2013-06-02

RNA Interference by the Numbers: Explaining Biology Through Enzymology: A Dissertation

Liang Meng Wee
University of Massachusetts Medical School

RNA INTERFERENCE BY THE NUMBERS: EXPLAINING BIOLOGY THROUGH ENZYMOLOGY

A Dissertation Presented

By

LIANG MENG WEE

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

DOCTOR OF PHILOSOPHY

June 2, 2013

INTERDISCIPLINARY GRADUATE PROGRAM

RNA INTERFERENCE BY THE NUMBERS: EXPLAINING BIOLOGY THROUGH ENZYMOLOGY

A Dissertation Presented

By

LIANG MENG WEE

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

_____
Phillip D. Zamore, Ph.D., Thesis Adviser

_____
William E. Theurkauf, Ph.D., Member of Committee

_____
Reid Gilmore, Ph.D., Member of Committee

_____
Sean P. Ryder, Ph.D., Member of Committee

_____
James R. Williamson, Ph.D., Member of Committee

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

_____
Anthony Carruthers, Ph.D., Chair of Committee

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

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

Interdisciplinary Graduate Program
June 2, 2013

DEDICATION

To my supportive parents and my lovely siblings.

Thank You!

ACKNOWLEDGEMENTS

First and foremost, my gratitude goes to my “boss” Phillip Zamore who kindly allowed me to join his laboratory, to spend time mentoring me and to provide all the resources making this entire learning experience possible. I appreciate and thank Phil for all the freedom and space and at the same time the needed guidance. And his generosity to send his students to numerous conferences and courses has widened our learning environment beyond that of the lab. More amazingly, Phil recruited, supervised and trained numerous other students and postdocs and built a successful lab. I also stand to benefit given that good science equal good environment and interactions. To these, I offer my respect and a sincere thank you to my boss!

I also have to thank my committee members, Tony Carruthers, Reid Gilmore, Sean Ryder and Bill Theurkauf for being such remarkable mentors. They are always there to help and to give advice. They have been instrumental in my career as a graduate student! Also to Douglas Turner, who has allowed me to work in his lab and provided guidance to teach me how to perform RNA melts. I really enjoyed the short sabbatical training.

I was fortunate to work with past members of the lab and there are a few that I would like to mention. Yukihide Tomari is a supportive and helpful mentor and often amazed me with his experimental ideas. His calmness is always desired as an antidote to my impatient character. Klaus Förstemann’s wide scientific knowledge and high work efficiencies have always impressed me and motivates me to strive to be as good as him. Hervé Seitz, who is a remarkably careful scientist at the bench and a thoughtful one behind his computer. He is sociable and always patient to help. Stefan Ameres is an efficient and consistent scientist and whose work is always very neat and beautiful. I am also very happy to be able to overlap with Alla Sigova, Christian Matranga, Du Tingting, Megha Ghildiyal and Elif Sarinay, my excellent seniors and peers who have also been great friends and bay mates. I must also thank Li Chengjian and Jennifer Broderick who have always been around and still are with me along my

graduate life. Finally thanks to members of the Argonaute group Jennifer Broderick, Fabián Flores-Jasso and William Salomon.

Tiffanie Covello is the most efficient and helpful lab administrator. She has helped in every way to help set up dates for the initial Qualifying to the multiple TRAC meetings and finally the Defense dates. I could not thank you enough! I must also thank Alicia Boucher for being my lab “shrink”, which really helps make me fell better during stressful times. Also to Gwen Farley for managing the lab and assist to acquire all reagents that I need. It has also been a pleasant time with her as my breakfast buddy. Thank you guys for the wonderful administrative support! A special mention also to my current roommate and colleague, Bo Han for being such a wonderful lunch and dinner buddy. Finally to all current lab mates who have made working in the Zamore lab a wonderful and memorable experience! Thank you all. You will all be missed!

Finally, special thanks to my parents who have been very patient and supportive for all these years! Especially when I am so far away from home. In particular, I have to thank my brother who has taken the responsibility to look after our parents. And also to my brother-in-law, my sister and their two kids for kindly inviting me to their home while they were here in the States. Thanks again for all your hard work and sacrifices!

Wee

April, 2013

ABSTRACT

Small silencing RNAs function in almost every aspect of cellular biology. Argonaute proteins bind small RNA and execute gene silencing. The number of Argonaute paralogs range from 5 in *Drosophila melanogaster*, 8 in *Homo sapiens* to an astounding 27 in *Caenorhabditis elegans*. This begs several questions: Do Argonaute proteins have different small RNA repertoires? Do Argonaute proteins behave differently? And if so, how are they functionally and mechanistically distinct?

To address these questions, we examined the thermodynamic, kinetic and functional properties of fly Argonaute1 (dAgo1), fly Argonaute2 (dAgo2) and mouse Argonaute2 (mAGO2). Our studies reveal that in fly, small RNA duplexes sort into Argonaute proteins based on their intrinsic structures: extensively paired siRNA duplex is preferentially sorted into dAgo2 while imperfectly paired miRNA duplex is channeled into dAgo1. The sorting of small RNA is uncoupled from its biogenesis. This is exemplified by mir-277, which is born a miRNA but its extensive duplex structure licenses its entry into dAgo2. In the Argonaute protein, the small RNA guide partitions into functional domains: anchor, seed, central, 3′ supplementary and tail. Of these domains, the seed initiates binding to target.

Both dAgo2 and mAGO2 (more closely related to and a surrogate for dAgo1 in our studies) bind targets at astonishing diffusion-limited rates (~10^7–

10^8 $M^{-1}s^{-1}$). The dissociation kinetics between dAgo2 and mAGO2 from their targets, however, are different. For a fully paired target, dAgo2 dissociates slowly ($t_{½}$ ~2 hr) but for a seed-matched target, dAgo2 dissociates rapidly ($t_{½}$ ~20 s). In comparison, mAGO2 does not discriminate between either targets and demonstrates an equivalent dissociation rate ($t_{½}$ ~20 min). Regardless, both dAgo2 and mAGO2 demonstrate high binding affinity to perfect targets with equilibrium dissociation constants, K_D ~4–20 pM. Functionally, we also showed that dAgo1 but not dAgo2 silence a centrally bulged target. By contrast, dAgo2 cleaved and destroyed perfectly paired targets 43-fold faster than dAgo1. In target cleavage, dAgo2 can tolerate mismatches, bulged and internal loop in the target but at the expense of reduced target binding affinities and cleavage rates.

Taken together, our studies indicate that small RNAs are actively sorted into different Argonaute proteins with distinct thermodynamic, kinetic and functional behaviors. Our quantitative biochemical analysis also allows us to model how Argonaute proteins find, bind and regulate their targets.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ΔG: Gibbs free energy change
Ago: Argonaute
A. thaliana: *Arabidopsis thaliana*
ASO: Antisense oligonucleotide
C3PO: Component 3 promoter of RISC
CAF1-CCR4-NOT1: Deadenylation complex
C. elegans: *Caenorhabditis elegans*
Chp1: Heterochromatin associated chromodomain protein
CLRC complex: Clr4-Rik1-Cul4, Cullin-dependent E3 ubiquitin ligase
CrPV: Cricket Paralysis virus
CTCF: CCCTC-binding factor
DCL: Dicer-like
DCP1-DCP2: Decapping enzyme complex
DGCR8: DiGeorge syndrome critical region 8
D. melanogaster: *Drosophila melanogaster* (fruit fly)
eIFX: Eukaryotic Initiation Factor X
endo-siRNA: Endogenous siRNA
EMCV: Encephalomyocarditis virus
EMS: Ethylmethanesulfonate
ER: Endoplasmic reticulum
ESCRT: Endosomal sorting complex required for transport
GAPDH: Glyceraldehyde-3-phosphate dehydrogenase
GERp95: Golgi ER protein 95 kDa
H3K9me: Methylated histone H3 at lysine 9
HCV: Hepatitis C virus
HEN1: Hua Enhancer 1
HITS-CLIP: High-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation
H. sapiens: *Homo sapiens*
HP1: Heterochromatin protein 1
HSL: Histone H4 stem-loop
IRES: Internal Ribosomal Entry Site
k_{cat}: Henri-Michaelis-Menten turnover rate
K_D: Equilibrium dissociation constant
K_M: Henri-Michaelis-Menten Constant
k_{off}: Dissociation rate constant
k_{on}: Association rate constant
Loqs: Loquacious
Lsm1–7: U6 snRNA-associated Sm-like protein 1–7

m^7GpppG: 7-methyl guanosine cap
miRNA: MicroRNA
MVB: Multivesicular bodies
NEM: N-ethyl maleimide
OB: Oligonucleotide and oligosaccharide binding
PABP: Poly(A) binding protein
PACT: Protein activator of the interferon-induced protein kinase, PKR
PAIP: PABP interacting protein
Pasha: Partner of Drosha
PAZ: PIWI/Argonaute/Zwille
PIWI: P-element induced wimpy testes
PKR: Protein kinase R
Pri-miRNA: Primary miRNA
Pre-miRNA: Precursor miRNA
R2D2: Two dsRNA binding domain (R2) and Dicer-2 associated (D2)
RCK/p54: Decapping activators
RDR6: RNA-dependent RNA polymerase6
RdRP: RNA-dependent RNA polymerase
RISC: RNA-induced silencing complex
RITS: RNA–induced initiation of transcriptional gene silencing
RNAi: RNA interference
S. cerevisiae: *Saccharomyces cerevisiae* (budding yeast)
siRNA: Small interfering RNA
S. pombe: *Schizosaccharomyces pombe* (fission yeast)
TE: Transposable element
TRBP: HIV Trans-activating response (TAR) RNA-binding protein
v_0: Initial velocity
XRN1: 3′–5′ exoribonuclease

COPYRIGHT INFORMATION

Chapter I: Introduction

A Little History About Small RNAs

The phenomenon of small RNA silencing was first observed when an exogenously introduced transgene in plants suppressed an endogenous gene that bore similar sequence (Napoli et al., 1990; van der Krol et al., 1990). Subsequently, Fire and Mello, in a landmark paper, showed that double stranded RNA (dsRNA) mediated the same silencing phenomenon in *C. elegans* formally described as RNA interference (RNAi; Fire et al., 1998). Their findings also explained an earlier puzzling observation that either sense or antisense RNA injected into worms was capable of repressing gene expression (Guo and Kemphues, 1995). This is because the sense or antisense RNA was converted into dsRNA and was routed to the RNAi pathway. Since its inception, RNAi was found to operate in numerous species such as plants, fungi, worm, protozoa, fruit fly and mammalian cells (Kennerdell and Carthew, 1998; Ngo et al., 1998; Waterhouse et al., 1998; Hamilton and Baulcombe, 1999; Lohmann et al., 1999; Sanchez Alvarado and Newmark, 1999; Wianny and Zernicka-Goetz, 2000; Caplen et al., 2001; Elbashir et al., 2001a; Volpe et al., 2002; Wienholds et al., 2003; Lee and Collins, 2006). Central to RNAi are small interfering RNAs (siRNA) processed from long dsRNAs that are introduced artificially or are virally derived (Fire et al., 1998; Hamilton and Baulcombe, 1999). In short, RNAi specifically denotes posttranscriptional regulation by siRNA that silences gene by a sequence homology-mediated process (Ghildiyal and Zamore, 2009).

Small RNAs, however, come in different flavors and the founding members *lin-4* and *let-7* were originally coined small temporal RNAs because their temporal expressions control developmental timings in worm (Lee et al., 1993; Wightman et al., 1993; Reinhart et al., 2000; Pasquinelli et al., 2000). A flurry of search by cloning and small scale Sanger sequencing identified numerous species of small temporal RNAs across many organisms (Bartel, 2004). Together with *lin-4* and *let-7*, these small temporal RNAs were rebranded as microRNAs (miRNA; Lee and Ambros, 2001; Lagos-Quintana et al., 2001; Lau et al., 2001). Fueled by high throughput sequencing, rare species of miRNAs were cloned and cataloged using specific guidelines (Ambros et al., 2003).

The advent of high throughput sequencing also ushered in the endogenous siRNAs (endo-siRNA; Yang and Kazazian, 2006; Ghildiyal et al., 2008; Czech et al., 2008; Tam et al., 2008; Watanabe et al., 2008; Okamura et al., 2008a; Okamura et al., 2008b; Chung et al., 2008). This was followed closely by the discovery of the germ-line specific piRNA (Aravin et al., 2001; Aravin et al., 2006; Vagin et al., 2006; Saito et al., 2006; Girard et al., 2006; Grivna et al., 2006; Lau et al., 2006; Batista et al., 2008; Grimson et al., 2008; Das et al., 2008). Both endo-siRNA and piRNA are predicted to regulate a limited number of protein-coding genes and were shown to target and silence transposons (Kim et al., 2009). In contrast, miRNAs were estimated to regulate 60% of all human protein coding genes (Bartel, 2009; Friedman et al., 2009). Not surprisingly,

Figure 1.1

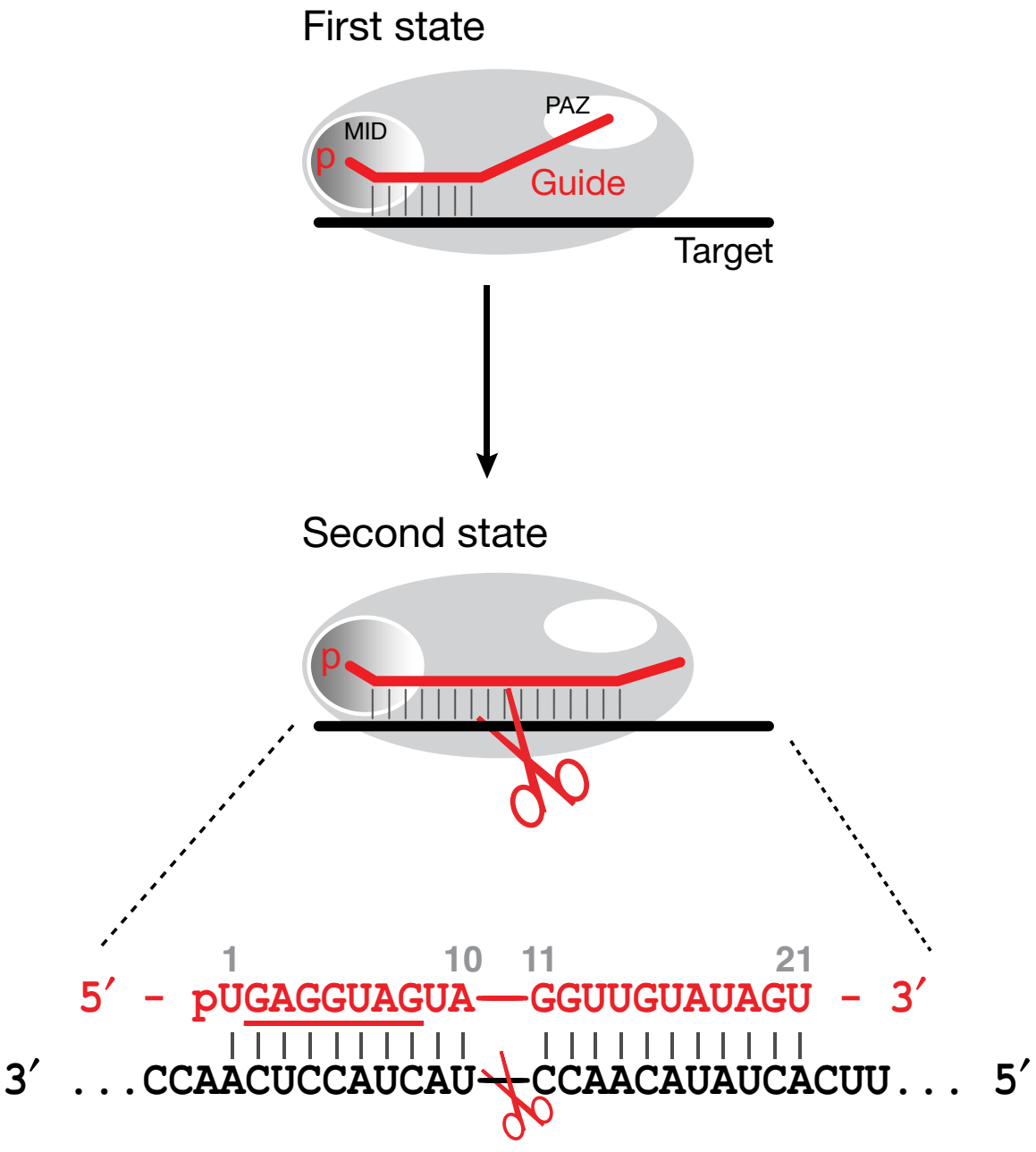

First state
PAZ
MID
p
Guide
Target
Second state
p
1
10
11
21
5′ - pUGAGGUAGUA—GGUUGUAUAGU - 3′
3′ ...CCAACUCCAUCAU—CCAACAUAUCACUU... 5′

Figure Legend 1.1. The guide defines the target cleavage site

The *let-7* guide (red) in the Argonaute protein (grey) binds target mRNA (black). Target cleavage (indicated by scissor) occurs across from nucleotide 10 (g10) and nucleotide 11 (g11) of the guide counting from its 5′ end.

small RNAs participate in numerous cellular processes such as transcription, assembling and maintaining chromatin structures, ensuring genome integrity and regulating mRNA stability. Accordingly, small RNAs function to maintain stem cells (Stadler and Ruohola-Baker, 2008), to promote cell differentiation (Li and Jin, 2009), to regulate cell proliferation (Suh and Blelloch, 2011), to fight off viral infection (Mlotshwa et.al., 2008) and to control cell death (Brennecke et. al., 2003; Ghildiyal and Zamore, 2009).

Components of Small RNA Silencing

Small Interfering RNA: A Key Component of RNAi

In plants, siRNAs mediate posttranscriptional silencing of viral RNAs (Hamilton and Baulcombe, 1999). Long dsRNA administered into fly, injected into worm, transfected into S2 cells or introduced in fly embryo lysate is processed into siRNAs by an RNase III enzyme, Dicer (Yang et al., 2000; Parrish et al., 2000; Hammond et al., 2000; Zamore et al., 2000; Elbashir et al., 2001b; Bernstein et al., 2001; Billy et al., 2001). These siRNAs are then loaded into Argonaute proteins, which execute RNAi (Elbashir et al., 2001a; Hammond et al., 2001; Hutvagner and Zamore, 2002). Biochemical studies revealed that during RNAi, Argonaute protein cleaves target mRNA between position 10 (t10) and position 11 (t11) with reference to the guide siRNA (Figure 1.1; Hammond et al., 2000; Elbashir et al., 2001c; Elbashir et al., 2001b). The use of siRNA for gene silencing in mammalian cells by RNAi was an important breakthrough as it

bypasses the need for long dsRNA, which activates the Protein kinase R (PKR) response (Elbashir et al., 2001a; Williams, 1999).

Apart from exogenously administered siRNA, naturally occurring endo-siRNA exists to repress both mRNA and transposons (Watanabe et al., 2006; Watanabe et al., 2008; Tam et al., 2008; Czech et al., 2008; Ghildiyal et al., 2008; Okamura and Lai, 2008; Chung et al., 2008; Okamura et al., 2008a; Kawamura et al., 2008). These endo-siRNAs can be processed from long dsRNA derived from convergent transcription of overlapping genes, from a gene whose transcript folds into hairpin dsRNA or from the pairing between the sense and the antisense transcript of a pseudogene (Figure 1.2; Watanabe et al., 2006; Watanabe et al., 2008; Tam et al., 2008; Czech et al., 2008; Ghildiyal et al., 2008; Okamura and Lai, 2008; Chung et al., 2008; Okamura et al., 2008a; Kawamura et al., 2008).

In fly, Dicer-2 with the help of its partner protein, R2D2 dices long dsRNA into siRNAs in a processive manner (Cenik et al., 2011; Welker et al., 2011). The siRNAs have a modal length of 21 nt that function mainly through Argonaute2 (dAgo2; Pham et al., 2004; Lee et al., 2004b; Tomari et al., 2004a; Cenik et al., 2011). In comparison, the biogenesis of endo-siRNAs requires in addition to Dicer-2, the loquacious protein isoform PD (loqs-PD; Hartig et al., 2009; Fukunaga et al., 2012). Dicer-2 and R2D2 also serve as the loading machinery to assign siRNAs into dAgo2 (Figure 1.2; Pham et al., 2004; Pham and Sontheimer,

2005; Tomari et al., 2004a). The analogous loading machinery for endo-siRNA is currently unknown. It appears that R2D2 is not required because in its absence, endo-siRNAs function remains unperturbed (Hartig et al., 2009). The processing of siRNAs by Dicer-2 and R2D2 is uncoupled from the loading of siRNAs into Argonaute proteins (Preall et al., 2006; Forstemann et al., 2007; Tomari et al., 2007). The exact mechanism how Dicer-2:R2D2 complex hands over siRNA to dAgo2 is currently unknown.

The Argonaute protein receives the siRNA duplex, removes the passenger strand and retains the single-stranded guide to form the active enzyme (Figure 1.2; Rand et al., 2005; Leuschner et al., 2006). The single-stranded guide strand in dAgo2 is then methylated on the 2′ hydroxyl group at its 3′ end by a methyltransferase HEN1/Pimet (Saito et al., 2007). In plants, however, HEN1 methylates siRNAs and miRNAs prior to their loading into Argonaute proteins (Chen et al., 2002; Park et al., 2002; Boutet et al., 2003; Yang et al., 2006; Huang et al., 2009).

MicroRNA: Tiny RNA with a Big Role in Gene Regulation

To date, the latest released of miRBase version 19, a registry that catalogs all miRNAs, reports a total of 21264 precursor miRNAs hairpins that can potentially express 25141 mature miRNAs in human, mouse, fly, worm and in *A. thaliana* (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008). In human, there are 1600 precursor miRNAs hairpins and 2042 mature miRNAs entries whereas in

Figure 1.2

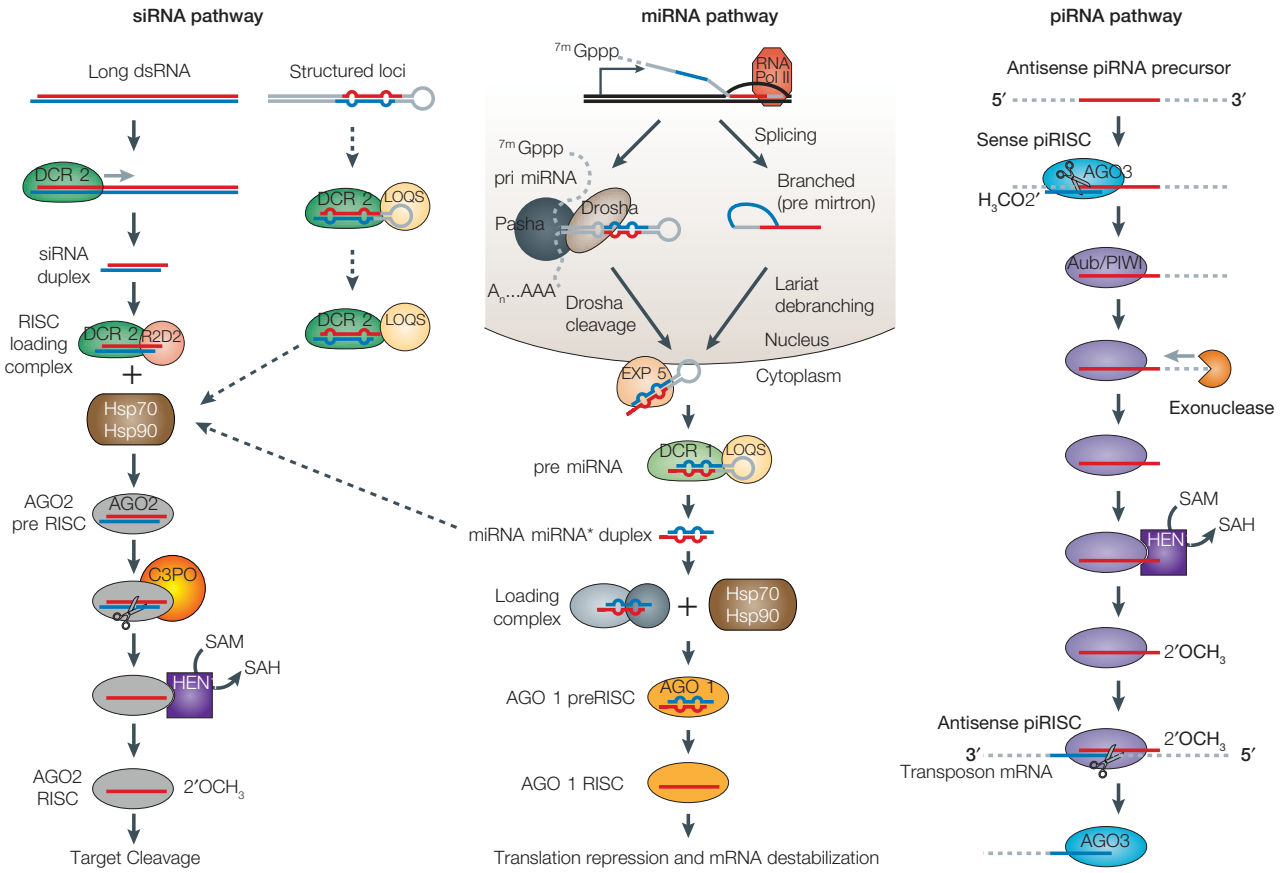


Figure Legend 1.2. The three main small RNA pathways

Adapted from Ghildiyal and Zamore, Nat. Rev. Genet., 2009.

fly, there are 238 precursor miRNAs hairpins and 426 mature miRNA entries.

In mammals, 61% of miRNA genes cluster within 50 kb of one another to produce megaclusters (Baskerville and Bartel, 2005; Chiang et al., 2010). In mouse, there are four known megaclusters: one cluster on chromosome 2 with 69 miRNA genes, one cluster on chromosome X with 18 miRNA genes and two clusters on chromosome 12 with 35 and 16 miRNA genes respectively (Calabrese et al., 2007). The remaining smaller clusters consist of an average of 2–7 miRNA genes (Calabrese et al., 2007; Chiang et al., 2010). Approximately 40% of miRNA genes originate from introns of annotated coding mRNAs while 20% of miRNA genes came from introns and exons of annotated non-coding mRNA (Rodriguez et al., 2004; Chiang et al., 2010). These miRNAs genes are presumably co-transcribed with their host genes (Rodriguez et al., 2004).

The primary miRNA transcript (pri-miRNA) is the earliest precursor of miRNA. In most cases, RNA Polymerase II and in rare cases, RNA Polymerase III is the enzyme that transcribes pri-miRNAs (Figure 1.2; Lee et al., 2004a; Cai et al., 2004; Borchert et al., 2006). In the nucleus, pri-miRNAs are recognized and processed into precursor miRNA transcript (pre-miRNA) by the microprocessor complex. The microprocessor complex consists of an RNase III enzyme Drosha and its partner protein DGCR8 in human or Pasha in fly (Figure 1.2; Lee et al., 2003; Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004; Auyeung et al., 2013). Pre-miRNAs are then exported from the nucleus

into the cytoplasm by exportin-5 (Figure 1.2; Zeng et al., 2005; Yi et al., 2003; Zeng and Cullen, 2004; Bohnsack et al., 2004; Lund et al., 2004). In the cytoplasm, another RNase III enzyme Dicer with its partner protein PACT or TRBP in human and loquacious (loqs) in fly cleaves the pre-miRNA and liberates the miRNA/miRNA* duplex (Figure 1.2; Jiang et al., 2005; Forstemann et al., 2005; Saito et al., 2005; Haase et al., 2005; Lee et al., 2006). Amazingly, some introns are spliced and debranched directly into pre-miRNAs that are therefore Drosha independent (Figure 1.2). These pre-miRNAs are then exported out of the nucleus into the cytoplasm and similarly processed by Dicer into miRNA/miRNA* duplex. These intronic-derived miRNAs are called mirtrons (Ruby et al., 2007; Okamura et al., 2007; Berezikov et al., 2007). Ultimately, mature miRNA is derived from either the 5′ or the 3′ arm of the hairpin precursor (Bartel, 2004).

Unlike siRNA that is usually 21 nt, miRNA has a modal length of 22 nt. Unlike siRNAs that pairs extensively to target mRNA, most miRNAs exhibit partial sequence complementarity to their targets (Bartel, 2009). In fly, miRNA functions mainly via Argonaute1 (dAgo1; Okamura et al., 2004). Finally, unlike siRNA in dAgo2, miRNA in dAgo1 is not methylated at its 3′ end.

Piwi-interacting RNA: Defender of the Germline

Small RNAs bound by PIWI Argonautes are known as piRNAs, short for PIWI interacting RNAs (Girard et al., 2006). Currently, the detailed mechanisms of

piRNAs biogenesis remain unknown. They originate from piRNA clusters found mainly in pericentromeric heterochromatin that contain remnants of transposable elements (TE; Brennecke et al., 2007). Some piRNAs originate from the 3′ UTR of protein coding genes and from TE rooted in euchromatic regions of the chromosome (Brennecke et al., 2007; Saito et al., 2009; Robine et al., 2009). Processive dicing of long dsRNA generates an siRNA population that is equally represented in both sense and antisense reads and demonstrates 21 nt phasing. In contrast, piRNA biogenesis does not require Dicer. The piRNA population exhibits strand bias and does not display phasing (Vagin et al., 2006). Compared to siRNAs and miRNAs, piRNAs are generally longer (~24–30 nt). This could either be due to the differences in their biogenesis or attributed to the protection footprints of their associated Argonaute proteins. Similar to siRNAs, piRNAs are methylated at their 3′ end by HEN1 (Figure 1.2; Kirino and Mourelatos, 2007a; Ohara et al., 2007; Kurth and Mochizuki, 2009; Kamminga et al., 2010; Kirino and Mourelatos, 2007b; Saito et al., 2007; Horwich et al., 2007).

In fly, piRNAs were first identified as products derived from the antisense transcripts of the *Suppressor of Stellate* (*Su(Ste)*) locus on the Y chromosome. These *Su(Ste)* piRNAs repressed the sense transcripts produced from the *Stellate (Ste)* locus on the X chromosome (Aravin et al., 2001). The piRNA machinery is mainly limited to gonadal tissues both in fly and animals and its main role is to silence selfish genetic elements and to maintain genomic integrity

critical for stable transgeneration inheritance (Figure 1.2; Williams and Rubin, 2002). Apart from understanding how piRNAs are made, current research strives to comprehend the functions of the abundant unique piRNA species and the mechanisms in which they silence their targets (Guzzardo et al., 2013).

Argonaute Protein is the Core Effector of RNAi

Argonaute proteins can be categorized into the Argonaute clade, the PIWI clade and the worm specific clade (Figure 1.3). AGO1 from *Arabidopsis thaliana* is the founding member of the Argonaute clade. The gene was isolated from an ethylmethanesulfonate (EMS) screen where its mutation in plants caused narrow leafs that resembled a squid and hence its name, Argonaute (Bohmert et al., 1998). Genetic studies in plant, worm and fly indicated that Argonaute proteins are important for organismal development and for RNAi (Wilson et al., 1996; Lin and Spradling, 1997; Cogoni and Macino, 1997; Cox et al., 1998; Tabara et al., 1999; Cox et al., 1998; Fagard et al., 2000; Kataoka et al., 2001; Grishok et al., 2001; Tijsterman et al., 2002; Vastenhouw et al., 2003).

In vitro biochemical characterization of RNAi using *Drosophila* S2 cell extract identified the RNA-induced silencing complex (RISC) that ablated target mRNA in a sequence-specific manner (Hammond et al., 2000). Further analysis determined that RISC contains Argonaute protein (Hammond et al., 2001). Subsequently, Argonaute protein was shown to be the enzyme in RISC directly responsible for the endonucleolytic cleavage of target mRNA and is Mg^{2+}-

dependent (Liu et al., 2004; Song et al., 2004; Schwarz et al., 2004).

Figure 1.3

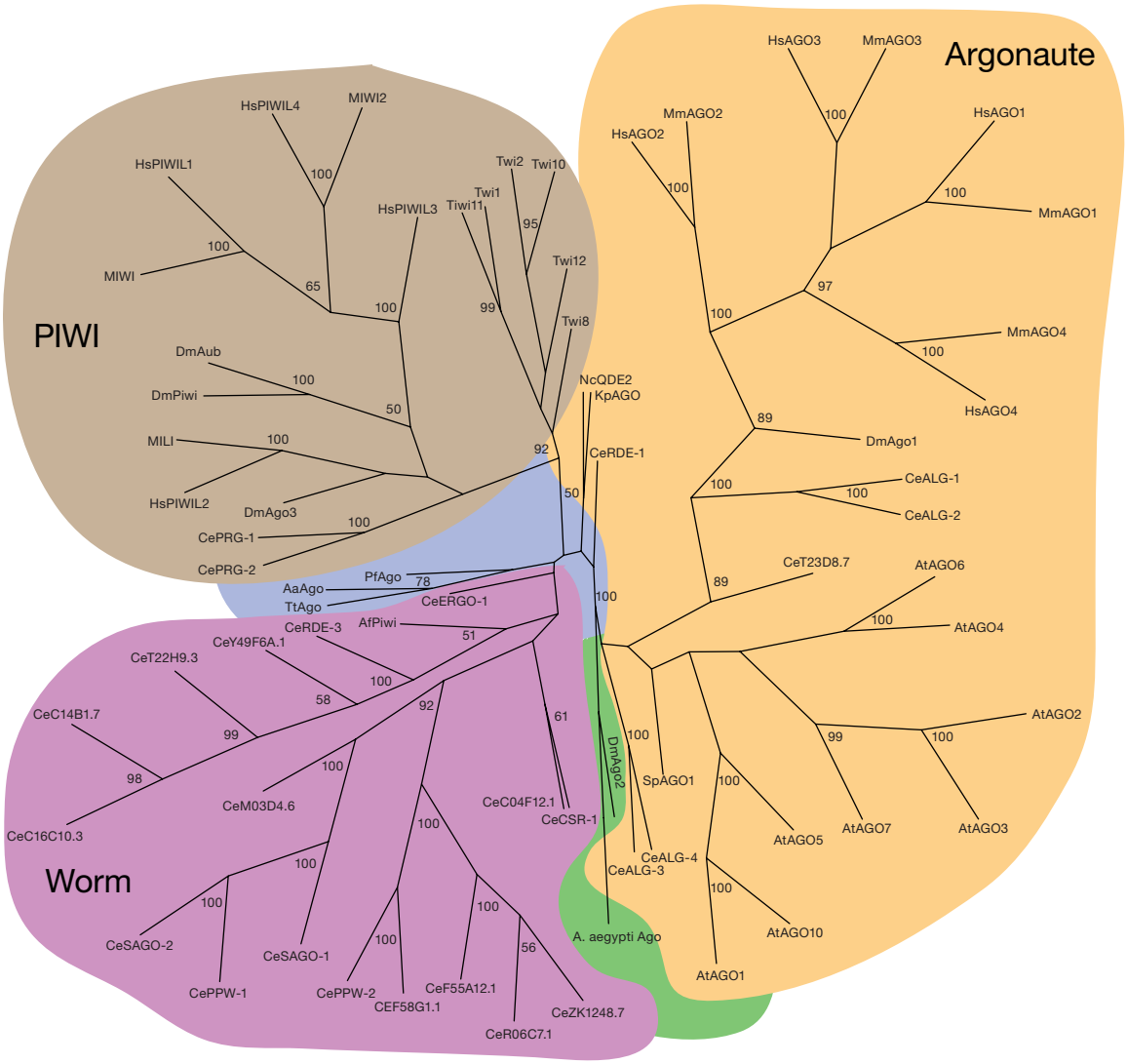


Figure Legend 1.3. Argonaute proteins are categorized into three main clades by phylogeny

Amino acid sequences of Argonaute proteins are aligned using ClustalX. Aligned sequences are fed into Phylip for bootstrapping, to calculate protein distance, and to construct consensus tree. Bootstrap percentages greater than 50% are indicated at the forks. Aa: *Aquifex aeolicus*, A. aegypti: *Aedes aegypti* Af: *Archaeoglobus fulgidus* At: *Arabidopsis thaliana*, Ce: *Caenorhabditis elegans*, Dm: *Drosophila melanogaster*, Hs: *Homo sapiens*, Kp: *Kluyveromyces polysporus* Mm: *Mus musculus*, Nc: *Neurospora crassa*, Pf: *Pyrococcus furiosus*, Sp: *Schizosaccharomyces pombe*, Tt: *Thermus thermophilus*, Twi: *Tetrahymena* Piwi Argonaute.

The Structure of Argonaute Protein

Argonaute Proteins Share Similar Architecture

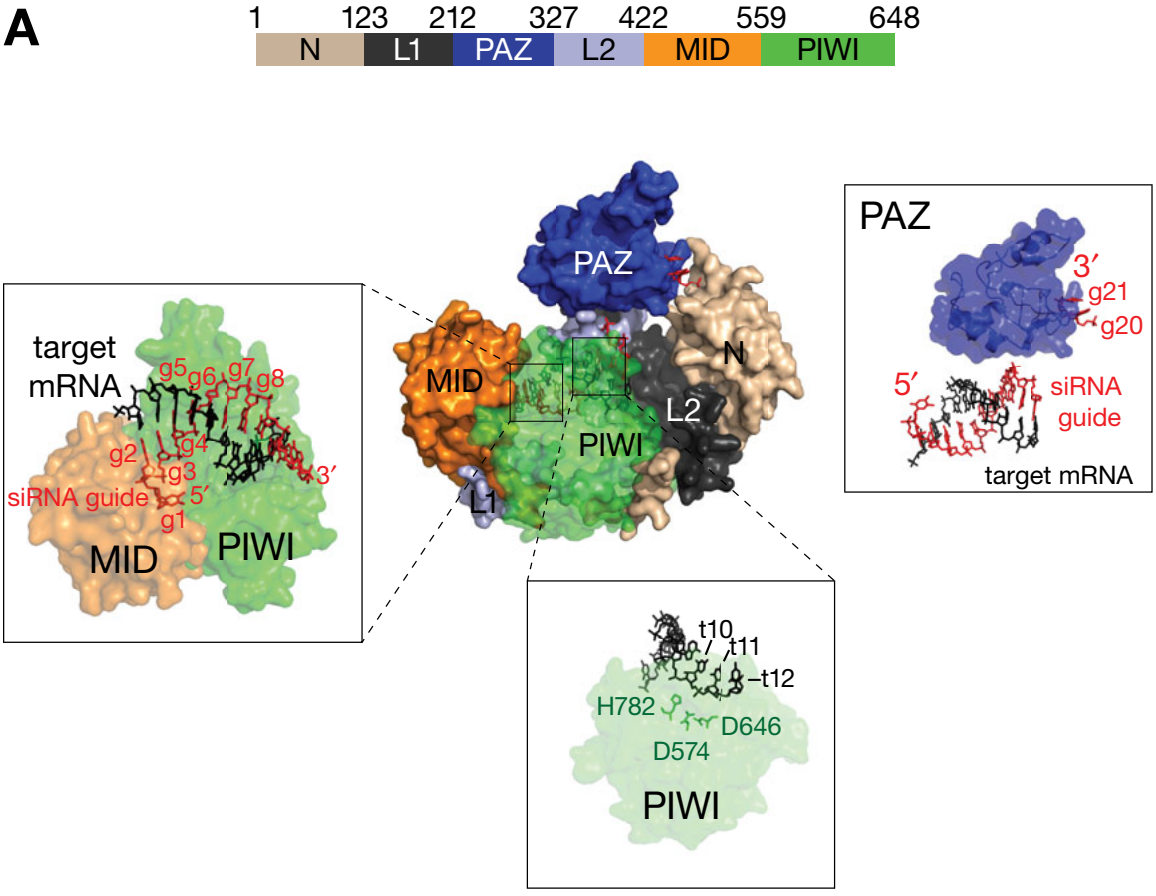
Eukaryotic Argonautes consist of four distinct domains: the N-terminal, the PAZ, the MID, and the PIWI domains (Figure 1.4A; Carmell et al., 2002; Hutvagner and Simard, 2008; Cenik and Zamore, 2011). The full structures of Argonaute proteins with all four domains came initially from eubacteria and archaebacteria (Table 1.1). Ironically, the biological functions of Argonaute proteins in these species are still not known (Parker et al., 2004; Parker et al., 2005; Song et al., 2004; Rivas et al., 2005; Ma et al., 2005; Yuan et al., 2006; Yuan et al., 2005; Rashid et al., 2007; Wang et al., 2008b; Wang et al., 2008a; Wang et al., 2009). Earlier attempts to solve the structures of eukaryotic Argonautes were successful only for the individual domains (Table 1.1; Lingel et al., 2003; Yan et al., 2003; Song et al., 2003; Lingel et al., 2004; Ma et al., 2004; Kiriakidou et al., 2007; Maiti et al., 2007; Tian et al., 2011b; Simon et al., 2011; Frank et al., 2010; Boland et al., 2010; Frank et al., 2011; Boland et al., 2011). Recently, the full-length structures of yeast and human AGO2 proteins were crystallized (Schirle and MacRae, 2012; Elkayam et al., 2012; Nakanishi et al., 2012). Comparing all current solved structures, archaebacterial, eubacterial and eukaryotic Argonaute proteins adopt similar folds. The human AGO2 structures were solved in the presence of the guide strands (Schirle and MacRae, 2012; Elkayam et al., 2012). The complete structure of mammalian Argonaute protein with both the guide

Table 1.1. Solved structures of Argonaute proteins

Structures of Argonaute proteins			
S/N	Descriptions	Organisms	References
1	Argonaute2 PAZ domain (NMR)	*Drosophila melanogaster*	(Lingel et al., 2003)
2	Argonaute1 PAZ domain (NMR)	*Drosophila melanogaster*	(Yan et al., 2003)
3	Argonaute2 PAZ domain (X-Ray)	*Drosophila melanogaster*	(Song et al., 2003)
4	Argonaute2 PAZ domain + RNA or DNA oligonucleotide (NMR)	*Drosophila melanogaster*	(Lingel et al., 2004)
5	PIWI	*Archaeoglobus fulgidus*	(Parker et al., 2004)
6	Argonaute2 PAZ domain + siRNA duplex (X-Ray)	*Homo sapiens*	(Ma et al., 2004)
7	Argonaute protein	*Pyrococcus furiosus*	(Song et al., 2004)
8	Argonaute protein + tungsten	*Pyrococcus furiosus*	(Rivas et al., 2005)
9	Argonaute protein	*Aquifex aeolicus*	(Yuan et al., 2005)
10	PIWI + siRNA guide	*Archaeoglobus fulgidus*	(Parker et al., 2005; Ma et al., 2005)
11	Argonaute protein + 22 or 26-mer siRNA	*Aquifex aeolicus*	(Yuan et al., 2006)
12	Argonaute protein	*Aquifex aeolicus*	(Rashid et al., 2007)
13	Argonaute protein + 21 nt DNA guide	*Thermus thermophilus*	(Wang et al., 2008)
14	Argonaute protein + 21 nt DNA guide + 20 nt RNA target	*Thermus thermophilus*	(Wang et al., 2008)
15	Argonaute protein + 21 nt DNA guide + 12, 15 and 19 nt RNA target	*Thermus thermophilus*	(Wang et al., 2009)
16	PIWI + DNA duplex	*Archaeoglobus fulgidus*	(Parker et al., 2009)
17	MID domain of Argonaute protein (QDE-2)	*Neurospora crassa*	(Boland et al., 2010)

18	MID domain of Argonaute2 + nucleoside monophosphate	*Homo sapiens*	(Frank et al., 2010)
19	Hili PAZ (free state) and Hiwi1 PAZ domains + 14-mer RNA duplex	*Homo sapiens*	(Tian et al., 2011)
20	MIWI PAZ domain + 8-mer RNA (NMR)	*Mus musculus*	(Simon et al., 2011)
21	MID and PIWI lobe of QDE-2	*Neurospora crassa*	(Boland et al., 2011)
22	MID domain of Argonaute2 + cap analogues	*Homo sapiens*	(Frank et al., 2011)
23	MID domain of Argonaute1	*Arabidopsis thaliana*	(Zha et al., 2012)
24	Argonaute protein + RNA guide	*Kluyveromyces polysporus*	(Nakanishi et al., 2012)
25	Argonaute2 protein + RNA guide	*Homo sapiens*	(Schirle and MacRae, 2012)
26	Argonaute2 protein + mir-20a RNA guide	*Homo sapiens*	(Elkayam et al., 2012)

Figure 1.4


A
1
123
212
327
422
559
648
N
L1
PAZ
L2
MID
PIWI
target
mRNA
siRNA guide
MID
PIWI
PAZ
L2
N
L1
PAZ
3′
g21
g20
5′
siRNA
guide
target mRNA
t10
t11
t12
H782
D646
D574
PIWI

B


seed
3′ supplementary
1
2
8
21
5′-pUGAGGUAGUAGGUUGUAUAGU-3′
13
16
anchor
central
tail

Figure Legend 1.4. Predicted structure of *Drosophila* Ago2

(A) The amino acid sequence of fly Ago2 is threaded using the I-TASSER online platform (Roy et al., 2010). The predicted structure is then superimposed onto the solved structure of *Thermus thermophilus* Argonaute with DNA guide and RNA target (Protein Data Bank accession code 3HO1). The additional panels zoom in on the seed, central and PAZ domains to reveal base pairing between the guide and target. (B) The functional domains of siRNA guide.

and target is much anticipated as it may reveal how mammalian Argonaute protein rearranges its structure to accommodate the target. Furthermore, it may disclose how mammalian Argonaute protein achieves RNA targeting specificity. We can predict that the Argonaute protein either makes specific contacts with the 2′ hydroxyl group of the ribose ring of the target RNA or that it favors an RNA:RNA helix of the guide:target duplex; and they are not mutually exclusive.

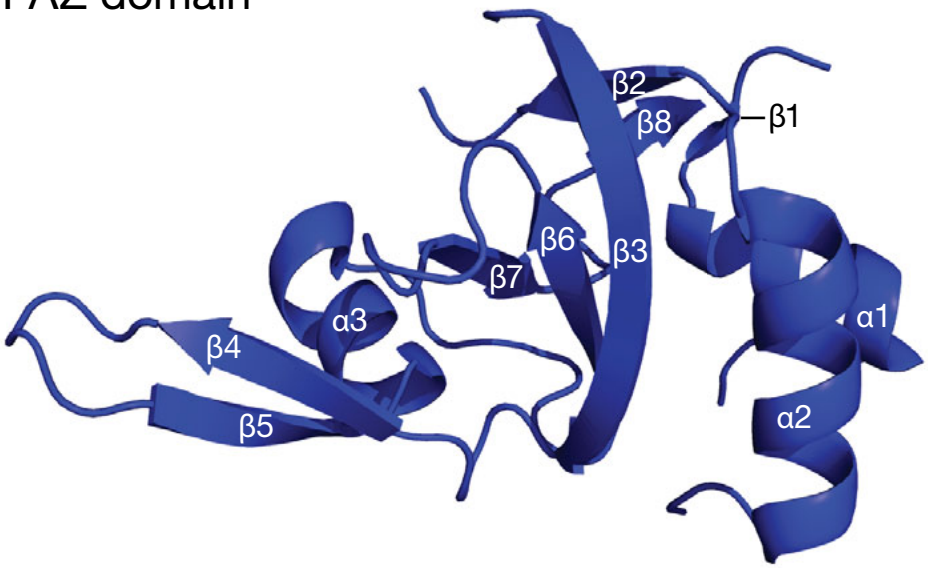
The PAZ Domain Binds the 3′ end of Small RNA Guide

The free structure of the PAZ (PIWI/Argonaute/Zwille) domain features a preformed deviant OB (oligonucleotide and oligosaccharide binding) fold that consists of a central 5-stranded β-barrel (β2–β3; β6–β8) flanked on one side by β1, α1, α2 and on the other by a β4, β5, α3 module (Figure 1.5; Lingel et al., 2003; Yan et al., 2003; Song et al., 2003).

The β4, β5, α3 module and the central β-barrel form a cleft-like structure lined and stabilized by conserved aliphatic and aromatic residues. In an RNA-bound form, these residues interact with the sugar rings of the RNA molecule by Van der Waals forces and stack along the faces of nucleotide bases (Lingel et al., 2004). Electrostatic interactions were also observed between the PAZ residues and the non-bridging oxygen of the oligonucleotides without making specific contacts with base edges. This explains why the PAZ domain can bind different oligonucleotide sequences (Ma et al., 2004).

Figure 1.5

PAZ domain



PAZ with guide strand

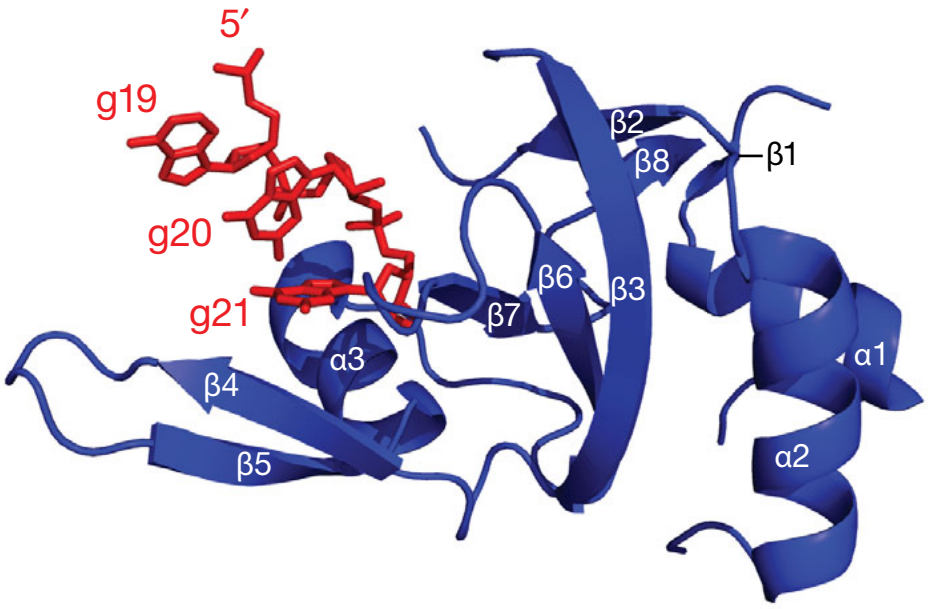


Figure Legend 1.5. PAZ domain of human AGO2

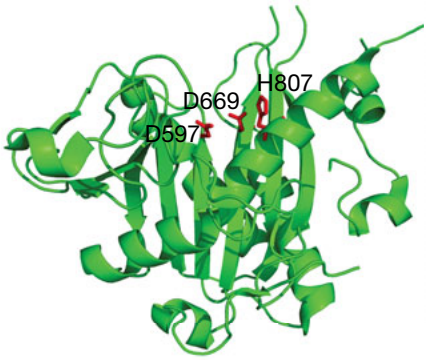
The PAZ domain from human AGO2 (Protein Data Bank accession code 4F3T) with (top) and without (bottom) mir-20a guide (g19–g20) displayed using Pymole. See text for detailed descriptions.

The PAZ domain binds single stranded RNA or double stranded RNA with 3′ 2 nt overhang and has limited tolerance for bulky addition at the 3′ end of RNA (Ma et al., 2004; Lingel et al., 2003; Yan et al., 2003). In PIWI Argonaute, the RNA binding pocket of the PAZ domain widen to accommodate the 2′-*O*-methyl group added to the 3′ end of piRNAs (Tian et al., 2011b; Simon et al., 2011). The PAZ domain alone binds RNA with low micromolar affinity whereas the full mammalian Argonaute protein binds miRNA at low nanomolar affinity (Tan et al., 2009; Lima et al., 2009). Therefore, the PAZ domain is neither the main nor the only site that holds onto the small RNA guide (Lingel et al., 2003; Yan et al., 2003; Song et al., 2003; Ma et al., 2004).

The Catalytic Residues Sit In The Highly Conserved PIWI Domain

The general architecture of the PIWI domain is composed of central β-sheets closed in on both sides by α-helices similar to that observed in HIV integrase, *E. coli* RuvC, RNase H1 and RNase H2 (Figure 1.6; Yang and Steitz, 1995; Parker et al., 2004; Rivas et al., 2005; Nowotny, 2009). The PIWI domain is also the most conserved domain among Argonaute proteins and it adopts an RNase H fold that harbors the catalytic residues responsible for target cleavage (Figure 1.6; Song et al., 2004; Liu et al., 2004). Target cleavage requires Mg^{2+} ions and can potentially be mediated by a two-metal-ion mechanism (Figure 1.7; Steitz and Steitz, 1993; Schwarz et al., 2004; Rivas et al., 2005). One Mg^{2+} ion produces the hydroxide ion from water molecule for in-line nucleophilic attack

Figure 1.6



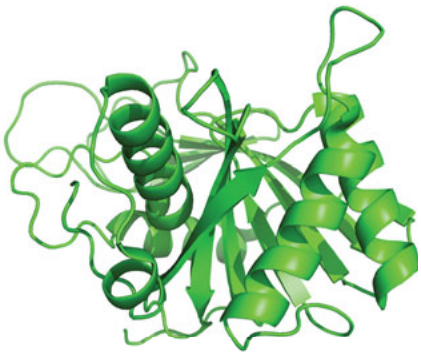
Argonaute PIWI domain
Homo sapiens (PDB: 4F3T)

RNase H2
Archaeoglobus fulgidus (PDB: 1I39)

RNase H1
Bacillus halodurans (PDB: IZB1)

Integrase
HIV (PDB: 1BIU)

RuvC
Escherichia coli (PDB: 1HJR)



PIWI domain
Pyrococcus furiosus (PDB: 1U04)

Figure Legend 1.6. PIWI domain of human AGO2

The PIWI domain from human AGO2 (Protein Data Bank accession code 4F3T) has similar folds to PIWI domain from *Pyrococcus furiosus*, RNase H1, RNase H2, HIV Integrase, and RuvC. The catalytic residues (D597, D669 and H807) in the PIWI domain of human AGO2 are displayed as red sticks. See text for detailed descriptions. Examples are taken from Nowotny, EMBO Rep, 2009.

Figure 1.7

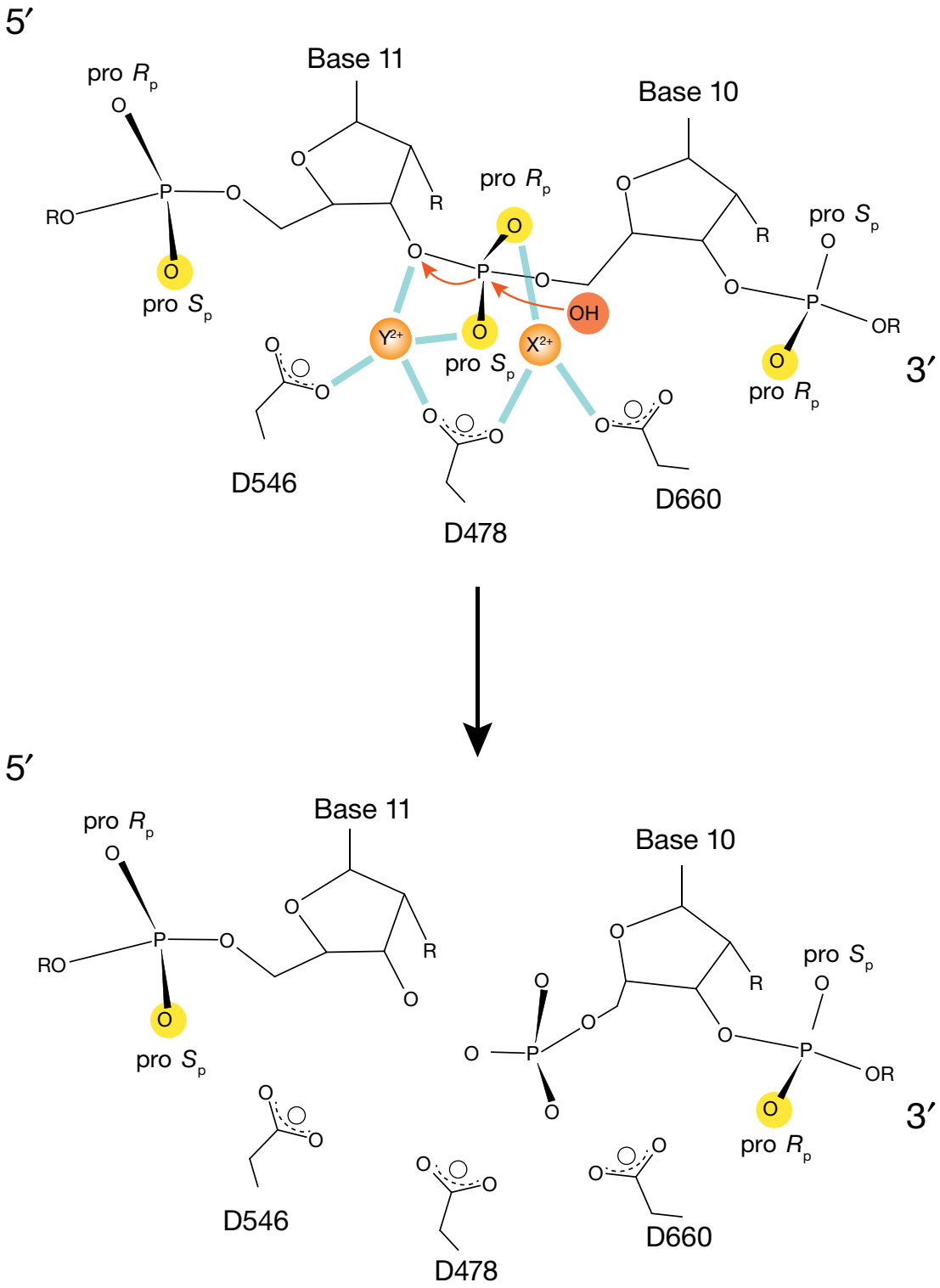

5′
pro R_p
Base 11
Base 10
pro R_p
pro S_p
RO
pro S_p
OH
Y^{2+}
X^{2+}
pro S_p
pro R_p
3′
D546
D478
D660
5′
pro R_p
Base 11
Base 10
RO
pro S_p
pro S_p
pro R_p
3′
D546
D478
D660

Figure Legend 1.7. Divalent metal ions mediate cleavage at the catalytic site

Two divalent metal ions positioned by three conserved aspartate residues facilitate target mRNA cleavage by phosphoryl transfer reaction. Metal ion X^{2+} activates a water molecule (OH^-) for in-line attack on the scissile phosphate and makes bond contact to the pro-R_p non-bridging oxygen. Metal ion Y^{2+} stabilizes the trigonal bipyramidal arrangement of the scissile phosphate intermediate through bond contact with the bridging and to the pro-S_p non-bridging oxygen. Metal Y^{2+} also facilitates the departure of the oxyanion intermediate. D546 coordinates metal ion Y^{2+} while D660 coordinates X^{2+} but D478 makes contact with both metal ions. For an RNA target, R = OH. In fly and human AGO2, a histidine residue replaces the third aspartate residue.

on the scissile phosphate while the other Mg^{2+} ion stabilizes the pentacovalent transition state and facilitates the leaving of the 3′ oxyanion group (Figure 1.7; Steitz and Steitz, 1993; Nowotny and Yang, 2009). Eventually, the cleaved RNA is left with a 5′ phosphate and a 3′ hydroxyl group (Martinez and Tuschl, 2004).

The residues in the PIWI domain make numerous contacts with the phosphodiester backbone of the guide strand; contacts among the protein residues and the nucleobases are scarce (Wang et al., 2008a). This therefore allows Argonaute proteins to capture RNA guides with different sequences. In the structure of human AGO2, residues in the PIWI domain interact with the 2′ hydroxyl group of the sugar moiety of the guide strand, which may explain why RNA instead of DNA guide is preferred. The guide strand interacts extensively with the PIWI and MID domains, which preorder the seed sequence (g2–g8) so that it can bind its target at a reduced entropic cost (Parker et al., 2009). In contrast, residues in the PIWI domain make minimal contact with the target thus allowing the target to dissociate after its cleavage (Parker et al., 2009; Ma et al., 2005; Wang et al., 2008a). The minimal contact with target also explains why Argonaute protein tolerated bulged loop in the target but not in the guide strand (Wang et al., 2008a).

The MID Domain Selects for Small RNA Guide

The MID domain adopts a Rossmann-like fold with 4 alternating α-helices and β-strands (Figure 1.8A). The β-strands form the domain core sandwiched by the

Figure 1.8

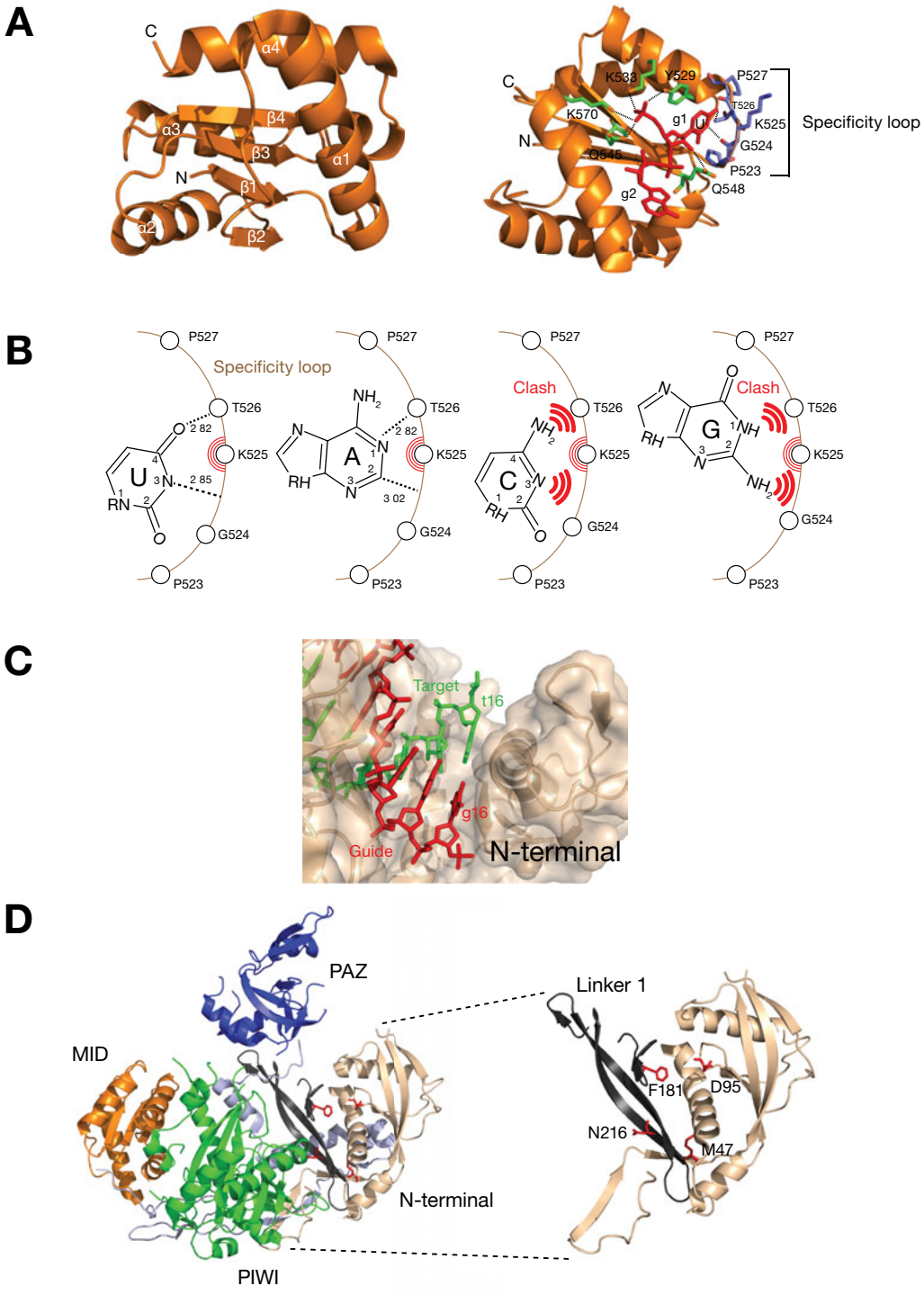

A
C
α4
β4
α3
β3
α1
N
β1
α2
β2
K533
Y529
P527
T526
K570
g1
K525
Specificity loop
N
G524
Q545
P523
g2
Q548
B
P527
Specificity loop
T526
K525
G524
P523
Clash
C
G
A
U
RH
RN
C
Target
t16
g16
Guide
N-terminal
D
PAZ
MID
Linker 1
F181
D95
N216
M47
N-terminal
PIWI

Figure Legend 1.8. The MID and N-terminal domains flank the 5′ and 3′ end of the siRNA guide respectively

(A) The 5′ phosphate of the guide nucleotide interacts with conserved residues of the MID domain. The first nucleotide is unpaired and is accommodated in the MID domain by the specificity loop. The MID domain is from human AGO2 (Protein Data Bank accession code 4F3T). (B) The MID domain accommodates uracil (U) and adenine (A) as the first nucleotide of the guide strand. The specificity loop clashes with cytosine (C) and guanine (G) as the first nucleotide of the guide strand. (C) The N terminal domains blocks guide:target (g16:t16) pairing in *Thermus thermophilus* Argonaute (Protein Data Bank accession code 3HK2). (D) Residues in N-terminal domain and linker 1 of human Ago2 required to separate siRNA duplex are shown as red sticks.

α-helices (Frank et al., 2010; Boland et al., 2010). In human AGO2, a specificity loop between β3 and α3 selects for guide strand that begins with either a uracil or an adenine but clashes with a cytosine or a guanine (Figure 1.8B; Frank et al., 2010). The specificity loops are also present in plant Argonaute proteins and most likely also exist in worm and fly Argonaute proteins given that first nucleotide bias are observed in small RNA guides isolated from Argonaute proteins in these species (Lau et al., 2001; Ghildiyal et al., 2008; Montgomery et al., 2008; Mi et al., 2008; Ghildiyal et al., 2009). It will be interesting to test if loop or domain swap among Argonaute proteins alter their first nucleotide preferences and their small RNA repertoires in worm and fly.

The MID and PIWI Domains Cooperate to Anchor the 5′ end of Small RNA Guide

Together, the MID and the PIWI domain constitute the PIWI fold (Figure 1.4A; Parker et al., 2004). The interface between the MID and the basic conserved pocket of the PIWI domain anchors the 5′ phosphate end of the siRNA guide. The 5′ phosphate of the first nucleotide (g1) and the phosphodiester group linking the second (g2) and the third (g3) nucleotides of the guide strand are coordinated by magnesium (Mg^{2+}) ions with the aid of several conserved residues in the PIWI fold. This configuration distorts the trajectory of g1 from the rest of the guide nucleotides and prevents base pairing to the corresponding t1 target nucleotide (Figures 1.4A and 1.8A; Parker et al., 2005; Ma et al., 2005).

Therefore, a first nucleotide mismatch will not affect target binding and cleavage by Argonaute proteins (Haley and Zamore, 2004; Wee et al., 2012). Furthermore, bioinformatics analysis of miRNA target indicates a preference for an adenosine at t1 regardless of its potential to pair with g1 of the guide (Lewis et al., 2005).

Taking reference from the 5′ anchor, Argonaute cleaved its target across from g10 and g11 of the guide (Figure 1.1; Elbashir et al., 2001c; Elbashir et al., 2001b; Haley and Zamore, 2004). In the absence of the 5′ phosphate on the guide strand, human AGO2 can cleave its target at the wrong site—1 nt upstream of the actual scissile phosphate. This suggests that the 5′ phosphate is required to ensure that the guide is securely attached to the MID-PIWI interface to prevent guide slippage so as to enforce cleavage site fidelity (Rivas et al., 2005).

The N-terminal Domain Pries Open siRNA duplex

The PIWI, the MID and the N-terminal domain form a crescent-like base (Figure 1.8D). The PAZ domain sits above the crescent connected by a three-stranded antiparallel β sheet stalk that defines the linker 1 domain (Figure 1.8D). The Argonaute structure of *Thermus thermophilus* in the presence of a DNA guide and an RNA target revealed that the N-terminal domain prohibits base pairing beyond nucleotide 16 (Figure 1.8C; Wang et al., 2009). In *Kluyveromyces polysporus* Argonaute, a modeled A-form duplex appeared unobstructed by the N-terminal domain (Nakanishi et al., 2012). These data suggest that the extent of

guide:target pairing may differ among Argonaute proteins of eubacterial and of eukaryotes origin.

Biochemical analysis of N-terminal mutants suggests that the N-terminal domain drives duplex separation during human AGO2 assembly (Figure 1.8D; Kwak and Tomari, 2012). In these mutants, passenger strand cleavage was compromised and remained trapped in the Argonaute protein. Subsequent cleavage of target and the released of the sliced mRNA, however, were not affected. These contradictory results between passenger strand and target mRNA removal from Argonaute proteins suggest that passenger strand and target when bound to the guide strand may be topologically different.

The Making of an Active Argonaute Protein

RISC Assembly in Fly Consumes ATP

The process of RISC assembly is well studied in fly (Kawamata and Tomari, 2010). It begins when Dicer-2 uses ATP to process long dsRNA into siRNAs (Figure 1.2; Zamore et al., 2000; Nykanen et al., 2001). The RISC loading complex (RLC) that consists of Dicer-2 with the help of its partner protein R2D2 accomplishes the loading of siRNAs into dAgo2 (Figure 1.2; Liu et al., 2003; Pham et al., 2004; Tomari et al., 2004a; Pham and Sontheimer, 2005). The processing and the loading of siRNAs into dAgo2, however, are uncoupled (Preall et al., 2006; Forstemann et al., 2007; Tomari et al., 2007). During loading, the RLC senses the thermodynamically asymmetric ends of the siRNA duplex

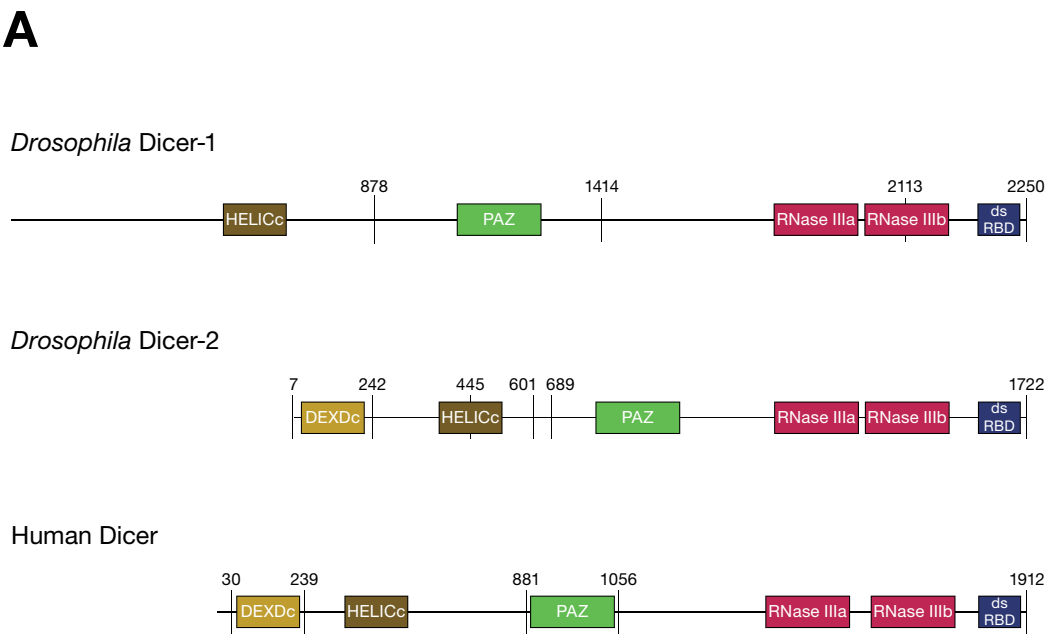
such that Dicer-2 binds the least stable end whereas R2D2 binds the more stable ends (Tomari et al., 2004b). This defines the orientation of the duplex when it gets transferred into dAgo2. The exact mechanism how the RLC hands over the siRNA duplex to dAgo2 is unknown. Eventually, the RNA with the less stable 5′ end lies closer to Dicer-2 and is retained in dAgo2 as the guide while the complementary passenger strand is discarded.

Unfortunately, there are no solved structures for mammalian Dicer or for its partner proteins. The general architecture of the Dicer came from electron micrograph (EM) that described an L-shaped particle consisting of a head, a body and a base with an extended arm (Figure 1.9; Lau et al., 2009; Wang et al., 2009; Lau et al., 2012). The PAZ domain sits in the head, the RNase III domains localize to the end of the body that connects to the helicase domains that occupy the base with the protruding arm (Figure 1.9B; Lau et al., 2012). Docking of the *Thermus thermophilus* Argonaute onto the EM of Dicer suggests that Argonaute binds close to the body of Dicer and spans the distance between the head and the base regions of Dicer (Figure 1.9B; Wang et al., 2009). It will be a challenge to capture and to crystallize snapshots of the interactions in order to reveal how Dicer hands over the siRNA duplex to Argonaute proteins.

