This is to certify that the
dissertation entitled

INTERACTION OF DICER, TRBP WITH SHORT INTERFERING RNAs: EFFECT ON SILENCING EFFICACY OF siRNAs

presented by

HEMANT K. KINI

has been accepted towards fulfillment
of the requirements for the

Ph.D degree in Chemical Engineering

Major Professor's Signature

11/30/09

Date

INTERACTION OF DICER, TRBP WITH SHORT INTERFERING RNAs: EFFECT ON SILENCING EFFICACY OF siRNAs

By

HEMANT K. KINI

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Chemical Engineering

2009

ABSTRACT

INTERACTION OF DICER, TRBP WITH SHORT INTERFERING RNAs: EFFECT ON SILENCING EFFICACY OF siRNAs

By

HEMANT K. KINI

Gene silencing by RNA interference (RNAi) is mediated by endogenous proteins and leads to target mRNA cleavage or translational inhibition. In the cytoplasm, the dsRNA binding proteins, TRBP and Dicer, recognize and bind the siRNA, and form the RNA induced silencing complex (RISC) loading complex (RLC). Argonaute 2, which forms the catalytic core of RISC, is then recruited to this protein complex. Ago2 then gets loaded with the guide strand, the strand complementary to the mRNA, and silences the target mRNA.

In this research, a novel role for human Dicer has been demonstrated. It was found that Dicer binds 21-nt ssRNAs, having higher affinity for those possessing a 5'-phosphate relative to a 5'-hydroxyl. Using liquid chromatography-mass spectrometry (LC/MS), the different Dicer domains binding the ssRNAs vs. the double stranded siRNAs were preliminarily identified. Based on these findings and prior findings, a model has been proposed for the loading of Ago2 with the single stranded guide strand.

Dicer along with TRBP recognizes, binds with the siRNAs and forms the RLC. This constitutes an important step in the RNAi pathway. To enhance silencing and avoid off

target effects, siRNAs are often designed with an intentional bias to make the end of the siRNA containing the guide strand 5'-end less stably hybridized relative to the end containing the passenger strand 5'-end. By studying a number of siRNAs with terminal mismatches, it was found that siRNA-TRBP binding is largely indicative of eventual silencing efficacy of the siRNAs and that this binding can be significantly reduced by terminal mismatches. TRBP-siRNA complex formation is more tolerant of internal mismatches than terminal mismatches. Terminal mismatches did however lead to a small increase in Dicer binding, though without a concomitant improvement in silencing activity. These results demonstrate that introduction of mismatches to control siRNA asymmetry may not always serve to improve target silencing and that care should be taken when designing siRNAs in this way.

Here, a previously unknown function of Dicer protein in its ability to bind ssRNAs, and the impact of TRBP protein on the silencing activity of asymmetric siRNAs with a terminal mismatch has been demonstrated. Overall this research has furthered the knowledge of the functional roles of two important proteins, Dicer and TRBP, in the RNAi pathway.

To my parents.

ACKNOWLEDGEMENTS

I am grateful to several people for their help, support, friendship, and guidance over these years as a graduate student. First and foremost I would like to offer my gratitude to my advisor, Pat Walton whose constant support and guidance was crucial to me and my research. I am thankful to Pat for not only inspiring me to think independently, but also giving me the freedom to plan and direct my own research projects. I would like to offer my special thanks to all the members of my Ph.D committee (Dr. Donna Koslowsky, Dr. Kris Chan, and Dr. Charles Petty) for their valuable insights and suggestions. To Dr. Kris Chan for her positive critique of my work, words of encouragement and insightful suggestions. Dr. Charles Petty who was also my masters thesis advisor, for getting me started on the path of scientific research.

I am truly grateful to the members of the Proteomics core (Doug Whitten) and the Mass spectrometry facility (Dr. Dan Jones and Dr. Lijun Chen) at MSU, for the training and ungrudging help they have provided me. Particularly to Dr. Dan Jones and Dr. Lijun Chen, who personally helped me with several of my mass spectrometry experiments and data analysis, despite their hectic schedules. I would also like to extend my special thanks to all the members of Chan and Walton labs, who were a great team to be part of. To Shireesh, Sheenu, Sachin, Sumit, Joe, Xuerui, Linxia, Hirosha, and Shengnan for their collaboration, support and camaraderie.

I would also like to thank the Chemical Engineering and Materials Science Department, the College of Engineering and, Office of International Students and Scholars (OISS) for their financial support. In particular, I would like to thank the staff from the Chemical

Engineering and Materials Science Department, Nancy, JoAnn and Eunice for their administrative support with cheerful demeanor.

To my dear friends who always made the Windy city a warm enclave for me Sandeep Gaonkar, Raj Naik, and rest of the Gaonkar, Naik and Torke clan. To Mrs. Marina Farhat, Gina Farhat, and Susan Farhat who always treated me like family. *Gratzi.* To all the friends to whom I am thankful for the good times, great memories, support and sage advice Afshan (Bawa), Bhavin, Sharad, Madhu, Srivatsan, Krishnan, Amit, Harish and Chetan.

Family and friends have been the quintessential rock of support for me through all these years. My utmost gratitude to my parents for their love, support and sacrifices they had to make, for me to be able to get here. To my brother, Ajay for not only for giving me a cute nephew (Tanay), but also for all his unerring love and support. Most importantly to my wife Akshata, for her love, understanding, support and trust which I can always rely on. I appreciate it more than you will ever know- thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Ago2	Argonaute 2
ATP	Adenosine triphosphate
Bi	Biotin
BSA	Bovine serum albumin
Da	Daltons
Dcr	Dicer
DNA	Deoxyribonucleic acid
dsRBD	Double-stranded RNA binding domain
dsRNA	Double-stranded RNA
DUF283	Domain of unknown function 283
EDTA	Ethylenediaminetetraacetic acid
eIF4E	Eukaryotic translation initiation factor 4E
EGFP	Enhanced green fluorescent protein
FBS	Fetal bovine serum
HeLa	Human cervical carcinoma cell line
HepG2	Hepatocellular carcinoma cell line
H1299	Non-small cell lung carcinoma cell line
HIV	Human immunodeficiency virus
kDa	kiloDalton
Loqs	Loquacious
m/z	mass to charge ratio of peptides
miRNA	microRNA
mRNA	Messenger RNA

nt	Nucleotide
PACT	PKR-activating protein
PAZ	Piwi Argonaute Zwille
PBS	Phosphate buffered saline
PKR	Protein kinase R
R2D2	Protein with two RNA domains (R2) and interacting with Dicer 2 (D2)
RISC	RNA-induced silencing complex
RLC	RISC-loading complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase	Ribonuclease
shRNA	Short hairpin RNA
siRNA	Short interfering RNA
ssRNA	Single-stranded RNA
TAR	Transactivating response
TBE	Tris-borate-EDTA
TRBP	TAR RNA binding protein
UV	Ultraviolet

CHAPTER I

RNAi pathway: initiation to application

INTRODUCTION

RNA interference (RNAi)

RNA interference (RNAi) is an evolutionarily conserved mode of gene regulation and host defense in eukaryotes (Cullen 2002; He and Hannon 2004). Short interfering RNAs (siRNAs) can be designed to specifically target and regulate the expression of any particular gene. However dsRNAs longer than 30bp activate the mammalian cell defense mechanism leading to induction of cytokine production and cell death (Williams 1999; Caplen *et al.* 2001; Elbashir *et al.* 2001a). Hence, 21-mer siRNAs are the most commonly used silencing agents for mammalian systems (Elbashir *et al.* 2001a). SiRNAs have been widely employed as tools to study gene function, regulating protein expression, and in therapeutic applications (Dorsett and Tuschl 2004; Whitehead *et al.* 2009). SiRNAs are similar in structure to endogenously encoded microRNAs (miRNAs), which regulate diverse cellular pathways and are key regulators of development and disease (Stefani and Slack 2008; Ghildiyal and Zamore 2009). Both of these active species share similar components and mechanisms of the RNAi pathway.

Initiation of RNAi: Dicer, miRNAs and siRNAs

RNAi can be initiated by exogenous or endogenous means. In endogenous regulation, primary miRNAs (pri-miRNAs) are produced in the nucleus by RNA polymerases II or III. These are then processed by the nuclear complex of Drosha, an RNase III family enzyme, and its dsRNA binding partner, which is known as Pasha in Drosophila (Denli *et al.* 2004) and DGCR8 in humans (Gregory *et al.* 2004; Han *et al.* 2004). This processing

yields precursor miRNAs (pre-miRNAs) (Lee *et al.* 2002; Zeng and Cullen 2003). Pre-miRNAs have 5'-phosphates and the 3' dinulceotide overhangs characteristic of RNase III processing (Basyuk *et al.* 2003; Lee *et al.* 2003). Once generated, pre-miRNAs are exported to the cytoplasm by Exportin-5-Ran-GTP (Yi *et al.* 2003). In the cytoplasm, Dicer, another RNase III family enzyme, further processes the pre-miRNA to the ~21-mer mature miRNAs (Lee *et al.* 2003). In a similar manner, Dicer can process any long dsRNA substrate to siRNAs with similar structure to miRNAs.

Human Dicer, which generates both siRNAs and miRNAs (Fig. 1.1), is a multi-domain protein with an N-terminal RNA helicase/ATPase domain, a domain of unknown function (DUF 283), a Piwi Argonaute Zwille (PAZ) domain, two RNase III domains (RIIIa and RIIIb), and a dsRNA binding domain (dsRBD) (Hammond 2005; Cook and Conti 2006). The RNase domains cleave the target dsRNA into siRNAs and miRNAs with characteristic 2 nt overhangs at each 3' end (Lee *et al.* 2002; Zhang *et al.* 2004). The C-terminal dsRBD and the PAZ domain orient the dsRNA for cleavage at the proper locations (Zhang *et al.* 2004).

Drosophila has two Dicers, Dcr1 and Dcr2, which have distinct roles in the miRNA and siRNA pathways (Lee *et al.* 2004; Farazi *et al.* 2008). Dcr1 requires the dsRNA binding protein Loquacious (Loqs) for processing of pre-miRNAs into mature miRNAs (Forstemann *et al.* 2005). Unlike Dcr1, Dcr2 alone can cleave long dsRNA substrates (Forstemann *et al.* 2005). Similar to the *Drosophila* Dicers, the human Dicer is also associated with a partner protein, the human immunodeficiency virus (HIV-1) transactivating response (TAR) RNA-binding protein (TRBP) (Chendrimada *et al.* 2005; Haase *et al.* 2005).

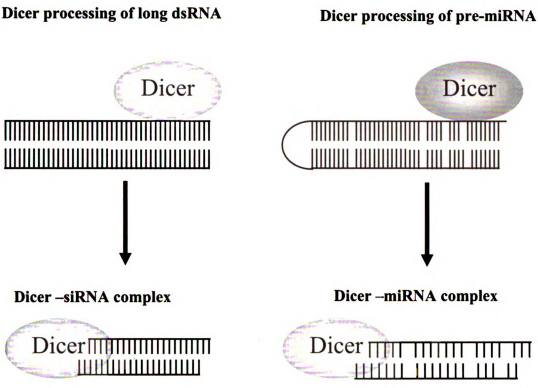


Figure 1.1 RNase III activity of Dicer.

Dicer processes long dsRNA or pre-miRNA substrates into siRNAs and miRNAs, respectively. Dicer is suggested to stay bound with the siRNAs/miRNAs it generates and facilitate their incorporation into RISC.

Initiation complex - RLC

Dicer generated siRNAs and miRNAs have a 5'-phosphate and 3'-hydroxyl (Zhang *et al.* 2002). Exogenously introduced, chemically synthesized siRNAs, which have a 5'-hydroxyl, upon entry into cells, are immediately phosphorylated by the human RNA kinase hClp1 (Nykanen *et al.* 2001; Weitzer and Martinez 2007). In the cytoplasm, siRNAs and miRNAs possessing a 5'-phosphate are then recognized by Dicer and TRBP to form the RNA induced silencing complex (RISC) loading complex (RLC).

TRBP

TRBP is a dsRNA binding protein known to stimulate the expression of HIV-1 virus (Duarte *et al.* 2000; Gatignol and Jeang 2000; Daher *et al.* 2001) and inhibit the activity of another dsRNA binding protein kinase R (PKR) (Daher *et al.* 2001). TRBP has two dsRNA binding domains and the C-terminal Medipal domain, which interacts with the other proteins, namely Merlin, Dicer, and PKR activating protein, to which TRBP is structurally similar (PACT) (Laraki *et al.* 2008).

However, they exert opposite regulatory effects on PKR, with PACT activating PKR (Patel and Sen 1998; Li *et al.* 2006) and TRBP inhibiting PKR (Daher *et al.* 2001). TRBP interacts with Dicer in processing pre-miRNAs and also in forming the RLC (Fig. 1.2) (Chendrimada *et al.* 2005; Haase *et al.* 2005). The RLC then recruits Argonaute 2 (Ago2) to form holo-RISC (Fig. 1.2).

RISC catalytic protein - Argonaute 2 (Ago2)

Ago2 belongs to the Argonaute protein family which derives its name from typical squid-like phenotype witnessed in plants lacking Argonaute proteins (Peters and Meister 2007). These proteins are characterized by the presence of PAZ (Piwi-Argonaute-Zwille) and PIWI (P-element induced wimpy testis) domains (Bohmert *et al.* 1998). These proteins are classified into the Ago subfamily and the Piwi subfamily (Peters and Meister 2007; Hutvagner and Simard 2008). While Ago proteins are ubiquitously expressed, Piwi proteins are restricted to germline cells (Carmell *et al.* 2007). Piwi proteins interact with another class of short interfering RNAs called piRNAs (Piwi-interacting RNAs) which

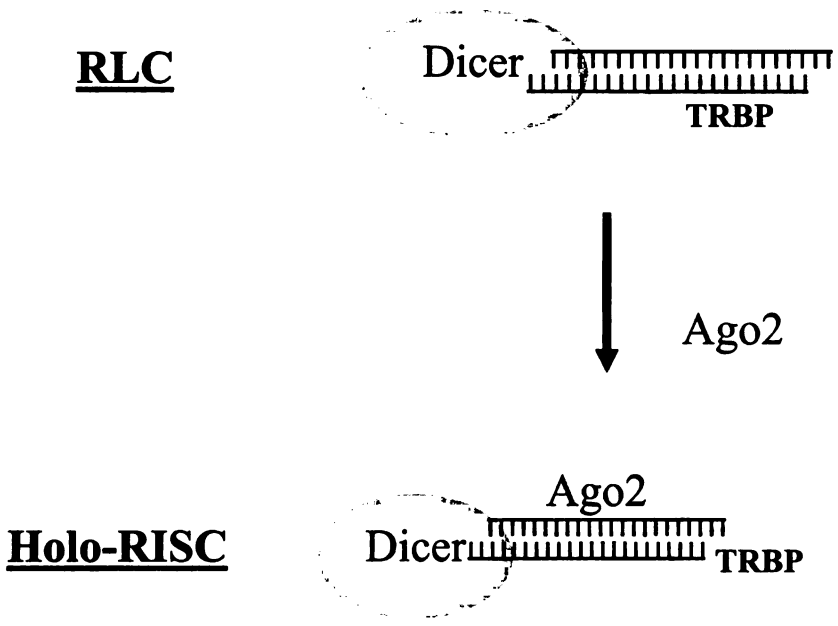


Figure 1.2 Formation of RLC and holo-RISC

Dicer and TRBP recognize siRNAs and form the RLC. Ago2 is then recruited to this complex to form holo-RISC.

are necessary for normal germline development (Das *et al.* 2008; Klattenhoff and Theurkauf 2008). Various Argonaute proteins have been shown to be involved in diverse cellular functions from embryonic development to transposon silencing (Lykke-Andersen *et al.* 2008; Siomi and Siomi 2008). The number of Argonaute genes varies between species with 27 in *C.elegans* to 8 in humans (Peters and Meister 2007). Out of these 8 human Argonaute proteins, Ago2 is the only protein possessing endonuclease activity (Meister *et al.* 2004). The crystal structure of the *Pyrococcus furiosus* Argonaute PIWI domain shows that this domain is very similar to RNase H, supporting the role of the PIWI domain as the active center of RISC (Song *et al.* 2004). While RNase H cleaves the RNA in a DNA-RNA duplex, Ago2 cleaves an RNA strand of an RNA-RNA duplex. For

silencing, this results in cleavage of the target mRNA when hybridized to the guide strand (Liu *et al.* 2004; Rivas *et al.* 2005). Due to its RNase activity, Ago2 is also referred to as 'Slicer' (as a counterpart to Dicer).

siRNA containing holo-RISC to guide strand programmed active RISC

Minimal RISC has been reconstituted *in vitro* using only purified Ago2, Dicer, and TRBP (MacRae *et al.* 2008). Activation occurs when one strand is destroyed during the process and thereby creating an active RISC. While either strand of a siRNA duplex can be part of active RISC, chemically synthesized siRNAs are typically biased such that only one strand is preferentially contained in active RISC (Schwarz *et al.* 2003). The initial four base pairs at each 5' end of the siRNA duplex are typically used to determine the stability of that end (Fig. 1.3). The strand with the relatively unstable 5' end becomes the guide strand and the complementary strand becomes the passenger strand. In *Drosophila*, the protein R2D2 senses the more stable 5' end of the siRNA duplex and binds to that end thereby positioning Dcr2 to the less stable end of the duplex, and forming the RLC (Tomari *et al.* 2004). TRBP, which is the human ortholog of R2D2, might play a role in sensing relative thermodynamic bias in humans (Haase *et al.* 2005; Pellino 2007). Following asymmetry sensing, the RLC then positions the less stable 5' end of the siRNA to anchor with the phosphate binding pocket of the Ago2 PIWI domain (Ma *et al.* 2005; Parker *et al.* 2005). Then Ago2 cleaves the passenger strand phosphodiester bond across the 10 and 11 nucleotides of the guide strand (Matranga *et al.* 2005; Leuschner *et al.* 2006) similar to the cleavage of target mRNA by the guide strand (Fig. 1.4) (Martinez *et al.* 2002). This mode of passenger strand cleavage works only for siRNA duplexes whose strands are perfectly base paired around the passenger strand cleavage site. For miRNAs,

which have internal mismatches, the programming of RISC with the guide strand is predicted to occur by a bypass mechanism which does not involve passenger strand cleavage (Matranga *et al.* 2005).

Active RISC and silencing of target mRNA

Purified Ago2, when provided double-stranded siRNAs, lacks the ability to form an active RISC programmed with the single-stranded guide strand (Rivas *et al.* 2005). However, Ago2 alone can combine with a ssRNA acting as a guide strand to form an active RISC (Rivas *et al.* 2005). *In vitro* experiments have shown that once Ago2 is loaded with the guide strand it dissociates from Dicer and TRBP (Fig. 1.4) (MacRae *et al.* 2008).

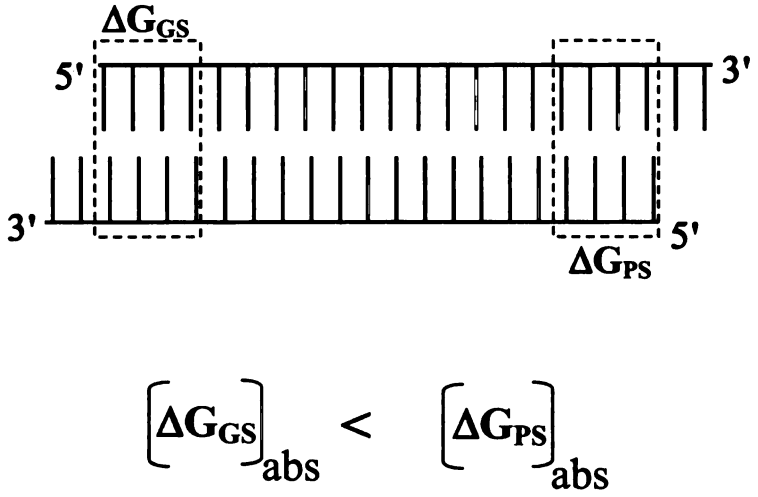


Figure 1.3 Relative stabilities of 5' ends determines the guide strand of a siRNA

First four base pairs at the 5' end contribute to the stability of that end. This schematic represents a case where the 5' end of the red strand is less stable than the 5' end of the blue strand. So the red strand is becomes the guide strand (guides RISC to the target mRNA) and the blue strand is designated the passenger strand (is subsequently destroyed).

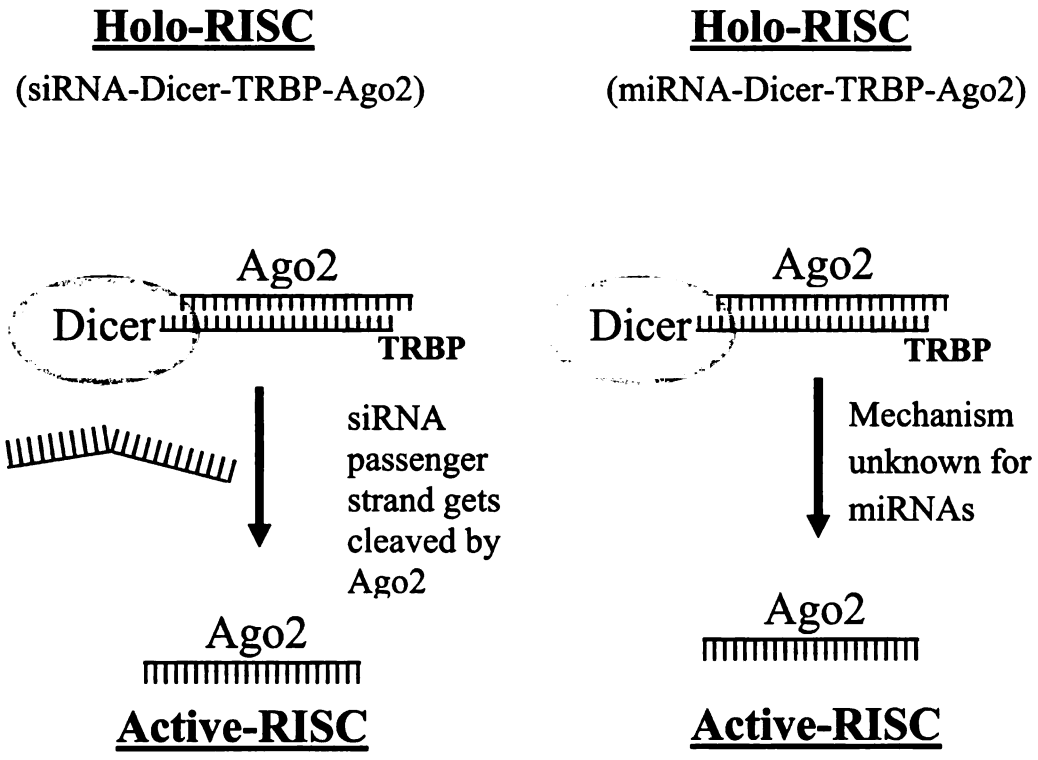


Figure 1.4 Transition from holo-RISC to active RISC

The passenger strand of the siRNA in the holo-RISC gets cleaved by Ago2 to form the active RISC programmed with the guide strand. The mechanism of unwinding/cleaving of the passenger strand for miRNAs is currently unknown.

Active RISC then binds with the target mRNA based on sequence complementarity. If there is sequence complementarity between the guide and the target around the central part of the guide strand, then Ago2 cleaves the target mRNA (Elbashir *et al.* 2001a; Holen *et al.* 2002) (Fig. 1.5). Following substrate release, active RISC is liberated and can engage in multiple cycles of target cleavage (Haley and Zamore 2004; Rivas *et al.* 2005)

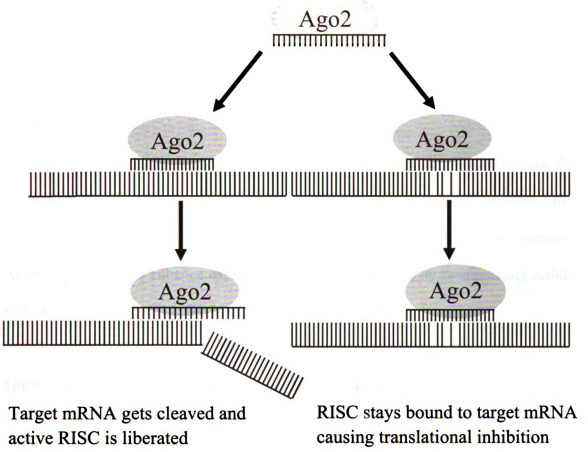


Figure 1.5 Target mRNA silencing by active RISC.

Active RISC programmed with the guide strand binds with the target mRNA, and cleaves the complementary target. Following the target cleavage active RISC becomes available for more catalytic activity. If the there are mismatches between the guide strand and the target mRNA then RISC stays bound to the target preventing the mRNA translation.

RNAi : discovery to application

Since its discovery in 1998 the use of RNAi has become a standard procedure for gene regulation. While there is still a lot of progress to be made related to siRNA design and efficacy (Tilesi *et al.* 2009), stability (Bramsen *et al.* 2009), controlled delivery (Fattal and Barratt 2009), reducing off-target effects (Jackson *et al.* 2006), immune response (Judge *et al.* 2005).

Ease of design and high specificity of siRNAs make them a convenient tool for, understanding gene functions and regulating their expression. SiRNAs offer novel therapeutic approaches to treat diseases by providing the ability to silence aberrantly expressed genes or to enhance expression of valuable proteins by manipulating cellular pathways (Zhu and Galili 2004; Martin and Caplen 2007).

Applications of RNAi in biotechnology and biomedicine

Metabolic flux analysis and genetic screens have identified cellular pathways influencing protein synthesis and cell survival (Tewari and Vidal 2003; Srivastava *et al.* 2007). Those pathways identified as being too active to achieve the desired phenotype can be specifically targeted and their activities reduced using RNAi. One specific application in which this has been demonstrated is in the expression of proteins by mammalian cells, which grow more slowly than bacterial cells, but provide proper folding and post-translational modification of the expressed proteins (Wurm 2004).

Since the approval of the tissue plasminogen activator as the first recombinantly expressed therapeutic protein from Chinese Hamster Ovary (CHO) cells, the CHO cell line and its variations have become the dominant mammalian system for recombinant protein synthesis (Kumar *et al.* 2008). Maximal protein production from CHO cells

requires high viability and initial rapid proliferation of the cells. Apoptosis from stress or other signals reduces the yield and purity of the recombinant protein product. As cells lacking both the proapoptotic proteins BAK and BAX exhibit considerably higher viability compared to cells expressing both these proteins (Wei *et al.* 2001), siRNAs were used to target these proteins in recombinant human IFN-γ expressing CHO cells. With BAK and BAX mRNAs reduced to 10% of their normal level, the CHO cultures showed 30-50% greater viable cell density, and 35% higher IFN-γ production relative to control cells (Lim *et al.* 2006).

