	17	0	0
NAP1L4 Nucleosome assembly protein 1-like 4	IPI00017763 (+2)	43 kDa	17	0	0
ASNA1 Arsenical pump-driving ATPase	IPI00013466	39 kDa	16	0	0
GSTM3 Glutathione S-transferase Mu 3	IPI00246975	27 kDa	16	0	0
C12orf10 UPF0160 protein MYG1	IPI00029444 (+2)	42 kDa	16	0	0
C7orf24 Uncharacterized protein C7orf24	IPI00031564	21 kDa	15	0	0
OTUB1 Putative uncharacterized protein DKFZp564E242	IPI00000581 (+2)	31 kDa	15	0	0

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
P4HB Protein disulfide-isomerase precursor	IPI00010796	57 kDa	15	0	0
YWHAZ 14-3-3 protein zeta/delta	IPI00021263	28 kDa	14	0	0
IGHV3OR16-13;IGHA1 IGHA1 protein	IPI00061977 (+11)	54 kDa	13	0	0
CUTA Isoform A of Protein CutA precursor	IPI00034319 (+2)	21 kDa	11	0	0
TGM2 Isoform 1 of Protein-glutamine gamma-glutamyltransferase 2	IPI00294578 (+1)	77 kDa	11	0	0
COL1A1 Collagen alpha-1(I) chain precursor	IPI00297646	139 kDa	11	0	0
BTF3 Isoform 1 of Transcription factor BTF3	IPI00221035 (+1)	22 kDa	11	0	0
YARS Tyrosyl-tRNA synthetase, cytoplasmic	IPI00007074	59 kDa	10	23	0
CHM Rab proteins geranylgeranyltransferase component A 1	IPI00028099	73 kDa	10	0	0
TXNDC17 Thioredoxin domain-containing protein 17	IPI00646689	14 kDa	9	0	0
PPM1G Protein phosphatase 1G	IPI00006167	59 kDa	9	0	0
TROVE2 Isoform Long of 60 kDa SS-A/Ro ribonucleoprotein	IPI00019450 (+3)	61 kDa	9	0	0
LCN1 Lipocalin-1 precursor	IPI00009650	19 kDa	9	0	0
LTF Growth-inhibiting protein 12	IPI00298860 (+3)	78 kDa	8	0	0
STRAP Serine-threonine kinase receptor-associated protein	IPI00294536	38 kDa	8	0	0
HDGF Hepatoma-derived growth factor	IPI00020956	27 kDa	8	0	0
THUMPD1 THUMP domain-containing protein 1	IPI00550243	39 kDa	7	0	0
SGTA Small glutamine-rich tetratricopeptide repeat-containing protein alpha	IPI00013949	34 kDa	7	0	0
TXNDC5;MUTED Thioredoxin domain-containing protein 5 precursor	IPI00171438 (+1)	48 kDa	7	0	0
DEFA1;LOC728358 Neutrophil defensin 1 precursor	IPI00005721 (+1)	10 kDa	5	2	0
BCCIP Isoform 1 of BRCA2 and CDKN1A-interacting protein	IPI00002203 (+3)	36 kDa	5	0	0
RAD23A UV excision repair protein RAD23 homolog A	IPI00008219	40 kDa	4	0	0
HDDC2 Isoform 2 of HD domain-containing protein 2	IPI00386751 (+3)	20 kDa	4	0	0
FHL3 Four and a half LIM domains protein 3	IPI00014399	31 kDa	2	0	0

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
FTO Isoform 1 of Protein fatso	IPI00028277 (+1)	58 kDa	2	0	0
UBA1 Ubiquitin-like modifier-activating enzyme 1	IPI00645078	118 kDa	0	181	0
CSE1L Isoform 1 of Exportin-2	IPI00022744 (+1)	110 kDa	0	171	0
TPM3 Isoform 2 of Tropomyosin alpha-3 chain	IPI00218319	29 kDa	0	167	0
HSP90AB1 Heat shock protein beta	IPI00411633	17 kDa	0	101	0
IPO5 importin 5	IPI00793443	126 kDa	0	84	0
TPM4 Isoform 1 of Tropomyosin alpha-4 chain	IPI00010779	29 kDa	0	79	0
CCT8 59 kDa protein	IPI00302925 (+1)	59 kDa	0	70	0
PFAS Phosphoribosylformylglycinamidine synthase	IPI00004534	145 kDa	0	67	0
TPM2 Tropomyosin 2	IPI00646748	33 kDa	0	65	0
TNPO1 transportin 1 isoform 1	IPI00024364 (+1)	102 kDa	0	64	0
RPSA 33 kDa protein	IPI00413108 (+1)	33 kDa	0	59	0
XPOT Exportin-T	IPI00306290	110 kDa	0	50	0
GSS Glutathione synthetase	IPI00010706	52 kDa	0	48	0
ANXA6 annexin VI isoform 2	IPI00002459 (+1)	75 kDa	0	42	0
RNPEP Uncharacterized protein RNPEP	IPI00647400	68 kDa	0	41	0
HSP90B1 Endoplasmin precursor	IPI00027230	92 kDa	0	38	0
SERPINB6 Putative uncharacterized protein DKFZp686I04222	IPI00413451 (+1)	46 kDa	0	35	0
SRM Spermidine synthase	IPI00292020	34 kDa	0	32	0
ANXA3 Annexin A3	IPI00024095	36 kDa	0	32	0
MAT2A S-adenosylmethionine synthetase isoform type-2	IPI00010157	44 kDa	0	31	0
ST13 Hsc70-interacting protein	IPI00032826 (+3)	41 kDa	0	29	0
CTPS CTP synthase 1	IPI00290142	67 kDa	0	28	0
PGLS 6-phosphogluconolactonase	IPI00029997	28 kDa	0	28	0