Heat Shock Proteins Assist in RISC Assembly

The loading of siRNA duplex into Argonaute protein requires ATP. Recent studies demonstrated that the heat shock chaperone machinery (hsc70/hsp90)

Figure 1.9


A
Drosophila Dicer-1
878
1414
2113
2250
HELICc
PAZ
RNase IIIa
RNase IIIb
ds RBD
Drosophila Dicer-2
7
242
445
601
689
1722
DEXDc
HELICc
PAZ
RNase IIIa
RNase IIIb
ds RBD
Human Dicer
30
239
881
1056
1912
DEXDc
HELICc
PAZ
RNase IIIa
RNase IIIb
ds RBD

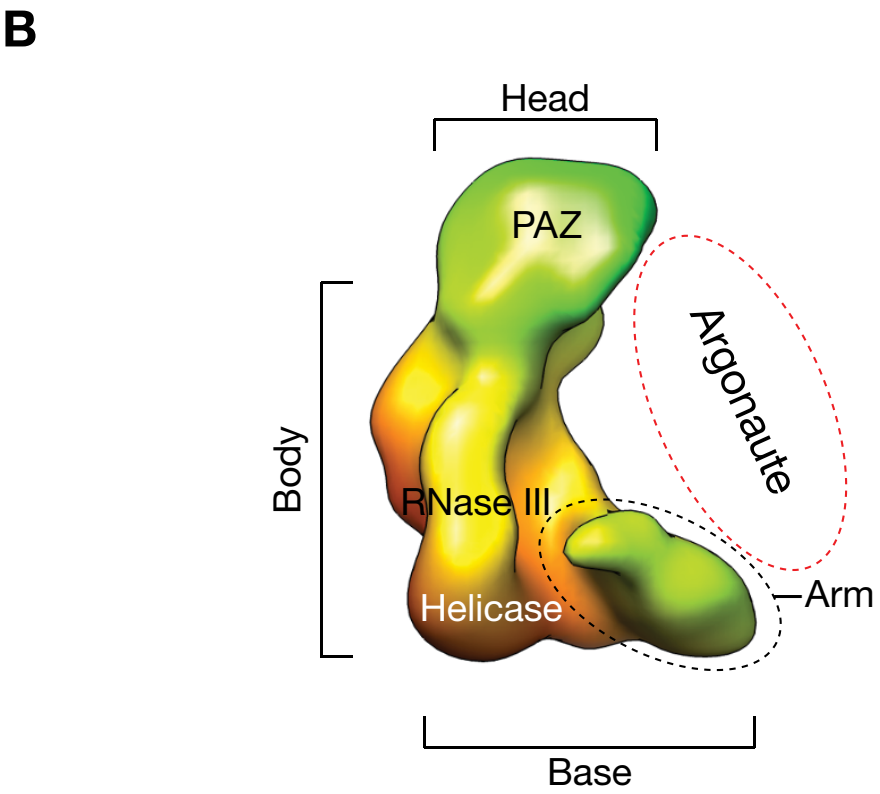

B
Head
PAZ
Argonaute
Body
RNase III
Helicase
Arm
Base

Figure Legend 1.9. Dicer adopts an L-shaped structure

(A) The domains organization of fly and human Dicer (B) Electron micrograph indicated that Dicer adopts an L-shaped structure that may interact with Argonaute protein that straddles between the head and the base regions of Dicer (Lau et al., 2009).

functions in the loading process (Figure 1.2; Johnston et al., 2010; Iki et al., 2010; Iwasaki et al., 2010; Miyoshi et al., 2010). Hsc70 and hsp90 contain ATPase domains that may explain why ATP is consumed during RISC assembly. Specific inhibitors of heat shock proteins such as geldanamycin, radicicol and 17-N-Allylamino-17-demethoxygeldanamycin (17AAG) inhibit interactions of heat shock proteins with Argonaute proteins. These inhibitors also prevent loading of siRNA or miRNA duplex into Argonaute proteins (Roe et al., 1999; Iki et al., 2010; Iwasaki et al., 2010; Miyoshi et al., 2010). By contrast, these inhibitors did not affect processes downstream of duplex loading: preassembled RISC in the presence of these inhibitors retains ability to remove passenger strand and to perform target cleavage (Iwasaki et al., 2010; Miyoshi et al., 2010). While components of the heat shock protein machinery interact with the N-terminal domain of Argonaute protein, the RNase III domains of Dicer-2 interact with the PIWI domain of Argonaute protein (Tahbaz et al., 2004; Tahbaz et al., 2001). Using the energy from ATP hydrolysis, heat shock proteins may pry open Argonaute protein so that it can receive the siRNA duplex from the RLC.

In the absence of heat shock proteins, Argonaute proteins are destabilized and mislocalized from P-bodies and stress granules (Tahbaz et al., 2001; Pare et al., 2009; Johnston et al., 2010). The loss of hsp90 also perturbed piRNA biogenesis and mutagenic transposons became derepressed (Specchia et al., 2010). The molecular explanations for these defects, however, are not known.

Future experiments will need to address if these defects stem from the failure to load small RNAs. At least one study suggests that an Argonaute protein free of small RNA guide is targeted for degradation by autophagy (Gibbings et al., 2012).

Passenger Strand Removal Signifies RISC Maturation

An Argonaute protein laden with siRNA duplex is also known as a pre-RISC and it matures into active Argonaute protein when it removes the passenger strand (Rand et al., 2005; Leuschner et al., 2006). Fly and human AGO2 proteins cleave and remove the passenger strand efficiently (Figure 1.1). Fly Ago1 has weak catalytic activity and human AGO1, AGO3 and AGO4 have lost their catalytic activities either remove the passenger strands of perfectly paired siRNA duplexes inefficiently or they fail to do so. Introducing mismatches in the siRNA duplex weakens the stability of the duplex and facilitates the eviction of the passenger strand (Kawamata et al., 2009; Yoda et al., 2009). Interestingly, the conversion of pre-RISC to mature RISC does not consume ATP.

Recent studies identified Component 3 promoter of RISC (C3PO) that helped remove cleaved passenger strand (Liu et al., 2009). C3PO is an octameric complex of 6 Translin and 2 Trax modules (Figure 1.10). It is an Mg^{2+}-dependent endoribonuclease with conserved catalytic residues that resides on the lumenal side of Trax (Liu et al., 2009; Tian et al., 2011a; Ye et al., 2011). In addition, the Trax-Translin complex also processes pre-tRNA (Li et al., 2012).

Figure 1.10

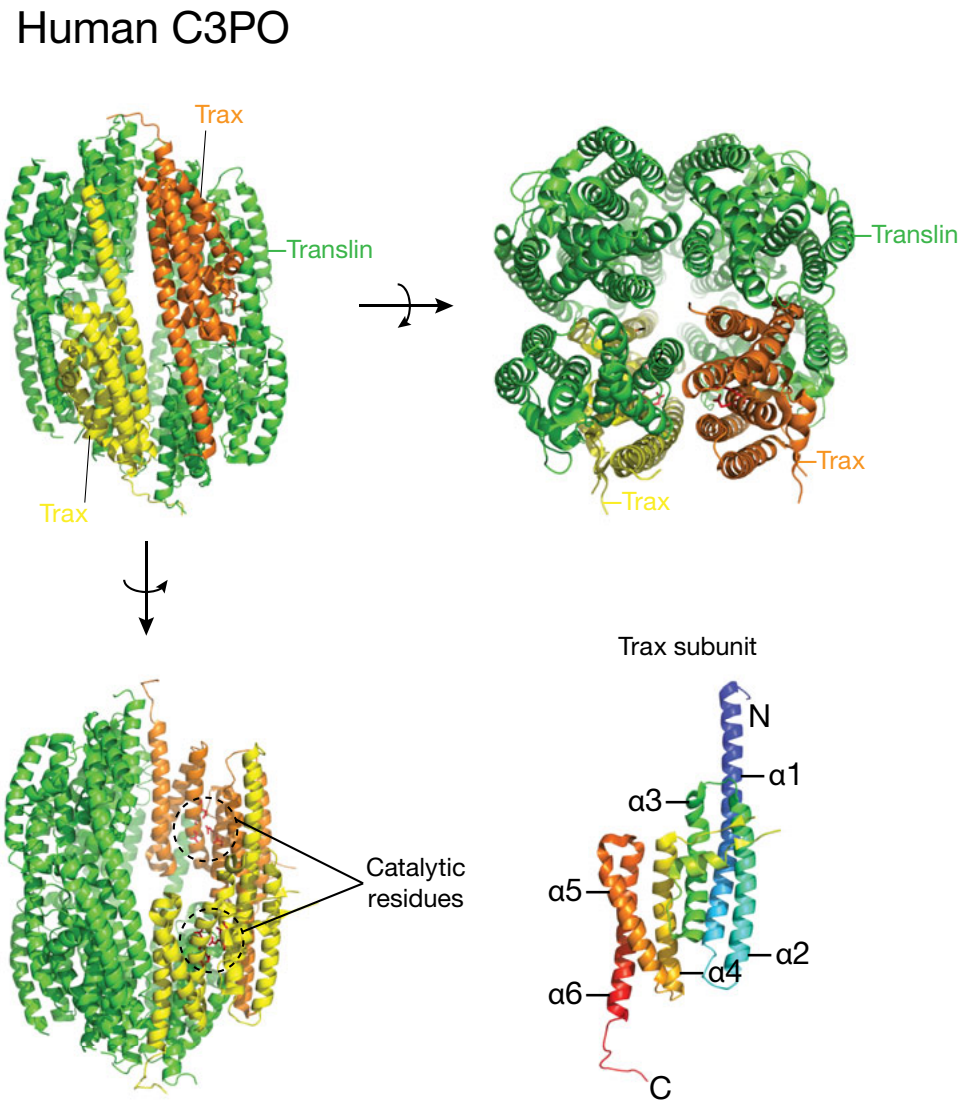

Human C3PO
Trax
Translin
Trax
Translin
Trax
Trax
Catalytic
residues
Trax subunit
N
α1
α3
α5
α2
α4
α6
C

Figure Legend 1.10. Human C3PO is an octameric complex

The octameric structure of C3PO consists of 2 Trax (yellow and orange) and 6 Translin subunits (green). The catalytic residues located on the Trax subunits are shown as red sticks and are found within the internal confines of C3PO. One Trax subunit that is made up of 6 α-helices is shown.

Most recent structural-functional insights into its role in RISC activation came from a close homolog of the Trax-Translin complex in *Archaeoglobus fulgidus* crystallized with bound RNA duplex (Parizotto et al., 2013). The study, however fails to address how the ~200 kDa mega molecular structure of C3PO that binds and cleaves both strand of the duplex is able to cooperate with Argonaute (~100 kDa) and to selectively remove only the passenger strand.

How does small RNA silencing work?

Mechanism of RNAi: The Two State Hypothesis

Based on both structural and biochemical data, it was proposed that Argonaute protein implements RNAi via a two-state model (Tomari and Zamore, 2005; Filipowicz, 2005). In the Argonaute protein, the 5′ phosphate of the single stranded guide is lodged between the interface of the MID and PIWI domain whereas its 3′ end is held by the PAZ domain (Figures 1.1 and 1.4A; Lingel et al., 2003; Yan et al., 2003; Song et al., 2003; Tomari et al., 2004b; Ma et al., 2004; Parker et al., 2005; Ma et al., 2005). The doubly anchored guide strand is constrained to initiate target binding using limited base nucleotides commonly known as the seed sequence (Figure 1.1).

Seed sequence spans nucleotide 2 (g2) to nucleotide 8 (g8) measured from the 5′ end of the guide strand (Bartel, 2009; Haley and Zamore, 2004; Ameres et al., 2007; Wee et al., 2012). The binding of target only in the seed sequence constitutes the first state. In addition to seed pairing, base pairing in the central

and 3′ supplementary region of the guide are crucial for target cleavage (Elbashir et al., 2001c; Haley and Zamore, 2004; Schwarz et al., 2006; Wee et al., 2012). Accordingly, base pairing has to propagate beyond the seed and into the 3′ end of the guide during RNAi. In the process, the 3′ end of the guide strand is expelled from the PAZ domain (Wang et al., 2008a; Wang et al., 2009). This is because the guide strand when fully paired is most likely topologically at odds with it being double anchored. Structural rearrangement of the Argonaute protein accompanied the repositioning of the guide strand in order to accommodate and to position the target for cleavage (Wang et al., 2009). After these remodeling steps, Argonaute protein achieves the second state (Figure 1.1).

Argonaute protein is a Mg^{2+}-dependent endonuclease that cleaves target mRNA across from g10 and g11 of the guide strand (Figure 1.1; Elbashir et al., 2001c; Elbashir et al., 2001b; Zamore et al., 2000; Martinez and Tuschl, 2004). Target cleavage generates target mRNA with 5′ phosphate and 3′ hydroxyl termini (Figure 1.7; Schwarz et al., 2004; Martinez and Tuschl, 2004). The 5′ fragment of a cleaved mRNA is removed from its 3′ hydroxyl end by the exosome whereas the 3′ fragment with its 5′ phosphate end is cleared by the XRN-1 exonuclease (Souret et al., 2004; Orban and Izaurralde, 2005; German et al., 2008; Chang et al., 2011). The efficient removal of cleaved target mRNA contributes to the proficiency of dAgo2 that undergoes multiple round of target

cleavage (Hutvagner and Zamore, 2002).

Partial Base Pairing Is a Hallmark of miRNA Targeting in Mammals

Extensive base pairing between mammalian miRNAs and their targets are rare. The seed sequence participates prominently in binding to five main types of seed-matched target sites (Figure 1.11). The miRNA seed sequence from g2–g8 that pairs with a complementary site on the target from t2–t8 and which also contains an adenosine at the t1 position (t1A) is a 8mer site (g2–g8:t2–t8; t1A). A guide:target pairing from g2–g8:t2–t8 and without a t1A denotes the 7mer-m8 site. A guide:target pairing, which involves g2–g7:t2–t7 and includes a t1A in the target represents the 7mer-A1 site. For guide:target pairing from g2–g7:t2:t7, it constitutes the 6mer site. Finally, in silico prediction identified the offset 6mer sites, g3–g8:t3–t8 that are selectively conserved above background (Friedman et al., 2009; Bartel, 2009). A less common miRNA target site with base pairing from g13–g16:t13–t16 in addition to seed pairing is termed seed plus 3′ supplementary pairing. In instances where the additional 3′ pairing enhances pairing of a mismatched seed, it is more appropriately termed seed plus 3′ compensatory pairing (Figure 1.11). The seed plus 3′ supplementary and the seed plus 3′ compensatory sites are estimated to be present in only 4.5% and 1.4% of selectively conserved sites respectively (Friedman et al., 2009).

The extent of base pairing between the guide and target mRNA in part dictates the modes of regulation by Argonaute proteins. A partially paired siRNA

Figure 1.11

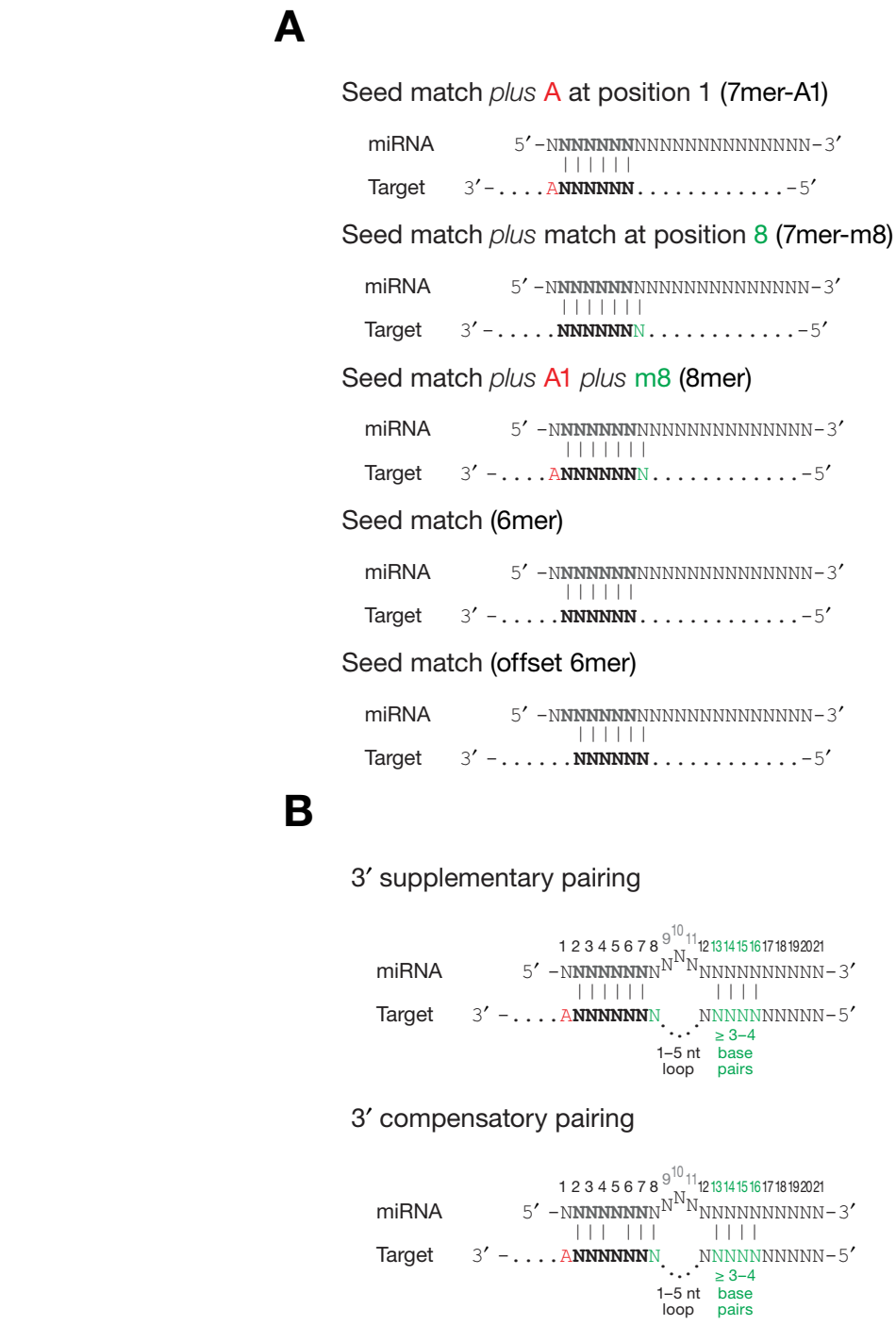


Figure Legend 1.11. Different types of miRNA targeting sites

behaves like a miRNA and causes translational repression instead of target cleavage (Figure 1.12B; Doench et al., 2003; Zeng et al., 2003). Plant miRNAs often pair extensively to target mRNAs and result in target cleavage by Argonaute proteins (Rhoades et al., 2002; Llave et al., 2002; Axtell et al., 2011). Some extensively paired plant miRNAs, however, can also repress mRNA translation (Chen, 2004; Brodersen et al., 2008; Lanet et al., 2009). In contrast, most animal miRNAs limit target base pairing to the seed sequence and therefore lack sufficient base pairing to silence genes by cleavage (Figure 1.12B; Yekta et al., 2008; Davis et al., 2005). Instead, mammalian Argonaute proteins either repress translation of mRNA or destroy mRNA by recruiting the cellular decay machineries (Figures 1.12C and 1.12D; Valencia-Sanchez et al., 2006; Parker and Sheth, 2007).

Predicting Targets of miRNA

Functional targets in plants are predicted with high confidence because they are often extensively paired to plant miRNAs. In animals, limited miRNA:target base pairings makes functional target prediction more daunting, less straightforward and prone to false positive hits (Figure 1.12B; Rhoades et al., 2002; Bartel, 2009). The target prediction algorithms include EMBL, EIMMo, TargetScan, miRanda, PicTar, PITA, DIANA-microT and mirWIP (Lewis et al., 2003; Lewis et al., 2005; Enright et al., 2003; John et al., 2004; Kiriakidou et al., 2004; Krek et al., 2005; Stark et al., 2005; Miranda et al., 2006; Lall et al., 2006; Gaidatzis et al., 2007;

Betel et al., 2008; Hammell et al., 2008; Friedman et al., 2009). Most of the target prediction tools emphasized on base complementarity to the seed sequence (Lewis et al., 2005; Krek et al., 2005; John et al., 2004). For instance, TargetScan require perfect seed pairing whereas PicTar, DIANA-microT and miRanda permit bulges and mismatches in seed pairing but with penalties imposed on the overall target prediction scores (Kiriakidou et al., 2004; Krek et al., 2005).

To increase prediction specificity, most algorithms consider target sites conservations across species. The tradeoff, however, is the loss in sensitivity, i.e. newly evolved miRNA targeting sites will be omitted. The free energy of binding (ΔG_B) between the miRNA and the target mRNA also contributes to the prediction score. MirWip and PITA, in addition, consider the free energy required to unfold target structure (ΔG_U) and therefore calculate the $\Delta\Delta G = \Delta G_B - \Delta G_U$ for miRNA:target pairing (Hammell et al., 2008; Kertesz et al., 2007). TargetScan also takes into account features that include 1) the positions of target sites in 3′ UTR 2) the AU content of miRNA binding sites 3) the presence of 3′ supplementary pairing and 4) the proximity of sites for cooperative regulation of target by Argonaute proteins (Doench and Sharp, 2004; Grimson et al., 2007; Broderick et al., 2011). The *rna22* prediction tool is the most radical among all prediction programs. It searches for pattern in a training set of uniquely represented miRNAs and then uses the statistically significant patterns to

identify potential targets. These potential targets are paired to candidate miRNAs and scores assigned to miRNA:target pairs that satisfy user-defined threshold of ΔG_B. *Rna22* does not consider site conservation, does not restrict target search to 3′ UTR and does not depend heavily on perfect seed complementarity (Miranda et al., 2006). It therefore claims to be more proficient at identifying newly evolved miRNA target sites (Miranda et al., 2006).

Accounting for phylogenetic relationships among the species compared when identifying miRNA sites based on conservations increased the prediction sensitivity (Friedman et al., 2009). In mammals, for the five miRNA site types (Figure 1.11), the number of selectively conserved predicted target sites stands at ~50,000 distributed across ~10,000 genes at a conservation threshold that yielded the highest signal above background. Of these conserved sites, each miRNA family was predicted to target an average of 534 ± 25. Only 7% of the targeted genes have multiple sites for the same miRNA family. This implies that each miRNA family regulates an average of 497 ± 49 conserved targets, which approaches the average number of conserved sites. On average, each target has ~4.2 miRNA binding sites. Taken together, ~60% of the human genes are targeted by miRNAs and ~72% of these targeted genes have multiple binding sites for different miRNA families (Friedman et al., 2009).

Silencing of mRNA by miRNA Proceeds by Multiple Mechanisms

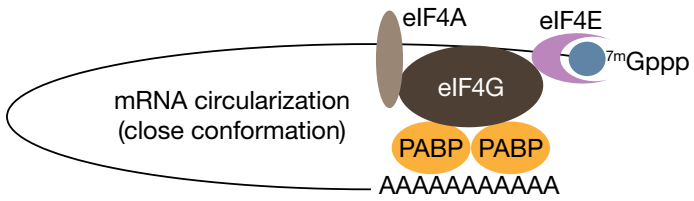
Initial studies in *C. elegans* showed that *lin-4* repressed its target mRNAs *lin-14*

and *lin-28* by inhibiting translational elongation (Olsen and Ambros, 1999; Seggerson et al., 2002). These claims were supported by the observations that targets remained associated with polysomes when repressed by miRNAs (Maroney et al., 2006; Nottrott et al., 2006; Petersen et al., 2006). These polysomes disassembled upon treatment with translational inhibitors (puromycin, hippuristanol and pactamycin) suggesting that they were translationally active before the repression (Maroney et al., 2006; Nottrott et al., 2006; Petersen et al., 2006). In further support of a post initiation mechanism, miRNAs also repressed the cap-independent translation of mRNA that uses either the Hepatitis C virus (HCV) or the Cricket Paralysis virus (CrPV) internal ribosomal entry site (IRES) elements (Petersen et al., 2006).

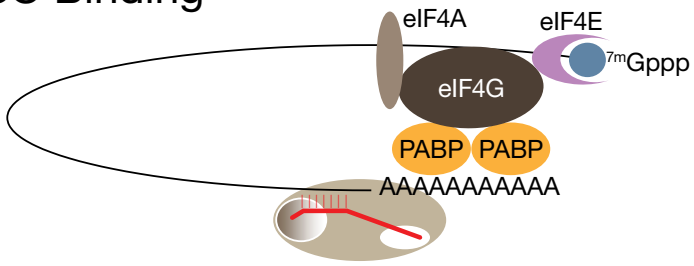
Proponents of a translational block at the initiation step by miRNA showed that mRNA that incorporated IRES from encephalomyocarditis virus (EMCV) resisted repression by miRNA (Pillai et al., 2005; Humphreys et al., 2005). Inhibition by miRNAs also resulted in a shift of the repressed target mRNAs that resided in the polysomes fractions into the lighter fractions (Pillai et al., 2005). This observation therefore contradicts earlier findings that mRNAs remained associated to polysomes when repressed by miRNAs. To explain the contradictions, these presumed polysomes observed during miRNA repression were believed to be pseudo-polysomes. Pseudo-polysomes are not canonical actively translating polysomes in that they lacked the 60S subunit

Figure 1.12

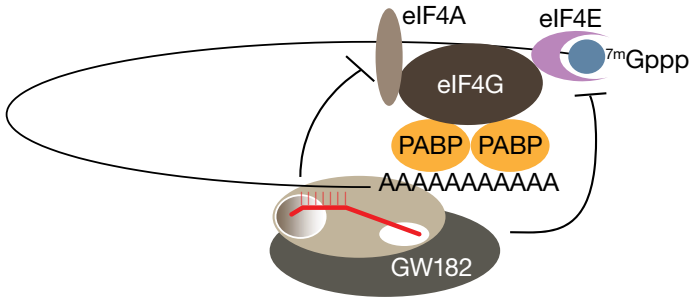
A. Active Translation



B. RISC Binding



C. Translational repression



D. mRNA decay

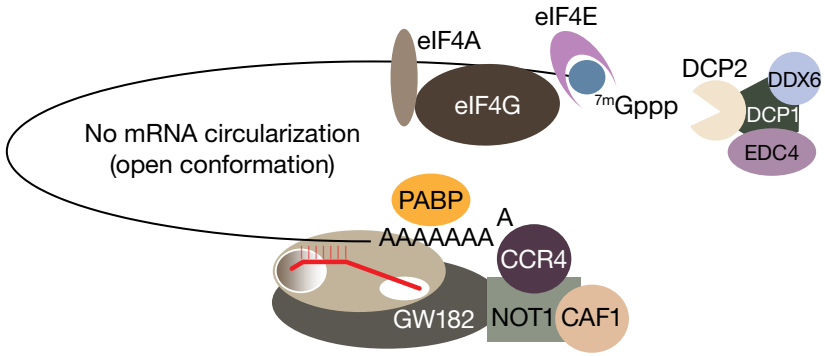


Figure Legend 1.12. Argonaute proteins repress and degrade target mRNA

Argonaute proteins bind and repress target mRNA by inhibiting the functions of eIF4E and eIF4A. Argonaute proteins can also recruit decapping enzymes and the exosome complex to destabilize mRNA targets.

(Thermann and Hentze, 2007; Wang et al., 2008). A more recent study showed that GW182 and poly(A) binding protein (PABP) are components of these pseudo-polysomes (Fukaya and Tomari, 2012).

In particular, some groups propose that translational inhibition at the initiation step by miRNA was ^{m7}GpppG cap-dependent (Figure 1.12C; Humphreys et al., 2005; Wang et al., 2006a; Wakiyama et al., 2007; Thermann and Hentze, 2007). In fly, dAgo2 competes with eIF4G for binding to the cap binding protein, eIF4E (Figure 1.12C; Iwasaki et al., 2009). The molecular mechanism by which this results in translational repression is not currently known. Interestingly, dAgo1 is found to repress translation in steps after cap recognition (Iwasaki et al., 2009). Subsequent work deduced and showed that dAgo1 and mammalian Argonaute proteins repress translation by inhibiting the function of eIF4A (Fukaya and Tomari, 2012; Ricci et al., 2012, Nucleic Acids Res; Meijer et al., 2013). eIF4A is an RNA helicase that helps unfold secondary structure at the 5′ UTR of mRNA to facilitate binding and scanning by small ribosomal subunit (Jackson et al., 2010). Interestingly, mRNAs with unstructured 5′ UTR that obviate the need for eIF4A2 are refractory to repression by miRNAs (Meijer et al., 2013).

In addition to translational repression, miRNAs trigger decapping and deadenylation of mRNA (Figure 1.12D; Rehwinkel et al., 2005; Behm-Ansmant et al., 2006a; Giraldez et al., 2006; Wu et al., 2006; Eulalio et al., 2009b).

Consequently, target repression by miRNAs was shown to downregulate mRNA levels (Krutzfeldt et al., 2005; Wu and Belasco, 2005; Lim et al., 2005; Bagga et al., 2005; Giraldez et al., 2006; Wu et al., 2006; Mishima et al., 2006; Hendrickson et al., 2009). Accordingly, numerous studies revealed that the effect of reduced protein expression from repression by miRNAs could be mostly accounted for by a corresponding decreased in mRNA levels (Forstemann et al., 2007; Baek et al., 2008; Selbach et al., 2008; Guo et al., 2010). These results implied that miRNAs act mainly to destroy mRNAs rather than to inhibit their translation.

These studies were based upon steady state measurements. Several time course experiments indicated that translational inhibition precedes mRNA degradation (Fabian et al., 2009; Bazzini et al., 2012; Djuranovic et al., 2012; Béthune et al., 2012). Furthermore, there is at least one example where miRNA-repressed mRNAs are reversibly desilenced in human cells subjected to stress. This suggested that the repressed mRNAs are not always destined for degradation (Bhattacharyya et al., 2006). In conclusion, target repression by miRNA can operate through multiple mechanisms. The cell types, the states of the cells and the Argonaute protein identities are some confounding factors that can explain the various mechanisms observed during miRNA-mediated mRNA silencing.

Figure 1.13

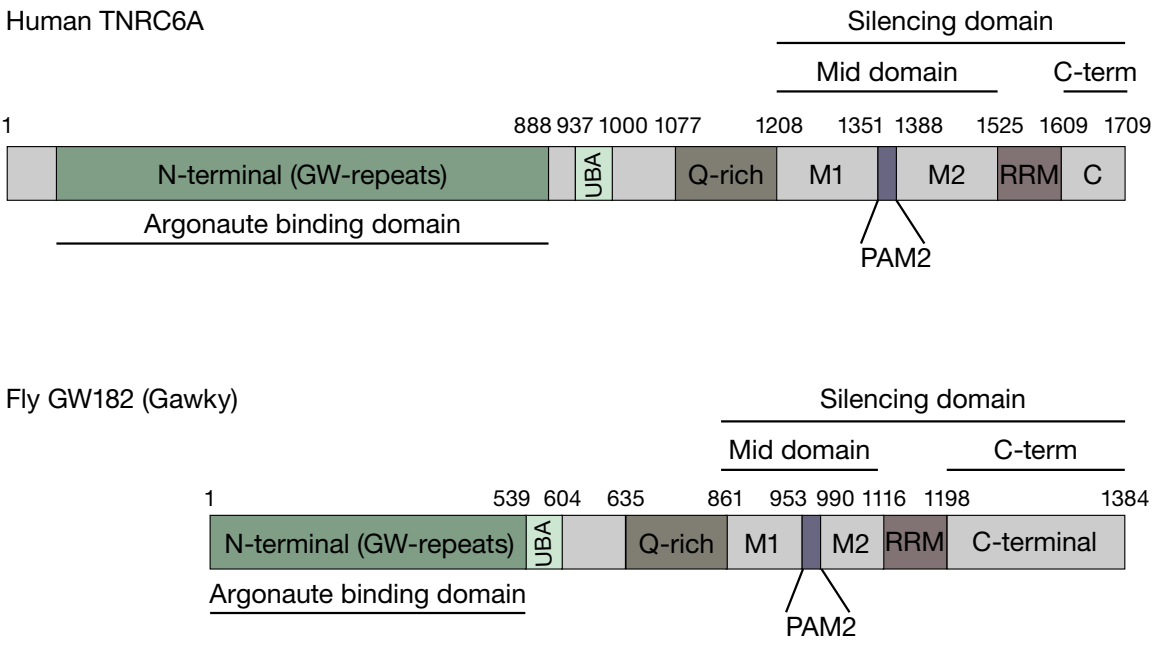


Figure Legend 1.13. Functional domains of GW182

GW182 Participates in miRNA-mediated mRNA Repression and Decay

GW182 is the autoantigen in patients afflicted with motor and sensory neuropathy (Eystathioy et al., 2002). There are three paralogs in vertebrates termed trinucleotide repeat-containing protein (TNRC) 6A, 6B and 6C, one ortholog in *D. melanogaster* also called Gawky and two orthologs in *C. elegans* known as AIN-1 and AIN-2 (Meister et al., 2005; Schneider et al., 2006; Rehwinkel et al., 2005). Argonaute protein recruits GW182 after it finds and binds its target mRNA (Behm-Ansmant et al., 2006b). The N-terminal region of GW182 interacts with Argonaute protein while its C-terminal silencing domain functions to repress and to destroy mRNA (Figure 1.13; Till et al., 2007; Takimoto et al., 2009; Eulalio et al., 2009a; Chekulaeva et al., 2010). Human AGO2 crystallized in the presence of tryptophan identified two binding pockets that purportedly will bind to the N-terminal domain of GW182 that is rich in glycine and tryptophan (Figure 1.14; Schirle and MacRae, 2012). In contrast, dAgo2 that functions in RNAi does not interact with GW182 and is predicted not to contain the binding pockets (Iwasaki et al., 2009).

GW182 is a resident protein of the P-bodies and therefore P-bodies are also refered to as GW-bodies (Filipowicz et al., 2008). GW182 ensures proper localization of other resident P-bodies components (i.e. CAF1-CCR4-NOT1 complex, DCP1, DCP2, LSm1–7, RCK/p54, XRN1) and maintains P-body structures (Liu et al., 2005b; Liu et al., 2005a; Sen and Blau, 2005; Behm-

Ansmant et al., 2006a). Argonaute proteins localized to P-bodies are presumably the sites where mRNA repression by miRNAs occur (Liu et al., 2005a; Sen and Blau, 2005; Liu et al., 2005b). However, disruption of P-bodies by LSm1 and LSm3 depletion did not affect target repression by Argonaute protein (Eulalio et al., 2007). This indicates that the localization of Argonaute proteins to P-bodies was a consequence rather than a cause of miRNA repression (Eulalio et al., 2007; Chu and Rana, 2006; Eulalio et al., 2009a). Of note, the persistence of submicroscopic foci that are functionally equivalent to P-bodies has yet to be ruled out.

The interactions between PABP that sits on the poly(A) tail and the scaffold protein eIF4G that lies close to the cap circularizes mRNA and stimulates its translation (Figure 1.12A; Jackson et al., 2010). The removal of poly(A) tail during miRNA-mediated repression precludes mRNA circularization, decreases translation efficiency and therefore scores as translational repression (Figure 1.12). Several evidences, however, also suggested that miRNA is capable of repressing mRNA translation even without mediating deadenylation of the mRNA. First, removing GW182 impaired repression of both poly(A) containing and poly(A) free mRNAs (Eulalio et al., 2008; Eulalio et al., 2009b). Second, poly(A) free mRNA equipped instead with 3′ histone H4 stem-loop structures (HSL) can still be repressed by miRNA (Eulalio et al., 2009b; Braun et al., 2011; Chekulaeva et al., 2011). Likewise, mRNA with internal poly (A) that resists

Figure 1.14

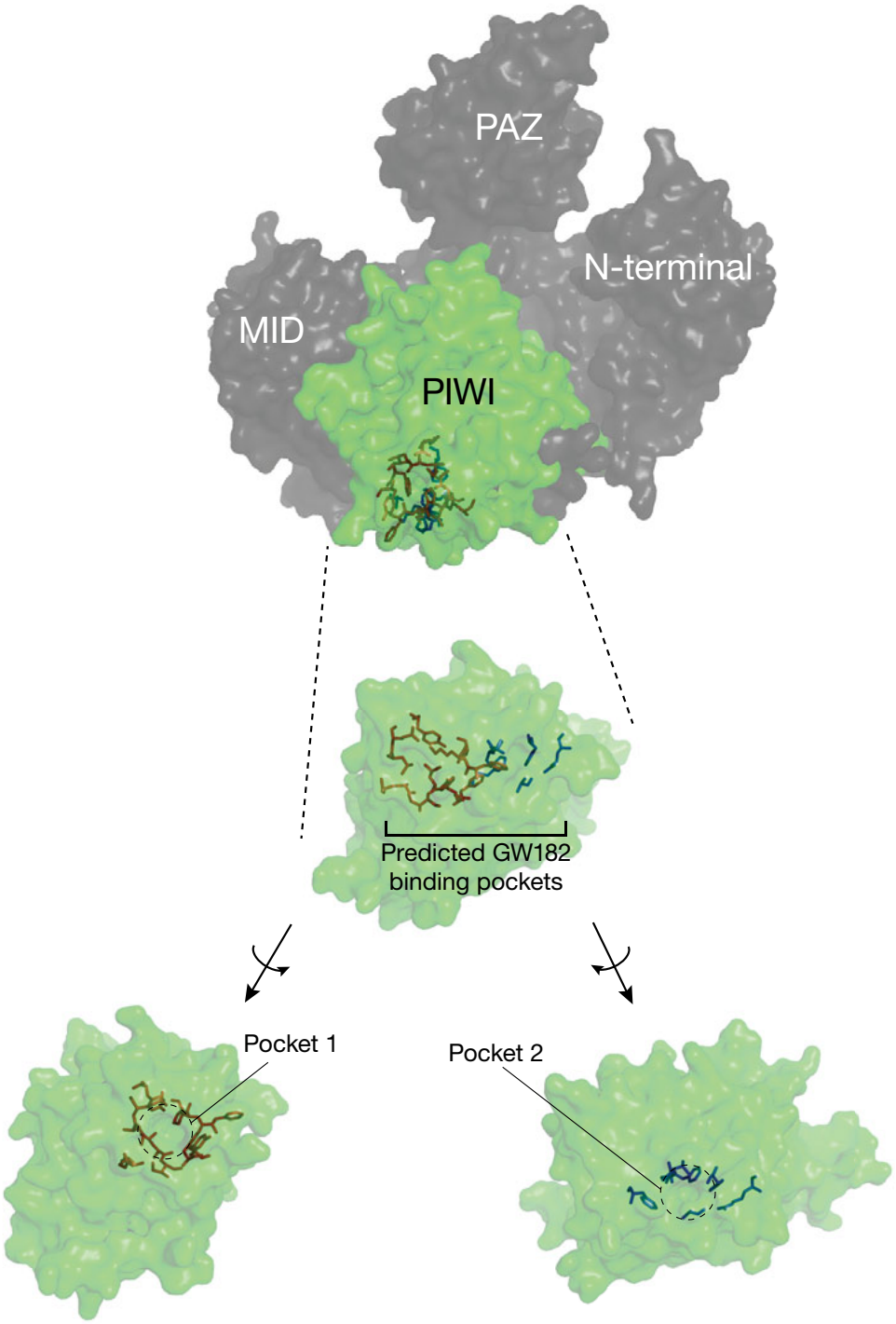

PAZ
N-terminal
MID
PIWI
Predicted GW182
binding pockets
Pocket 1
Pocket 2

Figure Legend 1.14. Human AGO2 contains potential GW182 binding pockets

Crystal structures of human AGO2 solved in the presence of tryptophan residues identify two potential binding pockets for GW182 in the PIWI domain. Residues surrounding the two binding pockets are shown in read and blue sticks respectively (Schirle and MacRae, 2012).

deadenylation was still efficiently silenced (Fukaya and Tomari, 2011; Fukaya and Tomari, 2012; Ricci et al., 2012). Finally, tethering a catalytically inactive deadenylation complex to poly(A) free mRNA is sufficient to repress mRNA (Cooke et al., 2010).

Indeed, it was recently demonstrated that GW182 recruits CAF1-CCR4-NOT1 complex to cause dissociation of PABP prior to the deadenylation of mRNA (Zekri et al., 2013). For mRNA that loses PABP from its poly(A) tail, it also fails to circularize and translates less efficiently. Eventually, GW182 recruits the decapping and the deadenylation enzymes to promote decay of the mRNA (Behm-Ansmant et al., 2006a; Chekulaeva et al., 2009; Braun et al., 2011; Fabian et al., 2009; Chekulaeva et al., 2011; Chen et al., 2009; Piao et al., 2010; Huntzinger et al., 2013). GW182 can interact directly with PABP and it was proposed that PABP is necessary to deadenylate repressed mRNA (Fabian et al., 2009; Huntzinger et al., 2013). This claim, however, has been challenged and remains debatable (Fabian et al., 2009; Fukaya and Tomari, 2011).

Figure 1.15

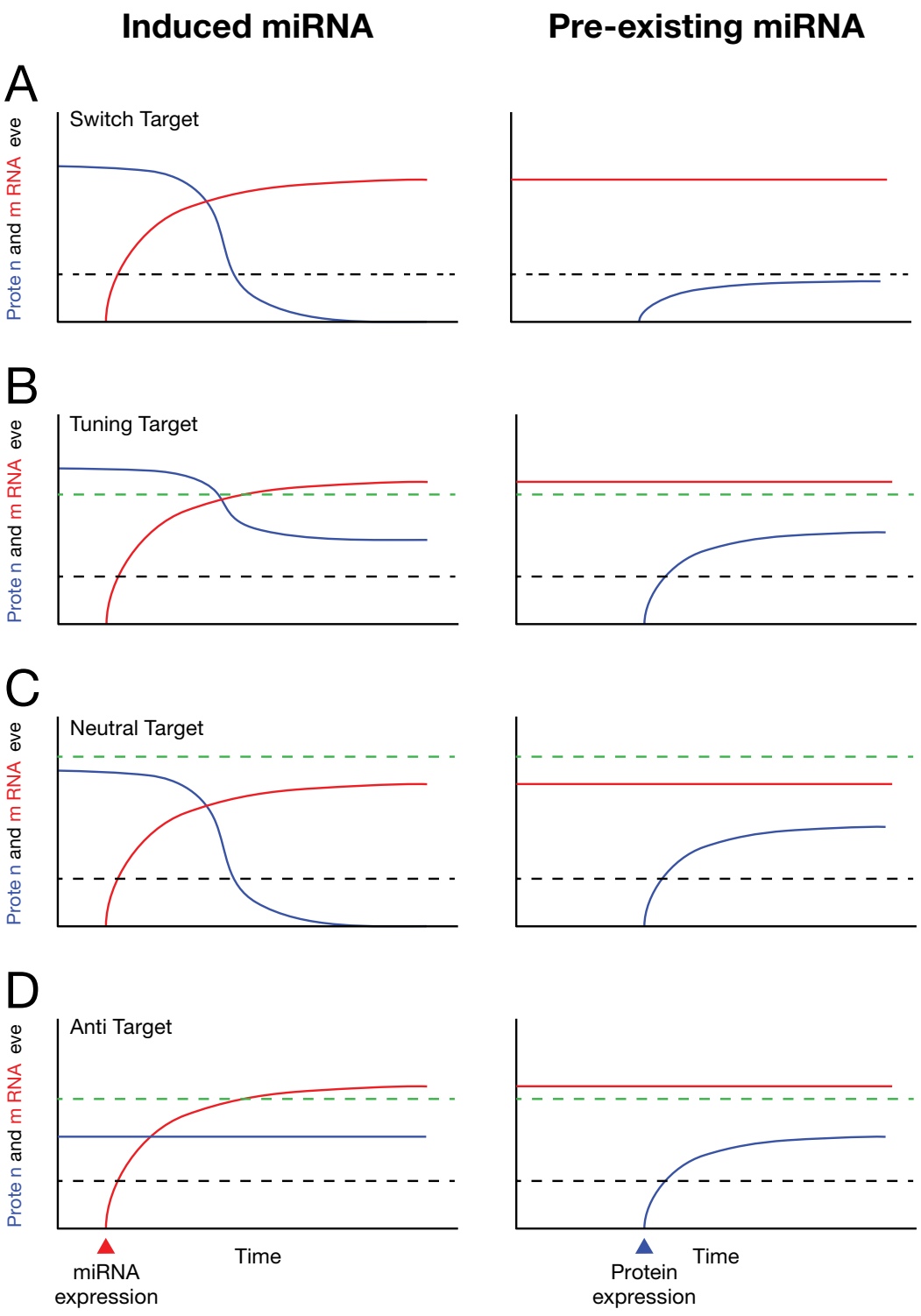

Induced miRNA
Pre-existing miRNA
A
Switch Target
B
Tuning Target
C
Neutral Target
D
Anti Target
Prote n and m RNA eve
miRNA expression
Protein expression
Time
Time

Figure Legend 1.15. Types of miRNA targets

Graphs depict protein expression under the control of pre-existing (right) or induced miRNAs (left). For B, C and D the lower black and the upper green dotted lines represent thresholds for functional and for excess undesirable protein levels respectively. For switch target in A, both thresholds coincide. The protein level of switch target has to be reduced below the same functional and undesirable level. In tuning target, the protein level is adjusted to optimal and functional level. Neutral target is regulated by miRNAs but the cell tolerates the high protein level even in the absence of miRNAs. Anti-target either looses the miRNA binding sites or is expressed mutually exclusively from targeting miRNAs. Therefore, its protein level does not change with miRNA expression. Graphs illustrate relative expression levels among miRNAs, proteins and the thresholds. Their absolute values will vary from cell to cell and is also cell-type specific. Graphs are adapted from Bartel and Chen, Nat. Rev. Genet., 2004.

A Need for Quantitative Modeling of Gene Silencing by Small RNA

In the preceding sections, I have highlighted the biogenesis and functions of siRNA, miRNA and piRNA. I have also described the structures of Argonaute proteins and discussed how we are beginning to address and learn about the underlying mechanisms of gene silencing. Our understanding of small RNA biology, however, is incomplete without quantitative measures. Thereupon, we can begin by phrasing our questions in more quantitative terms. For instance, how fast do Argonaute proteins bind their targets? What are the cellular Argonaute proteins and target mRNAs concentrations? How many copies of a particular mRNA are repressed given a defined amount of Argonaute protein? How many different distinct transcripts does an Argonaute protein with a particular guide strand silence? The answers to these questions will teach us about the mechanisms of small RNA silencing. They will also reveal the extent of mRNA repression required to generate meaningful regulation and to elicit a change in cellular physiology.

Several models posited to explain regulation of mRNAs by miRNAs justify the importance and relevance of quantitative science (Figure 1.15; Bartel and Chen, 2004). In the first model, miRNAs reduce gene expression to inconsequential level and is functionally equivalent to the gene being completely silenced. The heterochronic genes are classical examples of such tightly regulated mRNAs. These genes are functionally turned on or off during different

development stages akin to light switches and are thus called switch targets (Figure 1.15A; Lee et al., 1993; Reinhart et al., 2000; Wightman et al., 1993; Moss et al., 1997). In the second model, mRNAs are retained and kept at optimal levels for cellular functions. These are termed tuning targets (Figure 1.15B; Karres et al., 2007). In the third model, mRNA expression profile fluctuates in response to miRNA expression level but these changes are not functionally meaningful and do not produce any effect on cellular physiology. These are deemed neutral targets (Figure 1.15C). Finally, targets that avoid miRNAs by having mutually exclusive expression patterns or have lost miRNA-binding sites totally are the anti-targets (Figure 1.15D; Farh et al., 2005; Stark et al., 2005). A tuning target, however can behave like a switch target. For instance, miRNA that tunes the level of mRNA may just be sufficient to decrease its concentration below the functional threshold: the cell interprets the mRNA as being in the off state. Therefore, quantitative values for miRNA and mRNA expressions, its effect on the final concentration of repressed mRNA and its functional threshold value are vital information to test these models.

Rationales and Objectives

Different classes of small RNA and Argonaute proteins coexist in the cell. Does a specific class of small RNA associate with a particular Argonaute protein to become dedicated to specialized job? This thesis highlights work attempting 1) to understand if and how small RNAs are sorted into distinct Argonaute proteins 2) to describe the mechanisms of Argonaute protein paralogs by dissecting quantitatively their thermodynamic and kinetic properties using dAgo1, dAgo2 and mAGO2 as model proteins and 3) to reveal how their different mechanistic behaviors contribute to their specific functions. Not least, 4) to formulate a quantitative and reasonable model to explain how small RNA silencing functions in the cellular context using biochemically defined thermodynamic and kinetic parameters.

Chapter II: *Drosophila* microRNAs are sorted into functionally distinct Argonaute protein complexes after their production by Dicer-1

Disclaimer

This chapter was a product of a collaborative effort among the authors: Klaus Förstemann (KF), Michael Horwich (MH), Liang Meng Wee (Wee), Yukihide Tomari (YT) and Phillip Zamore (PDZ). KF and MH performed the experiments for figures 2.1 to 2.5. Wee and YT performed the experiments for figure 2.6.

SUMMARY

Small interfering RNAs (siRNAs) and microRNAs (miRNAs) guide distinct classes of RNA-induced silencing complexes (RISCs) to repress mRNA expression in biological processes ranging from development to antiviral defense. In *Drosophila*, separate but conceptually similar endonucleolytic pathways produce siRNAs and miRNAs. Here, we show that despite their distinct biogenesis, double-stranded miRNAs and siRNAs participate in a common sorting step that partitions them into Ago1- or Ago2-containing effector complexes. These distinct complexes silence their target RNAs by different mechanisms. MiRNA-loaded Ago2-RISC mediates RNAi, but only Ago1 is able to repress an mRNA with central mismatches in its miRNA-binding sites. Conversely, Ago1 cannot mediate RNAi, because it is an inefficient nuclease whose catalytic rate is limited by the dissociation of its reaction products. Thus, the two members of the *Drosophila* Ago sub-clade of Argonaute proteins are functionally specialized, but specific small RNAs classes are not restricted to associate with Ago1 or Ago2.

RESULTS

miR-277 is produced by Dcr-1, but loaded into Ago2

Like all known *Drosophila* miRNAs, miR-277 is produced by cleavage of its precursor by Dcr-1 acting with Loquacious (Loqs), rather than Dcr-2, which generates siRNAs (Lee et al., 2004b; Forstemann et al., 2005; Jiang et al., 2005; Saito et al., 2005). Both siRNAs and miRNAs are proposed to be loaded into Argonaute-containing effector complexes from double-stranded intermediates: guide/passenger strand duplexes for siRNAs and miRNA/miRNA* duplexes for miRNAs (Hutvagner and Zamore, 2002; Miyoshi et al., 2005; Rand et al., 2005; Matranga et al., 2005). The miR-277/miR-277* duplex is predicted to have more double-stranded character than typical miRNA/miRNA* duplexes, which are interrupted by mismatches and internal loops (Khvorova et al., 2003; Han et al., 2006). Thus, miR-277 has a miRNA/miRNA* duplex that resembles an siRNA. We asked if the resemblance of the miR-277/miR-277* duplex to an siRNA led to its being loaded into Ago2, rather than Ago1, in *Drosophila* cells. That is, is the biogenesis of a miRNA tightly coupled to its loading into Ago1 (Figure 2.1A)? Or are miRNAs, and perhaps siRNAs, sorted into distinct Ago proteins by a step unlinked to the Dicer that produced them (Figure 2.1B)?

To this end, we established stable lines of S2 cells expressing GFP mRNA alone, with a 3′ UTR containing one or two sites fully complementary to miR-277 or with four 3′ UTR sites complementary to miR-277 but bearing

Figure 2.1

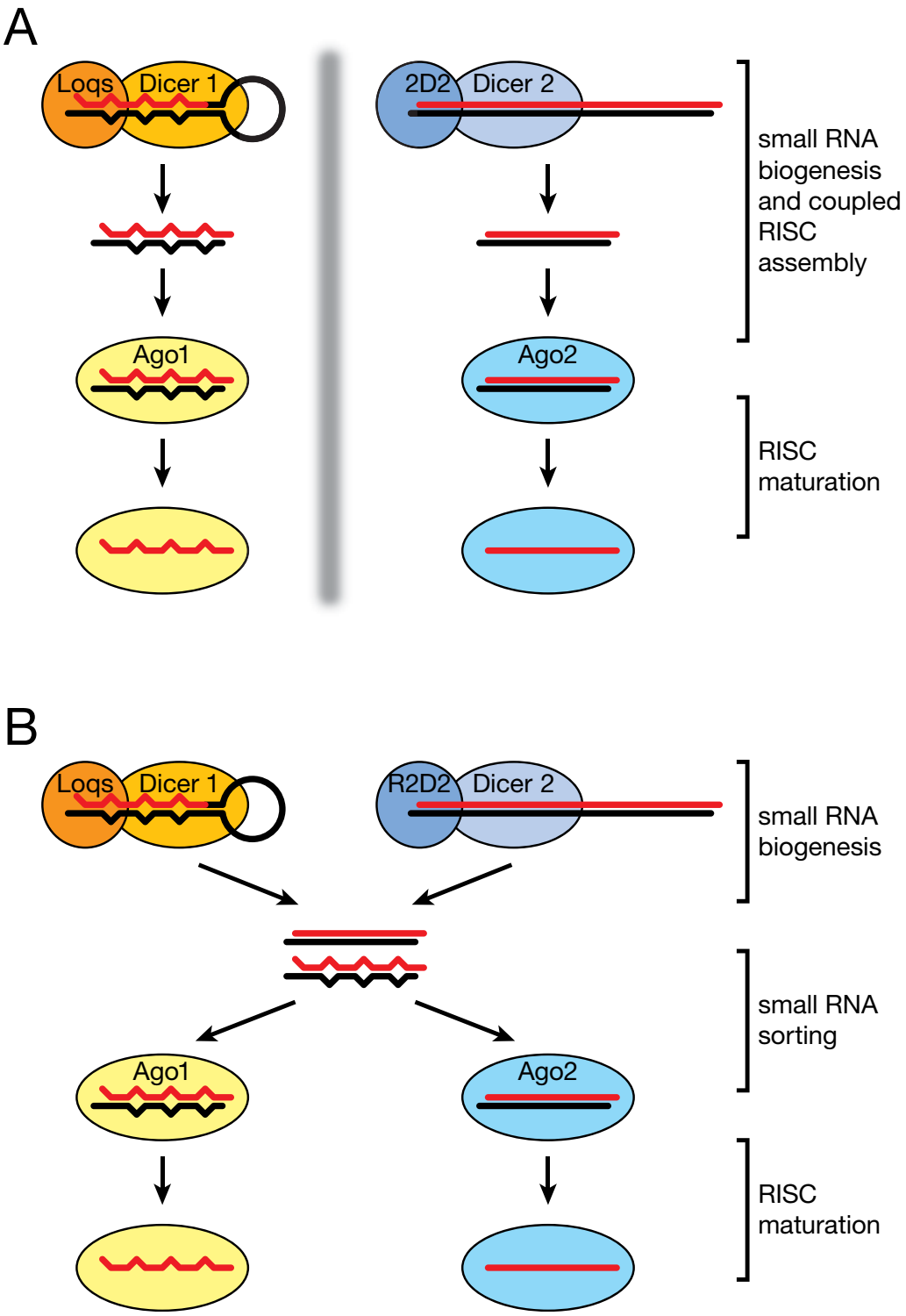

A
Loqs
Dicer 1
2D2
Dicer 2
small RNA biogenesis and coupled RISC assembly
Ago1
Ago2
RISC maturation
B
Loqs
Dicer 1
R2D2
Dicer 2
small RNA biogenesis
small RNA sorting
Ago1
Ago2
RISC maturation

Figure Legend 2.1. Two models for the miRNA and siRNA pathways in *Drosophila*.

(A) Small RNA biogenesis and RISC assembly are tightly coupled. miRNAs are exclusively loaded into Ago1 and siRNAs into Ago2. (B) Small RNA biogenesis and RISC assembly are independent. After their production, small RNA duplexes are proposed to be actively sorted into distinct Ago proteins solely according to their structures.

mismatches with miR-277 nucleotides 9, 10, and 11 (Figures 2.S1–2.S3). Surprisingly, repression of the reporter bearing two fully complementary miR-277-binding sites required Ago2 but not Ago1 (Figure 2.1). Treating the S2 cells with dsRNA to deplete Ago2 by RNAi increased GFP expression ~6-fold (Figures 2.2A and 2.2B). This agrees well with the extent of derepression observed with a miR-277-specific anti-sense oligonucleotide (ASO; Figure 2.S1), suggesting that without Ago2, the reporter is no longer repressed. Moreover, *ago1(RNAi)* increased repression of this reporter. Essentially identical data were obtained for a reporter containing a single miR-277-binding site (Figure 2.S3).

Expression of the miR-277-regulated reporter also increased when the cells were treated with dsRNA to deplete Drosha, the enzyme that excises pre-miRNAs from their primary transcripts or with dsRNA to deplete Dcr-1 or Loqs, which together convert pre-miRNA to miRNA/miRNA* duplexes (Figure 2.2A). RNAi directed against *ago1*, *ago2*, or *drosha* had no detectable effect on the expression of the GFP reporter lacking miR-277-binding sites.

We note that a control dsRNA was not inert with respect to Ago2-dependent silencing (i.e., RNAi), likely because it can compete with miR-277 for Ago2 loading. The idea that non-specific dsRNA can compete for Ago2 and other components of the Ago2-loading machinery is consistent with earlier reports that RNAi is a saturable process (Haley and Zamore, 2004). Thus, the most straightforward method to assess the significance of the effect of different

Figure 2.2

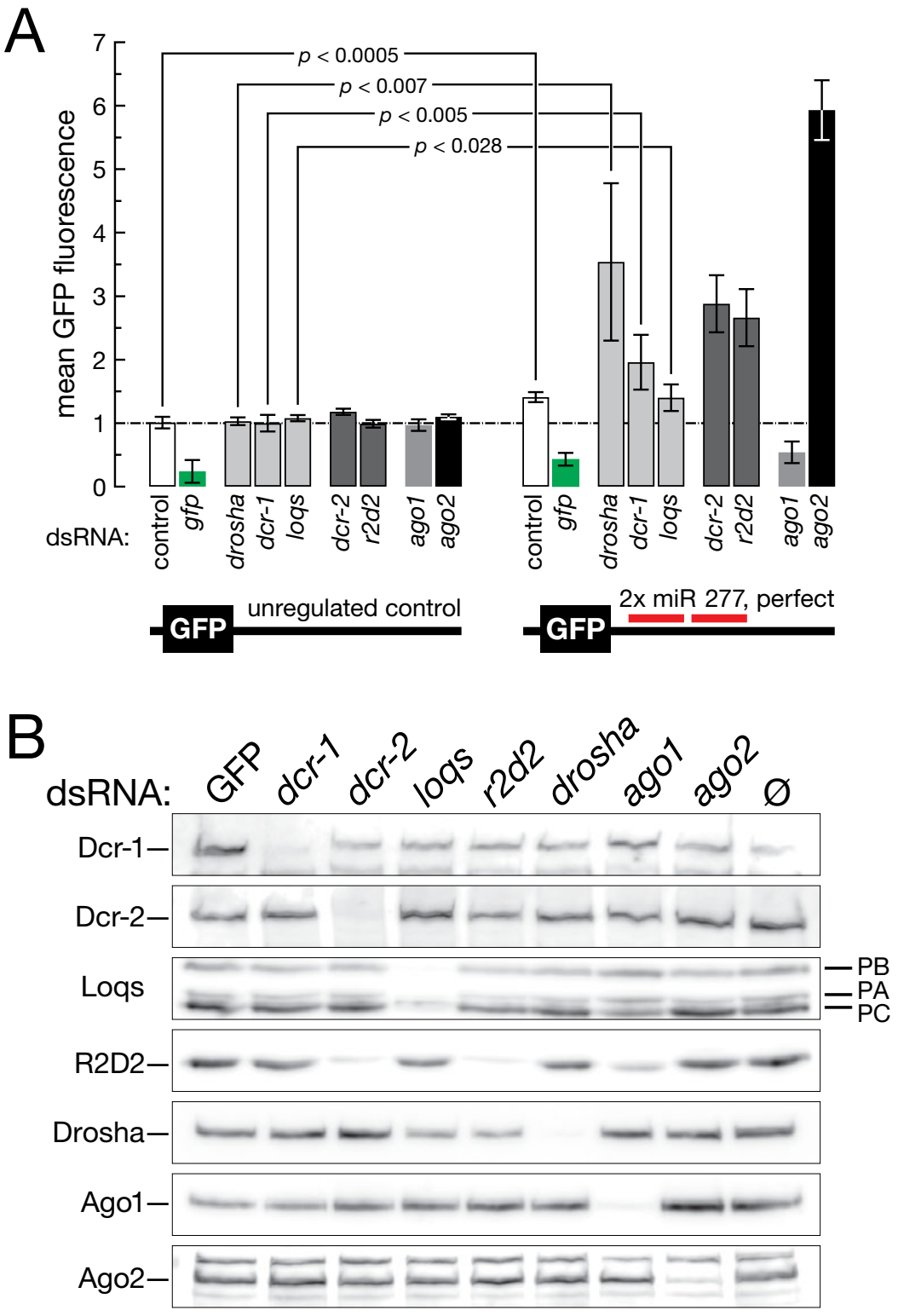

A
mean GFP fluorescence
p < 0.0005
p < 0.007
p < 0.005
p < 0.028
dsRNA:
control
gfp
drosha
dcr-1
loqs
dcr-2
r2d2
ago1
ago2
GFP
unregulated control
2x miR 277, perfect
B
dsRNA:
GFP
dcr-1
dcr-2
loqs
r2d2
drosha
ago1
ago2
Ø
Dcr-1
Dcr-2
Loqs
PB
PA
PC
R2D2
Drosha
Ago1
Ago2

Figure Legend 2.2. Components of both the miRNA and the RNAi pathways are required to silence a reporter with perfect matches to miR-277.

(A) Mean GFP fluorescence (average ± standard deviation for three or four trials). DsRNA-triggered RNAi was used to deplete the cells of the indicated protein. (B) Western blotting confirmed the extent and specificity of the RNAi-mediated depletion for each protein. *dcr-2(RNAi)* reduced the abundance of both Dcr-2 and R2D2, as previously reported {Liu et al., 2003, Science, 301, 1921-5}, but r2d2(RNAi) had no detectable effect on Dcr-2 abundance. The three isoforms of Loqs are indicated at the right of the Loqs panel. The bands above and below Ago2 correspond to cross-reacting proteins characteristically detected with this antibody.

dsRNAs on miR-277-directed repression of the perfect reporter in this experiment is not to compare the individual specific knock-down experiments to the control dsRNA, but rather to compare the change in GFP expression for the unregulated reporter to that observed for the perfect reporter for each RNAi knock-down. Analyzed this way, depletion of components of the miRNA biogenesis pathway clearly has a significant effect on miR-277-directed repression of the reporter: *dcr-1(RNAi)*, $p < 0.005$; *drosha(RNAi)*, $p < 0.007$; *loqs(RNAi)*, $p < 0.028$.

Together with previously published results (Forstemann et al., 2005), our data therefore suggest that miR-277 is produced by the standard miRNA pathway, but directs repression of a perfectly matched GFP reporter through Ago2.

Ago1 but not Ago2 mediates repression of mRNAs bearing bulged miR-277-binding sites

mRNAs containing miRNA-binding sites with perfect complementarity to specific miRNAs occur in animals, but are rare (Mansfield et al., 2004; Yekta et al., 2004; Davis et al., 2005). Instead, most miRNA are incompletely complementary to the mRNAs they repress. Typically, these miRNAs bind to multiple sites in the 3´ UTR of their targets. This mode of miRNA-directed repression can be recapitulated by engineering into the 3′ UTR of the reporter mRNA four, partially mismatched, miRNA-binding sites, each of which forms a

central bulge when paired to its cognate miRNA (Zeng et al., 2002; Doench et al., 2003).

We established stable lines of S2 cells expressing a GFP mRNA bearing four such sites (Figure 2.S1A). Repression of the reporter was modest, but required miR-277: transfection of a miR-277-specific, but not a control, ASO caused a small but statistically significant ($p < 0.003$) increase in GFP fluorescence (Figure 2.S1B). miR-277 is relatively abundant in S2 cells, which contain ~2,200 miR-277 molecules per cell (MH and PDZ, unpublished). Nonetheless, we wondered if the free pool of endogenous Ago1-loaded miR-277 was insufficient to repress expression of the bulged reporter.

We increased the expression of miR-277 by engineering stable S2 lines expressing both the GFP reporter and a 'mini' pri-miR-277 driven by the ubiquitin promoter. The resulting doubling of miR-277 expression caused a dramatic increase in the repression of the bulged GFP reporter, as evidenced by the more than 3-fold increase in GFP fluorescence observed when a miR-277-specific ASO was transfected into the cells (Figure 2.3A). Compared to the repression of this reporter by endogenous miR-277, the exogenous miR-277 increased repression of the bulged reporter by 230 percent (Figures 2.S1A and 2.3A). Repression was also enhanced, but to a smaller extent, for the reporter bearing two perfectly complementary miR-277-binding sites.

Figure 2.3

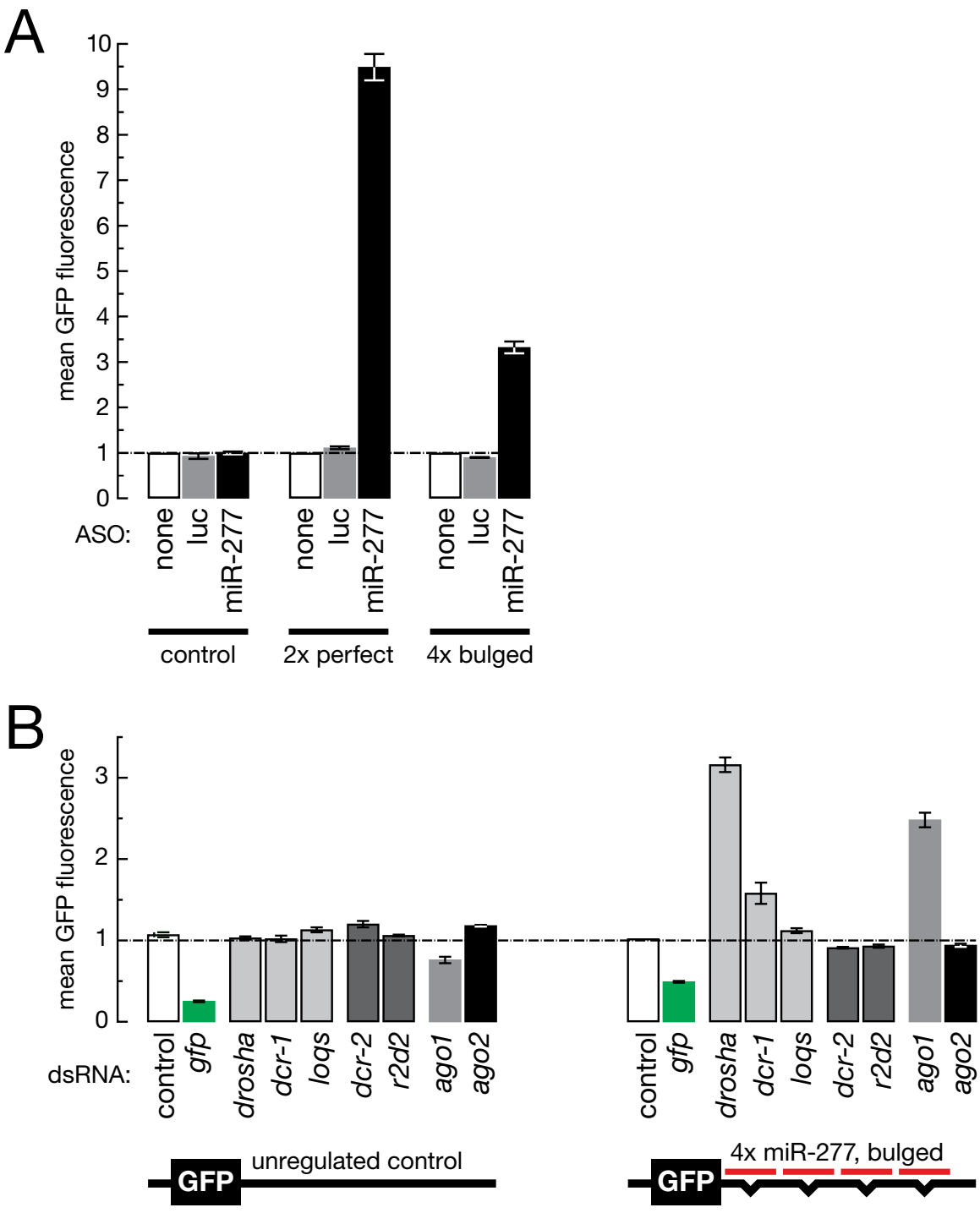

A
mean GFP fluorescence
10
9
8
7
6
5
4
3
2
1
0
ASO:
none
luc
miR-277
none
luc
miR-277
none
luc
miR-277
control
2x perfect
4x bulged
B
mean GFP fluorescence
3
2
1
0
dsRNA:
control
gfp
drosha
dcr-1
loqs
dcr-2
r2d2
ago1
ago2
control
gfp
drosha
dcr-1
loqs
dcr-2
r2d2
ago1
ago2
GFP
unregulated control
GFP
4x miR-277, bulged

Figure Legend 2.3. Only components of the miRNA pathway are required to silence a reporter bearing four imperfectly matched miR-277 target sites. (A) Over-expression of miR-277 from a mini-pri-miRNA transgene increased repression of the miR-277-regulated perfectly matched and bulged reporters. (B) Mean GFP fluorescence (average ± standard deviation for three or four trials). DsRNA-triggered RNAi was used to deplete the cells of the indicated protein.

For both the reporter bearing perfectly complementary miR-277-binding sites and the reporter with four bulged miR-277-binding sites, miR-277 reduced GFP expression by reducing the stability of the reporter mRNA, rather than by repressing GFP translation. We used qRT-PCR to measure the steady-state reporter mRNA abundance (Figure 2.S4A) and FACS to measure GFP protein abundance (Figure 2.S4B). For each stable cell line, we measured reporter mRNA and protein expression after transfection with a control ASO or a miR-277-specific ASO. Even when miR-277 was over-expressed, nearly all of the increased GFP protein expression observed when miR-277 was blocked could be accounted for by a corresponding increase in GFP mRNA expression. Figure 2.S4C reports the relative GFP protein expression normalized to the relative GFP mRNA expression. In all cases, when miR-277 was inhibited the ratio of relative protein expression to relative mRNA expression was close to one, indicating that most of the miR-277-directed reporter repression reflected mRNA destabilization rather than translational repression. However, for the bulged reporter, mRNA degradation might be tightly coupled to translational repression and therefore be a consequence, rather than a cause, of the decrease in protein production.

Silencing of the bulged reporter required Ago1 but not Ago2: *ago1(RNAi)* increased reporter expression, whereas *ago2(RNAi)* (Figure 2.2B) caused a small but statistically significant decrease in reporter expression ($p < 0.008$) (Figure

2.3B). While RNAi directed against *drosha*, *dcr-1*, or *loqs*—all genes required for miRNA biogenesis—increased expression of the bulged reporter, *dcr-2(RNAi)* and *r2d2(RNAi)*—both genes required to load small RNAs into Ago2, but not Ago1—caused a small but statistically significant ($p < 0.001$ and $p < 0.003$, respectively) increase in reporter silencing. These data suggest that (1) the Ago1 and Ago2 pathways compete for miR-277 and (2) Ago1 and Ago2 are functionally distinct and non-redundant, with Ago2 alone mediating small RNA-directed silencing of perfectly complementary target mRNAs (RNAi) and Ago1 mediating silencing of mRNAs with central mismatches in the target sites.

miR-277 accumulation requires Ago2

Our experiments in stable S2 reporter cell lines suggest that miR-277 is loaded predominantly into an Ago2-containing RISC and that Ago1 and Ago2 compete for miR-277 in cultured *Drosophila* cells. Moreover, they suggest that miR-277 repressed the reporter to which it was fully complementary as a component of an Ago2-RISC, but repressed the bulged reporter as a component of an Ago1-RISC. Supporting this view, the cellular concentration of miR-277 decreased when Ago2 was depleted by RNAi, but not when Ago1 was depleted (Figure 2.4A). The concentration of *bantam*, a miRNA shown previously to associate with Ago1 (Okamura et al., 2004), was reduced by *ago1(RNAi)*, but unaffected by *ago2(RNAi)*. Pre-*bantam* RNA was unaltered by either treatment (Figure 2.4A), supporting the idea that the loss of *bantam* in *ago1(RNAi)* S2 cells

Figure 2.4

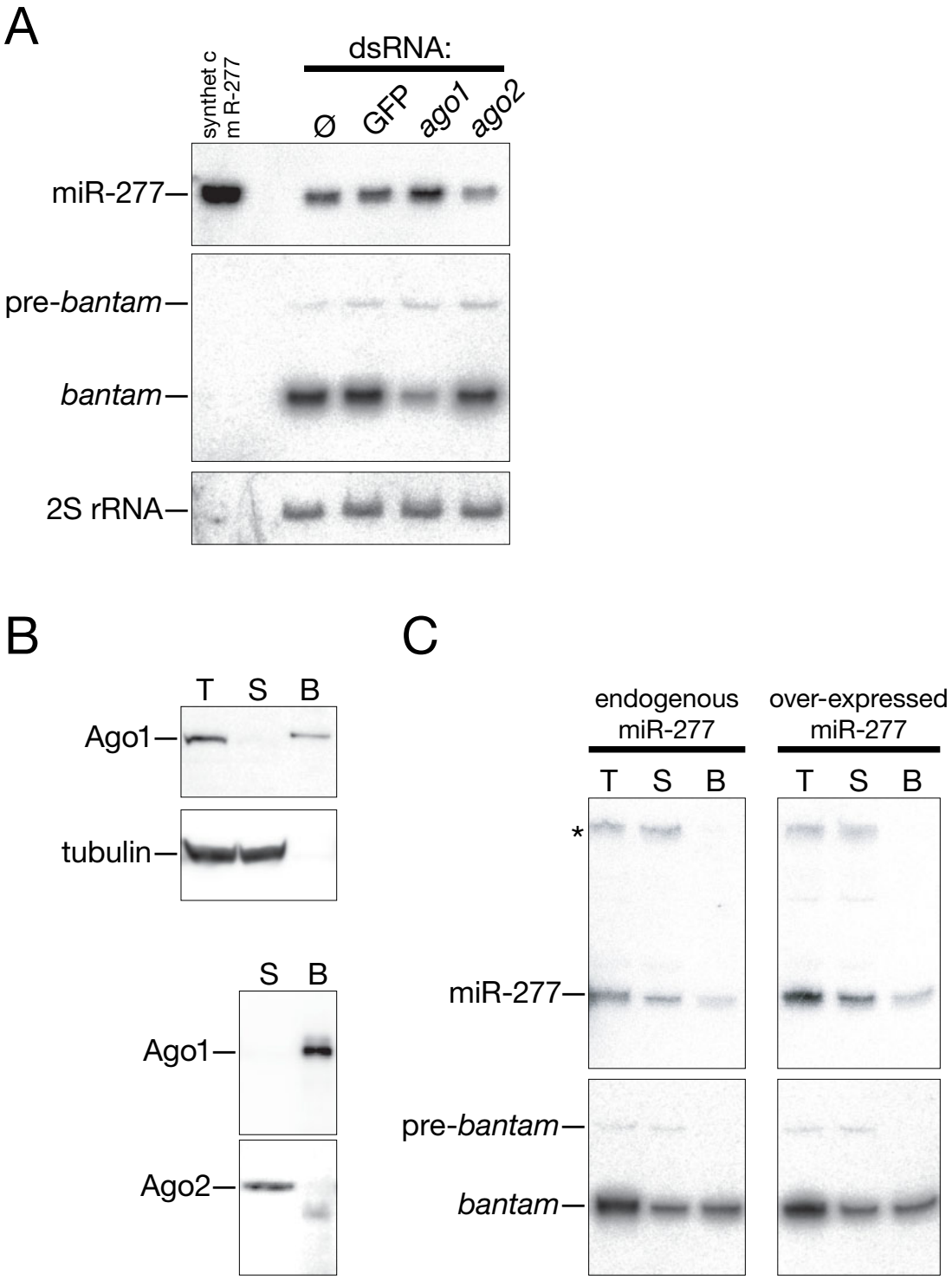

A
dsRNA:
synthet c m R-277
Ø
GFP
ago1
ago2
miR-277
pre-bantam
bantam
2S rRNA
B
T S B
Ago1
tubulin
S B
Ago1
Ago2
C
endogenous miR-277
over-expressed miR-277
T S B
T S B
*
miR-277
pre-bantam
bantam

Figure Legend 2.4. Most endogenous miR-277 is not associated with Ago1 in S2 cells.