Cellular pathways in plants can also be manipulated using RNAi to engineer and over-express mammalian antibodies and plant specific proteins. Plants have emerged as a safe and economical alternative to microbial and mammalian cell lines for expressing vaccines and therapeutic antibodies (Floss *et al.* 2007), in part because they have minimal risk of contamination from potential human pathogens and can be easily scaled up for mass production without the need for extensive purification steps (Daniell *et al.* 2001). For example, the aquatic plant *Lemna minor* has been developed for producing high yields of therapeutic recombinant proteins (Neuenschwander *et al.* 1991). These proteins are, however, susceptible to plant specific glycosylation by the genes α-1,3-fucosyltransferase and β-1,2-xylosyltransferase (Gomord *et al.* 2005). After silencing these enzymes by vector expressed shRNAs, their activities were reduced to the levels of negative control. When used to produce monoclonal antibodies against human CD30, a human cell surface receptor that is specifically upregulated in certain tumor cells such as Hodgkin and Reed-Sternberg cells, adult T cell leukemia, and embryonal carcinoma of the testis (Dong *et al.* 2003), the antibodies did not have any detectable plant specific

glycans and were more potent than antibodies expressed in mammalian cell lines (Cox *et al.* 2006).

Interestingly, storage organs such as potato tubers can also serve as bioreactors for mass production of human proteins and vaccines, as they provide superb environments for maintaining protein stability (Farran *et al.* 2002; Arntzen *et al.* 2005). A major problem for proteins expressed in tubers is patatin contamination (Arntzen *et al.* 2005). Patatins are a family of glycoproteins that constitute nearly 40 % of the soluble protein in potato tubers (Prat *et al.* 1990). shRNA vectors have been used to target the highly conserved gene (*pat3-k1*) of the patatin family, thereby reducing patatin expression by nearly 99% (Kim *et al.* 2008). The expressed glycoproteins did not then require purification steps to remove patatin, resulting in significant improvement in protein yield.

In cases where the native plant proteins are detrimental, RNAi has been shown to be active in reducing expression of the undesired protein. Peanut allergy is one of the most severe food allergies, affecting nearly 1% of the US population (Sicherer *et al.* 1999), with 86 % of cases resulting from reactions to the protein Ara h 2 (Koppelman *et al.* 2004). Knocking down the expression of Ara h 2 using plasmid-expressed shRNAs resulted in hypoallergenic seeds showing significant reduction in allergenic potency, as measured by IgE binding capacity, compared to wild-type peanuts (Dodo *et al.* 2008).

RNAi can also be coupled with metabolic engineering to provide several advantages over classical plant breeding techniques, such as control of spatial and temporal expression of genes of interest (Tang *et al.* 2007). These new methods of breeding are being investigated to improve the nutritional value of food, offset the loss of agricultural land, and help satisfy the food demand of the growing world population (Doos 2002). For

instance, cotton tissues express the protein Gossypol, which acts as a defense against insects and pathogens (Sunilkumar *et al.* 2006). However, Gossypol's toxicity to humans prevents the use of cotton seeds as a source of food in developing countries (Townsend and Llewellyn 2007). shRNAs targeting the enzyme δ-cadinene, which is crucial for Gossypol synthesis, reduced the Gossypol level in transgenic seeds by 99% compared to wild-type seeds, with the transgenic plant showing normal growth and development (Sunilkumar *et al.* 2006).

A similar approach was used to enhance the plant production of lysine, an essential amino acid important for human nutrition and also livestock growth. Lysine is found in limiting amounts in corn and other cereal grains (Singh *et al.* 2001), as it negatively regulates the activity of dihydropicolinate synthase (DHPS), the first enzyme in the lysine biosynthesis pathway (Tang *et al.* 2007). Plants engineered to express DHPS mutants insensitive to lysine showed increased lysine synthesis in all plant organs (Frankard *et al.* 1992). Unfortunately, elevated lysine levels caused abnormal tissue and flower development, which in turn reduced seed yield (Frankard *et al.* 1992). The quality of the seeds was also inferior due to defective post-germination lysine catabolism (Zhu and Galili 2004). Lys-ketoglutarate reductase (LKR) and saccharopine dehydrogenase (SDH) are key enzymes in the lysine catabolism pathway (Zhu and Galili 2004). In Arabidopsis plants, both the lysine content and seed quality were improved by seed-specific lysine over-expression using both the DHPS mutant and temporal shRNA silencing of the LKR and SDH enzymes during seed development (Zhu and Galili 2004). These seeds had about 25 mol % higher lysine content than wild-type seeds and also grew faster than

seeds from plants over-expressing lysine without LKR and SDH knockdown (Zhu and Galili 2004).

Alternatively, RNAi modified plants have been shown to be effective for therapeutic drug production. Morphinan alkaloids such as morphine, codeine, oripavine, and thebaine have direct therapeutic use and serve as intermediates for manufacture of synthetic analgesics used to treat opiate addiction (Allen *et al.* 2008). These alkaloids are only expressed in plants of the genus *Papaver* (Allen *et al.* 2008). Upregulating the flux through the alkaloid synthesis pathway by over-expression of the enzymes salutaridinol 7-O-acetyltransferase (SalAT) or codeinone reductase (COR) results in increased production of the alkaloids (Grothe *et al.* 2001; Allen *et al.* 2004). Morphine, however, remains a significant product of the pathway, impairing the expression and purification of thebaine and codeine, products of significantly greater therapeutic value. Targeting the COR transcript using shRNAs leads to almost complete inhibition of its enzymatic activity, resulting in the upstream accumulation of (S)-reticuline in the alkaloid biosynthesis pathway (Allen *et al.* 2004). (S)-reticuline is a valuable pharmaceutical intermediate that can easily be converted to salutaridine and then to thebaine (Page 2005). Engineering poppy plants in this manner is therefore a means to generate plants with significant medicinal value, while also possibly limiting their use in the production of illicit opioid drugs.

Summary, aims, and findings of this research

The application of RNAi technology to a multitude of platforms from bench side to expressing therapeutic proteins highlights it enormous potential. Hence the need to understand and optimize the technology is of critical importance. Several proteins are

involved in mediating this endogenous pathway of gene regulation and viral defense (Sontheimer 2005). Dicer and TRBP are two proteins, involved in both the initiation phase of the response and also the formation of the active complex.

Similar to siRNAs, 21-nt ssRNAs have also been used as silencing agents in Ago2 mediated silencing experiments (Schwarz *et al.* 2002; Rivas *et al.* 2005). It is not known if Dicer alone can bind 21-nt ssRNAs, as it binds siRNAs (Pellino *et al.* 2005). PIWI domain of Ago2 has been demonstrated to bind with the RNase III domain of Dicer (Tahbaz *et al.* 2004). Dicer binding with ssRNAs of this length would suggest its ability to interact and load Ago2 with the guide strand, which is also single stranded. Hence the primary aim of this research was to study if Dicer alone could bind with 21-nt ssRNAs. It was found that:

1. Dicer alone can bind both 21-nt ssRNAs and siRNAs independent of internal sequence and structure.

2. Only ssRNAs and siRNAs possessing a 5'-phosphate and not 5'-hydroxyl are bound by Dicer, which suggests that this binding is mediated by the Dicer PAZ domain.

3. Dicer binding with ssRNAs and siRNAs is stabilized by different Dicer domains, RNase IIIb domain for ssRNAs and dsRBD for siRNAs.

Based on these results a model has been proposed for the Dicer, Ago2 interaction leading to the loading of Ago2 with the single stranded guide strand.

Dicer along with TRBP and Ago2, contributes, to the formation of RISC with the siRNA, and selection of the guide strand to form active RISC. To enhance guide strand incorporation into RISC and reduce off target silencing by the siRNA passenger strand, siRNAs are routinely synthesized with a mismatch at the guide strand 5' end (Schwarz *et al.* 2006; DiFiglia *et al.* 2007). The effect of such mismatches on the activity of asymmetric siRNAs has not been studied. A comparative study of asymmetric siRNAs having a guide strand 5' end mismatch vs. asymmetric siRNAs without any mismatches, on their interaction with Dicer and TRBP, and also of their silencing activity was performed. The major findings were that:

1. Guide strand 5' end mismatch does not enhance the silencing efficacy of an asymmetric siRNA.
2. SiRNA-TRBP binding is largely indicative of eventual silencing efficacy of the siRNAs and that this binding can be significantly reduced by terminal mismatches.
3. Terminal mismatches led to a small increase in Dicer binding, as expected, however, this did not lead to an improvement in silencing activity.
4. Internal mismatch enhanced both Dicer and TRBP binding.

CHAPTER II

Human Dicer binding of ssRNAs and siRNAs

INTRODUCTION

Dicer is a member of the RNase III family of enzymes. There are three classes of RNase III enzymes (Fig. 2.1). Bacterial RNases, with only one RNase domain, comprise class I of the RNase III enzymes. The first RNase III family enzyme was isolated from *E.coli* and demonstrated divalent cation dependent catalysis of long dsRNAs (Robertson *et al.* 1968). These RNases act simultaneously on strands of dsRNAs by dimerization, each monomer possessing an active site cleaves one strand to generate duplexes ~ 10 bp in length. Class II enzymes, such as Drosha, have two endonuclease domains and one dsRBD. Human and *Drosophila* Dicers are more elaborate and, besides the dsRBD and endonuclease domains, also have the distinguishing PAZ and ATP helicase domains (Zamore 2001). Unlike the bacterial RNase which acts by dimerization, Dicer and Drosha act by intramolecular dimerization of their two endonuclease domains and, assisted by the dsRBD and PAZ domains, cleave the two strands of the substrate generating siRNAs and miRNAs.

Though Dicer has been shown to be part of RISC, its exact role in RISC is not well-understood. HeLa cell extracts stripped of Dicer protein by immunodepletion still show siRNA induced target mRNA cleavage activity (Martinez *et al.* 2002). Also, while Dicer is necessary for normal development of mice embryos (Bernstein *et al.* 2003; Harfe *et al.* 2005), mouse embryonic stem cells with an inactive Dicer still retained siRNA directed target mRNA cleavage (Kanellopoulou *et al.* 2005; Murchison *et al.* 2005). These results show that Dicer is not a necessary factor for siRNA induced silencing. However, Dicer

generated siRNAs from 27-mer dsRNAs have been shown to be better silencers compared to 21-nt siRNAs with identical base pair composition, possibly due to Dicer mediated incorporation of the siRNAs into the RISC (Kim *et al.* 2005). Also, Ago2 PIWI domain has been shown to bind directly with the RNase III domain of Dicer (Tahbaz *et al.* 2004). These results strongly support that Dicer, though not essential for RNAi, can play a role in facilitating siRNA processing and RNAi. This is further supported by *in vitro* studies that showed that Dicer alone is capable of binding 21-mer siRNAs (Pellino *et al.* 2005; Kini and Walton 2007).

Dicer processing of long dsRNA substrates without mismatches is much more efficient than substrates resembling pre-miRNAs, with internal mismatches (Soifer *et al.* 2008). The Dicer helicase domain attenuates the processing of substrates with internal mismatches (Ma *et al.* 2008). In the case of pre-miRNA processing, the helicase domain has been suggested to mediate a conformational change of the enzyme to facilitate product release (Ma *et al.* 2008; Soifer *et al.* 2008). Dicer may therefore have the ability to bind with ssRNAs, and be functional in the bypass mechanism suggested in the formation of active RISC programmed with single stranded miRNAs (Matranga *et al.* 2005).

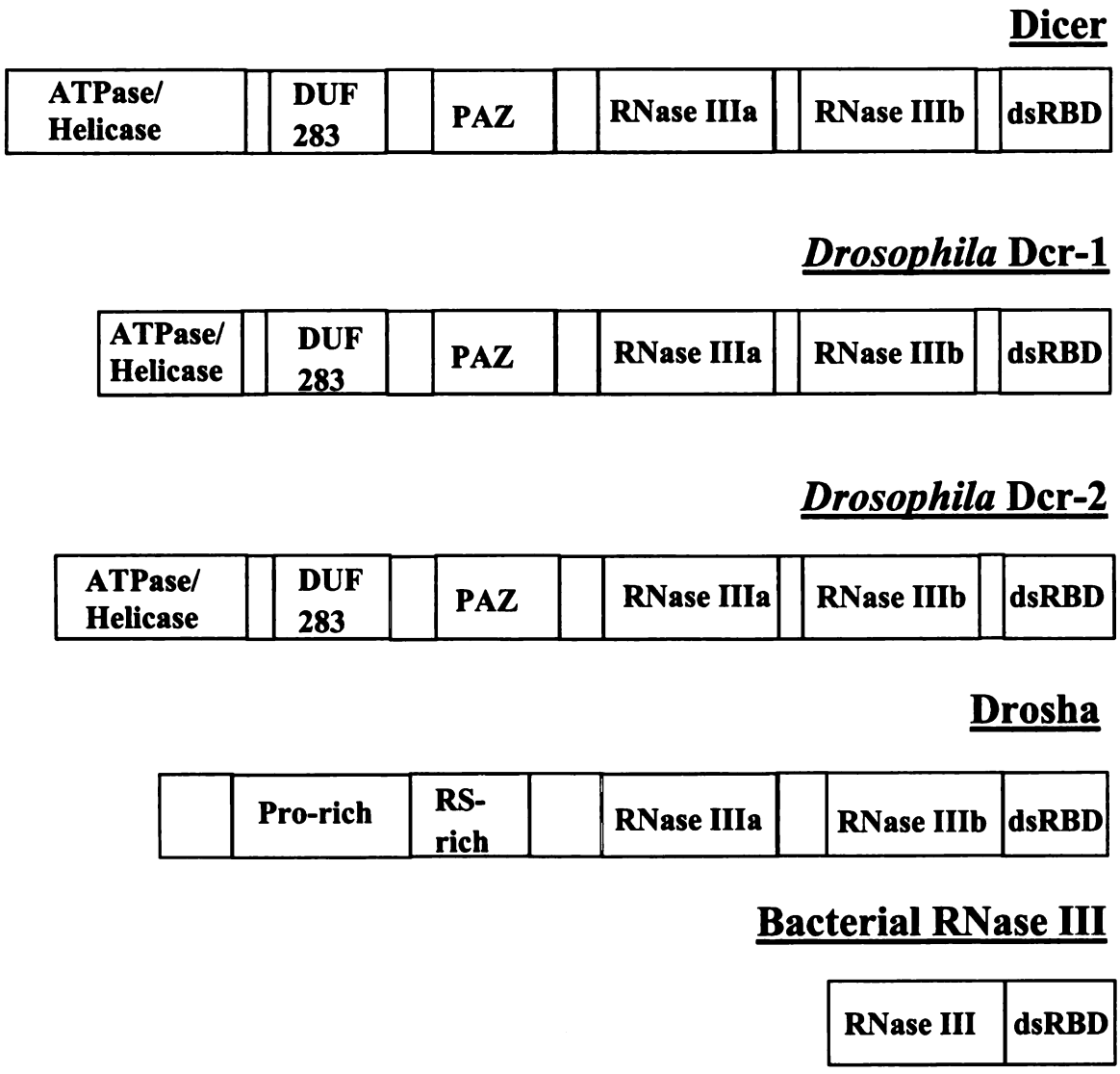


Figure 2.1 Representation of the various domains of RNase III family proteins.

RESULTS AND DISCUSSION

SsRNA-Dicer complex formation *in vitro*

Dicer is known to form stable complexes with dsRNAs and siRNAs (Provost *et al.* 2002; Zhang *et al.* 2002; Pellino *et al.* 2005), though for siRNAs some conflicting reports exist (Provost *et al.* 2002; Chendrimada *et al.* 2005). The focus of this study was to determine if Dicer alone could bind 21-nt ssRNAs. SsRNAs of this particular length were chosen as they are the same length as the individual strands of a siRNA duplex. Also 21-nt ssRNAs have been shown to be functional silencers in *Drosophila* (Schwarz *et al.* 2002).

For the binding studies, commercially available recombinant human Dicer was used. Silver staining and western blot analysis with monoclonal Dicer antibody for Dicer were done to confirm the size of Dicer protein (~ 210 kDa) and its identity, respectively (Fig. A1, A and B). Dicer functionality was confirmed by cleavage reactions with a 27-mer dsRNA substrate (Fig. A1 C).

To study whether Dicer would also form stable complexes with 21-nt ssRNAs, binding reactions were set up with 5'-^{33}P labeled ssRNAs and Dicer (Fig. 2.2A). As a control, complex formation between Dicer and an siRNA was verified (Fig. 2.2A, lane 2). Two bands were seen in the case of siRNA-Dicer binding under identical binding conditions, suggesting the possible presence of ssRNA in the siRNA preparation. Faster migration of the siRNA-Dicer complex relative to the ssRNA-Dicer complex is due to the increase in negative charge on the complex (due to the backbone phosphates of the second strand) with a relatively small change in molecular size/weight (Wu and Regnier 1993).

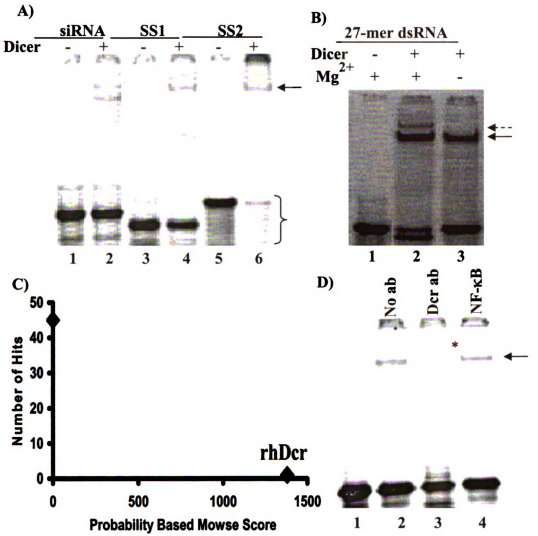


Figure 2.2 *In vitro* binding of Dicer with siRNA and ssRNAs.

(A) Electrophoretic mobility shift assay of the complex formed between Dicer and, siRNA (lane 2), structured ssRNA SS1 (lane 4), and unstructured ssRNA SS2 (lane 6). The arrow indicates the position of the Dicer-ssRNA complex. Curly brace indicates unbound siRNAs and ssRNAs. (B) Electrophoretic mobility shift assay of Dicer-27-mer dsRNA complex in the presence of Mg^{2+} (lane 2) and in the absence of Mg^{2+} (lane 3). The dotted arrow denotes the Dicer-siRNA complex whereas the solid arrow denotes the Dicer-27-mer dsRNA complex. (C) Mass spectrometry of the protein-ssRNA complex denoted by the arrow in (A) showed only the presence of Dicer in the complex. (D) Electrophoretic mobility shift assay of Dicer-SS2 complex in the presence of Dicer antibody (Lane 3) and antibody against NF-κB (Lane 4). The asterisk denotes the supershifted complex.

27-mer dsRNA substrate were carried out. After processing and binding, two bands remained, indicating Dicer complexes with the 27-mer duplex and the 21-mer siRNA product (Fig. 2.2B lane 2). The identities of these bands were confirmed by processing in the absence of divalent cations, which are not required for Dicer binding but are necessary for cleavage (Fig 2.2B, comparing lane 2 and 3) (Zhang *et al.* 2002). This experiment also provides evidence that Dicer forms stable complex with the siRNAs it generates, and may thereby enhance their incorporation into the RISC complex and also their functionality (Kim *et al.* 2005; Siolas *et al.* 2005) .

Because Dicer is naturally a dsRNA-binding protein, the secondary structure, or structure formed by intramolecular hybridization, of an ssRNA was expected to impact the ability of Dicer to bind it. So an ssRNA predicted to have secondary structure at 4°C (SS1) and a polyuridine ssRNA (SS2) that is presumably unable to form any significant secondary structure were tested. The reduced electrophoretic motility of unbound SS2 relative to unbound SS1 suggests the general absence of structure in SS2 (Fig. 2.2A, compare lanes 3 and 5), supporting the expectations from the T_M calculations (Table 2.1). Both SS1 and SS2 are bound stably by Dicer (Fig. 2.2A, lanes 4 and 6), as are other 21-nt ssRNA sequences (Table 2.1). Identical ssRNA-Dicer complexes were also formed with 12 and 15-nt ssRNAs (data not shown).

The ssRNA-Dicer complex formed by SS1 with Dicer was analyzed by mass spectrometry to confirm the presence of Dicer as the only protein component (Fig. 2.2C). To confirm further that the complex contained Dicer, an antibody supershift assay was performed. Adding Dicer antibody subsequent to Dicer-RNA binding did not produce

Table 2.1 Predicted ΔG and T_M for the sequences predicted using mfold (Zuker 2003).

Name	Sequence	ΔG, kcal/mole	T_M, °C
SS1	GUC ACA UUG CCC AAG UCU CTT	0.7	20
SS2	UUU UUU UUU UUU UUU UUU UTT	-	-
SS3	GCU AAA AAA AAA AAA AAA ATT	1.94	-19.3
SS4	GUC AAA AAA AAA AAA AAA ATT	1.94	-19.3
SS5	GCU GAC CCU GAA AUU GAU CTT	-0.61	32.2
SS6	GAG ACU UGG GCA AUG UGA CUT	-1.62	36.7

any shift in the complex, so assays were performed by prior incubation of the antibody with Dicer followed by addition of SS1. This implies that this antibody is a competitive inhibitor of the binding by Dicer of ssRNAs. In the presence of the Dicer antibody, the characterized complex is nearly completely retarded as compared to negative and non-specific antibody controls (Fig. 2.2D, compare lane 3 with lanes 2 and 4).

Possible contribution of the PAZ domain to ssRNA binding by Dicer

SiRNAs and miRNAs resulting from Dicer processing have characteristic 5'-phosphate and 3' overhangs (Fig 2.3B). Chemically synthesized siRNAs possess a 5'-hydroxyl (Fig 2.3C). Upon entry into cells, these siRNAs are immediately phosphorylated (Fig 2.3B) by the human RNA kinase hClp1 (Nykanen *et al.* 2001; Weitzer and Martinez 2007). SiRNAs with a chemical modification that prevents 5' phosphorylation are bound by Dicer with significantly lower affinity (Pellino *et al.* 2005; Chen *et al.* 2008). The presence of a 5' phosphate is an important structural determinant for the guide strand to be incorporated into RISC (Liu *et al.* 2004). Structural studies of the Ago2 PIWI domain have shown that a conserved metal binding site secures the 5'-phosphate

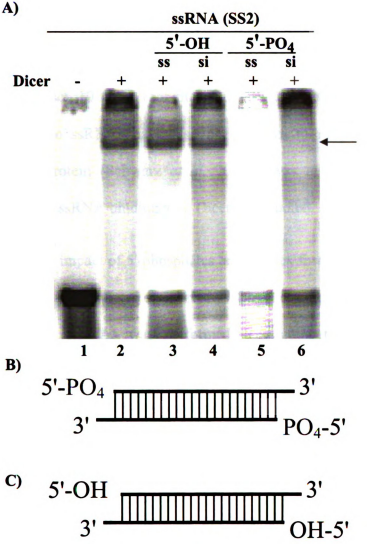


Figure 2.3 Impact of 5'-phosphate on Dicer binding affinity for ssRNAs and siRNAs.

Dicer-ssRNA binding reactions were performed in the presence of 100-fold excess unlabeled ssRNA (lanes 3 and 5) or siRNA (lanes 4 and 6). In lanes 3 and 4, competition was performed with 5'-hydroxyl RNAs. In lanes 5 and 6, competition was performed with 5'-phosphate RNAs. The arrow indicates the position of the Dicer-ssRNA complex. B) Schematic of siRNA structure with 5' –phosphate. C) Schematic of siRNA structure with 5'-hydroxyl.

(Parker *et al.* 2004; Parker *et al.* 2005; Wang *et al.* 2008). Anchoring of the 5' end is an important step that determines the position of the target mRNA to be cleaved. Following the formation of this complex, the target mRNA is cleaved between the nucleotides paired to bases 10 and 11 of the guide strand (Elbashir *et al.* 2001b; Elbashir *et al.* 2001c). Since ssRNAs have also been shown to induce silencing mediated by the Argonaute protein (Schwarz *et al.* 2002; Rivas *et al.* 2005), the impact of the 5'-phosphate on ssRNA binding with Dicer was studied.

To assess the impact of 5'-phosphates on complex formation between ssRNAs and Dicer, a competition assay between labeled ssRNAs in the presence of an excess of unlabeled ssRNAs and siRNAs was performed. The competing siRNAs and ssRNAs were chemically synthesized having a 5'-hydroxyl (Fig. 2.3C). At 100-fold excess concentration, neither 5'-hydroxyl ssRNAs nor 5'-hydroxyl siRNAs could displace 5'-phosphate ssRNA SS2 from its complex with Dicer (Fig. 2.3A, lanes 3 and 4). Then, the competing siRNAs and ssRNAs were 5'-phosphorylated using cold ATP (Fig. 2.3B). These phosphorylated siRNAs and ssRNAs effectively displaced bound SS2 (Fig. 2.3A, lanes 5 and 6). This suggests a role for Dicer in quality control of the molecules entering the RNAi pathway, by verifying the presence of 5'-phosphate on the silencing agents. Of the domains in Dicer, only the PAZ domain is known to possess a greater affinity for 5'-phosphate ssRNAs as compared to 5'-hydroxyl ssRNAs (Yan *et al.* 2003). While the 5'-phosphate determines whether the siRNAs and ssRNAs can enter the silencing pathway, the 3'-structure affects the activity of the siRNAs and dsRNAs. DsRNAs with 2-nt 3' overhangs are processed with higher efficiency by Dicer compared to blunt end duplexes, and this interaction is mediated by the Dicer PAZ domain (Zhang *et al.* 2004). Similarly

siRNAs with 3' overhangs have been shown to be better silencers compared to blunt end duplexes (Elbashir *et al.* 2001c). The PAZ domain, containing a conserved oligonucleotide/oligosaccharide binding fold, anchors the 3' overhangs of the duplexes (Song *et al.* 2003).

To explore if the higher affinity for 5'-phosphate substrates seen in ssRNA-Dicer binding was due to the PAZ domain, binding to SS2 with a 5'-phosphate ssRNA and a 3'-biotin was tested. It is known that this modification disturbs PAZ domain-RNA interactions (Yan *et al.* 2003). It was found that, Dicer has a lower affinity for this sequence relative to 3'-hydroxyl SS2 (Fig. 2.4), further suggesting a role for the PAZ domain in binding to the ssRNAs.

5' terminal nucleotide sequence dependence of Dicer binding

The Dicer-ssRNA binding experiments in the presence of competing siRNAs showed that the 5'-phosphate is preferred for binding with siRNAs and ssRNAs. Also siRNAs possessing a 5'-methoxy modification at one end of the duplex, which prevents the phosphorylation of that end, leads to the preferential selection of the other strand as the guide strand (Chen *et al.* 2008). These results indicate that Dicer interacts with the 5' end of both the ssRNAs and siRNAs. Also in *Drosophila*, it has been shown that during the formation of the RLC, TRBP binds at the 3' end of the guide strand and Dicer 2 binds to the 5' end of the guide strand. Dicer processing of long dsRNAs has been shown to be dependent on the terminal nucleotide sequence (Vermeulen *et al.* 2005). DsRNA substrates possessing an A at the terminal end were processed much more efficiently than those having a G or C.