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
ATP6V1A Vacuolar ATP synthase catalytic subunit A	IPI00007682	68 kDa	0	27	0
FKBP4 FK506-binding protein 4	IPI00219005	52 kDa	0	25	0
PPA1 Inorganic pyrophosphatase	IPI00015018	33 kDa	0	25	0
THOP1 Thimet oligopeptidase	IPI00549189 (+1)	79 kDa	0	24	0
BAT1 Isoform 2 of Spliceosome RNA helicase BAT1	IPI00641829 (+1)	51 kDa	0	22	0
NPEPPS Puromycin-sensitive aminopeptidase	IPI00026216	103 kDa	0	21	0
ABHD14B Isoform 1 of Abhydrolase domain-containing protein 14B	IPI00063827	22 kDa	0	20	0
GANAB Isoform 2 of Neutral alpha-glucosidase AB precursor	IPI00011454 (+1)	109 kDa	0	20	0
RPS21 40S ribosomal protein S21	IPI00017448	9 kDa	0	19	0
IPO9 Importin-9	IPI00185146	116 kDa	0	18	0
UAP1 Isoform AGX2 of UDP-N-acetylhexosamine pyrophosphorylase	IPI00000684 (+3)	59 kDa	0	16	0
HMGCS1 Hydroxymethylglutaryl-CoA synthase, cytoplasmic	IPI00008475	57 kDa	0	16	0
CDC37 Hsp90 co-chaperone Cdc37	IPI00013122	44 kDa	0	15	0
HSPA5 HSPA5 protein	IPI00003362	72 kDa	0	15	0
IPO4 Isoform 1 of Importin-4	IPI00156374 (+1)	119 kDa	0	14	0
HSPBP1 Isoform 1 of Hsp70-binding protein 1	IPI00100748	39 kDa	0	14	0
BPNT1 Isoform 1 of 3'(2'),5'-bisphosphate nucleotidase 1	IPI00410214	33 kDa	0	13	0
PRKCSH Glucosidase 2 subunit beta precursor	IPI00026154 (+1)	59 kDa	0	13	0
ARHGDIA Rho GDP-dissociation inhibitor 1	IPI00003815 (+3)	23 kDa	0	12	0
S100A6 Protein S100-A6	IPI00027463	10 kDa	0	11	0
UBA3 NEDD8-activating enzyme E1 catalytic subunit	IPI00328154	52 kDa	0	11	0
HNRNPF Heterogeneous nuclear ribonucleoprotein F	IPI00003881	46 kDa	0	10	0
TBCA Tubulin-specific chaperone A	IPI00217236 (+1)	13 kDa	0	10	0
APRT Adenine phosphoribosyltransferase	IPI00218693 (+1)	20 kDa	0	9	0

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
XRCC5 ATP-dependent DNA helicase 2 subunit 2	IPI00220834 (+1)	83 kDa	0	9	0
ENO2 Gamma-enolase	IPI00216171	47 kDa	0	8	0
PRMT1 HMT1 hnRNP methyltransferase-like 2 isoform 1	IPI00018522 (+2)	42 kDa	0	8	0
FAF1 Isoform Long of FAS-associated factor 1	IPI00070643 (+2)	74 kDa	0	8	0
HK2 Hexokinase-2	IPI00102864	102 kDa	0	8	0
APOD Apolipoprotein D precursor	IPI00006662	21 kDa	0	7	0
PFDN1 Prefoldin subunit 1	IPI00000051	14 kDa	0	7	0
MGC3207 Isoform 1 of Translation initiation factor eIF-2B subunit alpha/beta/delta-like protein	IPI00005948 (+1)	39 kDa	0	7	0
GLOD4 Uncharacterized protein C17orf25	IPI00007102 (+2)	55 kDa	0	7	0
PDXK Isoform 1 of Pyridoxal kinase	IPI00013004	35 kDa	0	7	0
RHOC Rho-related GTP-binding protein RhoC precursor	IPI00027434 (+3)	22 kDa	0	6	0
MAT2B Isoform 1 of Methionine adenosyltransferase 2 subunit beta	IPI00002324 (+2)	38 kDa	0	6	0
TBCD Isoform 4 of Tubulin-specific chaperone D	IPI00030774 (+1)	139 kDa	0	6	0
PDIA6 Isoform 2 of Protein disulfide-isomerase A6 precursor	IPI00299571 (+1)	54 kDa	0	6	0
GBE1 1,4-alpha-glucan-branching enzyme	IPI00296635	80 kDa	0	5	0
NDRG1 Protein NDRG1	IPI00022078	43 kDa	0	4	0
USP47 Isoform 2 of Ubiquitin carboxyl-terminal hydrolase 47	IPI00165528 (+1)	147 kDa	0	3	0
PEPD Xaa-Pro dipeptidase	IPI00257882 (+2)	55 kDa	0	3	0
HIST1H1C Histone H1.2	IPI00217465 (+2)	21 kDa	0	3	0
NSFL1C Isoform 2 of NSFL1 cofactor p47	IPI00022830 (+3)	37 kDa	0	3	0
UBA2 SUMO-activating enzyme subunit 2	IPI00023234	71 kDa	0	3	0
XPO5 Isoform 1 of Exportin-5	IPI00640703	136 kDa	0	3	0
POLR1C Isoform 1 of DNA-directed RNA polymerases I and III subunit RPAC1	IPI00005179 (+1)	39 kDa	0	3	0