(A) Northern analysis revealed that *ago2(RNAi)* reduced the steady-state abundance of miR-277, but not *bantam*, whereas *ago1(RNAi)* decreased the abundance of *bantam*, but not pre-*bantam* or miR-277. (B) Western blotting showed that immunoprecipitation of Ago1 depleted nearly all Ago1, but little or no Ago2, from S2 cell cytoplasmic extract. (C) The majority of *bantam*, but not pre-*bantam* co-immunoprecipitated with Ago1 (Northern analysis). In contrast, the majority of endogenous and of over-expressed miR-277 remained in the supernatant, unbound by Ago1. The asterisk marks non-specific hybridization of the probe with 5S rRNA. T: Total input; S: Supernatant; B: Bound.

reflects a failure to load the miRNA into its Ago1-RISC, rather than a defect in pre-miRNA processing, which would cause pre-*bantam* to accumulate.

Moreover, most *bantam* but only a minority of miR-277 is physically associated with Ago1 (Figures 2.4B and 2.4C). We immunoprecipitated Ago1 using a monoclonal antibody bound to agarose beads. Western blotting with the same antibody demonstrated that the overwhelming majority of Ago1, but little or no Ago2, was depleted from the supernatant and recovered with the beads (Figure 2.4B). By Northern blotting, more than half of *bantam*, but less than a third of miR-277, was recovered with the beads (Figure 2.4C).

In vivo, miR-277 is produced by Dcr-1, then loaded by Dcr-2 into Ago2

Both *dcr-1(RNAi)* and *dcr-2(RNAi)* increased GFP expression for the reporter mRNA bearing two fully complementary miR-277-binding sites (Figure 2.2A). While the effect of *dcr-1* dsRNA was anticipated, current models for the miRNA pathway in *Drosophila* do not predict a role for Dcr-2 in miRNA function. Moreover, *dcr-2(RNAi)* did not detectably alter the expression of components of the miRNA pathway, Dcr-1, Loqs, Drosha, or Ago1 (Figure 2.2B). We can imagine two explanations for the reduction in miR-277 function when Dcr-2 was depleted. Dcr-1 and Dcr-2 might both act in the production of miR-277, with each contributing to the conversion of pre-miR-277 to miR-277/miR-277* duplex. Alternatively, Dcr-1 alone might excise miR-277 from pre-miR-277, remanding the resulting miR-277/miR-277* duplex to the RISC-loading complex (RLC),

whose core constituent is the Dcr-2/R2D2 heterodimer and which is required to load siRNA duplexes into Ago2.

To distinguish between these two explanations, we examined in adult flies the expression of the same GFP reporter bearing two fully complementary miR-277-binding sites that we used in our S2 cell experiments. The *dcr-2*^{G31R} allele (Lee et al., 2004b) separates siRNA production from Ago2 loading because it selectively inactivates the nuclease function of Dcr-2. The GFP miR-277-reporters were expressed in transgenic flies heterozygous and homozygous for the *dcr*^{G31R} mutation. We also examined GFP reporter expression in *dcr2*^{L811fsX} mutant flies, which produce no Dcr-2 protein and can neither produce siRNA duplexes nor load them into Ago2. We prepared protein extracts from adult flies and measured GFP expression by Western blotting and fluorescence (Figure 2.5 and data not shown).

By both measures, expression of the reporter bearing two perfectly complementary miR-277-binding sites increased significantly in homozygous *dcr2*^{L811fsX} mutant flies, relative to that measured in extracts from their heterozygous siblings (Figure 2.5), corroborating our observation that expression of this reporter was increased in S2 cells treated with *dcr-2* dsRNA (Figure 2.3A). However, reporter expression was unaltered in homozygous *dcr-2*^{G31R} mutant flies, relative to their heterozygous siblings (Figure 2.5). Reporter expression similarly increased in flies lacking R2D2 (Figure 2.5); R2D2 acts together with

Figure 2.5

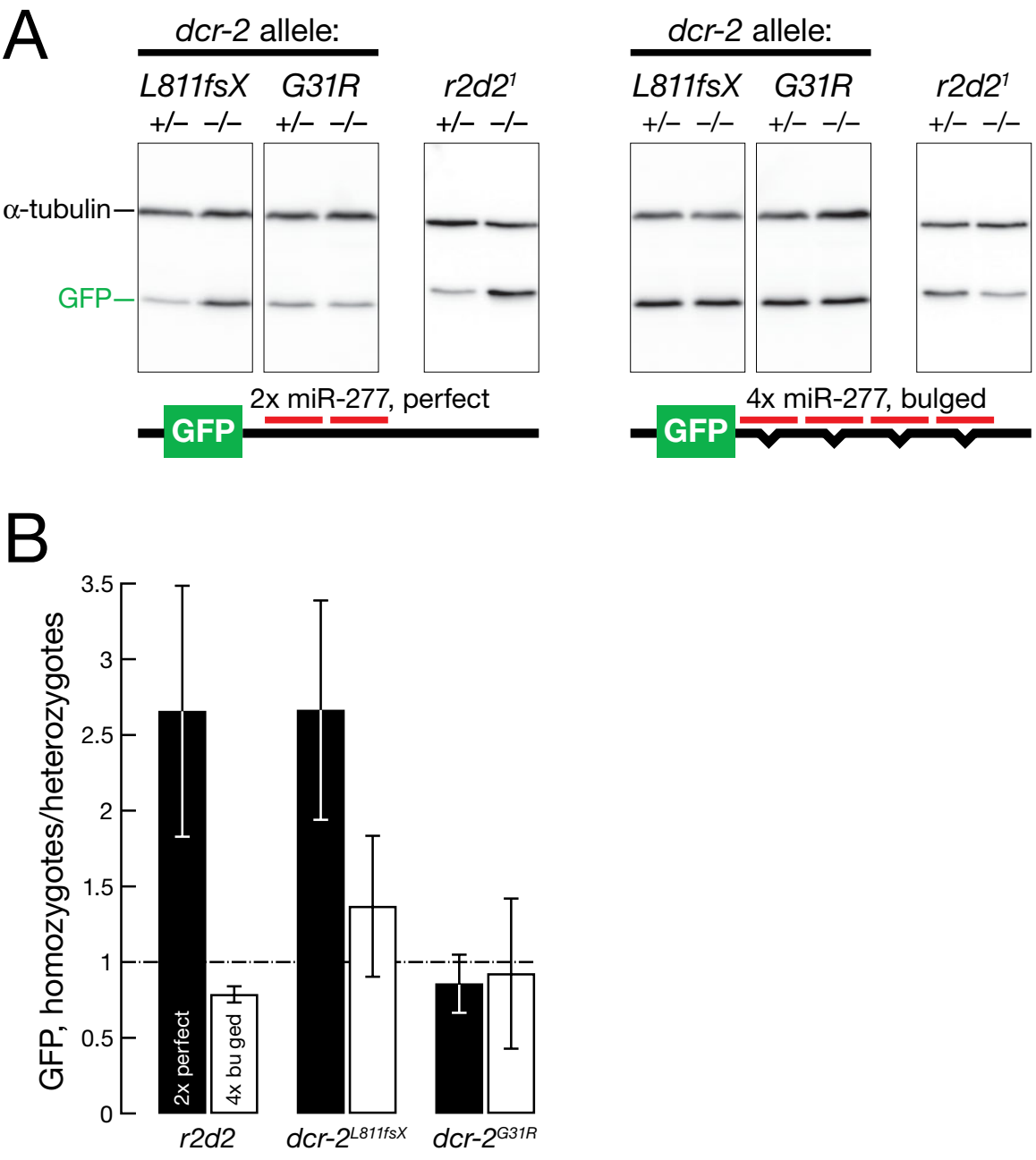

A
dcr-2 allele:
L811fsX
G31R
r2d2¹
+/– –/–
α-tubulin
GFP
2x miR-277, perfect
GFP
dcr-2 allele:
L811fsX
G31R
r2d2¹
+/– –/–
4x miR-277, bulged
GFP
B
GFP, homozygotes/heterozygotes
3.5
3
2.5
2
1.5
1
0.5
0
2x perfect
4x bu ged
r2d2
dcr-2L811fsX
dcr-2G31R

Figure Legend 2.5. In adult flies, repression of the miR-277 reporter via perfectly complementary sites requires the loading activity of Dcr-2 and R2D2, but repression via bulged sites does not.

(A) Representative Western blotting data for α-tubulin and GFP in total lysates from adult flies of the indicated heterozygous (+/–) and homozygous (–/–) mutant genotypes. (B) The average (± standard deviation) GFP expression in homozygous mutant flies, relative to heterozygotes, for three (*r2d2*) or four trials of the experiment in (A). The *dcr-2*^{L811fsX} mutant lacks detectable Dcr-2 protein, whereas the *dcr-2*^{G31R} point mutant produces a Dcr-2 protein that cannot dice long dsRNA, but can nonetheless load siRNA and miRNA/miRNA* duplexes into Ago2.

Dcr-2 to load Ago2, but is not required for siRNA production (Liu et al., 2003). We conclude that the requirement for Dcr-2 in miR-277-directed silencing of the GFP reporter bearing two fully complementary miR-277-binding sites reflects a role for Dcr-2 in loading miR-277 into Ago2, rather than in the conversion of pre-miR-277 into mature miR-277.

In contrast to the perfectly matched reporter, the GFP reporter bearing four bulged miR-277-binding sites was unaltered in flies homozygous for either the *dcr-2*^{L811fsX} null allele or the *dcr-2*^{G31R} separation-of-function allele. Thus, repression of this reporter in vivo does not require Ago2 loading, strong support for our conclusion that the bulged reporter is regulated by miR-277-programmed Ago1-RISC. In fact, we observed a small but statistically significant increase in the repression of the bulged reporter in flies homozygous for the *r2d2*^{1} allele (Figure 2.5B). These data suggest that as in vitro (Tomari et al., 2007) and in cultured cells (see above), Ago1 and Ago2 compete in vivo for loading with miR-277 and that in the absence of the Ago2-loading machinery, more miR-277-programmed Ago1-RISC is produced.

Ago1 cleaves target RNAs with low efficiency

Drosophila Ago1 retains the ability to catalyze endonucleolytic cleavage of a perfectly matched target RNA (Okamura et al., 2004; Miyoshi et al., 2005). Thus, it is surprising that for the reporter bearing two perfectly complementary miR-277-binding sites, the sub-population of miR-277 associated with Ago1 did not

detectably rescue the loss of silencing caused by depletion from S2 cells of Ago2 or the loss in adult flies of Dcr-2—a core component of the Ago2-loading machinery. To assess the molecular basis for the distinct functional capacities of Ago1 and Ago2, we analyzed in vitro the kinetics of target cleavage by each protein (Figure 2.6).

In *Drosophila*, the structure of a small RNA duplex governs into which Argonaute protein—Ago1 versus Ago2—it is loaded (Tomari et al., 2007). For the *let-7* miRNA sequence, an siRNA duplex containing *let-7* as its guide strand loads Ago2 almost exclusively, whereas the *let-7*/*let-7** duplex loads only Ago1. By adjusting the time allowed for RISC assembly, we generated approximately equal concentrations (4.6–4.7 nM) of Ago1- and Ago2-associated *let-7*. For each *let-7*-programmed RISC, we measured the rate of cleavage of a 5′ ^{32}P-radiolabled RNA target containing a single site with complete complementarity to *let-7* (Figure 2.6A).

Our data reveal two differences between *Drosophila* Ago1 and Ago2. First, Ago2 is a faster enzyme than Ago1: the initial rate of target cleavage for Ago2 was at least 12-fold greater than that of Ago1 (Figure 2.6A). Second, Ago1, unlike Ago2, failed to efficiently catalyze multiple rounds of target cleavage in vitro, even in the presence of ATP. That is, for Ago2-RISC, the rate of target cleavage was the same throughout the steady-state phase of the reaction, while the rate of target cleavage for Ago1-RISC was biphasic (Figures 2.6A).

Figure 2.6

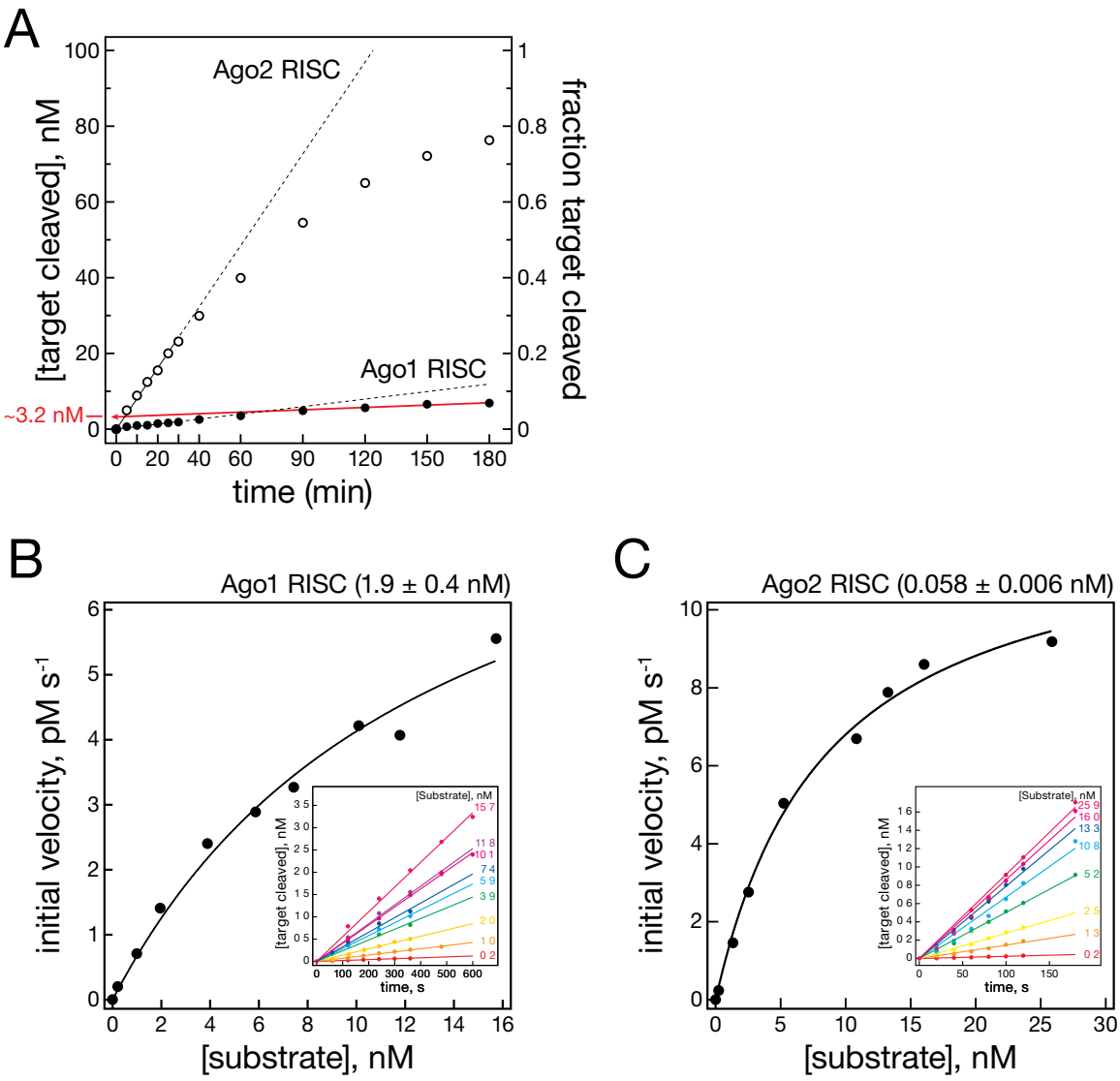

A
Ago2 RISC
Ago1 RISC
~3.2 nM
[target cleaved], nM
fraction target cleaved
time (min)
B
Ago1 RISC (1.9 ± 0.4 nM)
initial velocity, pM s-1
[substrate], nM
[Substrate], nM
[target cleaved], nM
time, s
C
Ago2 RISC (0.058 ± 0.006 nM)
initial velocity, pM s-1
[substrate], nM
[Substrate], nM
[target cleaved], nM
time, s

Figure Legend 2.6. Ago1 is a poor endonuclease.

(A) Distinct cleavage kinetics distinguish Ago1- and Ago2-RISC. At approximately equal enzyme concentrations, the initial velocity for Ago2-RISC was ~12-fold greater than that of Ago1-RISC. Cleavage by Ago2-RISC was linear throughout the reaction, as long as the substrate remained in vast excess, whereas cleavage by Ago1-RISC was biphasic, suggesting that product release is the rate-determining step. The RISC concentration estimated by burst analysis (~3.2 nM; red arrow) correlated well with that measured by 2′-*O*-methyl ASO affinity capture (~4.6 nM). (B) Pseudo-Michaelis-Menten and (C) Michaelis-Menten analyses of Ago1- and Ago2-RISC, respectively. Michaelis-Menten parameters are summarized in Table 2.1.

Such biphasic behavior was reported previously for both *Drosophila* and human Ago2-mediated target cleavage in the absence of ATP and suggests that product release is rate-determining for Ago2 under these conditions (Haley and Zamore, 2004; Rivas et al., 2005). The first phase of the Ago1 reaction in the presence of ATP likely corresponds to a pre-steady state period in which most Ago1 proteins have not yet released the product of their first target cleavage event. The second phase may correspond to a steady-state period in which the products of target cleavage are slowly released from Ago1. Supporting this view, the second phase fit well to a line whose *y*-intercept, ~3.2 nM, was similar to the amount of Ago1-RISC, ~4.7 nM, measured by affinity purification using an immobilized 2′-*O*-methyl *let-7* ASO.

The difference in efficiency between Ago1- and Ago2-catalyzed target cleavage might reflect a difference in the rate of catalysis or in the affinity for the target RNA. To distinguish between these two explanations, we performed a kinetic analysis of Ago1- and Ago2-RISC programmed with *let-7* in *Drosophila* embryo lysate. We estimated the amount of active *let-7* programmed RNAi enzyme complex from the size of the burst for Ago1-RISC in the presence of ATP and for Ago2-RISC by depleting ATP after RISC assembly. In the presence of ATP, Ago2-RISC conforms to a simple Michaelis-Menten scheme (Haley and Zamore, 2004):

Ago2-*let-7* + target RNA <—> [Ago2-*let-7*•target RNA] —> Ago2-*let-7* + cleaved target RNA

In contrast, Ago1-RISC appears to follow a more complex kinetic scheme, even in the presence of ATP:

Ago1-*let-7* + target RNA <—> [Ago1-*let-7*•target RNA] —>

[Ago1-*let-7*•cleaved target RNA] <—> Ago1-*let-7* + cleaved target RNA

Because Michaelis-Menten parameters are determined from the initial velocity of the enzyme observed at different concentrations of substrate (target RNA), we analyzed Ago1 as if it followed the same kinetic scheme as Ago2. This pseudo-Michaelis-Menten analysis allows the Ago1 K_M to retain the same meaning—an approximation of the affinity of the enzyme for its substrate—as that determined for Ago2. Our data (Figure 2.6B and Table 2.1) suggest that *let-7*-programmed Ago1 and Ago2 bind the *let-7* complementary sequence in the target RNA with nearly the same affinity. In vitro, Ago1 binds its target RNAs as well as Ago2, but cleaves them much more slowly than Ago2. This suggests that in vivo, Ago1 is too inefficient to silence a perfectly matched target by endonucleolytic cleavage.

Table 2.1. Kinetic analysis of *Drosophila* Ago1- and Ago2-RISC.

	K_M (nM)	V_{max} (nM s^{-1})	[RISC] (nM)	k_{cat} (s^{-1})	**k_{cat} (relative)**	k_{cat}/K_M (nM^{-1} s^{-1})	**k_{cat}/K_M (relative)**
Ago1	13.3 ± 3.2	0.0096 ± 0.0013	1.9 ± 0.4	0.005 ± 0.0013	**1**	0.4 ± 0.1	**1**
Ago2	8.4 ± 1.0	0.0125 ± 0.006	0.058 ± 0.006	0.215 ± 0.025	**43**	25.5 ± 4.3	**64**

DISCUSSION

In lysates from *Drosophila* embryos, in cultured *Drosophila* S2 cells, and in adult flies, miRNA can be loaded into both Ago1 and Ago2. Our data suggest that sorting miRNAs into Ago1- and Ago2-RISC generates silencing complexes with distinct functional capacities: Ago1-RISC represses expression of targets with which its guide miRNA matches only partially, whereas Ago2 silences fully matched target RNAs. These differences result, in part, from the surprisingly different catalytic efficiencies of Ago1 and Ago2: only Ago2 catalyzes robust, multiple-turnover target cleavage.

Why does *Drosophila* Ago1 retain its endonuclease activity?

In mammals, only Ago2 retains the ability to catalyze guide RNA-directed endonucleolytic cleavage of RNA; the three other mammalian Argonaute proteins, Ago1, Ago3, and Ago4, lack a functional active site that is presumed to have been present in the evolutionarily ancestral Argonaute protein. Why then has *Drosophila* Ago1 retained any endonuclease activity at all, if it is so inefficient at target cleavage that it cannot measurably contribute to small RNA-directed RNAi? One potential explanation is that the primary role of the Ago1 endonuclease activity is to facilitate loading of Ago1-RISC. That is, the predominant substrate for the Ago1 endonuclease is not target RNA, but rather miRNA* strands, and perhaps the occasional siRNA passenger strand. Because miRNA* strand cleavage would occur only in *cis* and only once per loaded

Ago1-RISC, efficient, multiple-turnover cleavage of target RNA would not be required.

Our data reveal an important biochemical difference between Ago2 and Ago1, but they do not explain the molecular basis for the inefficiency of Ago1-directed cleavage of target RNA. We can envision two explanations for the more than 40-fold lower k_{cat} of Ago1 compared to Ago2. First, the active site of Ago1 might be less well suited to catalyzing phosphodiester bond cleavage. Alternatively, Ago1 might be slow to assume a catalytically active conformation. In this second model, the rate of a conformational rearrangement would limit the speed of target RNA cleavage by Ago1. Such a conformational rearrangement of the siRNA guide has been proposed previously for Ago2 (Filipowicz, 2005; Tomari and Zamore, 2005).

Implications for the mechanism of guide strand choice

Neither the current genome sequence of *Drosophila melanogaster* nor GenBank in its entirety contains a *Drosophila* mRNA with complete complementarity to miR-277. Why then do flies load miR-277 into Ago2-RISC? Perhaps there are—yet unknown—viral RNAs targeted by Ago2-loaded miR-277. Such an innate immune response function has previously been proposed for miRNAs in mammals (Lecellier and Voinnet, 2004). Regardless of the biological purpose for loading miR-277 into Ago2, miR-277 provides an important in vivo test of the controversial proposal that the production of small RNA duplexes by Dicer is

uncoupled from the loading of Argonaute proteins (Aza-Blanc et al., 2003; Khvorova et al., 2003; Schwarz et al., 2003). That Dcr-2 and R2D2 act in vivo to load Ago2 with miR-277, a miRNA produced by Dcr-1 and Loqs, confirms previous in vitro data suggesting that both ends of a small RNA duplex are available for examination by the Ago2 loading machinery (Schwarz et al., 2003; Tomari et al., 2004b; Preall et al., 2006). Our results suggest that the miR-277/miR-277* duplex dissociates from Dcr-1 after the dicing of pre-miR-277 and is then bound by the Dcr-2/R2D2 heterodimer, which loads it into Ago2; Sontheimer and colleagues reached similar conclusions about small RNA loading from in vitro experiments that asked if dicer processing and Ago2-loading were coupled (Preall et al., 2006).

We reason that Ago1 loading is also uncoupled from dicing. In all animals, some miRNAs are found on the 5′ and other on the 3′ arm of their pre-miRNA stem-loops. In contrast, the geometry of Dcr-1 with respect to the two arms of the pre-miRNA stem is essentially the same for all miRNAs: Dcr-1 always makes staggered cuts that separate the pre-miRNA loop from the miRNA/miRNA* duplex. If Dcr-1 were to load miRNAs directly into Ago1, without first releasing the miRNA/miRNA* duplex, we would expect that all miRNAs would reside on the same arm of the pre-miRNA stem. The simplest explanation, and one most consistent with the partitioning of miR-277 into both Ago1- and Ago2-RISCs, is that miRNA/miRNA* duplexes are released from Dicer immediately after their

production, then rebound by the Ago1- and Ago2-loading machineries. Such a model allows both the terminal thermodynamics of the miRNA/miRNA* duplex to determine the mature miRNA strand (rather than its position within the pre-miRNA) and the pattern of mismatches within the duplex to determine how the miRNA partitions between Ago1 and Ago2.

Why are Ago1 and Ago2 functionally specialized?

In mammals, siRNAs produce off-target effects largely by acting like miRNAs (Jackson et al., 2003; Lim et al., 2005; Jackson et al., 2006). In flies, siRNAs loaded into Ago2 are believed to defend against viral infection (Wang et al., 2006b; Galiana-Arnoux et al., 2006). Virus-derived siRNAs might therefore trigger widespread, off-target silencing of host genes as flies mount an anti-viral RNAi response. The partitioning of siRNAs into Ago2-RISC appears to circumvent this problem, because silencing by *Drosophila* Ago2 requires greater complementarity between the siRNA and its target than silencing by Ago1. It is tempting to speculate that a similar functional specialization among Argonaute proteins has gone undetected in mammals.

EXPERIMENTAL PROCEDURES

Construction of cell lines with increased miR-277 expression

A 270 nt fragment of genomic DNA encompassing the miR-277 sequence was PCR-amplified from S2-cell genomic DNA using the oligonucleotides 5′-GCG GAT CCG GTA CCT ATA CAT ATA TAA CGA GGC CTA ACG-3′ and 5′-ATG CGG CCG CAA AAC AGT GTC TTA CAA ACA AGT GG-3′ and cloned *Bam*HI to *Not*I into pKF62, yielding a mini-pri-miR-277 transgene under the control of the ubiquitin promoter.

Cell culture and flow cytometry

Drosophila Schneider 2 (S2) cells were cultured at 28°C in Schneider's medium (Invitrogen, Carlsbad, CA, USA) supplemented with 10% (v/v) fetal bovine serum (Invitrogen). GFP expression plasmids were transfected with siLentfect (see below) at 1 mg of plasmid per well of a 24-well plate. For selection of stable transformants, 20 ng of phsNeo (Steller and Pirrotta, 1985) was co-transfected with 1 mg each GFP reporter plasmid. Five days after transfection, cells were split 1:5 into medium supplemented with 1.2 mg/ml G418 (Invitrogen) and diluted 1:5 every 7 days into G418-containing medium for three weeks, then serial dilutions were plated in a 96-well plate in growth medium supplemented with 1% (v/v) sterile-filtered conditioned medium. After two weeks, wells with a single colony of cells were expanded and analyzed by flow cytometry. Cell

clones that produced a single peak in the flow cytometer were retained as reporter lines.

For dsRNA transfection, cells were seeded at a density of 1 x 10^6 cells/ml in 24-well plates (500 ml/well) using Schneider's medium without G418. 100 ml Liposome/nucleic acid complexes (prepared by incubating at room temperature for 45 min 1.5 ml siLentfect [BioRAD, Hercules/CA, USA] or Dharmafect 4 [Darmacon, Lafayette, CO] and either 1 mg of dsRNA or 10 pmol of 3′ cholesterol-conjugated, 2′-*O*-methyl modified antisense-oligonucleotide [see Figure 2.S1B] in 100 ml Schneider's medium without serum) were added per well of a 24 well plate. After 6 days, cells were analyzed by FACS (BD FACScan flow cytometer; Becton Dickinson, Franklin Lakes, NJ). GFP-expression was quantified as the arithmetic mean of fluorescence (CellQuest; Becton Dickinson).

Anti-Dcr-1 and Ago2 antibodies

KLH-conjugated peptides (Dcr-1: CQGLIAKKD; R2D2: CSDEYESSKDKAMD) or the Ago2 PAZ-domain fused to glutathione-S-transferase (Lingel et al., 2003) were used to immunize rabbits (Covance Research Products, Denver, PA, USA or ProSci, Poway, CA, USA) or chickens (Gallus Immunotech, Cary, NC, USA), and the antiserum affinity-purified using immobilized peptide (Sulfolink, Pierce, Rockford, IL, USA) or NusA-Ago2-PAZ fusion protein (Aminolink plus, Pierce).

Western Blotting

Proteins from cultured S2 cells or from hand-dissected adult fly heads and thoraces were extracted in PBS containing 1% Triton X-100 (Sigma, St. Louis, MO, USA) and protease inhibitors (Complete without EDTA, Roche Molecular Biochemicals, Basel, Switzerland). To quantify myc_3GFP-expression in transgenic flies, 20 mg of total protein was resolved by electrophoresis through a 12% polyacrylamide/SDS gel and transferred to PVDF-membrane (Immobilon-P, Millipore, Billerica, MA, USA) by semi-dry transfer (BioRAD, Hercules, CA, USA) at 20 V for 120 min in 25 mM Tris (pH 8.3), 250 mM glycine, 10% (v/v) methanol as anode buffer and 20 mM CAPS (pH 11.0) as cathode buffer. The blot was incubated for 90 min at room temperature with purified monoclonal anti-myc 9E10 (Sigma #M4439) diluted 1:1000 in 25 mM Tris, 137.5 mM NaCl, 2.5 mM KCl, 0.02% (v/v) Tween-20 (Sigma) for 90 min at room temperature and then HRP-conjugated goat anti-rabbit secondary antibody diluted 1:1,000, then developed with Pierce SuperSignal West Dura kit (Pierce). HRP-conjugated rabbit anti-chicken secondary antibody (Gallus Immunotech), diluted 1:15,000, was used to detect the affinity purified anti-R2D2 chicken IgY. Western Blot images were acquired using a Fuji LAS-3000 (Fujifilm Life Sciences, Stamford, CT, USA) and quantified using ImageGauge (Fujifilm Life Sciences). α-tubulin was detected with DM1A antibody (Sigma #T6199) diluted 1:1,000. myc_3-GFP western blot signals were corrected using a standard curve for created by diluting extract from pKF63-transgenic flies into extract from yw flies.

For S2 cell proteins, 50 mg total protein was resolved by electrophoresis through an 8% polyacrylamide/SDS gel, transferred to PVDF, incubated with primary antibodies (1:2,000) overnight at 4°C, and secondary antibodies for 120 min at room temperature. Anti-Dcr-2 antibody was the kind gift of Qinghua Liu (Liu et al., 2003).

Ago1 and Ago2 target cleavage kinetics

Target cleavage reactions were performed essentially as described (Haley and Zamore, 2004; Haley et al., 2003). In Figure 7A, 50 nM *let-7* siRNA or *let-7*/*let-7** duplex was incubated with *Drosophila* 0–2 h embryo lysate for 2 min to program Ago2- (~4.7 nM) or 5 min for Ago1-RISC (~4.6 nM). In Figure 2.7B, 20 nM *let-7* siRNA or *let-7*/*let-7** duplex was incubated with lysate for 3 min or 8 min to program Ago2- or Ago1-RISC. For Ago2 cleavage in Figure 7B, RISC was diluted 10-fold in *N*-ethyl maleimide (NEM) treated embryo lysate (Nykanen et al., 2001; Haley and Zamore, 2004). *let-7* siRNA assembled little or no active RISC in *ago2*^{414} lysate whereas *let-7*/*let-7** was as active in *ago2*^{414} as in wild-type lysate, indicating that the *let-7* siRNA and *let-7*/*let-7** duplex are almost exclusively loaded into Ago2- and Ago1-RISC, respectively.

RISC assembly was stopped by treatment with NEM followed by DTT to quench unreacted NEM for both Ago1- (Figure 2.S4A) and Ago2-RISC (Nykanen et al., 2001). Control experiments (Figure 2.S4B) established that the biphasic kinetics of Ago1-RISC in the presence of ATP were not a consequence of

treatment with NEM. RISC concentration was estimated by 2′-*O*-methyl ASO affinity purification in Figure 2.6A and by the size of the pre-steady-state burst in Figure 6B (Haley and Zamore, 2004; Schwarz et al., 2003). The concentration of RNA target was 100 nM in Figure 7A and 0.5 to 100 nM in Figure 6B. Data were analyzed using IGOR 5 (WaveMetrics) and *Visual*Enzymics 2005 (Softzymics) software.

Construction of reporter plasmids and RNAi trigger dsRNAs

To create an expression vector for both cultured cells and transgenic flies, we PCR amplified the 3′ UTR and SV40 poly-A signal from plasmid pEGFP-N1 (Clontech, Mountain View, CA, USA) with oligonucleotides 5´-ATC ACT CTC GGC ATG GAC GAG-3′ and 5′-GTG AAT TCA TAC ATT GAT GAG TTT GGA C-3′ and inserted the resulting PCR product into pUbi-Casper2 (a kind gift of Dr. Siu Ing The) using the *Not*I and *Eco*RI restriction sites, creating vector pKF60. For the GFP-insert, we transferred a *Bam*HI-*Not*I fragment from pEGFP-N1 (Clontech) into pBluescript (Stratagene, La Jolla, CA, USA) cut with *Bam*HI/*Not*I, creating pKF20. Subsequently, we annealed the oligos 5′-CAT GGA ACA AAA ACT TAT TTC TGA AGA AGA CTT GGG-3′ and 5′-CAT GCC CAA GTC TTC TTC AGA AAT AAG TTT TTG TTC -3′, encoding a myc-tag, and ligated this DNA-fragment into *Nco*I-cut pKF20. After sequencing, one clone was selected that contained a triple insertion in the correct orientation (pKF30). From this plasmid, the myc_3-GFP-sequence was transferred as a *Bam*HI-*Not*I-fragment into pKF60,

resulting in plasmid pKF62. To remove an *Xba*I-site from the pCASPER2 polylinker, pKF62 was cut with *Xba*I, the ends treated with Klenow polymerase (New England Biolabs, Ipswitch, MA, USA), and the vector was re-ligated, creating pKF63. This plasmid was transformed into *dam*/*dcm* negative bacteria (strain GM2163, New England Biolabs), which rendered a second *Xba*I-site in the 3′-UTR, adjacent to the *Not*I-site, cleavable. To insert the miR-277 target sites, we annealed oligos 5′-GGC CTG TCG TAC CAG ATA GTG CAT TTA CAG TGT CGT ACC AGA TAG TGC ATT TA-3′ and 5′-CTA GTA AAT GCA CTA TCT GGT ACG ACA CTG TAA ATG CAC TAT CTG GTA CAG CA-3′ for the two perfectly matched sites, and oligos 5′-GGC CTG TCG TAC CAG AGG ATG CAT TTA CAG TGT CGT ACC AGA GGA TGC ATT TAT GTC GTA CCA GAG GAT GCA TTT ACA GTG TCG TAC CAG AGG ATG CAT TTA -3′ and 5′-CTA GTA AAT GCA TCC TCT GGT ACG ACA CTG TAA ATG CAT CCT CTG GTA CGA CAT AAA TGC ATC CTC TGG TAC GAC ACT GTA AAT GCA TCC TCT GGT ACG ACA-3′ for the four bulged sites, then ligated the DNA fragments into *Not*I-*Xba*I-cut pKF63, creating pKF67 and pKF68, respectively.

These pCASPER2-derived expression plasmids were used both for the generation of stable S2-cell lines and for the P-element-mediated genetic transformation of *Drosophila melanogaster* (Rubin and Spradling, 1982).

Constructs to make dsRNA directed against GFP, *dcr-1*, *dcr-2*, *loqs* and *drosha* were described previously (Forstemann et al., 2005). Templates for the

synthesis of dsRNA directed against *ago1* and *ago2* were generated by T/A-cloning PCR products generated using the oligonucleotides 5′-CGC ACC ATT GTG CAT CCT AAC GAG-3′ and 5′-GGG GAC AAT CGT TCG CTT TGC GTA-3′ for *ago2* and 5′-ATT TGA TTT CTA TCT ATG CAG CCA-3′ and 5′-GCC CTG GCC ATG GCA CCT GGC GTA-3′ for *ago1* into the modified Litmus28i vector described previously (Forstemann et al., 2005). The template for producing dsRNA-targeting *r2d2* was generated by PCR using oligonucleotides 5′-CGT AAT ACG ACT CAC TAT AGG CAT ACA CGG CTT GAT GAA GGA TTC-3′ and 5′-CGT AAT ACG ACT CAC TAT AGG TTG CTT GTG CTC GCT ACT TGC-3′. Templates for in vitro transcription were generated by PCR-amplification of each plasmid construct using a single primer corresponding to the T7 promoter (5′-CGT AAT ACG ACT CAC TAT AGG-3′) and dsRNA for knock-down was generated as described in (Haley et al., 2003).

ACKNOWLEDGMENTS

We thank Alicia Boucher for assistance with fly husbandry, Gwen Farley for technical assistance, and members of the Zamore lab for advice, suggestions, and critical comments on the text. PDZ is a W.M. Keck Foundation Young Scholar in Medical Research. This work was supported in part by grants from the National Institutes of Health to PDZ (GM62862 and GM65236) and post-doctoral fellowships from the Human Frontier Science Program to KF and YT. FACS core resources were supported by a National Institutes of Health grant (NIDDKD 5 P30 DK032520) to the University of Massachusetts Medical School Diabetes Endocrinology Research Center.

Figure 2.S1

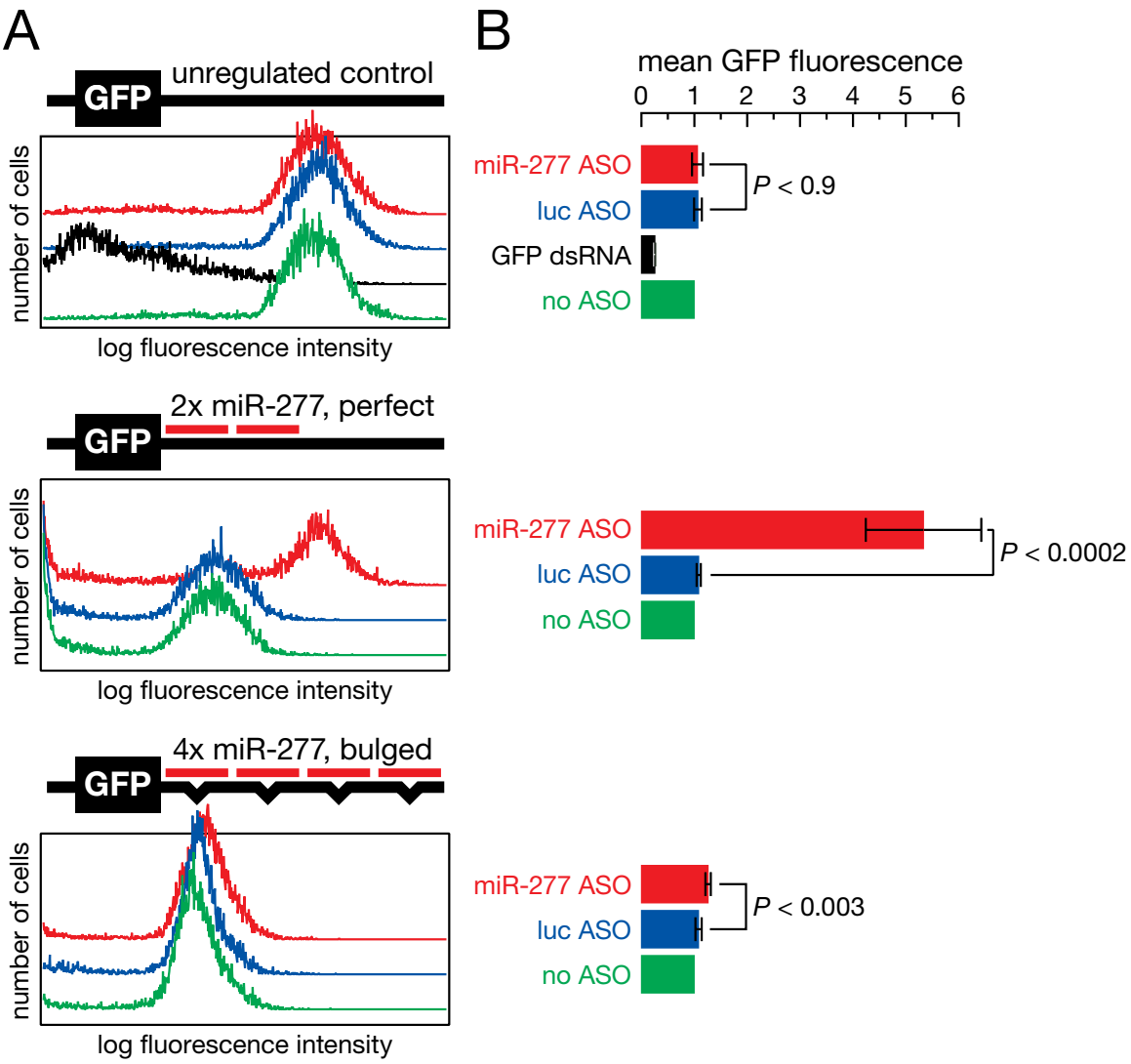

A
unregulated control
GFP
number of cells
log fluorescence intensity
2x miR-277, perfect
GFP
number of cells
log fluorescence intensity
4x miR-277, bulged
GFP
number of cells
log fluorescence intensity
B
mean GFP fluorescence
0 1 2 3 4 5 6
miR-277 ASO
luc ASO
GFP dsRNA
no ASO
P < 0.9
miR-277 ASO
luc ASO
no ASO
P < 0.0002
miR-277 ASO
luc ASO
no ASO
P < 0.003

Figure Legend 2.S1. Regulation of GFP reporter expression in cultured Drosophila S2 cells by endogenous miR-277.

Clonally derived stable cell lines were generated that expressed control GFP unregulated by miR-277, GFP bearing two miR-277-complementary sites in its 3′ untranslated region (UTR), and GFP bearing in its 3′ UTR four miR-277-complementary sites, each containing three mismatches to miR-277 at nucleotides 9, 10 and 11, producing a 'bulge.' Each cell line was transfected with a cholesterol conjugated 2′-*O*-methyl modified, antisense oligonucleotide (ASO) complementary to miR-277 or to an unrelated luciferase sequence. In cultured cells and in vivo, ASOs inhibit the function of miRNAs to which they are complementary, relieving repression of their mRNA targets (Krutzfeldt et al., 2005; Meister et al., 2004; Hutvagner et al., 2004}. As a control, the unregulated GFP reporter cell line was transfected with GFP dsRNA. GFP expression was quantified by flow cytometry. (A) Representative FACS profiles from a single experiment. (B) The average ± standard deviation for the mean fluorescence recorded in three trials. *P*-values were calculated using a two-sample T-test assuming equal variances.

Figure 2.S2

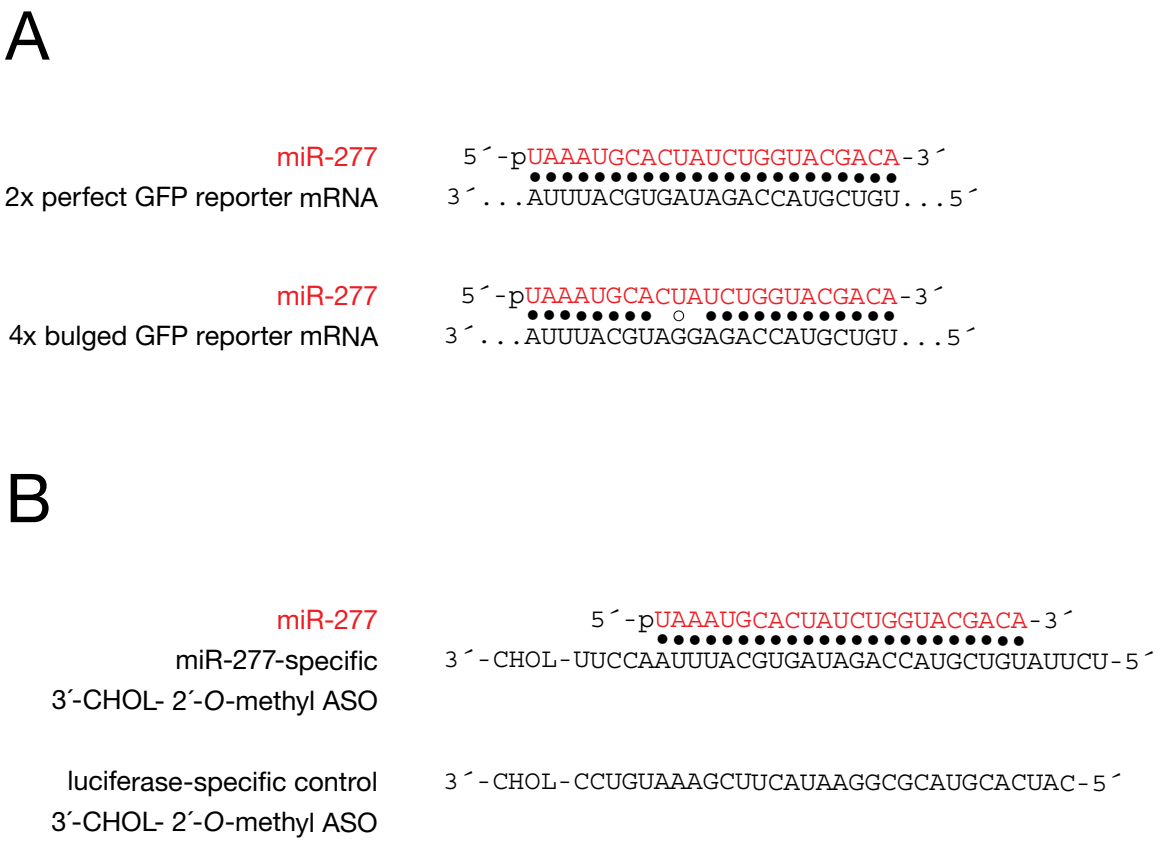

A
miR-277
5´-pUAAAUGCACUAUCUGGUACGACA-3´
2x perfect GFP reporter mRNA
3´...AUUUACGUGAUAGACCAUGCUGU...5´
miR-277
5´-pUAAAUGCACUAUCUGGUACGACA-3´
4x bulged GFP reporter mRNA
3´...AUUUACGUAGGAGACCAUGCUGU...5´
B
miR-277
5´-pUAAAUGCACUAUCUGGUACGACA-3´
miR-277-specific
3´-CHOL-UUCCAAUUUACGUGAUAGACCAUGCUGUAUUCU-5´
3´-CHOL- 2´-O-methyl ASO
luciferase-specific control
3´-CHOL-CCUGUAAAGCUUCAUAAGGCGCAUGCACUAC-5´
3´-CHOL- 2´-O-methyl ASO

Figure Legend 2.S2. siRNA, siRNA-binding-site, and antisense oligonucleotide (ASO) structures.

(A) Structure of the miR-277-binding sites in 3′ UTR of the perfectly complementary or the partially complementary GFP reporter mRNAs. (B) Structure of the 3′-cholesterol conjugated, 2′-*O*-methyl modified, antisense oligonucleotides (ASOs) used as a control (luciferase-specific) or used to inhibit miR-277 expression in cultured S2 cells. Every ribose 2′ hydroxyl in each ASO was replaced with a methoxy group.

Figure 2.S3

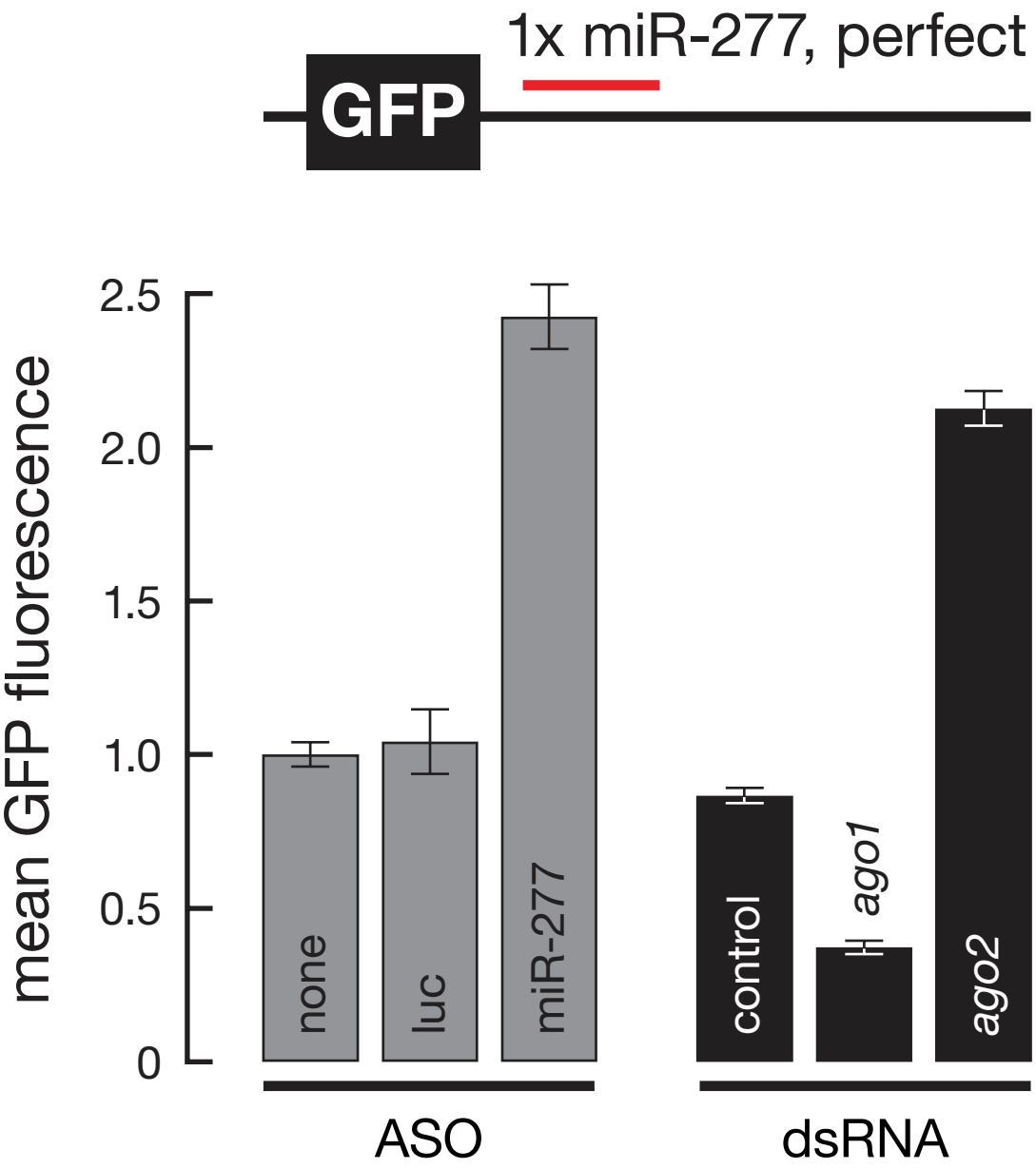

1x miR-277, perfect
GFP
mean GFP fluorescence
2.5
2.0
1.5
1.0
0.5
0
none
luc
miR-277
control
ago1
ago2
ASO
dsRNA

Figure Legend 2.S3. Endogenous miR-277-programmed Ago2, not Ago1, represses a GFP reporter containing a single, perfectly complementary site in its 3′ UTR.

At left, clonally derived S2 cells bearing the GFP reporter were transfected with the indicated antisense oligonucleotide (ASO), including an ASO complementary to miR-277. At right, the cells were transfected with dsRNA corresponding to the indicated gene. Each bar represents the average ± standard deviation for three independent experiments.

Figure 2.S4

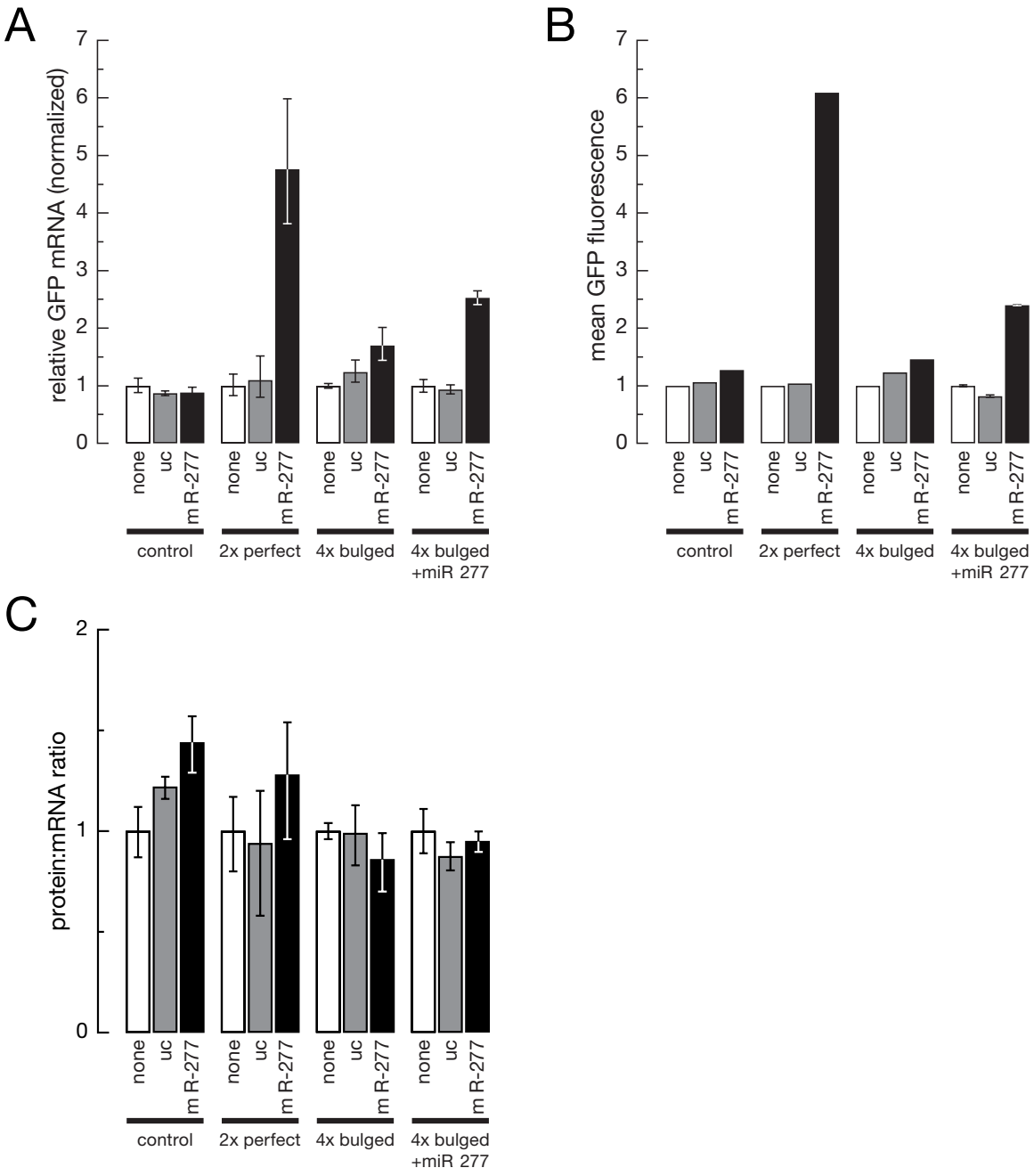

A
relative GFP mRNA (normalized)
0
1
2
3
4
5
6
7
none
uc
m R-277
control
2x perfect
4x bulged
4x bulged +miR 277
B
mean GFP fluorescence
C
protein:mRNA ratio

Figure Legend 2.S4. miR-277 regulates reporter mRNA steady state abundance.

Inhibition of miR-277 by a specific ASO increased both the mRNA abundance (A) and the production of GFP protein (B) for the perfect and bulged GFP reporters, but not the unregulated control, even when additional miR-277 ("+miR-277") was expressed from a transgene. After 72 hrs, mRNA abundance was measured by qRT-PCR and GFP protein expression measured by FACS analysis. Three replicate transfections were performed for the cell line containing the 4x bulged reporter and expressing additional miR-277; single transfections were performed for all other cell lines. For qRT-PCR, total RNA was extracted from ~107 cells with Trizol (Invitrogen, Carlsbad, CA). Reverse transcription was performed with Superscript II (Invitrogen) according to the manufacturer's instructions, using gene specific primers (forward primer and RT/reverse primer) for the GFP and RP49 coding regions: 5′-CCG CTT CAA GGG ACA GTA TCT G-3′ and 5′-ATC TCG CCG CAG TAA ACG C-3′ for RP49; 5′-TGT CGG GCA GCA GCA C-3′ and 5′-AAC GGC ATC AAG GTG AAC TTC-3′ for GFP. Relative GFP mRNA abundance was calculated using the 2-ΔΔCt method. Values were normalized to the no treatment control. Error bars represent the standard deviation of three PCR replicates. For the 4x bulge+miR277 sample, error bars represent standard deviation of the means of three independent transfection experiments. (C) miR-277 does not affect the

ratio of reporter GFP protein to RNA. For each cell line tested, the ratio of the mean GFP fluorescence to the relative GFP mRNA abundance is shown, normalized to the no treatment control.

Figure 2.S5

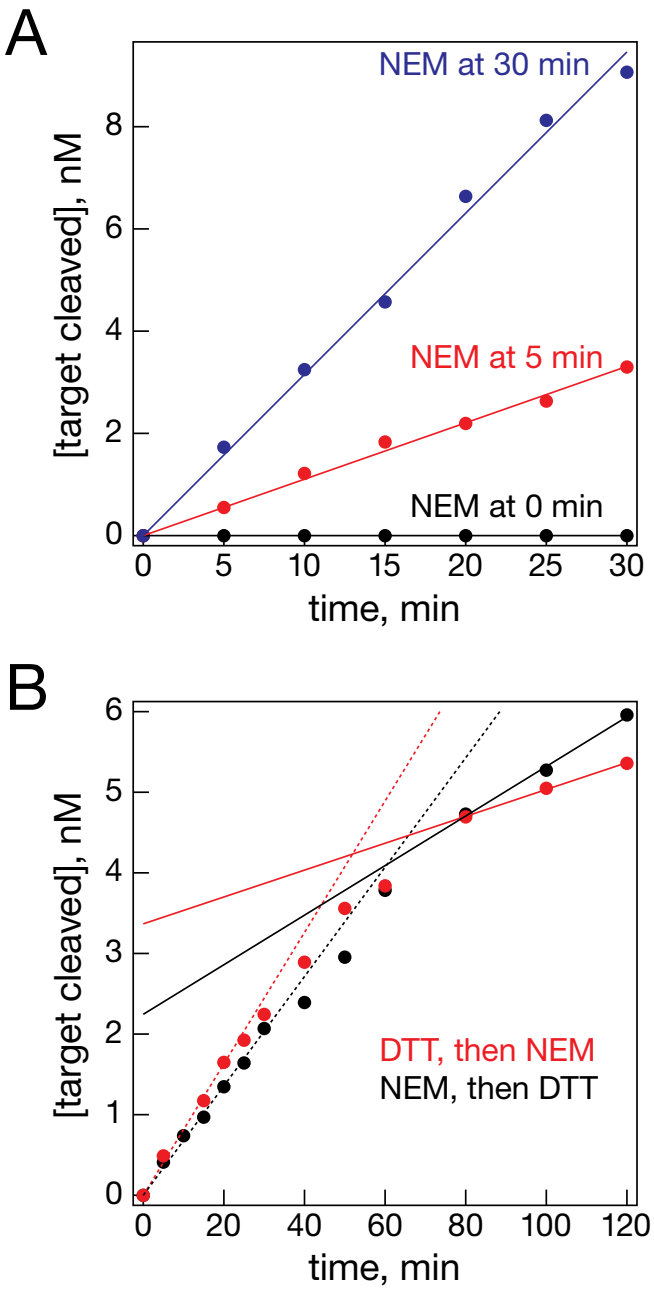

A
NEM at 30 min
NEM at 5 min
NEM at 0 min
[target cleaved], nM
time, min
0
2
4
6
8
0
5
10
15
20
25
30
B
DTT, then NEM
NEM, then DTT
[target cleaved], nM
time, min
0
1
2
3
4
5
6
0
20
40
60
80
100
120

Figure Legend 2.S5. *N*-ethyl maleimide (NEM) inactivates assembly of Ago1-RISC, but does not alter the kinetics of Ago1-mediated target cleavage.

(A) Ago1 was programmed with *let-7* in vitro using 50 nM *let-7*/*let-7** duplex, then the reaction treated with NEM at the indicated time. Unreacted NEM was quenched with DTT. The relative amount of Ago1-RISC assembled was then determined by measuring the rate of cleavage of 50 nM target RNA containing a single *let-7*-complementary site. (B) 50 nM *let-7*/*let-7** duplex was incubated with embryo lysate for 60 min to program Ago1-RISC, then the reaction treated with NEM, followed by DTT. As a control, DTT was added before the NEM. The cleavage of 200 nM target RNA by *let-7*-programmed Ago1-RISC was then assayed. Both reactions displayed the burst kinetics characteristic of Ago1-mediated target cleavage.

Chapter III: Argonaute Divides Its RNA Guide Into Domains with Distinct Functions and RNA-Binding Properties

Disclaimer

This chapter was made possible from a joint effort among the authors: Liang Meng Wee (Wee), Fabián Flores-Jasso (FF), William Salomon (WES) and Phillip Zamore (PDZ). FF with the help of WES devised the oligocapture-based method to purify fly Ago2 and mouse AGO2 from extracts.

SUMMARY

MicroRNAs (miRNAs) and small interfering RNAs (siRNAs) guide Argonaute proteins to silence mRNA expression. Argonaute binding alters the properties of an RNA guide, creating functional domains. We show that the domains established by Argonaute—the anchor, seed, central, 3′ supplementary, and tail regions—have distinct biochemical properties that explain the differences between how animal miRNAs and siRNAs bind their targets. Extensive complementarity between an siRNA and its target slows the rate at which fly Argonaute2 (Ago2) binds to and dissociates from the target. Highlighting its role in antiviral defense, fly Ago2 dissociates so slowly from extensively complementary target RNAs that essentially every fully paired target is cleaved. Conversely, mouse AGO2, which mainly mediates miRNA-directed repression, dissociates rapidly and with similar rates for fully paired and seed-matched targets. Our data narrow the range of biochemically reasonable models for how Argonaute-bound siRNAs and miRNAs find, bind, and regulate their targets.

RESULTS

To determine how siRNA:target pairing affects Ago2 function, we systematically altered the sequence of an siRNA whose guide strand corresponds to the *let-7* miRNA. We measured the rate of cleavage of a target RNA that was fully complementary to *let-7* for 45 variants of the siRNA. (Figure 3.S1). The use of a common target eliminated the influence on Ago2 activity of target site accessibility (Brown et al., 2005; Ameres et al., 2007; Long et al., 2007; Tafer et al., 2008).

Of 26 overlapping dinucleotide mismatches, 22 reduced the rate of target cleavage by *Drosophila* Ago2 (Figure 3.S1). To understand why some mismatches were tolerated but others were not, we determined the Michaelis-Menten parameters, K_M and k_{cat}, for 59 siRNA:target combinations comprising seven single-nucleotide mismatches, 21 dinucleotide mismatches, a contiguous g17–g21 mismatch, and 30 fully complementary siRNA:target pairs (Figures 3.1A, 3.1B and Table 3.S1). Each siRNA was assembled into Ago2-RNA-induced silencing complex (RISC) in *Drosophila* embryo lysate. Half the assembly reaction was used to measure the initial rates of cleavage for a mismatched target and half for a fully complementary target (Figures 3.1A and 3.S2). Because the RISC concentration was identical for the two targets, the change in k_{cat} attributable to the mismatches corresponded to mismatched

Figure 3.1

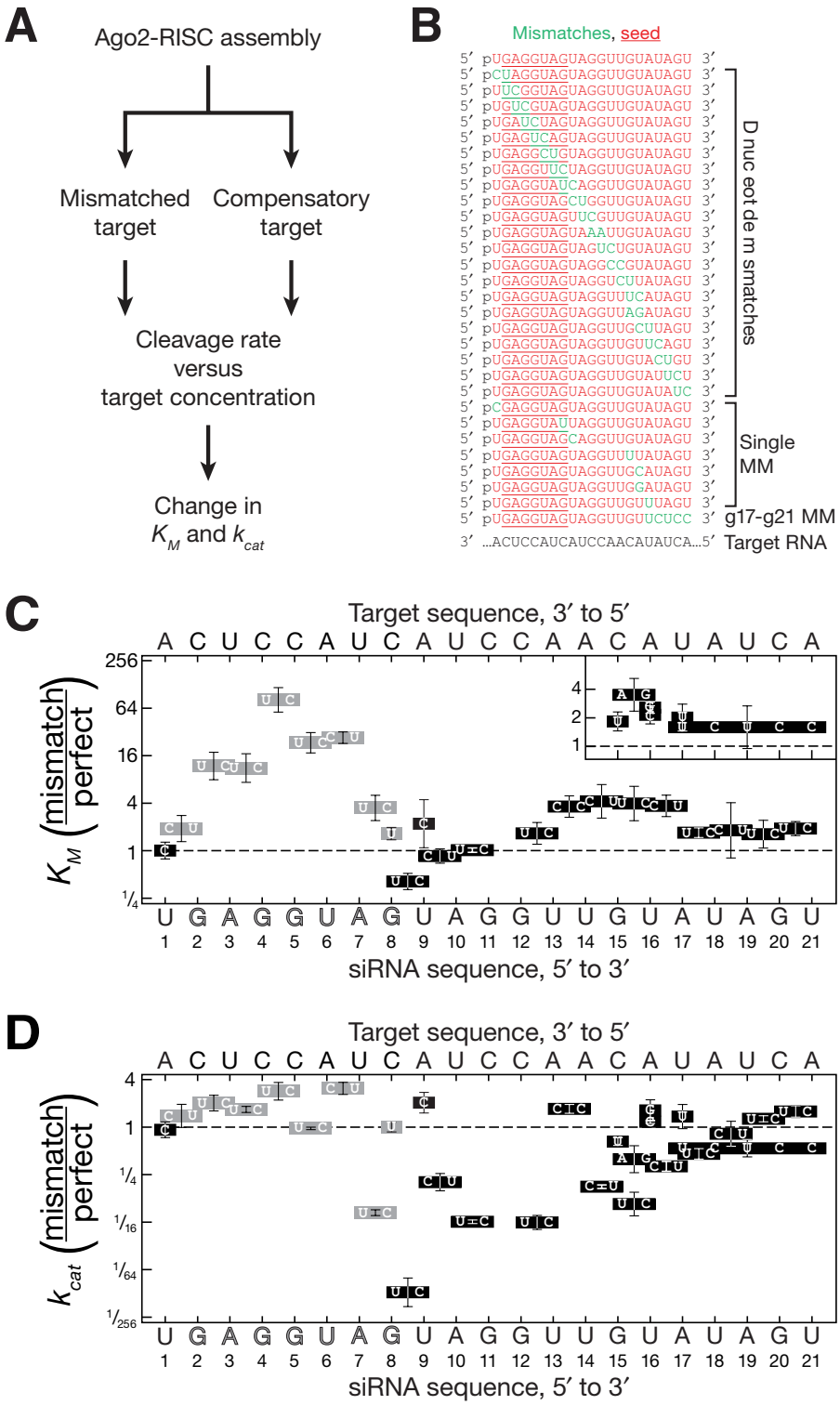

A
Ago2-RISC assembly
Mismatched target
Compensatory target
Cleavage rate versus target concentration
Change in K_M and k_{cat}
B
Mismatches, seed
Double mismatches
Single MM
g17-g21 MM
Target RNA
C
Target sequence, 3′ to 5′
siRNA sequence, 5′ to 3′
K_M (mismatch / perfect)
D
Target sequence, 3′ to 5′
siRNA sequence, 5′ to 3′
k_{cat} (mismatch / perfect)

Figure Legend 3.1. *Drosophila* Ago2 Divides an siRNA into Functional Domains

(A) Kinetic analysis scheme. (B) siRNAs (red) were used with a single target RNA to examine the effect of mismatches (green) on target cleavage. A fully complementary target analyzed in parallel for each siRNA was used to calculate the change (mismatched target/fully complementary target) in K_M (C) and k_{cat} (D). Mismatched bases are indicated in boxes. Gray: mismatches that disrupt seed pairing. Data are mean ± S.D. for ≥3 independent experiments. No cleavage was detected for g11g12 dinucleotide mutations (AA, UU, UA, and UC) or a UCU trinucleotide mutation at g15–g17. See also Figures 3.S1–3.S3 and Tables 3.S1 and 3.S2.

V_{max}/fully complementary V_{max}; similarly, the change in K_M equaled mismatched K_M/fully complementary K_M.

A g1 Mismatch Does Not Alter K_M or k_{cat}

In early studies of fly Ago2, a mismatch between siRNA nucleotide g1 and the corresponding t1 position of its target did not impair target cleavage (Haley and Zamore, 2004). Subsequent studies of archaeal (Ma et al., 2005; Parker et al., 2005) and eubacterial (Wang et al., 2008b; Wang et al., 2008a; Wang et al., 2009) Argonautes revealed that binding of the siRNA 5′ phosphate to Argonaute forces the first nucleotide to be unpaired (Ma et al., 2005; Parker et al., 2005; Wang et al., 2008b). Consistent with these findings, a g1C:t1A mismatch had no detectable effect on the K_M or k_{cat} of fly Ago2 (Figures 3.1C and 3.1D).

The Seed Sequence Behaves Like a Small Helix

Seed sequence mismatches increased K_M (Figure 3.1C). The effect of mismatches on K_M was not constant across the seed (Figure 3.1C); mismatches at the center of the seed (g4g5) increased K_M 82-fold, while the flanking dinucleotide mismatches (g2g3; g3g4; g5g6; and g6g7) increased K_M 11- to 27-fold. These data suggest that base pairs g4:t4 and g5:t5 lie at the center of a six or seven nucleotide RNA helix, because central mismatches should disrupt coaxial stacking more than mismatches closer to the ends of the helix. Dinucleotide and single mismatches at the seed periphery (g1g2; g7g8 and g8)

had the smallest effect, increasing K_M 1.5- to 3.5-fold. The small effect of peripheral seed mismatches helps explain how miRNAs can regulate their targets through some imperfectly seed-matching sites (Ha et al., 1996; Yekta et al., 2004) and through an "offset 6mer seed," in which seed pairing begins at g3 and extends to g8 (Friedman et al., 2009).

Dinucleotide mismatches in the seed were generally accompanied by a small increase in k_{cat}; central mismatches caused the greatest effect (e.g., 2.8-fold for a g4g5:t4t5 mismatch). Thus, seed mismatches decreased target binding but enhanced enzyme turnover, perhaps by accelerating release of the 3′ fragment of the cleaved target (Figure 3.1D).

Central Mismatches Perturb k_{cat}

Target cleavage requires that the center of the siRNA pair with its substrate (Elbashir et al., 2001c; Holen et al., 2002; Amarzguioui et al., 2003; Ding et al., 2003; Haley and Zamore, 2004). Central pairing positions the scissile phosphate of the target near the amino acid side chains that catalyze cleavage (Ma et al., 2005; Parker et al., 2005). Structures of eubacterial Argonaute bound to a DNA guide paired to RNA targets of different lengths suggest that base pairing at the center of the guide moves the three catalytic residues—and, presumably, the Mg^{2+} they bind—closer to the target (Wang et al., 2009). For yeast Argonaute, the rearrangement brings a fourth conserved glutamate into the catalytic site

(Nakanishi et al., 2012). For fly Ago2, mismatches spanning g8 to g12 all reduced target cleavage, albeit to widely varying extents (Figures 3.1D and 3.S1).

While single-nucleotide mismatches at g8 or g9 had little effect on K_M or k_{cat}, a g8g9 dinucleotide mismatch reduced k_{cat} by 93-fold (Figure 3.1D). Dinucleotide mismatches at g8g9 had a similar effect on k_{cat} for a *luciferase*-targeting siRNA (Figures 3.S3A–3.S3C; p-value = 1.7×10^{-8}; two-tailed, unpaired Student's t-test). The effects on k_{cat} of dinucleotide mismatches at g9g10 (5.0-fold reduction) and g10g11 (16-fold reduction) were more modest (Figure 3.1D). We saw no target cleavage for a g11g12 dinucleotide mismatch. (Our assay can detect ~500-fold decrease in k_{cat}.) Mismatches at positions g9g10 or g10g11 did not alter K_M. Our data support the idea that central pairing enables Ago2 to achieve a catalytically competent conformation but contributes little to target binding.

Only a Subset of 3′ Base Pairs Contribute to K_M or k_{cat}

Target pairing 3′ to the center of the small RNA has been proposed to enable Argonaute to achieve a catalytically competent conformation (Haley and Zamore, 2004). Consistent with this view, a dinucleotide mismatch at g12g13 reduced k_{cat} 16-fold, while dinucleotide mismatches at g14g15, g15g16, g16g17, or g17g18, as well as a single mismatch at g15, reduced k_{cat} 1.5- to 9.4-fold (Figure 3.1D). Similarly, a g15g16 dinucleotide mismatch in a *luciferase*-targeting siRNA

decreased k_{cat} 7.6-fold, compared to a fully matched target RNA (Figure 3.S3). A dinucleotide mismatch at g13g14 however, did not decrease k_{cat} (Figure 3.1D). We do not know why this dinucleotide mismatch alone among the six had no detectable effect. We note that this atypical dinucleotide mismatch (CC:AA) lies between a GG:CC (g11g12) dinucleotide and G:C pair (g15). These flanking base pairs may mitigate the helical disruption caused by the intervening pyrimidine:purine dinucleotide mismatch.

The effect on K_M of dinucleotide mismatches from g12 to g17 was qualitatively similar to mismatches in the seed sequence (Figure 3.1C). Pairing to miRNA bases g13–g16 ("3′ supplementary base pairing") is a computational hallmark of a high confidence miRNA-binding site (Brennecke et al., 2005; Grimson et al., 2007; Bartel, 2009; Friedman et al., 2009). We observed a small but significant increased in K_M for dinucleotide mismatches at g13g14 (3.6-fold, *p*-value = 0.022), g14g15 (4.2-fold, *p*-value = 0.017) and g15g16 (3- to 4-fold, *p*-value = 6.6×10^{-3}) and for a single-nucleotide mismatch at g16:t16 (3.4-fold, *p*-value = 4.7×10^{-3}; Figure 3.1C and Table 3.S1). A g15g16 dinucleotide mismatch also increased the K_M of the *luciferase* siRNA by 12-fold (*p*-value = 7.6×10^{-4}; Figure 3.S3B). Notably, the 7 nucleotide seed of this siRNA is predicted to pair more weakly with its target ($\Delta G_{seed(25°C)}$ = −7.7 kcal mol^{-1}) than the seed of the *let-7* siRNA ($\Delta G_{seed(25°C)}$ = −11.2 kcal mol^{-1}). Weaker seed pairing likely makes 3′ supplementary base pairing more important (Brennecke et al., 2005).