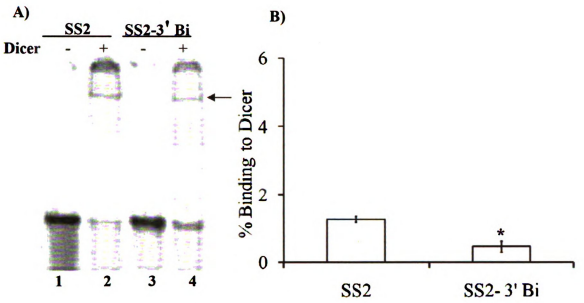


Figure 2.4 Dicer has lower affinity for ssRNAs having a 3'-biotin.

(A) Representative figure showing Dicer-ssRNA binding with ssRNA having a 3'-hydroxyl (lane 2) and ssRNA having a 3'-biotin (lane 4). B) Comparison of binding of Dicer to ssRNA having a 3'-biotin to that with a 3'-hydroxyl . Band intensities from lanes 2 and 4 normalized to the corresponding controls in lane 1 and 3 respectively. *, denotes the 2-tailed t-test comparison of Dicer binding to SS2 vs. SS2 -3' Bi for $p < 0.05$.

Therefore, it was tested whether terminal sequence, in addition to terminal structure, was important for Dicer binding to ssRNAs and siRNAs.

To look specifically at the influence of the terminal nucleotides on binding with Dicer, ssRNAs lacking stable secondary structures but with varying terminal nucleotide sequence at the 5' end were designed and tested (Table 2.2). It was observed that Dicer binding to these sequences varied with the terminal nucleotide sequence (Fig. 2.5). SS2-AA shows almost two fold higher binding with Dicer than the other sequences. These

results indicate that the Dicer interaction with ssRNAs is dependent on the terminal nucleotide sequence.

Table 2.2 Predicted ΔG and T_M for the sequences predicted using mfold (Zuker 2003).

Name	Sequence	ΔG, kcal/mole	T_M, °C
SS2	**UUU** UUU UUU UUU UUU UUU UTT	-	-
SS2-CC	**CCU** UUU UUU UUU UUU UUU UTT		
SS2-GG	**GGU** UUU UUU UUU UUU UUU UTT	2.5	-17.2
SS2-AA	**AAU** UUU UUU UUU UUU UUU UTT	2.6	-17.9

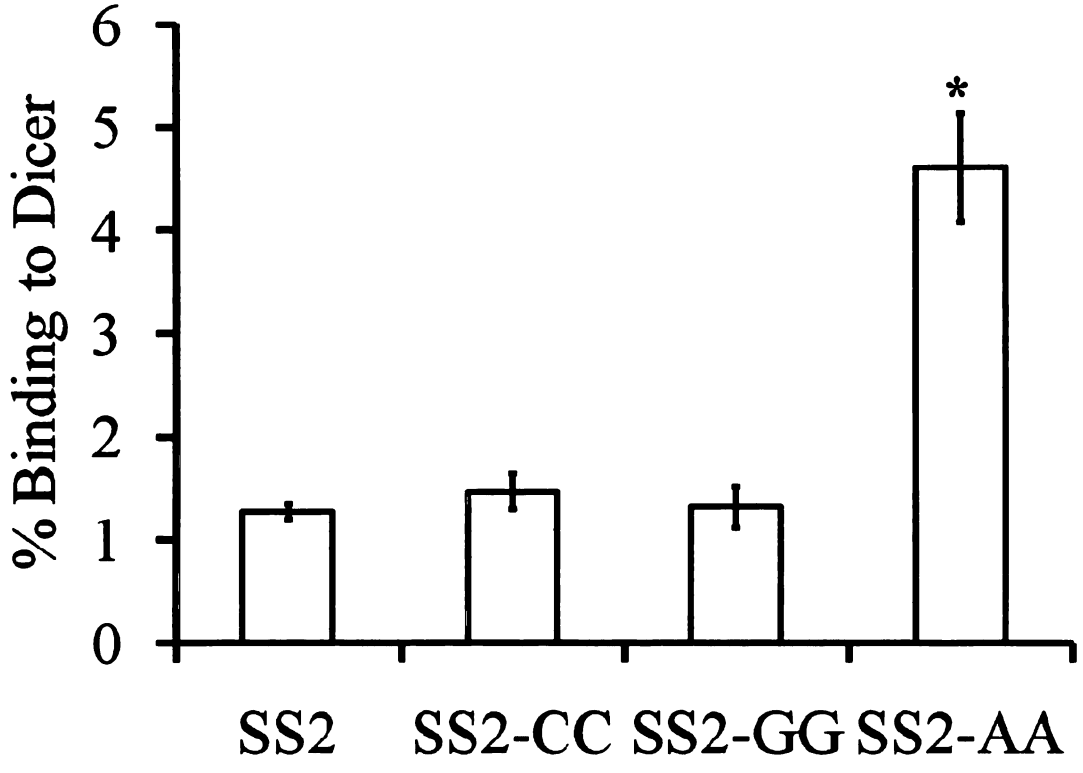


Figure 2.5 Dicer ssRNA binding is dependent on the terminal nucleotide sequence.

Percentage binding was calculated by normalizing intensity of siRNA-protein complex to the respective unbound ssRNA. *, denotes the 2-tailed t-test comparison of Dicer binding to different ssRNAs vs. SS2 for $p < 0.05$.

Divalent cation dependence of ssRNA-Dicer complex formation

While Dicer requires the presence of Mg^{2+} to be catalytically-active (Zhang *et al.* 2002), it has been shown to form stable complexes with 100-130 bp dsRNAs even in the absence of Mg^{2+} (Provost *et al.* 2002; Zhang *et al.* 2002). Structural studies with the *Aquifex aeolicus* RNase III has shown dsRNAs in a non-catalytic complex with the RNase protein, which is structurally different from the catalytic assembly (Blaszczyk *et al.* 2001; Blaszczyk *et al.* 2004). The PIWI domain of Ago2 has been shown to interact directly with the RNase domain of Dicer (Tahbaz *et al.* 2004). These studies show that the Dicer and other RNases can attain alternate structural conformation in a non-catalytic assembly and function primarily as dsRNA binding proteins in the RNAi pathway, and shuttle the silencers to the downstream proteins. In this regard, the effect of divalent cations on the interactions of 21-nt ssRNAs and 21-mer siRNAs with the human Dicer was studied.

To test if ssRNA-Dicer complex formation depends on the presence of divalent cations, binding reactions were performed in Mg^{2+}-free buffers. In the absence of Mg^{2+}, Dicer did not stably bind either ssRNA (Fig. 2.6A, lanes 6 and 9, arrow). Complex formation was restored when Mn^{2+} or Ca^{2+} were added to the buffer but not Co^{2+}, Ni^{2+}, or Zn^{2+} (Fig. 2.6B, solid arrow). Both in the presence and absence of divalent cations, ssRNAs appear to form smaller complexes (Fig. 2.6B, lane 2 and 8, dotted arrow), but this binding could not be assigned to a specific protein after analysis by mass spectrometry (Fig. A2). Similar divalent cation dependence was not observed for the formation of a stable siRNA-Dicer complex (Fig. 2.6A, lane 3). Comparing lanes 2 and 3 (Fig 2.6A),

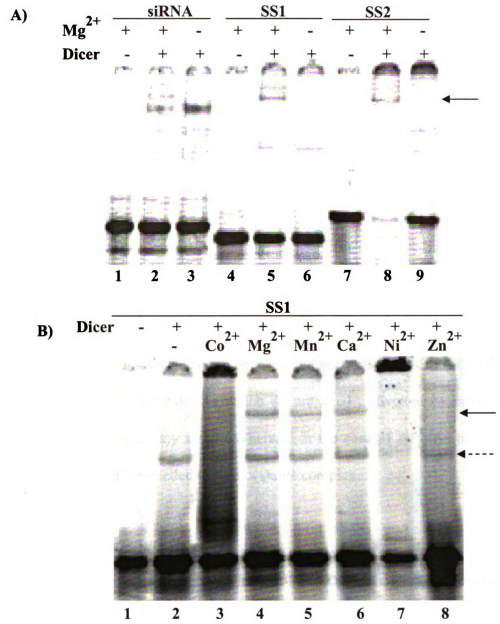


Figure 2.6 Divalent cation dependence of Dicer-ssRNA complex formation.

(A) Dicer-ssRNA complexes formed in the presence of Mg^{2+} (lanes 5 and 8) and not in its absence (lanes 6 and 9). SiRNA-Dicer complex formation was not cation dependent (lanes 2 and 3). (B) Dicer-ssRNA(SS1) complex formation can occur in the presence of Mg^{2+},Mn^{2+}, and Ca^{2+} but not Co^{2+}, Ni^{2+}, or Zn^{2+}.

lane 3 does not have the second slower moving complex seen at the top of the gel in lane 2. This is further proof to the presence of ssRNA in the second slower complex as it is not seen in the absence of Mg^{2+}.These results indicate that different Dicer domains are involved in forming stable complexes with the siRNAs vs. the ssRNAs. Beyond the recognition of the 5'-phosphate by the Dicer PAZ domains, other Dicer domains may be needed to form stable complexes with either the siRNAs or the ssRNAs. For *E. coli* RNase III, Mg^{2+} and Ca^{2+} act to stabilize complex formation of the enzyme with the bacteriophage T7 R1.1 RNA, which has a hairpin structure (Li and Nicholson 1996). Divalent cations not involved in catalytic activity have been shown to stabilize the interactions between the RNase and the phosphate backbone of the RNA (Ji 2006). In the case of siRNAs and other dsRNAs which do not need a divalent cation to bind with Dicer, the dsRBD might play a role. Whereas, in the case of ssRNAs one or both the RNase domains might be needed to form a stable complex.

SsRNA binding by *Giardia* Dicer

While the human Dicer is a relatively large (~210 kD), the *Giardia* Dicer consisting of 756 amino acids is one of the smallest Dicer proteins known (Macrae *et al.* 2006). It consists of one RNA binding domain, the PAZ domain, and two RNase III domains (Fig. 2.7A). Crystal structure of this enzyme with a dsRNA has been used to demonstrate the details of substrate processing by Dicer. The two metal ion containing RNase III domains secure the two scissile phosphates of the dsRNA duplex and cleave the substrate ~25 bp from the end secured by the PAZ domain. It was reasoned that, if the PAZ domain in

human Dicer contributes to the binding of ssRNAs, then *Giardia* Dicer which also has the PAZ domain should also be able to bind ssRNAs.

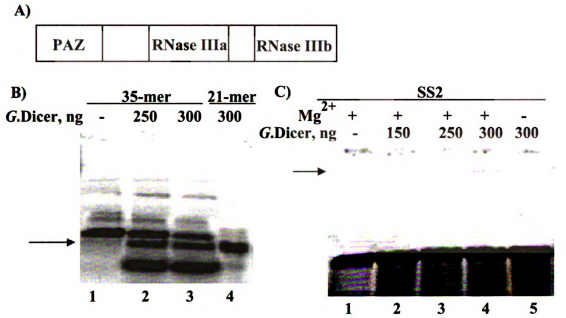


Figure 2.7 Schematic of the *Giardia* Dicer domain and its *in vitro* binding with ssRNAs.

(A) Domains of *Giardia* Dicer. B) Catalytic activity of *Giardia* Dicer. Cleavage products of *Giardia* Dicer enzyme, analyzed by native gel electrophoresis. Arrow indicates the ~25mer siRNA generated (lanes 2 and 3) from cleavage of 35mer dsRNA substrate (lane 1). C) Electrophoretic mobility shift assay of the complex formed between *Giardia* Dicer and SS2 in the presence (lanes 2-4) and absence of Mg^{2+} (lane 5). Arrow indicates the position of the *Giardia* Dicer-ssRNA complex.

Interestingly while the human Dicer can process 25 and 27-mer dsRNA substrates, the *G.* Dicer cannot process substrates of these lengths (data not shown). Hence the RNase activity of the *G.* Dicer was tested by the processing of a 35-bp substrate into ~25-mer siRNAs (Fig. 2.7B). Then, *in vitro* binding reactions with ssRNAs and *Giardia* Dicer were performed. Similar to human Dicer, *Giardia* Dicer also binds ssRNAs in Mg^{2+}

dependent manner *in vitro* (Fig. 2.7C, lanes 2-4). In the presence of EDTA, no stable Dicer-ssRNA complex is formed (Fig. 2.7C, lane 5).

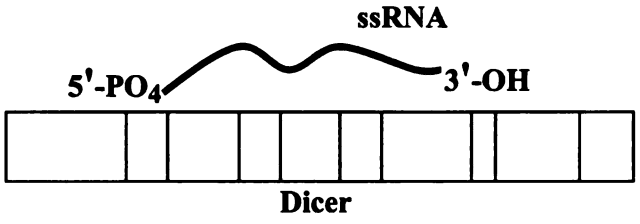
Identifying the Dicer domain/s involved in binding with the ssRNA and siRNAs

The binding studies showed that human Dicer can bind both siRNAs and ssRNAs. These interactions seemed to involve the PAZ domain and possibly other domains as well. The binding studies with the *Giardia* Dicer further pointed to the involvement of the PAZ domain. Binding studies in the absence of divalent cations discriminated between binding with siRNAs vs. ssRNAs, indicating the involvement of other Dicer domains in each case. The goal of these studies was to identify the Dicer domain directly in contact with the nucleic acid using photo crosslinking and mass spectrometry.

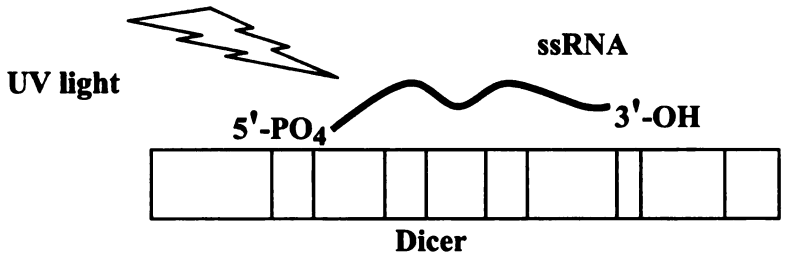
Photo crosslinking has been commonly used to study protein-nucleic acid complexes in diverse cellular processes such as transcription, translation, and DNA replication. UV crosslinking induces the formation of zero-length covalent bonds where protein amino acids and nucleic acid directly contact each other (Shetlar 1980; Bennett *et al.* 1994). These crosslinked complexes can then be used for structural analysis, characterization of the protein or nucleic acid. Mass spectrometry which provides better sensitivity and specificity compared to traditional methods such as Edman degradation (Merrill *et al.* 1984; Prasad *et al.* 1993) has been widely used to determine the protein molecular weight or sequencing the amino acids (Golden *et al.* 1999; Rieger *et al.* 2000).

A schematic of the experimental setup for the binding site studies is shown in Figure 2.8. After the binding reactions, the samples were UV crosslinked and proteolysed by Trypsin. This series of experiments was designed to identify the peptide fragments

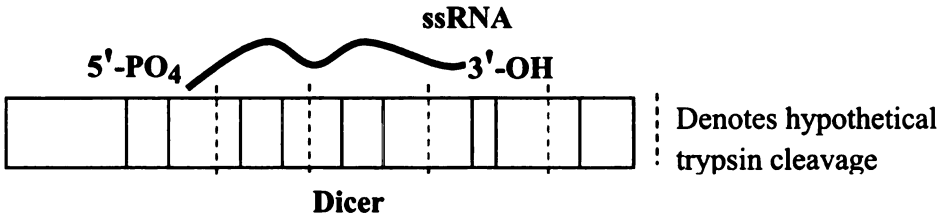
Step 1: Dicer-ssRNA binding reaction



Step 2: UV cross-linking the Dicer-ssRNA complex



Step 3: Proteolysis of UV cross-linked Dicer-ssRNA complex using Trypsin



Step 4: Liquid chromatography-mass spectrometry of proteolysed sample

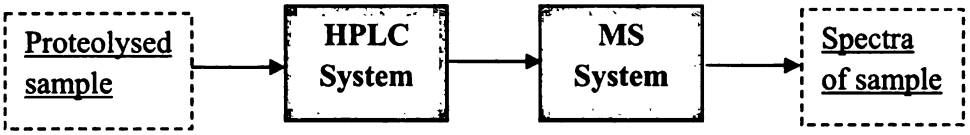


Figure 2.8 Sequence of steps to determine the Dicer domain/s binding with the ssRNA and siRNA

which are crosslinked to the ssRNA or the siRNAs. Such nucleic acid bound peptides would have a different mass to charge ratio (m/z), relative to the identical peptide from the Dicer protein alone. Due to the difference in 'm/z' the nucleic bound peptide would have a different time of flight than the peptide alone. Overlaying the spectra of the Dicer protein alone with the spectra of either Dicer-siRNA or Dicer-ssRNA would reveal the 'm/z' of such peptides. The amino acid sequence of these peptides can then be inferred from the corresponding predicted amino acid sequence (Appendix A, lists details of generating the predicted cleavage map) for the particular 'm/z', and thereby identify the Dicer domain binding to the nucleic acid.

Comparison of mass spectrometry generated spectrum peaks with predicted peaks for BSA

The experimental setup was tested by proteolysing BSA (Bovine serum albumin) protein using Trypsin. Mass spectrometry data was analyzed using Waters MassLynx software. Multiple charge peaks (m/z) from the BSA peptide fragments were compared to the multiple charge peaks predicted in the cleavage map of BSA for tryptic digest (details in Appendix A) for confirming the tryptic digest of BSA. Figures A3.A and A4.A list two of the doubly charged peaks of BSA peptides and the corresponding amino acid sequence (Fig.A3.B and A4.B).

Data analysis and predicting the Dicer domain Dicer-ssRNA complex and Dicer

Mass spectrometry analysis of the typtic digests of Dicer, Dicer-ssRNA and Dicer-siRNA was then performed. To maintain the same conditions for both the samples the Dicer

protein was also UV treated for 10 min prior to trypsinization. Peptide fragments were run through the HPLC column for 75 min and analyzed by the Q-TOF Ultima API system to yield the chromatograms for Dicer protein and Dicer-ssRNA complex (overlaid spectra Fig. 2.9A, individual chromatograms in Appendix A). The spectra from the overlaid chromatograms were analyzed, to identify similar Dicer peptide peaks which are shifted from the fragment peaks of the Dicer-ssRNA complex due to crosslinking to the ssRNA. Fig. 2.9B shows such a peak identified from the overlaid spectra. The spectrum of the Dicer-ssRNA fragments does not have a m/z peak corresponding to the Dicer protein peak of 917.8 (Fig.2.9B, denoted by the arrow). However there are no corresponding peaks, with these m/z predicted by the theoretical cleavage map of the Dicer protein (Fig. 2.9C). This could be due to incomplete cleavage of the protein or due to the presence of the detergent Triton in the Dicer preparation. Cleavage of the Dicer protein was conformed by the presence of several doubly and triply charged peaks indicative of typtic digest. Presence of Triton is the probable cause of variation in the m/z of the peptide fragments as pure BSA protein digested under identical condition had peptide fragments with the same m/z as predicted by the theoretical BSA cleavage map (Fig. A3 and A4). Detergents like Triton have been shown to interfere with peptide ionization in mass spectrometry (Bornsen *et al.* 1997; Norris *et al.* 2005). Removing the detergent from the samples was not a feasible option considering the low yields and microgram amount of protein required to be analyzed by LC followed by mass spectrometry (using Q-TOF).

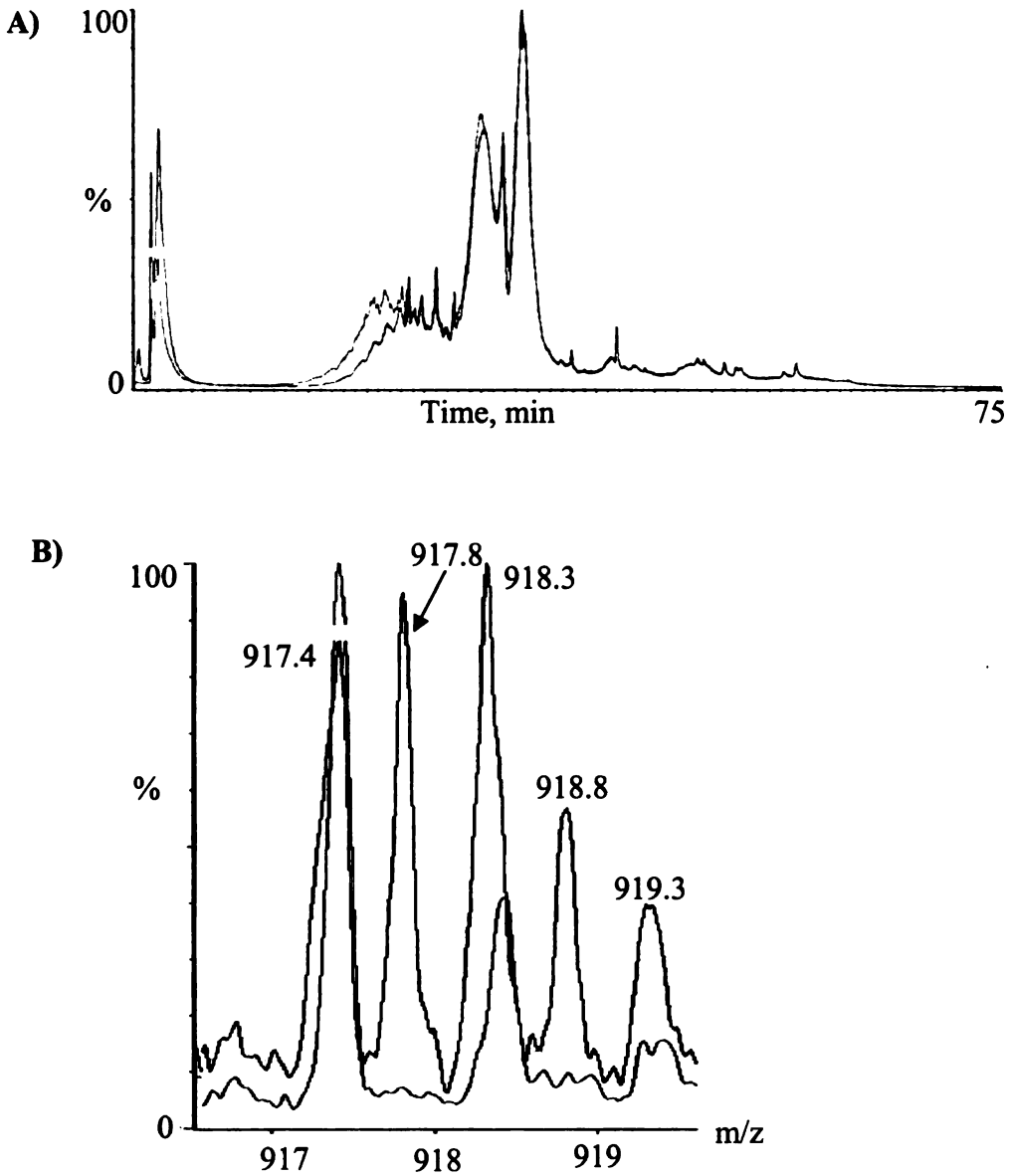


Figure 2.9 LC/MS chromatograms

A) Overlaid chromatograms of trypsin digested Dicer protein and Dicer-SS complex. B) Spectra of overlaid chromatograms from A) displaying the m/z vales for the Dicer-ssRNA fragments. C) Few predicted peptides resulting from tryptic digest of the Dicer protein, with m/z around 917.

C)

m/z (mi)	m/z (av)	Star	End	Sequence
916.441	917.0122^{+2}	1457	1472	ISLSPFSTTDSAYEWK
917.466	917.0654^{+3}	418	440	EKPETNFPSPFTNILCGIIPVER
917.843	918.4029^{+3}	1716	1741	QHSPGVLTDLRSALVNNTIFASLA
917.984	918.5891^{+2}	545	560	APISNYIMLADTDKIK

Figure 2.9 (Continued)

Identifying the Dicer domains binding siRNA and ssRNA by comparison of the triple spectra

Based on the known roles of RNA domains in Dicer, the PAZ and dsRBD possibly mediated the binding with siRNAs and likewise PAZ and RNase III domains might interact with the ssRNAs. While PAZ domain discriminates between phosphorylated and non-phosphorylated substrates for Dicer binding, it can bind with both siRNAs and ssRNAs having a 5'-phosphate. However different Dicer domain/s are involved in forming a stable complex with either the siRNAs or the ssRNAs. Results from the binding studies support this observation with ssRNAs requiring the presence of a divalent cation to bind with Dicer suggesting the involvement of RNase III domains. SiRNAs on the other hand bind with Dicer even in the absence of divalent cations suggesting the RNase III domains may not be needed in binding with the siRNAs. Also, the Dicer dsRBD, which has been shown to bind only with a dsRNA and not with ssRNA, RNA-

DNA hybrids or dsDNA, might mediate binding with the siRNAs (Bevilacqua and Cech 1996).

As unique Dicer domain/s are involved in the binding with the ssRNAs or the siRNAs. Overlaying and comparing the spectra for all the three cases: Dicer protein, Dicer-siRNA complex, and Dicer-ssRNA complex, would differentiate the Dicer domains involved in forming a stable complex with the siRNA or the ssRNA. Also comparing the three spectra with each other gives more confidence in identifying the right peaks by avoiding non-specific variation in the spectral peaks. Figure 2.10A shows the overlaid spectra for the three cases indicating the peaks of the ~ 917 m/z. While both the Dicer protein (Fig. 2.10A) and Dicer-siRNA complex (Fig. 2.10B) have a peak corresponding to m/z of 917.8, the Dicer-ssRNA complex does not (Fig. 2.9B). The predicted peak with the m/z of 917.8 has the amino acid sequence of QHSPGVLTDLRSALVNNTIFASLAVK. This sequence corresponds to the Dicer domain RNase-IIIb (Table 2.3). The other predicted peak with m/z value of 917.4 with amino acids from 418-440 (Fig 2.9C), was ignored as it does not correspond to one of the characterized domains of Dicer.

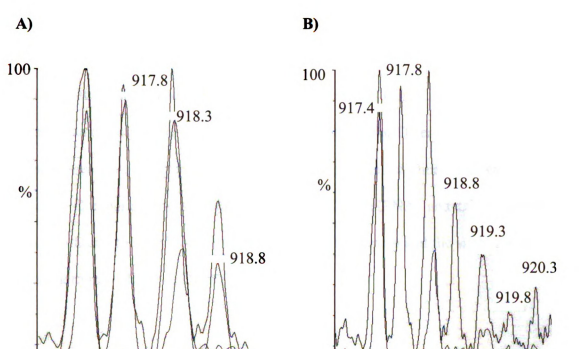


Figure 2.10 Overlaid spectra of proteolysed Dicer, Dicer-siRNA and Dicer-ssRNA.
The m/z values displayed are for A) the Dicer-siRNA fragments. and B) for the Dicer protein peptides.

Table 2.3 Amino acid sequence of human Dicer domains.

List of the amino acid sequence for the PAZ, dsRBD, RNase-IIIa and RNase-IIIb domains of Dicer. The bold underlined amino acids refer to the sequences indentified by mass spectrometry analyses to be crosslinked with either with the siRNA or the ssRNA.