Protein Name	Accession Number	Molecular Weight	D Spectral Count	C2 Spectral Count	C1 Spectral Count
EIF2S1 Eukaryotic translation initiation factor 2 subunit 1	IPI00219678	36 kDa	0	2	0
PPP2R2A Serine/threonine-protein phosphatase 2A 55 kDa regulatory subunit B alpha isoform	IPI00332511	52 kDa	0	2	0
GCLC Glutamate--cysteine ligase catalytic subunit	IPI00215768	73 kDa	0	2	0
PLS3 plastin 3	IPI00216694	71 kDa	0	2	0
APOA1BP Isoform 1 of Apolipoprotein A-I-binding protein precursor	IPI00168479 (+1)	32 kDa	0	2	0
TNPO3 Isoform 2 of Transportin-3	IPI00395694 (+3)	104 kDa	0	2	0

Table C.2. List of identified proteins and spectral counts observed in each of the complexes.

Appendix D

Visualizing the binding of target mRNA to siRNA-containing complexes and the subsequent cleavage of the target at the single molecule level

D.1 Introduction

In order to gain mechanistic insight into how the siRNA-containing complexes unwind siRNA duplexes, bind their target mRNA and cleave it, single-molecule fluorescence microscopy was be used. Observing these processes at the single-molecule level will shed light on the heterogeneity associated with these RNAi processes and will delineate their kinetics.

D.2 Materials and Method

Nucleic acid synthesis. siRNAs were synthesized by the HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory at the Yale University School of Medicine, having the same sequence as indicated in **Section 4.2** with a 5' phosphate (P) and a 3' amino group on a C7 carbon linker. RNA was purified as described and the 3' amine groups were labeled with Cy3 or Cy5 succinimidyl ester as described in **Section 3.2**. The target luciferase mRNA was synthesized as described in **Section 4.2**.

Periodate Oxidation of the 3' Terminus of target mRNA and reaction with fluorophore or biotin. The reaction was performed as described in the Handbook of RNA Biochemistry. 20nmol of RNA was incubated in a volume of 400 uL containing 2.5 mM NaIO4, 100 mM NaOAc (pH 5.0) for 50 min on ice. The reaction was then ethanol precipitaed. For the coupling reaction with

fluorophore or biotin, the RNA was dissolved in 400 uL of 100 mM NaOAc (pH 5.0) and 1 mM fluorophore or biotin mono-reactive hydrazide (10 mM stock solution in water) and incubated on ice overnight. The reaction was then ethanol-precipitated and redissolved in 50 uL double distilled water and analyzed on denaturing polyacrylamide gel.

Fluorophore/biotin lableling of the 5' end of target mRNA using EDC (1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride). The reaction was performed as described by ThermoScientific. 7.5-15 nmol (~60-120 µg) RNA was dissolved in ~10 µl of water. Biotin or fluorophore hydrazide was dissolved to a final concentration of 0.25 M in 10 µl of 0.1 M imidazole. 1.25 mg (6.52 µmol) of EDC was weighed out and 7.5 µl of the prepared oligonucleotide was added to EDC. 5 µl of the hydrazide containing solution was added to it immediately and vortexed until contents were completely dissolved. An additional 20 µl of 0.1 M imidazole, pH 6, was added and the reaction was incubated overnight at room temperature. The non-reacted EDC was removed by purifying the sample using the bio-rad Micro Bio-Spin 30 Columns. The sample was eluted in 25 uL of water and analyzed on a denaturing polyacrylamide gel.