Mismatches at the center of the g12–g17 region had the greatest effect on K_M, with a g14g15 dinucleotide mismatch increasing K_M 4.2-fold. Likely the g14g15 base pairs lie at the center of a small RNA helix, much as the g4g5 base pairs do for the seed.

The siRNA 3′ End Contributes Little to K_M or k_{cat}

The g17:t17 base pair marks the end of the 3′ supplementary binding site: a single nucleotide mismatch at g17 and dinucleotide mismatches at g18g19 and g19g20 caused no significant change in K_M or k_{cat}. A g17g18 dinucleotide mismatch decreased k_{cat} by 2.2-fold (p-value = 0.037) while a g17–g21 contiguous mismatch decreased k_{cat} by 1.9-fold (p-value = 0.024), but neither had an effect on K_M. In contrast, a trinucleotide mismatch within the 3′ supplementary region (g15–g17) completely inhibited target cleavage (Figures 3.1C, 3.1D and 3.S1).

Notably, a dinucleotide mismatch at g20g21 caused a modest increase in both K_M (1.9-fold, p-value = 1.7×10^{-3}) and k_{cat} (1.6-fold, p-value = 6.9×10^{-3}), consistent with earlier suggestions that terminal mismatches facilitate product release from plant and animal RISC (Tang et al., 2003; Haley and Zamore, 2004). We conclude that the final four nucleotides of the small RNA guide—the “tail”—form base pairs only after the target RNA is cleaved.

Mismatches that Reduce k_{cat} Reflect a Defect in Catalysis

Mismatches that reduce k_{cat} could reflect a defect in catalysis, product release, or regeneration of RISC to an active state. For these mismatches, we measured the initial rate of target cleavage (v_0) under conditions of enzyme excess. When [E] > [S], v_0 is largely uninfluenced by product release or enzyme regeneration because most RISCs cleave just a single molecule of target.

All mismatches that reduced the multiple turnover cleavage rate also decreased the rate when [E] > [S] (Figure 3.2 and Table 3.S1). Thus, a defect in the catalytic step suffices to explain the reduced k_{cat}. In fact, the effects of mismatches were greater when [E] > [S] than when [E] << [S], suggesting that the deleterious effect of mismatches on the inherent rate of target cleavage is partially offset by a favorable effect of mismatches on steps present only when each RISC catalyzes many successive rounds of target cleavage (Table 3.S1, relative k_{cat}/relative v_0). In other words, mismatches inhibited catalysis but promoted product release or enzyme regeneration. This was most pronounced for mismatches in the seed and 3′ supplementary region (Table 3.S1, relative k_{cat}/relative v_0), favoring the idea that mismatches in these domains promote product release, just as they facilitate the release of miRNA* from pre-Ago1-RISC in flies and humans (Tomari et al., 2007; Kawamata et al., 2009; Yoda et al., 2009).

Figure 3.2

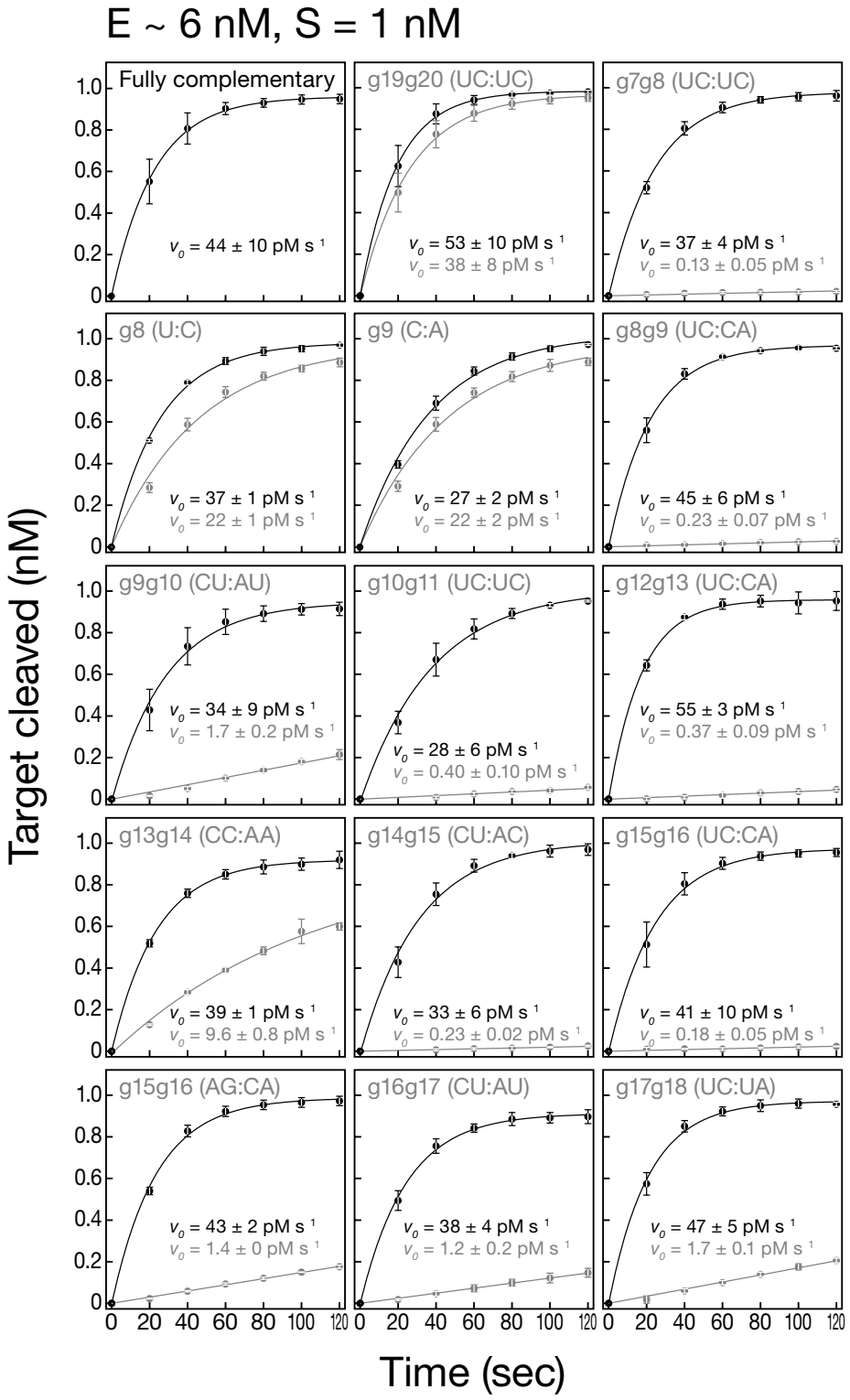

E ~ 6 nM, S = 1 nM
Fully complementary
v_0 = 44 ± 10 pM s $^{-1}$
g19g20 (UC:UC)
v_0 = 53 ± 10 pM s $^{-1}$
v_0 = 38 ± 8 pM s $^{-1}$
g7g8 (UC:UC)
v_0 = 37 ± 4 pM s $^{-1}$
v_0 = 0.13 ± 0.05 pM s $^{-1}$
g8 (U:C)
v_0 = 37 ± 1 pM s $^{-1}$
v_0 = 22 ± 1 pM s $^{-1}$
g9 (C:A)
v_0 = 27 ± 2 pM s $^{-1}$
v_0 = 22 ± 2 pM s $^{-1}$
g8g9 (UC:CA)
v_0 = 45 ± 6 pM s $^{-1}$
v_0 = 0.23 ± 0.07 pM s $^{-1}$
g9g10 (CU:AU)
v_0 = 34 ± 9 pM s $^{-1}$
v_0 = 1.7 ± 0.2 pM s $^{-1}$
g10g11 (UC:UC)
v_0 = 28 ± 6 pM s $^{-1}$
v_0 = 0.40 ± 0.10 pM s $^{-1}$
g12g13 (UC:CA)
v_0 = 55 ± 3 pM s $^{-1}$
v_0 = 0.37 ± 0.09 pM s $^{-1}$
g13g14 (CC:AA)
v_0 = 39 ± 1 pM s $^{-1}$
v_0 = 9.6 ± 0.8 pM s $^{-1}$
g14g15 (CU:AC)
v_0 = 33 ± 6 pM s $^{-1}$
v_0 = 0.23 ± 0.02 pM s $^{-1}$
g15g16 (UC:CA)
v_0 = 41 ± 10 pM s $^{-1}$
v_0 = 0.18 ± 0.05 pM s $^{-1}$
g15g16 (AG:CA)
v_0 = 43 ± 2 pM s $^{-1}$
v_0 = 1.4 ± 0 pM s $^{-1}$
g16g17 (CU:AU)
v_0 = 38 ± 4 pM s $^{-1}$
v_0 = 1.2 ± 0.2 pM s $^{-1}$
g17g18 (UC:UA)
v_0 = 47 ± 5 pM s $^{-1}$
v_0 = 1.7 ± 0.1 pM s $^{-1}$
Target cleaved (nM)
Time (sec)
0
0.2
0.4
0.6
0.8
1.0
0 20 40 60 80 100 120

Figure Legend 3.2. Mismatches that Impair k_{cat} Disrupt Catalysis but Promote Turnover

Target cleavage with [S] < [E]. Initial rates, v_0, for mismatched (gray) and fully complementary targets (black) were determined by fitting the data to a single exponential. Table 3.S1 lists the change in initial rates (mismatched vs. fully complementary). Data are mean ± S.D. for at ≥3 independent experiments.

The Standard Rules for RNA Base Pairing Apply to RISC

Might Ago create a special environment for seed base paring? To test if the standard rules for RNA base pairing apply, we used the change in K_M between mismatched and fully complementary siRNA:target pairs to calculate the free energy cost of mismatches in the seed. We compared this to the cost predicted by nearest neighbor analysis (Xia et al., 1998).

First, we tested whether nearest neighbor values determined in 1 M sodium, pH 7.0 (Schroeder and Turner, 2009), changed in our more physiological conditions, 100 mM potassium, 4 mM magnesium, pH 7.4. Values obtained in our conditions agreed well with the published data (Figures 3.S4A and 3.S4B and Table 3.S3). Second, an increase in K_M may reflect an increase in k_{cat}, because $K_M = (k_{off} + k_{cat})/k_{on}$. For mismatches in the seed and 3′ supplementary regions, we detected no correlated changes between K_M and k_{cat}, justifying our use of the change in K_M as a surrogate for relative K_D.

The free energy cost, $\Delta\Delta G_{25°C}$, calculated from the change in K_M for both seed (r = 0.93, p-value = 4.1×10^{-4}) and 3′ supplementary (r = 0.76, p-value = 2.0×10^{-3}) mismatches correlated well with the values predicted by the nearest neighbor values for RNA base pairing (Figures 3.S4C and 3.S4D). Thus, the relative contributions of each base pair in RISC are similar to those in an RNA:RNA duplex.

Ago2 Reduces the Affinity of a Guide RNA for Its Target

A key obstacle to measuring the binding affinity of Ago2-RISC has been the inability to purify Ago2 bound to a single siRNA guide sequence. We recently developed a simple and efficient method for purifying mature RISC assembled in *Drosophila* embryo lysate or mouse embryonic fibroblast S100 lysate (Figures 3.3A and 3.3B; Flores-Jasso et al., 2013). (Mouse and human AGO2 are 99% identical.) We used nitrocellulose filter binding to measure the binding affinity of both fly Ago2-RISC and mouse AGO2-RISC purified by this procedure (fly, Figures 3.3–3.5; mouse, Figure 3-6). RISC concentration was determined by quantitative Northern hybridization and pre-steady state analysis (Figures 3.S5A–3.S5C). To block cleavage, the target RNA contained a phosphorothioate linkage flanked by 2′-*O*-methyl ribose at positions t10 and t11 (Figures 3.S5D and 3.S5E). Stoichiometric titration showed that 0.81 fly Ago2-RISC and 1.4 mouse AGO2-RISCs bound each molecule of target, consistent with one RISC per target (Figures 3.3C, 3.6A, and 3.6B).

Fly Ago2- and mouse AGO2-RISC bound tightly to a fully complementary RNA (Figures 3.3D and 3.6C). Our K_M data and published Argonaute structures (Wang et al., 2009) suggest that 16–17 base pairs form between the guide and the target RNAs, yet the binding affinity of fly Ago2-RISC (K_D = 3.7 ± 0.9 pM, mean ± S.D.; $\Delta G_{25°C}$ ~−16 kcal mol^{-1}) and mouse AGO2-RISC (K_D = 20 ± 10 pM, mean ± S.D.; $\Delta G_{25°C}$ ~−15 kcal mol^{-1}; see below) for a fully complementary target

Figure 3.3

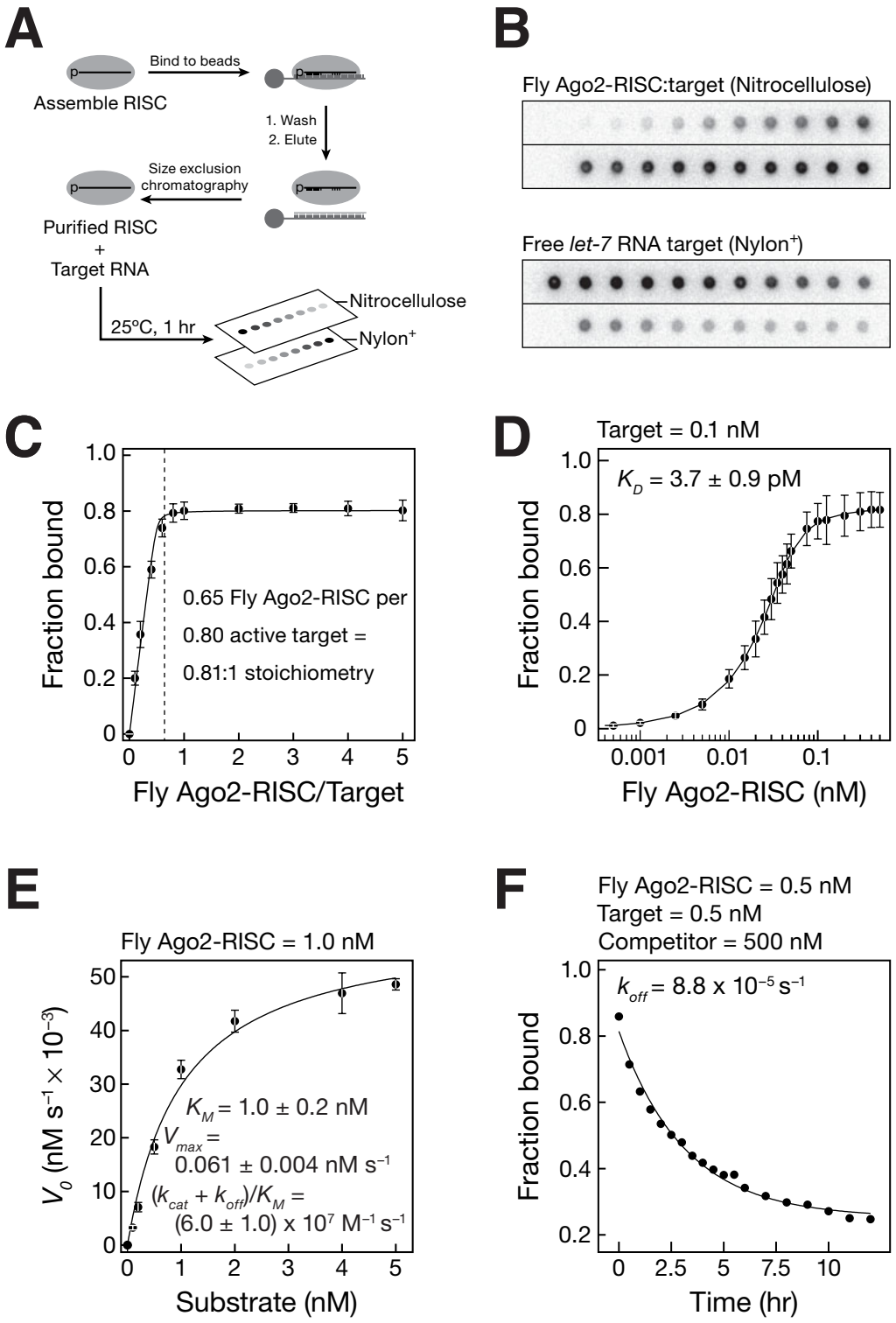

A
Bind to beads
Assemble RISC
1. Wash
2. Elute
Size exclusion chromatography
Purified RISC
+
Target RNA
25°C, 1 hr
Nitrocellulose
Nylon+
B
Fly Ago2-RISC:target (Nitrocellulose)
Free let-7 RNA target (Nylon+)
C
Fraction bound
0.65 Fly Ago2-RISC per
0.80 active target =
0.81:1 stoichiometry
Fly Ago2-RISC/Target
D
Target = 0.1 nM
KD = 3.7 ± 0.9 pM
Fraction bound
Fly Ago2-RISC (nM)
E
Fly Ago2-RISC = 1.0 nM
V0 (nM s−1 × 10−3)
KM = 1.0 ± 0.2 nM
Vmax =
0.061 ± 0.004 nM s−1
(kcat + koff)/KM =
(6.0 ± 1.0) x 10^7 M−1 s−1
Substrate (nM)
F
Fly Ago2-RISC = 0.5 nM
Target = 0.5 nM
Competitor = 500 nM
koff = 8.8 x 10−5 s−1
Fraction bound
Time (hr)

Figure Legend 3.3. Fly Ago2-RISC Binding

RISC was assembled and then purified using a partially complementary, tethered 2′-*O*-methyl oligonucleotide (A). Purified Ago2-RISC was then used in filter-binding assays (B). (C) Stoichiometric binding titration of target RNA with increasing amounts of purified fly Ago2-RISC. Data are mean ± S.D. (D) Equilibrium binding assays. Data are mean ± S.D. for 15 independent experiments using three preparations of fly Ago2-RISC. (E) Kinetics of purified fly Ago2-RISC using a 29 nt fully complementary target RNA. Data are mean ± S.D. for three independent experiments. (F) Dissociation rate for a fully complementary target RNA. See also Figures 3.S4, 3.S5 and Table 3.S3.

was comparable to that of a 10 bp RNA:RNA helix. Thus, Argonaute functions to *weaken* the binding of the 21 nt siRNA to its fully complementary target: without the protein, the siRNA, base paired from positions g2 to g17, is predicted to have a K_D ~3.0×10^{-11} pM ($\Delta G_{25°C}$ = –30.7 kcal mol^{-1}). Argonaute raises the K_D of the 16 bp RNA:RNA hybrid by a factor of >10^{11}.

K_M is not K_D

The K_D measured in our binding assay (3.7 ± 0.9 pM) was ~270-fold smaller than the K_M (1.0 ± 0.2 nM) determined using purified fly Ago2 (Figure 3.3E). By definition, $K_M = (k_{off} + k_{cat})/k_{on}$. When $k_{cat} << k_{off}$, $K_M \sim K_D$. To understand why K_M so dramatically underestimates the affinity of fly Ago2 for a fully complementary target, we measured k_{off} directly (Figure 3.3F). For fly Ago2-RISC, the dissociation rate constant, $k_{off} = 8.8 \times 10^{-5}$ s^{-1}, was much slower than the turnover rate, $k_{cat} = 6.1 \times 10^{-2}$ s^{-1}. Consequently, $K_M \sim k_{cat}/k_{on}$. Hence, for fly Ago2-RISC, K_M is not K_D.

In contrast, the K_D for mouse AGO2 (20 ± 10 pM) was only ~5-fold smaller than the K_M (0.10 ± 0.06 nM), because for mouse the dissociation rate ($k_{off} = 7.7 \times 10^{-4}$ s^{-1}), is comparable to k_{cat} (8.1×10^{-4} s^{-1}; Figures 3.6C and 3.6D). For mouse AGO2-RISC, $K_M \sim K_D$.

We used a competition assay to determine the contributions to binding of the anchor, seed, central, 3′ supplementary and tail regions of the siRNA. For

the fully complementary *let-7* target, this assay gave values similar to those measured in the direct binding assay: 10 ± 1 pM (Figure 3.4 and Table 3.S4) versus 3.7 ± 0.9 pM (Figure 3.3D) for fly and 36 ± 5 pM (Figure 3.6E and Table 3.S4) versus 20 ± 10 pM (Figure 3.6C) for mouse. Binding was specific: a non-complementary *luciferase* RNA target competed ~1600-fold less well for fly Ago2 (Figure 3.4) and ~100-fold less efficiently for mouse AGO2 (Figure 3.6E). Single-stranded sequences flanking the RISC-binding site in a target RNA have been reported to have no effect on the K_M of human AGO2-RISC (Ameres et al., 2007), and we detected no difference in binding between a 28 nt (K_D = 3.9 ± 0.9 pM) and a 21 nt (K_D = 3.6 ± 0.7 pM) competitor for fly Ago2 (Figure 3.4 and Table 3.S4).

The Fly Ago2 Seed Does not Tolerate GU Wobble Pairs

GU wobble pairs between miRNAs and their targets have been reported to be tolerated, and some miRNA target prediction algorithms permit GU wobbles even in the seed (John et al., 2004; Miranda et al., 2006; Kertesz et al., 2007). We measured the effect of seed GU wobble pairs on target binding by fly Ago2-RISC (Figure 3.4). A GU wobble at g4 decreased K_D by 30-fold; two GU wobbles (g2, g8) decrease K_D 40-fold (Figure 3.4). Two GU wobbles at the center of the seed (g4, g5) reduced binding 370-fold, and four GU wobbles (g2, g4, g5, g8) decreased binding 470-fold. We conclude that GU wobbles behave like

Figure 3.4

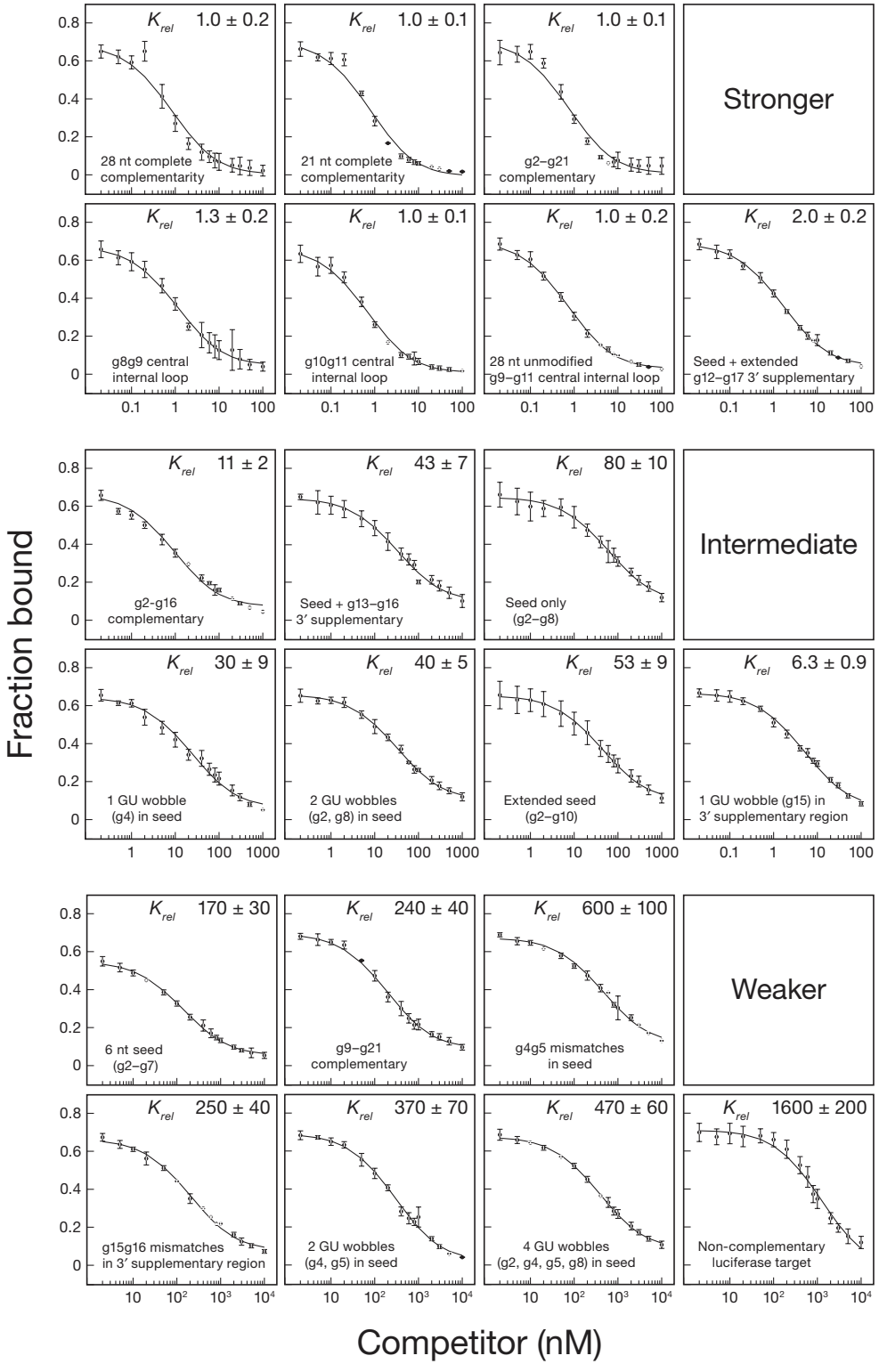

Fraction bound
Competitor (nM)
Stronger
Intermediate
Weaker
K_{rel} 1.0 ± 0.2
28 nt complete complementarity
K_{rel} 1.0 ± 0.1
21 nt complete complementarity
K_{rel} 1.0 ± 0.1
g2–g21 complementary
K_{rel} 1.3 ± 0.2
g8g9 central internal loop
K_{rel} 1.0 ± 0.1
g10g11 central internal loop
K_{rel} 1.0 ± 0.2
28 nt unmodified g9–g11 central internal loop
K_{rel} 2.0 ± 0.2
Seed + extended g12–g17 3′ supplementary
K_{rel} 11 ± 2
g2-g16 complementary
K_{rel} 43 ± 7
Seed + g13–g16 3′ supplementary
K_{rel} 80 ± 10
Seed only (g2–g8)
K_{rel} 30 ± 9
1 GU wobble (g4) in seed
K_{rel} 40 ± 5
2 GU wobbles (g2, g8) in seed
K_{rel} 53 ± 9
Extended seed (g2–g10)
K_{rel} 6.3 ± 0.9
1 GU wobble (g15) in 3′ supplementary region
K_{rel} 170 ± 30
6 nt seed (g2–g7)
K_{rel} 240 ± 40
g9–g21 complementary
K_{rel} 600 ± 100
g4g5 mismatches in seed
K_{rel} 250 ± 40
g15g16 mismatches in 3′ supplementary region
K_{rel} 370 ± 70
2 GU wobbles (g4, g5) in seed
K_{rel} 470 ± 60
4 GU wobbles (g2, g4, g5, g8) in seed
K_{rel} 1600 ± 200
Non-complementary luciferase target

Figure Legend 3.4. Fly Ago2-RISC Equilibrium Competition

The equilibrium dissociation constant of fly Ago2-RISC for the competitor, relative to that of a fully complementary target, is reported as the mean K_{rel} ± S.D. for ≥3 independent experiments. Data are mean ± S.D. See also Table 3.S4.

mismatches and are not tolerated in the seed. Our data explain earlier reports that GU wobbles interfere with Argonaute function (Doench and Sharp, 2004; Brennecke et al., 2005; Grimson et al., 2007) and suggest that GU pairs in the seed should not be allowed by miRNA target prediction algorithms.

Just Two-Thirds of siRNA Nucleotides Contribute to Binding for Fly Ago2

Mismatches at g1, g8g9, or g10g11 had little or no effect on binding. Likewise, a target lacking phosphorothioate and 2′-*O*-methyl modifications but mismatched with the siRNA from positions g9–g11 bound with an affinity similar to that of the fully complementary, modified RNA (K_{rel} = 1.0–1.3; Figure 3.4). A target complementary to only siRNA nucleotides g2–g16 bound just 11-fold less tightly than a target with complete, 21 nt complementarity. In contrast, a g4g5 dinucleotide mismatch in the seed weakened binding 600-fold; a g15g16 mismatch in the 3′ supplementary region reduced binding 250-fold (Figure 3.4).

Thus, more than a third of the nucleotides in an siRNA guide make little or no contribution to target binding. Supporting this view, a target RNA complementary to only g2–g8 (the seed) and g12–g17 (extended 3′ supplementary pairing) bound nearly as tightly as the fully complementary RNA (K_{rel} = 2.0 ± 0.2; Figure 3.4). Yet, a target complementary only to the seed and the 3′ supplementary region (g2–g8; g13–g16) bound 43-fold less tightly than the fully complementary target; a target complementary only to the seed bound

Figure 3.5

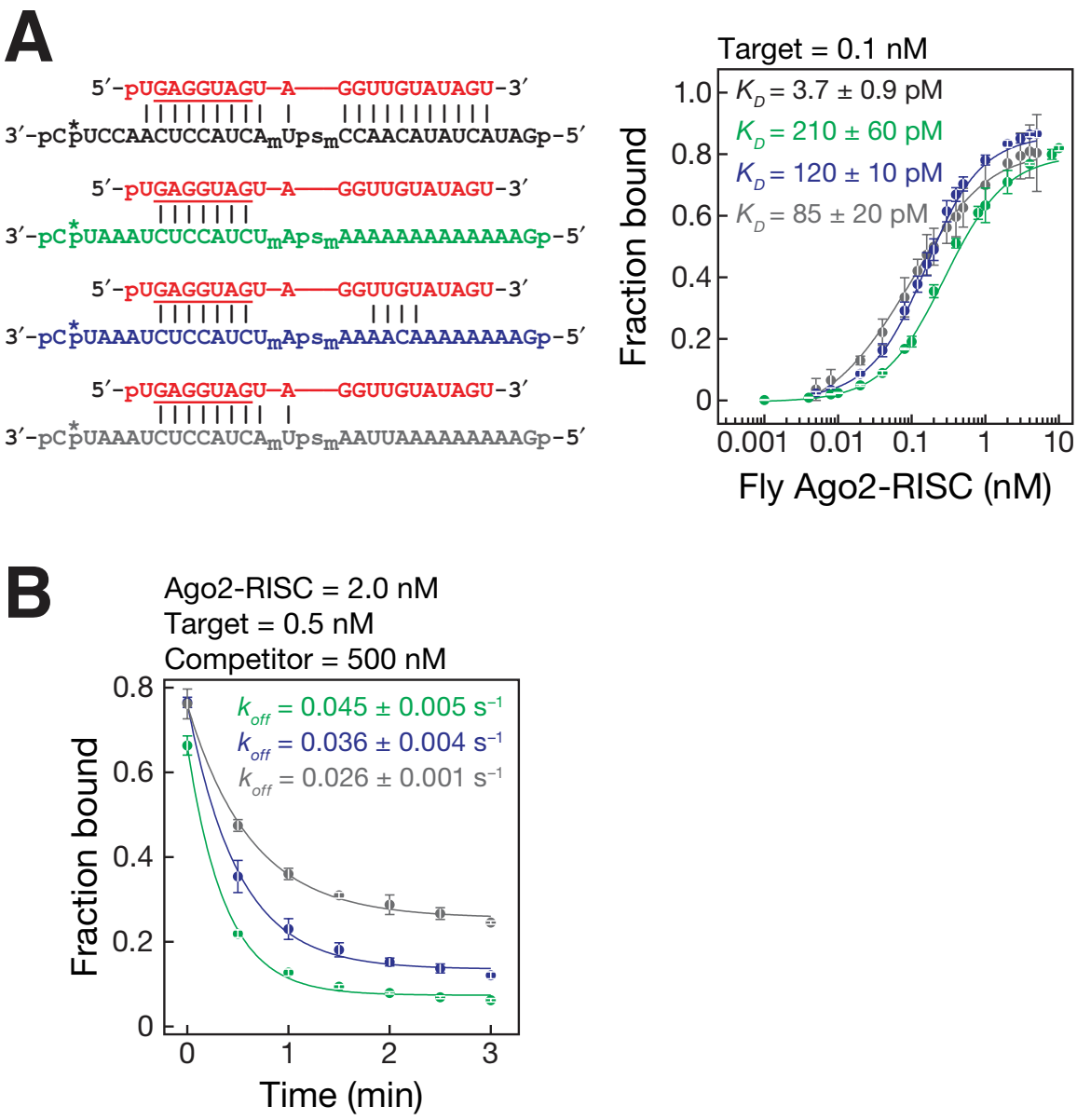

A
5′-pUGAGGUAGU–A——GGUUGUAUAGU-3′
3′-pC*pUCCAACUCCAUCAmUpsmCCAACAUAUCAUAGp-5′
5′-pUGAGGUAGU–A——GGUUGUAUAGU-3′
3′-pC*pUAAAUCUCCAUCUmApsmAAAAAAAAAAAAAGp-5′
5′-pUGAGGUAGU–A——GGUUGUAUAGU-3′
3′-pC*pUAAAUCUCCAUCUmApsmAAAACAAAAAAAAGp-5′
5′-pUGAGGUAGU–A——GGUUGUAUAGU-3′
3′-pC*pUAAAUCUCCAUCAmUpsmAAUUAAAAAAAAAGp-5′
Target = 0.1 nM
K_D = 3.7 ± 0.9 pM
K_D = 210 ± 60 pM
K_D = 120 ± 10 pM
K_D = 85 ± 20 pM
Fraction bound
1.0
0.8
0.6
0.4
0.2
0
0.001
0.01
0.1
1
10
Fly Ago2-RISC (nM)
B
Ago2-RISC = 2.0 nM
Target = 0.5 nM
Competitor = 500 nM
k_{off} = 0.045 ± 0.005 s^{-1}
k_{off} = 0.036 ± 0.004 s^{-1}
k_{off} = 0.026 ± 0.001 s^{-1}
Fraction bound
0.8
0.6
0.4
0.2
0
0
1
2
3
Time (min)

Figure Legend 3.5. Fly Ago2-RISC Binds Seed-Matched Targets at the Rate of Diffusion

(A) Binding and dissociation was analyzed for target RNAs (left) that were complementary (black) to the entire siRNA (red), the seed (green), the seed plus 3′ supplementary region (blue), or positions g2–g10 (gray). Asterisk: ^{32}P-radiolabel; subscript “m”: 2′-*O*-methyl ribose; “ps”: phosphorothioate linkage. (B) Dissociation rates for the RNAs in (A). Data are mean ± S.D. For the dissociation rate curve for the fully complementary RNA, see Figure 3.3F. See also Figure 3.S6.

80 times less tightly. Direct binding measurements yielded essentially the same results as the competition assay (Figure 3.5A). While the seed and 3′ supplementary regions supply much of the energy used by RISC to bind targets, nucleotides adjacent to the 3′ supplementary region also contribute to binding for fly Ago2-RISC.

For Fly Ago2-RISC, a 7 nt Seed Binds Better than a 6mer

Computational analysis in flies suggested that in the absence of 3′ supplementary pairing, 7 nt (g2–g8) but not 6 nt (g2–g7) seed complementarity can distinguish authentic miRNA-binding sites from chance complementarity (Brennecke et al., 2005), unlike in mammals, where both types of seed-matching sites have predictive power (Lewis et al., 2005). Intriguingly, fly Ago2-RISC bound a 6mer seed-matching target 2-fold less tightly than the 7mer seed (Figure 3.4). Since most miRNAs function through Ago1 in flies, it remains to be tested whether Ago1 behaves similarly.

Mouse AGO2 is Optimized for miRNA Regulation, not RNAi

Like fly Ago2, competition assays performed with mouse AGO2-RISC showed that central (g10g11) and terminal mismatches (g20g21) had no detectable effect on binding, whereas g4g5 seed mismatches reduced binding 40-fold (Figure 3.6E). Surprisingly, g15g16 mismatches did not impair binding for mouse AGO2-RISC (K_{rel} = 1.4 ± 0.6; Figure 3.6E). Moreover, direct binding assays found

Figure 3.6

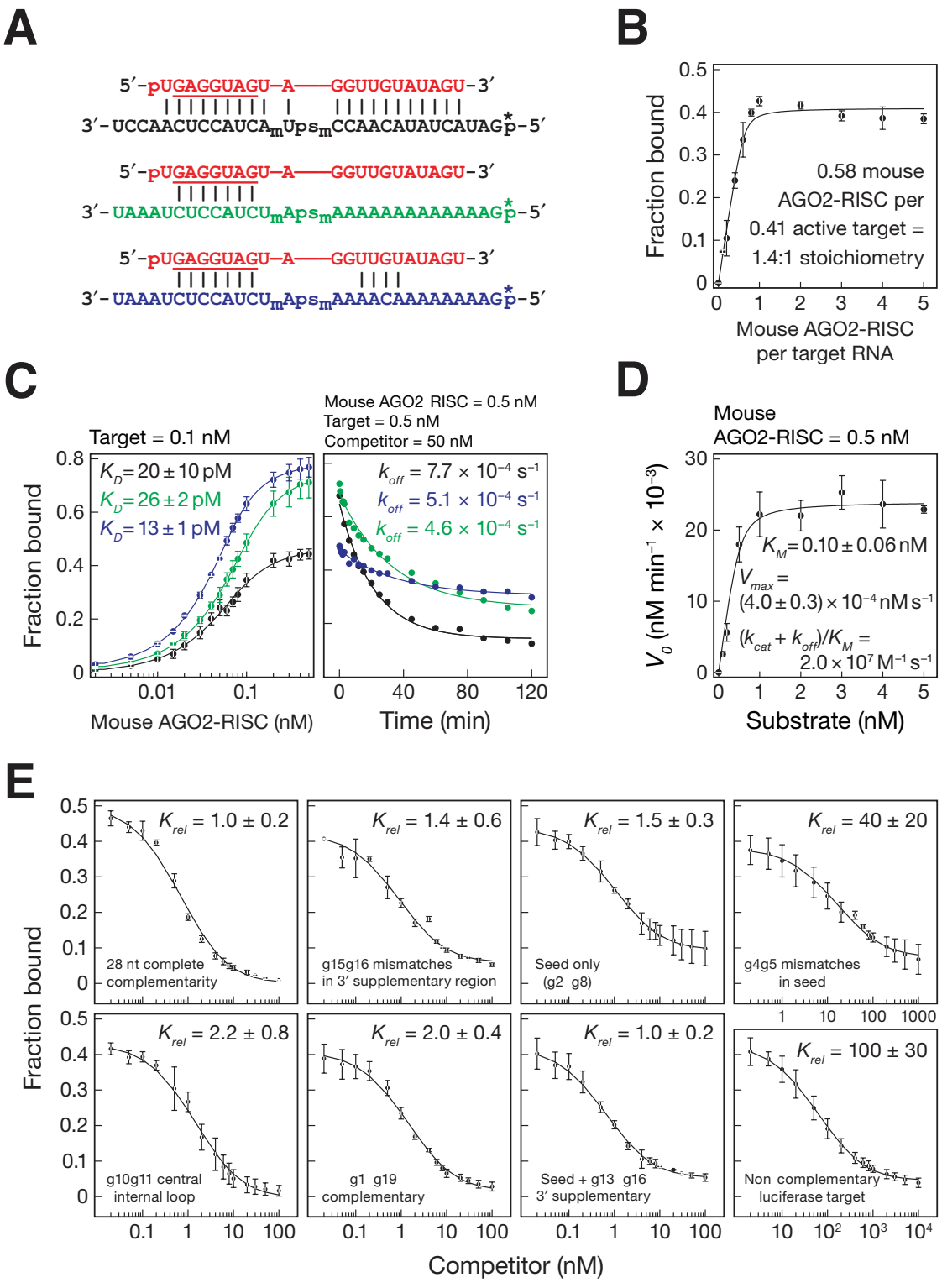

A
5′–pUGAGGUAGU–A——GGUUGUAUAGU–3′
3′–UCCAACUCCAUCA$_m$UPS$_m$CCAACAUAUCAUAGp*–5′
5′–pUGAGGUAGU–A——GGUUGUAUAGU–3′
3′–UAAAUCUCCAUCU$_m$APS$_m$AAAAAAAAAAAAAGp*–5′
5′–pUGAGGUAGU–A——GGUUGUAUAGU–3′
3′–UAAAUCUCCAUCU$_m$APS$_m$AAAACAAAAAAAAGp*–5′
B
Fraction bound
0.58 mouse AGO2-RISC per 0.41 active target = 1.4:1 stoichiometry
Mouse AGO2-RISC per target RNA
C
Target = 0.1 nM
Mouse AGO2 RISC = 0.5 nM
Target = 0.5 nM
Competitor = 50 nM
K_D = 20 ± 10 pM
K_D = 26 ± 2 pM
K_D = 13 ± 1 pM
k_{off} = 7.7 × 10^{-4} s^{-1}
k_{off} = 5.1 × 10^{-4} s^{-1}
k_{off} = 4.6 × 10^{-4} s^{-1}
Fraction bound
Mouse AGO2-RISC (nM)
Time (min)
D
Mouse AGO2-RISC = 0.5 nM
K_M = 0.10 ± 0.06 nM
V_{max} = (4.0 ± 0.3) × 10^{-4} nM s^{-1}
(k_{cat} + k_{off})/K_M = 2.0 × 10^7 M^{-1} s^{-1}
V_0 (nM min^{-1} × 10^{-3})
Substrate (nM)
E
K_{rel} = 1.0 ± 0.2
28 nt complete complementarity
K_{rel} = 1.4 ± 0.6
g15g16 mismatches in 3′ supplementary region
K_{rel} = 1.5 ± 0.3
Seed only (g2 g8)
K_{rel} = 40 ± 20
g4g5 mismatches in seed
K_{rel} = 2.2 ± 0.8
g10g11 central internal loop
K_{rel} = 2.0 ± 0.4
g1 g19 complementary
K_{rel} = 1.0 ± 0.2
Seed + g13 g16 3′ supplementary
K_{rel} = 100 ± 30
Non complementary luciferase target
Fraction bound
Competitor (nM)

Figure Legend 3.6. Mouse AGO2-RISC is Specialized for miRNA Regulation

(A) Binding and dissociation analyses for target RNAs that were complementary (black) to the entire siRNA (red), the seed (green) or the seed plus 3′ supplementary region (blue). (B) Stoichiometric binding titration with increasing amounts of mouse AGO2-RISC. (C) Equilibrium binding (left) and dissociation assays (right). Data are mean ± S.D. for ≥3 independent experiments. (D) Kinetics of purified mouse AGO2-RISC using a 28 nt fully complementary target. Data are mean ± S.D. for three independent experiments fitted to the quadratic equation for tight binding. (E) The equilibrium dissociation constant of mouse AGO2-RISC for the competitor, relative to that of a fully complementary target, is reported as the mean K_{rel} ± S.D. for ≥3 independent experiments.

no substantive difference in affinity between a seed-matching (K_D = 26 ± 2 pM) and a fully complementary target (20 ± 10 pM; Figure 3.6C). We did observe a small but significant (p-value = 3.2×10^{-4}) increase in affinity for a target with seed and 3′ supplementary pairing (K_D = 13 ± 1 pM), compared to the seed alone. The modest contribution of the 3′ supplementary region to target binding helps explain why in mammals less than 5% of evolutionarily conserved, predicted miRNA-binding sites include conserved 3′ pairing (Friedman et al., 2009). We conclude that seed complementarity and, to a far lesser extent, 3′ supplementary base pairing, provide all the binding energy tethering mouse AGO2-RISC to its targets. Our data suggest that evolution has optimized mammalian AGO2 for miRNA-based regulation. In contrast, fly Ago2 binds far more tightly to fully complementary targets than to those matching only the seed, as might be expected for an enzyme responsible for binding and destroying viral and transposon transcripts.

Essentially Every Target that is Fully Paired to Fly Ago2-RISC is Cleaved

To understand the molecular basis for the difference between mouse and fly Ago2-RISC, we measured the rate of dissociation of *let-7*-programmed fly Ago2-RISC for several prototypical RNA targets (Figure 3.5A).

Fly Ago2-RISC dissociated slowly from a fully complementary target: k_{off} = 8.8×10^{-5} s^{-1}, corresponding to a half-life ($t_{½}$) ~2.2 hr (Figure 3.3F). Given that

k_{cat} for *let-7*-programmed fly Ago2-RISC was 6.1×10^{-2} s^{-1} ($t_{½}$ ~11 s), essentially every fly Ago2-RISC that binds a target will slice it rather than dissociate from the uncleaved RNA (Figure 3.3E).

RISC dissociated far more rapidly from targets paired to the seed sequence (g2–g8) or the seed plus the 3′ supplementary region (g13–g16): k_{off} = 4.5×10^{-2} s^{-1} ($t_{½}$ ~15 s) for seed-matched and $k_{off} = 3.6 \times 10^{-2}$ s^{-1} ($t_{½}$ ~19 s) for seed plus 3′ supplementary pairing (Figure 3.5B). Such rapid dissociation from partially paired targets may minimize titration of RISC by seed-matching off-targets. Intriguingly, fly Ago2-RISC dissociated more slowly from a target that paired with an extended seed-match (g2–g10; $k_{off} = 2.6 \times 10^{-2}$ s^{-1}, $t_{½}$ ~27 s) than from a target complementary to both the seed and 3′ supplementary region (Figure 3.5B).

Mouse AGO2-RISC often Dissociates before It Cleaves

Mouse AGO2-RISC dissociated ~90-fold more slowly from a seed-matched target ($k_{off} = 5.1 \times 10^{-4}$ s^{-1}; $t_{½}$ ~23 min) than fly Ago2. Moreover, the mouse AGO2 dissociation rate constants for targets matching the seed, seed plus 3′ supplementary region ($k_{off} = 4.6 \times 10^{-4}$ s^{-1}; $t_{½}$ ~25 min), and the entire RNA guide ($k_{off} = 7.7 \times 10^{-4}$ s^{-1}; $t_{½}$ ~15 min) were quite similar (Figure 3.6C), consistent with their similar K_D values. Given that the k_{cat} for purified mouse AGO2-RISC was 8.1

$\times 10^{-4}$ s^{-1} ($t_{½}$ ~14 min; Figure 3.6D), a fully complementary target is as likely to dissociate as to be cleaved.

Our data also suggest that in both flies and mammals, the typical miRNA:Argonaute complex is in rapid equilibrium between the target-bound and unbound states, explaining why RNA-binding proteins can compete with miRNAs for overlapping binding sites (Bhattacharyya et al., 2006; Huang et al., 2007; Kedde et al., 2007; Elcheva et al., 2009; Takeda et al., 2009; Goswami et al., 2010; Jafarifar et al., 2011; Toledano et al., 2012).

miRNAs in RISC Find Their Targets at Rates That Approach That of Diffusion

We used our experimentally determined K_D and k_{off} to calculate k_{on} (= k_{off} /K_D), the bimolecular association rate constant for RISC binding its target. For both fly and mouse AGO2-RISC, k_{on} for targets matching only the seed and the seed plus the 3′ supplementary region were similar: k_{on} (seed) = 2.1×10^8 M^{-1} s^{-1} and k_{on} (seed plus 3′ supplementary) = 3.1×10^8 M^{-1} s^{-1} for fly Ago2; k_{on} (seed) = 2.0×10^7 M^{-1} s^{-1} and k_{on} (seed plus 3′ supplementary) = 3.6×10^7 M^{-1} s^{-1} for mouse AGO2. These rates suggest that miRNA-programmed Argonautes find their target RNAs near the limits of macromolecular diffusion (Hammes and Schimmel, 1970; Berg and von Hippel, 1985).

For fly Ago2-RISC, a dinucleotide mismatch that disrupts seed pairing (g4g5) reduced k_{on} (= k_{cat} + k_{off}/K_M) by ~30-fold (Table 3.S1) and increased k_{off} by

~40-fold (k_{on} = 7.0 × 10^5 M^{-1} s^{-1}; k_{off} = 3.6 ± 0.9 × 10^{-3} s^{-1}; Figures 3.S6A and 3.S6B). The K_D value (5.2 nM) calculated from these k_{on} and k_{off} values agrees well with the K_D (2.3 ± 0.6 nM) measured by equilibrium competition experiments (Figure 3.4 and Table 3.S4). Our data provide strong support for the idea that in flies seed pairing must precede the formation of base pairs between the target and the 3′ half of the siRNA.

Base Pairing Beyond the Seed Proceeds at a Slower Rate for Fly Ago2-RISC

In contrast, the calculated k_{on} (k_{on} = 2.4 × 10^7 M^{-1} s^{-1}) for fly Ago2-RISC binding a fully complementary target is >10 times slower than for a seed-matching target. For fully complementary targets in flies, (k_{cat} + k_{off})/K_M approximates k_{on} and should reflect the rate at which RISC attains a catalytically active conformation, i.e., pairing from g2 to g17. Calculating k_{on} from enzyme kinetics yields a similar value: 6.0 ×10^7 M^{-1} s^{-1} (Figure 3.3E). Taken together, our data suggest that seed pairing occurs more rapidly than the subsequent propagation of base pairs across the center of the siRNA and through the 3′ supplementary region.

We imagine that complete base pairing to fully complementary targets requires conformational rearrangement of the siRNA within fly Ago2-RISC. Structural studies of eubacterial and eukaryotic Argonautes support this idea. They reveal a conformational rearrangement of the protein near the center of the

guide when it is extensively paired to its target and release of its 3′ end from the PAZ domain of Argonaute (Wang et al., 2009; Boland et al., 2011). In this view, cleavage of a target by fly Ago2-RISC is not limited by the search for a complementary sequence among the RNAs in a cell, but rather by the rate at which the siRNA, bound to Argonaute, can form an additional ~8 base pairs beyond the seed.

In contrast, mouse AGO2-RISC associates with a fully paired target at a rate (k_{on} = 3.6 × 10^7 M^{-1} s^{-1}) indistinguishable from seed (k_{on} = 2.0 × 10^7 M^{-1} s^{-1}) or seed plus 3′ supplementary pairing (k_{on} = 3.6 × 10^7 M^{-1} s^{-1}). The association rate derived from enzyme kinetics corroborates these measurements: k_{on} = (k_{off} + k_{cat})/ K_M = 2.0 × 10^7 M^{-1} s^{-1} (Figure 3.6D). Thus, fly Ago2 binds rapidly through its seed, then completes pairing of its 3′ bases more slowly, while mouse AGO2 binds seed matching targets more slowly, so that the rate of propagating the helix to the 3′ half of the guide does not limit the rate of target cleavage.

Centrally Bulged Sites

Centrally bulged siRNAs are often used to model miRNA function in cultured mammalian cells (Zeng et al., 2002; Doench and Sharp, 2004; Broderick et al., 2011). This approach typically uses an asymmetric 3 × 2 internal loop at g9–g11. While we have not measured the binding of 3 × 2 asymmetric internal loops, our results with 3 × 3 symmetric internal loops are likely to be similar. Compared to

naturally occurring, seed-match sites, centrally bulged sites bind RISC 80-fold more tightly for fly Ago2 (Figure 3.4), suggesting an explanation why centrally bulged sites require a lower concentration of RISC to mediate reporter repression (Broderick et al., 2011).

While a target with g9–g11 mismatch bound fly Ago2-RISC as tightly as a fully complementary RNA (K_D = 3.0 ± 1.0 pM; Figure 3.4 and Table 3.S4), the mechanism of binding is clearly different from the fully paired target: its measured k_{off} value of 1.1 ± 0.1 × 10^{-3} s^{-1} and calculated k_{on} value of 3.1 × 10^{8} M^{-1} s^{-1} are ≥ 5–13-fold faster than the fully complementary target (Figures 3.S6A and 3.S6B). We propose that the g9–g11 mismatch bypasses an energetically unfavorable rearrangement that occurs for a fully complementary target RNA. Interestingly, the crystal structure of eubacterial Argonaute shows that both ends of the guide remain anchored in the presence of a g10g11 mismatch (Wang et al., 2008a).

DISCUSSION

Argonaute divides a small RNA guide into anchor, seed, central, 3′ supplementary, and tail functional domains (Figure 3.7). Nucleotides in the anchor (g1) and tail (g18–g21) facilitate Argonaute loading and help secure the siRNA or miRNA guide to Argonaute after the passenger or miRNA* strand has been removed. But these terminal domains are unlikely to base pair with a target RNA, even when pairing is predicted by their sequences. In contrast, central base pairing (g9–g12) between the guide and target is required for efficient target cleavage. Mismatches in this central region prevent RISC from attaining a catalytically competent conformation. For fly Ago2-RISC, achieving this conformation takes more time than seed pairing alone. Our data show that nearly every fly Ago2-RISC that reaches this conformation cleaves its RNA target rather than releasing it. For mouse AGO2-RISC, a slow catalytic rate often allows the target to escape before being sliced.

In contrast, most miRNA:Argonaute complexes rapidly bind to and dissociate from their RNA targets via their seed. Even when RISC binds a target through both its seed and 3′ supplementary regions, it dissociates nearly as rapidly as for seed-only binding. Thus, the properties of RISC are essentially the same for both the typical seed-only and the less common seed plus 3′ supplementary pairing targets. That the rates of association and dissociation are so similar for these two binding modes suggests that pairing between a target

Figure 3.7

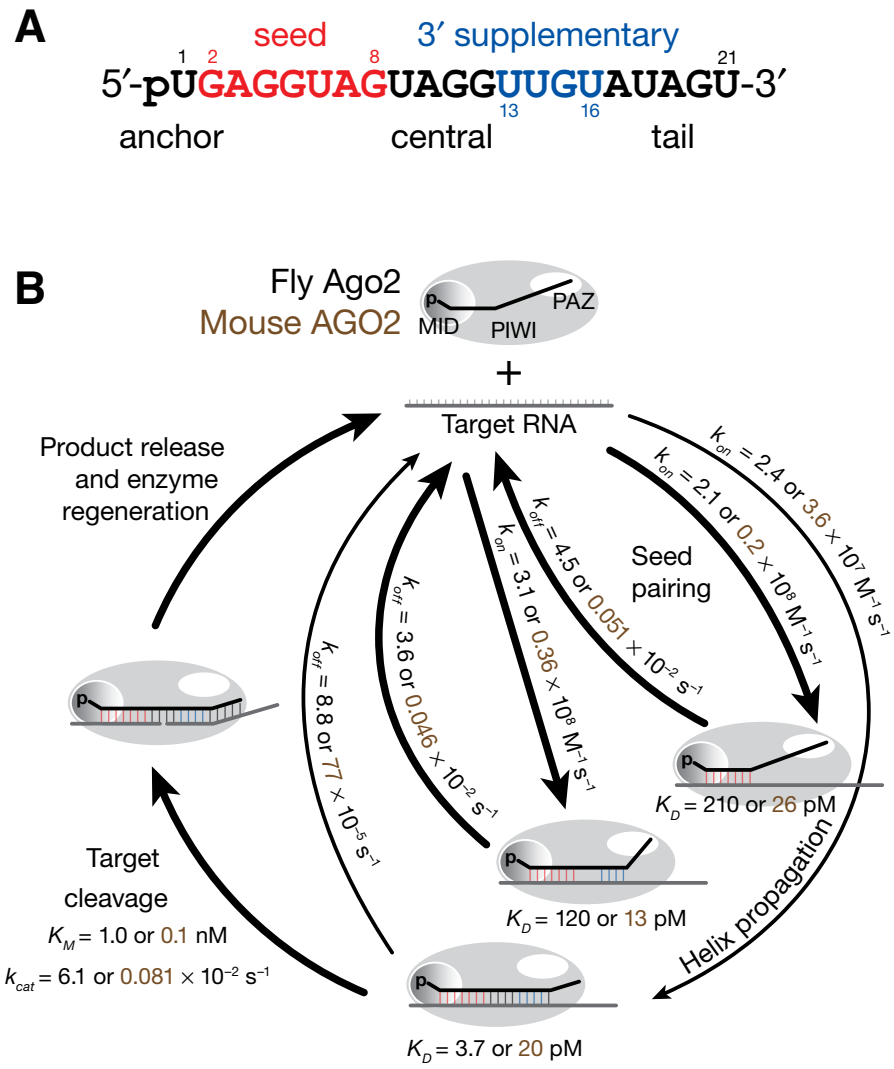

A
seed
3′ supplementary
1 2 8 13 16 21
5′-pUGAGGUAGUAGGUUGUAUAGU-3′
anchor
central
tail
B
Fly Ago2
Mouse AGO2
p
PAZ
MID
PIWI
+
Target RNA
Product release and enzyme regeneration
k_{on} = 2.4 or 3.6 × 10⁷ M⁻¹ s⁻¹
k_{on} = 2.1 or 0.2 × 10⁸ M⁻¹ s⁻¹
Seed pairing
k_{off} = 4.5 or 0.051 × 10⁻² s⁻¹
k_{on} = 3.1 or 0.36 × 10⁸ M⁻¹ s⁻¹
k_{off} = 3.6 or 0.046 × 10⁻² s⁻¹
k_{off} = 8.8 or 77 × 10⁻⁵ s⁻¹
K_D = 210 or 26 pM
K_D = 120 or 13 pM
K_D = 3.7 or 20 pM
Helix propagation
Target cleavage
K_M = 1.0 or 0.1 nM
k_{cat} = 6.1 or 0.081 × 10⁻² s⁻¹

Figure Legend 3.7. Model for RISC Function

(A) Loading of an siRNA or miRNA into Argonaute creates distinct functional domains in the RNA guide. (B) A model for RISC binding and cleavage of target RNA. See also Figure 3.S7.

and the 3′ supplementary region of a miRNA does not require winding the target RNA around the guide, side-stepping the topological problem that must be solved for siRNAs to direct RISC to cleave a target.

The finding that miRNAs use so little of their sequence to identify their regulatory targets surprised the biological community (Wightman et al., 1993; Reinhart et al., 2000; Lai, 2002). Our data show that miRNA-programmed RISC binds with a strength and binding site size similar to those of high affinity RNA-binding proteins (e.g., Hall and Stump, 1992; Zamore et al., 1999; Zearfoss et al., 2011; Wright et al., 2011). It is siRNA-programmed RISC whose behavior should surprise us: it binds highly complementary targets far *less* tightly than a comparable antisense RNA, because Argonaute reduces the contribution of most of its nucleotides to target binding.

What do the physical properties of RISC teach us about its cellular function? miRNAs and siRNAs are typically present in cells at dramatically different concentrations. For example, in flies in which the *white* gene is silenced by RNAi, the abundance of all antisense *white* siRNAs combined is less than that of any of 29 most abundant miRNAs (Ghildiyal et al., 2008). Previously, the ability of siRNAs to function at low abundance has been ascribed to the catalytic nature of RNAi (Fire et al., 1998; Montgomery et al., 1998; Hutvagner and Zamore, 2002). To achieve a concentration 10-fold greater than the K_D for siRNA-like binding (3.7 pM for fly Ago2-RISC) would require only ~5 molecules

of RISC in ovarian terminal filament cells (~200 μm3; Schneider, 1972) and ~11 molecules in a cultured S2 cell (~500 μm3; Sarikaya et al., 2012). Thus, even for Argonaute proteins with no endonuclease activity, small numbers of molecules of RISC can repress highly complementary targets; endonuclease activity is only needed when a small amount of RISC must repress a larger amount of target. The combination of high affinity and catalytic turnover helps explain why the siRNA-directed RNAi pathway provides an effective defense against viral infection in plants and invertebrate animals (Hamilton and Baulcombe, 1999; Wilkins et al., 2005; Galiana-Arnoux et al., 2006; Wang et al., 2006b).

Animal miRNAs nearly always repress their targets by binding rather than endonucleolytic cleavage. This explains why animal cells express miRNAs at such high levels. Recent data suggest that only the most abundant cellular miRNAs mediate target repression (Mullokandov et al., 2012). Our data provide a biochemical explanation for this observation.

Consider two abundant miRNAs in a cultured HeLa cell (~5,000 μm3; Cohen and Studzinski, 1967; Milo et al., 2010): miR-21 (4nM; Lim et al., 2003) and the *let-7* miRNA family, nine highly related miRNAs sharing a common seed sequence (~3 nM; Cole et al., 2009). Both miRNAs are present at a concentration greater than the K_D we measured for seed matched targets for fly (~210 pM) or mouse (~27 pM) Ago2-RISC. Assuming a mean target mRNA abundance of 10 molecules per cell and 50 different mRNA targets per miRNA,

miR-21 and *let-7* each regulate ~500 (170 pM) total target mRNA molecules per HeLa cell (Friedman et al., 2009). Under these conditions, nearly every miR-21 or *let-7* target mRNA (~95–99%) with an accessible seed match will be bound by the complementary miRNA-programmed RISC (Figure 3.S7).

Target repression by miRNAs can be reduced by the presence of competitor RNAs containing miRNA binding sites that titrate miRNA-RISC away from the mRNAs it regulates (Arvey et al., 2010; Garcia et al., 2011; Mukherji et al., 2011). The fundamental properties of RISC make specific predictions about how the activity of specific miRNAs can be inhibited by the expression of these competitor transcripts. The effect of such competitor RNAs reflects the concentration of both the miRNA and miRNA-binding sites (Ebert and Sharp, 2012), as well as the affinity of miRNA-RISC for those sites. For abundant miRNAs such as miR-21 or the *let-7* family, the expression of competitor RNAs containing miRNA binding sites—even highly complementary binding sites—will have little impact on the regulation of their target genes in flies or mammals. Doubling the expression of mRNAs repressed by miR-21, for example, would require ~7.8 nM seed only competitor and ~4.0 nM fully paired competitor for fly Ago2-RISC. For mouse AGO2-RISC, it would still require ~7.7 nM seed only competitor and ~7.2 nM of the fully paired competitor. Taken together, this translates to ~22,400 copies of seed only competitor and ~12,000–21,700 copies of fully paired competitor (Figure 3.S7). If the competitor contained one

miRNA-binding site, it would comprise 12–50% of all the mRNA in the cell (Islam et al., 2011).

In contrast, doubling the expression of the mRNA targets for an intermediate (mir-93; ~140 pM) or a low abundance miRNA (mir-24; 7.3 pM) would require just 600–800 additional seed-matching sites (Figure 3.S7). For mir-93 whose abundance confers the ability to bind to ~60% of all potential targets, the competitor must be as abundant as the sum of all the target mRNAs (~500 copies). Low abundance miRNAs like mir-24 are unlikely to contribute much biologically meaningful regulation, since they are present at a concentration less than their K_D for seed-matching targets in both flies and mammals: < 4% of miR-24 targets are expected to be bound by the miRNA at any given time. Using the conservative assumption that every bound miRNA-RISC completely represses an mRNA target, miR-24 is predicted to reduce the expression of the average seed-matched target by <4% (Figure 3.S7).

Thus, the proposal that “competing endogenous RNAs” (“ceRNAs”) sequester miRNAs, de-repressing the authentic targets of that miRNA (Salmena et al., 2011), applies only to a small subset of miRNAs whose cellular concentration and target abundance meet a narrow range of values. The miRNAs with the largest impact on gene expression—the most abundant miRNAs—are not predicted to be regulatable by endogenous, transcribed seed-matched competitor transcripts. Consistent with this view, viral and

experimental inhibition of specific miRNA function by transcribed RNA requires the use of extensively complementary miRNA binding sites that recruit a cellular pathway that actively degrades the targeted miRNA (Ebert et al., 2007; Loya et al., 2009; Ameres et al., 2010; Xie et al., 2012). Absent this target-directed, catalytic destruction of miRNAs, RNAs of ordinary abundance are unlikely to compete with mRNAs for binding abundant, biologically functional miRNAs.

EXPERIMENTAL PROCEDURES

General Methods

Target cleavage reactions were performed as described (Haley et al., 2003; Haley and Zamore, 2004) except using 4 mM Mg^{2+}. Cleavage targets (Table 3.S2) were prepared by in vitro transcription and capping (Haley et al., 2003). For binding, synthetic RNAs were 3′ ^{32}P-radiolabeled.

Binding, Competition, and Dissociation Assays

Ago2-RISC was assembled with *let-7* siRNA in 0–2 hr embryo lysate or S100 from immortalized *Ago2*$^{-/-}$ MEFs expressing mouse AGO2 (O'Carroll et al., 2007). Binding reactions were at 25°C for 1 hr; protein-RNA complexes were captured on nitrocellulose and unbound RNA on Nylon membranes under vacuum and washed with ice-cold buffer. Competition reactions were at 25°C for 1 hr (mouse) or 6 hr (fly). For k_{off}, Ago2-RISC was incubated with ^{32}P-radiolabeled RNA target for 1 hr, then competitor RNA was added and dissociation measured.

Kinetics

Target cleavage reactions were performed as described (Haley et al., 2003; Haley and Zamore, 2004) except using 4 mM Mg^{2+}. For each siRNA duplex, guide position 1 (g1) was unpaired from the corresponding base in the passenger strand to ensure efficient loading of the guide strand into Ago2 (Schwarz et al., 2003; Khvorova et al., 2003; Forstemann et al., 2007; Tomari et

al., 2007). *let-7* is not present in the *Drosophila* embryo lysate used in our studies. For multiple-turnover reactions, 20 nM siRNA duplex was incubated with 0–2 hr *Drosophila* embryo lysate at 25°C for 3 min to assemble fly Ago2-RISC; for single-turnover reactions, 50 nM siRNA duplex was incubated with embryo lysate for 90 min to assemble more fly Ago2-RISC (Haley and Zamore, 2004). Subsequently, Ago2 assembly was inactivated by cooling the reaction to 4°C for 3 min and adding 1.0 mM (f.c.) *N*-ethylmaleimide (NEM). After 10 min, 1.2 mM (f.c.) DTT was added to quench unreacted NEM. RISC concentration was adjusted by dilution in NEM-treated embryo lysate (Haley et al., 2003). Control experiments demonstrated that target cleavage required Ago2. Single-turnover reactions employed 0.1, 0.5 or 1 nM (f.c.) target RNA. Pre-steady state experiments were performed in 100 mM potassium acetate, 18 mM HEPES (pH 7.4), 3 mM magnesium acetate, 5 mM DTT, 0.01% (v/v) IGEPAL CA-630, 0.01 mg/ml baker's yeast tRNA using 100 nM substrate (f.c) at 25°C. For each time point, an aliquot of the reaction was quenched in 200 mM Tris-HCl (pH 7.5), 25 mM EDTA (pH 8.0), 300 mM NaCl, 2% (w/v) sodium dodecyl sulfate, 2 µg/µl Proteinase K and 0.2 µg/µl glycogen carrier. All samples were resolved by urea-denaturing polyacrylamide gel electrophoresis (Haley et al., 2003). Gels were dried, exposed to image plates, and then scanned and analyzed using an FLA-5000 or FLA9000 (GE Healthcare Bioscience, Pittsburgh, *PA*) and Image Gauge 4.22 software (Fujifilm, Tokyo).

siRNAs and target RNAs

Synthetic siRNA duplexes (Dharmacon, ThermoFisher Scientific, Lafayette, CO; Sigma-Aldrich Corp, St. Louis, MO) were deprotected, phosphorylated and gel purified (Table 3.S2A). RNA targets (Table S2B) were prepared by in vitro capping and transcription (Haley et al., 2003). Briefly, DNA transcription templates were generated by PCR (Table 3.S2C) from pGL2-Control vector (Promega, Madison, WI). RNA was transcribed with T7 RNA polymerase and gel purified. Next, 150 pmol RNA was incubated for 1.5 hr at 37°C in a 40 µl reaction containing 1.25 µM ^{32}P α-GTP (Perkin Elmer, Waltham, MA), 25 U guanylyl transferase (Epicentre, Madison, WI), 40 U RNase inhibitor (Promega, Madison, WI), 125 µM *S*-adenosylmethionine, 50 mM Tris-HCl (pH 8.0), 6 mM KCl and 1.25 mM $MgCl_2$. Next, 15 U guanylyl transferase, 1.2 mM GTP, a concentration in excess of the RNA substrate and greater than the K_M of the enzyme (Myette and Niles, 1996), 120 µM *S*-adenosylmethionine were added and incubation continued for 1.5 hr. Capped RNA was gel purified, and small aliquots were frozen in liquid nitrogen and stored at −80°C. Target RNAs were radiolabeled with 0.1 mM [5′-^{32}P] cytidine-3′,5′-bisphosphate, 0.5 U/µl T4 RNA ligase (Ambion, Life Technologies, Grand Island, NY), 0.3 U/µl SUPERase•In (Ambion) 10% (v/v) dimethylsulfoxide at 4°C overnight. RNA was gel purified and concentration determined by absorbance at 260 nm.

Measurement of ΔG of Base Pairing by Hyperchromicity Analysis

RNA hyperchromicity analyses were performed in 100 mM KCl, 18 mM sodium cacodylate (pH 7.4) and 4 mM $MgCl_2$. Equimolar amounts of the two RNAs were used for each duplex. Control experiments established that none of the individual RNA strands contained stable intramolecular structure. Absorbance at 260 nm was measured every 0.5°C from 10°C to 80°C at a heating rate of 1°C min^{-1} using a DU 640 spectrophotometer with a high performance temperature controller (Beckman Coulter, Indianapolis, IN). Data were fit to a two-state model using Meltwin (McDowell and Turner, 1996; Schroeder and Turner, 2009).

Nearest Neighbor Analysis

$\Delta G_{25°C} = -RT \ln(1/K_M)$, where R = 1.987 cal $K^{-1}mol^{-1}$ and T = 298.15K. Error was propagated using the quadratic sum of the partial uncertainties (Taylor, 1997). Theoretical $\Delta G_{25°C}$ was calculated based on nearest neighbors values and from using RNAstructure 5.3 (Xia et al., 1998; Reuter and Mathews, 2010). The seed was taken to correspond to g2–g8, and 3′ supplementary pairing to correspond to g12–g17.

Binding, Competition, and Dissociation Assays

Fly Ago2-RISC was assembled for 90 min with *let-7* siRNA (5′-pUGA GGU AGU AGG UUG UAU AGU-3′) using 0–2 hr embryo lysate. Mouse AGO2-RISC was assembled for 90 min with *let-7* siRNA using S100 extract. The S100 extract

was obtained from *Ago2*$^{-/-}$ mouse embryonic fibroblasts (MEFs) immortalized with SV40 large T-antigen that stably over-expressed mouse AGO2 (O'Carroll et al., 2007). Briefly, MEF cells were grown to confluence in 5% CO_2 at 37°C in DMEM (Invitrogen, Life Technologies) supplemented with 15% fetal bovine serum (PAA Laboratories, Inc., Dartmouth, MA) and 50 U/ml penicillin and streptomycin (Invitrogen, Life Technologies). S100 extract was prepared as described (Dignam et al., 1983) except that the cell pellet was washed three times in ice-cold PBS and once in buffer A that contains 10 mM HEPES–KOH (pH 7.9), 10 mM potassium acetate, 1.5 mM magnesium acetate, 0.5 mM DTT and EDTA-free protease inhibitor cocktail. Then, the pellet is resuspended in twice its volume with buffer A and incubated on ice for 20 minutes. This allows the cells to swell and subsequently lysed (with 20 strokes) using a Dounce homogenizer and a tight pestle (B type) on ice. The homogenate was centrifuged gently at 2,000 x *g* to remove nuclei and cell membranes. Next, 0.11 volume (that of the clarified supernatant from the low speed centrifugation) of buffer that consists of 300 mM HEPES-KOH (pH 7.9), 1.4 M potassium acetate, 30 mM magnesium acetate, 0.5 mM DTT and EDTA-free protease inhibitor cocktail was added. This was immediately followed by ultracentrifugation at 100,000 × *g* at 4 °C for 20 min where the supernatant constitutes the S100 extract. Ice-cold 80% (w/v) glycerol is then added to the S100 extract to achieve a 13% (w/v) final glycerol concentration. Finally, the S100 extract was, aliquoted,

frozen in liquid nitrogen, and stored at −80 °C. The protein concentration of the S100 extract was ~3–4 mg/ml. Assembled Ago2-RISC was captured with a 5′ biotinylated 2′-*O*-methyl-modified oligonucleotide that pairs with g2–g8 and g13–g16 of *let-7* seed (5′-Biotin-AUA GAC UGC GAC AAU AGC CUA CCU CCG AAC G-3′) and eluted using a DNA oligonucleotide bearing four 2′-*O*-methyl modifications (m; 5′-GGmU AmGG CTA TmUmG TCG CAG TCT AT-3′ (Flores-Jasso et. al., 2013). Fly Ago2-RISC eluate was further purified using Superdex 200 HR 10/300 GL (GE Healthcare Bioscience, Pittsburgh, PA) size exclusion column. Mouse AGO2-RISC eluate was subsequently purified using Mono S 5/50 GL (GE Healthcare) cationic exchanger. Column purified fly Ago2-RISC was concentrated by centrifugation (3,000 × *g*; Amicon Ultra-0.5 ml, Millipore, Billerica, MA) in 100 mM potassium acetate, 18 mM HEPES (pH 7.4), 3 mM magnesium acetate, 5 mM DTT, 0.01% (v/v) IGEPAL CA-630, 0.01 mg/ml baker's yeast tRNA (equilibration buffer). Column purified mouse AGO2-RISC was dialyzed using a 3 ml *Slide-A-Lyzer* cassette (Pierce, ThermoFisher Scientific) in equilibration buffer supplemented to a final concentration of 20% glycerol (v/v). Finally, mouse AGO2-RISC is concentrated by centrifugation. Binding reactions were incubated at 25°C for 1 hr. RNA binding was measured by capturing protein-RNA complexes on Protran nitrocellulose membrane (Whatman, GE Healthcare Bioscience, Pittsburgh, *PA)* and unbound RNA on a Nylon XL membrane (GE Healthcare Bioscience*) using a* Bio-Dot apparatus (Bio-

Rad, Hercules, CA). After applying the sample under vacuum, membranes were washed with ice-cold equilibration buffer. Membranes were air-dried and signals detected by phosphorimaging. Competition reactions were incubated at 25°C for 1 (mouse AGO2-RISC) or 6 hr (fly Ago2-RISC). To measure dissociation rate constants, 0.5–5 nM Ago2-RISC was incubated with 0.5 nM ^{32}P-radiolabeled RNA target for 1 hr, then 50 nM or 500 nM competitor RNA was added and dissociation measured by filter-binding.

Measuring the Concentration of Purified Ago2-RISC

RNA guide was extracted from Ago2-RISC and resolved along with concentration standards by 15% denaturing polyacrylamide gel electrophoresis. RNA was transferred from the gel to Hybond-NX membrane (Amersham Biosciences, Piscataway, NJ) at 20 V for 1 hr, cross-linked to the membrane with 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (Pall and Hamilton, 2008), and probed as described to detect the guide and standards (Table 3.S2A; Ameres et al., 2010; Han et al., 2011).