Domain	Amino acid sequence
PAZ (123 aa)	dsstldidfk fmedieksea rigipstkyt ketpfvfkle dyqdaviipr yrnfdqphrf yva**dvytdlt** plskfpspey etfaeyyktk ynldltnlnq plldvdhtss rlnlltprhlnqk
dsRBD (66 aa)	vpr**spvrellemepetak**fs paertydgkv rvtvevvgkg kfkgvgrsyr iaksaaarralrslka
RNase-IIIa (128 aa)	**dseqspsigy** ssrtlgpnpg lilqaltlsn asdgfnlerl emlgdsflkh aittylfctypd**ahegrlsy** mrskkvsncn lyrlgkkkgl psrmvvsifd ppvnwlppgy vvnqdksntdkwekdemt
RNase-IIIb (159)	fenfekkiny rfknkayllq afthasyhyn titdcyqrle flgdaildyl itkhlyedpr**ghspgvltdl rsalvnntif aslavky**dyh kyfkavspel fhviddfvqf qleknemqgm dselrrseed eekeedievp kamgdifesl agaiymdsg

One of the other shifted peaks identified was from the spectra of Dicer-siRNA complex, relative to the other two spectra, had m/z of 590.3 (Fig. 2.11A-C). The corresponding nearest theoretical peaks had m/z values of 590.6 and 590.9 (Fig. 2.11D). The peak 590.9 with amino acid ranging from 1842-1856 corresponds to the Dicer dsRBD, whereas the other peak does not correspond to any characterized domains of Dicer, and was hence not considered. Comparing the spectra of the Dicer, Dicer-siRNA, and Dicer-ssRNA the ssRNA was found to crosslink to the Dicer RNase IIIb domain and the siRNA was found to crosslink to the dsRBD of Dicer (Table 2.3).

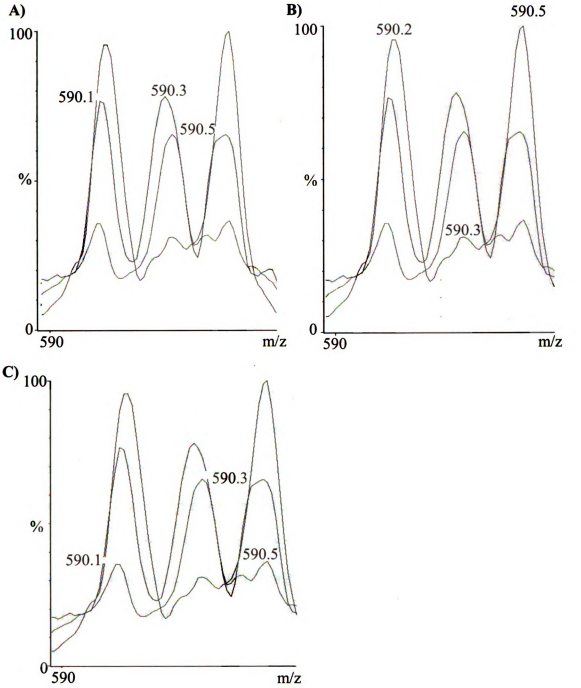


Figure 2.11 Overlaid spectra of proteolysed Dicer, Dicer-siRNA and Dicer-ssRNA.

The m/z values displayed are for A) the Dicer protein B) Dicer-siRNA fragments. and C) for the Dicer-ssRNA fragments. D) Few predicted peptides resulting from tryptic digest of the Dicer protein, with m/z around 590.6.

D)

m/z	m/z	Start	End	Sequence
590.6695^{+5}	590.6695^{+5}	1638	1661	CMFDHPDADKTLNHLI SGFENFEK
590.9696^{+3}	591.3515^{+3}	1842	1856	**SPVRELLEMEPETAK**
591.2849^{+2}	591.6683^{+2}	1081	1089	YPNLDFGWK
591.2984^{+2}	591.6424^{+2}	1091	1095	SIDSK

Figure 2.11 (Continued)

CONCLUSIONS

PAZ domain mediated ssRNA and siRNA binding by Dicer

A previously unknown function of the human Dicer has been demonstrated, in its ability to bind ssRNAs *in vitro*, independent of the internal sequence and secondary structure of the ssRNA. In all experiments, the ssRNAs tested were far shorter than hairpin structures with stems of at least 29 bp in length that are known to be Dicer substrates (Siolas *et al.* 2005). Nonetheless, the results from the competition experiments showed that Dicer recognition and binding to both siRNAs and ssRNAs maybe mediated by the Dicer PAZ domain. In the case of siRNAs and long dsRNAs with overhangs, both the RNA binding domains of Dicer, the PAZ domain and the dsRBD, are involved in binding the substrate (Zhang *et al.* 2004). The PAZ domain binds to the overhangs and positions the dsRBD and concomitantly the RNase III domains (Zhang *et al.* 2002; Pellino *et al.* 2005). It has been reported that the dsRBD can only bind dsRNAs (Bevilacqua *et al.* 1998; Ryter and Schultz 1998). 5'-phosphate dependence of the 21-nt ssRNAs for binding with Dicer

shown in this research is identical to the binding behavior of the 5-nt ssRNAs to the Ago2 PAZ domain (Yan *et al.* 2003). This suggests that ssRNA complex formation with Dicer is mediated at least in part by the PAZ domain.

Divalent cation dependence of ssRNA-Dicer complex

Dicer and other RNase III family enzymes need one Mg^{2+} ion for each of the RNase domains to be catalytically active (Sun *et al.* 2005). Moreover, recent structural evidence suggests that two metal ions may be required for maximum activity (Ji 2006; Macrae *et al.* 2006). The second metal ion is suggested to bind to the highly negative catalytic valley of RNase III enzymes and thereby stabilize the interaction between the enzyme and the negatively charged phosphate backbone of substrate RNA (Ji 2006). Stable binding of ssRNAs by Dicer depends upon the presence of divalent cations. Similarly divalent cations are also required by the *Giardia* Dicer for forming a stable complex with the ssRNAs. It may therefore be that the divalent cation dependence seen for ssRNA binding is related to an interaction of the RNase domains of Dicer with the ssRNAs. Such dependence would be akin to that seen in the formation of a stable Dicer-ssRNA complex for the *Giardia intestinalis* Dicer to its dsRNA substrate (Macrae *et al.* 2006). This effect may be more pronounced in Dicer binding due to its two RNase III domains. Each of these may simultaneously, and possibly cooperatively, contribute to the affinity of Dicer for ssRNAs.

Dicer domains stabilizing binding with ssRNAs and siRNAs

The PAZ domain mediates interactions with both siRNAs and ssRNAs through the 5' phosphate. However the divalent cation dependence of Dicer for binding with ssRNAs and not with siRNAs, showed other Dicer domains besides the PAZ domain are needed for forming stable complexes. Using LC/MS the amino acids from the RNase IIIb domain were found to bind with the ssRNAs (Fig. 2.10 A and B), supporting the results from the ssRNA-Dicer binding studies (Fig. 2.6). Also from the LC/MS studies, the Dicer dsRBD was found to crosslink with the siRNAs and not with the ssRNAs (Fig. 2.11 A, B and C). Crystallographic and mutation studies of *Aquifex aeolicus* RNase III domain has been shown to possess a catalytic and non-catalytic assembly with dsRNAs (Blaszczyk *et al.* 2001; Blaszczyk *et al.* 2004). The non catalytic assembly is suggested to be active in the transfer of the cleaved substrates and thereby different from the catalytic assembly involving the catalytic valley of the RNase III domain. Following the identification of RNase IIIb domain residues of Dicer in binding with the ssRNAs it was of interest to know if this binding was representative of catalytic or non catalytic assembly. Catalytic processing of the substrates by Dicer involves the dimerization of the RNase III domains a and b to form a catalytic valley for endonucleolytic cleavage (Zhang *et al.* 2004). From the LC/MS studies the ssRNAs were found to crosslink only with the RNase IIIb domain and not both the RNase III domains. This shows that ssRNA is not aligned between the two catalytic domains, and Dicer-ssRNA complex is suggestive of a non-catalytic assembly for shuttling of the substrate.

Proposed model for Dicer-ssRNA binding and Ago2-Dicer interaction

It is unclear if binding of ssRNAs to Dicer would occur *in vivo*, and, if so, what the biological relevance of such an interaction would be. One necessary feature of active silencers is the presence of a 5'-phosphate on the guide strand (Nykanen *et al.* 2001). SiRNAs possessing a 5' end that is impaired for phosphorylation, lack the ability to bind Dicer or induce RNAi (Nykanen *et al.* 2001; Pellino *et al.* 2005). The ability of Dicer to distinguish phosphorylated and non-phosphorylated targets may be a mechanism for controlling the loading of RISC (Fig. 2.12). It may also suggest a pathway by which ssRNAs can enact RNAi (Schwarz *et al.* 2002; Holen *et al.* 2003). In those cases, only ssRNAs possessing a 5'-phosphate, or that are 5'-phosphorylated upon entering the cell, act as potent inhibitors. These would be bound more tightly by Dicer, enhancing their association with other RISC proteins such as Ago 2, TRBP, and PACT (Chendrimada *et al.* 2005; Kok *et al.* 2007).

Formation of active RISC by Ago2 mediated cleavage of the siRNA passenger strand has been well established (Matranga *et al.* 2005; Leuschner *et al.* 2006; Kim *et al.* 2007). However this mode of passenger strand cleavage for active RISC formation has only been demonstrated for siRNA induced silencing and not for miRNAs.

A bypass mechanism not involving passenger strand cleavage has been proposed for miRNAs (Matranga *et al.* 2005). Minimal RISC comprising of Dicer, TRBP, and has been shown to possess the ability to dice double stranded pre-miRNAs into ~22 nt miRNAs, to sense the asymmetry of the miRNA duplex and load the RISC with the appropriate guide strand and ultimately to slice the target mRNA (MacRae *et al.* 2008). Ago2 alone cannot form active RISC from siRNAs or miRNAs (Rivas *et al.* 2005). Hence an interaction of Ago2 with Dicer or TRBP is necessary for loading of RISC with

the single stranded guide strand. Besides humans, Dicer-Ago2 association has also been reported in *C.elegans* and *Drosophila* (Tabara *et al.* 2002; Chendrimada *et al.* 2005; Jiang *et al.* 2005; Saito *et al.* 2005). These results suggest that Dicer may play a role in shuttling/loading single stranded miRNA guide strand to Ago2. During processing of pre-miRNA substrates, the Dicer helicase domain autoinhibits the catalytic cleavage of the substrates (Ma *et al.* 2008; Soifer *et al.* 2008). This regulation of Dicer action is witnessed only for pre-miRNA like substrates and not for fully paired dsRNA molecules (Soifer *et al.* 2008). Structural rearrangements of the substrates following cleavage have been observed in crystal studies with *Aquifex aeolicus* RNase III , indicating some form of product hand off (Gan *et al.* 2008). PIWI domain of human Ago2 has been shown to bind with the RNase IIIa domain of Dicer (Tahbaz *et al.* 2004). This research has shown the ability of Dicer to bind with ssRNAs which is mediated by the PAZ domain and stabilized by the RNase IIIb domain.

Based on these results a model has been proposed for Ago2 interaction with Dicer for loading of active RISC with the guide strand (Fig. 2.13). Dicer binds with ssRNAs possessing 5' -phosphate or with the miRNA guide strand following the unwinding of the duplex. Then Ago2 binds with the RNase IIIa domain of Dicer, causing a structural change and transfer of the guide strand from Dicer to Ago2. Following this the 5' end of the guide strand is anchored in the phosphate binding pocket of Ago2 PIWI domain and Ago2 gets loaded with the guide strand. In the case of siRNAs wherein the Dicer dsRBD is involved, the Dicer-siRNA complex most likely attains a conformation different from the Dicer-ssRNA complex, and Ago2 gets loaded with the guide strand by passenger strand cleavage.

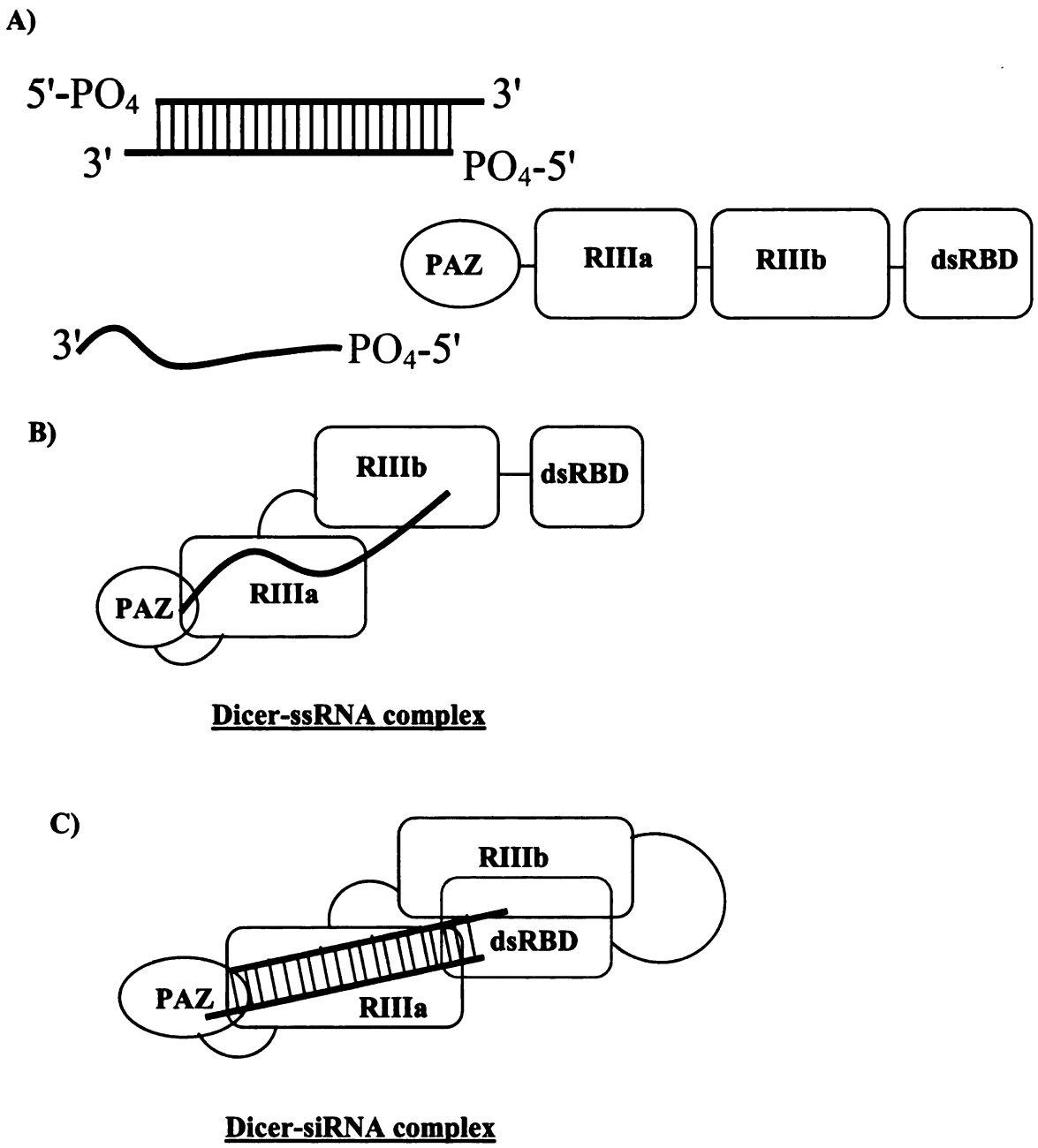
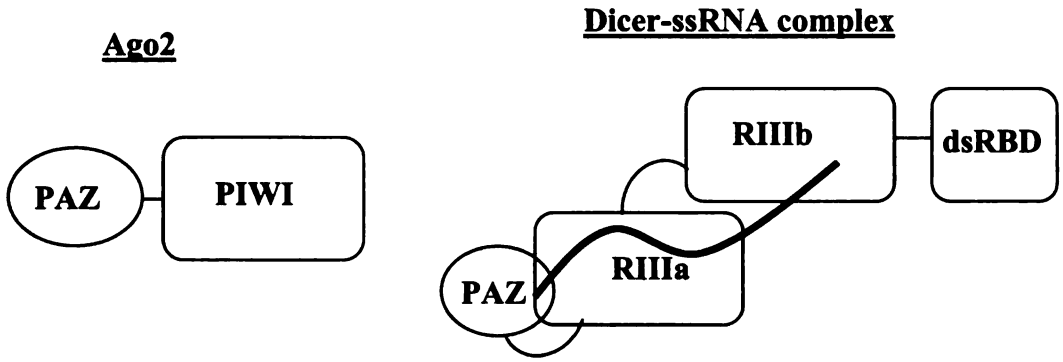


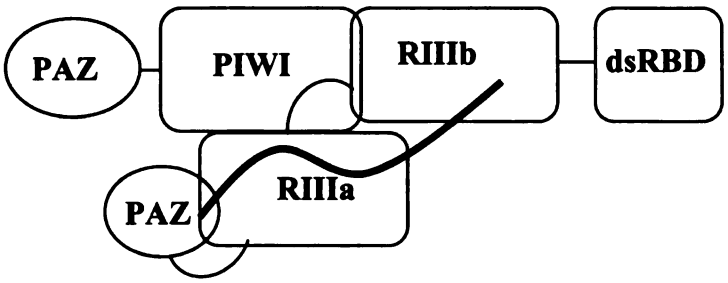
Figure 2.12 Schematic of Dicer interaction with siRNAs and ssRNAs

A) Initiation of Dicer interaction with the ssRNAs and siRNAs mediated by the PAZ domain. B) Dicer forms a complex with the ssRNAs stabilized by the RNase IIIb domain. C) Dicer forms a complex with the ssRNAs stabilized by the dsRBD.

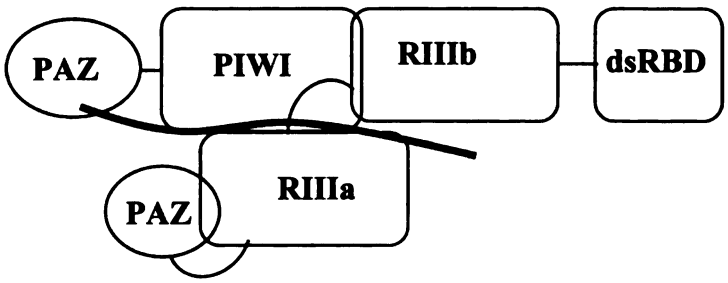
A)



B)



C)



D)

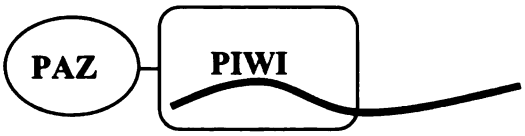


Figure 2.13 Schematic of Ago2 loading with the guide strand

CHAPTER III

Effect of siRNA terminal mismatches on TRBP and Dicer binding and silencing efficacy

INTRODUCTION

Short interfering RNAs (siRNAs) can be designed to target and regulate the expression of any gene of interest. Gene silencing by RNA interference (RNAi) is mediated by endogenous proteins resulting in target mRNA cleavage or translational inhibition (Filipowicz 2005). In the cytoplasm of human cells, the dsRNA binding proteins, TRBP and Dicer, recognize and bind with the siRNA and form the RNA induced silencing complex (RISC) loading complex (RLC) (Chendrimada *et al.* 2005; Haase *et al.* 2005; Diederichs and Haber 2007). Argonaute 2, the catalytic core of RISC (Liu *et al.* 2004; Rivas *et al.* 2005), is then recruited by the RLC to form holo-RISC (Maniataki and Mourelatos 2005). Although other proteins such as protein activator of protein kinase R (PACT) might also be associated with the formation of holo-RISC (Meister *et al.* 2005; Lee *et al.* 2006; Hock *et al.* 2007; Kok *et al.* 2007; Landthaler *et al.* 2008), *in vitro* experiments have demonstrated that TRBP, Dicer, and Ago 2 alone are capable of forming an active minimal RLC (MacRae *et al.* 2008).

Being double-stranded, either strand of the siRNA can be used as the guide strand of active RISC. Loading of both strands results in reduced silencing efficiency due to competition for RISC components and has the potential to result in off-target silencing (Jackson *et al.* 2006). Functional siRNAs and miRNAs have been shown to have greater asymmetries in their terminal hybridization stabilities compared to non-functional siRNAs (Khvorova *et al.* 2003; Schwarz *et al.* 2003; Reynolds *et al.* 2004). In *Drosophila*, the protein R2D2 binds to the more stable end of the siRNA duplex and

directs the binding of Dicer-2 to the other, less stable end. By this process, the guide strand is selected through the interaction of its 5'-end with Dicer-2 (Tomari *et al.* 2004; Bramsen *et al.* 2007). To ensure maximal specific silencing of the intended target, loading of the appropriate guide strand into RISC is critical. Improved understanding of the interactions of siRNAs with TRBP and Dicer will guide improved design of siRNA therapeutics.

In current applications, siRNAs are typically designed with an intentional bias, to maximize preferential selection of the appropriate guide strand, by making its 5'-end less stably hybridized than the other end (Reynolds *et al.* 2004; Schwarz *et al.* 2006). An end can be destabilized by introducing mismatches, wobble base pairs, or increasing A-U content (Reynolds *et al.* 2004). Some studies using siRNAs with a terminal mismatch showed improved activity (Hohjoh 2004; Ding *et al.* 2007), though not in all cases (Holen 2005; Hong *et al.* 2008). Typically, these studies used siRNAs that were initially found to be thermodynamically symmetric, with asymmetry subsequently induced by the mismatch. However, simultaneously changing sequence, structure, and asymmetry potentially disguises the impacts of multiple variables.

Thus, in this study, the effects of introducing a terminal mismatch to siRNAs with pre-existing thermodynamic asymmetry were investigated. In this way, the effects of structure and sequence changes were separated from changes in asymmetry. It was found that a terminal mismatch at the 5'-end of the known guide strand, which enhances the natural bias of the siRNA, adversely impacts its binding with TRBP and generally reduces its silencing activity. Unlike terminal mismatches, internal mismatches enhanced siRNA binding by both Dicer and TRBP. These results highlight the importance of

siRNA structure in impacting the interactions with RNAi pathway proteins and provide guidance for the design of highly active siRNAs.

RESULTS

Design of siRNAs and EGFP silencing efficiency

Designing siRNAs with an intentional bias in hybridization stability is intended to maximize proper guide strand selection and loading into RISC. This is beneficial both for achieving strong silencing and also for minimizing off target silencing by the passenger strand (Jackson *et al.* 2006). Thus, relative thermodynamic stability of the ends of the siRNA is an important design criterion for highly active siRNAs. Directing guide strand selection through chemical modifications has proven effective (Chen *et al.* 2008). However, typically, asymmetry is achieved through modification of either the passenger strand or the guide strand to generate a mismatch at the 5'-end of the guide strand (Holen *et al.* 2005; Ding *et al.* 2007). Asymmetric siRNAs generated by introducing a terminal mismatch to an initially symmetric siRNA were found to be more active than the *symmetric* siRNA (Table B1) (Ding *et al.* 2007). However, the goal here was to test whether introducing a mismatch to an already asymmetric siRNA would also improve the silencing efficiency of the siRNA.

An siRNA targeting position 396 of the enhanced green fluorescent protein (EGFP) mRNA was tested (Table B2) (Gredell *et al.* 2008). Using mfold (Mathews *et al.* 1999; Zuker 2003), the terminal stabilities of the siRNA was calculated (Table 3.1). For this siRNA, the known antisense strand 5'-end is located at the end predicted to be relatively thermodynamically unstable, as expected for correct loading into the RISC. Using this sequence as a basis, siRNAs with mismatches were generated by changing either the first

nucleotide of the guide strand, 396-**A**G, 396-**U**G, 396-**G**G, or the 19th nucleotide of the passenger strand, 396-C**A**, 396-C**U**, 396-C**C**, respectively (changed nucleotides are shown in bold; Table B2). Predicted free energies confirmed that the mismatches show increased bias relative to the fully-paired duplex (Table 3.1), which should enhance the likelihood for proper guide strand incorporation into RISC.

Table 3.1 Difference in end stabilities of the siRNAs.

The passenger strand 5'-end ΔG is -9.8 kcal/mol and -9.3 kcal/mol for all the variations of duplexes 396 and 306, respectively. Stability at each end was calculated using (Mathews *et al.* 1999; Zuker 2003), by summing up the contributions of base pairing energy of first four nearest neighbors and the overhang sequence.

Sequences	Guide strand 5'-end ΔG, kcal/mol	Difference in end stability ΔΔG, kcal/mol
396	-8.7	1.1
396-AG	-6.9	2.9
396-UG	-7.7	2.1
396-GG	-6.9	2.9
396-CA	-6.9	2.9
396-CU	-6.9	2.9
396-CC	-6.9	2.9
306	-7.1	2.8
306-CC	-5.4	4.5

These siRNAs were then used to silence EGFP in H1299 cells constitutively expressing EGFP (Liu *et al.* 2007). The silencing efficacy of the mismatched siRNAs was reduced, with the exception of 396-AG (Fig. 3.1). To confirm that this effect was not limited to

sequence 396, silencing by siRNA 306 (targeting position 306) and a corresponding mismatched sequence, 306-CC, was tested. Introducing a mismatch that increased the natural asymmetry of the duplex (Table 3.1) did not increase the silencing activity of the siRNA (Fig. 3.1). These results agree with prior studies in which siRNAs with terminal mismatches did not always function as predicted by terminal thermodynamic asymmetry (Table B3) (Holen *et al.* 2005; Hong *et al.* 2008). For selected siRNAs, the dose dependence of silencing was examined, to ensure that the differences among the siRNAs that were transfected at 10 nM were within the sensitive concentration range (Fig. B1).

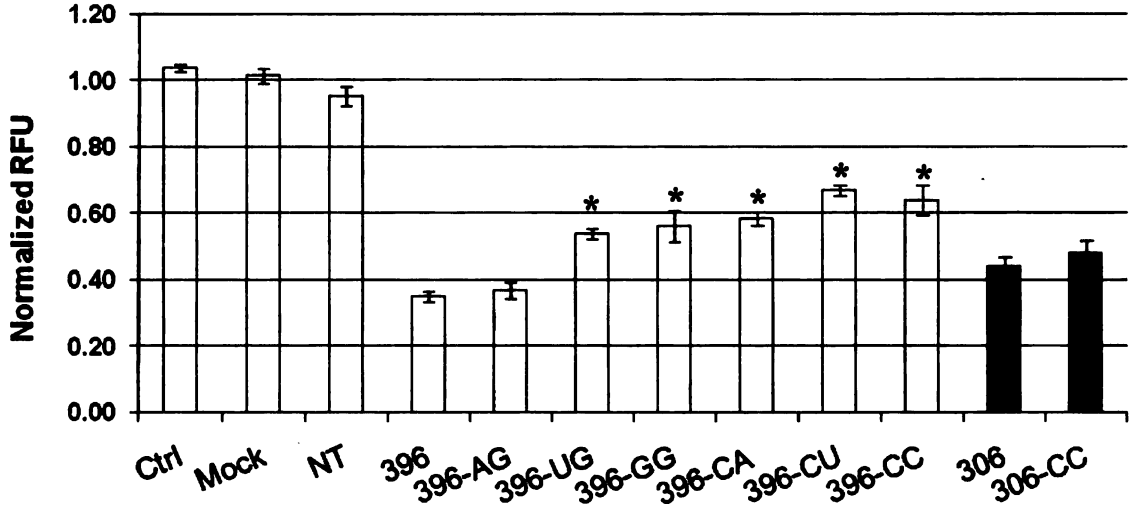


Figure 3.1 Effect of guide strand 5'-end mismatch on silencing efficacy of siRNAs.