Capping and Tailing of target mRNA with modified NTPs containing either biotin or fluorophore. The target mRNA was capped with 8[(6-Amino)Hexyl] Amino-GTP-Bio using the ScriptCap™ m^{7}G capping system, from Epicentre Biotechnologies, as described by the manufacturers protocol. The RNA was purified using phenol-chloroform extraction followed by ethanol precipitation. The purified RNA was dissolved in water and tailed with fluorophore modified ATP using the Poly(A) Polymerase Tailing Kit, from Epicentre Biotechnologies, as described by the manufacturers protocol. The RNA was purified using phenol-chloroform extraction followed by ethanol precipitation. The purified RNA was dissolved in water and analyzed on a denaturing polyacrylamide gel.

Transcriptional priming of target mRNA with Bio- GG. The target mRNA sequence was designed to start with GG right after the T7 promoter sequence as depicted in **Section 4.2**. This facilitated the addition of Bio-GG at the 5'- end of the target mRNA transcript when the target mRNA was transcribed as described

in **Section 4.2**. Bio-GG was added to the transcription reaction in a 1:10 ratio with NTPs.

Slicing assay observed using Single Molecule Total Internal Reflection Fluorescence Microscopy (sm-TIRFM). Single molecule experiments were performed on a custom-built microscope equipped for prism-type TIRF. 532 nm and 635 nm lasers were be used to excite Cy3 and Cy5 respectively. Appropriate filter sets were used to select for emission from each fluorophore. An ICCD camera (Princeton Instruments, I–Pentamax) was used for image acquisition at 100 ms time resolution. First, the presence of tethered mRNA labeled with biotin and fluorophore was verified and then the slicing reaction containing the siRNA with the HeLa cell extract prepared as described in **section 4.2** was flowed into the sample chamber and imaged. Data analysis was performed using custom-made MATLAB routines.

D.3 Results

Target mRNAs were best labeled with fluorophores at the 3'-terminus using periodate oxidation and with biotin at the 5'-terminus using EDC coupling.
Three different methods of target mRNA labeling was tested. We observed that the transcriptional priming of mRNA with biotin-GG yielded less than 10% biotinylated mRNA, and 5'- capping using modified GTP resulted in ~20% of biotinylated mRNA (**Fig D.1**). Tailing of target mRNA with modified ATP yielded 10-30%, however, there is limited control over the number of fluorophores being added to a single mRNA and hence we sought other labeling strategies as well. Incorporation of biotin using the EDC coupling method resulted in 30-60% labeling efficiency consistently (**Fig D.2**). The efficiency of biotinylation was tested by performing supershift assays with strepavidin or by estimating the bind efficiency of target mRNA to streptavidin coated beads. We also observed that periodate labeling of target mRNA with Alexa 488 hydrazide (**Fig D.1**) and Cy3/Cy5 hydrazide (**Fig D.2**) resulted in 10-20% labeling efficiency. We used Cy3 or Cy5 hydrazide as Alexa 488 is not compatible with the prism-type TIRF setup.

159

Target mRNA cleavage at the single molecule level. Slicing assays performed at the single molecule level as shown in the schematic (**Fig D.3**) using mRNAs capped with 8[(6-Amino)Hexyl] Amino-GTP-Bio in the 5`end and Poly-A tailed with ATP-Cy3 at the 3`end in the presence and absence of siRNA (cy5-guide strand) duplex showed target mRNA cleavage activity. However the rate of cleavage in the presence of specific siRNA was only ~1.5 fold faster than in the absence of siRNA (**Fig D.4**). Target mRNA cleavage assays performed with labeled with cy3 at the 3'-terminus using periodate oxidation and with biotin at the 5'-terminus using EDC coupling showed ds- siRNA (unlabeled) specific cleavage with extent of cleavage being maximum at 30 min similar to that observed in bulk assays (**Fig D.5**). The initial lag observed during the cleavage assay at the single molecule level can be attributed to the time take for the reaction to equilibrate from 4°C to room temperature.