Data Analysis and Kinetic Modeling

When [E] << [S], time courses were fit to $y = mx + b$; when [S] < [E], time courses were fit to $y = y_0 + Ae^{-kx}$, where the initial rate, $v_0 = Ak$ (Lu and Fei, 2003). Initial rates measured at different substrate concentrations were fit to the Michaelis-

Menten model using Visual Enzymics 2008 (Softzymics, Princeton, NJ) for Igor Pro 6.11,

$$v = \frac{V_{\max}[S_T]}{K_M + [S_T]}$$

or to the Morrison quadratic equation for tight binding,

$$v = V_{\max} \frac{\left([S_T] + [E_T] + K_M\right) - \sqrt{\left([S_T] + [E_T] + K_M\right)^2 - 4[E_T][S_T]}}{2[E_T]}$$

using Igor Pro 6.11 (WaveMetrics, Lake Oswego, OR). The rate of product formation was determined separately for each replicate, and significance determined using Student's two-tailed, two-sample, equal variance *t*-test (Excel, Microsoft, Seattle, WA). R 2.14.0 software was used for other statistical analyses. To obtain the enzyme concentration by pre-steady state analysis, data were fit using non-linear least square regression in Igor Pro 6 to the burst and steady state equation,

$$F(t) = E \times \frac{a^2}{(a+b)^2}\left(1 - e^{-(a+b)t}\right) + E \times \frac{ab}{(a+b)}t$$

where *F(t)* is target cleaved with time, *E* is the enzyme concentration, *a* and *b* are rate constants according to the following scheme,

$$E + S \xrightarrow{a} EI \xrightarrow{b} E + P$$

Because K_D < [RNA target], all binding data were fit to

$$f = \frac{([E_T] + [S_T] + K_D) - \sqrt{([E_T] + [S_T] + K_D)^2 - 4([E_T][S_T])}}{2[S_T]}$$

where *f* is fraction target bound, $[E_T]$ is total enzyme concentration, $[S_T]$ is total RNA target concentration, and K_D is the apparent equilibrium dissociation constant.

For competition assays, the apparent equilibrium dissociation constant, K_C, for the competitor RNAs was obtained by fitting the data to the normalized quadratic solution of the Lin and Riggs equation (Lin and Riggs, 1972; Weeks and Crothers, 1992),

$$\Theta = \frac{\left([E_T] + [S_T] + K_D + \frac{K_D[C_T]}{K_C}\right) - \sqrt{\left([E_T] + [S_T] + K_D + \frac{K_D[C_T]}{K_C}\right)^2 - 4([E_T][S_T])}}{2[S_T]}$$

where Θ is the fraction target bound in the presence of competitor RNA, [C_T], with an apparent dissociation constant of K_C. To obtain the dissociation rate constant, k_{off}, data were fit to $f = e^{-k_{off}t}$; $t_{½} = \ln(2)/k_{off}$.

Berkeley Madonna 8.3.18 (www.berkeleymadonna.com/index.html) was used to model target derepression of miRNA-RISC by target competitors for high (human miR-21), intermediate (human miR-93) and low (human miR-24) abundance miRNAs. The concentration of miR-21 RISC (4 nM) was calculated using the reported abundance of ~12,000 copies per HeLa cell, assuming a maximum cell volume of 5,000 μm^3 (Lim et al., 2003; Moran et al., 2010). The concentrations of miRNA-RISC for miR-93 and miR-24 were calculated, relative to miR-21, based on published ratios of sequencing reads (Cole et al., 2009).

ACKNOWLEDGEMENTS

We thank D. Turner, J. Chen, A. Caruthers, R. Gilmore, S. Ryder, B. Farley, O. Bilsel, S. Kathuria, and P. Gandhi for help and discussions; the Turner lab for use of equipment; A. Boucher, T. Covello, and G. Farley for technical support; and members of the Zamore lab for critical comments on the manuscript. This work was supported in part by National Institutes of Health grants GM62862 and GM65236.

Figure 3.S1

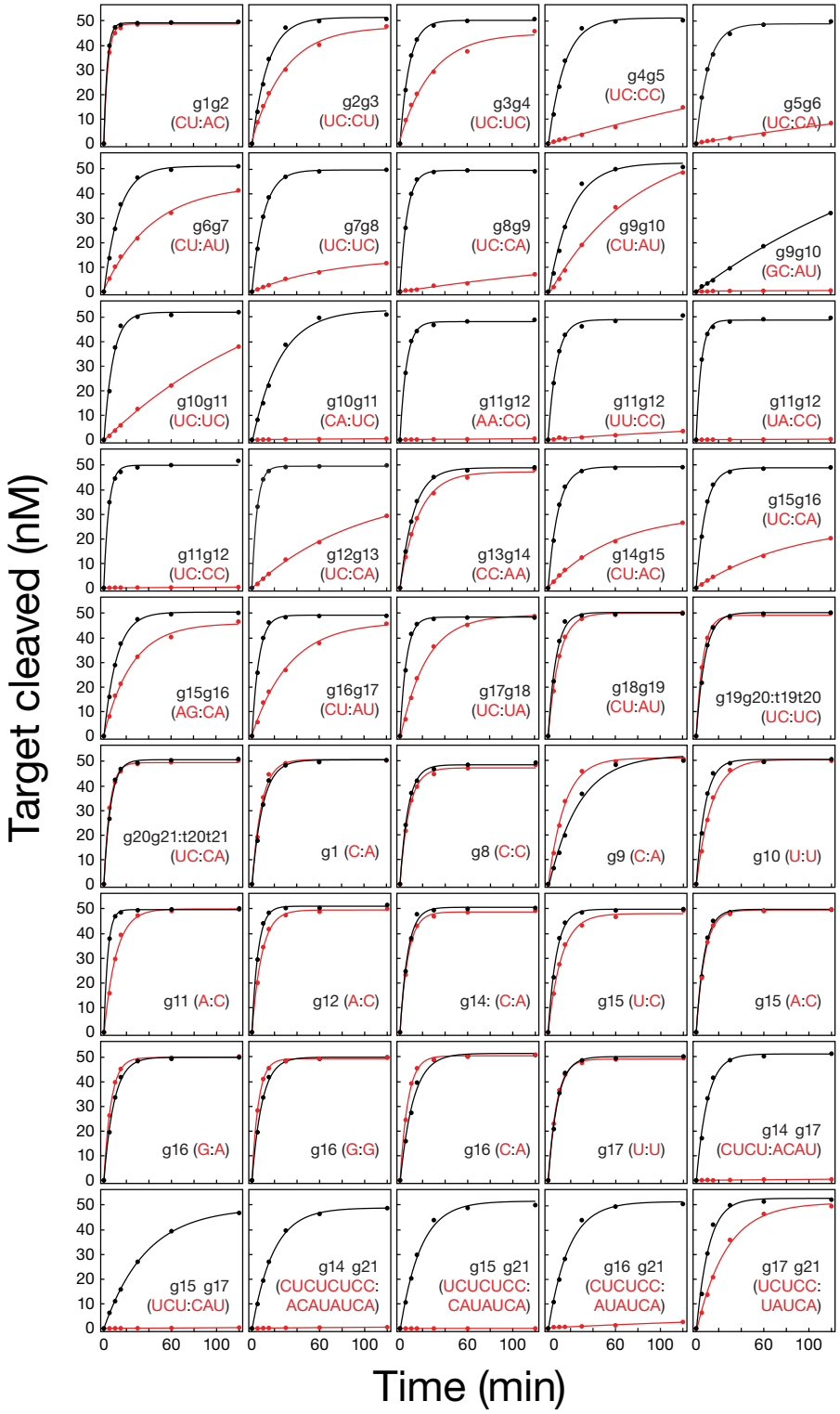

Target cleaved (nM)
Time (min)
g1g2 (CU:AC)
g2g3 (UC:CU)
g3g4 (UC:UC)
g4g5 (UC:CC)
g5g6 (UC:CA)
g6g7 (CU:AU)
g7g8 (UC:UC)
g8g9 (UC:CA)
g9g10 (CU:AU)
g9g10 (GC:AU)
g10g11 (UC:UC)
g10g11 (CA:UC)
g11g12 (AA:CC)
g11g12 (UU:CC)
g11g12 (UA:CC)
g11g12 (UC:CC)
g12g13 (UC:CA)
g13g14 (CC:AA)
g14g15 (CU:AC)
g15g16 (UC:CA)
g15g16 (AG:CA)
g16g17 (CU:AU)
g17g18 (UC:UA)
g18g19 (CU:AU)
g19g20:t19t20 (UC:UC)
g20g21:t20t21 (UC:CA)
g1 (C:A)
g8 (C:C)
g9 (C:A)
g10 (U:U)
g11 (A:C)
g12 (A:C)
g14: (C:A)
g15 (U:C)
g15 (A:C)
g16 (G:A)
g16 (G:G)
g16 (C:A)
g17 (U:U)
g14 g17 (CUCU:ACAU)
g15 g17 (UCU:CAU)
g14 g21 (CUCUCUCC: ACAUAUCA)
g15 g21 (UCUCUCC: CAUAUCA)
g16 g21 (CUCUCC: AUAUCA)
g17 g21 (UCUCC: UAUCA)
0
10
20
30
40
50
60
100

Figure Legend 3.S1. Target Cleavage by *let-7*-Programmed Fly Ago2-RISC, Related to Figure 3.1

Red, mismatched targets; Black, fully complementary. Substrate was in excess over enzyme (S = 50 nM, E ~3 nM).

Figure 3.S2

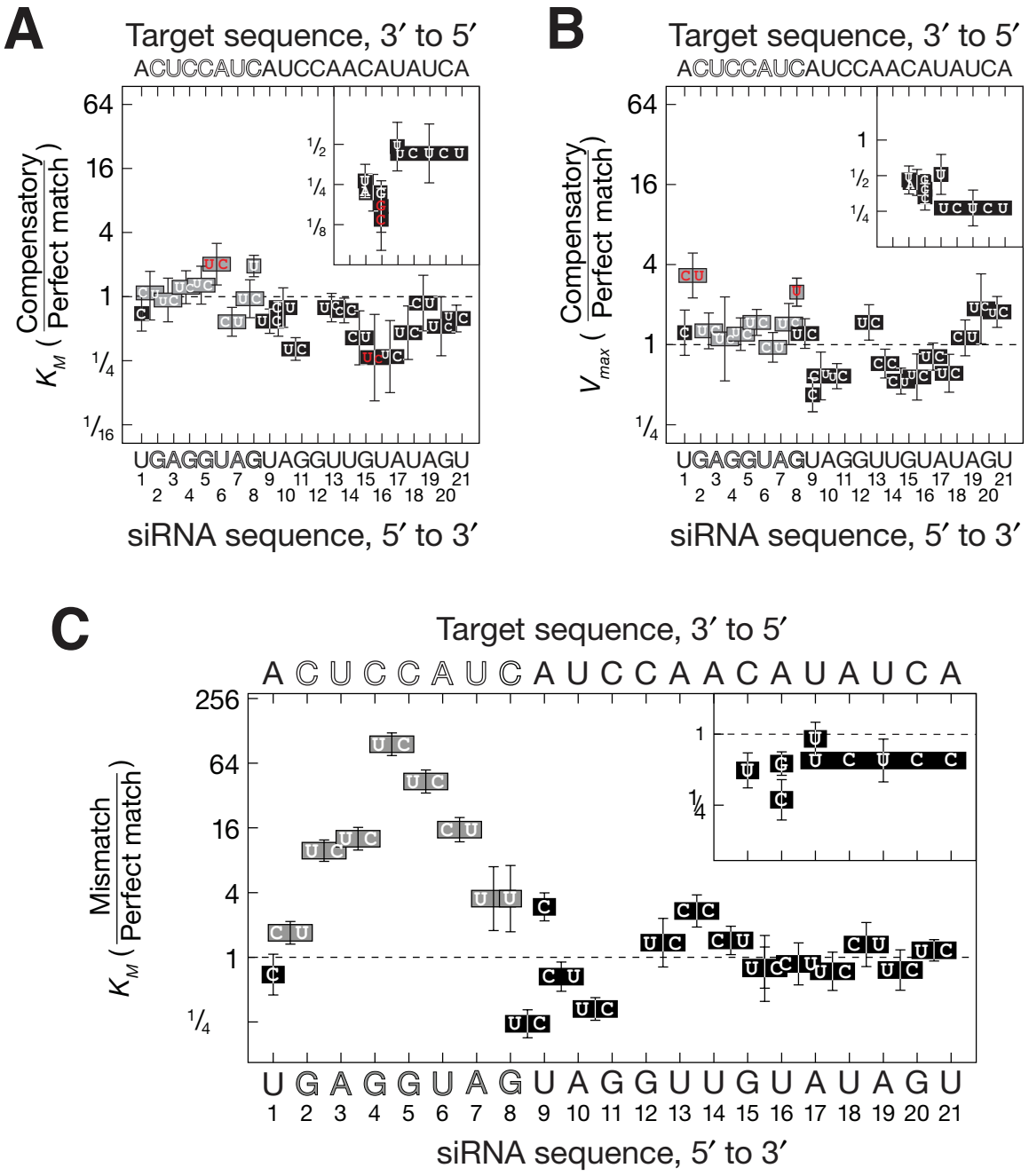

A
Target sequence, 3′ to 5′
ACUCCAUCAUCCAACAUAUCA
Compensatory / Perfect match
K_M
64
16
4
1
1/4
1/16
1/2
1/4
1/8
UGAGGUAGUAGGUUGUAUAGU
1 3 5 7 9 11 13 15 17 19 21
2 4 6 8 10 12 14 16 18 20
siRNA sequence, 5′ to 3′
B
Target sequence, 3′ to 5′
ACUCCAUCAUCCAACAUAUCA
Compensatory / Perfect match
V_{max}
64
16
4
1
1/4
1
1/2
1/4
UGAGGUAGUAGGUUGUAUAGU
1 3 5 7 9 11 13 15 17 19 21
2 4 6 8 10 12 14 16 18 20
siRNA sequence, 5′ to 3′
C
Target sequence, 3′ to 5′
A C U C C A U C A U C C A A C A U A U C A
Mismatch / Perfect match
K_M
256
64
16
4
1
1/4
1
1/4
U G A G G U A G U A G G U U G U A U A G U
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
siRNA sequence, 5′ to 3′

Figure Legend 3.S2. Michaelis-Menten Kinetics of Fly Ago2-RISC Using Compensatory Targets, Related to Figure 3.1

(A) Relative K_M for fly Ago2-RISC comparing the individual fully complementary targets for all the *let-7* derivatives to the original fully complementary target for the parental *let-7* siRNA guide. Values that differed significantly (*p*-value <0.05) from the parental *let-7* target are highlighted in red.

(B) As in (A) but for relative $V_{max.}$

(C) Values for the K_M of mismatched *let-7* variant siRNAs relative to the K_M of the parental *let-7* for its fully complementary target RNA.

Values are reported as mean ± standard deviation for three independent experiments.

Figure 3.S3

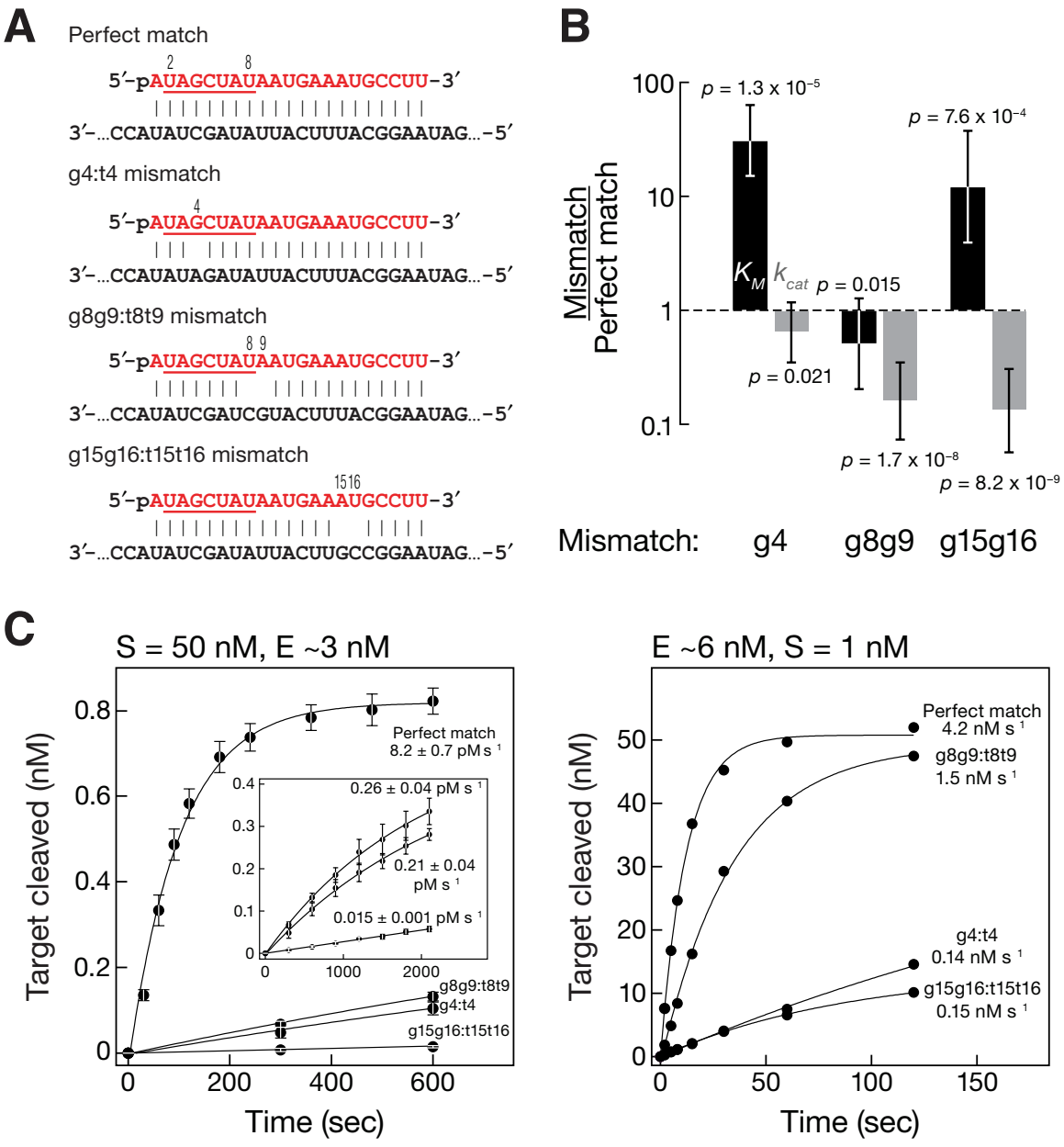

A
Perfect match
5′-pAUAGCUAUAAUGAAAUGCCUU-3′
3′-...CCAUAUCGAUAUUACUUUACGGAAUAG...-5′
g4:t4 mismatch
5′-pAUAGCUAUAAUGAAAUGCCUU-3′
3′-...CCAUAUAGAUAUUACUUUACGGAAUAG...-5′
g8g9:t8t9 mismatch
5′-pAUAGCUAUAAUGAAAUGCCUU-3′
3′-...CCAUAUCGAUCGUACUUUACGGAAUAG...-5′
g15g16:t15t16 mismatch
5′-pAUAGCUAUAAUGAAAUGCCUU-3′
3′-...CCAUAUCGAUAUUACUUGCCGGAAUAG...-5′
B
Mismatch / Perfect match
p = 1.3 x 10−5
p = 7.6 x 10−4
KM
kcat
p = 0.015
p = 0.021
p = 1.7 x 10−8
p = 8.2 x 10−9
Mismatch: g4 g8g9 g15g16
C
S = 50 nM, E ~3 nM
Target cleaved (nM)
Time (sec)
Perfect match
8.2 ± 0.7 pM s 1
0.26 ± 0.04 pM s 1
0.21 ± 0.04 pM s 1
0.015 ± 0.001 pM s 1
g8g9:t8t9
g4:t4
g15g16:t15t16
E ~6 nM, S = 1 nM
Target cleaved (nM)
Time (sec)
Perfect match
4.2 nM s 1
g8g9:t8t9
1.5 nM s 1
g4:t4
0.14 nM s 1
g15g16:t15t16
0.15 nM s 1

Figure Legend 3.S3. Fly Ago2-RISC Loaded With *Luciferase* siRNA, Related to Figure 3.1

(A) Scheme for a *Renilla reniformis* luciferase siRNA (red) pairing with its targets.

(B) Michaelis-Menten parameters for *luciferase* siRNA for mismatched targets, relative to the corresponding fully complementary target RNA. Mean K_M and k_{cat} values (± standard deviation) are from eight independent experiments. *P*-values calculated using two-tailed Student's *t*-test.

(C) Target cleavage by fly Ago2-RISC loaded with *luciferase* siRNA using enzyme (left) and substrate excess (right). Initial rates are reported to the right of each curve. Data represent mean ± standard deviation for four independent experiments.

Figure 3.S4

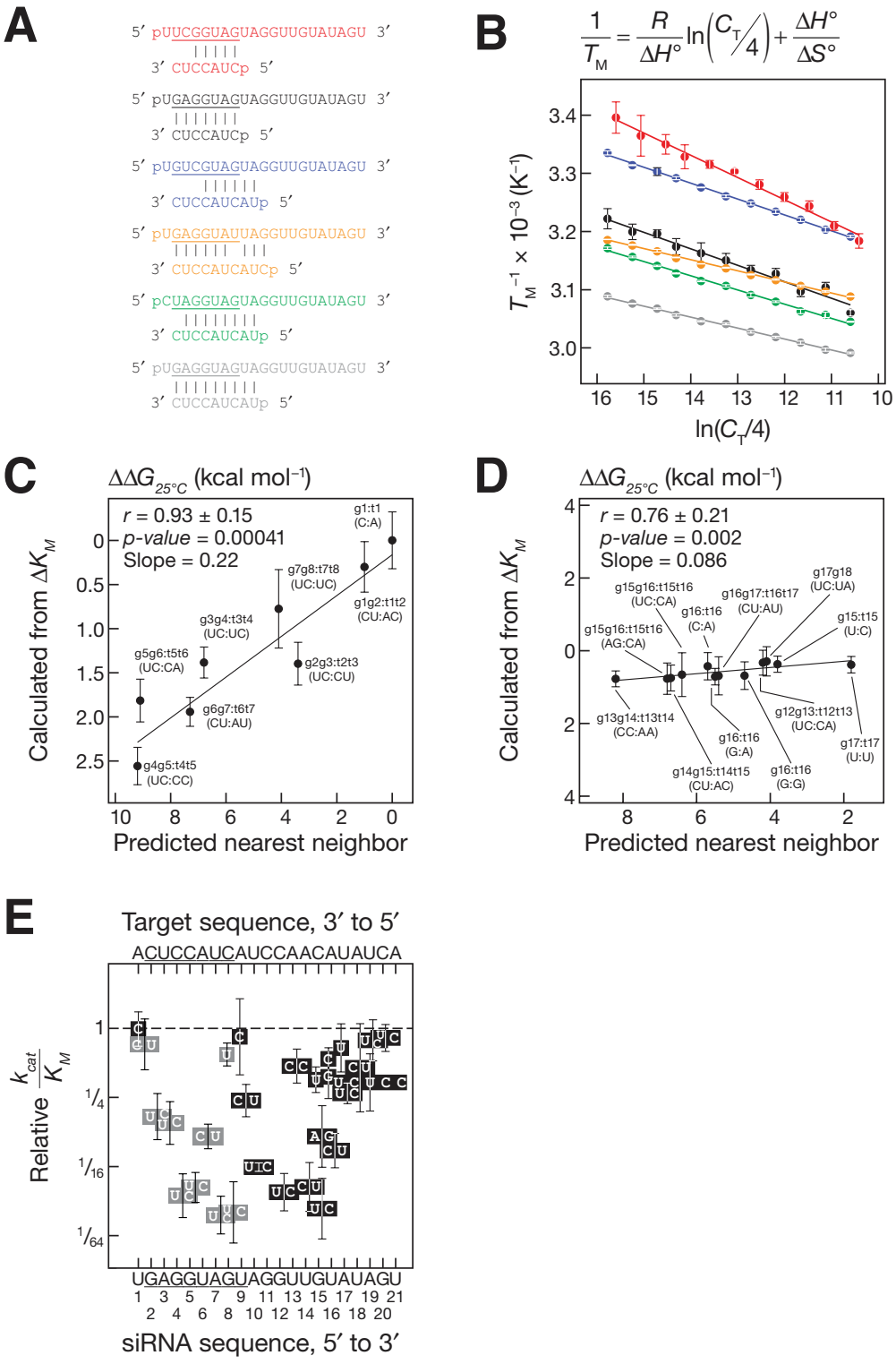

A
5′ pUUCGGUAGUAGGUUGUAUAGU 3′
3′ CUCCAUCp 5′
5′ pUGAGGUAGUAGGUUGUAUAGU 3′
3′ CUCCAUCp 5′
5′ pUGUCGUAGUAGGUUGUAUAGU 3′
3′ CUCCAUCAUp 5′
5′ pUGAGGUAUUAGGUUGUAUAGU 3′
3′ CUCCAUCAUCp 5′
5′ pCUAGGUAGUAGGUUGUAUAGU 3′
3′ CUCCAUCAUp 5′
5′ pUGAGGUAGUAGGUUGUAUAGU 3′
3′ CUCCAUCAUp 5′
B
$\frac{1}{T_M} = \frac{R}{\Delta H°}\ln\left(C_T/4\right) + \frac{\Delta H°}{\Delta S°}$
$T_M^{-1} \times 10^{-3}$ (K^{-1})
ln(C_T/4)
C
ΔΔ$G_{25°C}$ (kcal mol^{-1})
Calculated from ΔK_M
Predicted nearest neighbor
r = 0.93 ± 0.15
p-value = 0.00041
Slope = 0.22
g1:t1 (C:A)
g1g2:t1t2 (CU:AC)
g2g3:t2t3 (UC:CU)
g3g4:t3t4 (UC:UC)
g4g5:t4t5 (UC:CC)
g5g6:t5t6 (UC:CA)
g6g7:t6t7 (CU:AU)
g7g8:t7t8 (UC:UC)
D
ΔΔ$G_{25°C}$ (kcal mol^{-1})
Calculated from ΔK_M
Predicted nearest neighbor
r = 0.76 ± 0.21
p-value = 0.002
Slope = 0.086
g15g16:t15t16 (UC:CA)
g16g17:t16t17 (CU:AU)
g17g18 (UC:UA)
g16:t16 (C:A)
g15:t15 (U:C)
g15g16:t15t16 (AG:CA)
g12g13:t12t13 (UC:CA)
g13g14:t13t14 (CC:AA)
g16:t16 (G:A)
g17:t17 (U:U)
g14g15:t14t15 (CU:AC)
g16:t16 (G:G)
E
Target sequence, 3′ to 5′
ACUCCAUCAUCCAACAUAUCA
Relative $\frac{k_{cat}}{K_M}$
UGAGGUAGUAGGUUGUAUAGU
siRNA sequence, 5′ to 3′

Figure Legend 3.S4. Base Pairing in Fly Ago2-RISC Obeys the Standard Rules for RNA, Related to Figure 3.3

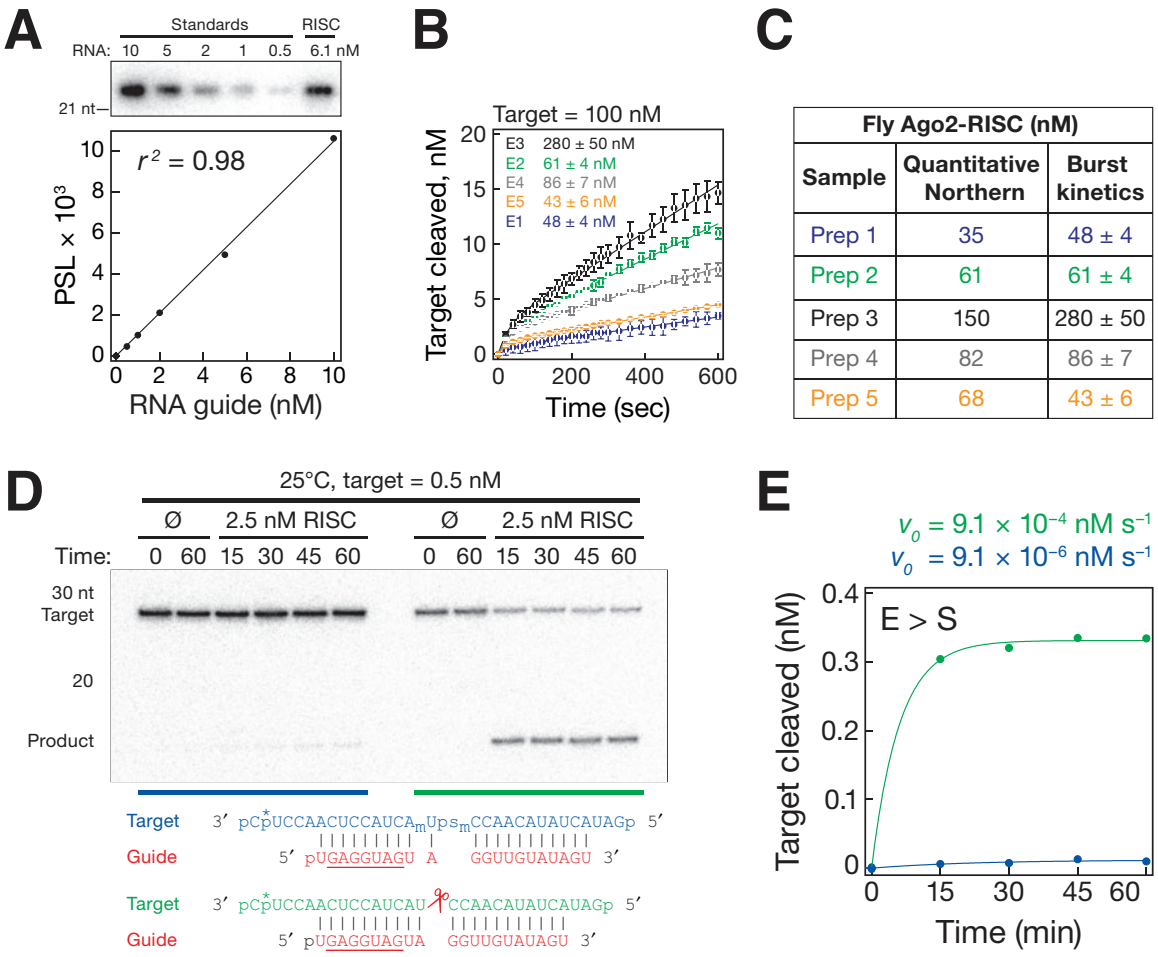
(A) RNA duplexes used to measure the effect of mismatches on the strength of base pairing under standard RNAi reaction conditions.

(B) Van't Hoff plots for the RNA duplexes in (A). C_T is the total single-strand RNA concentration.

(C,D) Comparison of the $\Delta\Delta G_{25°C}$ derived from K_M and the $\Delta\Delta G_{25°C}$ calculated by nearest neighbor analysis between fully complementary targets and targets with mismatches in the seed, positions g2–g8 (C) and in the g12–g18 region (D).

(E) Change in k_{cat}/K_M for all mismatches introduced into the *let-7* siRNA guide. Values are mean ± standard deviation.

Figure 3.S5



Fly Ago2-RISC (nM)		
Sample	**Quantitative Northern**	**Burst kinetics**
Prep 1	35	48 ± 4
Prep 2	61	61 ± 4
Prep 3	150	280 ± 50
Prep 4	82	86 ± 7
Prep 5	68	43 ± 6

Figure Legend 3.S5. Purified Fly Ago2-RISC Binds a Fully Complementary Target RNA Tightly and Specifically, Related to Figure 3.3

(A) Quantitative Northern hybridization to detect the *let-7* siRNA guide strand present in purified fly Ago2-RISC and to detect the corresponding *let-7* RNA concentration standards. Above, the gel image quantified to yield the graph below. PSL (background subtracted) are the units reported by the phosphorimager.

(B) Pre-steady-state "burst" kinetics to determine the amount of active *let-7* RISC present in five independent preparations of purified fly Ago2-RISC. Values are mean ± standard deviation.

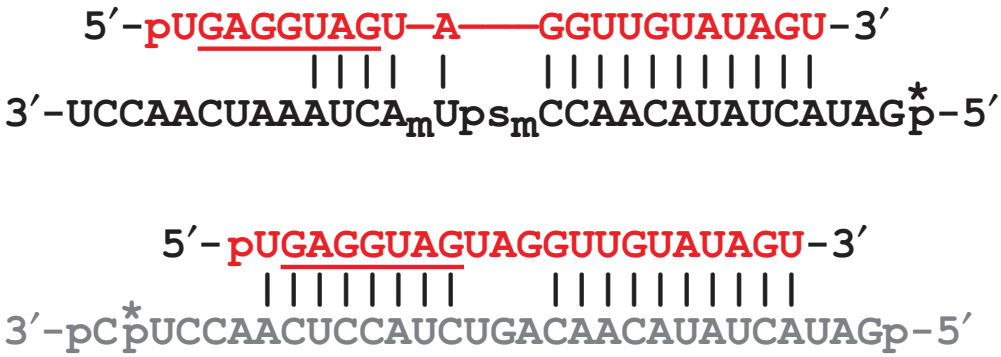
(C) Concentrations of fly Ago2-RISC obtained by quantitative Northern hybridization and pre-steady-state kinetics for the five independent preparations in (A) and (B).

(D) Time course of cleavage by fly Ago2-RISC using modified (blue) and unmodified (green) target RNA. A subscript "m" denotes 2′-*O*-methyl ribose; "ps" indicates a phosphorothioate linkage.

(E) Cleaved target (product) generated with time for the reactions shown in (D).

Figure 3.S6

A



B

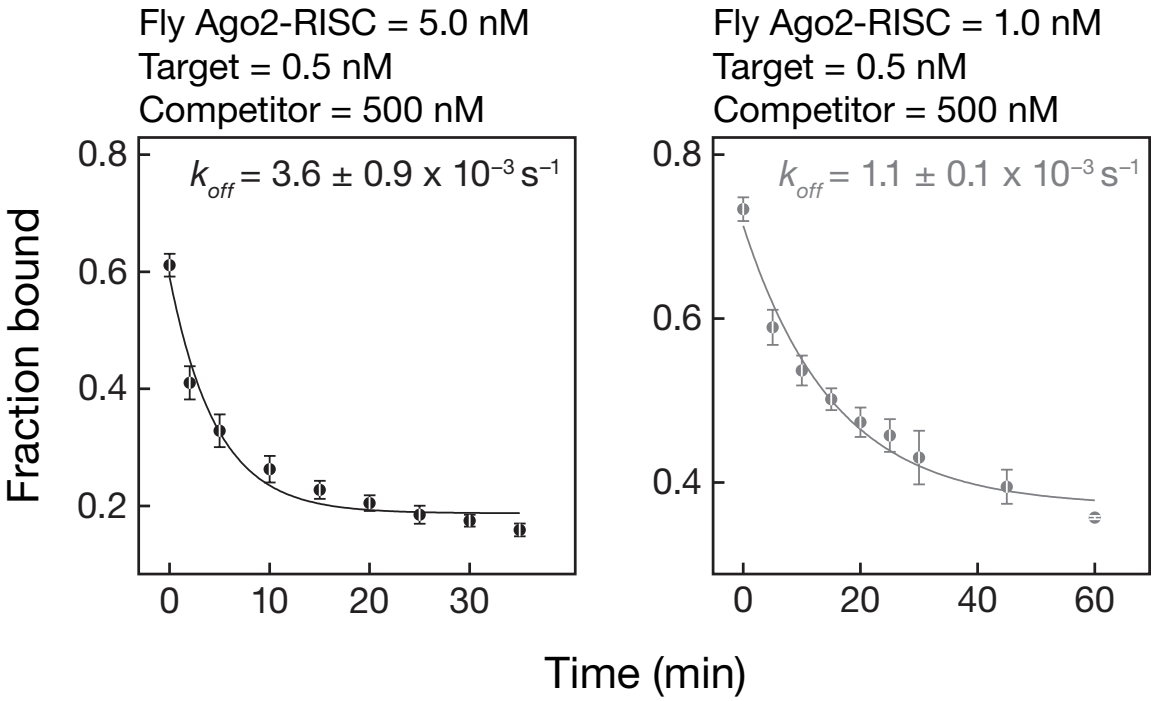


Figure Legend 3.S6. Mismatches Promote Target Dissociation, Related to Figure 3.5

(A) Target RNA bearing a g4g5:t4t5 mismatch (black) or g9–g11:t9–t11 mismatch (grey). An asterisk indicates the position of the ^{32}P-radiolabel; the subscript “m” indicates 2′-*O*-methyl ribose; “ps” indicates a phosphorothioate linkage.

(B) Dissociation rate curves for the targets shown in (A).

Data represent mean ± standard deviation for three independent experiments.

Figure 3.S7

A

$E + S \underset{k_{1r}}{\overset{k_{1f}}{\rightleftharpoons}} ES$ $\quad d(ES)/dt = k_{1f} \times (E_T - ES - EC) \times (S_T - ES) - (k_{1r} \times ES)$

$E + C \underset{k_{2r}}{\overset{k_{2f}}{\rightleftharpoons}} EC$ $\quad d(EC)/dt = k_{2f} \times (E_T - ES - EC) \times (C_T - EC) - (k_{2r} \times EC)$

E = Free miRNA-RISC, E_T = Total miRNA-RISC
S = Free target, S_T = Total target = 0.17 nM
C = Free competitor, C_T = Total competitor
ES = miRNA-RISC:target complex
EC = miRNA-RISC:competitor complex

	fly Ago2	mouse AGO2
$k_{1f} = k_{2f} =$	0.21 nM^{-1} s^{-1}	0.020 nM^{-1} s^{-1}
$k_{1r} = k_{2r} =$	0.045 s^{-1}	5.0 x 10^{-4} s^{-1}
$k_{2f} =$	0.024 nM^{-1} s^{-1}	0.036 nM^{-1} s^{-1}
$k_{2r} =$	8.8 x 10^{-5} s^{-1}	7.7 x 10^{-4} s^{-1}

B

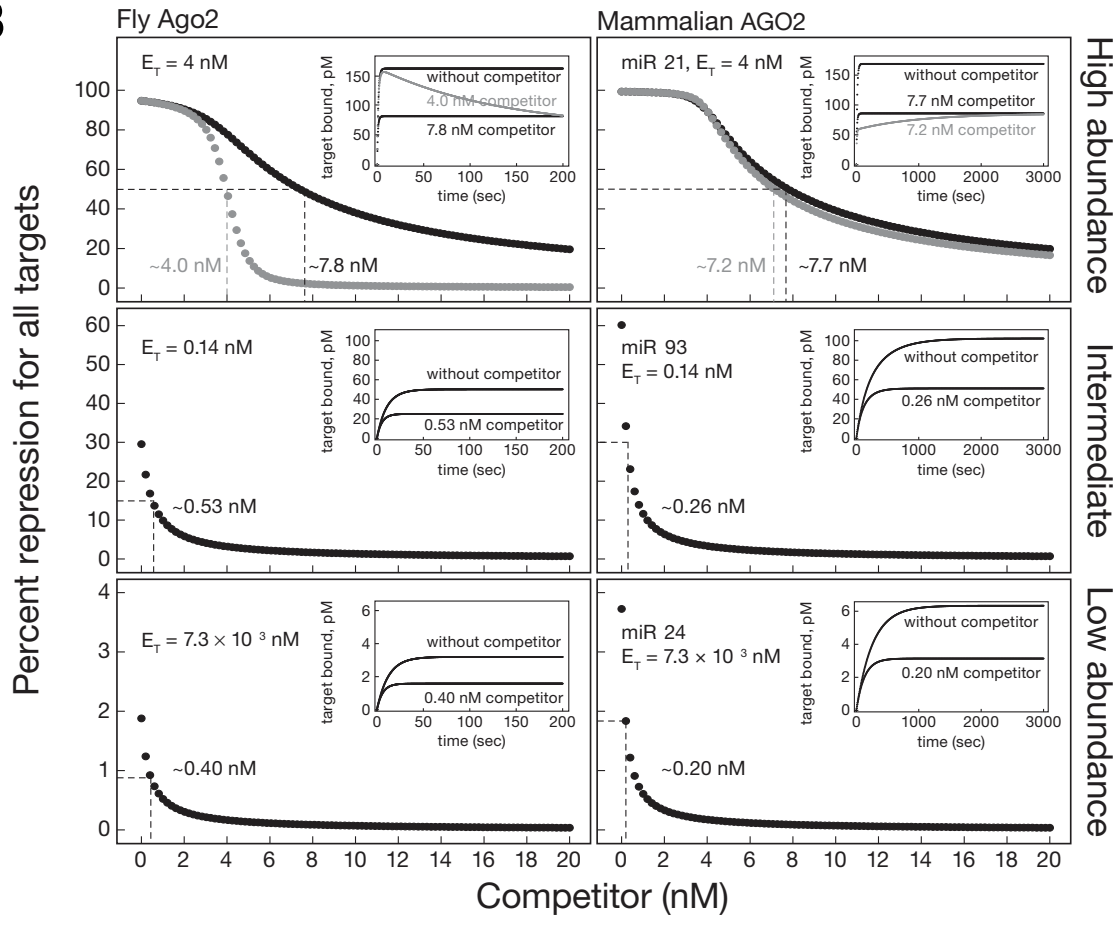


Figure Legend 3.S7. Modeling Target Repression by miRNA, Related to Figure 3.7

(A) Reaction equations and rates.

(B) Predicted seed match or fully complementary (grey) competitor concentrations required to relieve 50% of repression by Ago2-RISC. Simulations used the k_{on} and k_{off} values measured for fly (left panels) or mouse AGO2 (right panels). We assume every bound miRNA-RISC fully represses its target. Insets show concentration of target bound with and without competitor.

Table 3.S1. Michaelis-Menten parameters for target cleavage by *Drosophila* Ago2-RISC (mean ± S.D.; $n \geq 3$ except $n = 2$ for mismatch positions marked with an asterisk).

let–7-programmed Ago2-RISC							
	Mismatched target		Fully complementary target				
Mismatch position	K_M (nM)	V_{max} (nMs^{-1})	K_M (nM)	V_{max} (nMs^{-1})	$\frac{\text{relative } k_{cat}}{\text{relative } K_M}$	Single–turnover relative v_0, $\frac{\text{mismatched}}{\text{fully complementary}}$	$\frac{\text{relative } k_{cat}}{\text{relative } v_0}$
g1 (C:A)	17 ± 8	0.22 ± 0.09	17 ± 5	0.23 ± 0.08	0.99 ± 0.34	NA	NA
g1-g2 (CU:AC)	42 ± 5	0.80 ± 0.10	26 ± 12	0.63 ± 0.20	0.72 ± 0.37	NA	NA
g2-g3 (UC:CU)	250 ± 10	0.48 ± 0.12	23 ± 10	0.24 ± 0.05	0.17 ± 0.08	NA	NA
g3-g4 (UC:UC)	320 ± 40	0.35 ± 0.24	31 ± 8	0.21 ± 0.14	0.15 ± 0.06	NA	NA
g4-g5 (UC:CC)	2400 ± 300	0.63 ± 0.15	32 ± 11	0.23 ± 0.04	0.035 ± 0.015	NA	NA
g5-g6 (UC:CA)	1100 ± 100	0.27 ± 0.02	51 ± 20	0.28 ± 0.02	0.041 ± 0.012	NA	NA
g6-g7 (CU:AU)	390 ± 70	0.55 ± 0.03	15 ± 3	0.18 ± 0.03	0.11 ± 0.03	NA	NA
g7-g8 (UC:UC)	88 ± 60	0.023 ± 0.008	24 ± 8	0.27 ± 0.08	0.024 ± 0.009	0.0035 ± 0.0011	23 ± 7
g8 (U:C)	74 ± 15	0.43 ± 0.10	49 ± 1	0.47 ± 0.05	0.59 ± 0.14	0.61 ± 0.06	1.5 ± 0.3
g9 (C:A)	30 ± 10	0.15 ± 0.04	16 ± 5	0.079 ± 0.016	0.84 ± 0.64	0.80 ± 0.01	2.4 ± 0.7
g8-g9 (UC:CA)	6.0 ± 1.6	0.0018 ± 0.0006	15 ± 2	0.23 ± 0.03	0.025 ± 0.015	0.0052 ± 0.0017	2.1 ± 1.3
g9-g10 (CU:AU)	17 ± 6	0.021 ± 0.005	20 ± 7	0.11 ± 0.04	0.23 ± 0.08	0.052 ± 0.008	3.8 ± 1.1
g10-g11 (UC:UC)	8.3 ± 1.2	0.0069 ± 0.0003	8 ± 1	0.11 ± 0.00	0.062 ± 0.005	0.016 ± 0.001	4.0 ± 0.4

g12-g13 (UC:CA)	34 ± 18	0.017 ± 0.007	20 ± 4	0.28 ± 0.06	0.037 ± 0.014	0.0068 ± 0.0013	9.2 ± 2.6
g13-g14 (CC:AA)	68 ± 23	0.23 ± 0.05	19 ± 3	0.14 ± 0.02	0.47 ± 0.16	0.25 ± 0.02	6.9 ± 1.0
g15 (U:C)	12 ± 4	0.11 ± 0.02	7 ± 1	0.17 ± 0.03	0.36 ± 0.09	NA	NA
g16 (G:G)*	13 ± 7	0.24 ± 0.06	4.3 ± 1.6	0.15 ± 0.03	0.60 ± 0.11	NA	NA
g16 (G:A)*	14 ± 1	0.21 ± 0.01	4.3 ± 1.6	0.15 ± 0.03	0.38 ± 0.16	NA	NA
g16 (C:A)	7.0 ± 2.8	0.15 ± 0.01	3.4 ± 1.7	0.12 ± 0.02	0.54 ± 0.12	NA	NA
g14-g15 (CU:AC)	36 ± 10	0.018 ± 0.002	10 ± 6	0.10 ± 0.01	0.042 ± 0.020	0.0071 ± 0.0011	25 ± 4
g15-g16 (UC:CA)	20 ± 9	0.011 ± 0.004	6.7 ± 6.1	0.11 ± 0.04	0.027 ± 0.016	0.0045 ± 0.0010	24 ± 9
g15-g16 (AG:CA)	20 ± 14	0.064 ± 0.024	5.4 ± 1.2	0.17 ± 0.02	0.11 ± 0.07	0.033 ± 0.000	12 ± 5
g17 (U:U)	23 ± 4	0.25 ± 0.08	12 ± 4	0.19 ± 0.06	0.67 ± 0.32	NA	NA
g16-g17 (CU:AU)	22 ± 10	0.049 ± 0.006	6.9 ± 5.1	0.15 ± 0.01	0.086 ± 0.030	0.031 ± 0.001	10 ± 2
g17-g18 (UC:UA)	19 ± 8	0.053 ± 0.019	12 ± 6	0.11 ± 0.03	0.27 ± 0.06	0.037 ± 0.007	12 ± 3
g18-g19 (CU:AU)	33 ± 16	0.17 ± 0.03	22 ± 12	0.21 ± 0.05	0.45 ± 0.40	NA	NA
g19-g20 (UC:UC)	19 ± 9	0.45 ± 0.26	13 ± 8	0.35 ± 0.20	0.79 ± 0.32	0.73 ± 0.03	1.8 ± 0.2
g20-g21 (UC:CA)	29 ± 1	0.52 ± 0.03	16 ± 3	0.33 ± 0.05	0.82 ± 0.22	NA	NA
g17-g21 (UCUCC:UAUCA)	15 ± 7	0.053 ± 0.016	11 ± 5	0.10 ± 0.03	0.34 ± 0.19	NA	NA
Parental *let-7*	NA	NA	25 ± 6	0.19 ± 0.04	NA	NA	NA

Renilla reniformis luciferase siRNA-programmed Ago2-RISC							
g4 (G:A)	1700 ± 800	0.27 ± 0.11	55 ± 20	0.39 ± 0.08	0.021 ± 0.019	0.021 ± 0.007	30 ± 21
g8g9 (UA:CG)	31 ± 15	0.068 ± 0.029	55 ± 20	0.39 ± 0.08	0.31 ± 0.38	0.023 ± 0.015	6.8 ± 6.9

g15g16 (AU:GC)	660 ± 430	0.058 ± 0.023	55 ± 20	0.39 ± 0.08	0.011 ± 0.015	0.0021 ± 0.0013	62 ± 65

Table 3.S2A, related to Figure 3.1. Synthetic siRNAs used in this study.

Description	Sequence (seed, guide, passenger, altered sequences)
let-7 guide strand sequence. Completely complementary to perfect target	pUGAGGUAGUAGGUUGUAUAGU
Passenger strand to *let-7* guide and creates frayed siRNA with UC mismatch at g1 of preceding guide strand	pUAUACAACCUACUACCUCCUU
Passenger strand to *let-7* guide and creates frayed siRNA with UU mismatch at g1 of preceding guide strand	pUAUACAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g1g2:t1t2 when paired to *let-7* perfect target	pCUAGGUAGUAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUACUACCUAUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g2g3:t2t3 when paired to *let-7* perfect target	pUUCGGUAGUAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUACUACCGAUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g3g4:t3t4 when paired to *let-7* perfect target	pUGUCGUAGUAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUACUACGACUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g4g5:t4t5 when paired to *let-7* perfect target	pUGAUCUAGUAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUACUAGAUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g5g6:t5t6 when paired to *let-7* perfect target	pUGAGUCAGUAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUACUGACUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g6g7:t6t7 when paired to *let-7* perfect target	pUGAGGCUGUAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUACAGCCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g7g8:t7t8 when paired to *let-7* perfect target	pUGAGGUUCUAGGUUGUAUAGU

Passenger strand to preceding guide	pUAUACAACCUAGAACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g8g9:t8t9 when paired to *let-7* perfect target	pUGAGGUAUCAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUGAUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g9g10:t9t10 when paired to *let-7* perfect target	pUGAGGUAGCUGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCAGCUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g9g10:t9t10 when paired to *let-7* perfect target	pUGAGGUAGGCGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCGCCUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g10g11:t10t11 when paired to *let-7* perfect target	pUGAGGUAGUUCGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACGAACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g10g11:t10t11 when paired to *let-7* perfect target	pUGAGGUAGUCAGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACUGACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	pUGAGGUAGUAAAUUGUAUAGU
Passenger strand to preceding guide	pUAUACAAUUUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	pUGAGGUAGUAUUUUGUAUAGU
Passenger strand to preceding guide	pUAUACAAAAUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	pUGAGGUAGUAUAUUGUAUAGU
Passenger strand to preceding guide	pUAUACAAUAUACUACCUCUUU

Modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	pUGAGGUAGUAUCUUGUAUAGU
Passenger strand to preceding guide	pUAUACAAGAUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g12g13:t12t13 when paired to *let-7* perfect target	pUGAGGUAGUAGUCUGUAUAGU
Passenger strand to preceding guide	pUAUACAGACUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g13g14:t13t14 when paired to *let-7* perfect target	pUGAGGUAGUAGGCCGUAUAGU
Passenger strand to preceding guide	pUAUACGGCCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g14g15:t14t15 when paired to *let-7* perfect target	pUGAGGUAGUAGGUCUUAUAGU
Passenger strand to preceding guide	pUAUAAGACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g15g16:t15t16 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUUCAUAGU
Passenger strand to preceding guide	pUAUGAAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g15g16:t15t16 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUAGAUAGU
Passenger strand to preceding guide	pUAUCUAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g16g17:t16t17 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGCUUAGU
Passenger strand to preceding guide	pUAAGCAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g17g18:t17t18 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGUUCAGU
Passenger strand to preceding guide	pUGAACAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g18g19:t18t19 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGUACUGU

Passenger strand to preceding guide	pAGUACAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g19g20:t19t20 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGUAUUCU
Passenger strand to preceding guide	pAAUACAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 2 nt mismatches at g20g21:t20t21 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGUAUAUC
Modified *let-7* guide strand that creates 1 nt mismatch at g1:t1 when paired to *let-7* perfect target	pCGAGGUAGUAGGUUGUAUAGU
Modified *let-7* guide strand that creates 1 nt mismatch at g14:t14 when paired to *let-7* perfect target	pUGAGGUAGUAGGUCGUAUAGU
Passenger strand to preceding guide	pUAUACGACCUACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g15:t15 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUUUAUAGU
Passenger strand to preceding guide	pUAUAAAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g15:t15 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUAUAUAGU
Passenger strand to preceding guide	pUAUAUAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g16:t16 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGGAUAGU
Passenger strand to preceding guide	pUAUCCAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g16:t16 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGCAUAGU
Passenger strand to preceding guide	pUAUGCAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g17:t17 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGUUUAGU

Passenger strand to preceding guide	pUAAACAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g9:t9 when paired to *let-7* perfect target	pUGAGGUAGCAGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCUGCUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g10:t10 when paired to *let-7* perfect target	pUGAGGUAGUUGGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACCAACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g11:t11 when paired to *let-7* perfect target	pUGAGGUAGUAAGUUGUAUAGU
Passenger strand to preceding guide	pUAUACAACUUACUACCUCUUU
Modified *let-7* guide strand that creates 1 nt mismatch at g12:t12 when paired to *let-7* perfect target	pUGAGGUAGUAGAUUGUAUAGU
Passenger strand to preceding guide	pUAUACAAUCUACUACCUCUUU
Modified *let-7* guide strand that creates 3 nt mismatches at g15–g17:t15–t17 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUUCUUAGU
Passenger strand to preceding guide	pUAAGAAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 4 nt mismatches at g14–g17:t14–t17 when paired to *let-7* target	pUGAGGUAGUAGGUCUCUUAGU
Passenger strand to preceding guide	pUAAGAGACCUACUACCUCUUU
Modified *let-7* guide strand that creates 5 nt mismatches at g17–g21:t17–t21 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGUUCUCC
Passenger strand to preceding guide	pAGAACAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 6 nt mismatches at g16–g21:t16–t21 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUGCUCUCC
Passenger strand to preceding guide	pAGAGCAACCUACUACCUCUUU

Modified *let-7* guide strand that creates 7 nt mismatches at g15–g21:t15–t21 when paired to *let-7* perfect target	pUGAGGUAGUAGGUUUCUCUCC
Passenger strand to preceding guide	pAGAGAAACCUACUACCUCUUU
Modified *let-7* guide strand that creates 8 nt mismatches at g14-g21:t14–t21 when paired to *let-7* perfect target	pUGAGGUAGUAGGUCUCUCUCC
Passenger strand to preceding guide	pAGAGAGACCUACUACCUCUUU
Renilla reniformis luciferase derived siRNA guide strand	pAUAGCUAUAAUGAAAUGCCUU
Passenger strand to preceding guide	pGGCAUUUCAUUAUAGCUACUU
RNA that is exactly complementary to *let-7* guide used as Northern probe	pACUAUACAACCUACUACCUCA

Description	**Sequence**
7 nt target strand for RNA melt	pCUACCUC
9 nt target strand for RNA melt	pUACUACCUC
10 nt target strand for RNA melt	pCUACUACCUC

Description	**Sequence** (m, 2′-*O*-methyl ribose; ps; phosphorothioate; complementary to guide; mismatch)
Complete complementary target to *let-7*	pGAUACUAUACAACmCpsmUACUACCUCAACCU
Complete complementary target to *let-7* (unmodified and cleavable)	pGAUACUAUACAACCUACUACCUCAACCU

21 nt complete complementary target to *let-7*	pACUAUACAACmCpsmUACUACCUCA
Target complementary to *let-7* with g9–g11:t9–t11 central internal loop	pGAUACUAUACAACAGUCUACCUCAACCU
Target to *let-7* with g2–g21 complementarity	pGAUACUAUACAACmCpsmUACUACCUCUACCU
Target to *let-7* with g1–g19 complementarity	pGAUUAUAUACAACmCpsmUACUACCUCAACCU
Target to *let-7* with g4g5:t4t5 seed internal loop	pGAUACUAUACAACmCpsmUACUAAAUCAACCU
Target to *let-7* with g8g9:t8t9 central internal loop	pGAUACUAUACAACmCpsmUUAUACCUCAACCU
Target to *let-7* with g10g11:t10t11 central internal loop	pGAUACUAUACAACmGpsmAACUACCUCAACCU
Target to *let-7* with g15g16:t15t16 3′ supplementary region internal loop	pGAUACUAUCGAACmCpsmUACUACCUCAACCU
Target to *let-7* with g2–g16:t2–t16 complementarity	pGAUUAGCCACAACmCpsmUACUACCUCAACCU
Target to *let-7* with g9–g21:t9–t21 complementarity	pGAUACUAUACAACmCpsmUAGAUGGAGAAAAU
Target to *let-7* with seed only pairing (g2–g8:t2–t8)	pGAAAAAAAAAAAmApsmAUCUACCUCUAAAU
Target to *let-7* with seed + 3′ supplementary pairing (g2–g8:t2–t8; g13–g16:t13–t16)	pGAAAAAAACAAAmApsmAUCUACCUCUAAAU
Target to *let-7* with extended seed pairing (g2–g10:t2–t10)	pGAAAAAAAAUUAmApsmUACUACCUCUAAAU
Target to *let-7* with seed + extended 3′ supplementary pairing (g2–g8:t2–t8; g12–g17:t12–t17)	pGAAAAAUUACAACmApsmAUCUACCUCUAAAU

Target to *let-7* with centered pairing (g4–g15:t4–t15)	pGAUGGAAACCAACmCpsmUACUACCAGUACCU
Perfect target to *Renilla reniformis* luciferase	pGAUAAGGCAUUUCmApsmUUAUAGCUAUACCU
Target to *let-7* with 1 GU wobble in the seed at g4	pGAUACUAUACAACmCpsmUACUACUUCAACCU
Target to *let-7* with 1 GU wobble in the 3′ supplementary region at g15	pGAUACUAUAUAACmCpsmUACUACCUCAACCU
Target to *let-7* with 2 GU wobble in the seed at g2 and g8	pGAUACUAUACAACmCpsmUAUUACCUUAACCU
Target to *let-7* with 2 GU wobble in the seed at g4 and g5	pGAUACUAUACAACmCpsmUACUAUUUCAACCU
Target to *let-7* with 4 GU wobble in the seed at g2, g4, g5 and g8	pGAUACUAUACAACmCpsmUAUUAUUUUAACCU

Table 3.S2B, related to Figure 3.1. DNA oligonucleotides used in this study.

Description	Sequence (siRNA pairing site, change from *let-7* complementary target)
Forward primer containing T7 promoter consensus sequence for making DNA template used in transcription of *let-7* based targets.	GCG TAA TAC GAC TCA CTA TAG GGT CAC ATC TCA TCT ACC TCC
Reverse primer for making *let-7* guide perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making *let-7* passenger strand perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTA TAC AAC CTA CTA CCT CTT TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g1g2:t1t2 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TCT AGG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g2g3:t2t3 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTT CGG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g3g4:t3t4 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG TCG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g4g5:t4t5 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG ATC TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g5g6:t5t6 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGT CAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G

Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g6g7:t6t7 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG CTG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g7g8:t7t8 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TTC TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g8g9:t8t9 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAT CAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making first compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g9g10:t9t10 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG CTG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making second compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g9g10:t9t10 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG GCG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making first compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g10g11:t10t11 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TTC GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making second compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g10g11:t10t11 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TCA GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making first compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAA ATT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for second complementary target to p21- modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAT TTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G

Reverse primer for making third compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAT ATT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making fourth compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAT CTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g12g13:t12t13 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG TCT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g13g14:t13t14 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GCC GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g14g15:t14t15 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTC TTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making first compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g15g16:t15t16 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT TCA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making second compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g15g16:t15t16 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT AGA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g16g17:t16t17 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GCT TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g17g18:t17t18 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTT CAG TAT CCA GAG GAA TTC ATT ATC AGT G

Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g18g19:t18t19 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTA CTG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g19g20:t19t20 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTA TTC TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 2 nt mismatches at g20g21:t20t21 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTA TAT CAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g1:t1 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TCG AGG TAG TAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g14:t14 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTC GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making first compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g15:t15 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT TTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making second compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g15:t15 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT ATA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for first compensatory target to p21-976 modified *let-7* guide strand that creates 1 nt mismatch at g16:t16 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GGA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making second compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g16:t16 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GCA TAG TAT CCA GAG GAA TTC ATT ATC AGT G

Reverse primer for making compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g17:t17 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTT TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g9:t9 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG CAG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g10:t10 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TTG GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g11:t11 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAA GTT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 1 nt mismatch at g12:t12 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG ATT GTA TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 3 nt mismatches at g15–g17:t15–t17 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT TCT TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 4 nt mismatches at g14–g17:t14–t17 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTC TCT TAG TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 5 nt mismatches at g17–g21:t17–t21 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GTT CTC CAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 6 nt mismatches at g16–g21:t16–t21 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT GCT CTC CAT CCA GAG GAA TTC ATT ATC AGT G

Reverse primer for making compensatory target to modified *let-7* guide strand that creates 7 nt mismatches at g15–g21:t15–t21 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTT TCT CTC CAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making compensatory target to modified *let-7* guide strand that creates 8 nt mismatches at g14–g21:t14–t21 when paired to *let-7* perfect target	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TTG AGG TAG TAG GTC TCT CTC CAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making complementary target to *Renilla reniformis* luciferase derived siRNA guide strand	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TAT AGC TAT AAT GAA ATG CCT TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making target to *Renilla reniformis* luciferase derived siRNA guide strand with 1 nt mismatch at g4:t4	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TAT ATC TAT AAT GAA ATG CCT TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making target to *Renilla reniformis* luciferase derived siRNA guide strand with 2 nt mismatches at g8g9:t8t9	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TAT AGC TAG CAT GAA ATG CCT TAT CCA GAG GAA TTC ATT ATC AGT G
Reverse primer for making target to *Renilla reniformis* luciferase derived siRNA guide strand with 2 nt mismatches at g15g16:t15t16	CCC ATT TAG GTG ACA CTA TAG ATT TAC ATC GCG TTG AGT GTA GAA CGG TTG TAT AAA AGG TAT AGC TAT AAT GAA CGG CCT TAT CCA GAG GAA TTC ATT ATC AGT G

Table 3.S2C, related to Figure 3.1. RNA transcripts used in this study.

Description	Sequence (siRNA pairing site, change from *let-7* complement)
Transcript with perfect target site to *let-7* guide strand	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU ACU ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with perfect target site to *let-7* passenger strand	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **AAA GAG GUA GUA GGU UGU AUA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g1g2:t1t2 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU ACU ACC UAG** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g2g3:t2t3 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU ACU ACC GAA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g3g4:t3t4 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU ACU ACG ACA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g4g5:t4t5 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUCCCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU ACU AGA UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g5g6:t5t6 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU ACU GAC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G

Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g6g7:t6t7 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU ACA GCC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g7g8:t7t8 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU AGA ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g8g9:t8t9 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCU GAU ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g9g10:t9t10 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCA GCU ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
First transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g10g11:t10t11 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CCG CCU ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Second transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g10g11:t10t11 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CGA ACU ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
Third transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g10g11:t10t11 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA CUG ACU ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G
First transcript with compensatory target site to *let-7* guide strand that creates 2 nt mismatches at g11g12:t11t12 when paired to *let-7* perfect target	GGG UCA CAU CUC AUC UAC CUC CCG GUU UUA AUG AAU ACG AUU UUG UAC CAG AGU CCU UUG AUC GUG ACA AAA CAA UUG CAC UGA UAA UGA AUU CCU CUG GAU **ACU AUA CAA UUU ACU ACC UCA** ACC UUU UAU ACA ACC GUU CUA CAC UCA ACG CGA UGU AAA UCU AUA GUG UCA CCU AAA UGG G

Table 3.S3, **Related to Figure 3.3**. Thermodynamic parameters from optical melts of RNA duplexes. T_M values are for 1.0×10^{-4} M total single-strand concentration.

	T_M^{-1} versus ln(C_T/4)				Average of Individual Fits				Nearest Neighbor Prediction			
Guide:Target Duplex	$\Delta G_{25°C}$ (kcal mol^{-1})	ΔH (kcal mol^{-1})	ΔS (cal K^{-1} mol^{-1})	T_M (°C)	$\Delta G_{25°C}$ (kcal mol^{-1})	ΔH (kcal mol^{-1})	ΔS (cal K^{-1} mol^{-1})	T_M (°C)	$\Delta G_{25°C}$ (kcal mol^{-1})	ΔH (kcal mol^{-1})	ΔS (cal K^{-1} mol^{-1})	T_M (°C)
||||| CUCCAUCp	−8.7 ± 0.3	−50 ± 8	−150 ± 30	39 ± 0	−8.8 ± 0.3	−54 ± 6	−150 ± 20	40 ± 1	−8.5	−48	−130	39
||||||| CUCCAUCp	−12 ± 1	−71 ± 7	−200 ± 20	52 ± 0	−12 ± 1	−62 ± 7	−170 ± 20	53 ± 1	−12	−66	−180	55
|||||| CUCCAUCAUp	−9.8 ± 0	−73 ± 1	−210 ± 0	40 ± 0	−10 ± 0	−64 ± 0	−180 ± 0	42 ± 0	−8.3	−50	−140	37
|||||| ||| CUCCAUCAUCp	−15 ± 0	−100 ± 1	−300 ± 0	51 ± 0	−14 ± 0	−99 ± 2	−280 ± 10	51 ± 0	−13	−77	−220	52
|||||||| CUCCAUCAUp	−14 ± 0	−82 ± 2	−230 ± 10	56 ± 0	−14 ± 0	−76 ± 1	−210 ± 0	57 ± 0	−13	−69	−190	58
||||||||| CUCCAUCAUp	−18 ± 0	−106 ± 3	−290 ± 10	61 ± 0	−17 ± 0	−93 ± 1	−260 ± 0	62 ± 0	−16	−84	−230	64

Table 3.S4, **Related to Figure 3.5 and Figure 3.7.** Equilibrium competition parameters for fly Ago2-RISC and mouse AGO2-RISC (mean ± S.D.; n = 3).

***let–7*-programmed, affinity-purified *Drosophila melanogaster* Ago2-RISC**			
Target	**K_C (pM)**	**K_{rel}**	**$K_{D(adjusted)}$ (pM)**
28 nt complete complementarity	10 ± 1	1.0 ± 0.2	4.0 ± 1.0
21 nt complete complementarity	9.0 ± 0	1.0 ± 0.1	4.0 ± 1.0
28 nt unmodified g9–g11 central internal loop	9.0 ± 2.0	1.0 ± 0.2	3.0 ± 1.0
g2–g21 complementary	8.7 ± 0.8	1.0 ± 0.1	4.0 ± 1.0
g4g5 mismatches in seed	(5.6 ± 0.8) x 10^3	600 ± 100	(2.3 ± 0.7) x 10^3
g8g9 central internal loop	13 ± 1	1.3 ± 0.2	5.0 ± 2.0
g10g11 central internal loop	7.6 ± 0.4	1.0 ± 0.1	3.1 ± 0.9
g15g16 mismatches in 3′ supplementary region	(2.4 ± 0.3) × 10^3	250 ± 40	(1.0 ± 0.3) × 10^3
g2–g16 complementary	(1.1 ± 0.1) × 10^2	11 ± 2	40 ±10
g9–g21 complementary	(2.3 ± 0.2) × 10^3	240 ± 40	(1.0 ± 0.3) × 10^3
Seed only (g2–g8)	(7.3 ± 0.9) × 10^2	100 ± 10	(3.0 ± 0.9) × 10^2

Seed plus g13–g16 3′ supplementary	$(4.1 \pm 0.4) \times 10^2$	43 ± 7	$(1.7 \pm 0.5) \times 10^2$
Extended seed (g2–g10)	$(5.1 \pm 0.6) \times 10^2$	53 ± 9	$(2.1 \pm 0.6) \times 10^2$
Seed plus extended g12–g17 3′ supplementary	19 ± 1	2.0 ± 0.2	8.0 ± 2.0
6 nt seed (g2–g7)	$(1.6 \pm 0.2) \times 10^3$	170 ± 30	$(7 \pm 2) \times 10^2$
Non-complementary luciferase target	$(1.5 \pm 0.1) \times 10^4$	1600 ± 200	$(6.0 \pm 2.0) \times 10^3$
1 GU wobble (g4) in seed	$(2.9 \pm 0.8) \times 10^2$	30 ± 9	$(1.2 \pm 0.5) \times 10^2$
1 GU wobble (g15) in 3′ supplementary region	60 ± 5	6.3 ± 0.9	24 ± 7
2 GU wobbles (g2, g8) in seed	$(3.8 \pm 0.3) \times 10^2$	40 ± 5	$(1.5 \pm 0.5) \times 10^2$
2 GU wobbles (g4, g5) in seed	$(3.5 \pm 0.5) \times 10^3$	370 ± 70	$(1.4 \pm 0.5) \times 10^3$
4 GU wobbles (g2, g4, g5, g8) in seed	$(4.5 \pm 0.1) \times 10^3$	470 ± 60	$(1.8 \pm 0.5) \times 10^3$
***let*–7-programmed, affinity-purified *Mus musculus* AGO2-RISC**			
Target	**K_C (pM)**	**K_{rel}**	**$K_{D(\text{adjusted})}$ (pM)**
28 nt complete complementarity	36 ± 5	1.0 ± 0.2	20 ± 10

g10g11 central internal loop	80 ± 30	2.2 ± 0.8	50 ± 30
g15g16 mismatches in 3′ supplementary region	50 ± 20	1.4 ± 0.6	30 ± 20
g1–g19 complementary	70 ± 10	2.0 ± 0.4	40 ± 20
Seed only (g2–g8)	50 ± 10	1.5 ± 0.3	30 ± 20
Seed plus g13–g16 3′ supplementary	34 ± 5	1.0 ± 0.2	20 ± 10
g4g5 mismatches in seed	$(1.3 \pm 0.7) \times 10^3$	40 ± 20	$(1.0 \pm 0.6) \times 10^3$
Non-complementary luciferase target	$(3.2 \pm 0.8) \times 10^3$	100 ± 30	$(2.0 \pm 1.0) \times 10^3$

Chapter IV: Argonaute Exhibits Selective Tolerance for Mismatches, Bulge and Internal Loops

Disclaimer

This chapter has not been published.

SUMMARY

Argonaute protein divides its small RNA guide into functional domains: anchor, seed, central, 3′ supplementary and tail. The anchor and tail domains secure the guide strand to the Argonaute, while the central and 3′ supplementary regions facilitate target cleavage by a catalytically competent Argonaute. The seed sequence specifies the target and initiates binding. For some types of targets, 3′ supplementary pairing enhances target binding. Accordingly, seed mismatches in the seed or 3′ supplementary domain decrease binding affinity to target. We find that increasing the number of GC base pairs compensates for these mismatches. We also examined the effect on Argonaute activity of nucleotide insertions in the target sequence complementary to each domain of the guide strand. We find that *Drosophila* Argonaute2 exhibits remarkable resilience for insertions in the target sequence complementary to the seed: target cleavage could be detected for a small RNA guide paired with a fully complementary target RNA containing as many as six nucleotides inserted in the seed-matching sequence. In contrast, insertions in the target across from the central or 3′ supplementary domains abolished target cleavage. Further experiments will address how nucleotide insertions perturb the thermodynamic and kinetic properties of Argonaute proteins. Our findings highlight the tolerance of Argonaute for target mismatches or internal nucleotide insertions, and suggest

that Argonaute proteins may bind more types of target sequences than previously appreciated.

RESULTS

A GC-rich Seed Reduces the Assembly of *Drosophila* Ago2-RISC

Seed sequence GC-content has been suggested to influence RISC activity (Doench and Sharp, 2004; Ui-Tei et al., 2008; Garcia et al., 2011). To test this idea directly, we synthesized *let-7*-GC, a variant of the *let-7* siRNA in which four A or U seed nucleotides were changed to G or C (Figure 4.1A). To distinguish the effects of altering seed sequence from altering seed thermodynamic stability, we synthesized *let-7*-IC, in which G residues at positions 4 and 6 were replaced with inosine (I). I:C pairs are slightly less thermodynamically stable than A:U pairs. We also tested two additional *let-7* variants: *let-7*-A, in which position 5 of the *let-7*-GC siRNA was changed from C to A and *let-7*-AI in which the seed sequence changes of *let-7*-IC and *let-7*-A were combined (Figure 4.1A).

let-7-AI and *let-7*-A cleaved their fully matching targets at a rate comparable to that of the parental *let-7* siRNA, whereas *let-7*-GC and *let-7*-IC cleaved more slowly (Figure 4.1B, left panel). When the reactions were performed at lower RISC concentration, differences in cleavage efficiencies were more noticeable: *let-7* had the fastest initial rate followed by *let-7*-AI, *let-7*-A, *let-7*-IC and then *let-7*-GC (Figures 4.S1A and 4.S1B). The passenger strands of the siRNA did not cleave their complementary targets; therefore all siRNAs retained their thermodynamic asymmetry, loading only the *let-7* strand as guide (Figures 4.1B and 4.S1B).