EGFP-expressing H1299 cells were transfected with either an siRNA targeting the EGFP mRNA or a non-targeting (NT) siRNA at a final concentration of 10 nM. Fluorescence was measured 24 hours after transfection. Mean and standard deviation are shown for each condition. *, denotes the 2-tailed t-test comparison of silencing efficacy of the siRNAs with guide strand 5'-end mismatch vs. siRNA 396 ($p < 0.05$). Control and mock refer to untreated cells and cells treated with the transfection reagent alone, respectively. White bars denote siRNAs based on siRNA 396 and gray bars for siRNAs based on siRNA 306.

Effect of TRBP or Dicer knockdown on the silencing efficacy of mismatched siRNAs

It was hypothesized that the reduction in the function of the mismatched siRNAs was a consequence of impaired interactions with TRBP and/or Dicer. While both proteins are part of the RLC and holo-RISC and necessary for optimum silencing, RNAi induced target silencing has been demonstrated in the absence of either Dicer (Martinez *et al.* 2002; Kanellopoulou *et al.* 2005; Murchison *et al.* 2005) or TRBP (Haase *et al.* 2005). Further, unlike the *Drosophila* RNAi pathway wherein the R2D2 binding is a necessary precursor step for Dicer-2 binding (Tomari *et al.* 2004), in humans, Dicer alone can bind siRNAs (Pellino *et al.* 2005; Kini and Walton 2007) .

To study the effect of these two proteins on the functionality of the siRNAs with and without a terminal mismatch, either TRBP or Dicer protein was knocked down in H1299 cells (Fig. B2). After knockdown of TRBP, silencing of EGFP by the fully-paired duplex was reduced from over 65% to under 37%, with only one mismatched sequence being statistically significantly affected (396-UG, 46% reduced to 34%) (Fig. 3.2A). Notably, even with TRBP knocked down, siRNA 396-AG maintained essentially the same silencing capacity as in the presence of TRBP, actually becoming the most active of all the siRNAs under these conditions (Fig. 3.2A). In contrast, after knockdown of Dicer, only the silencing efficacy of 396-AG was significantly reduced (63% reduced to 47%), making it significantly worse than the fully-paired duplex in this case (Fig. 3.2B). The functionality of 396, along with all of the other mismatched sequences, was relatively unaffected by the reduction of Dicer protein (Fig. 3.2B). TRBP-siRNA binding has been shown to be more critical in the formation of the RLC than Dicer-siRNA binding (Chendrimada *et al.* 2005; Katoh and Suzuki 2007). Among all the sequences

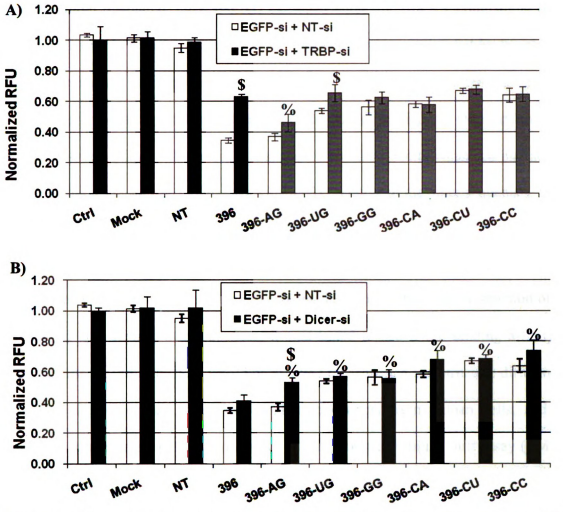


Figure 3.2 Effect of TRBP or Dicer knockdown on the silencing efficacy of the EGFP targeting siRNAs.

EGFP expressing H1299 cells were co-transfected with EGFP targeting siRNAs and either a non-targeting (NT) siRNA (white bars), TRBP targeting siRNA (A, gray bars), or Dicer targeting siRNA (B, black bars). Total, final siRNA concentrations were 20 nM (10 nM per siRNA). Fluorescence was measured 24 hours after transfection. Mean and standard deviation are shown for each condition. $ denotes the 2-tailed t-test comparison of silencing efficacy of the grey columns (EGFP-si + TRBP-si) vs. the white columns (EGFP-si + NT-si) ($p < 0.05$) (A) or of the black columns (EGFP-si + Dicer-si) vs. the white columns (EGFP-si + NT-si) ($p < 0.05$) (B); % denotes the 2-tailed t-test comparison of silencing efficacy of the EGFP targeting siRNAs co-transfected with, TRBP targeting siRNA vs. siRNA 396 ($p < 0.05$) (each grey vs. 396-grey) (A) or Dicer targeting siRNA vs. siRNA 396 ($p < 0.05$) (each black vs. 396-black) (B). Control, mock, and NT refer to untreated cells, cells treated with the transfection reagent alone, and cells transfected with the NT siRNA instead of an EGFP-targeting siRNA, respectively.

with a terminal mismatch, only 396-AG exhibited silencing efficacy comparable to the fully-paired 396. In addition, the activity of 396-AG was most impacted by knockdown of Dicer. Dicer has been shown to prefer adenosine nucleotides at the terminal position during Dicer processing of long double stranded RNAs to siRNAs (Vermeulen *et al.* 2005). Thus, the unique behavior of this sequence could be due to enhanced interactions with Dicer and reduced need to interact with TRBP, relative to the other sequences.

Effect of guide strand 5'-end mismatch on TRBP and Dicer binding

Having seen variability in the impact of silencing TRBP and Dicer on the function of the fully-paired and mismatched sequences, it was tested to see if the binding affinity of these proteins for the sequences would be affected by sequence and structure differences. Radiolabeled siRNA was added to cytoplasmic extracts from human cells, and the complexes formed were detected via native electrophoretic mobility shift assay (EMSA) (Fig. 3.3A). This G-C rich sequence was used, as it was shown to form easily discernable bands in the extracts (data not shown). As seen previously (Pellino *et al.* 2005), putative Dicer-siRNA complex formation in H1299 cell lysates was detected (Fig. 3.3A, dashed arrow, lane 1; substantially equivalent data obtained with HepG2 and HeLa extracts not shown). To confirm the presence of Dicer in the complex, binding experiments were performed in the presence of Dicer antibody, TRBP antibody, and nuclear factor κB, NF-κB antibody (Fig. 3.3A, dashed arrow; compare lane 2 to lanes 1, 3, and 4), akin to prior experiments with purified Dicer protein (Kini and Walton 2007). As expected, the band was shifted in the presence of the Dicer antibody but not the other antibodies. It was

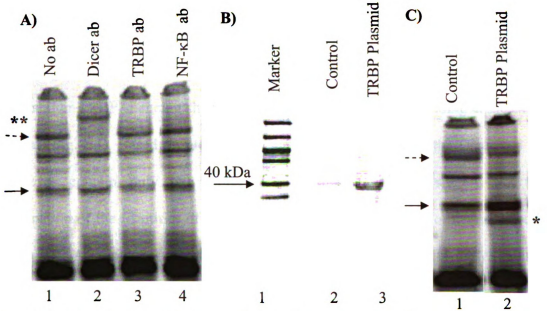


Figure 3.3 Characterizing siRNA-TRBP and siRNA-Dicer complexes

A) Electrophoretic mobility shift assay (EMSA) of siRNA-protein complexes formed in H1299 cell extracts (Lane 1), in the presence of Dicer antibody (Lane 2), in the presence of TRBP antibody (Lane 3), or a control antibody against NF-κB (nuclear transcription factor kappaB) (Lane 4). Broken arrow indicates position of the siRNA-Dicer complex; ** indicates the migration of the shifted siRNA- Dicer complex; and solid arrow indicates the position of the siRNA-TRBP complex. B) Western blot analysis shows TRBP overexpression in H1299 cells transfected with TRBP plasmid (Lane3) compared to control cells (Lane 2). C) EMSA of the siRNA-protein complexes formed in H1299 cell extracts with TRBP overexpression (Lane 2) and in control cells (Lane 1).

further tested to confirm the location of any TRBP containing bands, if these could be visualized. In the presence of TRBP antibody, a reduction in the signal from a band at the appropriate position for the expected molecular weight of TRBP was noticed (Fig. 3.3A, solid arrow; compare lane 3 to lanes 1, 2, and 4; intensity reduced ~ 40% by gel quantification compared to lane 1), but a shifted complex was not detected. To verify the presence of TRBP in this siRNA-protein complex, TRBP in H1299 cells was overexpressed, and the increase in expression confirmed by western blot (Fig. 3.3B; compare lanes 2 and 3). Incubating the radiolabeled siRNAs with TRBP-overexpressed

lysates showed a significant increase in the formation of the putative TRBP complex (Fig. 3.3C; compare lanes 1 and 2), strongly supporting the antibody shift results suggesting the presence of TRBP in this complex. Binding reactions performed in extracts after TRBP silencing showed a concomitant reduction in binding at the expected location (Fig. B3). Based on molecular weight, it seems that both the Dicer and TRBP complexes contain only one molecule each of protein and siRNA. As further confirmation of the identities of the complexes, it was shown that formation of both the protein-siRNA complexes was improved by the presence of ATP in the extracts (Fig. 3.4A and B), as shown previously (Pellino *et al.* 2005). Another siRNA-containing complex of unknown identity was also seen in these extracts (Fig. 3.3C, *), which may be a result of the response of the cell to the presence of the plasmid and/or the excess TRBP.

Identical binding reactions were carried out with siRNA 396 and the mismatched siRNAs in H1299 cell extracts (Fig. 3.5A and B). Analysis of protein complexes formed by these siRNAs with TRBP showed that binding to TRBP was significantly lower for the siRNAs with a single terminal mismatch (Fig. 3.5A and B), including 396-AG. This trend agrees closely with the results from the TRBP silencing experiments, where the fully matched sequence appeared to depend most on the function of TRBP. The trend in TRBP binding was verified with two other siRNAs, 306 and 274 (Gredell *et al.* 2008), with both an outright mismatch, 306-CC (Fig. B4), and a U-G wobble, 274-UG (Fig.B5). There was no consistent behavior in Dicer binding for the different sequences (Fig. 3.5A and B), even including 396-AG.

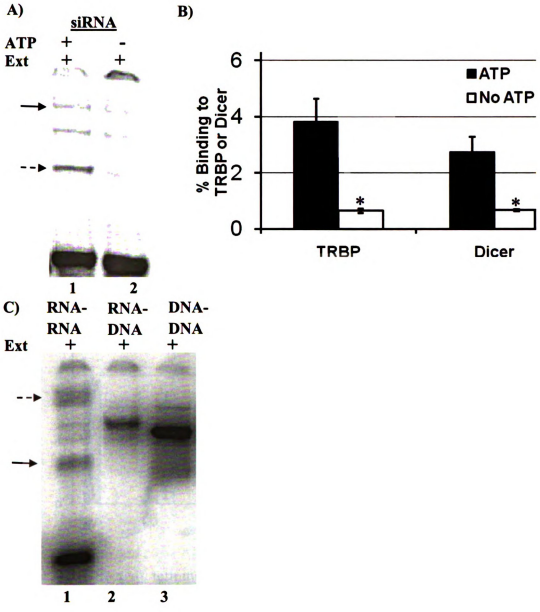


Figure 3.4 Characterizing Dicer, TRBP compmlexes in H1299 cell extracts

A) Dicer and TRBP complex formation with siRNA in H1299 cell extracts is enhanced by the presence of ATP. B) Quantification of EMSA gel images. Percentage binding was calculated by normalizing the intensities of siRNA-protein complex bands to the respective unbound siRNAs (control lanes not shown). Mean and standard deviation are shown for triplicate binding experiments. C) EMSA of Dicer and TRBP complexes formed in H1299 cell extracts with siRNAs (lane 1), and RNA-DNA heteroduplex (lane 2), and a DNA-DNA duplex (lane 3). Dicer and TRBP complexes form only with the RNA-RNA (siRNA) duplex.

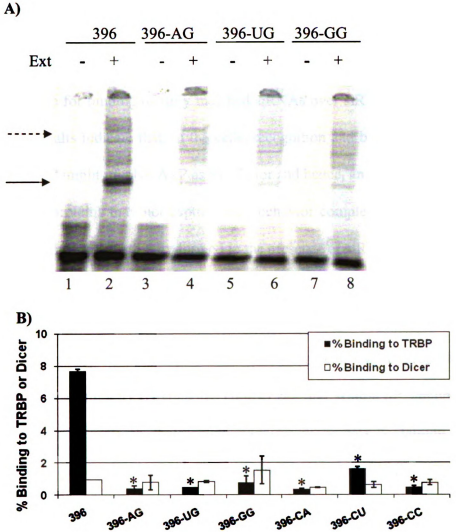


Figure 3.5 Effect of terminal mismatch at guide strand 5'-end on siRNA-TRBP and siRNA-Dicer complex formation.

A) EMSA of siRNA-TRBP and siRNA-Dicer complexes formed in H1299 cell extracts with siRNAs 396 (lane 2), 396-AG (lane 4), 396-UG (lane 6), and 396-GG (lane 8). Separate gels containing other siRNAs not shown. B) Quantification of EMSA gel images. Percentage binding was calculated by normalizing the intensity of siRNA-protein complexes to the siRNA not exposed to extract (e.g., complexes in lane 2 vs. free siRNA in lane 1). Mean and standard deviation are shown for triplicate binding experiments. * denotes the 2-tailed t-test comparison of TRBP binding of different siRNAs vs. siRNA 396 ($p < 0.05$); $ denotes the 2-tailed t-test comparison of Dicer binding of different siRNAs vs. siRNA 396 ($p < 0.05$).

Recent work done by our group using purified TRBP protein has shown that it can bind siRNAs in an ATP-independent manner (Gredell JA, Dittmer MJ, and Walton SP, unpublished data). In those studies, TRBP protein by itself did not show a strong preference for binding of fully matched siRNAs over siRNAs with a terminal mismatch. These results indicate that, in the cells, recognition and binding of the siRNAs by Dicer and TRBP might involve ATP as a cofactor and hence, an *in vitro* assay using the purified proteins may not capture their behavior completely. That said, both Dicer and TRBP complexes were only formed in the presence of siRNAs and not RNA-DNA hetero duplex or DNA-DNA duplex (Fig. 3.4C), similar to the results with recombinant TRBP protein *in vitro*. The sensitivity of TRBP binding to the terminal modifications suggests that it primarily binds at the siRNA termini, corroborating its proposed role as a sensor for siRNA asymmetry (Gredell JA, Dittmer MJ, and Walton SP, unpublished data).

It has also been shown that an immunopurified complex containing Dicer, TRBP, and Ago2 has the ability to process pre-miRNAs, form active RISC upon selection of a guide strand, and direct Ago2-mediated silencing (Gregory *et al.* 2005; Maniataki and Mourelatos 2005). Active RISC formed from Dicer processed pre-miRNAs was 10-fold more active than that formed from mature miRNAs targeting the same sequence (Gregory *et al.* 2005). This is different from the activity of *in vitro* constituted RLC consisting of only Dicer, TRBP and Ago2 (MacRae *et al.* 2008). Silencing activity of the RISC formed from the *in vitro* complex is similar for both pre-miRNAs or miRNAs (MacRae *et al.* 2008), suggesting in cells there might be other cellular cofactors associated with the RLC and RISC that affect their function. Studying proteins such as MOV10 (Moloney leukemia virus 10 homologue) (Meister *et al.* 2005; Hock *et al.* 2007), TNRC6B

(trinucleotide repeat containing 6B) (Meister *et al.* 2005), and RHA (DEAH box polypeptide 9) (Hock *et al.* 2007), which are associated with Ago2 may elucidate the differences between *in vitro* and *in vivo* RLC/RISC formation and function.

Effect of siRNA structure and composition on siRNA protein complexes

The impact of terminal and also selected internal mismatches on the interactions of Dicer and TRBP with siRNAs was further studied (Table B2). G-C rich sequence (used in Fig. 3.3), was used to give the cleanest readout for changes that occurred in the formation of the complexes. A single or double mismatch at one end of the duplex seemed to decrease TRBP binding slightly, though not significantly (Fig. 3.6). For Dicer, binding was improved slightly with a single mismatch, while weakened by the double mismatch. Again, neither of these changes was statistically significant. Simultaneous single or double mismatches at both ends of the duplex significantly reduced binding by TRBP, echoing what was seen with mismatches at only one end. As with one terminal mismatch, binding by Dicer was improved for simultaneous single mismatches but reduced for double mismatches. In all cases, terminal mismatches reduced TRBP binding, as above (Fig. 3.5), strongly suggesting that terminal mismatches should be avoided to generate siRNAs with maximal activity.

The efficiency of Dicer processing of long dsRNAs is known to depend on the overhang length of the substrates, with 2-3 nt overhangs being highly favorable compared to overhangs longer than 3-nt (Vermeulen *et al.* 2005). In addition, the PAZ (PIWI Argonaute Zwille) domain, which Dicer possesses, is known to mediate binding with dsRNAs and siRNAs through 3'-overhangs (Lingel *et al.* 2003; Song *et al.* 2003; Lingel *et al.* 2004). Binding affinity of the human Ago2 PAZ domain to a siRNA duplex

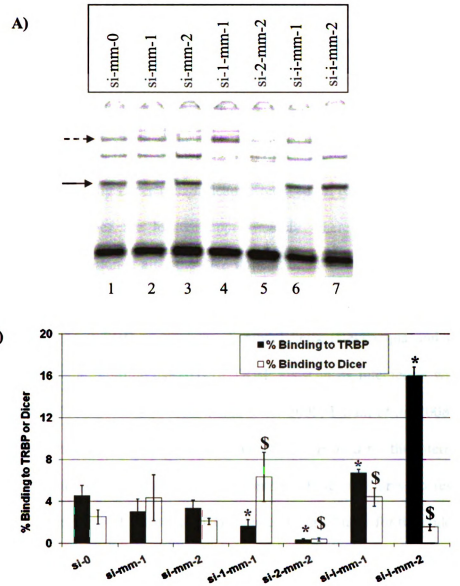


Figure 3.6 Effect of terminal and internal mismatches on siRNA-TRBP and siRNA-Dicer complexes.

A) EMSA of siRNA-TRBP and siRNA-Dicer complexes formed in H1299 cell extracts with siRNAs of varying terminal and internal structures (Fig. B6). Broken and solid arrows indicate the migration of the siRNA-Dicer and siRNA-TRBP complexes, respectively. B) Quantification of EMSA gel images. Percentage binding was calculated by normalizing intensity of siRNA-protein complex to the respective unbound siRNAs (control lanes not shown). Mean and standard deviation are shown for triplicate binding experiments. * denotes the 2-tailed t-test comparison of TRBP binding of different siRNAs vs. siRNA si-0 ($p < 0.05$); $ denotes the 2-tailed t-test comparison of Dicer binding of different siRNAs vs. siRNA si-0 ($p < 0.05$).

has been shown to be reduced 5-fold to 50-fold by increasing the overhang length from 2 nt to, 4 nt and 10 nt respectively (Ma *et al.* 2004). Thus, the assays using cellular extracts seem to accurately demonstrate the natural function of the proteins.

Both proteins showed higher affinity for a duplex with one internal mismatch (Fig. 3.6, si-i-mm-1). Binding by TRBP improves with two internal mismatches (Fig. 3.6, si-i-mm-2), while binding by Dicer is significantly reduced. In relation to Dicer binding, the two internal mismatches are located approximately where the dsRBD (dsRNA binding domain) gets positioned after the PAZ domain binds to one end of the duplex (Lingel *et al.* 2003; Song *et al.* 2003), thus the reduction in binding affinity may result from the inability of the dsRBD to bind the disrupted helix (Bevilacqua and Cech 1996) It is possible that the multiple dsRBDs of TRBP assist in its interaction with the sequences with internal mismatches (Chendrimada *et al.* 2005; Laraki *et al.* 2008). However, it is not immediately clear why the binding would be *improved* for the internally mismatched sequence relative to the fully matched control. These structures do resemble miRNAs, and it may be that both Dicer and TRBP have higher affinity for the endogenous silencers as compared to exogenous siRNAs. Also, functional siRNAs tend to have lower internal stability than non-functional siRNAs, particularly at positions 1-6 and 10-15 (with position 1 being the 5'-end of the guide strand) (Khvorova *et al.* 2003), exactly where the mismatches are located in this study, which may be due to some still uncharacterized function of TRBP in RNAi.

CONCLUSIONS

TRBP, Dicer binding of siRNAs and functionality of the siRNAs

In this research the interactions of siRNAs possessing terminal mismatches with TRBP and Dicer were characterized, and how these interactions impact their silencing activity. Primarily, it was found that, for an asymmetric siRNA, introducing a terminal mismatch that further reduces the stability of the guide strand 5'-end does not enhance the functionality of the siRNAs. Based on comparison of the binding and silencing results, it is opined that reduced TRBP binding is a likely reason for reduced silencing by mismatched siRNAs. That said, it appears that Dicer binding can impact the silencing efficiency of some siRNAs in a terminal sequence dependent manner. It is interesting to note that all of the mismatches were located at the end where Dicer preferentially should bind based on the current model for RISC formation and siRNA asymmetry sensing (Tomari *et al.* 2004). Nonetheless, the binding by TRBP is more dramatically and consistently affected by the mismatches. This assay does not discriminate the location to which either TRBP or Dicer bind on the siRNA. It is possible that TRBP can associate with equal likelihood with either end of the siRNA but that its dissociation rate is faster with the less-stable end. As such, the mismatches likely enhance this dissociation rate and hence reduce the overall average affinity of TRBP for the mismatched siRNA relative to the fully-paired sequence. Alternatively, this could be a reflection of the importance of the TRBP-Dicer interaction in binding to siRNAs, which would also help to explain the differences between binding with only purified TRBP or Dicer versus binding in extracts. It may also suggest that the role of human Dicer in selecting the guide strand and generating active RISC is more prominent than that of *Drosophila* Dicer-2, which is

controlled by R2D2 binding rather than actively participating in determining which end to bind (Tomari *et al.* 2004). Future work examining internal and terminal modifications should further identify design rules for enhancing the activity of siRNA duplexes and also provide for a better understanding of the roles of TRBP and Dicer in controlling siRNA silencing activities.

CHAPTER IV

Summary and future directions

Being an effective and convenient tool to regulate gene expression, RNAi has been widely employed to study gene function and is being pursued as a therapeutic strategy. Hence the understanding of the details of the mechanism of silencing in humans is of great importance to realize its full potential and particularly if it is to be used as a means of therapy. While a great deal of mechanistic details has been learned from the pioneering work done in *Drosophila* and *C. Elegans*, there are still several important details yet to be unraveled. Some of these pertain to off-target silencing by siRNAs, the role of the guide strand seed region, terminal stability of siRNAs and selection of the guide strand in RISC, proteins involved in asymmetry sensing in humans, and how the RISC gets programmed with the guide strand in the case of miRNAs without passenger strand cleavage.

Dicer and TRBP are the proteins which are primarily responsible for recognizing siRNAs and initiating the formation of RISC. Besides formation of RISC, both proteins are involved in other cellular processes as well. Though they share some known and proposed roles with the corresponding proteins in *Drosophila*, the human proteins are more complex and might have other functions beyond those that are currently known or can be inferred from their counterparts in other organisms.

Major contributions of this research

- A novel function for both human and *Giardia* Dicer protein has been found, i.e., ability to bind ssRNAs.

- While both ssRNAs and siRNAs interact with the Dicer PAZ domain, their stable complexes are mediated by different domains: the RNase IIIb domain for ssRNAs and dsRBD for siRNA.
- Based on the findings of this research and the current understanding of the pathway, a model has been proposed for the interaction between Dicer and Ago2 for loading of Ago2 with the guide strand.
- The effect of terminal mismatches on TRBP binding has been correlated to the activity of siRNAs.

Future Work

Characterizing other RNAi pathway complexes in cell extracts

In vitro silencing experiments to knock down EGFP and GAPDH mRNAs as measured by qPCR have been used to demonstrate the cleavage competency of cell extracts. Complexes formed by siRNAs with Dicer and TRBP proteins in H1299, HepG2 and HeLa cytoplasmic cell extracts have been characterized. It has also been shown that immunopurified complexes containing Dicer, TRBP, and Ago2 have the ability to process pre-miRNAs, form active RISC with the guide strand, and direct Ago2-mediated silencing (Gregory *et al.* 2005; Maniataki and Mourelatos 2005). Active RISC formed from Dicer-processed pre-miRNAs was 10-fold more active than that formed from identical miRNAs (Gregory *et al.* 2005). This is different from the activity of *in vitro* constituted RISC consisting of only Dicer, TRBP and Ago2 (MacRae *et al.* 2008). Silencing activity of this RISC is similar for both pre-miRNAs or miRNAs (MacRae *et al.* 2008), suggesting in cells there might be other cellular cofactors associated with the RISC proteins that affect its function. Studying proteins such as MOV10 (Moloney

leukemia virus 10 homologue) (Meister *et al.* 2005; Hock *et al.* 2007), TNRC6B (trinucleotide repeat containing 6B) (Meister *et al.* 2005) , and RHA (DEAH box polypeptide 9) (Hock *et al.* 2007), which are associated with Ago2, may elucidate the differences between *in vitro* and *in vivo* RISC function. Other sequence and structural modifications could be used to study their effect on interactions with the proteins of the RNAi pathway as the siRNA progresses from a duplex to single stranded guide strand.

Is it RNAi ?

Results from this research have shown that both TRBP and Dicer can bind only to RNA-RNA duplex, and not to a RNA-DNA or DNA-DNA duplex. (Fig. 3.4C). These results are supported by earlier reports demonstrating the necessity of A-form of helix formed by a RNA-RNA duplex in the siRNAs and between guide RNA and mRNA for effective silencing (Chiu and Rana 2002). Furthermore, A-form helix formed by the guide strand and mRNA is necessary for endonucleolytic cleavage by Ago2 (Parker *et al.* 2004; Parker *et al.* 2005). However there are reports with duplexes with DNA substitutions in the sense and antisense sequence inducing moderate to active silencing, comparable to siRNAs with the same sequence (Hohjoh 2002; Ui-Tei *et al.* 2008). It is possible that there are other proteins and mechanisms different from RNAi mediating the silencing in these cases. The *in vitro* cell extract system can be used to identify the protein complexes formed by these modified duplexes to understand the mode of silencing. This strategy can also be employed to study silencing by siRNAs with modifications which alter the basic structure of the duplex, 19 base pairs, 5'-phosphate and 3' overhangs. Such shorter siRNAs have been tested and found to be effective silencers (Hohjoh 2004; Chang *et al.* 2009).

siRNAs with terminal and internal modifications

Among all the sequences with a terminal mismatch only 396-AG exhibited silencing efficacy comparable to 396 (Fig. 3.1). It appears that Dicer binding does impact the silencing efficiency of siRNAs in a terminal sequence dependent manner. While the terminal mismatches seem to have an adverse impact on TRBP, one internal mismatch seems to enhance both the TRBP and Dicer binding (Fig. 3.6). Future work with siRNAs having internal and terminal modifications should identify design rules for enhancing the activity of siRNA duplexes and also provide for a better understanding of the role of TRBP and Dicer in controlling siRNA silencing activity. Identifying siRNAs that are functional in the absence of TRBP might be useful for targeting genes in cell types with low TRBP expression, such as astrocytes (Gatignol *et al.* 2005; Christensen *et al.* 2007).