D.4. Conclusion

Target mRNAs labeleling with fluorophores at the 3'-terminus using periodate oxidation and with biotin at the 5'-terminus using EDC coupling proved to be the most efficient method in our hands. We observe ds- siRNA specific cleavage activity at the single molecule level. Further investigation of target mRNA cleavage and possibly strand unwinding will help us understand the RNAi pathway better.
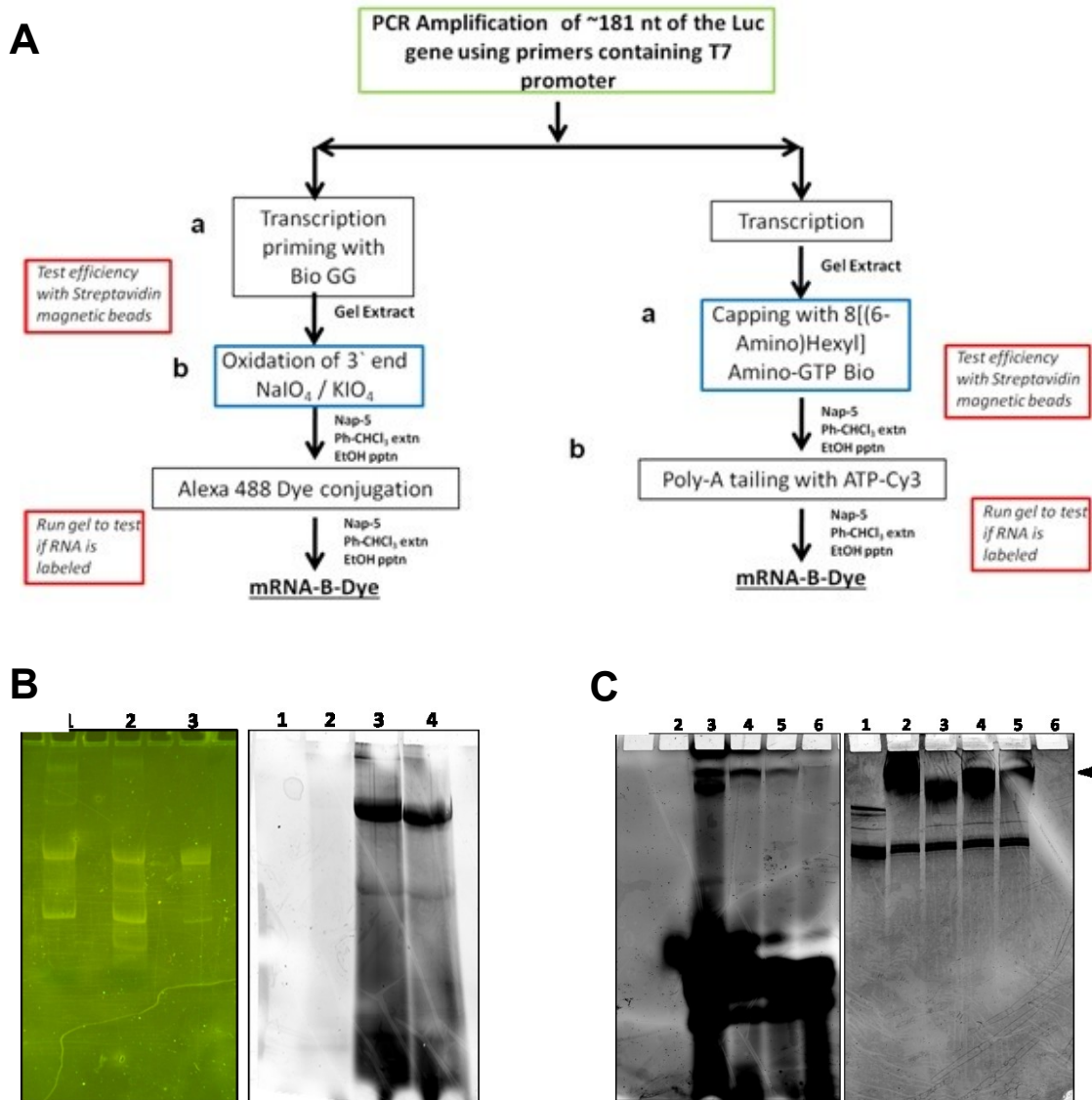
Fig D. 1. Methods of labeling target mRNA with biotin and fluorophore. (A) Schematic of two methods used for labeling target luciferase mRNA. **(B)** SYBR gold stained non-denaturing polyacrylamide gel electrophoresis of mRNA labeled with