Figure 4.1

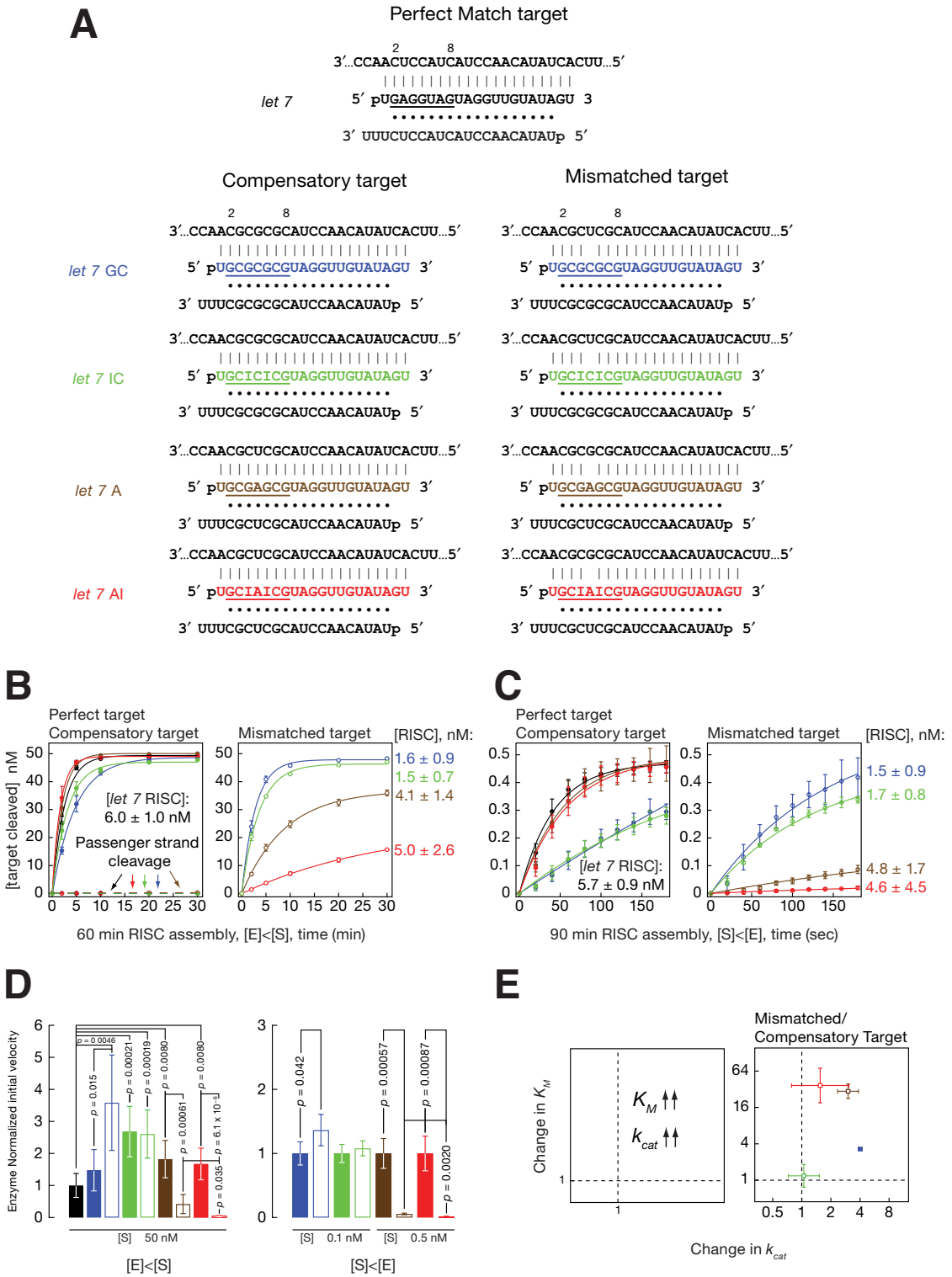

A
Perfect Match target
3′…CCAACUCCAUCAUCCAACAUAUCACUU…5′
let 7
5′ pUGAGGUAGUAGGUUGUAUAGU 3
3′ UUUCUCCAUCAUCCAACAUAUp 5′
Compensatory target
Mismatched target
let 7 GC
let 7 IC
let 7 A
let 7 AI
B
Perfect target
Compensatory target
Mismatched target
[RISC], nM:
[target cleaved] nM
[let 7 RISC]:
6.0 ± 1.0 nM
Passenger strand
cleavage
1.6 ± 0.9
1.5 ± 0.7
4.1 ± 1.4
5.0 ± 2.6
60 min RISC assembly, [E]<[S], time (min)
C
Perfect target
Compensatory target
Mismatched target
[RISC], nM:
[let 7 RISC]:
5.7 ± 0.9 nM
1.5 ± 0.9
1.7 ± 0.8
4.8 ± 1.7
4.6 ± 4.5
90 min RISC assembly, [S]<[E], time (sec)
D
Enzyme Normalized initial velocity
p = 0.0046
p = 0.015
p = 0.00021
p = 0.00019
p = 0.0080
p = 0.00061
p = 0.0080
p = 0.035
p = 6.1 x 10⁻⁵
[S] 50 nM
[E]<[S]
p = 0.042
p = 0.00057
p = 0.00087
p = 0.0020
[S] 0.1 nM
[S] 0.5 nM
[S]<[E]
E
Mismatched/
Compensatory Target
Change in K_M
K_M ↑↑
k_{cat} ↑↑
Change in k_{cat}

Figure Legend 4.1. The seed is sensitive to base pairing disruption

(A) Pairing schematics of *let-7* based seed modified siRNAs. The top shows pairing between the parent *let-7* guide strand and its perfect target. At the bottom are aligned pairings of seed modified siRNAs to compensatory (left) and mismatched (right) targets. (B) Time course target cleavage by Ago2 against compensatory or perfect (left panel, solid circles) and mismatched targets (right panel, open circles) under condition where [E]<[S]. Ago2 concentrations are indicated to the immediate right of the second panel. All data points are from 3 independent experiments ± one standard deviation. (C) Similar to (B) except that experiment was conducted when [S]<[E]. (D) Left panel displays initial velocities for compensatory targets (colored solid bars) and mismatched (open bar) for reactions in (B) normalized to that of the parent *let-7* against its perfect target (solid black bar). The rates were corrected for the concentrations of Ago2 assembled by each siRNA duplex. Right panel displays the initial velocities of mismatched targets normalized to the initial velocities of compensatory targets for reactions shown in (C). (E) Panels depict the change in K_M (vertical axis) against the change in k_{cat} (horizontal axis) for mismatched targets relative to their compensatory targets.

The rate of target cleavage could reflect the affinity of RISC for its target—believed to be determined largely by the seed sequence, RISC concentration or both. To test whether the *let-7* variant siRNAs differed in their abilities to assemble into dAgo2, we determined the concentration of mature dAgo2-RISC for each siRNA. During assembly of dAgo2-RISC, the siRNA passenger strand is destroyed (Matranga et al., 2005; Rand et al., 2005; Leuschner et al., 2006; Kawamata et al., 2009; Yoda et al., 2009). Thus, the difference in abundance of guide and passenger strands after incubation in *Drosophila* embryo lysate reflects the amount of guide loaded into dAgo2 and thus the concentration of mature dAgo2-RISC (Figures 4.S1C and 4.S1D). The concentration of active RISC assembled differed among the siRNA variants; *let-7*-GC and *let-7*-IC assembled only one-quarter as much RISC as the parent *let-7* siRNA (Figures 4.1B and 4.1C). After adjusting the initial rates of cleavage for each siRNA to reflect the concentration of dAgo2-RISC assembled, all siRNAs cleaved their fully matching target RNA at rates comparable to or faster than the parent *let-7* siRNA (Figure 4.1D, left). These data suggest that high GC seed sequences reduce the efficiency of dAgo2-RISC assembly that might provide an explanation for an earlier report that linked high overall GC content to poor siRNA activity (Ui-Tei et al., 2004; Reynolds et al., 2004).

Table 4.1. Michaelis-Menten parameters of *Drosophila* Ago2-RISC (mean ± standard deviation; $n \geq 3$ except *, where $n = 2$).

let–7 siRNA and targets							
	Mismatched target		Compensatory target				
Position (Mismatch)	K_M (nM)	V_{max} (nMs^{-1})	K_M (nM)	V_{max} (nMs^{-1})	$\frac{\Delta k_{cat}}{\Delta K_M}$	Single-turnover Δv_0	$\frac{\Delta k_{cat}}{\Delta v_0}$
let-7-GC (C:U)	169 ± 43	0.237 ± 0.075	52 ± 11	0.0600 ± 0.0071	1.228 ± 0.078	2.01 ± 0.29	2.00 ± 0.31
let-7-IC (C:U)	136 ± 15	0.172 ± 0.074	173 ± 13	0.1800 ± 0.0057	0.89 ± 0.51	1.64 ± 0.29	0.64 ± 0.26
let-7-A (A:G)	940 ± 180	0.42 ± 0.13	33.1 ± 9.8	0.137 ± 0.027	0.101 ± 0.038	0.05517 ± 0.00028	55 ± 13
let-7-AI (A:G)	1090 ± 310	0.37 ± 0.17	43 ± 20	0.41 ± 0.24	0.042 ± 0.040	0.0159 ± 0.0060	98 ± 76
let-7-3′-GC (CC:UU)	23.9 ± 7.0	0.135 ± 0.026	32 ± 14	0.061 ± 0.028	2.5 ± 2.2	0.073 ± 0.011	35 ± 13
let-7-3′-I (CC:UU)	20.7 ± 6.0	0.053 ± 0.024	20.7 ± 6.4	0.132 ± 0.035	0.38 ± 0.12	0.0816 ± 0.0092	4.8 ± 1.1
let-7-3′-A (AA:GG)	28 ± 13	0.086 ± 0.012	18.3 ± 3.5	0.07000 ± 0.00033	0.83 ± 0.29	0.283 ± 0.023	4.36 ± 0.69
let-7-3′-AI (AA:GG)	7.5 ± 3.7	0.0053 ± 0.0019	22.8 ± 8.3	0.200 ± 0.098	0.090 ± 0.019	0.00889 ± 0.00078	3.19 ± 0.57
Parental *let-7*	NA	NA	25.1 ± 5.8	0.189 ± 0.041	NA	NA	NA

The Seed has Limited Tolerance for Mismatches

The extensive interaction between the phosphodiester backbone of the RNA guide and the Argonaute protein in the seed sequence imposes a geometric constraint on the nucleotide. The enclosed Argonaute limits the available space for bases of the guide RNA to flip out of the helix. This may sensitize Argonaute to mismatches in the seed (Parker et al., 2005; Ma et al., 2005; Wang et al., 2008a). A pyrimidine-pyrimidine (g5C:t5U) mismatch in the seed of *let-7*-GC increased K_M 3.3-fold (*p-value* = 9.9 x 10^{-3})—i.e., reduced binding affinity—but simultaneously increased k_{cat} 4.0-fold (*p-value* = 0.015)—i.e., accelerated enzymatic turnover (Figure 4.1E). The increase in k_{cat} suggests that g5C:t5U mismatch amidst the GC-rich seed of *let-7*-GC may enhance product release just as dinucleotide seed mismatches in *let-7* improve turnover (Figure 3.1D).

let-7-IC binds its complementary target with reduced affinity with a K_M of 173 ± 13 nM—6.8-fold greater than the K_M for *let-7* against its perfect matching target. This is comparable to the K_M of 136 ± 15 nM in the presence of g5C:t5U mismatch (Table 4.1). This indicates that *let-7*-IC has as weak a seed pairing as when g5C:t5U mismatch was introduced. Moreover, the introduction of g5C:t5U mismatch did not alter the k_{cat} of *let-7*-IC (Figure 4.1E). Taken together, the modest effect of g5C:t5U mismatch renders a $\Delta k_{cat}/\Delta K_M$ value of 0.89 ± 0.51. In addition, the initial velocities normalized to enzyme concentrations indicate that *let-7*-IC is more proficient at target cleavage than the parent *let-7*: the weaker

seed pairing of *let-7*-IC confers better turnover of dAgo2-RISC (Figure 4.1D). In all, a single pyrimidine-pyrimidine g5C:t5U mismatch in *let-7*-GC and *let-7*-IC were well accommodated by dAgo2.

With a purine-purine (g5A:t5G) mismatch, Michaelis-Menten analysis indicates a larger decrease in binding affinity than a g5C:t5U mismatch. For *let-7*-A, K_M increased by 30-fold (*p-value* = 1.3×10^{-6}) and for *let-7*-AI, K_M increased by 37-fold (*p-value* = 1.3×10^{-5}) relative to their compensatory targets (Figure 4.1E). In contrast, for *let-7*-A, k_{cat} increased by 3.0-fold (*p-value* = 7.4×10^{-4}) and for *let-7*-AI, k_{cat} did not change significantly (Figure 4.1E). The 3.0-fold increase in k_{cat} for *let-7*-A underscores the earlier notion that mismatch in the seed facilitates product dissociation (Figure 3.1D). In the case of *let-7*-AI, despite its fully complementary seed pairing, target binding with inosines:cytosine base pairing is weak and may be efficient at product release and thus g5A:t5G mismatch did not further increase k_{cat} (Figure 4.1E).

Having inosines (*let-7*-AI) in lieu of guanosines (*let-7*-A) exacerbated impaired target cleavage caused by a g5A:t5G mismatch: *let-7*-AI cleaves 3.7-fold slower than *let-7*-A (*p-value* = 0.0020) when [S] < [E] and 8.4-fold slower (*p-value* = 0.035) when [E] << [S] (Figures 4.1B–4.1D). The presence of G:C pairs that flank the g5A:t5G mismatch reduces the disruption to seed base pairing. Consequently, *let-7*-A has a $\Delta k_{cat}/\Delta K_M$ value of 0.1 ± 0.04 whereas *let-7*-AI has a lower $\Delta k_{cat}/\Delta K_M$ value of 0.04 ± 0.04. We conclude that dAgo2 tolerates

purine:purine mismatches in the seed sequence less well than pyrimidine:pyrimidine mismatches and that more stable G:C base pairs flanking the site of the mismatch better compensate for non-complementary positions than do weaker A:U or I:C pairs.

More Stable 3′ Base Pairing Slows dAgo2 Turnover

Target cleavage by dAgo2 requires nearly continuous base pairing from the seed to guide position 17. Early experiments in plants and animals suggested that extensive guide:target base pairs in the 3′ half of a miRNA or siRNA slowed target cleavage, likely because dissociation of the 3′ fragment of the target from dAgo2 was rate determining (Haley and Zamore, 2004; Tang et al., 2003). Our detailed analysis supports this idea. We analyzed four *let-7* variants bearing changes that varied the pairing strength of the nucleotides participating in 3′ supplementary pairing between the siRNA and its target RNA. In *let-7*-3′-GC, G and C replaced six A or U nucleotides from g13–g19; *let-7*-3′-I was the same as *let-7*-3′-GC except that nucleotides g14 and g17 were I; *let-7*-3′-A was the same as *let-7*-3′-GC except that nucleotides g15 and g16 were A; *let-7*-3′-AI was the same as *let-7*-3′-GC except that nucleotides g14 and g17 were I and g15 and g16 were A (Figure 4.2A). All four siRNAs assembled roughly equal amounts of dAgo2-RISC (Figures 4.2B and 4.2C).

Figure 4.2

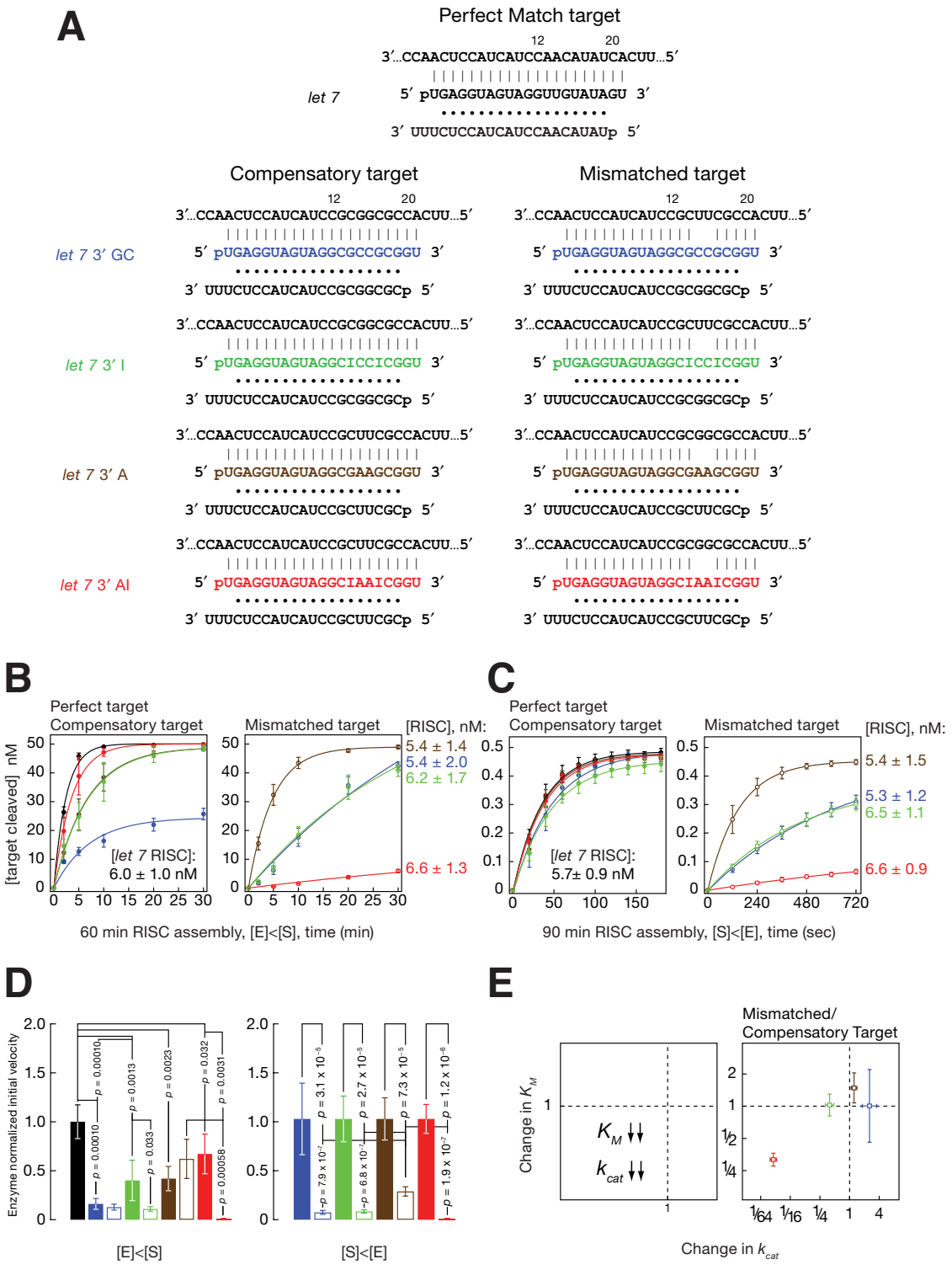

A
Perfect Match target
12 20
3′...CCAACUCCAUCAUCCAACAUAUCACUU...5′
let 7
5′ pUGAGGUAGUAGGUUGUAUAGU 3′
3′ UUUCUCCAUCAUCCAACAUAUp 5′
Compensatory target
Mismatched target
let 7 3′ GC
let 7 3′ I
let 7 3′ A
let 7 3′ AI
B
Perfect target
Compensatory target
Mismatched target
[RISC], nM:
5.4 ± 1.4
5.4 ± 2.0
6.2 ± 1.7
6.6 ± 1.3
[target cleaved] nM
[let 7 RISC]:
6.0 ± 1.0 nM
60 min RISC assembly, [E]<[S], time (min)
C
Perfect target
Compensatory target
Mismatched target
[RISC], nM:
5.4 ± 1.5
5.3 ± 1.2
6.5 ± 1.1
6.6 ± 0.9
[let 7 RISC]:
5.7± 0.9 nM
90 min RISC assembly, [S]<[E], time (sec)
D
Enzyme normalized initial velocity
p = 0.00010
p = 0.00010
p = 0.0013
p = 0.033
p = 0.0023
p = 0.032
p = 0.0031
p = 0.00058
p = 3.1 x 10−5
p = 7.9 x 10−2
p = 2.7 x 10−5
p = 6.8 x 10−2
p = 7.3 x 10−5
p = 1.2 x 10−6
p = 1.9 x 10−7
[E]<[S]
[S]<[E]
E
Mismatched/
Compensatory Target
Change in KM
KM
kcat
Change in kcat

Figure Legend 4.2. GC-rich base pairing in 3′ region of siRNA limits turnover of Ago2-RISC

(A) Pairing schematics of *let-7* based 3′ supplementary region modified siRNAs. The top shows pairing between the parent *let-7* guide strand and its perfect target. The bottom depicts aligned pairings of modified siRNAs to compensatory (left) and mismatched (right) targets. (B) Time course of target cleavage by Ago2-RISC of fully matched (left panel, filled circles) and mismatched targets (right panel, opened circles) under conditions where [E]<[S]. RISC concentrations are indicated to the right of the second panel. All data points are derived from 3 independent experiments ± one standard deviation. (C) Similar to (B) except that experiment is conducted when [S]<[E]. (D) Left panel displays initial velocities for compensatory (colored solid bars) and mismatched (open bar) targets for reactions in (B) normalized to that of the parent *let-7* against its perfect target (solid black bar). The rates were corrected for the concentrations of Ago2 assembled by each siRNA duplex. Right panel displays the initial velocities of mismatched targets normalized to the initial velocities of compensatory targets for reactions shown in (C). (E) Panels depict the change in K_M (vertical axis) against the change in k_{cat} (horizontal axis) for mismatched targets relative to their compensatory targets.

The amount of product produced per unit time per molecule of enzyme for the siRNAs against their compensatory targets when [E] << [S] revealed rates that indeed differ based on the strength of 3′ supplementary pairing. The rates, relative to that of the parent *let-7*, were 0.16 for *let-7*-3′-GC (*p-value* = 1.0×10^{-4}), 0.40 for *let-7*-3′-I (*p-value* = 1.3×10^{-3}), 0.41 for *let-7*-3′-A (*p-value* = 2.3×10^{-3}) and 0.67 for *let-7*-3′-AI (*p-value* = 0.032; Figure 4.2D, left panel, solid colored bars). A similar order of cleavage competence among the siRNAs was observed from initial rates obtained with less dAgo2 assembled (Figures 4.S2A and 4.S2B). From Michaelis-Menten kinetics, the K_M values among the parent *let-7* and modified siRNAs were comparable, indicating that the disparate cleavage rates were not caused by the differences in target binding affinities (Table 4.1). Under conditions where [S] < [E], however, the initial rates for all siRNAs were broadly similar (Figure 4.2C, left panel). This indicates that the first cycle of event from target binding until and including the catalytic step proceeded at similar rates across all siRNAs, unaffected by the strength of base pairing in the 3′ region. Taken together, the data point to steps after catalysis such as product release or enzyme regeneration that contributed to the different cleavage rates observed under multiple turnover cleavage reactions. This is best exemplified by *let-7*-3′-GC being the least efficient siRNA when [E]<<[S] and the simplest explanation is that extensive GC base pairing in the 3′ region of the

siRNA guide binds the cleaved product tightly, slows its released and limits recycling of dAgo2.

Supporting this view, weakening base pairing in the 3′ region of the siRNA guide should facilitate product release and “rescue” the slow cleavage under multiple turnover reaction, which we tested with mismatch at g15g16. Expectedly, g15g16:t15t16 mismatch in all 3′ variant siRNAs reduced the cleavage rates under substrate limiting condition that promote single turnover reaction: an observation that agrees with earlier findings that guide-target pairing in the 3′ region is necessary for target cleavage (Figure 3.1D; Haley and Zamore, 2004; Schwarz et al., 2006). Accordingly, the initial rates for all modified siRNAs with g15g16:t15t16 mismatch were reduced significantly in comparisons with rates for their compensatory targets (Figures 4.2C and 4.2D, right panels). In contrast, the initial rates for all siRNAs with g15g16:t15t16 mismatch improved—at varying degrees—relative to their compensatory targets when the ratio of substrate to enzyme concentration increased to allow for multiple cycles of target cleavage (Figures 4.2C–4.2D, 4.S2A and 4.S2B).

To elaborate, the g15g16:t15t16 mismatch in *let-7*-3′-GC reduced the initial rate by 14-fold (*p-value* = 3.1 x 10^{-5}) when [E]>[S]. Then, when [E]<[S], the initial rates between mismatched and compensatory targets became comparable and when [E]<<[S], the initial rate for mismatched target was 3.9-fold greater (*p-value* = 0.0011) than that of the compensatory target (Figure

4.S2C). Moreover, Michaelis-Menten analysis of *let-7*-3′-GC with g15g16:t15t16 mismatch showed no change in K_M while k_{cat} increased by 2.5-fold (*p-value* = 0.0026) relative to its compensatory target (Figure 4.2E and 4.S2D–4.S2F). Taken together, g15g16:t15t16 mismatch in *let-7*-3′-GC disrupted catalysis and reduced the cleavage rate. However, under multiple cycle of cleavage, the reduced rate was offset by enhanced dAgo2-RISC turnover. For maximal activity under enzyme limited conditions, dAgo2 must satisfy pairing in the 3′ region sufficient for cleavage and must not compromise efficient enzymatic turnover. Finally, between *let-7*-GC that pairs with a GC-rich seed and *let-7*-3′-GC that pairs with a GC-rich 3′ supplementary site, *let-7*-3′-GC exhibits a slower steady state rate (Figures 4.S2G–4.S2I). In fact, the burst kinetics of *let-7*-GC are comparable to that of the parent *let-7*. Taken together, these data suggest that GC-rich 3′ supplementary pairing but not GC-rich seed pairing limit turnover of dAgo2-RISC.

A Stable 3′ Base Pairing Accommodates Purine-Purine Mismatches

Given that strong flanking base pairs can partially stabilize a purine-purine mismatch in the seed, we investigate whether the same is also true in the 3′ supplementary binding site. Under both conditions where [E]<<[S] and [S]<[E], *let-7*-3′-AI with purine-purine mismatches cleaved 6.3–14-fold slower than *let-7*-3′-GC and *let-7*-3′-I with pyrimidine-pyrimidine mismatches at g15g16:t15t16

(Figures 4.2B and 4.2C). Michaelis Menten kinetics indicated that k_{cat} for *let-7*-3′-AI is 31-fold lower than the k_{cat} of *let-7*-3′-GC and 11-fold lower than the k_{cat} of *let-7*-3′-I with g15g16:t15t16 mismatch (Figure 4.2E). Relative to its compensatory target, g15g16:t15t16 mismatch in *let-7*-3′-AI diminished the cleavage rate by 74-fold (*p-value* = 3.1 x 10^{-3}) when [E]<<[S] (Figures 4.2B and 4.2D) and 110-fold (*p-value* = 1.2 x 10^{-6}) when [S]<[E] (Figures 4.2C and 4.2D). In comparison, *let-7*-3′-A tolerates the same purine-purine g15g16:t15t16 mismatch and exhibited k_{cat} that is only1.6-fold lower than the k_{cat} of *let-7*-3′-GC but 1.9-fold higher than the k_{cat} of *let-7*-3′-I (Figures 4.2B–4.2E). Relative to its compensatory target, g15g16:t15t16 mismatch in *let-7*-3′-A gave a comparable cleavage rate when [E]<<[S] but reduced cleavage rate 3.6-fold (*p-value* = 7.3 x 10^{-5}) when [S]<[E] (Figures 4.2B–4.2D). The ability to handle the g15g16:t15t16 mismatch by *let-7*-3′-A could be attributed to the imino hydrogen bonded G:A base pairs flanked at both ends by Watson-Crick G:C base pairs (Xia et al., 1997).

Indeed, between *let-7*-3′-A and *let-7*-3′-AI that differed only in the base pairs flanking the g15g16:t15t16 mismatch, the stronger G:C base pairs bestow on *let-7*-3′-A 68-fold (*p-value* = 5.8 x 10^{-4}) higher initial rates when [E]<<[S] and 39-fold (*p-value* = 1.9 x 10^{-7}) higher activity when [S]<[E] than the weaker I:C base pairs in *let-7*-3′-AI (Figure 4.2D). For that reason, K_M or k_{cat} did not change for *let-7*-3′-A whereas k_{cat} declined 35-fold (*p-value* = 0.026) for *let-7*-3′-AI

(Figure 4.2E). Consequently, *let-7*-3′-A registered a $\Delta k_{cat}/\Delta K_M$ value of 0.83 ± 0.29 while *let-7*-AI exhibited almost 10-fold lower $\Delta k_{cat}/\Delta K_M$ value of 0.090 ± 0.019 (Table 4.1). Like the seed, the more stable G:C base pairs flanking the site of the mismatch better compensate for non-complementary positions than do the weaker I:C pairs.

In contradiction to earlier findings that the 3′ region serves as ancillary binding site, none of the modified siRNAs that contained g15g16 target mismatches led to significant increased in K_M (Figure 4.2E). A plausible explanation is that the surrounding G:C base pairs, that extend 3–4 nucleotides on both sides of the mismatches, can stabilize pairing and prevent any minimal increase in K_M. The adjoining G:C base pairs however were not sufficient to avert the decrease in catalysis, which most likely resulted from perturbed pairing geometry at g15g16:t15t16. Taken together, G:C base pairing in the 3′ region of the siRNA guide protects against a decrease in target binding affinity and buffers against catalytic insult due to mismatches.

Fly Ago2 Tolerates Bulge Loops in the Target Selectively

Unpaired nucleotides on one strand of a double helix constitute a bulge loop. Argonaute protein interacts with the seed region of the guide strand more extensively than with the target. Accordingly Argonaute proteins tolerate a bulge loop on the target but not on the guide strand in the seed region (Wang et al., 2008a). To study the effect of a bulge loop on Argonaute activity in detail, we

Figure 4.3

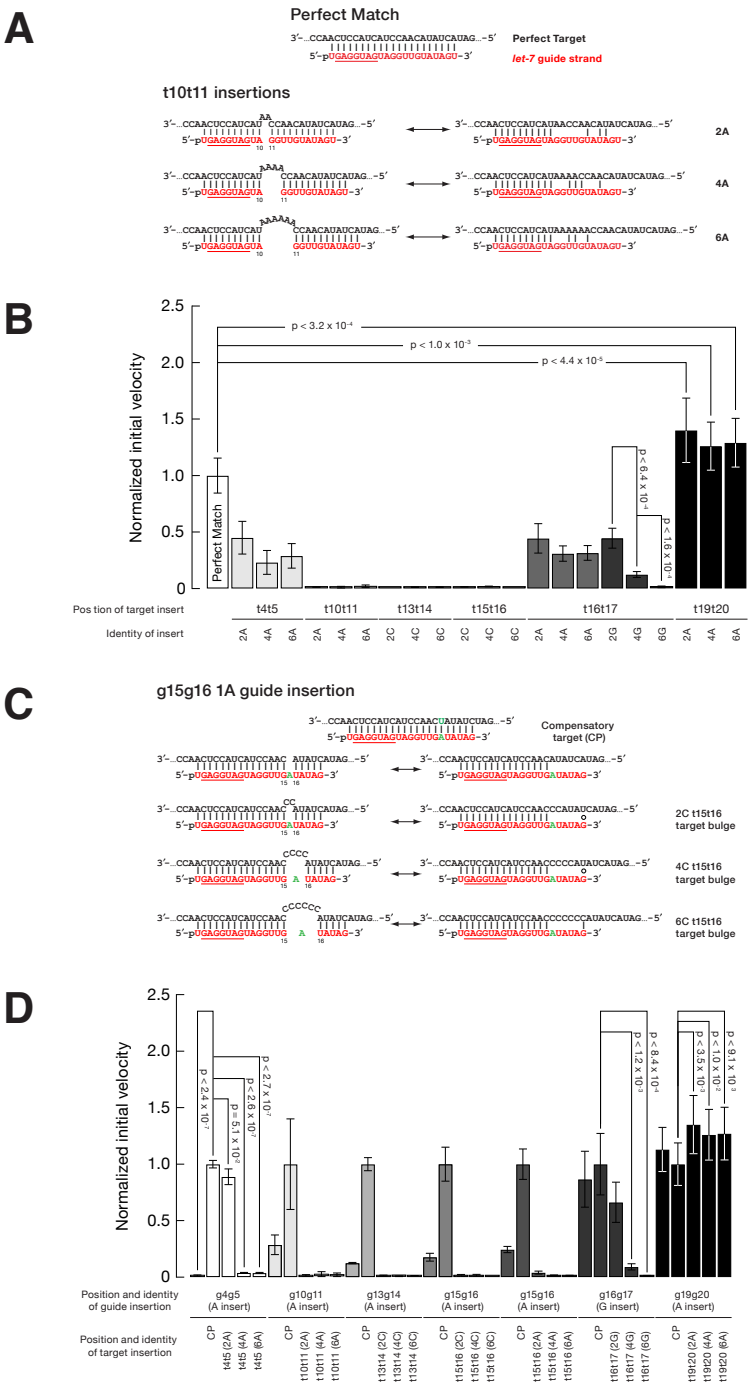

A
Perfect Match
Perfect Target
let-7 guide strand
t10t11 insertions
2A
4A
6A
B
Normalized initial velocity
Perfect Match
p < 3.2 x 10⁻⁴
p < 1.0 x 10⁻³
p < 4.4 x 10⁻⁵
p < 6.4 x 10⁻⁴
p < 1.6 x 10⁻³
Pos tion of target insert
Identity of insert
t4t5
t10t11
t13t14
t15t16
t16t17
t19t20
C
g15g16 1A guide insertion
Compensatory target (CP)
2C t15t16 target bulge
4C t15t16 target bulge
6C t15t16 target bulge
D
Normalized initial velocity
p < 2.4 x 10⁻⁷
p = 5.1 x 10⁻³
p < 2.6 x 10⁻⁷
p < 2.7 x 10⁻⁷
p < 1.2 x 10⁻³
p < 8.4 x 10⁻⁴
p < 3.5 x 10⁻³
p < 1.0 x 10⁻²
p < 9.1 x 10⁻³
Position and identity of guide insertion
g4g5 (A insert)
g10g11 (A insert)
g13g14 (A insert)
g15g16 (A insert)
g15g16 (A insert)
g16g17 (G insert)
g19g20 (A insert)
Position and identity of target insertion
CP
t4t5 (2A)
t4t5 (4A)
t4t5 (6A)
t10t11 (2A)
t10t11 (4A)
t10t11 (6A)
t13t14 (2C)
t13t14 (4C)
t13t14 (6C)
t15t16 (2C)
t15t16 (4C)
t15t16 (6C)
t15t16 (2A)
t15t16 (4A)
t15t16 (6A)
t16t17 (2G)
t16t17 (4G)
t16t17 (6G)
t19t20 (2A)
t19t20 (4A)
t19t20 (6A)

Figure Legend 4.3. Ago2-RISC tolerates bulge loop in guide and target selectively

(A) Pairing schematics of *let-7* to its perfect target (top). The bottom depicts pairings of *let-7* to target with 2, 4 or 6 nucleotides bulge loops between t10t11. There exist at least 2 possible pairing profiles between *let-7* and targets with a bulge loop. (B) Normalized target cleavage of Ago2-RISC against targets with bulge loops at various positions. (C) Pairing schematics of *let-7* with an adenosine bulge between g15g16 to its compensatory target (top) and to targets having 2, 4 or 6 nucleotides bulge loops between t15t16. (D) Similar to (B) except against *let-7* guide with adenosine bulge. Target cleavage is conducted under conditions where [E]<[S]. All data points are derived from 3 independent experiments ± one standard deviation.

introduced a 2, 4 or 6-nucleotide bulge loops at various positions on the target (Figures 4.3A and 4.3B). The cleavage activities of dAgo2 under multiple turnover condition where [E]<[S] are scored for various target bulge loops and normalized against that of a perfect target (Figure 4.3B).

Bulge loops placed between t10 to t15 inhibit dAgo2 cleavage. A bulge loop in the middle of the seed region or between t16t17 was tolerated but reduced dAgo2 activities by ~2-fold (Figure 4.3B). In contrast, a bulge loop positioned between t19t20 did not affect target cleavage by dAgo2. In fact, a modest but significant increase in target cleavage was observed when compared to the perfect target (*p-value* < 1.0×10^{-3}). This agrees with earlier observations that mismatches in the tail region of the guide promote turnover (Figure 3.1D; Tang et al., 2003; Haley and Zamore, 2004; Wee et al., 2012). The effects of 2, 4 or 6 nucleotides bulge loops placed between t4t5, t16t17 and t19t20 on target cleavage were similar (Figure 4.3B). In sum, dAgo2 tolerates bulge loops in the seed, 3′ supplementary and tail regions but not in the central region.

Fly Ago2 Demonstrates Limited Tolerance to Single Bulge in the Guide Strand

A single adenosine bulge positioned between g4–g16 impaired or reduced dAgo2 cleavage activities to ~20% that of a perfect target (Figure 4.3C and 4.3D). In the presence of a compensatory target, full cleavage activity of dAgo2

was restored. A guide bulge in the tail region between g16g17 and g19g20, however, had no effect on dAgo2 cleavage (Figure 4.3D). For guide bulge that reduced dAgo2 activities, pairing to targets with bulge loops across the guide bulge, with the exception of the seed region, did not rescue cleavage activity (Figure 4.3D). In fact, increasing the bulge loop from 2 to 6 nucleotides between t16t17 resulted in incremental decrease in dAgo2 activity most likely an effect of increased stearic hindrance.

Placing a two-nucleotide bulge loop across the single adenosine bulge between g4g5 restored dAgo2 activity almost to that of a perfect target (Figure 4.3D). The pairing to a target with two-nucleotide bulge we reasoned shifted the single nucleotide bulge to the target strand that was originally on the guide strand when paired to a perfect target. Increasing the target bulge loop to 4 and 6 nucleotides did not restore dAgo2 cleavage most likely due to stearic effect (Figure 4.3D).

Bulge Loop in Target did not Alter the Molecular Ruler that Defines Cleavage Site

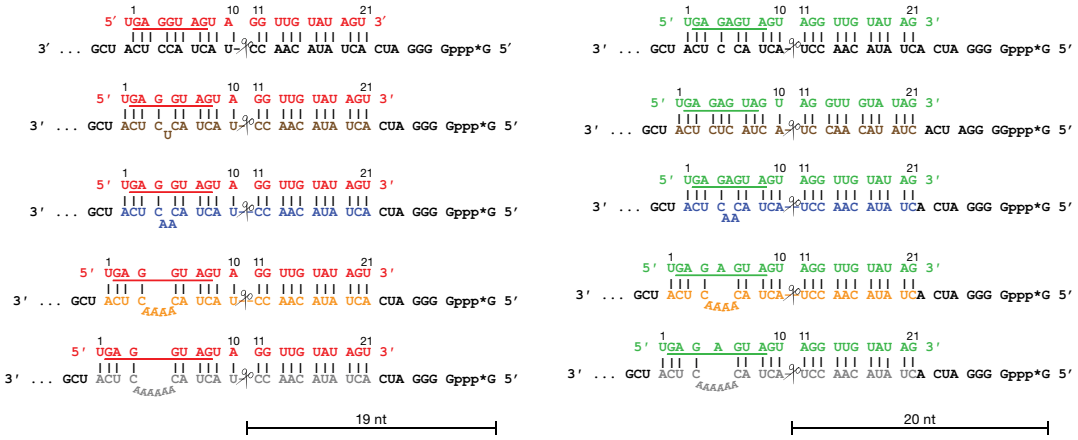
Fly Ago2 cleaves its target across from g10g11 of the guide. The 5′ end of the guide strand defines the start of the molecular ruler (Elbashir et al., 2001c; Haley and Zamore, 2004). The 5′-phosphate secures the guide on the Argonaute scaffold and without the 5′-phosphate, dAgo2 exhibits less precision in where it cuts its target (Rivas et al., 2005). To find out if a bulge loop in the seed region

affects the cleavage position, we subjected a series of short 5′ P^{32} cap-labeled targets with incremental bulge nucleotides to cleavage by dAgo2. These targets and their products are resolved in a denaturing gel to identify the cleavage sites (Figure 4.4A and 4.4B).

Amazingly, dAgo2 cleaved at the same site—across from g10 and g11 of the guide strand—despite the targets having 2, 4 or 6 nucleotides inserted between t4t5 (Figure 4.4B). As a trade-off, cleavage activity is diminished as the bulge was increased from 2 to 6 nucleotides (Figure 4.4B). Expectedly, an adenosine bulge in the guide strand between g4g5 abrogates dAgo2 cleavage. In the presence of target with bulge loop, however, we were able to rescue target cleavage. Likewise, cleavage activities were decreased with increased loop size (Figure 4.4B). As proposed earlier, a target bulge converts the detrimental guide bulge to an innocuous mismatched pair and permits cleavage (Figure 4.3D). Furthermore, the adenosine bulge in the guide strand meant that an additional nucleotide is included within the region of the molecular ruler. Therefore, the predicted cleavage site is now shifted 1 nt downstream of the original site. Indeed, with a bulge loop in the guide, dAgo2 generated a product that is 20 nt in length—one nt longer than the canonical product (Figures 4.4A and 4.4B).

Figure 4.4

A


19 nt
20 nt

B

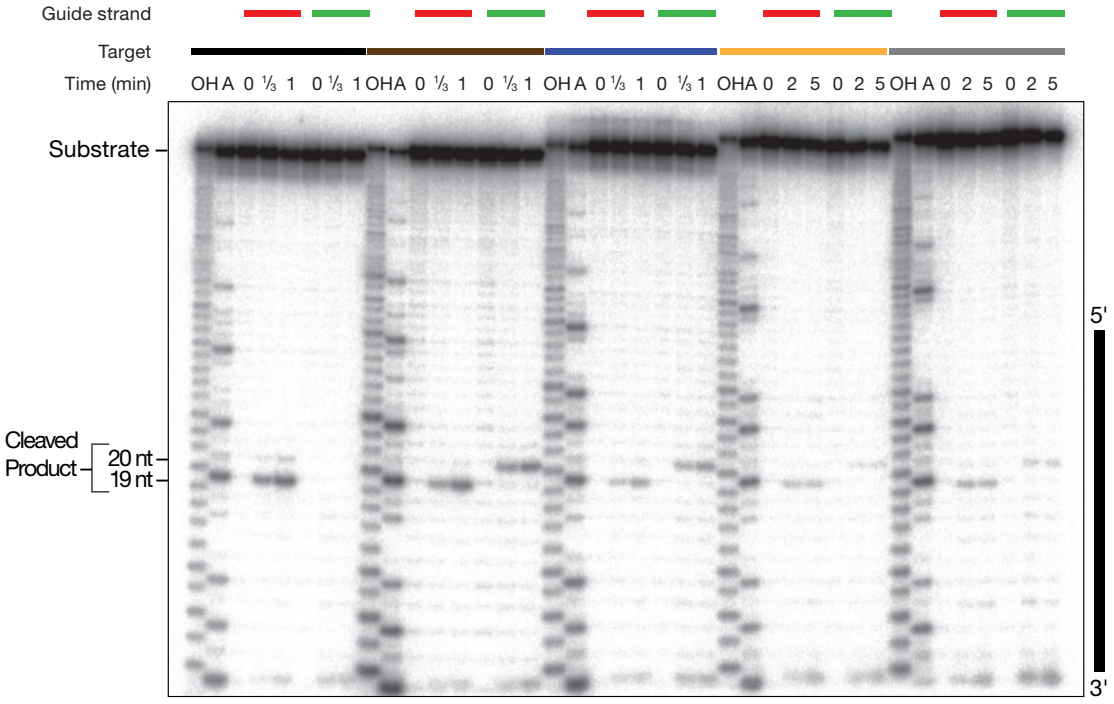

Guide strand
Target
Time (min)
OH A 0 ⅓ 1 0 ⅓ 1 OHA 0 ⅓ 1 0 ⅓ 1 OH A 0 ⅓ 1 0 ⅓ 1 OHA 0 2 5 0 2 5 OH A 0 2 5 0 2 5
Substrate
Cleaved Product
20 nt
19 nt
5'
3'

Figure Legend 4.4. Ago2-RISC retains the same cleavage site with target bulge loop in the seed region

(A) Pairing schematics of *let-7* to target with 2, 4 or 6 nucleotides bulge loops between t4t5 (left). On the right, it shows pairing between *let-7* with an adenosine bulge between g4g5 to target with 2, 4 or 6 nucleotides bulge loops between t4t5. (B) Denaturing polyacrylamide gel that shows the cleavage products for guide:target interactions depicted in (A) after 30 sec or 1 min. OH: Base hydrolysis, A: RNaseA digestion. On the right of the gel, it shows where the guide binds the target.

Bulge Loop in the Seed Region of the Target has Minimal Effect on K_M

The seed initiates target binding and mismatches in the seed reduces the binding affinity of Argonaute protein for its target (Figure 3.1C; Haley and Zamore, 2004; Bartel, 2009; Ameres et al., 2007; Wee et al., 2012). *Let-7* microRNA and members of the family were shown to bind imperfectly to their endogenous target in *Caenorhabditis elegans*, forming mismatches, bulges and loopouts (Reinhart et al., 2000; Slack et al., 2000; Lin et al., 2003; Vella et al., 2004; Grosshans et al., 2005; Johnson et al., 2005). To understand how bulges in the seed will affect target binding by Argonaute proteins, we measured the Michaelis-Menten constant when a 2-nucleotide bulge loop was introduced between t4t5 and t15t16 (Figure 4.5A).

A target bulge loop at t4t5 increased K_M by 2.1-fold (p-value = 0.018) and decreased k_{cat} by 3.3-fold (p-value = 4.1×10^{-4}) relative to the perfect target (Figure 4.5A). When the target bulge loop at t4t5 is placed across the guide bulge loop at g4g5, K_M increased by 6.1-fold, (p-value = 1.1×10^{-3}) whereas k_{cat} showed no significant change with respect to the compensatory target (Figure 4.5A). We do not know why a different effect on k_{cat} was observed between seed target bulge with and without the guide bulge. We speculate that the absence of an effect on k_{cat} might be offset by an increased in product release when the occurrence of the guide and target bulge created a mismatched pairing in the seed (Figures 3.1D and 4.5A). Finally, a target bulge loop placed between t16t17

Figure 4.5

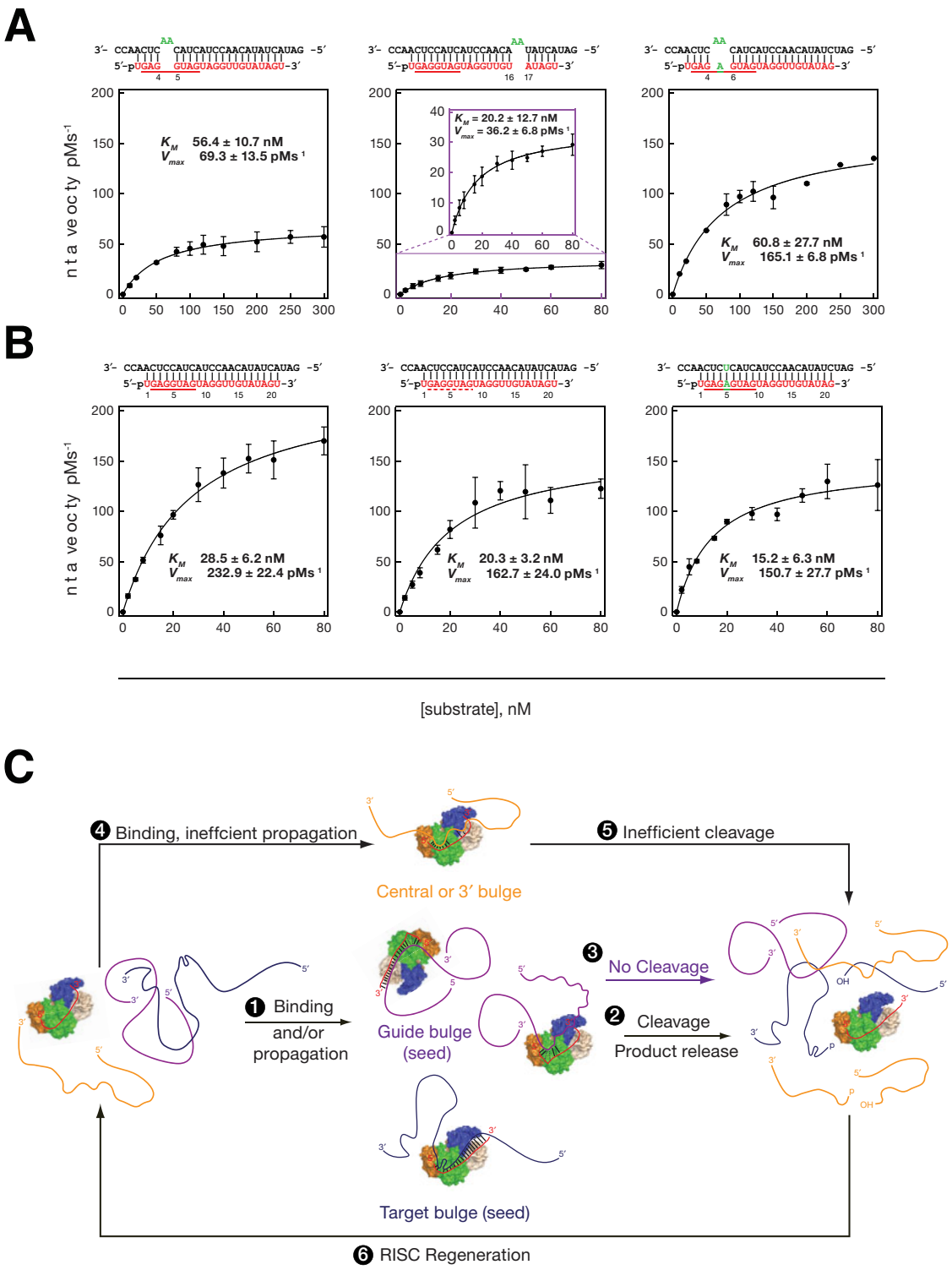

A
B
C
3′– CCAACUC AA CAUCAUCCAACAUAUCAUAG –5′
5′–pUGAG GUAGUAGGUUGUAUAGU–3′
K_M 56.4 ± 10.7 nM
V_{max} 69.3 ± 13.5 pMs⁻¹
3′– CCAACUCCAUCAUCCAACA AA UAUCAUAG –5′
5′–pUGAGGUAGUAGGUUGU AUAGU–3′
K_M = 20.2 ± 12.7 nM
V_{max} = 36.2 ± 6.8 pMs⁻¹
3′– CCAACUC AA CAUCAUCCAACAUAUCUAG –5′
5′–pUGAG A GUAGUAGGUUGUAUAG–3′
K_M 60.8 ± 27.7 nM
V_{max} 165.1 ± 6.8 pMs⁻¹
3′– CCAACUCCAUCAUCCAACAUAUCAUAG –5′
5′–pUGAGGUAGUAGGUUGUAUAGU–3′
K_M 28.5 ± 6.2 nM
V_{max} 232.9 ± 22.4 pMs⁻¹
3′– CCAACUCCAUCAUCCAACAUAUCAUAG –5′
5′–pUGAGGUAGUAGGUUGUAUAGU–3′
K_M 20.3 ± 3.2 nM
V_{max} 162.7 ± 24.0 pMs⁻¹
3′– CCAACUCUCAUCAUCCAACAUAUCUAG –5′
5′–pUGAGAGUAGUAGGUUGUAUAG–3′
K_M 15.2 ± 6.3 nM
V_{max} 150.7 ± 27.7 pMs⁻¹
n t a ve oc ty pMs⁻¹
[substrate], nM
❹ Binding, ineffcient propagation
❺ Inefficient cleavage
Central or 3′ bulge
❶ Binding and/or propagation
Guide bulge (seed)
❸ No Cleavage
❷ Cleavage
Product release
Target bulge (seed)
❻ RISC Regeneration

Figure Legend 4.5. Target bulge loop increases K_M and reduces k_{cat} moderately

(A) Michaelis-Menten kinetics of *let-7* to target with 2 nucleotides bulge loop between t4t5 (left panel); at t16t17 (middle panel) and t4t5 with a guide bulge at g4g5 (right panel). (B) Michaelis-Menten kinetics of *let-7* to its perfect or compensatory target. All data points are derived from 3 independent experiments ± one standard deviation. (C) Model of how fly Ago2-RISC finds, binds, cleaves and releases its bulge loop targets.

exhibited no change in K_M but k_{cat} dropped by 4.5-fold (p-value = 9.2×10^{-4}) (Figure 4.5A). In sum, dAgo2 endures a 2-nucleotide target bulge in the seed with minimal effects on its target binding affinity and cleavage rate. By contrast, placing a 2-nucleotide target bulge in the 3′ supplementary region between t16t17 did not perturb target binding but it had a strong effect on target cleavage activity. This agrees with earlier observation that 3′ supplementary pairing generates the ideal geometry for target cleavage to transpire (Figure 3.1D; Haley and Zamore, 2004; Schwarz et al., 2006; Wee et al., 2012).

Conclusion

The extent of target complementarity of guide to target RNA determines how well Argonaute proteins find, bind and regulate the target RNA. In this chapter, we show that the presence of GC base pairs that create a more stable guide:target mRNA interaction moderate the effect of mismatches that increase K_M (lower binding affinity) and reduce k_{cat} (lower target cleavage efficiency) of dAgo2 for its target mRNA. The presence of GC base pairs, however, can also reduce the rate of product release and therefore the turnover of dAgo2. In addition, we also demonstrate that dAgo2 tolerates a bulge loop introduced in the target but not in the guide. This observation agrees with structural data that show more extensive interactions between Argonaute protein and the guide RNA than between Argonaute protein and the target mRNA (Wang et al., 2008a; Elkayam et al., 2012). A bulge loop placed in the center of the seed region in the

target was tolerated by dAgo2 that exhibits a modest 2-fold increase in K_M but a more severe 3.5-fold decrease in k_{cat}. Furthermore, dAgo2 maintains the precise cleavage site on the target in the presence of the seed bulge loop. Future experiments will need to determine the mechanism how a bulge loop in the seed can affect the rate of catalysis that occurs in the central region of the guide (Wee et al., 2012). Furthermore, we need to address how the guide:target duplex is placed in the Argonaute protein in the presence of the bulge loop. Is the bulge loop extruded out into the solvent exposed side of the Argonaute protein? Will this solvent exposed bulge loop be accessible? The answer to the preceding will have a greater implication given that bulge nucleotide has been reported to function as protein:RNA contact site (Peattie et al., 1981).

Figure 4.S1

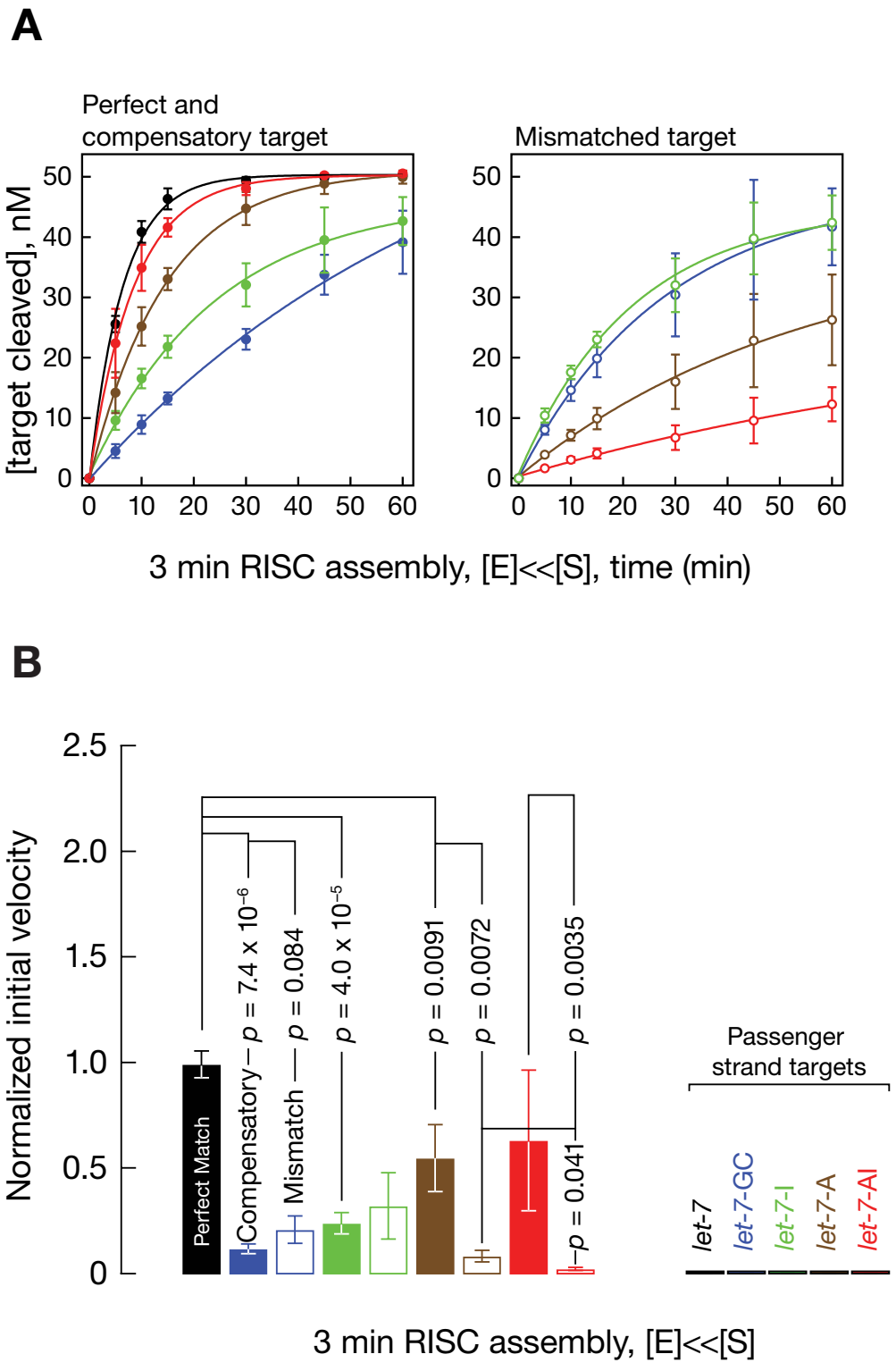

A
Perfect and
compensatory target
Mismatched target
[target cleaved], nM
0
10
20
30
40
50
0
10
20
30
40
50
60
3 min RISC assembly, [E]<<[S], time (min)
B
Normalized initial velocity
0
0.5
1.0
1.5
2.0
2.5
Perfect Match
Compensatory
Mismatch
p = 7.4 x 10−6
p = 0.084
p = 4.0 x 10−5
p = 0.0091
p = 0.0072
p = 0.0035
p = 0.041
Passenger
strand targets
let-7
let-7-GC
let-7-I
let-7-A
let-7-AI
3 min RISC assembly, [E]<<[S]

Figure 4.S1

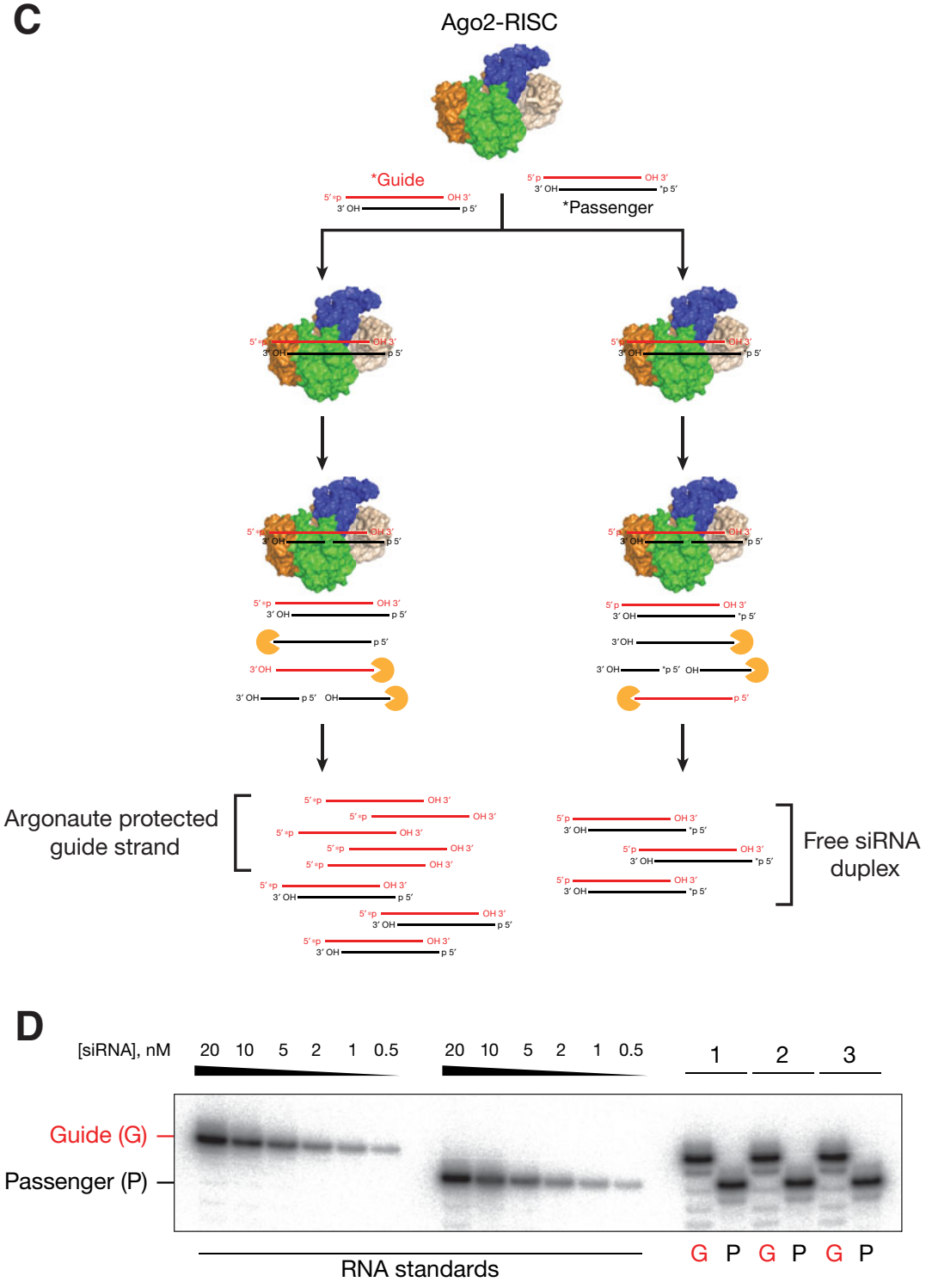

C
Ago2-RISC
*Guide
*Passenger
Argonaute protected guide strand
Free siRNA duplex
D
[siRNA], nM
20 10 5 2 1 0.5
20 10 5 2 1 0.5
1 2 3
Guide (G)
Passenger (P)
RNA standards
G P G P G P

Figure Legend 4.S1. siRNA with high GC content in the seed buffers the effect of mismatches but are selected against for RISC loading

(A) Time course cleavage of compensatory, perfect (left panel, solid circles) and mismatched targets (right panel, opened circles). Data points are means ± standard deviations for 3 independent experiments. (B) Bar graph shows initial velocities for (A) normalized to rate of *let-7* against its perfect target (black bar). Normalized rates for compensatory targets are solid bars and for mismatched targets, opened bars. To the right are normalized cleavage rates by passenger strands against their compensatory targets. The data represent averages of three independent experiments ± standard deviation. *P*-values were calculated using two-tailed Student's *t*-test assuming equal variances. (C) Ago2-RISC was assembled with siRNA duplexes that were either guide or passenger strand labeled at the 5′ end. Guide strand incorporated into Ago2-RISC was stabilized while cleaved passenger strand was degraded. Quantifying the difference in abundance between guide and passenger strands gives guide-loaded active Ago2-RISC concentration. (D) On the left are titrated standards and on the right are gel-resolved 5′ end labeled guide and passenger strands from 3 independent assembly reactions.

Figure 4.S2

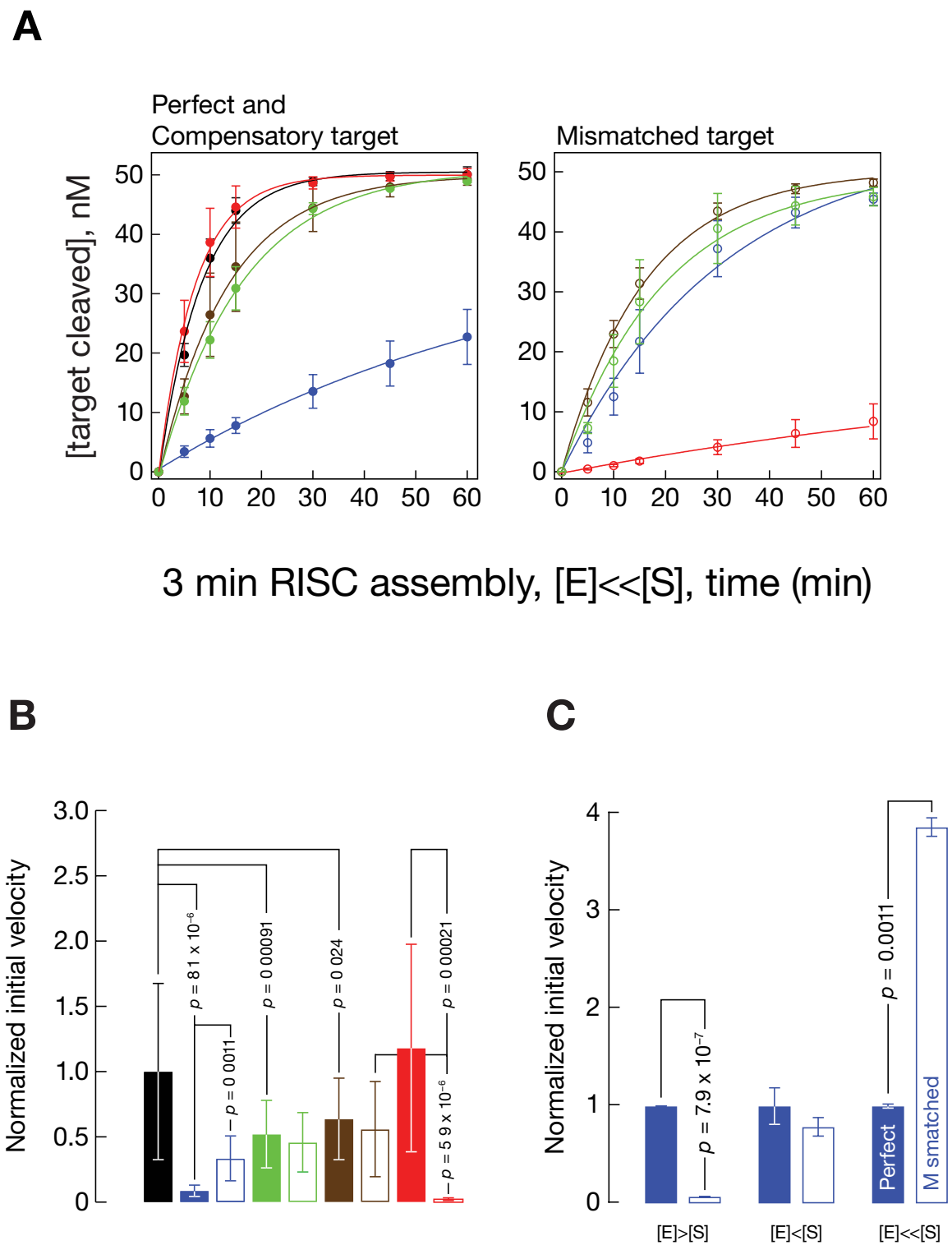

A
Perfect and
Compensatory target
Mismatched target
[target cleaved], nM
0
10
20
30
40
50
60
3 min RISC assembly, [E]<<[S], time (min)
B
Normalized initial velocity
0
0.5
1.0
1.5
2.0
2.5
3.0
p = 8 1 x 10⁻⁶
p = 0 0011
p = 0 00091
p = 0 024
p = 0 00021
p = 5 9 x 10⁻⁶
C
Normalized initial velocity
0
1
2
3
4
p = 7.9 x 10⁻⁷
p = 0.0011
Perfect
M smatched
[E]>[S]
[E]<[S]
[E]<<[S]

Figure 4.S2

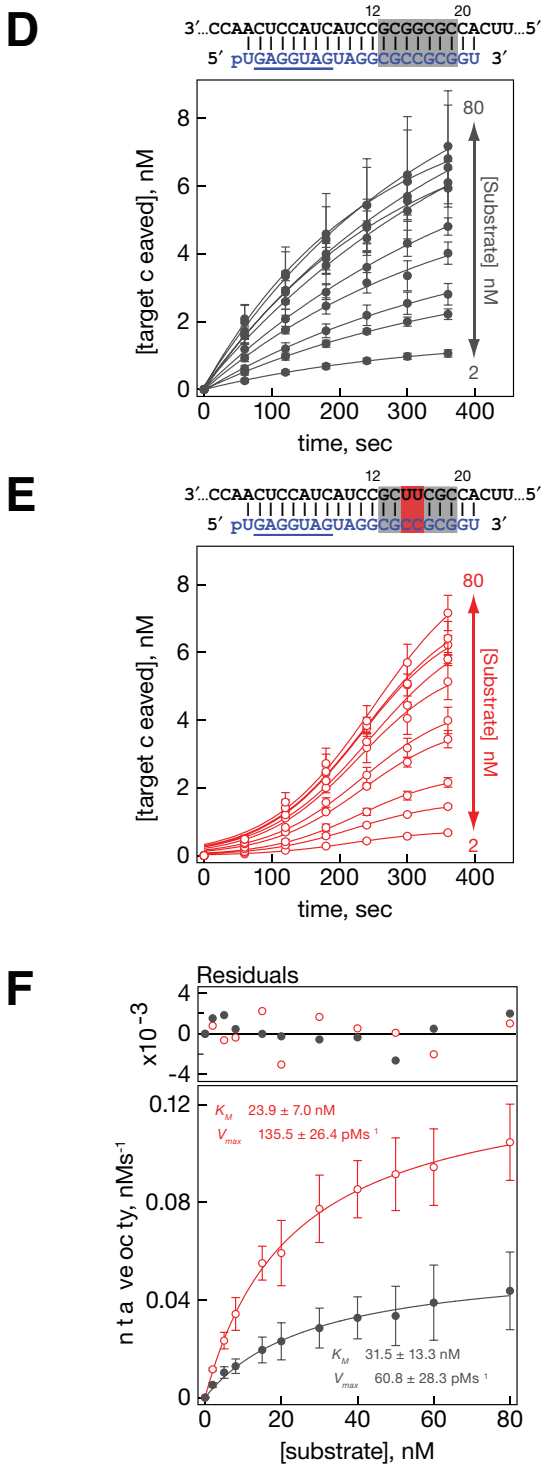

D
12
20
3′...CCAACUCCAUCAUCCGCGGCGCCACUU...5′
5′ pUGAGGUAGUAGGCGCCGCGGU 3′
80
[Substrate] nM
2
[target c eaved], nM
time, sec
E
12
20
3′...CCAACUCCAUCAUCCGCUUCGCCACUU...5′
5′ pUGAGGUAGUAGGCGCCGCGGU 3′
80
[Substrate] nM
2
[target c eaved], nM
time, sec
F
Residuals
x10 -3
K_M 23.9 ± 7.0 nM
V_{max} 135.5 ± 26.4 pMs ¹
K_M 31.5 ± 13.3 nM
V_{max} 60.8 ± 28.3 pMs ¹
n t a ve oc ty, nMs-1
[substrate], nM

Figure 4.S2

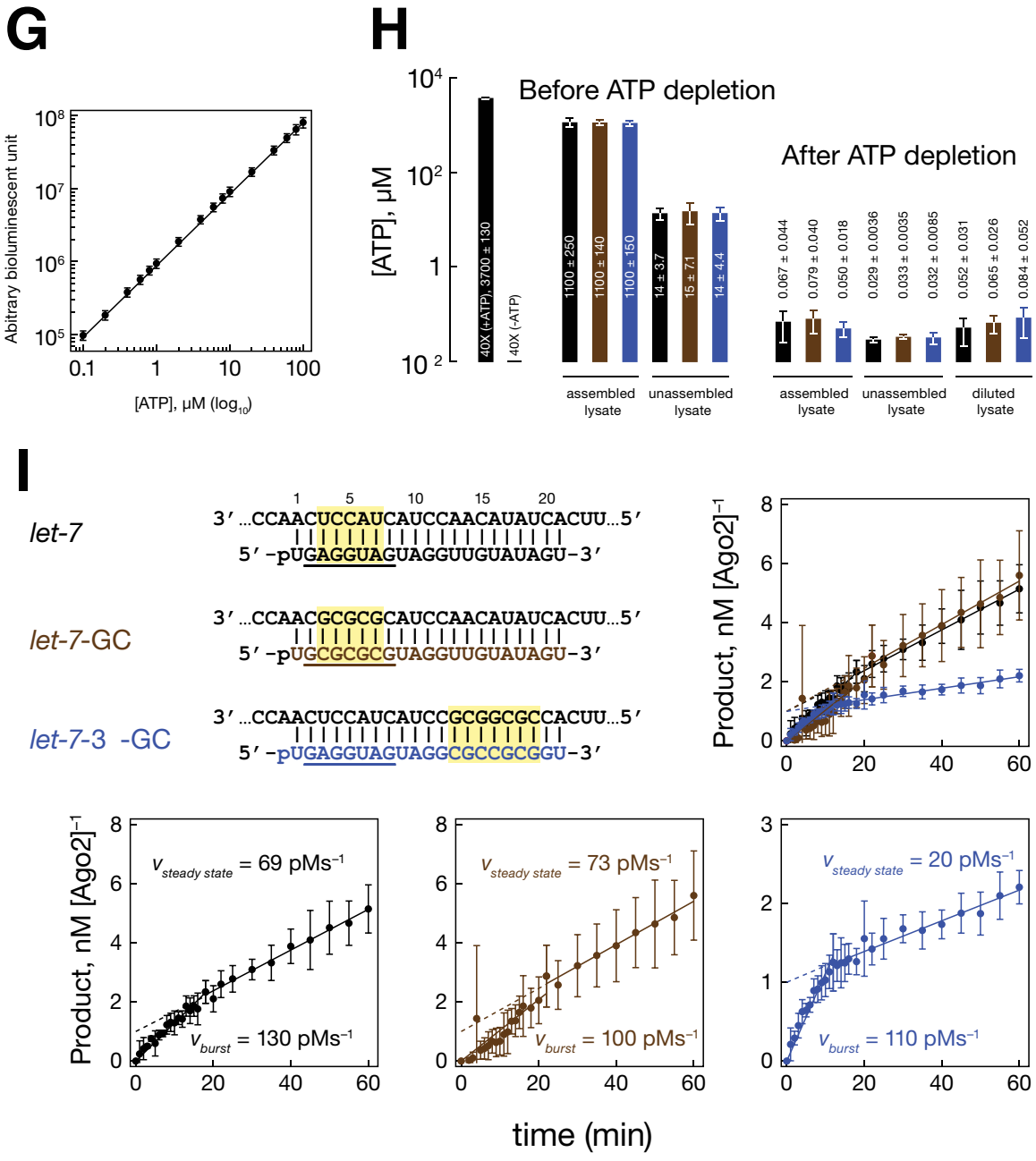

G
Abitrary bioluminescent unit
[ATP], μM (log10)
H
[ATP], μM
Before ATP depletion
After ATP depletion
40X (+ATP), 3700 ± 130
40X (-ATP)
1100 ± 250
1100 ± 140
1100 ± 150
14 ± 3.7
15 ± 7.1
14 ± 4.4
0.067 ± 0.044
0.079 ± 0.040
0.050 ± 0.018
0.029 ± 0.0036
0.033 ± 0.0035
0.032 ± 0.0085
0.052 ± 0.031
0.065 ± 0.026
0.084 ± 0.052
assembled lysate
unassembled lysate
assembled lysate
unassembled lysate
diluted lysate
I
let-7
let-7-GC
let-7-3 -GC
3′...CCAACUCCAUCAUCCAACAUAUCACUU...5′
5′-pUGAGGUAGUAGGUUGUAUAGU-3′
3′...CCAACGCGCGCAUCCAACAUAUCACUU...5′
5′-pUGCGCGCGUAGGUUGUAUAGU-3′
3′...CCAACUCCAUCAUCCGCGGCGCCACUU...5′
5′-pUGAGGUAGUAGGCGCCGCGGU-3′
Product, nM [Ago2]−1
v steady state = 69 pMs−1
v burst = 130 pMs−1
v steady state = 73 pMs−1
v burst = 100 pMs−1
v steady state = 20 pMs−1
v burst = 110 pMs−1
time (min)

Figure Legend 4.S2. High GC base pairing in 3′ supplementary region of siRNA guide restricts turnover of Ago2-RISC

(A) Time course of cleavage against compensatory, perfect (left panel, solid circles) and mismatched targets (right panel, opened circles). Data points are means ± standard deviations for at least 3 independent experiments. (B) Bar graph shows the initial velocities for (A) normalized to that of parental *let-7* against its perfect target (black bar). Normalized initial velocities for compensatory targets are represented by solid bars and that of mismatched targets by opened bars. *P*-values were calculated using a two-tailed Student's *t*-test assuming equal variances. (C) Normalized initial velocities for *let-7*-GC against mismatched or compensatory targets from 3 independent experiments ± propagated errors. *P*-values were calculated using a two-tailed Student's *t*-test. (D) Target cleavage by Ago2-RISC with GC-rich base pairing in the 3′ region of the siRNA at different substrate concentrations were fitted to single exponential kinetics. The plot shows data points derived from 4 independent sets of experiments ± one standard deviation. (E) Target cleavage by Ago2-RISC with GC-rich base pairing and g15g16:t15t16 mismatch in the 3′ region of the siRNA at different substrate concentrations were fitted to the logistic function. The plot shows data points derived from 3 independent sets of experiments ± one standard deviation. (F) Michaelis-Menten kinetics of initial rates from (D) and (E). All data points are derived from 5 independent experiments ± one standard

deviation. (G) Standard curve for ATP concentrations. (H) ATP concentrations of assembled and unassembled lysates were determined from the standard curve in (A) for pre (left) and post-ATP depleted samples (right). (I) Burst kinetics of *let-7*, *let-7*-GC and *let-7*-3′-GC against their perfect or compensatory targets per nM Ago2-RISC (A) in the presence of excess targets and minimal ATP ([ATP] ≤ 0.1 μM).