Binding studies with Dicer mutants

The Dicer helicase domain has been observed to regulate the processing of pre-miRNA like substrates and not dsRNA substrates (Ma *et al.* 2008; Soifer *et al.* 2008). So the structural rearrangement of the substrate protein complex must be different for a Dicer-miRNA complex vs. a Dicer-siRNA complex. It would be interesting to know if mutations in the helicase domain that enhance pre-miRNA processing would also affect the ability of Dicer to bind with ssRNAs. This would provide further evidence to the involvement of Dicer in the bypass mechanism of Ago2 loading.

APPENDIX A

Characterizing Dicer protein preparation

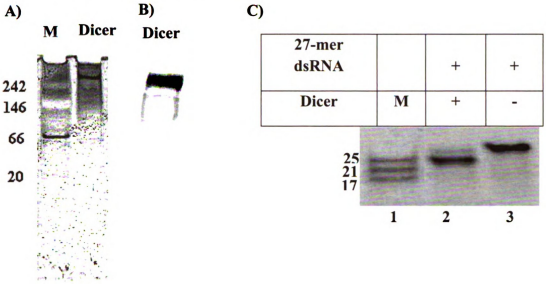


Figure A1 Characterizing ssRNA-protein complexes

Dicer protein visualized by A) Silver staining B) Western blot with monoclonal Dicer antibody. C) Dicer cleavage of a 27-mer dsRNA visualized by native gel electrophoresis

Mass spectrometry analysis of unknown complexes formed by ssRNAs

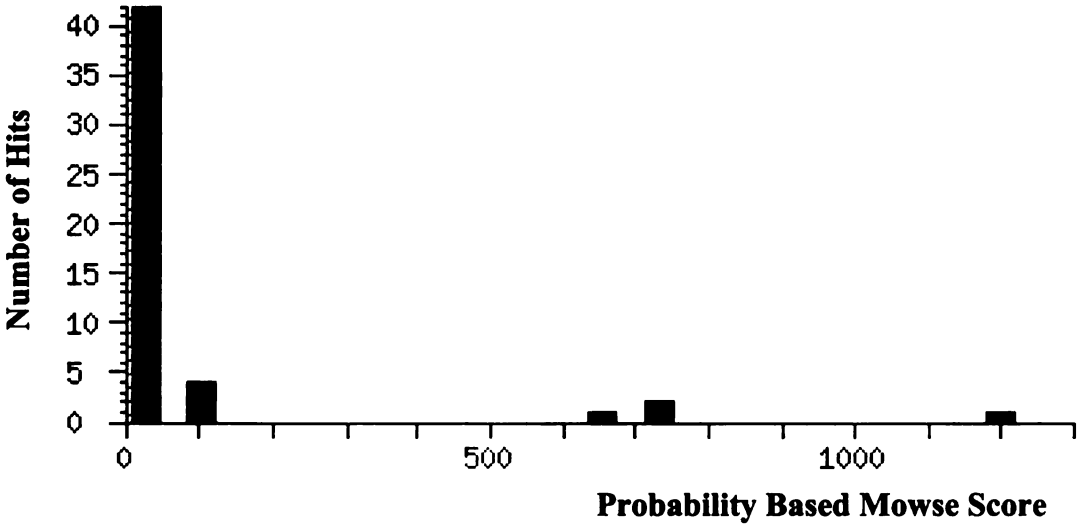


Figure A2 Mass spectrometry of the faster moving complexes formed by the ssRNA (denoted by the dotted arrows in Figure 2.6).

Proteins with significant hits identified by mass spectroscopy

keratin 1 (Homo sapiens), keratin 10 (Homo sapiens), keratin 9 (Homo sapiens), Keratin, type II cytoskeletal 2 epidermal (Cytokeratin-2e) (K2e), cytokeratin 8, keratin 5 (Homo sapiens], keratin K7, type II, epithelial, 55K - human keratin, 65K type II cytoskeletal - human, TPA: Hornerin (Homo sapiens), NuMA protein (Homo sapiens), alpha 1A adrenoceptor isoform 2b (Homo sapiens), KIAA1481 protein (Homo sapiens), putative (Homo sapiens)

Generating theoretical cleavage map for tryptic digest of a protein

Mass spectrometry data analysis: BSA

The cleavage map were generated from the University of California, San Fransisco ProteinProspector website :

http://prospector.ucsf.edu/cgi-bin/msform.cgi?form=msdigest

Following options were chosen for tryptic digest of BSA (accession number: P02769):

Database: SwissProt

Digest: Trypsin

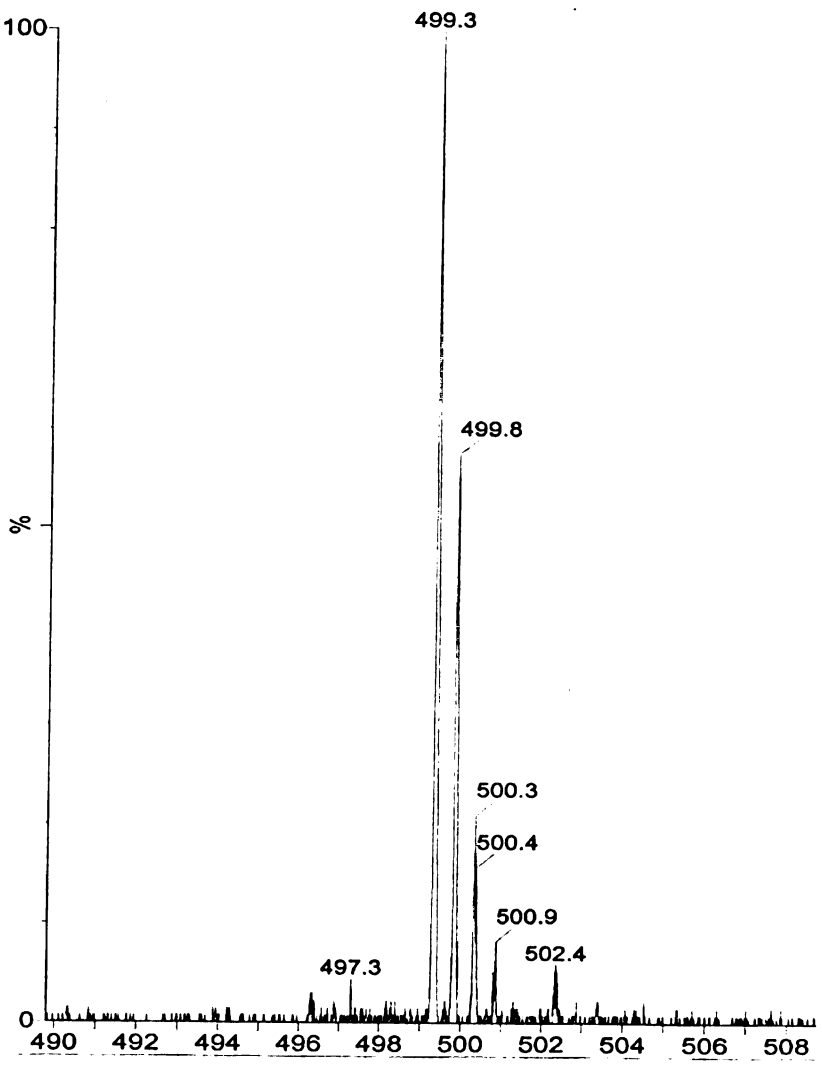
Missed clevages: 1

End modification: carbamidomethyl

Peptide mass range: 250-4000

Multiple charges to be displayed

A)



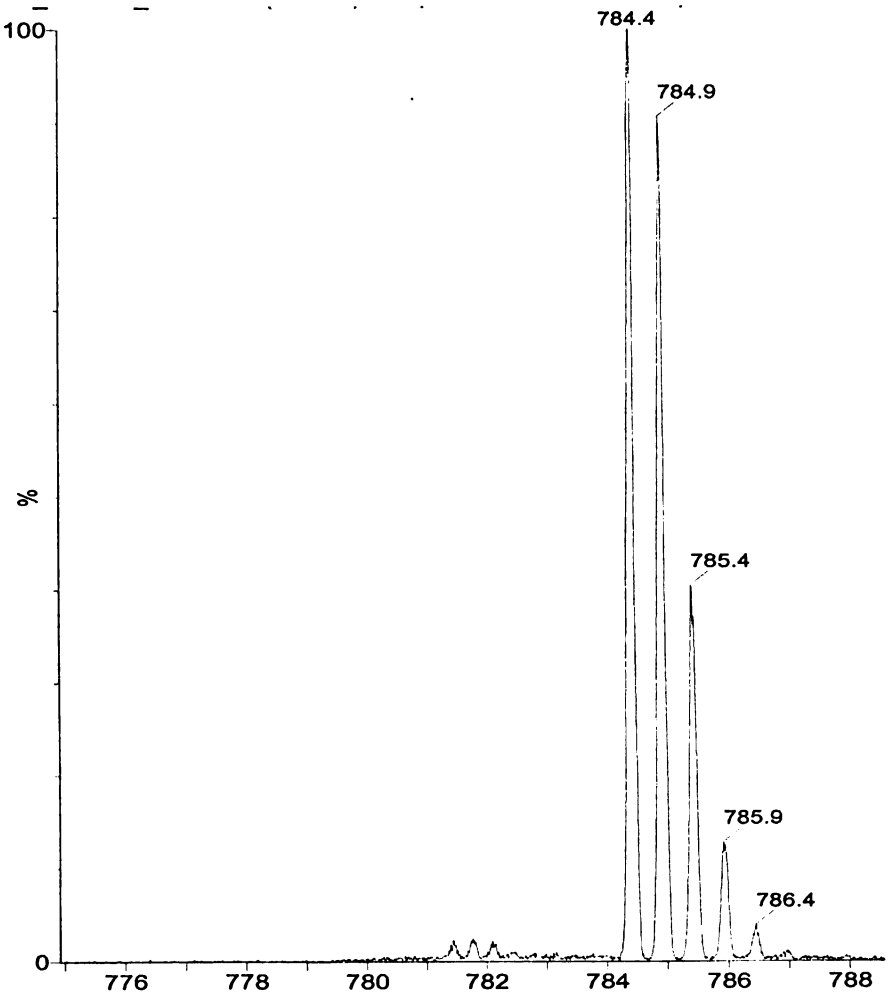
B)

m/z (mi)	m/z (av)	Start	End	Sequence
496.6363^{+3}	496.9465^{+3}	549	561	(K)QTALVELLKHKPK(
499.3001^{+2}	**499.6112^{+2}**	**549**	**557**	**(K)QTALVELLK(H)**
501.2982^{+2}	501.6096^{+2}	233	241	(R)ALKAWSVAR(L)
502.3118^{+3}	502.6234^{+3}	549	561	(K) QTALVELLKHPK(A)

Figure A3 Spectra of trypsinized BSA protein

A) A doubly charged peak of BSA peptide, and B) the corresponding peptide sequence in bold.

A)



B)

m/z (mi)	m/z (av)	Start	End	Sequence
784.3750^{+2}	784.8721^{+2}	347	359	(K)DAFLGSFLYEYSR(R)
785.7181^{+3}	784.2168^{+3}	402	420	(K)HLVDEPQNLIKQNC(L)
788.8874^{+2}	789.4074^{+2}	139	151	(K)LKPDPNTLC(A)
789.4716	789.9529	257	263	(K) LVTDLTKV)

Figure A4 Spectra of trypsinized BSA protein A) A doubly charged peak of BSA peptide, and B) the corresponding peptide sequence in bold.

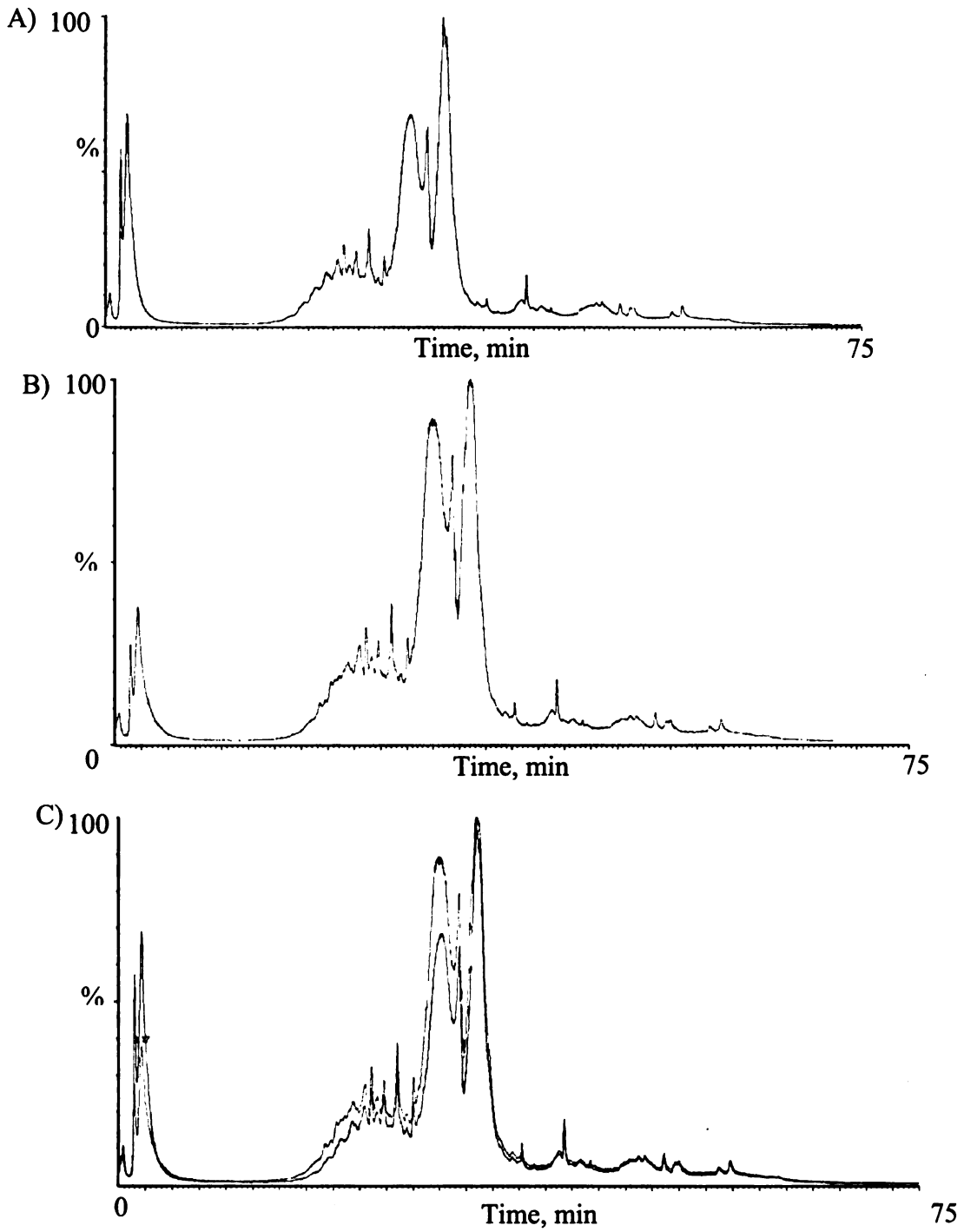


Figure A5 LC/MS chromatograms Chromatograms of A) proteolysed Dicer B) Dicer-siRNA complex and C) overlaid chromatograms from A and B.

Methods and materials

RNA and 5'-end labeling

HPLC purified RNAs were purchased from Invitrogen (Carlsbad, CA) or Dharmacon (Lafayette, CO). Lyophilized RNAs were resuspended in 100 mM concentration Tris-EDTA (TE, pH 8.0) and stored at -80°C. RNAs were 5'-end labeled with ^{33}P-γ-ATP (Perkin Elmer Life and Analytical Sciences, Boston, MA) using T4 polynucleotide kinase (New England Biolabs, Ipswich, MA). Labeled strands were purified from unincorporated label by G-25 Sephadex columns (Roche Applied Science, Indianapolis, IN). For nonisotopic labeling, unlabeled ATP (Roche Applied Science, Indianapolis, IN) was substituted for ^{33}P-γ-ATP. The ssRNA and siRNA sequences used were: siRNA, 5'–GCUGACCCUGAAGUUCAUCUU-3' (Sense strand), 5'-GAUGAACUUCAGGGUCAGCUU-3' (Anti-sense strand); ssRNA-SS1 5' –GUCACAUUGCCCAAGUCUCTT-3'; ssRNA- SS2, 5'- UUUUUUUUUUUUUUUUUUUTT-3'. SsRNA with the 3'-biotin had the same sequence as SS2 but with a 3'-biotin

Characterization of recombinant human Dicer purity

Recombinant human Dicer (Invitrogen) was used in all the binding reactions. Full length Dicer protein was visualized by Silver Staining Kit (Bio-Rad) (Figure A1A) and confirmed by Western blot (Figure A1B). For western blot analysis, proteins were initially electrophoresed at 150 V for 1.5 hours at 4°C on 4-20 % TBE gels (Bio-Rad) and then transferred to nitrocellulose membranes for 1h at 100V, followed by incubation with mouse monoclonal antibody for Dicer overnight at 4°C. Blots were washed and then

incubated with HRP-linked secondary antibodies (Pierce Biotechnology, IL, USA) for 1 h. After an additional wash, the blots were developed with Pierce SuperSignal West Femto Maximum Sensitivity Substrate (Pierce Biotechnology) and imaged on Chemidoc XRS imager (Bio-Rad). Dicer activity was assessed through a cleavage assay using a 27-mer dsRNA substrate being converted to ~ 21-mer siRNAs (Figure A1C, compare Lanes 2 and 3).

Dicer-RNA binding assays

Dicer binding assays were carried out in 30 mM Tris-HCl (pH=8.0), 250 mM NaCl, 2.5 mM $MgCl_2$, and 0.02 mM EDTA. Labeled RNA, and 0.5 U of Dicer (Invitrogen) were incubated at 4°C for 2 hours in 10 µl reaction volumes. For antibody supershift assays, either Rabbit polyclonal antibody to Dicer (Abcam) or control antibody was incubated with Dicer at 4°C for 3 hours after which the labeled RNA was added and incubated for 2 additional hours. Samples were electrophoresed at 150 V for 1.5 hours at 4°C on 4-20 % TBE gels (Bio-Rad). Gels were dried under vacuum at 80°C, exposed to storage phosphor screens, and imaged on a Storm 860 imager (GE Healthcare/Amersham Biosciences, Piscataway, NJ). Binding reactions to test the divalent cation dependence were carried out with the appropriate divalent cation substituting for Mg^{2+} in the binding buffer. Mass spectrometry was performed on gel purified protein by the Michigan State University Research Technology Support facility.

In-solution trypsinization

1. Protein was dissolved in 50 mM ammonium bicarbonate of Tris buffer (pH 8) at a concentration of approximately 1-10 μM.
2. Dithiothreitol (DTT), tris-carboxyethylphosphine (TCEP), or tributylphosphine was added to 10 mM final concentration and heated at 95 ^{o}C for 10 minutes and cooled to room temperature.
3. Freshly made iodoacetamide was added to 200 μM final concentration and incubated for 1 hour at room temperature preferably in a dark drawer or wrapped in aluminum foil.
4. DTT was added to 1 mM to quench the ioodacetamide.
5. After 10 mins, 30 ng of freshly diluted trypsin (Sequencing grade, Promega) prepared in 50 mM ammonium bicarbonate, pH 8 was added, and incubated at 37 ^{o}C for 4-24 hours. An equal volume of 5% Formic acid was added and sonicated in water bath for 10 minutes to quench the digestion, and redissolved in a minimal volume of water/acetonitrile (90:10 v/v) containing 1% formic acid for LC/MS analysis.

LC/MS

Hypurity Aquastar Pioneer (Thermo Fisher Scientific) column was used in the Waters 2795 separation module. The mobile phase had a mixture of 0.1 % Formic acid (A) in water and acetonitrile (B). The following gradient was used during the run :

Time, min	0	2	40	45	55	70	70.01	75
B	1	1	85	85	99	99	1	1
A	99	99	15	15	1	1	99	99

Flow through from the column was then passed through Q-TOF Ultima API) system (Waters) under positive electron spray having a cone and capillary voltage of 35 V and 3 KV respectively.

APPENDIX B

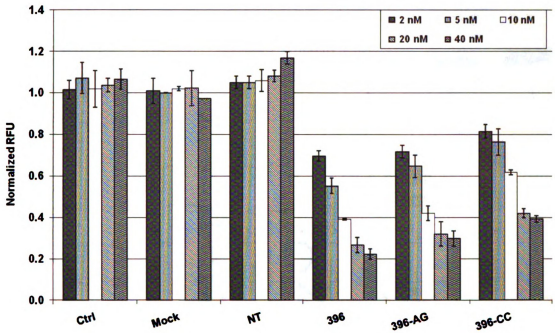


Figure B1 EGFP silencing efficacy of siRNAs at different concentrations

EGFP-expressing H1299 cells were transfected with siRNAs targeting the EGFP mRNA or a non targeting (NT) siRNA at final concentrations of 2, 5, 10, 20 or 40 nM. Fluorescence was measured 24 hours after transfection. Mean and standard deviation are shown for 12 wells for each condition. Control, mock, and NT refer to untreated cells, cells treated with the transfection reagent alone, and cells transfected with the NT siRNA, respectively.

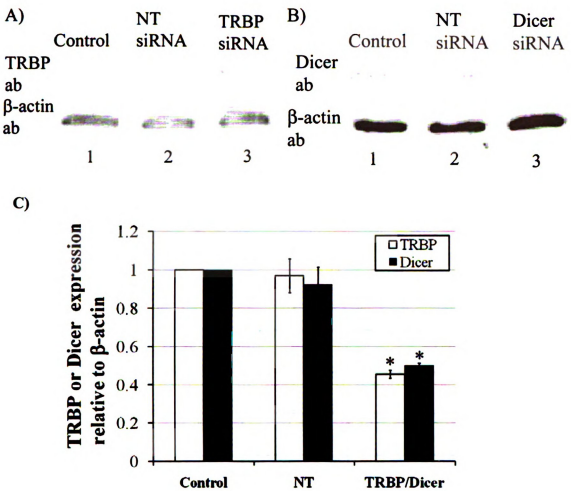


Figure B2 Western blot analysis of TRBP and Dicer levels in H1299 cells.

A) Comparison of TRBP knockdown (lane 3) relative to control (lane 1) and transfection with a non-targeting siRNA (lane 2). B) Comparison of Dicer knockdown (lane 3) relative to control (lane 1) and transfection with a non-targeting siRNA (lane 2). C) Band intensities were quantified using Bio-Rad Quantity One software. Quantification of TRBP or Dicer expression relative to β-actin; control ratio normalized to 1. * denotes the 2-tailed t-test comparison of TRBP or Dicer expression relative to control ($p < 0.05$). Mean and standard deviation are shown for triplicate silencing experiments.

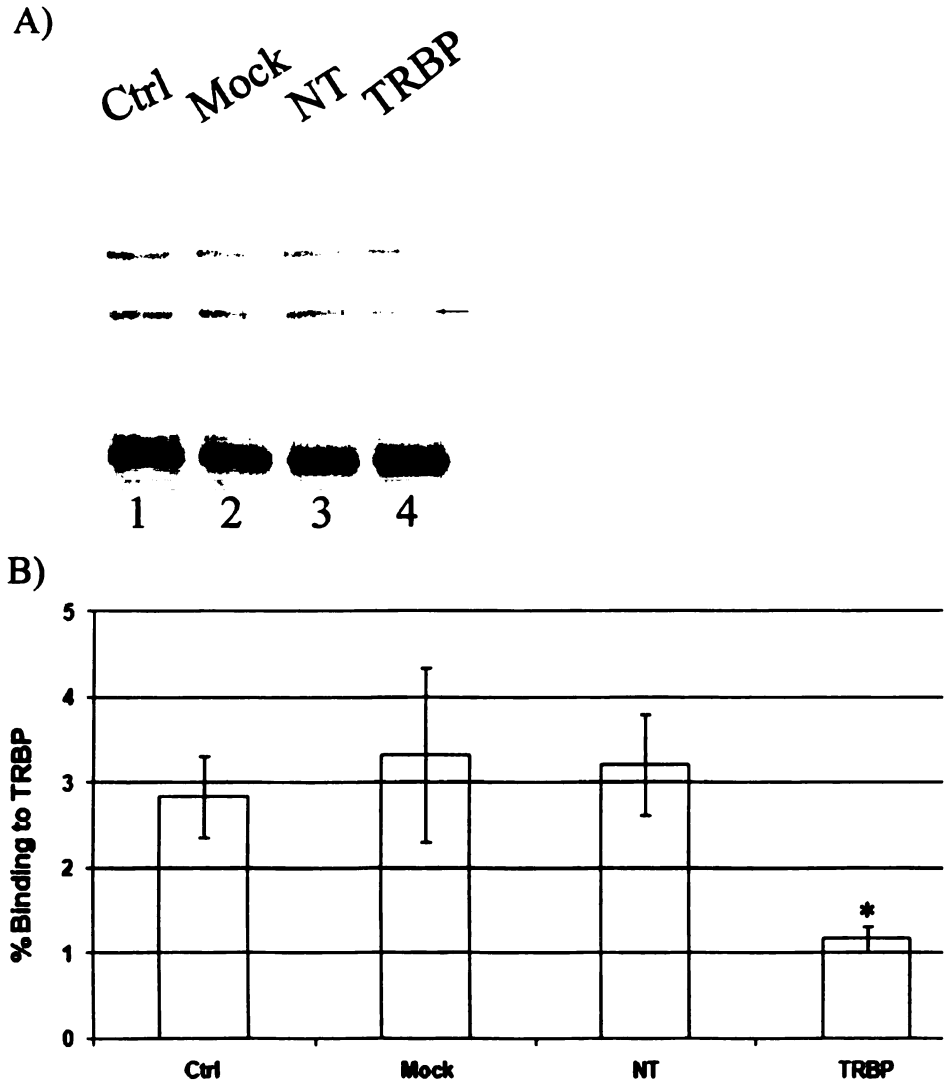


Figure B3 Characterization of siRNA-TRBP complex formation after TRBP knockdown

A) EMSA of siRNA-TRBP complex formed in H1299 cell extracts with siRNAs. Control, mock, NT, and TRBP refer to extracts from untreated cells, cells treated with the transfection reagent alone, cells transfected with the NT siRNA, and cells transfected with the TRBP-targeting siRNA, respectively. Arrow denotes the position of siRNA-TRBP complex. B) Quantification of gel images. Percentage binding was calculated by normalizing the intensity of siRNA-protein complexes to the siRNA not exposed to extract (e.g., complexes in Lane 2 vs. unbound siRNA (not shown)). Mean and standard deviation are shown for triplicate binding experiments.