BioGG in the 5'- end and supershifted in the presence of streptavidin (Left). Lanes: 1. mRNA plus Streptavidin, 2. mRNA-BioGG plus Streptavidin, 3. mRNA only control. 488 nm fluorescent scan of non-denaturing polyacrylamide gel electrophoresis of mRNA labeled with Bio-GG in the 5'- end and Alexa 488 at the 3'-end (Right). Lanes: 1. Loading dye only control, 2. Alexa 488 only control, 3. mRNA with Alexa 488, mRNA- bio-GG- Alexa 488. **(C)** 532 nm (left) and 488 nm (SYBR Gold scan, right) fluorescent scan of denaturing polyacrylamide gel electrophoresis of mRNAs Poly-A tailed with ATP-Cy3. Lanes: 1. mRNA only, 2. mRNA – ATP, 3. mRNA – ATP cy3: ATP- 1:1, 4. mRNA – ATP cy3: ATP- 1:4, 5. mRNA – ATP cy3: ATP- 1:10, 6. Loading dye only.
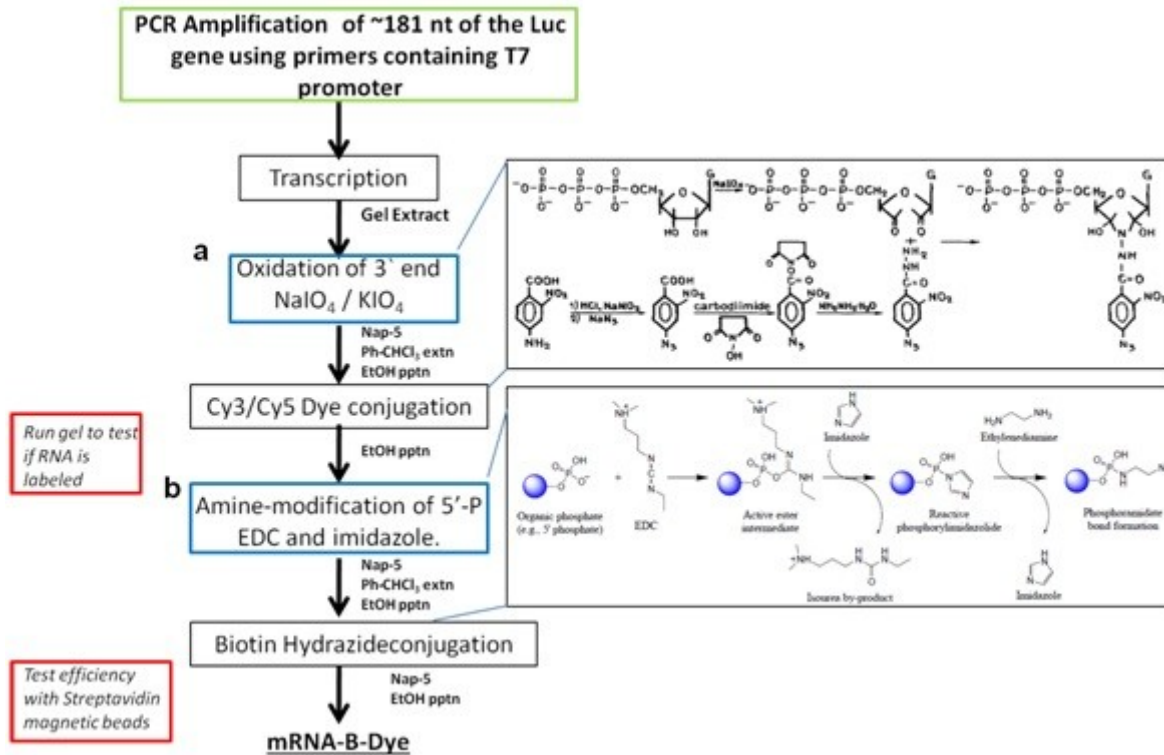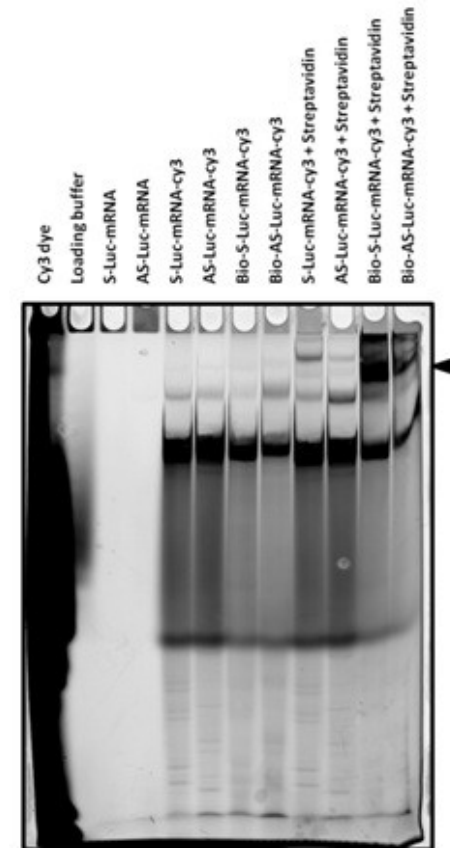
Fig D. 2. Labeling target mRNA with biotin and fluorophore. (A) Schematic of labeling target luciferase mRNA. **(B)** 532 nm fluorescent scan of denaturing polyacrylamide gel electrophoresis of mRNAs labeled with cy3 at the 3'-terminus using periodate oxidation and with biotin at the 5'-terminus using EDC coupling. The black arrow represents the supershift in the presence of streptavidin.
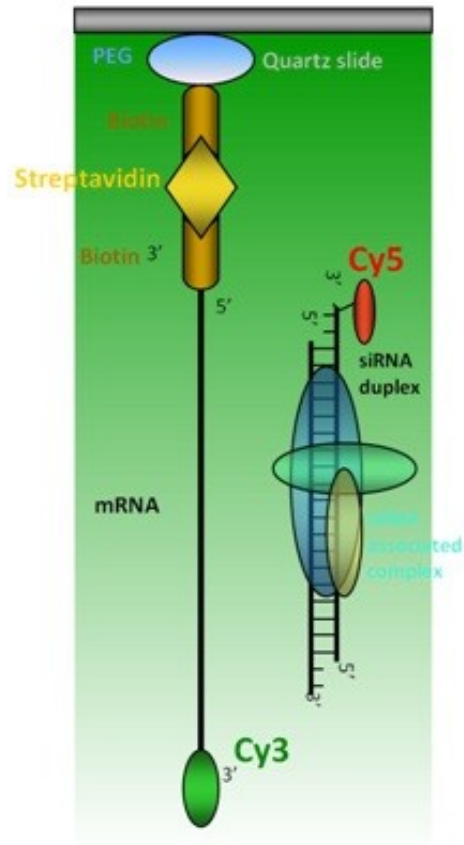
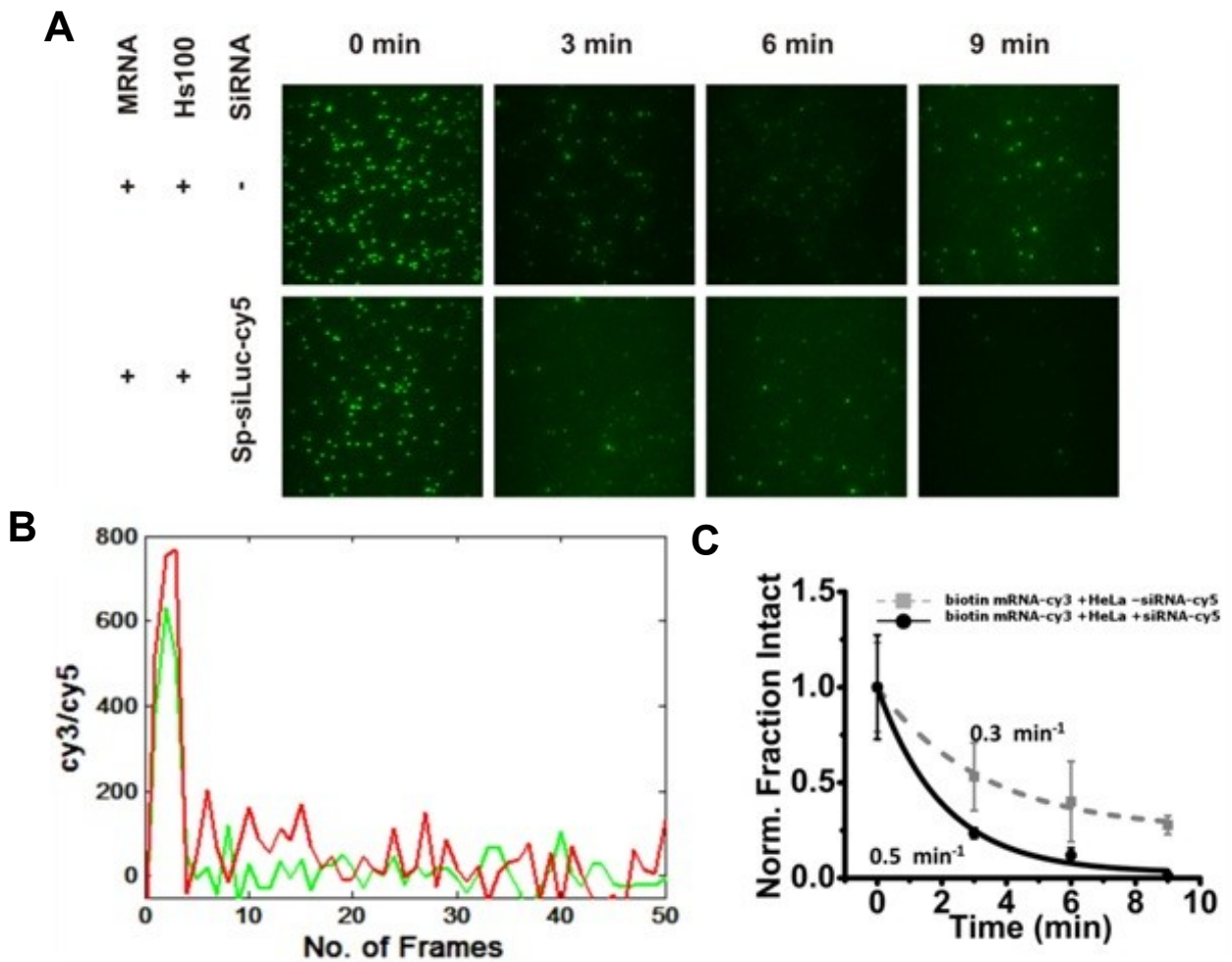**Fig D. 3. Schematic- Reaction setup for slicing assays performed at the single molecule level.**