Chapter V: Conclusions, Discussion and Future Directions

Perspectives

Small RNA bound to Argonaute protein regulates gene expression by sequence-based homology. Despite its name as small RNA, it has big role in almost every aspect of cellular processes. We can categorize small RNAs into siRNA, miRNA and piRNA that associate with specific Argonaute proteins with distinct functions. How these small RNAs distribute into different Argonaute proteins and the detailed thermodynamic, kinetic and functional behaviors of various Argonaute proteins, however, were not well studied. These gaps in our understanding led to an incomplete overall perception of gene regulation by small RNAs and their underlying mechanisms. We began to address this problem by studying dAgo1 and mAGO2 that function mainly in the miRNA pathway and dAgo2 that participates primarily in the siRNA pathway. The following sections in this chapter summarize our work, highlight the progress made in the field and examine some of the future challenges.

Small RNAs Sort into Fly Argonaute Proteins with Distinct Roles

In Chapter II, we demonstrated that small RNAs are actively sorted between dAgo1 and dAgo2. The loading of small RNAs into Argonaute proteins is independent of how they are synthesized. Instead, the duplex structures of small RNAs determine how they are allocated into either dAgo1 or dAgo2. In terms of functional consequences, we and others revealed that dAgo1, but not dAgo2 repressed a centrally bulged target (Forstemann et al., 2007; Okamura et al.,

2009). A separate study, however, contradicts this finding and showed that both dAgo1 and dAgo2 can repress translation of centrally bulged mRNA but by different mechanisms (Iwasaki et al., 2009). Unfortunately, a quantitative comparison of the efficiency of translational repression in vivo between dAgo1 and dAgo2 has not been conducted. Therefore, whether dAgo2 result in meaningful translational repression in vivo is still an open question.

On the other hand, dAgo2 whose main role is to eliminate foreign RNA during an antiviral response cleaves a perfect target ~43-fold faster than dAgo1 (Wang et al., 2006b; Forstemann et al., 2007). Taken together, these data suggest that in fly: 1) small RNAs are actively sorted into their specific Argonaute hosts and 2) dAgo1 and dAgo2 have evolved differently to perform specific tasks: dAgo1 for the miRNA pathway and dAgo2 for the siRNA pathway (Okamura et al., 2004).

Determinants and Effects of Small RNA Sorting

Duplex structure of small RNA

In fly, Dicer-2/R2D2 complex acts as gatekeeper that shuttles highly paired duplex into dAgo2. This is because Dicer-2/R2D2 shows higher binding affinity to a perfectly paired duplex than to a mismatched duplex (Tomari et al., 2007). When viewed with respect to the 5′ end of the guide strand, base pairing at position 9 is critical to license its entry into dAgo2 (Ghildiyal et al., 2009; Okamura et al., 2009). In contrast, an siRNA duplex with central mismatches is

preferentially loaded into dAgo1 (Tomari et al., 2007). Similarly in *C*. *elegans*, a perfectly paired duplex is channeled into RDE-1 that performs RNAi whereas a duplex with central mismatches gets routed to ALG-1 and ALG-2 that function in miRNA regulation (Steiner et al., 2007). All 4 human Argonaute proteins, however, show no preferences for small RNA duplex structures (Meister et al., 2004; Liu et al., 2004; Yoda et al., 2009; Czech and Hannon, 2011). This may be because the loading of small RNA duplexes into human Argonaute proteins is Dicer independent (Betancur and Tomari, 2012).

It is tempting to suggest that the sorting process may have coevolved with the biochemical characteristics of the Argonaute protein. When a pre-RISC converts into a mature RISC, it discards the passenger strand. Fly Ago2 is a robust slicer and can easily cleave an extensively paired duplex and renders it single stranded. On the contrary, dAgo1, *C. elegans* ALG-1, ALG-2 and the four human Argonaute proteins that have relatively weaker catalytic activities than dAgo2 may benefit from a mismatched duplex: the passenger strand can easily dissociate without cleavage (Kawamata et al., 2009; Yoda et al., 2009).

First Nucleotide Identity of Small RNA

The first nucleotide identity of the small RNA guide influences its sorting into Argonaute protein (Mi et al., 2008; Tomari et al., 2007; Forstemann et al., 2007; Steiner et al., 2007; Ghildiyal et al., 2009; Czech et al., 2009; Okamura et al., 2009; Ameres et al., 2011). A close analysis of small RNAs bound to fly

Argonaute proteins revealed a first nucleotide bias. Fly Ago1 selects for small RNA that begins with uridine while dAgo2 selects for small RNA that starts with cytidine (Ghildiyal et al., 2009; Okamura et al., 2009; Ameres et al., 2011). Likewise, different plant Argonaute proteins prefer different first nucleotide bases (Mi et al., 2008). For instance, plant AGO1 selects for uridine, AGO2 and AGO4 select for adenosine while AGO5 selects for cytidine (Mi et al., 2008; Frank et al., 2012). The phenomenon of first nucleotide selection is supported and explained by structural studies of human AGO2. The specificity loop in the MID domain of human AGO2 can accommodate only uridine and adenine but discriminates against cytidine and guanosine as the first nucleotide of the guide (Frank et al., 2010; Elkayam et al., 2012).

Length of Small RNA

Argonaute proteins associate with small RNAs with distinct length. Fly Ago1 binds preferentially 22 nt small RNA whereas fly Ago2 binds 21 nt small RNA (Ameres et al., 2011). Likewise, plant AGO1 associates with 21–22 nt small RNA while AGO4, AGO5, AGO6 and AGO9 select for 24 nt small RNAs (Vazquez et al., 2008; Mallory and Vaucheret, 2010; Wu et al., 2010; Chellappan et al., 2010). Regardless, plant AGO1 and AGO4 exhibit similar binding affinities for 21 nt and 24 nt small RNAs. In plant, different Dicer isoforms produce small RNAs of different lengths (Voinnet, 2009). Furthermore, the sorting of the 21 nt and the 24 nt small RNA correlates instead with the Dicer isoforms that make them (Wu et

al., 2010). Therefore, in plant, Dicer determines the length of small RNA and may dictate its sorting into Argonaute protein. It is not known if Dicer or the length of small RNA governs its sorting into distinct human or fly Argonaute proteins.

After loading into Argonaute proteins, small RNAs are subjected to further modifications. Small RNA loaded into dAgo2 is 2′-*O*-methylated at its 3′ end and is therefore protected from nucleases activities (Li et al., 2005; Pelisson et al., 2007; Horwich et al., 2007; Saito et al., 2007). By contrast, small RNA guides in dAgo1, mAGO2 and human AGO2 are not modified. In the presence of an extensively paired target, these unprotected small RNA guides get tailed and trimmed; and may be destabilized and removed from the Argonaute proteins (Ameres et al., 2010; Cazalla et al., 2010; Ameres et al., 2011). As a result, post-loading modification contributes to the final repertoire of Argonaute-associated small RNAs.

During the development of mammalian nervous system, miRNAs are preferentially loaded into human AGO2. Concomitantly, these miRNAs are trimmed but are also protected from extensive nuclease digestions by the Argonaute protein. The functional significance of these shorter miRNAs, however, is not known (Juvvuna et al., 2012). At least in plants, the 22 nt but not the 21 nt small RNA guides confer RDR6-dependent synthesis of secondary small RNAs (Wang et al., 2011; Cuperus et al., 2010). Furthermore, PIWI, Ago3 and Aubergine in fly associate with piRNA whose length is longer than that of

siRNA ad miRNA (Guzzardo et al., 2013). The longer piRNA may reflect the different protection footprints offered by the PIWI clade Argonaute proteins. This argues that the distinct length of small RNA may result as a consequence and not a cause of sorting into specific Argonaute protein.

Different Argonautes: Different Functional Output

The seed sequence of the guide determines the target but the Argonaute protein can decide if and how the target is to be silenced. To illustrate, in plants, swapping a uridine for an adenosine as the first nucleotide redirects a miRNA from AGO1 to AGO2. Unlike AGO1, AGO2 with the same miRNA fails to silence its target genes (Mi et al., 2008). Likewise, in fly, dAgo1 and dAgo2 regulates imperfectly paired versus perfectly paired targets differently despite having the same miRNA loaded (Forstemann et al., 2007; Okamura et al., 2009; Iwasaki et al., 2009). Therefore, small RNA sorting adds an additional layer of control to gene silencing: both the identities of small RNA guides and Argonaute proteins matter.

Argonaute Proteins are Eager Target Seekers

Of note, mAGO2 is more similar to dAgo1 (74%) and to all four human Argonaute proteins (79%–99%) than to dAgo2 (30%) based on amino acid sequence identities (Table 5.1). In Chapter III, we assessed the performance of dAgo2 in target cleavage when subjected to differently paired targets by Henri-

Table 5.1. Amino acid sequence identity matrix of Argonaute proteins. The percentage amino acid sequence identities shared among Argonaute proteins were calculated using Clustal2.1.

	HsAgo2	MmAgo4	HsAgo1	MmAgo1	HsAgo4	MmAgo2	HsAgo3	MmAgo3	DmAgo1	DmAgo2
HsAgo2	100	99	82	82	79	79	80	80	74	30
MmAgo4	99	100	82	82	79	79	80	80	74	30
HsAgo1	82	82	100	100	84	83	85	85	74	31
MmAgo1	82	82	100	100	84	83	85	85	74	31
HsAgo4	79	79	84	84	100	99	83	83	73	31
MmAgo2	79	79	83	83	99	100	82	83	72	30
HsAgo3	80	80	85	85	83	82	100	99	73	30
MmAgo3	80	80	85	85	83	83	99	100	73	30
DmAgo1	74	74	74	74	73	72	73	73	100	27
DmAgo2	30	30	31	31	31	30	30	30	27	100

Percent shared amino acid sequence identity

Michaelis-Menten kinetics. We also determined quantitatively how dAgo2 and mAGO2 (a surrogate for dAgo1) bind and cleave their targets. In summary, our results revealed that Argonaute protein establishes functional domains in the small RNA guide: the anchor, the seed, the central, the 3′ supplementary and the tail domains. These domains determine how the Argonaute protein finds, binds and regulates its target. Our studies revealed that the seed region initiates target binding and both dAgo2 and mAGO2 bind their target with high picomolar affinity.

The preordering of the seed reduces the entropic penalty for guide target pairing and primes the Argonaute protein for target binding with an astounding bimolecular association rate (~10^7–10^8 $M^{-1}s^{-1}$) that is close to being diffusion-limited (Parker et al., 2009; Wee et al., 2012; Berg and von Hippel, 1985; Hammes and Schimmel, 1970). By comparison, the association rate for naked 7mer RNA-RNA duplex is at least 10-fold slower (~10^6 $M^{-1}s^{-1}$; Cisse et al., 2012).

Fly Ago2 Behaves and Functions Differently from Mouse AGO2

When bound to a perfect target, dAgo2 exhibits slow dissociation kinetics with a $t_{½}$ of ~2 hr. Compared to a $t_{½}$ of ~11 s for target cleavage, it suggests that dAgo2 always cleaves and destroys the target it binds. In contrast, for a seed-matched or a centrally bulged target—typical of miRNA binding site—dAgo2 dissociates from the target rapidly (seed-matched target: $t_{½}$ ~15 s and centrally bulged target: $t_{½}$ ~10 min). This may in part explains why dAgo2 fails to regulate

centrally bulged targets: dAgo2 cannot remain stably bound to generate meaningful repression (Forstemann et al., 2007). The upside is that dAgo2 will only associate briefly but dissociate rapidly from all possible ~1800, 7mer seed binding sites calculated based on the size of the fly transcriptome (~30 MB; Daines et al., 2011). Therefore, these fortuitous sites will not severely distract dAgo2 from its authentic and extensively paired target sites. Furthermore, the very low equilibrium dissociation constant (K_D = 4 pM) of dAgo2 for a perfect target indicates that dAgo2 can bind even a single molecule of viral RNA in a cell volume of ~1000 μm^3 (~2 pM). In short, dAgo2 is optimized to function in antiviral defense: it hunts and eliminates perfectly paired viral RNA efficiently (Wang et al., 2006b; Nayak et al., 2010).

In contrast, mAGO2 demonstrates similar dissociation rates towards both seed-matched and perfectly matched targets ($t_{½}$ ~15–25 min). For a seed-matched target, mAGO2 (dissociation $t_{½}$ ~23 min) remains more stably bound than dAgo2 ($t_{½}$ ~15 s). The lackluster cleavage activity of mAGO2 ($t_{½}$ ~14 min) suggests that a fully paired target has equal chance to be cleaved or to escape. This piece of data implies that for most genes that mAGO2 and perhaps dAgo1 silenced, the main route is not by target cleavage. The immediate goal therefore is to test if dAgo1 behaves similarly to mAGO2. If we assumed that dAgo1 behaves like mAGO2, we have to explain why mAGO2/dAgo1 (average dissociation $t_{½}$ ~20 min) can repress a target with multiple centrally bulged sites

Figure 5.1

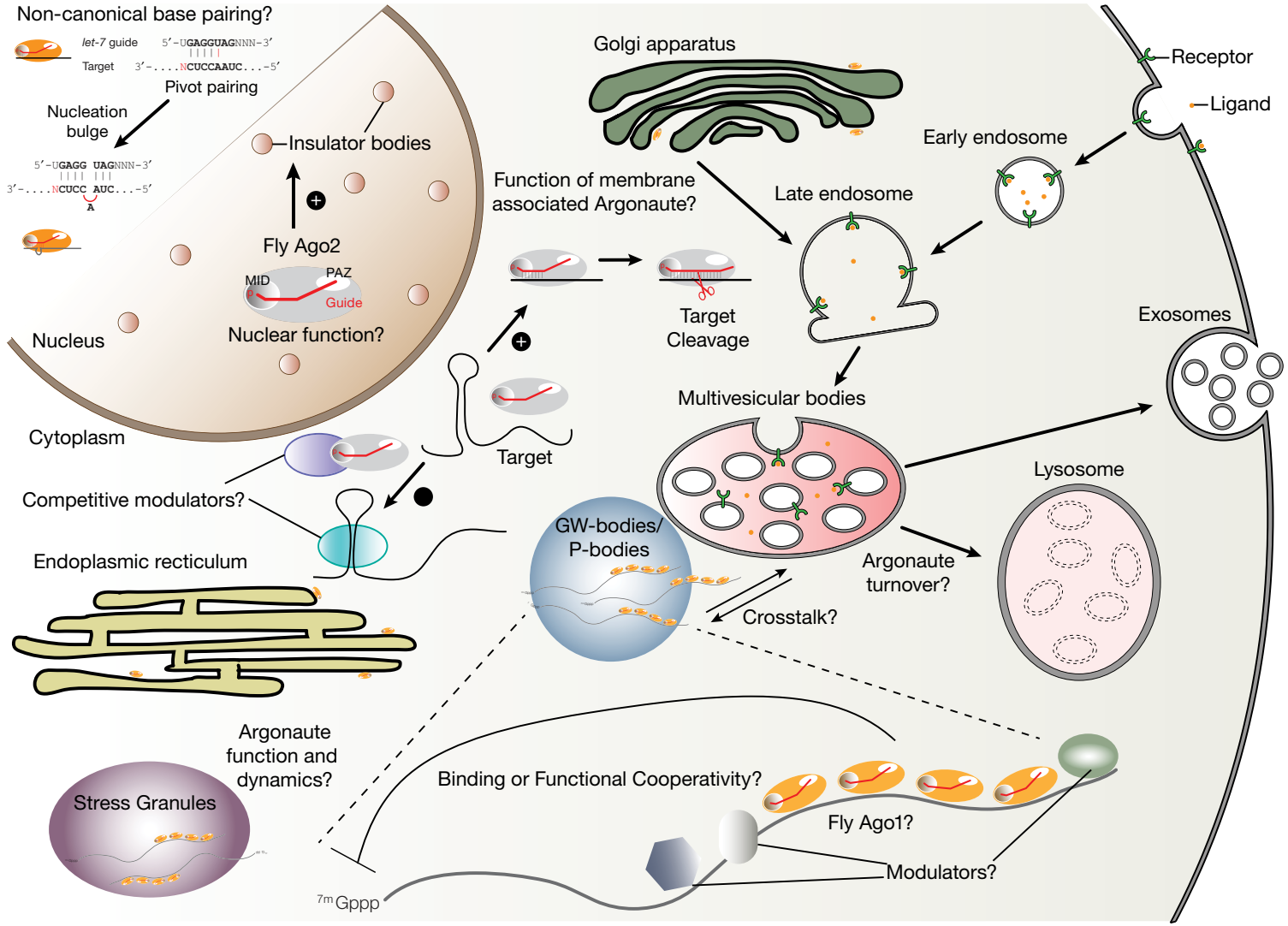

Non-canonical base pairing?
let-7 guide
Target
Pivot pairing
Nucleation bulge
Insulator bodies
Fly Ago2
MID
PAZ
Guide
Nuclear function?
Nucleus
Cytoplasm
Golgi apparatus
Receptor
Ligand
Early endosome
Function of membrane associated Argonaute?
Late endosome
Target Cleavage
Exosomes
Multivesicular bodies
Target
Competitive modulators?
Endoplasmic recticulum
GW-bodies/ P-bodies
Lysosome
Argonaute turnover?
Crosstalk?
Argonaute function and dynamics?
Stress Granules
Binding or Functional Cooperativity?
Fly Ago1?
Modulators?
7m Gppp

Figure Legend 5.1. Cellular localization and function of Argonaute proteins

Argonaute proteins function mostly in the cytoplasm but their roles in the nucleus are recently being recognized. In the cytosol, Argonaute proteins are either free in solution or are closely associated with the cellular membrane network, P-bodies, and stress granules. However, we still do not understand if these different cellular localizations represent different populations of Argonaute proteins with distinct functions. During target silencing, there may be more unidentified modulators that inhibit or promote Argonaute protein activity. Finally, few studies have addressed how the activities of Argonaute proteins terminate and how these enzymes turnover. Note that the cellular structures and molecules are not drawn to scale.

whereas dAgo2 that dissociates only two times faster ($t_{½}$ ~10 min) does not (Forstemann et al., 2007; Wee et al., 2012). There are least four explanations. First, dAgo1 and dAgo2 are simply structurally and therefore functionally different. Second, Argonaute proteins need to stay on their targets for at least 10 min to elicit functional regulation. Third, our dissociation rate measurements are limited to a single target site and may underestimate the duration that Argonaute proteins can stay bound to target mRNA with multiple sites as a result of cooperative interactions. Finally, perhaps only dAgo1/mAGO2 but not dAgo2 can bind and interact in a cooperative fashion (see next section).

The K_D of mAGO2 for a seed-matched target (26 pM) is similar to that for a perfect target (20 pM). This suggests that mAGO2 may not utilize all bases of the guide for pairing given a fully complementary target: mAGO2 limits paring to the seed. Therefore, we can predict that there will be no selective pressure for target sites to maintain base pairing beyond the seed sequence. Indeed, most predicted miRNA targets show only seed complementarity (Lai and Posakony, 1998; Bartel, 2009). Only less than 5% of preferentially conserved targets of miRNAs contained 3′ supplementary pairing in addition to seed pairing (Friedman et al., 2009). Finally, unlike siRNAs in dAgo2, miRNAs loaded into mAGO2 or dAgo1 are not 3′ methylated and are thus not protected from tailing and trimming and destabilization when paired to a perfect target (Ameres et al., 2010; Cazalla et al., 2010). Collectively, these observations suggest that mAGO2

and dAgo1 are not designed to regulate fully complementary targets and they behave and function differently from dAgo2.

Cooperativity among Argonaute Proteins

Our binding studies of dAgo2 and mAGO2 utilized target mRNAs with a single binding site (Wee et al., 2012). A miRNA-RISC, however, represses target in a cooperative fashion (Doench and Sharp, 2004; Broderick et al., 2011). On average, each mRNA target possesses 4.2 target sites in its 3′ UTR and is often complementary to multiple miRNA families (Friedman et al., 2009). This opens up the possibility that the assembly of multiple Argonaute proteins on a single mRNA accompanies effective mRNA repression. An intriguing challenge is to identify the step(s) during miRNA targeting responsible for the observed cooperative effect on target repression. An immediate achievable goal is to see if we can detect cooperative binding by Argonaute proteins to mRNA targets with multiple binding sites. Alternatively, target binding is non-cooperative but rather Argonaute proteins cooperate to repress target directly or they cooperate to recruit downstream protein factors for the task (Figure 5.1). Equally importantly, we need to test if cooperativity works for all Argonaute proteins or is limited to a subset that is known to function in the miRNA pathway such as dAgo1 and mAGO2. Of note, cooperativity with regard to multicomponent assembly is a common theme in cellular biology (Williamson, 2008).

Argonaute Proteins Bind Noncanonical Targets

Is a perfect seed pairing compulsory for Argonaute protein to bind and regulate its target? An earlier study surveyed bulged and internal loops in target repression by mAGO2 and human AGO2 and incorporated their results to predict imperfectly paired miRNA targets (Kiriakidou et al., 2004). The study, however, transfected high nanomolar concentration of siRNA into the cell, which is non-physiological and failed to account for the final cellular RISC concentrations. In Chapter IV, our preliminary results showed that dAgo2 tolerates mismatches, bulged and internal loops in the seed with moderately reduced target binding affinities or target cleavage rates. In our target cleavage readout with dAgo2 that necessitates extensive guide:target base pairings, these highly paired duplex may compensate for the mismatches, bulged and internal loops. That said, how our observations with dAgo2 translate to target regulations by dAgo1 or mAGO2 without extensive target pairing has not been vigorously tested.

By means of HITS-CLIP (High-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation) in mouse brain samples, it was discovered that ~27% of all Argonaute:mRNA interactions do not involve perfect seed match to miRNA (Figure 5.1; Licatalosi et al., 2008; Chi et al., 2009). In a follow up study that focused on mir-124:target interactions, it was found that ~15% of mir-124 targets formed a G bulged in the target across from position 5 and 6 of

the guide strand (Figure 5.1; Chi et al., 2012). The effect on K_D of mouse Argonaute proteins for these seed bulged targets remains to be determined. For highly abundant miRNA such as mir-124, its cellular concentration may still be greater than the K_D for the seed bulged target. That said, Argonaute proteins are expected to find, bind and leads to regulation of seed bulged targets. If binding to non-canonical targets, the number of functional miRNA targets may increase beyond those that fulfill the criterion of perfect seed pairing.

In an extreme example, mAGO2 binds target even without miRNA in mouse embryonic stem cells (Leung et al., 2011). It is not known how mAGO2 binds target mRNAs without the help of miRNAs. It is not hard to imagine that Argonaute protein itself or through its interacting protein binds to these mRNAs independent of miRNAs. These targets may also reflect artifacts produced during crosslinking. Functional studies suggest otherwise. Among targets that associate with Argonaute proteins with and without miRNAs, a G-rich motif is enriched. On its own, the G-motifs do not cause target repression. The G-motif, however, enhances target repression by miRNAs that bind to neighboring targeting sites (Leung et al., 2011).

Location, location, location! Where Do Argonaute Proteins Function?

Argonaute Proteins and the Membrane System

The small RNA pathways are intricately linked to the cellular membrane network in the cytoplasm. Earlier studies described mAGO2 as GERp95 (Golgi ER

protein 95 kDa) because it associated with the endoplasmic reticulum (ER) or the Golgi apparatus depending on cell-type (Figure 5.1; Cikaluk et al., 1999; Tahbaz et al., 2001). Neither was mAGO2 present within the lumen of the membrane system nor was it a transmembrane protein. Instead, it was proposed to be a peripheral membrane protein that exhibited high tolerance to trypsin/chymotrypsin digestion even in the presence of membrane disrupting detergent (Cikaluk et al., 1999). Interestingly, it was recently shown that human AGO2 loaded with small RNA was resistant to digestion by thermolysin at a concentration that digested non-loaded human AGO2 completely (Elkayam et al., 2012). This suggests that membrane-associated protease resistant mAGO2 may be small RNA-bound and active.

Early endosomes mature into MVB (multivesicular bodies) with defined intraluminal vesicles (Figure 5.1). This process requires the ESCRT (endosomal sorting complex required for transport) factors. The mature MVB either fuses with lysosome to degrade its cargo or it fuses with the plasma membrane to recycle its contents (Hanson and Cashikar, 2012). ESCRT mutants in *Drosophila* that cannot make MVB impair gene silencing by dAgo1. In contrast, blocking MVB turnover stimulates silencing by dAgo1 (Lee et al., 2009). Together, these findings suggest that miRNA functions in fly require MVB. Likewise, target repression in cultured monocytes by human AGO2 necessitates MVB formation (Gibbings et al., 2009).

It was found that the presence of MVB promotes loading of small RNAs into Argonaute proteins (Lee et al., 2009). Given that RISC assembly employs heat shock proteins, it will be interesting to find out if heat shock proteins concentrate at MVB (Czech and Hannon, 2011). MVB are also closely associated with GW-bodies, which are presumably the sites of actions of Argonaute proteins (Lee et al., 2009; Gibbings et al., 2009). Currently, it is not known if MVB and GW-bodies are functionally linked. It is also not known if the functions of Argonaute proteins are restricted to these membrane structures or associated GW-bodies.

P-bodies and Stress Granules Join the Fray

GW-bodies lack P-bodies components and were claimed to be distinct from P-bodies (Gibbings et al., 2009). However, GW-bodies and P-bodies are often assumed equivalent and are used synonymously (Filipowicz et al., 2008). There is therefore a need to characterize the components and to standardize the definitions of these subcellular structures. Here, I assume that GW-bodies and P-bodies are equivalent subcellular structures (Figure 5.1).

Translational inhibition causes Argonaute proteins to appear in P-bodies and stress granules (Leung et al., 2006). Intriguingly, heat shock proteins are also needed for the synthesis of P-bodies and stress granules (Matsumoto et al., 2011). One study showed that Argonaute protein localizes to P-bodies without the need for miRNAs and they turnover slowly. By contrast, Argonaute protein

localized to stress granule required miRNAs and they turnover rapidly (Leung et al., 2006). Further studies must address whether Argonaute proteins in P-bodies and stress granules are functionally distinct as reflected by their different turnover kinetics (Leung et al., 2006).

Most Argonaute proteins are dispersed in the cytoplasm with only a meager ~1.3% localized to P-bodies (Leung et al., 2006). The freely dispersed Argonaute proteins may reflect the starting population that is highly dynamic, capable of moving at the rate of diffusion to find their targets. Upon target binding and during target repression, these Argonaute proteins may then coalesce into visible P-bodies and stress granules. Given that target repression by Argonaute did not required its localization to the P-bodies (Eulalio et al., 2007; Chu and Rana, 2006), the freely dispersed sub-microscopic structure may also be Argonaute:target complexes that are functionally equivalent to those found in P-bodies but are simply below the limit of detection.

Single molecule tracking of Argonaute protein using labeled miRNA identified at least 2 distinct populations of Argonaute proteins with different intracellular mobility in the cytosol (Pitchiaya et al., 2012). The two populations were believed to correspond to the faster diffusing monomeric Argonaute:target complex and slower higher order multimeric Argonaute:target complexes in P-bodies. Unfortunately, the current setup did not have sufficient time resolution to distinguish the rapidly diffusing target-free Argonaute protein from free unloaded

miRNAs (Pitchiaya et al., 2012). In a separate study using HEK293 cells, Argonaute proteins were found to exist as different complexes in the nucleus and in the cytoplasm (Figure 5.1). The nuclear Argonaute is derived from the cytoplasmic pool, is smaller and demonstrates faster diffusion rate (Ohrt et al., 2008). In light of these data, several immediate questions come to mind. What constitutes the different Argonaute-RISC populations? Are they assembly intermediates? Are their functions distinct? What are their detailed kinetic properties? Do their different kinetic behaviors support their specific roles? The ultimate goal is to be able to correlate kinetically distinct Argonaute proteins to their components, functions and localizations.

Does Argonaute protein moonlight in the nucleus?

Argonaute proteins are present in the nucleus and they function in the nucleus too. The best studied is Argonaute1 in *Schizosaccharomyces pombe* (SpAgo1) that silences genes cotranscriptionally at centromeres, telomeres and the mating-type loci (Noma et al., 2004; Buhler et al., 2006). SpAgo1 with Tas3 and Chp1 forms the RITS complex that targets nascent transcript to silence the gene in a Dicer and RdRP-dependent manner (Verdel et al., 2004; Motamedi et al., 2004). The RITS complex then recruits the CLRC complex that deposits histone H3 lysine-9 methylation (H3K9me), a repressive mark, which is bound by Swi6 (*Drosophila* HP1 homolog; Bannister et al., 2001; Hall et al., 2002; Volpe et al., 2002; Hong et al., 2005; Horn et al., 2005; Jia et al., 2005; Li et al., 2005; Thon et

al., 2005; Bayne et al., 2010). Swi6 helps spread and maintain the H3K9me heterochromatic mark (Nakayama et al., 2001). Chp1 can also bind H3K9me to enhance localization of RITS and CLRC complexes to the silenced loci to reinforce and to propagate the repressive mark.

In *C. elegans*, a nuclear Argonaute protein NRDE1 associates with endogenous siRNA and trimethylates H3K9 to direct nuclear gene silencing (Buckley et al., 2012). Another *C. elegans* Argonaute protein NRDE3, shuttles into the nucleus when it is loaded with a small RNA guide. In the nucleus, it binds nascent transcripts and silences gene expression by inhibiting Pol II occupancy and activity (Guang et al., 2008; Guang et al., 2010). NRDE3 is not equipped with catalytic residues and therefore does not silence a gene by endonucleolytic cleavage. In contrast, the worm Argonaute protein, CSR-1 exhibits cleavage activity in vitro but it does not cleave its perfectly matched target in vivo (Aoki et al., 2007). Transcriptional profiling of *csr-1* mutant worm shows no significant difference in global gene expression to wild type worm suggests that CSR-1 does not degrade its target genes (Claycomb et al., 2009). Instead, CSR-1 functions to organize chromosomal domains to ensure proper chromosomal segregation (Claycomb et al., 2009; van Wolfswinkel et al., 2009). It is puzzling why Argonaute proteins that retain conserved catalytic residues and can cleave perfectly paired targets in vitro do not appear to function by cleavage in vivo.

PIWI Argonautes in fly contain catalytic residues and can cleave extensively paired target mRNAs (Gunawardane et al., 2007). The nuclear localized PIWI protein, however, silences transposons independent of its slicer activity (Gunawardane et al., 2007). Instead, PIWI established repressive chromatin marks to silence genes transcriptionally much like the way SpAgo1 represses the repeat elements in fission yeast (Sienski et al., 2012; Rozhkov et al., 2013; Le Thomas et al., 2013; Huang et al., 2013). It is not clear if PIWI, like the RITS complex in yeast, is responsible for recruiting deacetylation and methylation enzymes to modify histone tails. Nor is it well understood if PIWI recruits HP1 (heterochromatin protein 1), a Swi6 homolog, to bind methylated histone.

Chromatin insulators are protein-DNA barriers that function to restrict the spread of silent chromatin or to prevent enhancer-promoter interactions (Van Bortle and Corces, 2013). In *D. melanogaster*, a zinc-finger DNA binding protein, CTCF together with CP190 bind the *Fab-8* DNA element of the *Abd-B* locus in the bithorax complex to form an insulator element (Gerasimova et al., 2007). The association among thousands of these insulator elements creates higher order looping of chromosomes and establishes a limited number of nuclear foci termed insulator bodies. *Drosophila* Ago2 but not Dicer-2 is required to make these insulator bodies (Figure 5.1). Moreover, the catalytic activity of dAgo2 is dispensable in generating insulator bodies. These findings reveal a nuclear role

of dAgo2 to form higher order chromosomal structure without the need for small RNAs and is independent of its canonical RNAi function (Moshkovich et al., 2011). In conclusion, data from yeast, worm and fly indicate that Argonaute proteins do function in the nucleus.

Modulators of Argonaute Proteins

Numerous proteins interact directly with Argonaute proteins to modulate their functions. For example, GW182 binds to Argonaute protein and assists in miRNA-mediated gene silencing (Rehwinkel et al., 2005; Till et al., 2007; Eulalio et al., 2008). The human GW182 ortholog, TNRC6A assists to shuttle Argonaute protein between the nucleus and the cytoplasm (Nishi et al., 2012). In fission yeast, two proteins Arb1 and Arb2 inhibit the release of the passenger strand and keep the siRNA double-stranded in Argonaute (Buker et al., 2007). Arb1 and Arb2 therefore prevent the formation of active Argonaute proteins. In *C. elegans*, the TRIM-NHL protein, NHL-2 interacts with Argonaute protein and enhances repression of target mRNA (Hammell et al., 2009). *Belle* (*vasa* paralog) identified in an RNAi screen is shown to be required for gene silencing. *Belle* associates with both dAgo1 and dAgo2 in S2 cell lysate. In the eye cells of fly that lacks *belle*, RNAi fails to occur (Zhou et al., 2008; Pek and Kai, 2011). In *Tetrahymena*, Giw1p senses the mature state of the Argonaute protein Twi1p and Gwi1P help localize mature Twi1p to the nucleus to initiate DNA elimination (Noto et al., 2010). The Tudor family of proteins interacts with Argonaute proteins of the PIWI

Table 5.2. Argonaute interacting proteins

Protein	Type of protein	Function	References
Belle/DDX 3	DEAD-box RNA helicase	1) Interacts with RNAi components (DmAgo1, DmAgo2, FMRP and VIG) and is required for RNA silencing 2) Required for cell viability and germline development 3) Interacts with Ago2 to localize hCAP-H/Barr to chromosome and promote chromosome segregation	(Johnstone et al., 2005; Zhou et al., 2008; Pek and Kai, 2011)
RpL5, RpL11, RpL21, RpS7, RpL22, and Rp49	Ribosomal protein	Interacts with Ago2, VIG and FMR	(Ishizuka et al., 2002; Zhou et al., 2008)
dFXR	*Drosophila* homolog of FMRP (Fragile X Mental Retardation Protein); RNA binding protein that contains KH (hnRNP K Homology) domain and RGG box	Function in RNAi. Actual mechanism unknown. RNAi slightly impaired when protein is knocked down	(Caudy et al., 2002; Ishizuka et al., 2002)
VIG	RNA binding protein that contains a RGG box	Function in RNAi. Actual mechanism unknown. RNAi slightly impaired when protein is knocked down	(Caudy et al., 2002)
p68	DEAD-box helicase	Direct interaction with Argonaute has not been shown. Present in a complex with Argonaute and dFMRP. Requires for RNA silencing	(Ishizuka et al., 2002)
RNA helicase A/ DHX9/NDHII	DEAH-box RNA helicase	Requires for RISC assembly	(Robb and Rana, 2007)
RCK/p54	DEAD-box RNA	Interacts with human Ago1	(Chu and

	helicase	and Ago2. Functions in miRNA mediated translational repression	Rana, 2006)
Hsp90	Heat shock protein	Assembles RISC; stabilizes Argonaute protein; targets human Ago2 to stress granules and p-bodies; its absence affects piRNA level and causes transposon derepression	(Tahbaz et al., 2001; Pare et al., 2009; Specchia et al., 2010; Iki et al., 2010; Iwasaki et al., 2010; Miyoshi et al., 2010)
GW182/Gawky	Glycine and tryptophan (GW) repeats containing protein. Paralogs in vertebrates includes TNRC6A/GW182, TNRC6B, and TNRC6C	Interacts with Argonaute proteins and mediates translational repression and degradation of mRNA targets	(Ding et al., 2005; Schneider et al., 2006; Liu et al., 2005; Jakymiw et al., 2005; Rehwinkel et al., 2005)
NHL-2, TRIM32, Brat/Mei-P26	TRIM-NHL protein	Brat and Mei-P26 binds *drosophila* Ago1; Mei-P26 inhibits miRNA pathway to regulate cell proliferation. NHL-2 interacts with worm Argonaute proteins, ALG-1 and ALG-2 and facilitates miRNA:target interaction. TRIM32 binds mouse AGO1 and increases miRNA activity.	(Neumuller et al., 2008; Hammell et al., 2009; Schwamborn et al., 2009)
Arb1 and Arb2		Inhibits passenger strand removal from siRNA duplex bound by yeast Ago1, therefore inhibiting Ago1 maturation and slicing activity. Required for heterochromatin assembly in yeast.	(Buker et al., 2007)
Gemin3 and Gemin4	DEAD-box helicase	Exists as Gemin3-Gemin4-Argonaute complex. Function unknown.	(Hutvagner and Zamore, 2002; Mourelatos et al., 2002)

clade and is required for gametogenesis in fly and mouse. Currently, not much is known about the mechanistic roles of Tudor proteins in these processes (Lim and Kai, 2007; Vagin et al., 2009; Nishida et al., 2009; Zhang et al., 2011; Anand and Kai, 2011).

There are also several examples of RNA binding proteins that bind in the vicinity of the miRNA binding sites either to promote or to antagonize Argonaute activities (Bhattacharyya et al., 2006; Kedde et al., 2007; Elcheva et al., 2009; Goswami et al., 2010; Kedde et al., 2010; Jafarifar et al., 2011; Toledano et al., 2012). Mechanistically, these RNA binding proteins can bind to Argonaute protein and deter its interaction with target mRNAs (Bhattacharyya et al., 2006). By having close or overlapping binding sites to that of the miRNA, the RNA binding proteins may physically obstruct and preclude Argonaute binding (Kedde et al., 2007; Elcheva et al., 2009; Goswami et al., 2010; Jafarifar et al., 2011). Alternatively, these RNA binding proteins can create highly structured target mRNAs that discourage Argonaute binding or promote Argonaute dissociation. By contrast, these RNA binding proteins can also increase accessibility of the miRNA target sites to Argonaute proteins (Kedde et al., 2010).

Biochemical and genomic screens in *C elegans* and in *D. melanogaster* uncovered many other proteins that interact with Argonaute proteins or are implicated in small RNA silencing (Table 5.2; Mourelatos et al., 2002; Hutvagner and Zamore, 2002; Caudy et al., 2002; Ishizuka et al., 2002; Caudy et al., 2003;

Meister et al., 2005; Kim et al., 2005; Duchaine et al., 2006; Hock et al., 2007; Zhou et al., 2008). In addition, by means of comparing and clustering proteins with similar phylogenetic profiles across 86 eukaryotic genomes, a set of 60 candidate proteins in *C. elegans* were identified that may functions in small RNA silencing pathway (Tabach et al., 2013). Currently, the precise molecular roles of these candidate interacting proteins are still mysteries.

Argonaute Beware! Decoys Among Real Targets

The concept of target decoy was first observed in plants where a target mimic binds and sequesters miRNA away from its canonical target thus relieving it from repression (Franco-Zorrilla et al., 2007). In animals, pseudogenes are proposed to be the doppelgänger of canonical targets that serve as target decoys for miRNAs (Seitz, 2009). Subsequently, the concept of target decoys was recast as the prevalent and pervasive networks of competitive endogenous RNAs (ceRNA; Poliseno et al., 2010; Salmena et al., 2011; Tay et al., 2011; Karreth et al., 2011). In the ceRNA hypothesis, ceRNAs along with the canonical targets are predicted to share common miRNA binding sites and therefore are believed to communicate with one another by means of shared miRNA regulators. The biggest flaw of the ceRNA hypothesis, elegant as it may seem, is that it lacks sound and quantitative experimental proof. It is at odds with fundamental facts on mRNA and miRNA concentrations and the principles governing mRNA regulation by miRNA. To elaborate, it was claimed that

PTENP1 acts as the target decoy of *PTEN* mRNA (Poliseno et al., 2010). In the cell line that was used to test the ceRNA hypothesis, *PTENP1* is present at ~1% that of *PTEN* thus making it highly incapable to effectively sequester all miRNAs away from *PTEN* (Ebert and Sharp, 2010; Ebert and Sharp, 2012). Furthermore, only the few most highly abundant miRNAs are functional in the cells (Mullokandov et al., 2012). To effectively compete with canonical target for the abundant and functional miRNAs would require that all transcripts in the cell be used as target decoys (Wee et al., 2012). This is highly questionable!

To gain further insight into the biological relevance of target decoy, we modelled target repression by high, intermediate and low abundance miRNAs using our biochemically-defined target binding parameters of Argonaute proteins in Chapter III. In the model, we assume that each miRNA has 50 different targets and each target is present in 10 copies, which is equivalent to a total of ~500 targets in the cell. For high abundance miRNAs such as *let-7* and mir-21 (~20, 000 copies; ~3–4 nM in a HeLa cell volume of 5000 μm^3), nearly every molecule of its seed-matched targets (~95–99%) are bound and repressed. To achieve 50% derepression of these seed-matched targets of *let-7* or mir-21, a staggering ~22,000 copies of target decoys is required. This assumes that each target decoy has one miRNA-binding site. Twenty-two thousand copies of target decoys is equivalent to 50% of all mRNAs in the cell (Islam et al., 2011). Therefore, the idea that target decoys also known as ceRNAs exist to bind and

titrate these abundant miRNAs away from the real targets seems highly improbable.

In the other extreme, only ~380 copies of target decoys are required to attain 50% derepression of real targets of low abundance miRNA (7.3 pM) such as mir-24. Our model indicates that mir-24 at such low abundance only binds <4% of all targets and therefore seems unlikely to contribute to meaningful repression of target mRNAs. This leaves us with miRNA of intermediate abundance (~140 pM) such as mir-93. At this cellular concentration, mir-93 elicits 60% target repression. Moreover, the cell needs only to produce ~500 copies of target decoys to counteract the target repression of mir-93 by 50%. Taken together, our model suggests that target decoy is functionally relevant only for miRNAs and targets whose cellular concentrations satisfy a limited range of values: miRNAs whose abundance contribute to functional repression and can be quenched by reasonable amount of target decoys.

Intriguingly, nature has figured out a way to inhibit highly abundant miRNAs using circular RNAs (Salzman et al., 2012). Two reports identified target decoys of mir-7 in neuronal tissues that exist as circular RNAs. Amazingly, each circular decoy harbors seventy-four mir-7 binding sites. Among these sites, 63 are conserved binding sites for mir-7 in at least 1 out of the 32 vertebrate species compared (Memczak et al., 2013). Present at ~280 molecules per cell (~0.09 nM in a cell volume of 5000 μm^3), it is capable of binding to a total upper

limit of 20,000 copies of mir-7. Being circular, these target decoys have their ends protected and rendered unavailable to the cellular mRNA decay machinery as well (Memczak et al., 2013; Hansen et al., 2013). Furthermore, none of the partially paired target sites demonstrate base pairings beyond position 12 of the guide strand thus restricting target cleavage by Argonaute proteins to the minimum. These circular RNA decoys are therefore able to quench even the most abundant miRNAs in the cells while retaining great stability.

Conclusions

In this thesis, we characterized the thermodynamics, kinetics and functions of Argonaute proteins to learn more about the mechanisms of small RNA silencing. Argonaute proteins are core effectors that capture and divide small RNA guides into functional domains. In turn, the sequences of small RNAs are like address codes that pilot Argonaute proteins to targets with complementary base sequences. Argonaute proteins are robust enzymes that tolerate mismatches, bulged and internal loops in target mRNA. This increases the number of potential targets in addition to those predicted based on perfect seed pairing to the small RNA guide. An additional level of flexibility and complexity involves active sorting of small RNA guides into distinct Argonaute proteins with defined kinetic and binding properties. Therefore, small RNA guide dictates not only which targets are being regulated and also how these targets are regulated. Our quantitative analysis of the binding and cleavage kinetics of Argonaute proteins provide mechanistic explanation as to how and why dAgo2 is designed to achieve RNAi whereas dAgo1/mAGO2 is devised to accomplish miRNA-based target repression. Finally, using biochemically established binding parameters, we can propose a reasonable model to explain how Argonaute proteins find, bind and regulate their targets.

We also discussed some main issues that still need to be addressed: 1) What other factors modulate Argonaute functions and how do they achieve it

mechanistically? 2) Where do Argonaute proteins regulate their targets in the cell? 3) Are there different populations of Argonaute proteins that can be categorized based on their distinct thermodynamic and kinetic behaviors, their locations and interacting modulators? 4) How do Argonaute proteins interact and work with other intracellular components? i.e. subcellular structures, target decoys and RNA binding proteins. 5) Can we delineate the temporal order from RISC assembly to when it executes its function in the cell? 6) Will the temporal information correlate with distinct cellular kinetics and locations of Argonaute proteins? Overcoming these challenges and answering these questions will extend our understanding of these remarkable biological innovations.

Future Directions

K_M is not a good reflection of K_D

Our current model for fly Ago2 that binds and cleaves a perfect target is given by

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P$$

where fly Ago2 being the enzyme, E binds its substrate, S to form the enzyme:substrate complex, ES before cleaving the substrate to give free enzyme E and the cleaved product, P (Forstemann et al., 2007). The association rate constant is denoted by k_1, the dissociation rate constant by k_{-1} and the catalytic rate is given by k_{cat}. Accordingly the Henri-Michaelis-Menten constant, K_M is given by

$$K_M = \frac{k_{-1} + k_{cat}}{k_1} \quad (1)$$

And

$$\frac{k_{cat}}{K_M} = \frac{k_{cat} k_1}{k_{-1} + k_{cat}} \quad (2)$$

In the event that $k_{cat} << k_{-1}$, the definition of K_M reduces to k_{-1}/k_1. This will be equivalent to the equilibrium dissociation constant, K_D.

$$K_D = \frac{k_{-1}}{k_1} \quad (3)$$

We showed that the K_M of fly Ago2 for a perfect target is 25 ± 6 nM while its k_{cat} is 0.02–0.06 s^{-1}. In our recent binding studies, we demonstrated that fly Ago2 binds a perfect target at a K_D = 3.7 ± 0.9 pM. We measured an off-rate, k_{-1} of 8.8 x 10^{-5} s^{-1} and a calculated on-rate, k_1 of 2.4 x 10^7 $M^{-1}s^{-1}$. From the definition of K_M (equation 1), we can also calculate k_1 given by k_{cat}/K_M (equation 2) since k_{-1} is very small and we arrive at a similar k_1 value of 5.9 x 10^7 $M^{-1}s^{-1}$. Of note, $k_{cat}>>k_{-1}$ and K_D is at least 6000-fold smaller than K_M. Therefore K_M is not equal to K_D!

k_{cat} and k_{on} vary for different mismatches

We can rewrite equation 1 as follows,

$$K_M = \frac{k_{-1}}{k_1} + \frac{k_{cat}}{k_1}$$

$$K_M = K_D + \frac{k_{cat}}{k_1} \quad (4)$$

If both k_{cat} and k_1 values do not change for all the different mismatched targets tested (Figure 3.1), we expect that a plot of K_M against K_D will correlate with a y-intercept at k_{cat}/k_1 (equation 4). In this scenario, changes in K_M strictly reflect changes in k_{-1} with a gradient of $1/k_1$ (equation 4). In our survey of mismatched pairing between guide and target, k_{cat} and k_{on} were altered either alone or in combination across all mismatches tested (Figures 3.1 and 5.2 and Appendix II). Therefore K_M does not correlate with K_D or with V_{max}/k_{cat} (Figure 5.2).

Figure 5.2

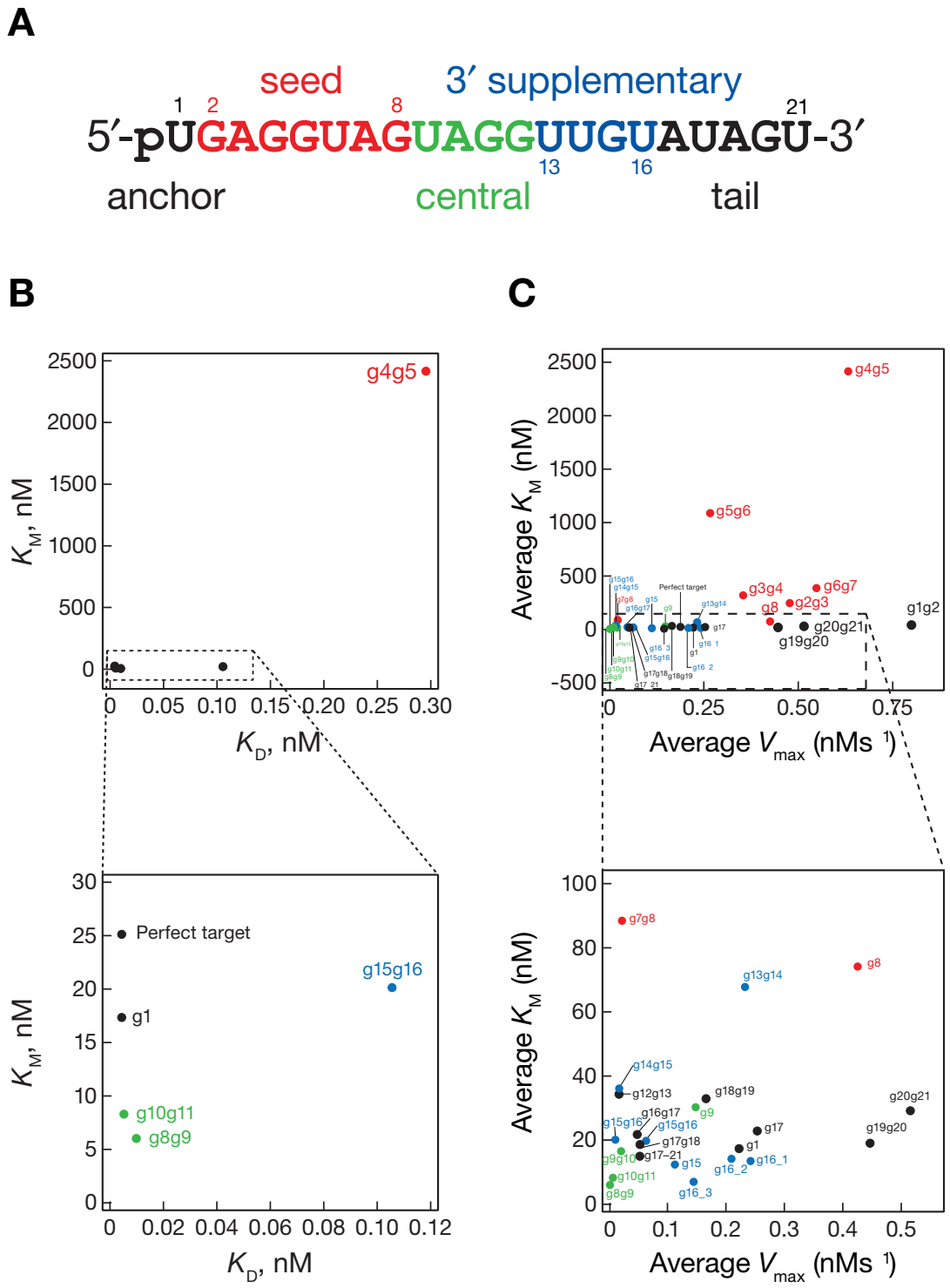

A
seed
3′ supplementary
1 2 8 13 16 21
5′-pUGAGGUAGUAGGUUGUAUAGU-3′
anchor
central
tail
B
K_M, nM
K_D, nM
g4g5
Perfect target
g15g16
g1
g10g11
g8g9
C
Average K_M (nM)
Average V_{max} (nMs $^{-1}$)
g4g5
g5g6
g3g4
g6g7
g8
g2g3
g1g2
g20g21
g19g20
g7g8
g13g14
g14g15
g12g13
g18g19
g16g17
g9
g15g16
g17g18
g17
g1
g17–21
g15
g16_2
g16_1
g10g11
g8g9
g16_3

Figure Legend 5.2. Changes in k_{cat} and k_{on} changes K_M

(A) Domains of the siRNA guide (B) K_M versus K_D (C) K_M versus V_{max}. On the assumption that the concentration of fly Ago2-RISC assembled is the same for every siRNA, the K_M versus k_{cat} plot will have the same distribution of data points.

Binding of a perfect target involves at least two distinct steps

For a seed-only target, fly Ago2 binds 10-fold faster than a perfect target with an on-rate, k_1 = 2.1 x 10^8 $M^{-1}s^{-1}$. Also for a seed-only target, fly Ago2 dissociates faster with a k_{-1} value of 0.045 s^{-1}. As such, we can rewrite our kinetic model as

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} ES^* \xrightarrow{k_{cat}} E + P$$

In this model, ES complex refers to the seed binding step and ES* refers to the complex that is fully paired and restructured and is therefore competent to cleave target mRNA. According to this model, we can calculate the value of k_2 = 0.006–0.018 s^{-1} (Appendix III). Compared to the k_{cat} for target cleavage of 0.02–0.06 s^{-1}, it suggests that k_2 is either smaller than or similar to k_{cat} and is predicted to be the rate limiting step for fly Ago2 in target cleavage. This rate-limiting step includes both the propagation of base pairing beyond the seed and the structural rearrangement of fly Ago2.

Refining the kinetic model for *Drosophila* Ago2 in target cleavage

We can further refine our kinetic model to represent a more realistic mechanism for target cleavage by fly Ago2. As such, we shall incorporate the rates that describe product release. Here, I will highlight two possible models as follow,

$$E + S \xleftrightarrow{K_S} ES \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES^* \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E^*P \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} EP \xleftrightarrow{K_P} E + P$$

In this first model, product is released only when fly Ago2 reverts to its inactive conformation E. We also assume that the steps for seed binding and for product release are in rapid equilibrium with a dissociation rate constant K_S and K_P respectively.

$$E + S \xleftrightarrow{K_S} ES \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES^* \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E^*P \underset{k_{3}}{\overset{k_3}{\rightleftharpoons}} E^* + P \xleftrightarrow{K_P} E + P$$

The second model (above) predicts that product release can occur while fly Ago2 is still in its active state E*. In this model, the steps for seed binding and the reversion of active fly Ago2 into its inactive form are in rapid equilibrium with dissociation rate constants K_S and K_P respectively. Accordingly, for the first model, we can show that

$$K_M = \frac{K_S\left(k_{-1}k_{-2} + k_3k_2\right)}{k_1k_{-2} + k_1k_2 + k_{-1}k_{-2} + k_3k_2}$$

$$k_{cat} = \frac{k_2k_{-2}k_1}{k_1k_{-2} + k_1k_2 + k_{-1}k_{-2} + k_3k_2}$$

$$\frac{k_{cat}}{K_M} = \frac{k_2k_{-2}k_1}{K_S\left(k_{-1}k_{-2} + k_3k_2\right)}$$

And for the second model,

$$K_M = \frac{K_S\left(k_{-1}k_3 + k_2k_3 + k_2k_{-3} + k_{-1}k_{-3}\right)}{k_{-1}k_3 + k_2k_3 + k_1k_2 + k_1k_3}$$

$$k_{cat} = \frac{k_2k_1k_3}{k_{-1}k_3 + k_2k_3 + k_1k_2 + k_1k_3}$$

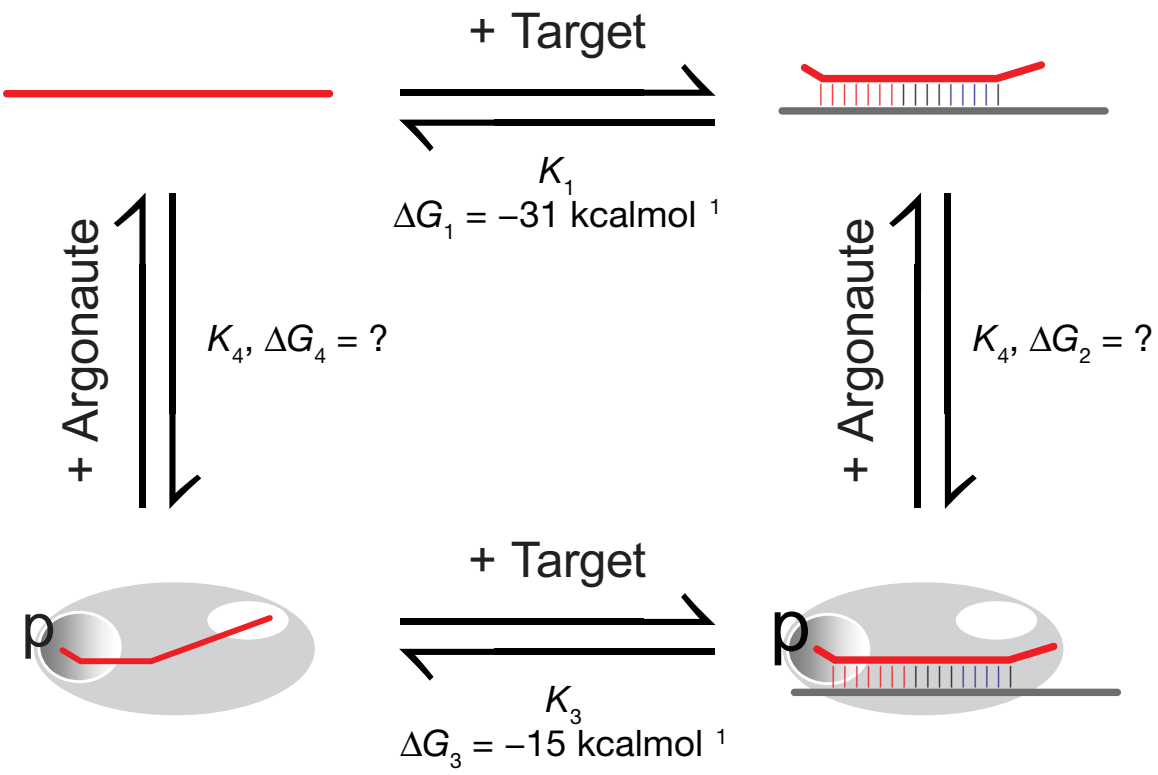
$$\frac{k_{cat}}{K_M} = \frac{k_2k_1k_3}{K_S\left(k_{-1}k_3 + k_2k_3 + k_2k_{-3} + k_{-1}k_{-3}\right)}$$

Each model provides unique solutions for K_M and k_{cat} in the absence of product and therefore offers the potential in future studies to distinguish which model (if any) most clearly predicts experimental behavior. The more complete solutions for these schema (Appendix IV and V) indicate that product may interact either E or E* in different ways. This indicates that product inhibition studies may provide an additional tool to test these models.

Drosophila Ago2 prefers single-stranded RNA to double-stranded RNA

The binding of RNA by Argonaute protein can be described by a thermodynamic cycle (Figure 5.3). There are two routes that will lead to the final state of Argonaute protein bound to both guide and target mRNA. In the first path, Argonaute protein is first loaded with guide RNA with an equilibrium constant K_4 and then binds its target with an equilibrium constant K_3 (Figure 5.3).

Figure 5.3



$$K_1 K_2 = K_3 K_4$$

$$\Delta G_1 + \Delta G_2 = \Delta G_3 + \Delta G_4$$

Figure Legend 5.3. Thermodynamic cycle for RNA binding by fly Ago2

Alternatively, guide RNA binds the target mRNA with an equilibrium constant K_1 before associating with Argonaute protein with an equilibrium constant K_2 (Figure 5.3). The overall thermodynamic to form the final state is independent of the paths taken and hence $K_1K_2 = K_3K_4$. In terms of free energies, both pathways must sum up to the same amount: $\Delta G_1 + \Delta G_2 = \Delta G_3 + \Delta G_4$.

Case Study: Fly Ago2

Fly Ago2 binds a perfect target with a K_D of 4 pM. This corresponds to ΔG_3 = −15 kcalmol^{-1} at 25°C ($\Delta G = -RTln1/K_D$). In the absence of Argonaute protein, a fully paired guide and target RNA has a free energy of −31 kcalmol^{-1}. This, however, is an overestimate of the free energy of binding when the duplex is incorporated into Argonaute protein given that only ~two-thirds of the base nucleotide in the guide is used for target binding (Figure 3.4). Given that seed (g2–g8) and extended 3′ supplementary pairing (g12–g17) gives almost the same binding affinity as a perfect target, we assume that the free energy of binding of guide and target in the absence of Argonaute protein would be more accurate reflected by just seed and extended 3′ supplementary pairing (Figure 3.4). To obtain the overall binding free energy for such pairing, we calculate the free energy of seed pairing $\Delta G_{(seed)}$ = −11.5 kcalmol^{-1} and for extended 3′ supplementary pairing $\Delta G_{(3' \text{ supplementary})}$ = −5.3 kcalmol^{-1} and sum them to arrive at

$\Delta G_1 \cong -17$ kcalmol^{-1}. When calculating $\Delta G_{(seed)}$ and $\Delta G_{(3' supplementary)}$, the initiation penalties are included.

Currently, we do not have free energy data of fly Ago2 for single-stranded guide or for double-stranded RNA. We can, however, make certain predictions. Given that $\Delta G_1 + \Delta G_2 = \Delta G_3 + \Delta G_4$ and that $\Delta G_1 = -31$ kcalmol^{-1} and $\Delta G_3 = -15$ kcalmol^{-1} we expect that ΔG_2 (the binding of duplex RNA) will be energetically less favorable than ΔG_4 (the binding of single-stranded guide). This is at least true for *Thermus thermophilus* Argonaute protein for single-stranded and double-stranded DNA (personal communication with Samson Jolly). The K_D of *Thermus thermophilus* Argonaute protein for single-stranded and double-stranded DNA are 0.5 nM ($\Delta G_4 = -12.7$ kcalmol^{-1}) and 40 nM ($\Delta G_3 = -10.8$ kcalmol^{-1}). The less favorable loading of RNA duplex will also explain why its loading into Argonaute proteins requires assistance from hsp90/hsc70 and ATP (Johnston et al., 2010; Iki et al., 2010; Iwasaki et al., 2010; Miyoshi et al., 2010). Moreover, the loading of duplex that remains fully paired is probably discouraged given that ~two-thirds of the guide nucleotides are used in pairing (see above).

BIBLIOGRAPHY

Amarzguioui, M., Holen, T., Babaie, E., and Prydz, H. (2003). Tolerance for mutations and chemical modifications in a siRNA. Nucleic Acids Res. *31*, 589-595.

Ambros, V., Bartel, B., Bartel, D. P., Burge, C. B., Carrington, J. C., Chen, X., Dreyfuss, G., Eddy, S. R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. (2003). A uniform system for microRNA annotation. RNA *9*, 277-279.

Ameres, S. L., Horwich, M. D., Hung, J. H., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P. D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science *328*, 1534-1539.

Ameres, S. L., Hung, J. H., Xu, J., Weng, Z., and Zamore, P. D. (2011). Target RNA-directed tailing and trimming purifies the sorting of endo-siRNAs between the two Drosophila Argonaute proteins. RNA *17*, 54-63.

Ameres, S. L., Martinez, J., and Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. Cell *130*, 101-112.

Anand, A., and Kai, T. (2011). The tudor domain protein Kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila. EMBO J.

Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in C. elegans. EMBO J. *26*, 5007-5019.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M. J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J. J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., and Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A. A., Naumova, N. M., Tulin, A. V., Vagin, V. V., Rozovsky, Y. M., and Gvozdev, V. A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline. Curr. Biol. *11*, 1017-1027.

Arvey, A., Larsson, E., Sander, C., Leslie, C. S., and Marks, D. S. (2010). Target mRNA abundance dilutes microRNA and siRNA activity. Mol Syst Biol *6*, 363.

Auyeung, V. C., Ulitsky, I., McGeary, S. E., and Bartel, D. P. (2013). Beyond Secondary Structure: Primary-Sequence Determinants License Pri-miRNA Hairpins for Processing. Cell *152*, 844-858.

Axtell, M. J., Westholm, J. O., and Lai, E. C. (2011). Vive la difference: biogenesis and evolution of microRNAs in plants and animals. Genome Biol *12*, 221.

Aza-Blanc, P., Cooper, C. L., Wagner, K., Batalov, S., Deveraux, Q. L., and Cooke, M. P. (2003). Identification of modulators of TRAIL-induced apoptosis via RNAi-based phenotypic screening. Mol. Cell. *12*, 627-637.

Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008). The impact of microRNAs on protein output. Nature

Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A. E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell *122*, 553-563.

Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

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

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

Bartel, D. P., and Chen, C. Z. (2004). Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. Nat. Rev. Genet. *5*, 396-400.

Baskerville, S., and Bartel, D. P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA *11*, 241-247.

Batista, P. J., Ruby, J. G., Claycomb, J. M., Chiang, R., Fahlgren, N., Kasschau, K. D., Chaves, D. A., Gu, W., Vasale, J. J., Duan, S., Conte, D. J., Luo, S., Schroth, G. P., Carrington, J. C., Bartel, D. P., and Mello, C. C. (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans. Mol. Cell. *31*, 67-78.

Bayne, E. H., White, S. A., Kagansky, A., Bijos, D. A., Sanchez-Pulido, L., Hoe, K.-L., Kim, D.-U., Park, H.-O., Ponting, C. P., Rappsilber, J., and Allshire, R. C. (2010). Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity. Cell *140*, 666-677.

Bazzini, A. A., Lee, M. T., and Giraldez, A. J. (2012). Ribosome Profiling Shows That miR-430 Reduces Translation Before Causing mRNA Decay in Zebrafish. Science

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

Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. (2006b). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. Cold Spring Harb. Symp. Quant. Biol. *71*, 523-530.

Berezikov, E., Chung, W. J., Willis, J., Cuppen, E., and Lai, E. C. (2007). Mammalian mirtron genes. Mol. Cell. *28*, 328-336.

Berg, O. G., and von Hippel, P. H. (1985). Diffusion-controlled macromolecular interactions. Annu Rev Biophys Biophys Chem *14*, 131-160.

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

Betancur, J. G., and Tomari, Y. (2012). Dicer is dispensable for asymmetric RISC loading in mammals. RNA *18*, 24-30.

Betel, D., Wilson, M., Gabow, A., Marks, D. S., and Sander, C. (2008). The microRNA.org resource: targets and expression. Nucleic Acids Res. *36*, D149-53.

Béthune, J., Artus-Revel, C. G., and Filipowicz, W. (2012). Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. EMBO Rep.

Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell *125*, 1111-1124.

Billy, E., Brondani, V., Zhang, H., Muller, U., and Filipowicz, W. (2001). Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. Proc. Natl. Acad. Sci. U S A *98*, 14428-14433.

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*, 170-180.

Bohnsack, M. T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA *10*, 185-191.

Boland, A., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2011). Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. Proc. Natl. Acad. Sci. U S A

Boland, A., Tritschler, F., Heimstadt, S., Izaurralde, E., and Weichenrieder, O. (2010). Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. EMBO Rep. *11*, 522-527.

Borchert, G. M., Lanier, W., and Davidson, B. L. (2006). RNA polymerase III transcribes human microRNAs. Nat. Struct. Mol. Biol. *13*, 1097-1101.

Boutet, S., Vazquez, F., Liu, J., Beclin, C., Fagard, M., Gratias, A., Morel, J. B., Crete, P., Chen, X., and Vaucheret, H. (2003). Arabidopsis HEN1: a genetic link between endogenous miRNA controlling development and siRNA controlling transgene silencing and virus resistance. Curr. Biol. *13*, 843-848.

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

Brennecke, J., Hipfner, D. R., Stark, A., Russell, R. B., and Cohen, S. M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. Cell *113*, 25-36.

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., and Hannon, G. J. (2007). Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell *128*, 1089-1103.

Brennecke, J., Stark, A., Russell, R. B., and Cohen, S. M. (2005). Principles of microRNA-target recognition. PLoS Biol. *3*, e85.

Broderick, J. A., Salomon, W. E., Ryder, S. P., Aronin, N., and Zamore, P. D. (2011). Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing. RNA *17*, 1858-1869.

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., and Voinnet, O. (2008). Widespread translational inhibition by plant miRNAs and siRNAs. Science *320*, 1185-1190.

Brown, K. M., Chu, C. Y., and Rana, T. M. (2005). Target accessibility dictates the potency of human RISC. Nat. Struct. Mol. Biol. *12*, 469-470.

Buckley, B. A., Burkhart, K. B., Gu, S. G., Spracklin, G., Kershner, A., Fritz, H., Kimble, J., Fire, A., and Kennedy, S. (2012). A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. Nature

Buhler, M., Verdel, A., and Moazed, D. (2006). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. Cell *125*, 873-886.

Buker, S. M., Iida, T., Buhler, M., Villen, J., Gygi, S. P., Nakayama, J., and Moazed, D. (2007). Two different Argonaute complexes are required for siRNA generation and heterochromatin assembly in fission yeast. Nat. Struct. Mol. Biol. *14*, 200-207.

Cai, X., Hagedorn, C. H., and Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA *10*, 1957-1966.

Calabrese, J. M., Seila, A. C., Yeo, G. W., and Sharp, P. A. (2007). RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. Proc. Natl. Acad. Sci. U S A *104*, 18097-18102.

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*, 9742-9747.

Carmell, M. A., Xuan, Z., Zhang, M. Q., and Hannon, G. J. (2002). The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. Genes Dev. *16*, 2733-2742.

Caudy, A. A., Ketting, R. F., Hammond, S. M., Denli, A. M., Bathoorn, A. M., Tops, B. B., Silva, J. M., Myers, M. M., Hannon, G. J., and Plasterk, R. H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature *425*, 411-414.

Caudy, A. A., Myers, M., Hannon, G. J., and Hammond, S. M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. Genes Dev. *16*, 2491-2496.

Cazalla, D., Yario, T., and Steitz, J. A. (2010). Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. Science *328*, 1563-1566.

Cenik, E. S., Fukunaga, R., Lu, G., Dutcher, R., Wang, Y., Tanaka Hall, T. M., and Zamore, P. D. (2011). Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. Mol. Cell. *42*, 172-184.

Cenik, E. S., and Zamore, P. D. (2011). Argonaute proteins. Curr. Biol. *21*, R446-9.

Chang, J. H., Xiang, S., Xiang, K., Manley, J. L., and Tong, L. (2011). Structural and biochemical studies of the 5'→3' exoribonuclease Xrn1. Nat. Struct. Mol. Biol. *18*, 270-276.

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

Chekulaeva, M., Parker, R., and Filipowicz, W. (2010). The GW/WG repeats of Drosophila GW182 function as effector motifs for miRNA-mediated repression. Nucleic Acids Res. *38*, 6673-6683.

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

Chellappan, P., Xia, J., Zhou, X., Gao, S., Zhang, X., Coutino, G., Vazquez, F., Zhang, W., and Jin, H. (2010). siRNAs from miRNA sites mediate DNA methylation of target genes. Nucleic Acids Res. *38*, 6883-6894.

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

Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. Science *303*, 2022-2025.

Chen, X., Liu, J., Cheng, Y., and Jia, D. (2002). HEN1 functions pleiotropically in Arabidopsis development and acts in C function in the flower. Development *129*, 1085-1094.

Chi, S. W., Hannon, G. J., and Darnell, R. B. (2012). An alternative mode of microRNA target recognition. Nat. Struct. Mol. Biol. *19*, 321-327.

Chi, S. W., Zang, J. B., Mele, A., and Darnell, R. B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature

Chiang, H. R., Schoenfeld, L. W., Ruby, J. G., Auyeung, V. C., Spies, N., Baek, D., Johnston, W. K., Russ, C., Luo, S., Babiarz, J. E., Blelloch, R., Schroth, G. P., Nusbaum, C., and Bartel, D. P. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev. *24*, 992-1009.

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

Chung, W. J., Okamura, K., Martin, R., and Lai, E. C. (2008). Endogenous RNA interference provides a somatic defense against Drosophila transposons. Curr. Biol. *18*, 795-802.

Cikaluk, D. E., Tahbaz, N., Hendricks, L. C., DiMattia, G. E., Hansen, D., Pilgrim, D., and Hobman, T. C. (1999). GERp95, a membrane-associated protein that

belongs to a family of proteins involved in stem cell differentiation. Mol. Biol. Cell *10*, 3357-3372.

Cisse, I. I., Kim, H., and Ha, T. (2012). A rule of seven in Watson-Crick base-pairing of mismatched sequences. Nat. Struct. Mol. Biol.

Claycomb, J. M., Batista, P. J., Pang, K. M., Gu, W., Vasale, J. J., van Wolfswinkel, J. C., Chaves, D. A., Shirayama, M., Mitani, S., Ketting, R. F., Conte, D. J., and Mello, C. C. (2009). The Argonaute CSR-1 and its 22G-RNA cofactors are required for holocentric chromosome segregation. Cell *139*, 123-134.

Cogoni, C., and Macino, G. (1997). Isolation of quelling-defective (qde) mutants impaired in posttranscriptional transgene-induced gene silencing in Neurospora crassa. Proc. Natl. Acad. Sci. U S A *94*, 10233-10238.

Cohen, L. S., and Studzinski, G. P. (1967). Correlation between cell enlargement and nucleic acid and protein content of HeLa cells in unbalanced growth produced by inhibitors of DNA synthesis. J. Cell. Physiol. *69*, 331-339.

Cole, C., Sobala, A., Lu, C., Thatcher, S. R., Bowman, A., Brown, J. W., Green, P. J., Barton, G. J., and Hutvagner, G. (2009). Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. RNA *15*, 2147-2160.

Cooke, A., Prigge, A., and Wickens, M. (2010). Translational repression by deadenylases. J. Biol. Chem. *285*, 28506-28513.

Cox, D. N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev. *12*, 3715-3727.

Cuperus, J. T., Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Burke, R. T., Takeda, A., Sullivan, C. M., Gilbert, S. D., Montgomery, T. A., and Carrington, J. C. (2010). Unique functionality of 22-nt miRNAs in triggering RDR6-dependent siRNA biogenesis from target transcripts in Arabidopsis. Nat. Struct. Mol. Biol.

Czech, B., and Hannon, G. J. (2011). Small RNA sorting: matchmaking for Argonautes. Nat. Rev. Genet. *12*, 19-31.

Czech, B., Malone, C. D., Zhou, R., Stark, A., Schlingeheyde, C., Dus, M., Perrimon, N., Kellis, M., Wohlschlegel, J. A., Sachidanandam, R., Hannon, G. J.,

and Brennecke, J. (2008). An endogenous small interfering RNA pathway in Drosophila. Nature *453*, 798-802.