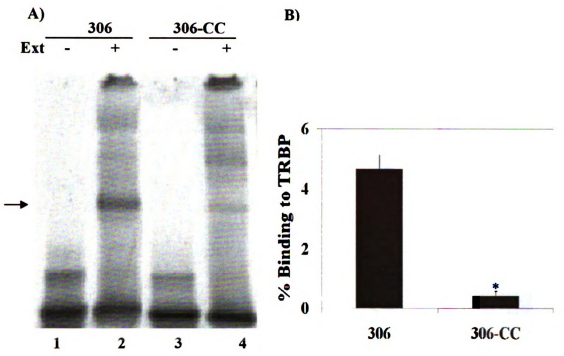


Figure B4 Effect of a terminal mismatch at the guide strand 5' end on siRNA-TRBP complex formation.

A) EMSA of siRNA-TRBP formed in H1299 cell extracts with siRNAs 306 (lanes 1 & 2) and 306-CC (lanes 3 & 4). B) Quantification of EMSA gel images. Percentage binding was calculated by normalizing the intensity of the siRNA-TRBP complex (lanes 2 and 4) to the respective unbound siRNAs (lanes 1 and 3). Mean and standard deviation are shown for triplicate binding experiments. * denotes the 2-tailed t-test comparison of TRBP binding of siRNA 306-CC vs. 306 ($p < 0.05$).

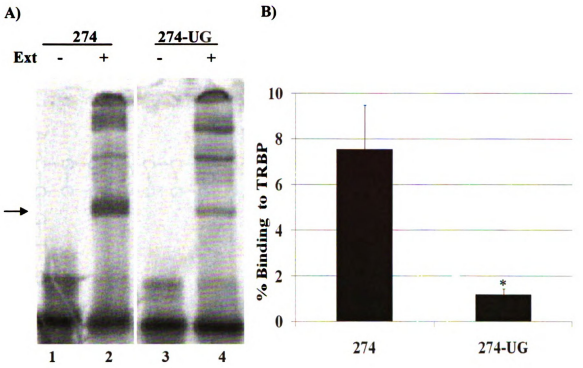


Figure B5 Effect of a terminal mismatch at the guide strand 5' end on siRNA-TRBP complex formation.

A) EMSA of siRNA-TRBP complexes formed in H1299 cell extracts with siRNAs 274 (lanes 1 & 2) and 274-UG (lanes 3 & 4). B) Quantification of EMSA gel images. Percentage binding was calculated by normalizing intensity of siRNA-TRBP complex (lanes 2 and 4) to the respective unbound siRNAs (lanes 1 and 3). Mean and standard deviation are shown for triplicate binding experiments. * denotes the 2-tailed t-test comparison of TRBP binding of siRNA 274-UG vs. 274 ($p < 0.05$).

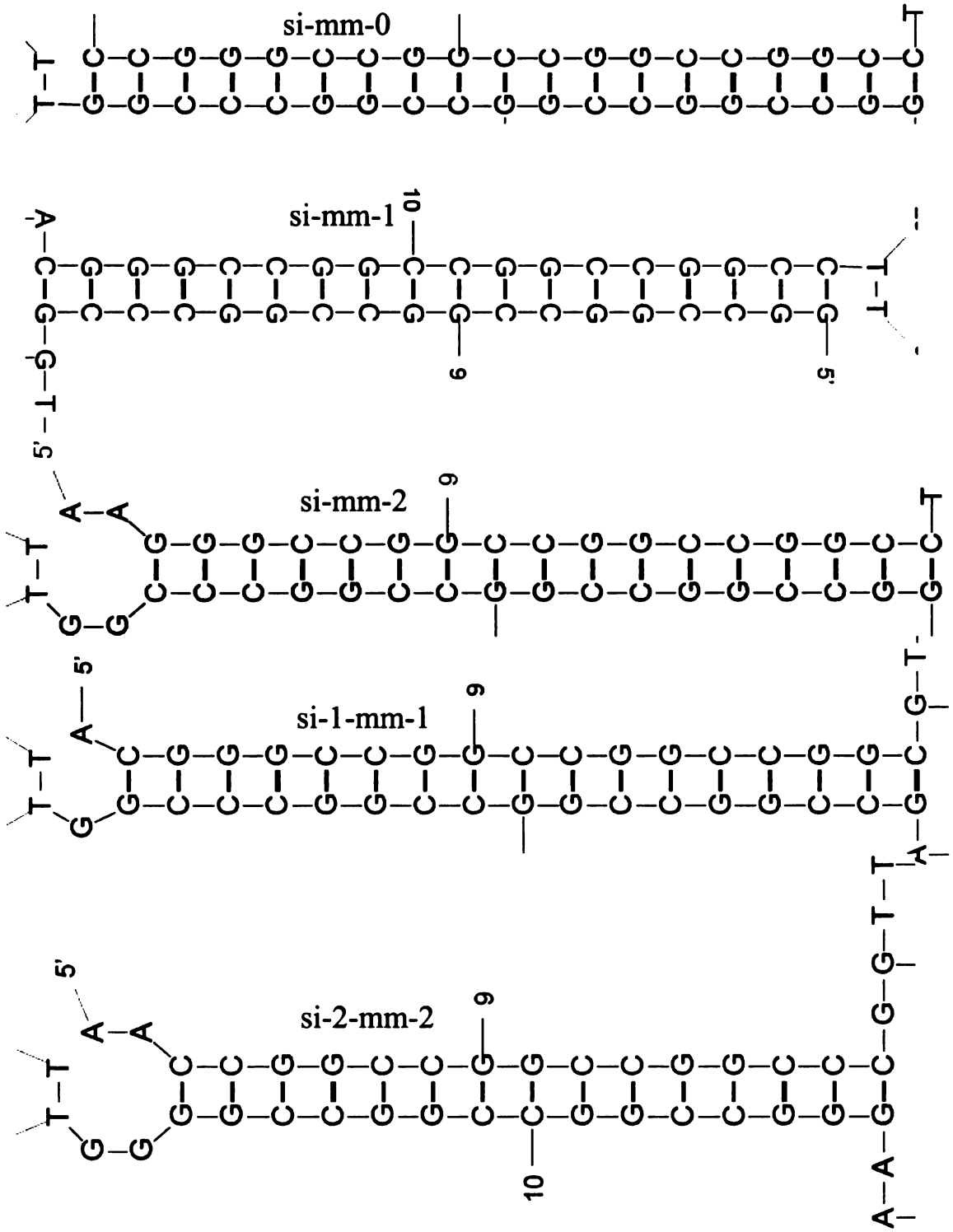


Figure B6 Structure of siRNAs with terminal and internal mismatches

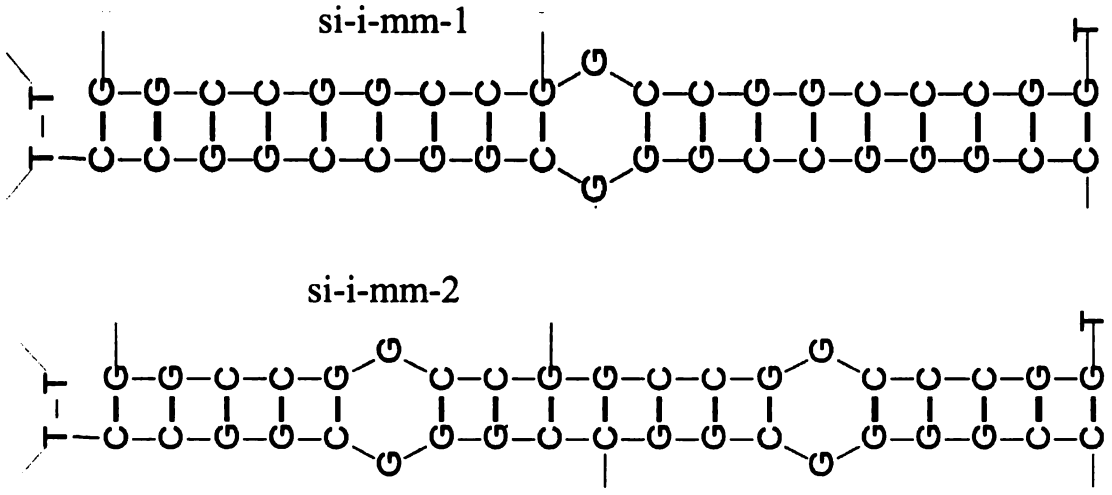


Figure B6 (Continued)

Terminal stability analysis of siRNAs with terminal mismatches

Table B1 SiRNAs with terminal modifications (Ding *et al.* 2007)

Name	ΔΔG, kcal/mol	Antisense sequence (5' - 3')	Sense sequence (5' - 3')
P11	0.1	GAGACUUGGGCAAUGUGACT	GUCACAUUGCCCAAGUCUCTT
A1	2.2	GAGACUUGGGCAAUGUGACT	GUCACAUUGCCCAAGUCUATT

Table B2 Sequence of siRNAs used to target EGFP

Name	Antisense sequence(5' - 3')	Sense sequence (5' - 3')
si-396	CAGGAUGUUGCCGUCCUCCTT	GGAGGACGGCAACAUCCUGTT
si-396-AG	AAGGAUGUUGCCGUCCUCCTT	GGAGGACGGCAACAUCCUGTT
si-396-UG	UAGGAUGUUGCCGUCCUCCTT	GGAGGACGGCAACAUCCUGTT
si-396-GG	GAGGAUGUUGCCGUCCUCCTT	GGAGGACGGCAACAUCCUGTT
si-396-CA	CAGGAUGUUGCCGUCCUCCTT	GGAGGACGGCAACAUCCUATT
si-396-CU	CAGGAUGUUGCCGUCCUCCTT	GGAGGACGGCAACAUCCUUTT
si-396-CC	CAGGAUGUUGCCGUCCUCCTT	GGAGGACGGCAACAUCCUCTT
si-306	CUUGUAGUUGCCGUCGUCCTT	GGACGACGGCAACUACAAGTT
si-306-CC	CUUGUAGUUGCCGUCGUCCTT	GGACGACGGCAACUACAAATT
si-274	UGCGCUCCUGGACGUAGCCTT	GGCUACGUCCAGGAGCGCATT
si-274-UG	UGCGCUCCUGGACGUAGCCTT	GGCUACGUCCAGGAGCGCGTT
Dicer-siRNA	UUUGUUGCGAGGCUGAUUCTT	GAAUCAGCCUCGCAACAAATT
TRBP-siRNA	GCUGCCUAGUAUAGAGCAATT	UUGCUCUAUACUAGGCAGCTT
NT or si-0	CCGGGCCGGCCGGCCGGCCTT	GGCCGGCCGGCCGGCCCGGTT
si-mm-1	ACGGGCCGGCCGGCCGGCCTT	GGCCGGCCGGCCGGCCCGGTT
si-mm-2	AAGGGCCGGCCGGCCGGCCTT	GGCCGGCCGGCCGGCCCGGTT
si-1-mm-1	ACGGGCCGGCCGGCCGGCGTT	AGCCGGCCGGCCGGCCCGGTT
si-2-mm-2	AAGGGCCGGCCGGCCGGGGTT	AACCGGCCGGCCGGCCCGGTT
si-i-mm-1	CCGGGCCGGGCGGCCGGCCTT	GGCCGGCCGGCCGGCCCGGTT
si-i-mm-2	CCGGGGCGGCCGGGCGGCCTT	GGCCGGCCGGCCGGCCCGGTT
RNA-RNA	GGCUACGUCCAGGAGCGCAUU	UGCGCUCCUGGACGUAGCCUU
RNA-DNA	GGCUACGUCCAGGAGCGCAUU	TGCGCTCCTGGACGTAGCCTT
DNA-DNA	GGCTACGTCCAGGAGCGCATT	TGCGCTCCTGGACGTAGCCTT

Table B3 siRNAs with terminal modifications (Holen *et al.* 2005)

Name	ΔΔG, kcal/mol	% Silencing	Antisense sequence (5' - 3')	Sense sequence (5' - 3')
Fe775	1.3	~70	GCAGCACAAUGAUGAGUGCAA	GCACUCAUCAUUGUGCUGCU U
m1Aas	4.7	~70	GCAGCACAAUGAUGAGUGCAA	ACACUCAUCAUUGUGCUGCU U
m1Cas	4.7	~70	GCAGCACAAUGAUGAGUGCAA	CCACUCAUCAUUGUGCUGCU U
m1Uas	4.7	~70	GCAGCACAAUGAUGAGUGCAA	UCACUCAUCAUUGUGCUGCU U

Calculating terminal stability of the siRNAs

Terminal stability (ΔG, kcal/mol) at each end of the siRNA duplex was calculated using Mfold (Mathews *et al.* 1999; Zuker 2003) by summing up the base pairing energy of the initial four base pairs at that 5' end (Schwarz *et al.* 2003) (Fig. B7). Differential end stability (ΔΔG, kcal/mol) was calculated by taking the difference in thermodynamic stabilities at each end.

Sample calculations for differential end stability of a siRNA duplex :

5'- ACGCUGAACUUGUGGCCGUTT - 3' Antisense strand(AS)
5'- ACGGCCACAAGUUCAGCGUTT – 3' Sense strand (SS)

Step 1:

Go the URL mentioned below,

http://dinamelt.bioinfo.rpi.edu/results/twostate/081201/151031/

Step 2:

In the left box paste the AS sequence from 5' to 3'

In the right box paste the SS sequence from 5' to 3' and submit it to the Mfold server.

Step 3:

After Mfold hybridizes the sequences it leads to the page with T_M, siRNA structure, Thermodynamic details etc. Click on the Thermodynamic details. Sum the ΔG for the rows 2-5 of the table (Fig. B8.A) to get the terminal stability of the antisense end ΔG_{AS}, kcal/mol.

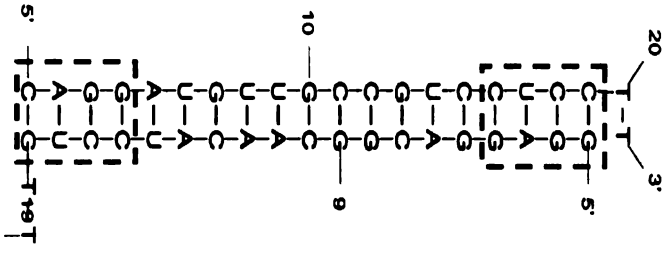


Figure B7 Structure of siRNA.

The dotted box shows bases used in calculating stability at that end of the duplex

Step 4:

Repeat Steps 2 through 3 by pasting the SS sequence in the left box and the AS sequence in the right box for Step 2 to calculate the terminal stability of the other end of the siRNA, ΔG_{SS}, kcal/mol (Fig. B8.B).

Step 6:

Differential end stability, $\Delta\Delta G = \Delta G_{AS} - \Delta G_{SS}$.

$$=-9.2 - (-10.4)$$

$$=-1.2 \text{ kcal/mol}$$

A)

Structural element	ΔG	Information
External loop	-0.1	2 ss bases & 1 closing helices
Stack	-2.1	External closing pair is A^{1}-U^{19}
Stack	-2.0	External closing pair is C^{2}-G^{18}
Stack	-2.9	External closing pair is G^{3}-C^{17}
Stack	-3.4	External closing pair is G^{4}-G^{16}

B)

Structural element	ΔG	Information
External loop	-0.1	2 ss bases & 1 closing helices
Stack	-2.1	External closing pair is A^{1}-U^{19}
Stack	-2.0	External closing pair is C^{2}-G^{18}
Stack	-3.4	External closing pair is G^{3}-C^{17}
Stack	-1.7	External closing pair is C^{4}-G^{16}

Figure B8 Mfold 2 state hybridization server webpage listing thermodynamic details.

Tables show contribution and base pairing energy of first four base pairs at A) Antisense strand end and B) Sense strand end respectively.

Methods and materials

General methods

siRNAs were purchased from Thermo Scientific Dharmacon (Lafayette, CO). Lyophilized RNAs were resuspended to 100 μM in TE (pH 8.0) and stored at -80^{o}C. RNAs were 5'-labeled with ^{33}P-γ-ATP (Perkin–Elmer Life and Analytical Sciences, Boston, MA) using T4 polynucleotide kinase (New England Biolabs, Ipswich, MA). Labeled strands were purified from unincorporated label using G-25 Sephadex columns (Roche Applied Science, Indianapolis, IN). Cell cytoplasmic extracts were prepared as described (Lee *et al.* 1995). Binding reactions in cell extracts with radiolabeled siRNAs were performed as described (Pellino *et al.* 2005). All binding reactions are performed for 1 h at 37°C. The competency of all extracts for *in vitro* silencing was tested by measuring EGFP mRNA transcript levels in H1299 cell cytoplasmic extracts prior to and following addition of siRNAs (data not shown). EMSAs was performed as previously described (Kini and Walton 2007) and quantified using a Storm 860 imager (Amersham/GE Healthcare, Piscataway, NJ). Percent binding was calculated by normalizing the intensity of the siRNA-protein complex (Fig. 3.5A, lanes 2, 4, 6 and 8, complexes indicated by arrows) to that of the respective unbound siRNA (Fig. 3.5A, lanes 1, 3, 5 and 7). Sequences of all RNAs used in these studies are listed in Table B2. ATP depletion experiments were carried out in binding buffer lacking ATP and containing glucose and hexokinase without creatine phosphate and creatine kinase (Pellino *et al.* 2005).

Cell transfection and EGFP quantification

Human lung carcinoma cells (H1299) constitutively expressing EGFP were generously provided by Dr. Jorgen Kjems, Department of Molecular Biology, University of Aarhus, Denmark. They were maintained in Dulbecco's modified Eagle's medium complemented with 10 % (vol/vol) fetal bovine serum (Invitrogen), 100 mg/ml of penicillin, and 100 units/ml streptomycin (Invitrogen). 24 h before transfection, cells were seeded at 50,000 cells/well in 24-well plates in antibiotic free media for siRNA transfection or seeded at 400,000 cells/well in 6-well plates for TRBP plasmid DNA transfection. Cells were transfected using Lipofectamine 2000 (Invitrogen) (0.8 μL for siRNA transfection and 3 μL for plasmid transfection), according to the manufacturer's recommendations. SiRNA or TRBP plasmid DNA was diluted with Opti-MEM (Invitrogen) followed by addition of Lipofectamine and complex formation. SiRNAs were used at final concentrations of 10 nM and TRBP plasmid DNA at 1 μg. When two siRNAs were transfected simultaneously, the final, total siRNA concentration was 20 nM. Cells were treated with this transfection medium for 4 h at 37^{o}C after which the transfection medium was replaced with regular cell culture medium. 24 h after transfection, the culture medium was aspirated and EGFP levels were quantified as described (Gredell *et al.* 2008). It was previously confirmed that transfection efficiency using established protocols provides essentially uniform siRNA loading across the different siRNA treatments (Gredell *et al.* 2008). For EGFP quantification, the fluorescence of each well of the 24-well plates was measured in 9 locations using a Gemini fluorescence plate reader (Molecular Devices, Sunnyvale, CA). The mean fluorescence for each well was calculated from these 9

values. The average fluorescence for a condition was calculated as the mean of multiple wells (typically 3-4) on the same plate. Relative fluorescence units (RFU) (Fig. 3.1 and 3.2) were calculated by normalizing the multi-well average fluorescence for each condition to the multi-well average fluorescence of mock transfected wells from the same plate. At least three wells from at least six different 24-well plates were measured for each condition ($n \geq 18$).

Western Blots

Cells were collected 24 h after plasmid or siRNA transfection. SDS-loading buffer was added to samples and heat denatured at 95^{o}C for 5 minutes. The samples were immediately placed on ice, and the proteins were resolved on 4-20% gradient sodium dodecyl sulfate-polyacrylamide gel (Bio-Rad) at 150V for 90 minutes. Proteins were then transferred to a PVDF membrane at 100V for 1 h. The membrane was then incubated with blotting grade milk (Bio-Rad) for 1 h and then incubated overnight at 4^{o}C with either TRBP antibody (Abnova, Walnut, Ca) or Dicer antibody (Abcam, Cambridge, MA) at 1:1000 dilution. Blots were then washed with TBS-Tween and incubated with the horseradish peroxidase-conjugated secondary antibody, and the proteins were detected using SuperSignal West Femto Maximum Sensitivity Substrate (Pierce Biotechnology, IL). β-Actin was used as the loading control. Dicer and TRBP knockdown levels were measured by normalizing to the control cells (no transfection). Images were collected using a ChemiDoc XRS (Bio-Rad, Hercules, CA), and band intensities were quantified using Bio-Rad Quantity One software. Dicer and TRBP knockdown were quantified by a ratio of ratios. Dicer and TRBP levels were normalized to β-Actin loading control for

each treatment, and these ratios were then normalized to the ratio for control cells (no transfection).

Free Energy Calculations

Terminal stability (ΔG, kcal/mol) at each end of the siRNA duplex was calculated using mfold (Mathews *et al.* 1999; Zuker 2003) by summing the nearest neighbor contributions for the first five nucleotides (four nearest neighbor energies) at the 5' end (as in (Schwarz *et al.* 2003)). Differential end stability (ΔΔG, kcal/mol) was calculated by taking the difference in thermodynamic stabilities at each end. For example, si-396 has guide strand sequence of 5'-CAGGAUGUUGCCGUCCUCCTT-3' and passenger strand sequence of 5'-GGAGGACGGCAACAUCCUGTT-3'. Base pairing energies for the duplex were predicted using the mfold 2-state hybridization server with RNA selected and default parameters. The four nearest neighbors at the guide strand 5' end, CA:GU, AG:UC, GG:CC, and GA:CU, have a cumulative base pairing energy of -8.7 kcal/mol. The four nearest neighbors at the passenger strand 5' end, GG:CC, GA:CU, AG:UC, and GG:CC, have a cumulative base pairing energy of -9.8 kcal/mol. Consequently the differential end stability, i.e., the thermodynamic asymmetry, for the duplex is ΔΔG = 1.1 kcal/mol. Positive values of ΔΔG indicate that the sequence is asymmetric in favor of the appropriate guide strand.

qPCR

Total RNA was extracted using RNeasy Plus Mini Total RNA Purification Kit (QIAGEN) according to manufacturer's instructions. cDNA was synthesized using iScript Select cDNA Synthesis Kit (BIO-RAD laboratories) with a final volume of 20 μL

including 4 μL of 5X reaction mix, 1 μL of reverse transcriptase and 1 μg of RNA in every reaction tube. Quantitative real time PCR was performed using iQ™ SYBR® Green Supermix (BIO-RAD laboratories) on the synthesized cDNA with 12.5 μL of iQ SYBR green super mix, 5 μL primers(100nM) of GFP (forward CACCTACGGCAAGCTGACCCTGAA, reverse CCCTTCAGCTCGATGCGGTTCAC) and the , 0.5 μL of water and 2.0 μL of cDNA with each biological sample having three replicates. The DNA was denatured at 95^0C initially for 9 min and amplification was performed for 50 cycles with a denaturing temperature of 95^0C for 0.20 min, an annealing temperature of 57^0C for 1 min and an extension temperature of 72^0C for 0.30 min at each cycle. The mRNA expression levels of GFP were calculated using the threshold cycle value and by normalizing to the housekeeping gene, β-Actin.

REFERENCES

Allen, R.S., Miller, J.A., Chitty, J.A., Fist, A.J., Gerlach, W.L., and Larkin, P.J. 2008. Metabolic engineering of morphinan alkaloids by over-expression and RNAi suppression of salutaridinol 7-O-acetyltransferase in opium poppy. *Plant Biotechnol J* **6**(1): 22-30.

Allen, R.S., Millgate, A.G., Chitty, J.A., Thisleton, J., Miller, J.A., Fist, A.J., Gerlach, W.L., and Larkin, P.J. 2004. RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. *Nat Biotechnol* **22**(12): 1559-1566.

Arntzen, C., Plotkin, S., and Dodet, B. 2005. Plant-derived vaccines and antibodies: potential and limitations. *Vaccine* **23**(15): 1753-1756.

Basyuk, E., Suavet, F., Doglio, A., Bordonne, R., and Bertrand, E. 2003. Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Res* **31**(22): 6593-6597.

Bennett, S.E., Jensen, O.N., Barofsky, D.F., and Mosbaugh, D.W. 1994. UV-catalyzed cross-linking of Escherichia coli uracil-DNA glycosylase to DNA. Identification of amino acid residues in the single-stranded DNA binding site. *J Biol Chem* **269**(34): 21870-21879.

Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. 2003. Dicer is essential for mouse development. *Nat Genet* **35**(3): 215-217.

Bevilacqua, P.C. and Cech, T.R. 1996. Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry* **35**(31): 9983-9994.

Bevilacqua, P.C., George, C.X., Samuel, C.E., and Cech, T.R. 1998. Binding of the protein kinase PKR to RNAs with secondary structure defects: role of the tandem A-G mismatch and noncontiguous helixes. *Biochemistry* **37**(18): 6303-6316.

Blaszczyk, J., Gan, J., Tropea, J.E., Court, D.L., Waugh, D.S., and Ji, X. 2004. Noncatalytic assembly of ribonuclease III with double-stranded RNA. *Structure* **12**(3): 457-466.

Blaszczyk, J., Tropea, J.E., Bubunenko, M., Routzahn, K.M., Waugh, D.S., Court, D.L., and Ji, X. 2001. Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. *Structure* **9**(12): 1225-1236.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. 1998. AGO1 defines a novel locus of Arabidopsis controlling leaf development. *EMBO J* **17**(1): 170-180.

Bornsen, K.O., Gass, M.A., Bruin, G.J., von Adrichem, J.H., Biro, M.C., Kresbach, G.M., and Ehrat, M. 1997. Influence of solvents and detergents on matrix-assisted laser desorption/ionization mass spectrometry measurements of proteins and oligonucleotides. *Rapid Commun Mass Spectrom* **11**(6): 603-609.

Bramsen, J.B., Laursen, M.B., Damgaard, C.K., Lena, S.W., Babu, B.R., Wengel, J., and Kjems, J. 2007. Improved silencing properties using small internally segmented interfering RNAs. *Nucleic Acids Res* **35**(17): 5886-5897.

Bramsen, J.B., Laursen, M.B., Nielsen, A.F., Hansen, T.B., Bus, C., Langkjaer, N., Babu, B.R., Hojland, T., Abramov, M., Van Aerschot, A., Odadzic, D., Smicius, R., Haas, J., Andree, C., Barman, J., Wenska, M., Srivastava, P., Zhou, C., Honcharenko, D., Hess, S., Muller, E., Bobkov, G.V., Mikhailov, S.N., Fava, E., Meyer, T.F., Chattopadhyaya, J., Zerial, M., Engels, J.W., Herdewijn, P., Wengel, J., and Kjems, J. 2009. A large-scale chemical modification screen identifies design rules to generate siRNAs with high activity, high stability and low toxicity. *Nucleic Acids Res.*

Caplen, N.J., Parrish, S., Imani, F., Fire, A., and Morgan, R.A. 2001. Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. *Proc Natl Acad Sci U S A* **98**(17): 9742-9747.

Carmell, M.A., Girard, A., van de Kant, H.J., Bourc'his, D., Bestor, T.H., de Rooij, D.G., and Hannon, G.J. 2007. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* **12**(4): 503-514.