Figure D.4. Target mRNA cleavage activity. (A) Representative images depicting the distribution of single-immobilized target mRNA molecules over the camera field of view at various time-points post HeLa cell extract +/- siRNA addition. (B) Representative intensity trajectory of an siRNA bound mRNA. The loss of signal from both the siRNA (Cy5/ red solid line) and mRNA (Cy3/ green solid line) simultaneously may depict a cleavage event. (C) Kinetics of cleavage. Grey dash line and squares represent fraction mRNA intact in the sample that was only treated with the extract. Black solid line and circles represents data points from the sample that was treated with both the siRNA and the extract. For these experiments mRNA was capped with 8[(6-Amino)Hexyl] Amino-GTP Bio in the 5`end and Poly-A tailed with ATP-Cy3 at the 3`end. mRNA was first immoblized on the PEG passivated streptavidin coated surface. Hela cell extract was incubated with/without siRNA at 4 C for 30 min, 37 ^{o}C for 15 min and added

to the surface containing the immobilized mRNA. 30 continuous frames (at 10 frames per second) were taken every 3mins. Three different areas on the slides were imaged at every time-point.
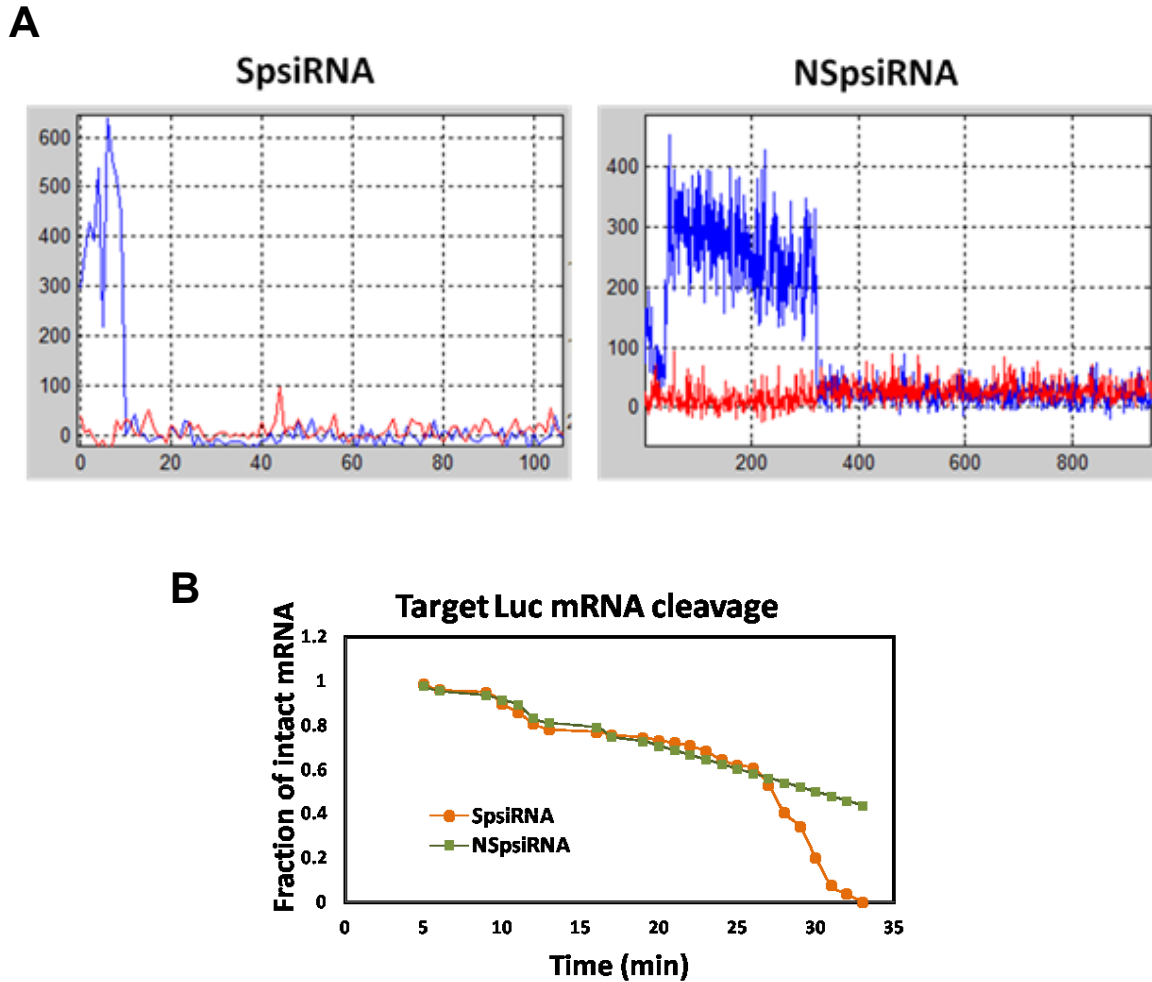
Figure D.5. Target mRNA cleavage activity. (A) Representative intensity trajectory of mRNA cleavage using specific (SpsiRNA) and non-specific siRNA (NSpsiRNA). The loss of signal from mRNA (Cy3/ blue solid line) depicts a cleavage event. (C) Kinetics of cleavage. Orange line and round represent fraction mRNA intact in the sample that was treated with the extract containing SpsiRNA. Green line and squares represents data points from the sample that was treated with the NSpsiRNA and the extract. For these experiments mRNA was labeled with Cy3 at the 3'-terminus using periodate oxidation and with biotin at the 5'-terminus using EDC coupling. mRNA was first immoblized on the PEG passivated streptavidin coated surface. Hela cell extract was incubated with SpsiRNA or NSpsiRNA at 4 ^{o}C for 30 min and added to the surface containing the immobilized mRNA. 30 continuous frames (at 10 frames per second) were taken every min. The same area on the slide was imaged at every time-point.