Chang, C.I., Yoo, J.W., Hong, S.W., Lee, S.E., Kang, H.S., Sun, X., Rogoff, H.A., Ban, C., Kim, S., Li, C.J., and Lee, D.K. 2009. Asymmetric shorter-duplex siRNA structures trigger efficient gene silencing with reduced nonspecific effects. *Mol Ther* **17**(4): 725-732.

Chen, P.Y., Weinmann, L., Gaidatzis, D., Pei, Y., Zavolan, M., Tuschl, T., and Meister, G. 2008. Strand-specific 5'-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. *RNA* **14**(2): 263-274.

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**(7051): 740-744.

Chiu, Y. and Rana, T.M. 2002. RNAi in Human Cells: Basic Structural and Functional Features of Small Interfering RNA. *Molecular Cell* **10**: 549-561.

Christensen, H.S., Daher, A., Soye, K.J., Frankel, L.B., Alexander, M.R., Laine, S., Bannwarth, S., Ong, C.L., Chung, S.W., Campbell, S.M., Purcell, D.F., and Gatignol, A. 2007. Small interfering RNAs against the TAR RNA binding protein, TRBP, a Dicer cofactor, inhibit human immunodeficiency virus type 1 long terminal repeat expression and viral production. *J Virol* **81**(10): 5121-5131.

Cook, A. and Conti, E. 2006. Dicer measures up. *Nat Struct Mol Biol* **13**(3): 190-192.

Cox, K.M., Sterling, J.D., Regan, J.T., Gasdaska, J.R., Frantz, K.K., Peele, C.G., Black, A., Passmore, D., Moldovan-Loomis, C., Srinivasan, M., Cuison, S., Cardarelli, P.M., and Dickey, L.F. 2006. Glycan optimization of a human monoclonal antibody in the aquatic plant Lemna minor. *Nat Biotechnol* **24**(12): 1591-1597.

Cullen, B.R. 2002. RNA interference: antiviral defense and genetic tool. *Nat Immunol* **3**(7): 597-599.

Daher, A., Longuet, M., Dorin, D., Bois, F., Segeral, E., Bannwarth, S., Battisti, P.L., Purcell, D.F., Benarous, R., Vaquero, C., Meurs, E.F., and Gatignol, A. 2001. Two dimerization domains in the trans-activation response RNA-binding protein (TRBP) individually reverse the protein kinase R inhibition of HIV-1 long terminal repeat expression. *J Biol Chem* **276**(36): 33899-33905.

Daniell, H., Streatfield, S.J., and Wycoff, K. 2001. Medical molecular farming: production of antibodies, biopharmaceuticals and edible vaccines in plants. *Trends Plant Sci* **6**(5): 219-226.

Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., Matthews, N., Berezikov, E., Ketting, R.F., Tavare, S., and Miska, E.A. 2008. Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the Caenorhabditis elegans germline. *Mol Cell* **31**(1): 79-90.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**(7014): 231-235.

Diederichs, S. and Haber, D.A. 2007. Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* **131**(6): 1097-1108.

DiFiglia, M., Sena-Esteves, M., Chase, K., Sapp, E., Pfister, E., Sass, M., Yoder, J., Reeves, P., Pandey, R.K., Rajeev, K.G., Manoharan, M., Sah, D.W., Zamore, P.D., and Aronin, N. 2007. Therapeutic silencing of mutant huntingtin with siRNA attenuates striatal and cortical neuropathology and behavioral deficits. *Proc Natl Acad Sci U S A* **104**(43): 17204-17209.

Ding, H., Liao, G., Wang, H., and Zhou, Y. 2007. Asymmetrically designed siRNAs and shRNAs enhance the strand specificity and efficacy in RNAi. *Journal of RNAi and Gene Silencing* **4**(1): 269-280.

Dodo, H.W., Konan, K.N., Chen, F.C., Egnin, M., and Viquez, O.M. 2008. Alleviating peanut allergy using genetic engineering: the silencing of the immunodominant allergen Ara h 2 leads to its significant reduction and a decrease in peanut allergenicity. *Plant Biotechnol J* **6**(2): 135-145.

Dong, L., Hulsmeyer, M., Durkop, H., Hansen, H.P., Schneider-Mergener, J., Ziegler, A., and Uchanska-Ziegler, B. 2003. Human CD30: structural implications from epitope mapping and modeling studies. *J Mol Recognit* **16**(1): 28-36.

Doos, B.R. 2002. Population growth and loss of arable land. *Global Environmental Change-Human and Policy Dimensions* **12**(4): 303-311.

Dorsett, Y. and Tuschl, T. 2004. siRNAs: applications in functional genomics and potential as therapeutics. *Nat Rev Drug Discov* **3**(4): 318-329.

Duarte, M., Graham, K., Daher, A., Battisti, P.L., Bannwarth, S., Segeral, E., Jeang, K.T., and Gatignol, A. 2000. Characterization of TRBP1 and TRBP2. Stable stem-loop structure at the 5' end of TRBP2 mRNA resembles HIV-1 TAR and is not found in its processed pseudogene. *J Biomed Sci* **7**(6): 494-506.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. 2001a. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**(6836): 494-498.

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

Elbashir, S.M., Martinez, J., Patkaniowska, A., Lendeckel, W., and Tuschl, T. 2001c. Functional anatomy of siRNAs for mediating efficient RNAi in Drosophila melanogaster embryo lysate. *Embo J* **20**(23): 6877-6888.

Farazi, T.A., Juranek, S.A., and Tuschl, T. 2008. The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development* **135**(7): 1201-1214.

Farran, I., Sanchez-Serrano, J.J., Medina, J.F., Prieto, J., and Mingo-Castel, A.M. 2002. Targeted expression of human serum albumin to potato tubers. *Transgenic Res* **11**(4): 337-346.

Fattal, E. and Barratt, G. 2009. Nanotechnologies and controlled release systems for the delivery of antisense oligonucleotides and small interfering RNA. *Br J Pharmacol.*

Filipowicz, W. 2005. RNAi: the nuts and bolts of the RISC machine. *Cell* **122**(1): 17-20.

Floss, D.M., Falkenburg, D., and Conrad, U. 2007. Production of vaccines and therapeutic antibodies for veterinary applications in transgenic plants: an overview. *Transgenic Res* **16**(3): 315-332.

Forstemann, K., Tomari, Y., Du, T., Vagin, V.V., Denli, A.M., Bratu, D.P., Klattenhoff, C., Theurkauf, W.E., and Zamore, P.D. 2005. Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol* **3**(7): e236.

Frankard, V., Ghislain, M., and Jacobs, M. 1992. Two Feedback-Insensitive Enzymes of the Aspartate Pathway in Nicotiana sylvestris. *Plant Physiol* **99**(4): 1285-1293.

Gan, J., Shaw, G., Tropea, J.E., Waugh, D.S., Court, D.L., and Ji, X. 2008. A stepwise model for double-stranded RNA processing by ribonuclease III. *Mol Microbiol* **67**(1): 143-154.

Gatignol, A. and Jeang, K.T. 2000. Tat as a transcriptional activator and a potential therapeutic target for HIV-1. *Adv Pharmacol* **48**: 209-227.

Gatignol, A., Laine, S., and Clerzius, G. 2005. Dual role of TRBP in HIV replication and RNA interference: viral diversion of a cellular pathway or evasion from antiviral immunity? *Retrovirology* **2**: 65.

Ghildiyal, M. and Zamore, P.D. 2009. Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10**(2): 94-108.

Golden, M.C., Resing, K.A., Collins, B.D., Willis, M.C., and Koch, T.H. 1999. Mass spectral characterization of a protein-nucleic acid photocrosslink. *Protein Sci* **8**(12): 2806-2812.

Gomord, V., Chamberlain, P., Jefferis, R., and Faye, L. 2005. Biopharmaceutical production in plants: problems, solutions and opportunities. *Trends Biotechnol* **23**(11): 559-565.

Gredell, J.A., Berger, A.K., and Walton, S.P. 2008. Impact of target mRNA structure on siRNA silencing efficiency: A large-scale study. *Biotechnol Bioeng.*

Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. 2005. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* **123**(4): 631-640.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**(7014): 235-240.

Grothe, T., Lenz, R., and Kutchan, T.M. 2001. Molecular characterization of the salutaridinol 7-O-acetyltransferase involved in morphine biosynthesis in opium poppy Papaver somniferum. *J Biol Chem* **276**(33): 30717-30723.

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 Rep* **6**(10): 961-967.

Haley, B. and Zamore, P.D. 2004. Kinetic analysis of the RNAi enzyme complex. *Nat Struct Mol Biol* **11**(7): 599-606.

Hammond, S.M. 2005. Dicing and slicing: the core machinery of the RNA interference pathway. *FEBS Lett* **579**(26): 5822-5829.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**(24): 3016-3027.

Harfe, B.D., McManus, M.T., Mansfield, J.H., Hornstein, E., and Tabin, C.J. 2005. The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. *Proc Natl Acad Sci U S A* **102**(31): 10898-10903.

He, L. and Hannon, G.J. 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* **5**(7): 522-531.

Hock, J., Weinmann, L., Ender, C., Rudel, S., Kremmer, E., Raabe, M., Urlaub, H., and Meister, G. 2007. Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. *EMBO Rep* **8**(11): 1052-1060.

Hohjoh, H. 2002. RNA interference (RNA(i)) induction with various types of synthetic oligonucleotide duplexes in cultured human cells. *FEBS Lett* **521**(1-3): 195-199.

Hohjoh, H. 2004. Enhancement of RNAi activity by improved siRNA duplexes. *FEBS Lett* **557**(1-3): 193-198.

Holen, T. 2005. Mechanisms of RNAi: mRNA cleavage fragments may indicate stalled RISC. *Journal of RNAi and Gene Silencing* **1**(1): 21-25.

Holen, T., Amarzguioui, M., Babaie, E., and Prydz, H. 2003. Similar behaviour of single-strand and double-strand siRNAs suggests they act through a common RNAi pathway. *Nucleic Acids Res* **31**(9): 2401-2407.

Holen, T., Amarzguioui, M., Wiiger, M.T., Babaie, E., and Prydz, H. 2002. Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res* **30**(8): 1757-1766.

Holen, T., Moe, S.E., Sorbo, J.G., Meza, T.J., Ottersen, O.P., and Klungland, A. 2005. Tolerated wobble mutations in siRNAs decrease specificity, but can enhance activity in vivo. *Nucleic Acids Res* **33**(15): 4704-4710.

Hong, J., Wei, N., Chalk, A., Wang, J., Song, Y., Yi, F., Qiao, R.P., Sonnhammer, E.L., Wahlestedt, C., Liang, Z., and Du, Q. 2008. Focusing on RISC assembly in mammalian cells. *Biochem Biophys Res Commun* **368**(3): 703-708.

Hutvagner, G. and Simard, M.J. 2008. Argonaute proteins: key players in RNA silencing. *Nature reviews* **9**(1): 22-32.

Jackson, A.L., Burchard, J., Schelter, J., Chau, B.N., Cleary, M., Lim, L., and Linsley, P.S. 2006. Widespread siRNA "off-target" transcript silencing mediated by seed region sequence complementarity. *RNA* **12**(7): 1179-1187.

Ji, X. 2006. Structural basis for non-catalytic and catalytic activities of ribonuclease III. *Acta Crystallogr D Biol Crystallogr* **62**(Pt 8): 933-940.

Jiang, F., Ye, X., Liu, X., Fincher, L., McKearin, D., and Liu, Q. 2005. Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila. *Genes Dev* **19**(14): 1674-1679.

Judge, A.D., Sood, V., Shaw, J.R., Fang, D., McClintock, K., and MacLachlan, I. 2005. Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. *Nat Biotechnol* **23**(4): 457-462.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* **19**(4): 489-501.

Katoh, T. and Suzuki, T. 2007. Specific residues at every third position of siRNA shape its efficient RNAi activity. *Nucleic Acids Res* **35**(4): 1-14.

Khvorova, A., Reynolds, A., and Jayasena, S.D. 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**(2): 209-216.

Kim, D.H., Behlke, M.A., Rose, S.D., Chang, M.S., Choi, S., and Rossi, J.J. 2005. Synthetic dsRNA Dicer substrates enhance RNAi potency and efficacy. *Nat Biotechnol* **23**(2): 222-226.

Kim, K., Lee, Y.S., and Carthew, R.W. 2007. Conversion of pre-RISC to holo-RISC by Ago2 during assembly of RNAi complexes. *RNA* **13**(1): 22-29.

Kim, Y.S., Lee, Y.H., Kim, H.S., Kim, M.S., Hahn, K.W., Ko, J.H., Joung, H., and Jeon, J.H. 2008. Development of patatin knockdown potato tubers using RNA interference (RNAi) technology, for the production of human-therapeutic glycoproteins. *BMC Biotechnol* **8**: 36.

Kini, H.K. and Walton, S.P. 2007. In vitro binding of single-stranded RNA by human Dicer. *FEBS Lett* **581**(29): 5611-5616.

Klattenhoff, C. and Theurkauf, W. 2008. Biogenesis and germline functions of piRNAs. *Development* **135**(1): 3-9.

Kok, K.H., Ng, M.H., Ching, Y.P., and Jin, D.Y. 2007. Human TRBP and PACT Directly Interact with Each Other and Associate with Dicer to Facilitate the Production of Small Interfering RNA. *J Biol Chem* **282**(24): 17649-17657.

Koppelman, S.J., Wensing, M., Ertmann, M., Knulst, A.C., and Knol, E.F. 2004. Relevance of Ara h1, Ara h2 and Ara h3 in peanut-allergic patients, as determined by immunoglobulin E Western blotting, basophil-histamine release and intracutaneous testing: Ara h2 is the most important peanut allergen. *Clin Exp Allergy* **34**(4): 583-590.

Kumar, N., Gammell, P., Meleady, P., Henry, M., and Clynes, M. 2008. Differential protein expression following low temperature culture of suspension CHO-K1 cells. *BMC Biotechnol* **8**: 42.

Landthaler, M., Gaidatzis, D., Rothballer, A., Chen, P.Y., Soll, S.J., Dinic, L., Ojo, T., Hafner, M., Zavolan, M., and Tuschl, T. 2008. Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA* **14**(12): 2580-2596.

Laraki, G., Clerzius, G., Daher, A., Melendez-Pena, C., Daniels, S., and Gatignol, A. 2008. Interactions between the double-stranded RNA-binding proteins TRBP and PACT define the Medipal domain that mediates protein-protein interactions. *RNA Biol* **5**(2): 92-103.

Lee, K., Zerivitz, K., and Akusjarvi, G. 1995. *Small-Scale preparation of nuclear extracts from mammalian cells*. Academic Press, London.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**(6956): 415-419.

Lee, Y., Hur, I., Park, S.Y., Kim, Y.K., Suh, M.R., and Kim, V.N. 2006. The role of PACT in the RNA silencing pathway. *EMBO J* **25**(3): 522-532.

Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. 2002. MicroRNA maturation: stepwise processing and subcellular localization. *Embo J* **21**(17): 4663-4670.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. 2004. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**(1): 69-81.

Leuschner, P.J., Ameres, S.L., Kueng, S., and Martinez, J. 2006. Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep* **7**(3): 314-320.

Li, H. and Nicholson, A.W. 1996. Defining the enzyme binding domain of a ribonuclease III processing signal. Ethylation interference and hydroxyl radical footprinting using catalytically inactive RNase III mutants. *Embo J* **15**(6): 1421-1433.

Li, S., Peters, G.A., Ding, K., Zhang, X., Qin, J., and Sen, G.C. 2006. Molecular basis for PKR activation by PACT or dsRNA. *Proc Natl Acad Sci U S A* **103**(26): 10005-10010.

Lim, S.F., Chuan, K.H., Liu, S., Loh, S.O., Chung, B.Y., Ong, C.C., and Song, Z. 2006. RNAi suppression of Bax and Bak enhances viability in fed-batch cultures of CHO cells. *Metab Eng* **8**(6): 509-522.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. 2003. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* **426**(6965): 465-469.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. 2003. 2004. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11**(6): 576-577.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**(5689): 1437-1441.

Liu, X., Howard, K.A., Dong, M., Andersen, M.O., Rahbek, U.L., Johnsen, M.G., Hansen, O.C., Besenbacher, F., and Kjems, J. 2007. The influence of polymeric properties on chitosan/siRNA nanoparticle formulation and gene silencing. *Biomaterials* **28**(6): 1280-1288.

Lykke-Andersen, K., Gilchrist, M.J., Grabarek, J.B., Das, P., Miska, E., and Zernicka-Goetz, M. 2008. Maternal Argonaute 2 is essential for early mouse development at the maternal-zygotic transition. *Mol Biol Cell* **19**(10): 4383-4392.

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

Ma, J.B., Ye, K., and Patel, D.J. 2004. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**(6989): 318-322.

Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. 2005. Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature* **434**(7033): 666-670.

MacRae, I.J., Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A. 2008. In vitro reconstitution of the human RISC-loading complex. *Proc Natl Acad Sci U S A* **105**(2): 512-517.

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. 2006. Structural basis for double-stranded RNA processing by Dicer. *Science* **311**(5758): 195-198.

Maniataki, E. and Mourelatos, Z. 2005. A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev* **19**(24): 2979-2990.

Martin, S.E. and Caplen, N.J. 2007. Applications of RNA interference in mammalian systems. *Annu Rev Genomics Hum Genet* **8**: 81-108.

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.

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J Mol Biol* **288**(5): 911-940.

Matranga, C., Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. 2005. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **123**(4): 607-620.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. 2004. Human Argonaute2 Mediates RNA Cleavage Targeted by miRNAs and siRNAs. *Mol Cell* **15**(2): 185-197.

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

Merrill, B.M., Williams, K.R., Chase, J.W., and Konigsberg, W.H. 1984. Photochemical Cross-Linking of the Escherichia-Coli Single-Stranded DNA-Binding Protein to Oligodeoxynucleotides - Identification of Phenylalanine-60 as the Site of Cross-Linking. *Journal of Biological Chemistry* **259**(17): 850-856.

Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. 2005. Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci U S A* **102**(34): 12135-12140.

Neuenschwander, U., Suter, M., and Brunold, C. 1991. Regulation of Sulfate Assimilation by Light and O-Acetyl-l-Serine in Lemna minor L. *Plant Physiol* **97**(1): 253-258.

Norris, J.L., Hangauer, M.J., Porter, N.A., and Caprioli, R.M. 2005. Nonacid cleavable detergents applied to MALDI mass spectrometry profiling of whole cells. *J Mass Spectrom* **40**(10): 1319-1326.

Nykanen, A., Haley, B., and Zamore, P.D. 2001. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell* **107**(3): 309-321.

Page, J.E. 2005. Silencing nature's narcotics: metabolic engineering of the opium poppy. *Trends Biotechnol* **23**(7): 331-333.

Parker, J.S., Roe, S.M., and Barford, D. 2004. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *Embo J* **23**(24): 4727-4737.

Parker, J.S., Roe, S.M., and Barford, D. 2005. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* **434**(7033): 663-666.

Patel, R.C. and Sen, G.C. 1998. PACT, a protein activator of the interferon-induced protein kinase, PKR. *Embo J* **17**(15): 4379-4390.

Pellino, J.L. 2007. Dicer's Role in RISC Assembly during RNA Silencing. In, pp. 216. Northwestern University, Chicago.

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

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

Prasad, R., Kumar, A., Widen, S.G., Casasfinet, J.R., and Wilson, S.H. 1993. Identification of Residues in the Single-Stranded DNA-Binding Site of the 8-Kda Domain of Rat DNA Polymerase-Beta by Uv Cross-Linking. *Journal of Biological Chemistry* **268**(30): 22746-22755.

Prat, S., Frommer, W.B., Hofgen, R., Keil, M., Kossmann, J., Koster-Topfer, M., Liu, X.J., Muller, B., Pena-Cortes, H., Rocha-Sosa, M., and et al. 1990. Gene expression during tuber development in potato plants. *FEBS Lett* **268**(2): 334-338.

Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B., and Radmark, O. 2002. Ribonuclease activity and RNA binding of recombinant human Dicer. *Embo J* **21**(21): 5864-5874.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S., and Khvorova, A. 2004. Rational siRNA design for RNA interference. *Nat Biotechnol* **22**(3): 326-330.

Rieger, R.A., McTigue, M.M., Kycia, J.H., Gerchman, S.E., Grollman, A.P., and Iden, C.R. 2000. Characterization of a cross-linked DNA-endonuclease VIII repair complex by electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom* **11**(6): 505-515.

Rivas, F.V., Tolia, N.H., Song, J.J., Aragon, J.P., Liu, J., Hannon, G.J., and Joshua-Tor, L. 2005. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* **12**(4): 340-349.

Robertson, H.D., Webster, R.E., and Zinder, N.D. 1968. Purification and properties of ribonuclease III from Escherichia coli. *J Biol Chem* **243**(1): 82-91.

Ryter, J.M. and Schultz, S.C. 1998. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *Embo J* **17**(24): 7505-7513.

Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. 2005. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. *PLoS Biol* **3**(7): e235.

Schwarz, D.S., Ding, H., Kennington, L., Moore, J.T., Schelter, J., Burchard, J., Linsley, P.S., Aronin, N., Xu, Z., and Zamore, P.D. 2006. Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genet* **2**(9): e140.

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**: 199-208.

Schwarz, D.S., Hutvagner, G., Haley, B., and Zamore, P.D. 2002. Evidence that siRNAs function as guides, not primers, in the Drosophila and human RNAi pathways. *Mol Cell* **10**(3): 537-548.

Shetlar, M.D. 1980. Photochemical and Free-Radical Initiated Reactions of 1,3-Dimethylthymine with Ethanol. *Photochemistry and Photobiology* **32**(5): 587-592.

Sicherer, S.H., Munoz-Furlong, A., Burks, A.W., and Sampson, H.A. 1999. Prevalence of peanut and tree nut allergy in the US determined by a random digit dial telephone survey. *J Allergy Clin Immunol* **103**(4): 559-562.

Singh, J., Sharp, P.J., and Skerritt, J.H. 2001. A new candidate protein for high lysine content in wheat grain. *Journal of the Science of Food and Agriculture* **81**(2): 216-226.

Siolas, D., Lerner, C., Burchard, J., Ge, W., Linsley, P.S., Paddison, P.J., Hannon, G.J., and Cleary, M.A. 2005. Synthetic shRNAs as potent RNAi triggers. *Nat Biotechnol* **23**(2): 227-231.

Siomi, H. and Siomi, M.C. 2008. Interactions between transposable elements and Argonautes have (probably) been shaping the Drosophila genome throughout evolution. *Curr Opin Genet Dev* **18**(2): 181-187.

Soifer, H.S., Sano, M., Sakurai, K., Chomchan, P., Saetrom, P., Sherman, M.A., Collingwood, M.A., Behlke, M.A., and Rossi, J.J. 2008. A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. *Nucleic Acids Res* **36**(20): 6511-6522.

Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. 2003. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* **10**(12): 1026-1032.

Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**(5689): 1434-1437.

Sontheimer, E.J. 2005. Assembly and function of RNA silencing complexes. *Nature reviews* **6**(2): 127-138.

Srivastava, S., Li, Z., Yang, X., Yedwabnick, M., Shaw, S., and Chan, C. 2007. Identification of genes that regulate multiple cellular processes/responses in the context of lipotoxicity to hepatoma cells. *BMC Genomics* **8**: 364.

Stefani, G. and Slack, F.J. 2008. Small non-coding RNAs in animal development. *Nature reviews* **9**(3): 219-230.

Sun, W., Pertzev, A., and Nicholson, A.W. 2005. Catalytic mechanism of Escherichia coli ribonuclease III: kinetic and inhibitor evidence for the involvement of two magnesium ions in RNA phosphodiester hydrolysis. *Nucleic Acids Res* **33**(3): 807-815.

Sunilkumar, G., Campbell, L.M., Puckhaber, L., Stipanovic, R.D., and Rathore, K.S. 2006. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. *Proc Natl Acad Sci U S A* **103**(48): 18054-18059.

Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. 2002. The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in C. elegans. *Cell* **109**(7): 861-871.

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 Rep* **5**(2): 189-194.

Tang, G.L., Galili, G., and Zhuang, X. 2007. RNAi and microRNA: breakthrough technologies for the improvement of plant nutritional value and metabolic engineering. *Metabolomics* **3**(3): 357-369.

Tewari, M. and Vidal, M. 2003. RNAi on the apoptosis TRAIL: the mammalian cell genetic screen comes of age. *Dev Cell* **5**(4): 534-535.

Tilesi, F., Fradiani, P., Socci, V., Willems, D., and Ascenzioni, F. 2009. Design and validation of siRNAs and shRNAs. *Curr Opin Mol Ther* **11**(2): 156-164.

Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P.D. 2004. A protein sensor for siRNA asymmetry. *Science* **306**(5700): 1377-1380.

Townsend, B.J. and Llewellyn, D.J. 2007. Reduced terpene levels in cottonseed add food to fiber. *Trends Biotechnol* **25**(6): 239-241.

Ui-Tei, K., Naito, Y., Zenno, S., Nishi, K., Yamato, K., Takahashi, F., Juni, A., and Saigo, K. 2008. Functional dissection of siRNA sequence by systematic DNA substitution: modified siRNA with a DNA seed arm is a powerful tool for mammalian gene silencing with significantly reduced off-target effect. *Nucleic Acids Res* **36**(7): 2136-2151.

Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W.S., Karpilow, J., and Khvorova, A. 2005. The contributions of dsRNA structure to Dicer specificity and efficiency. *Rna* **11**(5): 674-682.

Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D.J. 2008. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**(7219): 209-213.

Wei, M.C., Zong, W.X., Cheng, E.H., Lindsten, T., Panoutsakopoulou, V., Ross, A.J., Roth, K.A., MacGregor, G.R., Thompson, C.B., and Korsmeyer, S.J. 2001. Proapoptotic BAX and BAK: a requisite gateway to mitochondrial dysfunction and death. *Science* **292**(5517): 727-730.

Weitzer, S. and Martinez, J. 2007. The human RNA kinase hClp1 is active on 3' transfer RNA exons and short interfering RNAs. *Nature* **447**(7141): 222-226.

Whitehead, K.A., Langer, R., and Anderson, D.G. 2009. Knocking down barriers: advances in siRNA delivery. *Nat Rev Drug Discov* **8**(2): 129-138.

Williams, B.R. 1999. PKR; a sentinel kinase for cellular stress. *Oncogene* **18**(45): 6112-6120.

Wu, D. and Regnier, F.E. 1993. Native protein separations and enzyme microassays by capillary zone and gel electrophoresis. *Anal Chem* **65**(15): 2029-2035.

Wurm, F.M. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol* **22**(11): 1393-1398.

Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. 2003. Structure and conserved RNA binding of the PAZ domain. *Nature* **426**(6965): 468-474.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**(24): 3011-3016.

Zamore, P.D. 2001. Thirty-three years later, a glimpse at the ribonuclease III active site. *Mol Cell* **8**(6): 1158-1160.

Zeng, Y. and Cullen, B.R. 2003. Sequence requirements for micro RNA processing and function in human cells. *RNA* **9**(1): 112-123.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. 2002. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *Embo J* **21**(21): 5875-5885.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. 2004. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**(1): 57-68.

Zhu, X. and Galili, G. 2004. Lysine metabolism is concurrently regulated by synthesis and catabolism in both reproductive and vegetative tissues. *Plant Physiol* **135**(1): 129-136.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**(13): 3406-3415.