Czech, B., Zhou, R., Erlich, Y., Brennecke, J., Binari, R., Villalta, C., Gordon, A., Perrimon, N., and Hannon, G. J. (2009). Hierarchical rules for Argonaute loading in Drosophila. Mol. Cell. *36*, 445-456.

Daines, B., Wang, H., Wang, L., Li, Y., Han, Y., Emmert, D., Gelbart, W., Wang, X., Li, W., Gibbs, R., and Chen, R. (2011). The Drosophila melanogaster transcriptome by paired-end RNA sequencing. Genome Res. *21*, 315-324.

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.

Davis, E., Caiment, F., Tordoir, X., Cavaille, J., Ferguson-Smith, A., Cockett, N., Georges, M., and Charlier, C. (2005). RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. Curr. Biol. *15*, 743-749.

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*, 231-235.

Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. *11*, 1475-1489.

Ding, H., Schwarz, D. S., Keene, A., Affar el, B., Fenton, L., Xia, X., Shi, Y., Zamore, P. D., and Xu, Z. (2003). Selective silencing by RNAi of a dominant allele that causes amyotrophic lateral sclerosis. Aging Cell *2*, 209-217.

Ding, L., Spencer, A., Morita, K., and Han, M. (2005). The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in C. elegans. Mol. Cell. *19*, 437-447.

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

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

Doench, J. G., and Sharp, P. A. (2004). Specificity of microRNA target selection in translational repression. Genes Dev. *18*, 504-511.

Duchaine, T. F., Wohlschlegel, J. A., Kennedy, S., Bei, Y., Conte, D. J., Pang, K., Brownell, D. R., Harding, S., Mitani, S., Ruvkun, G., Yates, J. R. r., and Mello, C. C. (2006). Functional proteomics reveals the biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. Cell *124*, 343-354.

Ebert, M. S., Neilson, J. R., and Sharp, P. A. (2007). MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods *4*, 721-726.

Ebert, M. S., and Sharp, P. A. (2010). Emerging roles for natural microRNA sponges. Curr. Biol. *20*, R858-61.

Ebert, M. S., and Sharp, P. A. (2012). Roles for MicroRNAs in Conferring Robustness to Biological Processes. Cell *149*, 515-524.

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*, 494-498.

Elbashir, S. M., Lendeckel, W., and Tuschl, T. (2001b). RNA interference is mediated by 21- and 22-nucleotide RNAs. Genes Dev. *15*, 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*, 6877-6888.

Elcheva, I., Goswami, S., Noubissi, F. K., and Spiegelman, V. S. (2009). CRD-BP protects the coding region of betaTrCP1 mRNA from miR-183-mediated degradation. Mol. Cell. *35*, 240-246.

Elkayam, E., Kuhn, C.-D., Tocilj, A., Haase, A. D., Greene, E. M., Hannon, G. J., and Joshua-Tor, L. (2012). The Structure of Human Argonaute-2 in Complex with miR-20a. Cell

Enright, A. J., John, B., Gaul, U., Tuschl, T., Sander, C., and Marks, D. S. (2003). MicroRNA targets in Drosophila. Genome Biol *5*, R1.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol. Cell. Biol. *27*, 3970-3981.

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

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

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

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

Fabian, M. R., Mathonnet, G., Sundermeier, T., Mathys, H., Zipprich, J. T., Svitkin, Y. V., Rivas, F., Jinek, M., Wohlschlegel, J., Doudna, J. A., Chen, C. Y., Shyu, A. B., Yates, J. R. r., Hannon, G. J., Filipowicz, W., Duchaine, T. F., and Sonenberg, N. (2009). Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. Mol. Cell. *35*, 868-880.

Fagard, M., Boutet, S., Morel, J. B., Bellini, C., and Vaucheret, H. (2000). AGO1, QDE-2, and RDE-1 are related proteins required for post-transcriptional gene silencing in plants, quelling in fungi, and RNA interference in animals. Proc. Natl. Acad. Sci. U S A *97*, 11650-11654.

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

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

Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat. Rev. Genet. *9*, 102-114.

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

Flores-Jasso, C. F., Salomon, W. E., and Zamore, P. D. (2013). Rapid and specific purification of Argonaute-small RNA complexes from crude cell lysates. RNA *19*, 1-9.

Forstemann, K., Horwich, M. D., Wee, L., Tomari, Y., and Zamore, P. D. (2007). Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. Cell *130*, 287-297.

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*, e236.

Franco-Zorrilla, J. M., Valli, A., Todesco, M., Mateos, I., Puga, M. I., Rubio-Somoza, I., Leyva, A., Weigel, D., Garcia, J. A., and Paz-Ares, J. (2007). Target mimicry provides a new mechanism for regulation of microRNA activity. Nat. Genet. *39*, 1033-1037.

Frank, F., Fabian, M. R., Stepinski, J., Jemielity, J., Darzynkiewicz, E., Sonenberg, N., and Nagar, B. (2011). Structural analysis of 5'-mRNA-cap interactions with the human AGO2 MID domain. EMBO Rep. *12*, 415-420.

Frank, F., Hauver, J., Sonenberg, N., and Nagar, B. (2012). Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. EMBO J.

Frank, F., Sonenberg, N., and Nagar, B. (2010). Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. Nature *465*, 818-822.

Friedman, R. C., Farh, K. K., Burge, C. B., and Bartel, D. P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. *19*, 92-105.

Fukaya, T., and Tomari, Y. (2012). MicroRNAs Mediate Gene Silencing via Multiple Different Pathways in Drosophila. Mol. Cell.

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

Fukunaga, R., Han, B. W., Hung, J. H., Xu, J., Weng, Z., and Zamore, P. D. (2012). Dicer Partner Proteins Tune the Length of Mature miRNAs in Flies and Mammals. Cell

Gaidatzis, D., van Nimwegen, E., Hausser, J., and Zavolan, M. (2007). Inference of miRNA targets using evolutionary conservation and pathway analysis. BMC Bioinformatics *8*, 69.

Galiana-Arnoux, D., Dostert, C., Schneemann, A., Hoffmann, J. A., and Imler, J. L. (2006). Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila. Nat. Immunol. *7*, 590-597.

Garcia, D. M., Baek, D., Shin, C., Bell, G. W., Grimson, A., and Bartel, D. P. (2011). Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat. Struct. Mol. Biol. *18*, 1139-1146.

Gerasimova, T. I., Lei, E. P., Bushey, A. M., and Corces, V. G. (2007). Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. Mol. Cell. *28*, 761-772.

German, M. A., Pillay, M., Jeong, D. H., Hetawal, A., Luo, S., Janardhanan, P., Kannan, V., Rymarquis, L. A., Nobuta, K., German, R., De Paoli, E., Lu, C., Schroth, G., Meyers, B. C., and Green, P. J. (2008). Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat. Biotechnol.

Ghildiyal, M., Seitz, H., Horwich, M. D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E. L., Zapp, M. L., Weng, Z., and Zamore, P. D. (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science *320*, 1077-1081.

Ghildiyal, M., Xu, J., Seitz, H., Weng, Z., and Zamore, P. D. (2009). Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. RNA

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

Gibbings, D. J., Ciaudo, C., Erhardt, M., and Voinnet, O. (2009). Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. Nat. Cell Biol. *11*, 1143-1149.

Gibbings, D., Mostowy, S., Jay, F., Schwab, Y., Cossart, P., and Voinnet, O. (2012). Selective autophagy degrades DICER and AGO2 and regulates miRNA activity. Nat. Cell Biol.

Giraldez, A. J., Mishima, Y., Rihel, J., Grocock, R. J., Van Dongen, S., Inoue, K., Enright, A. J., and Schier, A. F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science *312*, 75-79.

Girard, A., Sachidanandam, R., Hannon, G. J., and Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature *442*, 199-202.

Goswami, S., Tarapore, R. S., Teslaa, J. J., Grinblat, Y., Setaluri, V., and Spiegelman, V. S. (2010). MicroRNA-340-mediated degradation of microphthalmia-associated transcription factor mRNA is inhibited by the coding region determinant-binding protein. J. Biol. Chem. *285*, 20532-20540.

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*, 235-240.

Griffiths-Jones, S., Grocock, R. J., van Dongen, S., Bateman, A., and Enright, A. J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res. *34*, D140-4.

Griffiths-Jones, S., Saini, H. K., van Dongen, S., and Enright, A. J. (2008). miRBase: tools for microRNA genomics. Nucleic Acids Res. *36*, D154-8.

Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engele, P., Lim, L. P., and Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol. Cell. *27*, 91-105.

Grimson, A., Srivastava, M., Fahey, B., Woodcroft, B. J., Chiang, H. R., King, N., Degnan, B. M., Rokhsar, D. S., and Bartel, D. P. (2008). Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature *455*, 1193-1197.

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

Grivna, S. T., Pyhtila, B., and Lin, H. (2006). MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. Proc. Natl. Acad. Sci. U S A *103*, 13415-13420.

Grosshans, H., Johnson, T., Reinert, K. L., Gerstein, M., and Slack, F. J. (2005). The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in C. elegans. Dev. Cell *8*, 321-330.

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

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

Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M. C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila. Science *315*, 1587-1590.

Guo, H., Ingolia, N. T., Weissman, J. S., and Bartel, D. P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835-840.

Guo, S., and Kemphues, K. J. (1995). par-1, a gene required for establishing polarity in C. elegans embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell *81*, 611-620.

Guzzardo, P. M., Muerdter, F., and Hannon, G. J. (2013). The piRNA pathway in flies: highlights and future directions. Curr. Opin. Genet. Dev.

Ha, I., Wightman, B., and Ruvkun, G. (1996). A bulged lin-4/lin-14 RNA duplex is sufficient for Caenorhabditis elegans lin-14 temporal gradient formation. Genes Dev. *10*, 3041-3050.

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*, 961-967.

Haley, B., Tang, G., and Zamore, P. D. (2003). In vitro analysis of RNA interference in Drosophila melanogaster. Methods *30*, 330-336.

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

Hall, I. M., Shankaranarayana, G. D., Noma, K., Ayoub, N., Cohen, A., and Grewal, S. I. (2002). Establishment and maintenance of a heterochromatin domain. Science *297*, 2232-2237.

Hall, K. B., and Stump, W. T. (1992). Interaction of N-terminal domain of U1A protein with an RNA stem/loop. Nucleic Acids Res. *20*, 4283-4290.

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

Hammell, C. M., Lubin, I., Boag, P. R., Blackwell, T. K., and Ambros, V. (2009). nhl-2 Modulates microRNA activity in Caenorhabditis elegans. Cell *136*, 926-938.

Hammell, M., Long, D., Zhang, L., Lee, A., Carmack, C. S., Han, M., Ding, Y., and Ambros, V. (2008). mirWIP: microRNA target prediction based on microRNA-containing ribonucleoprotein-enriched transcripts. Nat. Methods *5*, 813-819.

Hammes, G. G., and Schimmel, P. R. (1970). 2 Rapid Reactions and Transient States. The Enzymes *2*, 67-114.

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

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

Han, B. W., Hung, J. H., Weng, Z., Zamore, P. D., and Ameres, S. L. (2011). The 3'-to-5' Exoribonuclease Nibbler Shapes the 3' Ends of MicroRNAs Bound to Drosophila Argonaute1. Curr. Biol.

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*, 3016-3027.

Han, J., Lee, Y., Yeom, K. H., Nam, J. W., Heo, I., Rhee, J. K., Sohn, S. Y., Cho, Y., Zhang, B. T., and Kim, V. N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887-901.

Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. Nature

Hanson, P. I., and Cashikar, A. (2012). Multivesicular body morphogenesis. Annu Rev Cell Dev Biol *28*, 337-362.

Hartig, J. V., Esslinger, S., Bottcher, R., Saito, K., and Forstemann, K. (2009). Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. EMBO J. *28*, 2932-2944.

Hendrickson, D. G., Hogan, D. J., McCullough, H. L., Myers, J. W., Herschlag, D., Ferrell, J. E., and Brown, P. O. (2009). Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. PLoS Biol. *7*, e1000238.

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*, 1052-1060.

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*, 1757-1766.

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

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

Horwich, M. D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P. D. (2007). The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. Curr. Biol. *17*, 1265-1272.

Huang, J., Liang, Z., Yang, B., Tian, H., Ma, J., and Zhang, H. (2007). Derepression of microRNA-mediated protein translation inhibition by apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3G (APOBEC3G) and its family members. J. Biol. Chem. *282*, 33632-33640.

Huang, X. A., Yin, H., Sweeney, S., Raha, D., Snyder, M., and Lin, H. (2013). A Major Epigenetic Programming Mechanism Guided by piRNAs. Dev. Cell

Huang, Y., Ji, L., Huang, Q., Vassylyev, D. G., Chen, X., and Ma, J. B. (2009). Structural insights into mechanisms of the small RNA methyltransferase HEN1. Nature *461*, 823-827.

Humphreys, D. T., Westman, B. J., Martin, D. I., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. Proc. Natl. Acad. Sci. U S A *102*, 16961-16966.

Hutvagner, G., and Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. Nat. Rev. Mol. Cell. Biol. *9*, 22-32.

Hutvagner, G., and Zamore, P. D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science *297*, 2056-2060.

Iki, T., Yoshikawa, M., Nishikiori, M., Jaudal, M. C., Matsumoto-Yokoyama, E., Mitsuhara, I., Meshi, T., and Ishikawa, M. (2010). In Vitro Assembly of Plant RNA-Induced Silencing Complexes Facilitated by Molecular Chaperone HSP90. Mol. Cell.

Ishizuka, A., Siomi, M. C., and Siomi, H. (2002). A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev. *16*, 2497-2508.

Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J.-B., Lönnerberg, P., and Linnarsson, S. (2011). Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res. *21*, 1160-1167.

Iwasaki, S., Kawamata, T., and Tomari, Y. (2009). Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression. Mol. Cell. *34*, 58-67.

Iwasaki, S., Kobayashi, M., Yoda, M., Sakaguchi, Y., Katsuma, S., Suzuki, T., and Tomari, Y. (2010). Hsc70/Hsp90 Chaperone Machinery Mediates ATP-Dependent RISC Loading of Small RNA Duplexes. Mol. Cell.

Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003). Expression profiling reveals off-target gene regulation by RNAi. Nat. Biotechnol. *21*, 635-637.

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*, 1179-1187.

Jackson, R. J., Hellen, C. U., and Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell. Biol. *11*, 113-127.

Jafarifar, F., Yao, P., Eswarappa, S. M., and Fox, P. L. (2011). Repression of VEGFA by CA-rich element-binding microRNAs is modulated by hnRNP L. EMBO J. *30*, 1324-1334.

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

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

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*, 1674-1679.

John, B., Enright, A. J., Aravin, A., Tuschl, T., Sander, C., and Marks, D. S. (2004). Human MicroRNA targets. PLoS Biol. *2*, e363.

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., and Slack, F. J. (2005). RAS is regulated by the let-7 microRNA family. Cell *120*, 635-647.

Johnston, M., Geoffroy, M. C., Sobala, A., Hay, R., and Hutvagner, G. (2010). HSP90 Protein Stabilizes Unloaded Argonaute Complexes and Microscopic P-bodies in Human Cells. Mol. Biol. Cell

Johnstone, O., Deuring, R., Bock, R., Linder, P., Fuller, M. T., and Lasko, P. (2005). Belle is a Drosophila DEAD-box protein required for viability and in the germ line. Dev. Biol. *277*, 92-101.

Juvvuna, P. K., Khandelia, P., Lee, L. M., and Makeyev, E. V. (2012). Argonaute identity defines the length of mature mammalian microRNAs. Nucleic Acids Res. *40*, 6808-6820.

Kamminga, L. M., Luteijn, M. J., den, B. M. J., Redl, S., Kaaij, L. J., Roovers, E. F., Ladurner, P., Berezikov, E., and Ketting, R. F. (2010). Hen1 is required for oocyte development and piRNA stability in zebrafish. EMBO J. *29*, 3688-3700.

Karres, J. S., Hilgers, V., Carrera, I., Treisman, J., and Cohen, S. M. (2007). The conserved microRNA miR-8 tunes atrophin levels to prevent neurodegeneration in Drosophila. Cell *131*, 136-145.

Karreth, F. A., Tay, Y., Perna, D., Ala, U., Tan, S. M., Rust, A. G., DeNicola, G., Webster, K. A., Weiss, D., Perez-Mancera, P. A., Krauthammer, M., Halaban, R., Provero, P., Adams, D. J., Tuveson, D. A., and Pandolfi, P. P. (2011). In vivo identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. Cell *147*, 382-395.

Kataoka, Y., Takeichi, M., and Uemura, T. (2001). Developmental roles and molecular characterization of a Drosophila homologue of Arabidopsis Argonaute1, the founder of a novel gene superfamily. Genes Cells *6*, 313-325.

Kawamata, T., Seitz, H., and Tomari, Y. (2009). Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. Nat. Struct. Mol. Biol.

Kawamata, T., and Tomari, Y. (2010). Making RISC. Trends Biochem. Sci.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T. N., Siomi, M. C., and Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature *453*, 793-797.

Kedde, M., Strasser, M. J., Boldajipour, B., Vrielink, J. A., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P. M., van Duijse, J., Orom, U. A., Lund, A. H., Perrakis, A., Raz, E., and Agami, R. (2007). RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. Cell *131*, 1273-1286.

Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J. A., Elkon, R., and Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. Nat. Cell Biol. *12*, 1014-1020.

Kennerdell, J. R., and Carthew, R. W. (1998). Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. Cell *95*, 1017-1026.

Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. Nat. Genet. *39*, 1278-1284.

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

Kim, J. K., Gabel, H. W., Kamath, R. S., Tewari, M., Pasquinelli, A., Rual, J. F., Kennedy, S., Dybbs, M., Bertin, N., Kaplan, J. M., Vidal, M., and Ruvkun, G. (2005). Functional genomic analysis of RNA interference in C. elegans. Science *308*, 1164-1167.

Kim, V. N., Han, J., and Siomi, M. C. (2009). Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell. Biol. *10*, 126-139.

Kiriakidou, M., Nelson, P. T., Kouranov, A., Fitziev, P., Bouyioukos, C., Mourelatos, Z., and Hatzigeorgiou, A. (2004). A combined computational-experimental approach predicts human microRNA targets. Genes Dev. *18*, 1165-1178.

Kiriakidou, M., Tan, G. S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P. T., and Mourelatos, Z. (2007). An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell *129*, 1141-1151.

Kirino, Y., and Mourelatos, Z. (2007a). Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. Nat. Struct. Mol. Biol. *14*, 347-348.

Kirino, Y., and Mourelatos, Z. (2007b). The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. RNA *13*, 1397-1401.

Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., and Rajewsky, N. (2005). Combinatorial microRNA target predictions. Nat. Genet. *37*, 495-500.

Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K. G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005). Silencing of microRNAs in vivo with 'antagomirs'. Nature *438*, 685-689.

Kurth, H. M., and Mochizuki, K. (2009). 2'-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in Tetrahymena. RNA *15*, 675-685.

Kwak, P. B., and Tomari, Y. (2012). The N domain of Argonaute drives duplex unwinding during RISC assembly. Nat. Struct. Mol. Biol. *19*, 145-151.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853-858.

Lai, E. C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat. Genet. *30*, 363-364.

Lai, E. C., and Posakony, J. W. (1998). Regulation of Drosophila neurogenesis by RNA:RNA duplexes? Cell *93*, 1103-1104.

Lall, S., Grun, D., Krek, A., Chen, K., Wang, Y. L., Dewey, C. N., Sood, P., Colombo, T., Bray, N., Macmenamin, P., Kao, H. L., Gunsalus, K. C., Pachter, L., Piano, F., and Rajewsky, N. (2006). A genome-wide map of conserved microRNA targets in C. elegans. Curr. Biol. *16*, 460-471.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Curr. Biol. *14*, 2162-2167.

Lanet, E., Delannoy, E., Sormani, R., Floris, M., Brodersen, P., Crete, P., Voinnet, O., and Robaglia, C. (2009). Biochemical evidence for translational repression by Arabidopsis microRNAs. Plant Cell *21*, 1762-1768.

Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858-862.

Lau, N. C., Seto, A. G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D. P., and Kingston, R. E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

Lau, P. W., Guiley, K. Z., De, N., Potter, C. S., Carragher, B., and MacRae, I. J. (2012). The molecular architecture of human Dicer. Nat. Struct. Mol. Biol. *19*, 436-440.

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

Le Thomas, A., Rogers, A. K., Webster, A., Marinov, G. K., Liao, S. E., Perkins, E. M., Hur, J. K., Aravin, A. A., and Toth, K. F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev. *27*, 390-399.

Lecellier, C. H., and Voinnet, O. (2004). RNA silencing: no mercy for viruses? Immunol. Rev. *198*, 285-303.

Lee, R. C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862-864.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, S. R., and Collins, K. (2006). Two classes of endogenous small RNAs in Tetrahymena thermophila. Genes Dev. *20*, 28-33.

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*, 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*, 522-532.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., and Kim, V. N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. *23*, 4051-4060.

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

Lee, Y. S., Pressman, S., Andress, A. P., Kim, K., White, J. L., Cassidy, J. J., Li, X., Lubell, K., Lim do, H., Cho, I. S., Nakahara, K., Preall, J. B., Bellare, P., Sontheimer, E. J., and Carthew, R. W. (2009). Silencing by small RNAs is linked to endosomal trafficking. Nat. Cell Biol. *11*, 1150-1156.

Leung, A. K., Calabrese, J. M., and Sharp, P. A. (2006). Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. Proc. Natl. Acad. Sci. U S A *103*, 18125-18130.

Leung, A. K., Young, A. G., Bhutkar, A., Zheng, G. X., Bosson, A. D., Nielsen, C. B., and Sharp, P. A. (2011). Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. Nat. Struct. Mol. Biol. *18*, 237-244.

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*, 314-320.

Lewis, B. P., Burge, C. B., and Bartel, D. P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15-20.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.

Li, F., Goto, D. B., Zaratiegui, M., Tang, X., Martienssen, R., and Cande, W. Z. (2005). Two novel proteins, dos1 and dos2, interact with rik1 to regulate heterochromatic RNA interference and histone modification. Curr. Biol. *15*, 1448-1457.

Li, J., Yang, Z., Yu, B., Liu, J., and Chen, X. (2005). Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. Curr. Biol. *15*, 1501-1507.

Li, L., Gu, W., Liang, C., Liu, Q., Mello, C. C., and Liu, Y. (2012). The translin-TRAX complex (C3PO) is a ribonuclease in tRNA processing. Nat. Struct. Mol. Biol.

Li, X., and Jin, P. (2010). Roles of small regulatory RNAs in determining neuronal identity. Nat. Rev. Neurosci. *11*, 329-338.

Licatalosi, D. D., Mele, A., Fak, J. J., Ule, J., Kayikci, M., Chi, S. W., Clark, T. A., Schweitzer, A. C., Blume, J. E., Wang, X., Darnell, J. C., and Darnell, R. B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature *456*, 464-469.

Lim, A. K., and Kai, T. (2007). Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster. Proc. Natl. Acad. Sci. U S A *104*, 6714-6719.

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature *433*, 769-773.

Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B., and Bartel, D. P. (2003). The microRNAs of Caenorhabditis elegans. Genes Dev. *17*, 991-1008.

Lima, W. F., Wu, H., Nichols, J. G., Sun, H., Murray, H. M., and Crooke, S. T. (2009). Binding and cleavage specificities of human Argonaute2. J. Biol. Chem. *284*, 26017-26028.

Lin, H., and Spradling, A. C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary. Development *124*, 2463-2476.

Lin, S. Y., Johnson, S. M., Abraham, M., Vella, M. C., Pasquinelli, A., Gamberi, C., Gottlieb, E., and Slack, F. J. (2003). The C elegans hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. Dev. Cell *4*, 639-650.

Lin, S. Y., and Riggs, A. D. (1972). Lac repressor binding to non-operator DNA: detailed studies and a comparison of eequilibrium and rate competition methods. J. Mol. Biol. *72*, 671-690.

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

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

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

Liu, J., Rivas, F. V., Wohlschlegel, J., Yates, J. R. r., Parker, R., and Hannon, G. J. (2005a). A role for the P-body component GW182 in microRNA function. Nat. Cell Biol. *7*, 1261-1266.

Liu, J., Valencia-Sanchez, M. A., Hannon, G. J., and Parker, R. (2005b). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. Nat. Cell Biol. *7*, 719-723.

Liu, Q., Rand, T. A., Kalidas, S., Du, F., Kim, H. E., Smith, D. P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science *301*, 1921-1925.

Liu, Y., Ye, X., Jiang, F., Liang, C., Chen, D., Peng, J., Kinch, L. N., Grishin, N. V., and Liu, Q. (2009). C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. Science *325*, 750-753.

Llave, C., Xie, Z., Kasschau, K. D., and Carrington, J. C. (2002). Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. Science *297*, 2053-2056.

Lohmann, J. U., Endl, I., and Bosch, T. C. (1999). Silencing of developmental genes in Hydra. Dev. Biol. *214*, 211-214.

Long, D., Lee, R., Williams, P., Chan, C. Y., Ambros, V., and Ding, Y. (2007). Potent effect of target structure on microRNA function. Nat. Struct. Mol. Biol. *14*, 287-294.

Loya, C. M., Lu, C. S., Van Vactor, D., and Fulga, T. A. (2009). Transgenic microRNA inhibition with spatiotemporal specificity in intact organisms. Nat Methods *6*, 897-903.

Lu, W. P., and Fei, L. (2003). A logarithmic approximation to initial rates of enzyme reactions. Anal. Biochem. *316*, 58-65.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.

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*, 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*, 666-670.

Maiti, M., Lee, H. C., and Liu, Y. (2007). QIP, a putative exonuclease, interacts with the Neurospora Argonaute protein and facilitates conversion of duplex siRNA into single strands. Genes Dev. *21*, 590-600.

Mallory, A., and Vaucheret, H. (2010). Form, function, and regulation of ARGONAUTE proteins. Plant Cell *22*, 3879-3889.

Mansfield, J. H., Harfe, B. D., Nissen, R., Obenauer, J., Srineel, J., Chaudhuri, A., Farzan-Kashani, R., Zuker, M., Pasquinelli, A. E., Ruvkun, G., Sharp, P. A., Tabin, C. J., and McManus, M. T. (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nat. Genet. *36*, 1079-1083.

Maroney, P. A., Yu, Y., Fisher, J., and Nilsen, T. W. (2006). Evidence that microRNAs are associated with translating messenger RNAs in human cells. Nat. Struct. Mol. Biol. *13*, 1102-1107.

Martinez, J., and Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. Genes Dev. *18*, 975-980.

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*, 607-620.

Matsumoto, K., Minami, M., Shinozaki, F., Suzuki, Y., Abe, K., Zenno, S., Matsumoto, S., and Minami, Y. (2011). Hsp90 is involved in the formation of P-bodies and stress granules. Biochem. Biophys. Res. Commun. *407*, 720-724.

McDowell, J. A., and Turner, D. H. (1996). Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of (rGAGGUCUC)2 by two-dimensional NMR and simulated annealing. Biochemistry *35*, 14077-14089.

Meijer, H. A., Kong, Y. W., Lu, W. T., Wilczynska, A., Spriggs, R. V., Robinson, S. W., Godfrey, J. D., Willis, A. E., and Bushell, M. (2013). Translational Repression and eIF4A2 Activity Are Critical for MicroRNA-Mediated Gene Regulation. Science *340*, 82-85.

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

Meister, G., Landthaler, M., Dorsett, Y., and Tuschl, T. (2004). Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. RNA *10*, 544-550.

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*, 2149-2155.

Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S. D., Gregersen, L. H., Munschauer, M., Loewer, A., Ziebold, U., Landthaler, M., Kocks, C., le Noble, F., and Rajewsky, N. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature

Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., Chen, S., Hannon, G. J., and Qi, Y. (2008). Sorting of small RNAs into Arabidopsis argonaute complexes is directed by the 5' terminal nucleotide. Cell *133*, 116-127.

Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers--the database of key numbers in molecular and cell biology. Nucleic Acids Res. *38*, D750-3.

Miranda, K. C., Huynh, T., Tay, Y., Ang, Y. S., Tam, W. L., Thomson, A. M., Lim, B., and Rigoutsos, I. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell *126*, 1203-1217.

Mishima, Y., Giraldez, A. J., Takeda, Y., Fujiwara, T., Sakamoto, H., Schier, A. F., and Inoue, K. (2006). Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. Curr. Biol. *16*, 2135-2142.

Mlotshwa, S., Pruss, G. J., and Vance, V. (2008). Small RNAs in viral infection and host defense. Trends Plant Sci *13*, 375-382.

Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M. C. (2005). Slicer function of Drosophila Argonautes and its involvement in RISC formation. Genes Dev. *19*, 2837-2848.

Miyoshi, T., Takeuchi, A., Siomi, H., and Siomi, M. C. (2010). A direct role for Hsp90 in pre-RISC formation in Drosophila. Nat. Struct. Mol. Biol. *17*, 1024-1026.

Montgomery, M. K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U S A *95*, 15502-15507.

Montgomery, T. A., Howell, M. D., Cuperus, J. T., Li, D., Hansen, J. E., Alexander, A. L., Chapman, E. J., Fahlgren, N., Allen, E., and Carrington, J. C. (2008). Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. Cell *133*, 128-141.

Moran, U., Phillips, R., and Milo, R. (2010). SnapShot: key numbers in biology. Cell *141*, 1262-1262.e1.

Moshkovich, N., Nisha, P., Boyle, P. J., Thompson, B. A., Dale, R. K., and Lei, E. P. (2011). RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. Genes Dev. *25*, 1686-1701.

Moss, E. G., Lee, R. C., and Ambros, V. (1997). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell *88*, 637-646.

Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. Genes Dev. *16*, 720-728.

Mukherji, S., Ebert, M. S., Zheng, G. X., Tsang, J. S., Sharp, P. A., and van Oudenaarden, A. (2011). MicroRNAs can generate thresholds in target gene expression. Nat. Genet. *43*, 854-859.

Mullokandov, G., Baccarini, A., Ruzo, A., Jayaprakash, A. D., Tung, N., Israelow, B., Evans, M. J., Sachidanandam, R., and Brown, B. D. (2012). High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries. Nat Methods

Myette, J. R., and Niles, E. G. (1996). Domain structure of the vaccinia virus mRNA capping enzyme. Expression in Escherichia coli of a subdomain possessing the RNA 5'-triphosphatase and guanylyltransferase activities and a kinetic comparison to the full-size enzyme. J. Biol. Chem. *271*, 11936-11944.

Nakanishi, K., Weinberg, D. E., Bartel, D. P., and Patel, D. J. (2012). Structure of yeast Argonaute with guide RNA. Nature *486*, 368-374.

Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D., and Grewal, S. I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science *292*, 110-113.

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

Nayak, A., Berry, B., Tassetto, M., Kunitomi, M., Acevedo, A., Deng, C., Krutchinsky, A., Gross, J., Antoniewski, C., and Andino, R. (2010). Cricket paralysis virus antagonizes Argonaute 2 to modulate antiviral defense in Drosophila. Nat. Struct. Mol. Biol. *17*, 547-554.

Neumuller, R. A., Betschinger, J., Fischer, A., Bushati, N., Poernbacher, I., Mechtler, K., Cohen, S. M., and Knoblich, J. A. (2008). Mei-P26 regulates microRNAs and cell growth in the Drosophila ovarian stem cell lineage. Nature

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proc. Natl. Acad. Sci. U S A *95*, 14687-14692.

Nishi, K., Nishi, A., Nagasawa, T., and Ui-Tei, K. (2012). Human TNRC6A is an Argonaute-navigator protein for microRNA-mediated gene silencing in the nucleus. RNA

Nishida, K. M., Okada, T. N., Kawamura, T., Mituyama, T., Kawamura, Y., Inagaki, S., Huang, H., Chen, D., Kodama, T., Siomi, H., and Siomi, M. C. (2009). Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines. EMBO J. *28*, 3820-3831.

Noma, K., Sugiyama, T., Cam, H., Verdel, A., Zofall, M., Jia, S., Moazed, D., and Grewal, S. I. (2004). RITS acts in cis to promote RNA interference-mediated transcriptional and post-transcriptional silencing. Nat. Genet. *36*, 1174-1180.

Noto, T., Kurth, H. M., Kataoka, K., Aronica, L., DeSouza, L. V., Siu, K. W., Pearlman, R. E., Gorovsky, M. A., and Mochizuki, K. (2010). The Tetrahymena argonaute-binding protein Giw1p directs a mature argonaute-siRNA complex to the nucleus. Cell *140*, 692-703.

Nottrott, S., Simard, M. J., and Richter, J. D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. Nat. Struct. Mol. Biol. *13*, 1108-1114.

Nowotny, M. (2009). Retroviral integrase superfamily: the structural perspective. EMBO Rep. *10*, 144-151.

Nowotny, M., and Yang, W. (2009). Structural and functional modules in RNA interference. Curr. Opin. Struct. Biol. *19*, 286-293.

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

O'Carroll, D., Mecklenbrauker, I., Das, P. P., Santana, A., Koenig, U., Enright, A. J., Miska, E. A., and Tarakhovsky, A. (2007). A Slicer-independent role for Argonaute 2 in hematopoiesis and the microRNA pathway. Genes Dev. *21*, 1999-2004.

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007). The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. Nat. Struct. Mol. Biol. *14*, 349-350.

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

Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E. C. (2008a). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. Nat. Struct. Mol. Biol. *15*, 581-590.

Okamura, K., Chung, W. J., Ruby, J. G., Guo, H., Bartel, D. P., and Lai, E. C. (2008b). The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature *453*, 803-806.

Okamura, K., Hagen, J. W., Duan, H., Tyler, D. M., and Lai, E. C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. Cell *130*, 89-100.

Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M. C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev. *18*, 1655-1666.

Okamura, K., and Lai, E. C. (2008). Endogenous small interfering RNAs in animals. Nat. Rev. Mol. Cell. Biol. *9*, 673-678.

Okamura, K., Liu, N., and Lai, E. C. (2009). Distinct mechanisms for microRNA strand selection by Drosophila Argonautes. Mol. Cell. *36*, 431-444.

Olsen, P. H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev. Biol. *216*, 671-680.

Orban, T. I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. RNA *11*, 459-469.

Pall, G. S., and Hamilton, A. J. (2008). Improved northern blot method for enhanced detection of small RNA. Nat Protoc *3*, 1077-1084.

Pare, J. M., Tahbaz, N., Lopez-Orozco, J., LaPointe, P., Lasko, P., and Hobman, T. C. (2009). Hsp90 regulates the function of argonaute 2 and its recruitment to stress granules and P-bodies. Mol. Biol. Cell *20*, 3273-3284.

Parizotto, E. A., Lowe, E. D., and Parker, J. S. (2013). Structural basis for duplex RNA recognition and cleavage by Archaeoglobus fulgidus C3PO. Nat. Struct. Mol. Biol. *20*, 380-386.

Park, W., Li, J., Song, R., Messing, J., and Chen, X. (2002). CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. Curr. Biol. *12*, 1484-1495.

Parker, J. S., Parizotto, E. A., Wang, M., Roe, S. M., and Barford, D. (2009). Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. Mol. Cell. *33*, 204-214.

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*, 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*, 663-666.

Parker, R., and Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. Mol. Cell. *25*, 635-646.

Parrish, S., Fleenor, J., Xu, S., Mello, C., and Fire, A. (2000). Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. Mol. Cell. *6*, 1077-1087.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., Hayward, D. C., Ball, E. E., Degnan, B., Muller, P., Spring, J., Srinivasan, A., Fishman, M., Finnerty, J., Corbo, J., Levine, M., Leahy, P., Davidson, E., and Ruvkun, G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89.

Peattie, D. A., Douthwaite, S., Garrett, R. A., and Noller, H. F. (1981). A “bulged” double helix in a RNA-protein contact site. Proc. Natl. Acad. Sci. U S A *78*, 7331-7335.

Pek, J. W., and Kai, T. (2011). DEAD-box RNA helicase Belle/DDX3 and the RNA interference pathway promote mitotic chromosome segregation. Proc. Natl. Acad. Sci. U S A

Pelisson, A., Sarot, E., Payen-Groschene, G., and Bucheton, A. (2007). A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the Drosophila ovary. J. Virol. *81*, 1951-1960.

Petersen, C. P., Bordeleau, M. E., Pelletier, J., and Sharp, P. A. (2006). Short RNAs repress translation after initiation in mammalian cells. Mol. Cell. *21*, 533-542.

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

Pham, J. W., and Sontheimer, E. J. (2005). Molecular requirements for RNA-induced silencing complex assembly in the Drosophila RNA interference pathway. J. Biol. Chem. *280*, 39278-39283.

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

Pillai, R. S., Bhattacharyya, S. N., Artus, C. G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. Science *309*, 1573-1576.

Pitchiaya, S., Androsavich, J. R., and Walter, N. G. (2012). Intracellular single molecule microscopy reveals two kinetically distinct pathways for microRNA assembly. EMBO Rep.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., and Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature *465*, 1033-1038.

Preall, J. B., He, Z., Gorra, J. M., and Sontheimer, E. J. (2006). Short interfering RNA strand selection is independent of dsRNA processing polarity during RNAi in Drosophila. Curr. Biol. *16*, 530-535.

Rand, T. A., Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. Cell *123*, 621-629.

Rashid, U. J., Paterok, D., Koglin, A., Gohlke, H., Piehler, J., and Chen, J. C. (2007). Structure of Aquifex aeolicus argonaute highlights conformational flexibility of the PAZ domain as a potential regulator of RNA-induced silencing complex function. J. Biol. Chem. *282*, 13824-13832.

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

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.

Reuter, J. S., and Mathews, D. H. (2010). RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics *11*, 129.

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

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, D. P. (2002). Prediction of plant microRNA targets. Cell *110*, 513-520.

Ricci, E. P., Limousin, T., Soto-Rifo, R., Rubilar, P. S., Decimo, D., and Ohlmann, T. (2012). miRNA repression of translation in vitro takes place during 43S ribosomal scanning. Nucleic Acids Res.

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*, 340-349.

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

Robine, N., Lau, N. C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M. D., and Lai, E. C. (2009). A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr. Biol. *19*, 2066-2076.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. Genome Res. *14*, 1902-1910.

Roe, S. M., Prodromou, C., O'Brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1999). Structural basis for inhibition of the Hsp90 molecular chaperone by the antitumor antibiotics radicicol and geldanamycin. J. Med. Chem. *42*, 260-266.

Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. *5*, 725–738.

Rozhkov, N. V., Hammell, M., and Hannon, G. J. (2013). Multiple roles for Piwi in silencing Drosophila transposons. Genes Dev.

Rubin, G. M., and Spradling, A. C. (1982). Genetic transformation of Drosophila with transposable element vectors. Science *218*, 348-353.

Ruby, J. G., Jan, C. H., and Bartel, D. P. (2007). Intronic microRNA precursors that bypass Drosha processing. Nature *448*, 83-86.

Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M. C. (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature *461*, 1296-1299.

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*, e235.

Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M. C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. Genes Dev. *20*, 2214-2222.

Saito, K., Sakaguchi, Y., Suzuki, T., Suzuki, T., Siomi, H., and Siomi, M. C. (2007). Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi- interacting RNAs at their 3' ends. Genes Dev. *21*, 1603-1608.

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P. P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell *146*, 353-358.

Salzman, J., Gawad, C., Wang, P. L., Lacayo, N., and Brown, P. O. (2012). Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. PLoS ONE *7*, e30733.

Sanchez Alvarado, A., and Newmark, P. A. (1999). Double-stranded RNA specifically disrupts gene expression during planarian regeneration. Proc. Natl. Acad. Sci. U S A *96*, 5049-5054.

Sarikaya, D. P., Belay, A. A., Ahuja, A., Dorta, A., Green, D. A. n., and Extavour, C. G. (2012). The roles of cell size and cell number in determining ovariole number in Drosophila. Dev. Biol. *363*, 279-289.

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

Schneider, I. (1972). Cell lines derived from late embryonic stages of Drosophila melanogaster. J Embryol Exp Morphol *27*, 353-365.

Schneider, M. D., Najand, N., Chaker, S., Pare, J. M., Haskins, J., Hughes, S. C., Hobman, T. C., Locke, J., and Simmonds, A. J. (2006). Gawky is a component of cytoplasmic mRNA processing bodies required for early Drosophila development. J. Cell Biol. *174*, 349-358.

Schroeder, S. J., and Turner, D. H. (2009). Optical melting measurements of nucleic acid thermodynamics. Methods Enzymol. *468*, 371-387.

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*, 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., Tomari, Y., and Zamore, P. D. (2004). The RNA-induced silencing complex is a Mg2+-dependent endonuclease. Curr. Biol. *14*, 787-791.

Seggerson, K., Tang, L., and Moss, E. G. (2002). Two genetic circuits repress the Caenorhabditis elegans heterochronic gene lin-28 after translation initiation. Dev. Biol. *243*, 215-225.

Seitz, H. (2009). Redefining microRNA targets. Curr. Biol. *19*, 870-873.

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58-63.

Sen, G. L., and Blau, H. M. (2005). Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. Nat. Cell Biol. *7*, 633-636.

Sienski, G., Dönertas, D., and Brennecke, J. (2012). Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell *151*, 964-980.

Simon, B., Kirkpatrick, J. P., Eckhardt, S., Reuter, M., Rocha, E. A., Andrade-Navarro, M. A., Sehr, P., Pillai, R. S., and Carlomagno, T. (2011). Recognition of 2'-O-Methylated 3'-End of piRNA by the PAZ Domain of a Piwi Protein. Structure *19*, 172-180.

Slack, F. J., Basson, M., Liu, Z., Ambros, V., Horvitz, H. R., and Ruvkun, G. (2000). The lin-41 RBCC gene acts in the C. elegans heterochronic pathway between the let-7 regulatory RNA and the LIN-29 transcription factor. Mol. Cell. *5*, 659-669.

Schneider, M. D., Najand, N., Chaker, S., Pare, J. M., Haskins, J., Hughes, S. C., Hobman, T. C., Locke, J., and Simmonds, A. J. (2006). Gawky is a component of cytoplasmic mRNA processing bodies required for early Drosophila development. J. Cell Biol. *174*, 349-358.

Schwamborn, J. C., Berezikov, E., and Knoblich, J. A. (2009). The TRIM-NHL protein TRIM32 activates microRNAs and prevents self-renewal in mouse neural progenitors. Cell *136*, 913-925.

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*, 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*, 1434-1437.

Souret, F. F., Kastenmayer, J. P., and Green, P. J. (2004). AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. Mol. Cell. *15*, 173-183.

Specchia, V., Piacentini, L., Tritto, P., Fanti, L., D'Alessandro, R., Palumbo, G., Pimpinelli, S., and Bozzetti, M. P. (2010). Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. Nature

Stadler, B. M., and Ruohola-Baker, H. (2008). Small RNAs: keeping stem cells in line. Cell *132*, 563-566.

Stark, A., Brennecke, J., Bushati, N., Russell, R. B., and Cohen, S. M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. Cell *123*, 1133-1146.

Steiner, F. A., Hoogstrate, S. W., Okihara, K. L., Thijssen, K. L., Ketting, R. F., Plasterk, R. H., and Sijen, T. (2007). Structural features of small RNA precursors determine Argonaute loading in Caenorhabditis elegans. Nat. Struct. Mol. Biol. *14*, 927-933.

Steitz, T. A., and Steitz, J. A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc. Natl. Acad. Sci. U S A *90*, 6498-6502.

Steller, H., and Pirrotta, V. (1985). A transposable P vector that confers selectable G418 resistance to Drosophila larvae. EMBO J. *4*, 167-171.

Suh, N., and Blelloch, R. (2011). Small RNAs in early mammalian development: from gametes to gastrulation. Development *138*, 1653-1661.

Tabach, Y., Billi, A. C., Hayes, G. D., Newman, M. A., Zuk, O., Gabel, H., Kamath, R., Yacoby, K., Chapman, B., Garcia, S. M., Borowsky, M., Kim, J. K., and Ruvkun, G. (2013). Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature *493*, 694-698.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123-132.

Tafer, H., Ameres, S. L., Obernosterer, G., Gebeshuber, C. A., Schroeder, R., Martinez, J., and Hofacker, I. L. (2008). The impact of target site accessibility on the design of effective siRNAs. Nat. Biotechnol. *26*, 578-583.

Tahbaz, N., Carmichael, J. B., and Hobman, T. C. (2001). GERp95 belongs to a family of signal-transducing proteins and requires Hsp90 activity for stability and Golgi localization. J. Biol. Chem. *276*, 43294-43299.

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*, 189-194.

Takeda, Y., Mishima, Y., Fujiwara, T., Sakamoto, H., and Inoue, K. (2009). DAZL relieves miRNA-mediated repression of germline mRNAs by controlling poly(A) tail length in zebrafish. PLoS ONE *4*, e7513.

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

Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M., and Hannon, G. J. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature *453*, 534-538.

Tan, G. S., Garchow, B. G., Liu, X., Yeung, J., Morris, J. P. t., Cuellar, T. L., McManus, M. T., and Kiriakidou, M. (2009). Expanded RNA-binding activities of mammalian Argonaute 2. Nucleic Acids Res. *37*, 7533-7545.

Tang, G., Reinhart, B. J., Bartel, D. P., and Zamore, P. D. (2003). A biochemical framework for RNA silencing in plants. Genes Dev. *17*, 49-63.

Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S. M., Ala, U., Karreth, F., Poliseno, L., Provero, P., Di Cunto, F., Lieberman, J., Rigoutsos, I., and Pandolfi, P. P. (2011). Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell *147*, 344-357.

Taylor, J. R. (1997). An Introduction to Error Analysis : The Study of Uncertainties in Physical Measurements (Sausalito, Calif.: University Science Books).

Thermann, R., and Hentze, M. W. (2007). Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. Nature *447*, 875-878.

Thon, G., Hansen, K. R., Altes, S. P., Sidhu, D., Singh, G., Verhein-Hansen, J., Bonaduce, M. J., and Klar, A. J. (2005). The Clr7 and Clr8 directionality factors and the Pcu4 cullin mediate heterochromatin formation in the fission yeast Schizosaccharomyces pombe. Genetics *171*, 1583-1595.

Tian, Y., Simanshu, D. K., Ascano, M., Diaz-Avalos, R., Park, A. Y., Juranek, S. A., Rice, W. J., Yin, Q., Robinson, C. V., Tuschl, T., and Patel, D. J. (2011a). Multimeric assembly and biochemical characterization of the Trax-translin endonuclease complex. Nat. Struct. Mol. Biol.

Tian, Y., Simanshu, D. K., Ma, J. B., and Patel, D. J. (2011b). Inaugural article: Structural basis for piRNA 2'-O-methylated 3'-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains. Proc. Natl. Acad. Sci. U S A *108*, 903-910.

Tijsterman, M., Okihara, K. L., Thijssen, K., and Plasterk, R. H. (2002). PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of C. elegans. Curr. Biol. *12*, 1535-1540.

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

Toledano, H., D'Alterio, C., Czech, B., Levine, E., and Jones, D. L. (2012). The let-7-Imp axis regulates ageing of the Drosophila testis stem-cell niche. Nature *485*, 605-610.

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

Tomari, Y., Du, T., and Zamore, P. D. (2007). Sorting of Drosophila small silencing RNAs. Cell *130*, 299-308.

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

Tomari, Y., and Zamore, P. D. (2005). Perspective: machines for RNAi. Genes Dev. *19*, 517-529.

Ui-Tei, K., Naito, Y., Takahashi, F., Haraguchi, T., Ohki-Hamazaki, H., Juni, A., Ueda, R., and Saigo, K. (2004). Guidelines for the selection of highly effective siRNA sequences for mammalian and chick RNA interference. Nucleic Acids Res. *32*, 936-948.

Ui-Tei, K., Naito, Y., Nishi, K., Juni, A., and Saigo, K. (2008). Thermodynamic stability and Watson-Crick base pairing in the seed duplex are major

determinants of the efficiency of the siRNA-based off-target effect. Nucleic Acids Res. *36*, 7100-7109.

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P. D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.

Vagin, V. V., Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., Girard, A., Sachidanandam, R., Hannon, G. J., and Aravin, A. A. (2009). Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. Genes Dev. *23*, 1749-1762.

Valencia-Sanchez, M. A., Liu, J., Hannon, G. J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev. *20*, 515-524.

Van Bortle, K., and Corces, V. G. (2013). The role of chromatin insulators in nuclear architecture and genome function. Curr. Opin. Genet. Dev.

van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N., and Stuitje, A. R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell *2*, 291-299.

van Wolfswinkel, J. C., Claycomb, J. M., Batista, P. J., Mello, C. C., Berezikov, E., and Ketting, R. F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell *139*, 135-148.

Vastenhouw, N. L., Fischer, S. E., Robert, V. J., Thijssen, K. L., Fraser, A. G., Kamath, R. S., Ahringer, J., and Plasterk, R. H. (2003). A genome-wide screen identifies 27 genes involved in transposon silencing in C. elegans. Curr. Biol. *13*, 1311-1316.

Vazquez, F., Blevins, T., Ailhas, J., Boller, T., and Meins, F. J. (2008). Evolution of Arabidopsis MIR genes generates novel microRNA classes. Nucleic Acids Res. *36*, 6429-6438.

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

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

Voinnet, O. (2009). Origin, biogenesis, and activity of plant microRNAs. Cell *136*, 669-687.

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

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

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

Wang, B., Yanez, A., and Novina, C. D. (2008). MicroRNA-repressed mRNAs contain 40S but not 60S components. Proc. Natl. Acad. Sci. U S A *105*, 5343-5348.

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

Wang, X. B., Jovel, J., Udomporn, P., Wang, Y., Wu, Q., Li, W. X., Gasciolli, V., Vaucheret, H., and Ding, S. W. (2011). The 21-Nucleotide, but Not 22-Nucleotide, Viral Secondary Small Interfering RNAs Direct Potent Antiviral Defense by Two Cooperative Argonautes in Arabidopsis thaliana. Plant Cell *23*, 1625-1638.

Wang, X. H., Aliyari, R., Li, W. X., Li, H. W., Kim, K., Carthew, R., Atkinson, P., and Ding, S. W. (2006b). RNA interference directs innate immunity against viruses in adult Drosophila. Science *312*, 452-454.

Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T., and Patel, D. J. (2008a). Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature *456*, 921-926.

Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G. S., Tuschl, T., and Patel, D. J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature *461*, 754-761.

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

Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., Minami, N., and Imai, H. (2006). Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. Genes Dev. *20*, 1732-1743.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., Surani, M. A., Sakaki, Y., and Sasaki, H. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature *453*, 539-543.

Waterhouse, P. M., Graham, M. W., and Wang, M. B. (1998). Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proc. Natl. Acad. Sci. U S A *95*, 13959-13964.

Wee, L. M., Flores-Jasso, C. F., Salomon, W. E., and Zamore, P. D. (2012). Argonaute Divides Its RNA Guide into Domains with Distinct Functions and RNA-Binding Properties. Cell *151*, 1055-1067.

Weeks, K. M., and Crothers, D. M. (1992). RNA binding assays for Tat-derived peptides: implications for specificity. Biochemistry *31*, 10281-10287.

Welker, N. C., Maity, T. S., Ye, X., Aruscavage, P. J., Krauchuk, A. A., Liu, Q., and Bass, B. L. (2011). Dicer's Helicase Domain Discriminates dsRNA Termini to Promote an Altered Reaction Mode. Mol. Cell. *41*, 589-599.

Wianny, F., and Zernicka-Goetz, M. (2000). Specific interference with gene function by double-stranded RNA in early mouse development. Nat. Cell Biol. *2*, 70-75.

Wienholds, E., Koudijs, M. J., van Eeden, F. J., Cuppen, E., and Plasterk, R. H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nat. Genet. *35*, 217-218.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.

Wilkins, C., Dishongh, R., Moore, S. C., Whitt, M. A., Chow, M., and Machaca, K. (2005). RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. Nature *436*, 1044-1047.

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

Williams, R. W., and Rubin, G. M. (2002). ARGONAUTE1 is required for efficient RNA interference in Drosophila embryos. Proc. Natl. Acad. Sci. U S A *99*, 6889-6894.

Williamson, J. R. (2008). Cooperativity in macromolecular assembly. Nat Chem Biol *4*, 458-465.

Wilson, J. E., Connell, J. E., and Macdonald, P. M. (1996). aubergine enhances oskar translation in the Drosophila ovary. Development *122*, 1631-1639.

Wright, J. E., Gaidatzis, D., Senften, M., Farley, B. M., Westhof, E., Ryder, S. P., and Ciosk, R. (2011). A quantitative RNA code for mRNA target selection by the germline fate determinant GLD-1. EMBO J. *30*, 533-545.

Wu, L., and Belasco, J. G. (2005). Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. Mol. Cell. Biol. *25*, 9198-9208.

Wu, L., Fan, J., and Belasco, J. G. (2006). MicroRNAs direct rapid deadenylation of mRNA. Proc. Natl. Acad. Sci. U S A *103*, 4034-4039.

Wu, L., Zhou, H., Zhang, Q., Zhang, J., Ni, F., Liu, C., and Qi, Y. (2010). DNA methylation mediated by a microRNA pathway. Mol. Cell. *38*, 465-475.

Xia, T., McDowell, J. A., and Turner, D. H. (1997). Thermodynamics of nonsymmetric tandem mismatches adjacent to G.C base pairs in RNA. Biochemistry *36*, 12486-12497.

Xia, T., SantaLucia, J. J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998). Thermodynamic parameters for an expanded

nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry *37*, 14719-14735.

Xie, J., Ameres, S. L., Friedline, R., Hung, J. H., Zhang, Y., Xie, Q., Zhong, L., Su, Q., He, R., Li, M., Li, H., Mu, X., Zhang, H., Broderick, J. A., Kim, J. K., Weng, Z., Flotte, T. R., Zamore, P. D., and Gao, G. (2012). Long-term, efficient inhibition of microRNA function in mice using rAAV vectors. Nat Methods *9*, 403-409.

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*, 468-474.

Yang, D., Lu, H., and Erickson, J. W. (2000). Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in Drosophila embryos. Curr. Biol. *10*, 1191-1200.

Yang, N., and Kazazian, H. H. (2006). L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. Nat. Struct. Mol. Biol. *13*, 763-771.

Yang, W., and Steitz, T. A. (1995). Recombining the structures of HIV integrase, RuvC and RNase H. Structure *3*, 131-134.

Yang, Z., Ebright, Y. W., Yu, B., and Chen, X. (2006). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. Nucleic Acids Res. *34*, 667-675.

Ye, X., Huang, N., Liu, Y., Paroo, Z., Huerta, C., Li, P., Chen, S., Liu, Q., and Zhang, H. (2011). Structure of C3PO and mechanism of human RISC activation. Nat. Struct. Mol. Biol. *18*, 650-657.

Yekta, S., Shih, I. H., and Bartel, D. P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594-596.

Yekta, S., Tabin, C. J., and Bartel, D. P. (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. Nat. Rev. Genet. *9*, 789-796.

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*, 3011-3016.

Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., and Tomari, Y. (2009). ATP-dependent human RISC assembly pathways. Nat. Struct. Mol. Biol.

Yuan, Y. R., Pei, Y., Chen, H. Y., Tuschl, T., and Patel, D. J. (2006). A potential protein-RNA recognition event along the RISC-loading pathway from the structure of A. aeolicus Argonaute with externally bound siRNA. Structure *14*, 1557-1565.

Yuan, Y. R., Pei, Y., Ma, J. B., Kuryavyi, V., Zhadina, M., Meister, G., Chen, H. Y., Dauter, Z., Tuschl, T., and Patel, D. J. (2005). Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. Mol. Cell. *19*, 405-419.

Zamore, P. D., Bartel, D. P., Lehmann, R., and Williamson, J. R. (1999). The PUMILIO-RNA interaction: a single RNA-binding domain monomer recognizes a bipartite target sequence. Biochemistry *38*, 596-604.

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

Zearfoss, N. R., Clingman, C. C., Farley, B. M., McCoig, L. M., and Ryder, S. P. (2011). Quaking regulates Hnrnpa1 expression through its 3' UTR in oligodendrocyte precursor cells. PLoS Genet. *7*, e1001269.

Zekri, L., Kuzuoglu-Ozturk, D., and Izaurralde, E. (2013). GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. EMBO J. *32*, 1052-1065.

Zeng, Y., and Cullen, B. R. (2004). Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. Nucleic Acids Res. *32*, 4776-4785.

Zeng, Y., Wagner, E. J., and Cullen, B. R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol. Cell. *9*, 1327-1333.

Zeng, Y., Yi, R., and Cullen, B. R. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc. Natl. Acad. Sci. U S A *100*, 9779-9784.

Zeng, Y., Yi, R., and Cullen, B. R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. EMBO J. *24*, 138-148.

Zha, X., Xia, Q., and Adam Yuan, Y. (2012). Structural insights into small RNA sorting and mRNA target binding by Arabidopsis Argonaute Mid domains. FEBS Lett.

Zhang, Z., Xu, J., Koppetsch, B. S., Wang, J., Tipping, C., Ma, S., Weng, Z., Theurkauf, W. E., and Zamore, P. D. (2011). Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. Mol. Cell. *44*, 572-584.

Zhou, R., Hotta, I., Denli, A. M., Hong, P., Perrimon, N., and Hannon, G. J. (2008). Comparative analysis of argonaute-dependent small RNA pathways in Drosophila. Mol. Cell. *32*, 592-599.

Appendix I: Equations and Derivations

Binding Equilibrium

Quadratic (Morrison) Equation for Tight Binding

Henri-Michaelis-Menten Kinetics

Briggs Haldane Steady State Kinetics

Henri-Michaelis-Menten Kinetics with Limited Substrate

Henri-Michaelis-Menten Kinetics for Tight Binding Substrate

Competitive Inhibition

Non-Competitive Inhibition

Uncompetitive Inhibition

Quadratic Equations for Competitions

Dissociation Rate Constant

Dissociation Half-Time

BINDING EQUILIBRIUM

Given the following binding scheme,

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES$$

Where E = free enzyme, S = free substrate, k_1 is the bimolecular rate constant with unit $M^{-1}s^{-1}$ and k_{-1} is the unimolecular dissociation rate constant with unit s^{-1}. Assuming a pure reversible binding reaction, the rate of formation of ES complex can be written as,

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES]$$

At equilibrium,

$$\frac{d[ES]}{dt} = 0$$

$$k_1[E][S] - k_{-1}[ES] = 0$$

$$k_1[E][S] = k_{-1}[ES]$$

$$\frac{[E][S]}{[ES]} = \frac{k_{-1}}{k_1} = K_D$$

And K_D is defined as the equilibrium dissociation rate constant. According to the law of mass action, the total enzyme concentration (E_T) is the sum of the free enzyme concentration (E) and the enzyme:substrate complex (ES) where

$$E_T = E + ES$$

Therefore the free enzyme concentration can be defined as

$$E = E_T - ES$$

From the following equation,

$$k_1[\mathrm{E}][\mathrm{S}] - k_{-1}[\mathrm{ES}] = 0$$

Substituting for E, we have

$$k_1[\mathrm{E_T} - \mathrm{ES}][\mathrm{S}] - k_{-1}[\mathrm{ES}] = 0$$

$$k_1[\mathrm{E_T}][\mathrm{S}] - k_1[\mathrm{ES}][\mathrm{S}] - k_{-1}[\mathrm{ES}] = 0$$

$$k_1[\mathrm{E_T}][\mathrm{S}] - k_1[\mathrm{ES}][\mathrm{S}] - k_{-1}[\mathrm{ES}] = 0$$

$$[\mathrm{ES}]\left(k_1[\mathrm{S}] + k_{-1}\right) = k_1[\mathrm{E_T}][\mathrm{S}]$$

Expressing ES relative to E_T,

$$\frac{[\mathrm{ES}]}{[\mathrm{E_T}]} = \frac{k_1[\mathrm{S}]}{k_1[\mathrm{S}] + k_{-1}}$$

$$\frac{[\mathrm{ES}]}{[\mathrm{E_T}]} = \frac{[\mathrm{S}]}{[\mathrm{S}] + \frac{k_{-1}}{k_1}}$$

$$\frac{[\mathrm{ES}]}{[\mathrm{E_T}]} = \frac{[\mathrm{S}]}{[\mathrm{S}] + K_D}$$

When K_D = [S],

$$\frac{[\mathrm{ES}]}{[\mathrm{E_T}]} = \frac{[\mathrm{S}]}{[\mathrm{S}] + [\mathrm{S}]} = \frac{1}{2}$$

In other words, K_D defines the substrate concentration at which half the total enzyme concentration exists as ES complex. This assumes that S is present in excess such that free S = initial S concentration (S_T).

Quadratic Equation for Tight Binding

From the same binding scheme as above,

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES$$

Similarly at equilibrium,

$$k_1[E][S] - k_{-1}[ES] = 0$$

In the event of tight binding, however, E ≠ ET and S ≠ S_T such that

$$E = E_T - ES$$
$$S = S_T - ES$$

We can rewrite the equation at equilibrium as

$$k_1\left([E_T] - [ES]\right)\left([S_T] - [ES]\right) - k_{-1}[ES] = 0$$

$$k_1\left([E_T][S_T] - [E_T][ES] - [ES][S_T] + [ES]^2\right) - k_{-1}[ES] = 0$$

$$[E_T][S_T] - [E_T][ES] - [ES][S_T] + [ES]^2 - K_D[ES] = 0$$

$$[ES]^2 - \left([E_T] + [S_T] + K_D\right)[ES] + [E_T][S_T] = 0$$

Using the quadratic solution to the equation in the form of

$$ax^2 + bx + c = 0$$

And that the solution of x is given by

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Where

$$x = [\mathrm{ES}]$$
$$a = 1$$
$$b = -\left([\mathrm{E_T}] + [\mathrm{S_T}] + K_D\right)$$
$$c = [\mathrm{E_T}][\mathrm{S_T}]$$

Therefore,
∞

$$[\mathrm{ES}] = \frac{\left([\mathrm{E_T}] + [\mathrm{S_T}] + K_D\right) \pm \sqrt{\left([\mathrm{E_T}] + [\mathrm{S_T}] + K_D\right)^2 - 4[\mathrm{E_T}][\mathrm{S_T}]}}{2}$$

There should only be one solution to [ES] and we can rationalize to arrive at one out of two the possible solutions. At very low binding affinity, i.e. $K_D = \infty$, [ES] approaches 0 because it would be unlikely to form ES complex. Following this reasoning, we should expect the signage to be a minus sign instead of a plus sign. Therefore the solution can be rewritten as

$$[\mathrm{ES}] = \frac{\left([\mathrm{E_T}] + [\mathrm{S_T}] + K_D\right) - \sqrt{\left([\mathrm{E_T}] + [\mathrm{S_T}] + K_D\right)^2 - 4[\mathrm{E_T}][\mathrm{S_T}]}}{2}$$

Expressing as *f*: the ratio of ES to total substrate concentration, $\mathrm{S_T}$—as in filter binding—we can recast the equation as

$$f = \frac{\left([\mathrm{E_T}] + [\mathrm{S_T}] + K_D\right) - \sqrt{\left([\mathrm{E_T}] + [\mathrm{S_T}] + K_D\right)^2 - 4[\mathrm{E_T}][\mathrm{S_T}]}}{2[\mathrm{S_T}]}$$

HENRI-MICHAELIS-MENTEN KINETICS

The scheme is an extension of the binding equilibrium where ES complex can be converted to free enzyme E and the formation of product P such that

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

The Michaelis-Menten equation derived in 1913 is based on several assumptions that were used in the Henri's equation in 1903. Henceforth, it is also known as the Henri-Michaelis-Menten equation. These assumptions are

1) The enzyme (E) is a catalyst
2) Only a single substrate (S) is present and forms only a single enzyme:substrate complex
3) Reciprocally, the enzyme:substrate complex (ES) can break down in a reverse reaction to form the starting free enzyme and substrate
4) Substrate is present in excess such that the free substrate concentration at time *t* is equivalent to its initial concentration
5) The interaction between free enzyme and free substrate is very rapid, hence the derivation of the equation is said to assume rapid equilibrium
6) In contrast, the conversion of enzyme:substrate complex into free enzyme and product (P) is slow and therefore is the rate limiting step
7) The measurements is taken at early time points and therefore the reverse reaction of product with the free enzyme is negligible

The rapid equilibrium also known as quasi-equilibrium assumptions meant that only the early components are in equilibrium. i.e.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES$$

Based on the law of mass action,

$$E_T = E + ES$$

Also at equilibrium,

$$k_1[E][S] = k_{-1}[ES]$$

Such that

$$K_S = \frac{k_{-1}}{k_1} = \frac{[E][S]}{[ES]}$$

The rate of the reaction is given by

$$v = k_p[ES]$$

Dividing the velocity equation by E_T,

$$\frac{v}{[E_T]} = \frac{k_p[ES]}{[E] + [ES]}$$

With rearrangement and substituting for free E (equilibrium equation),

$$\frac{v}{[E_T]} = \frac{k_p[ES]}{\frac{K_S[ES]}{[S]} + [ES]}$$

$$\frac{v}{[E_T]} = \frac{k_p[S]}{K_S + [S]}$$

$$v = \frac{k_p[E_T][S]}{K_S + [S]}$$

At maximum enzyme concentration, E_T,

$$v = k_p[E_T]$$

$$v = V_{max}$$

Finally, we arrive at the Henri-Michaelis-Menten equation

$$v = \frac{V_{max}[S]}{K_S + [S]}$$

BRIGGS-HALDANE STEADY STATE KINETICS

The Briggs-Haldane kinetic is a modified version of the Henri-Michaelis-Menten kinetic where a steady state instead of a rapid equilibrium approach is assumed

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

Likewise, we can define the rate of formation of ES complex as

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_2[ES]$$

Similarly, assuming the S is in excess and that free S = S_T, then at equilibrium

$$\frac{d[ES]}{dt} = 0$$

$$k_1[E][S] - k_{-1}[ES] - k_2[ES] = 0$$

Substituting E, and rearranging the equation,

$$k_1[E_T - ES][S] - k_{-1}[ES] - k_2[ES] = 0$$

$$k_1[E_T][S] - k_1[ES][S] - k_{-1}[ES] - k_2[ES] = 0$$

$$\left(k_1[S] + k_{-1} + k_2\right)[ES] = k_1[E_T][S]$$

$$[ES] = \frac{k_1[E_T][S]}{k_1[S] + k_{-1} + k_2}$$

$$[ES] = \frac{[E_T][S]}{[S] + \left(\frac{k_{-1} + k_2}{k_1}\right)}$$

And that

$$\left(\frac{k_{-1} + k_2}{k_1} \right) = K_M$$

Note that K_M from steady state approach of Briggs Haldane ≠ K_S from the rapid equilibrium approach of Henri-Michaelis-Menten.

Substituting for K_M, the equation assumes a final form of

$$[\mathrm{ES}] = \frac{[\mathrm{E_T}][\mathrm{S}]}{[\mathrm{S}] + \left(K_M \right)}$$

The velocity of the reaction can be interpreted as the function of enzyme concentration and the efficiency at which the enzyme makes the product determined by k_2. k_2 is the new kinetic parameter and in Henri-Michaelis-Menten kinetics is also known as the k_{cat}. That said,

$$v = k_2[\mathrm{ES}]$$

or

$$v = \frac{k_2[\mathrm{E_T}][\mathrm{S}]}{[\mathrm{S}] + \left(K_M \right)}$$

And when all enzyme are converted to ES complex, we should expect maximum velocity for product formation, therefore

$$V_{max} = k_2[\mathrm{E_T}]$$

We can rewrite the equation in the final form as

$$v = \frac{V_{max}[\mathrm{S}]}{[\mathrm{S}] + K_M}$$

This will describe a reaction curve with the shape of a rectangular hyperbola, where V_{max} limits the maximum velocity and K_M defines the substrate concentration at which reaction will proceed at half maximal velocity. For instance, when K_M = [S],

$$v = \frac{V_{max}[\mathrm{S}]}{[\mathrm{S}]+[\mathrm{S}]} = \frac{V_{max}}{2}$$

Alternatively, we can view the equation as a product between V_{max} and the fraction of ES complex relative to E_T (see equilibrium binding in earlier section) except that the K_D is being replaced by K_M.

$$v = V_{max} \bullet \frac{[\mathrm{S}]}{[\mathrm{S}]+K_M}$$

Intuitively, this means that the velocity of the reaction is determined by the percentage of total enzyme (determined by [S] and K_M) multiply by the velocity of product formation if all enzymes were available (V_{max}).

It is also important to stress that $K_D \neq K_M$.

$$K_D = \frac{k_{-1}}{k_1}$$

$$K_M = \frac{k_{-1}+k_2}{k_1}$$

However, if we assume rapid equilibrium, such that $k_2 << k_{-1}$, then $K_M \sim K_D$

Michaelis-Menten kinetics with limited substrate

At high [S], the reaction can be view as a pseudo first order reaction. As mentioned earlier,

$$v = k_2[\mathrm{ES}]$$

Therefore, at high [S], the reaction rate can be viewed solely as a function of [ES], where we can eliminate the S term. The rate constant is defined by k_2 with unit of s^{-1}. From here onwards, I will refer k_2 as k_{cat}.

In contrast, when substrate is limiting, from

$$v = \frac{V_{max}[\mathrm{S}]}{[\mathrm{S}] + K_M}$$

We get

$$v = \frac{V_{max}[\mathrm{S}]}{K_M}$$

$$v = \frac{k_{cat}[\mathrm{E_T}][\mathrm{S}]}{K_M}$$

Replacing the kinetic constant, such that

$$k_S = \frac{k_{cat}}{K_M}$$

$$v = k_S[\mathrm{E_T}][\mathrm{S}]$$

k_S is the second-order **specificity rate constant** with unit of $M^{-1}s^{-1}$ {1983, Arch Biochem Biophys, 224, 732-40}.

It follows that,

$$V_{max} = k_{cat}[\mathrm{E_T}]$$

$$V_{max} = k_S K_M[\mathrm{E_T}]$$

HENRI-MICHAELIS-MENTEN KINETICS FOR TIGHT BINDING SUBSTRATE

The principle for deriving the quadratic equation for Henri-Michaelis-Menten kinetics is similar to that of the tight binding reaction at equilibrium (see above). When we have tight binding, we violate the assumption that free substrate remains unchanged i.e. in excess. Likewise, we begin with the following kinetic scheme,

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E + P$$

At equilibrium,

$$\frac{d[ES]}{dt} = 0$$

Such that

$$k_1[E][S] - k_{-1}[ES] - k_2[ES] = 0$$

To account for the changed in free enzyme and free substrate concentration, where

$$[E] = [E_T] - [ES]$$

$$[S] = [S_T] - [ES]$$

We rewrite the change of [ES] at equilibrium as

$$k_1\Big(\big([E_T] - [ES]\big)\big([S_T] - [ES]\big)\Big) - k_{-1}[ES] - k_2[ES] = 0$$

With rearrangement,

$$k_1\Big(\big([E_T]-[ES]\big)\big([S_T]-[ES]\big)\Big)-k_{-1}[ES]-k_2[ES]=0$$

$$k_1[E_T][S_T]-k_1[E_T][ES]-k_1[ES][S_T]+k_1[ES]^2-k_{-1}[ES]-k_2[ES]=0$$

$$k_1[ES]^2-\big(k_1[E_T]+k_1[S_T]+k_{-1}+k_2\big)[ES]+k_1[E_T][S_T]=0$$

$$[ES]^2-\left([E_T]+[S_T]+\frac{k_{-1}+k_2}{k_1}\right)[ES]+[E_T][S_T]=0$$

$$[ES]^2-\big([E_T]+[S_T]+K_M\big)[ES]+[E_T][S_T]=0$$

This takes the form of the quadratic solution for [ES], where

$$\begin{aligned}a&=1\\ b&=-\big([E_T]+[S_T]+K_M\big)\\ c&=[E_T][S_T]\end{aligned}$$

Given that velocity of the reaction is given by

$$v=k_3[ES]$$

Replacing the solved solution of [ES],

$$v=k_3\frac{\big([E_T]+[S_T]+K_M\big)-\sqrt{\big([E_T]+[S_T]+K_M\big)^2-4[E_T][S_T]}}{2}$$

Also, we know that

$$V_{max} = k_3[E_T]$$

$$k_3 = \frac{V_{max}}{[E_T]}$$

Another form of the velocity equation is given by

$$v = V_{max} \bullet \frac{\left([E_T]+[S_T]+K_M\right) - \sqrt{\left([E_T]+[S_T]+K_M\right)^2 - 4[E_T][S_T]}}{2[E_T]}$$

This is known as the velocity equation for tight binding or the Morrison quadratic equation for tight binding {Morrison, 1969, Biochim Biophys Acta, 185, 269-86}.

COMPETITIVE INHIBITION

In the following scheme,

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

$$E + I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EI$$

We can define the equilibrium dissociation constant for the two reactions as

$$K_1 = \frac{k_{-1}}{k_1}$$

$$K_2 = \frac{k_{-2}}{k_2}$$

The rate of reaction is give by

$$v = k_3[ES]$$

Divide both sides of the rate equation by E_T gives

$$\frac{v}{[E_T]} = \frac{k_3[ES]}{[E] + [ES] + [EI]}$$

Defining [ES] and [EI] complex in terms of free E and S,

$$[ES] = \frac{[E][S]}{K_1}$$

$$[EI] = \frac{[E][I]}{K_2}$$

The rate equation will be

$$\frac{v}{[E_T]} = \frac{k_3 \frac{[E][S]}{K_1}}{[E] + \frac{[E][S]}{K_1} + \frac{[E][I]}{K_2}}$$

Rearranging the equation,

$$\frac{v}{[E_T]} = \frac{k_3 \frac{[S]}{K_1}}{1 + \frac{[S]}{K_1} + \frac{[I]}{K_2}}$$

$$\frac{v}{[E_T]} = \frac{k_3[S]}{K_1 + [S] + \frac{K_1[I]}{K_2}}$$

$$\frac{v}{[E_T]} = \frac{k_3[S]}{[S] + K_1\left(1 + \frac{[I]}{K_2}\right)}$$

$$v = \frac{k_3[E_T][S]}{[S] + K_1\left(1 + \frac{[I]}{K_2}\right)}$$

Defining

$$K_I = K_1\left(1 + \frac{[I]}{K_2}\right)$$

Our final equation will be

$$v = \frac{V_{max}[S]}{[S] + K_I}$$

Alternatively, we can derive the velocity equation starting from the equation

$$v = k_3[\mathrm{ES}]$$

From the law of mass action,

$$[\mathrm{E_T}] = [\mathrm{E}] + [\mathrm{ES}] + [\mathrm{EI}]$$

And that

$$[\mathrm{E}] = \frac{K_1[\mathrm{ES}]}{[\mathrm{S}]}$$
$$[\mathrm{EI}] = \frac{[\mathrm{E}][\mathrm{I}]}{K_2}$$

The equation for mass action is now written as

$$[\mathrm{E_T}] = \frac{K_1[\mathrm{ES}]}{[\mathrm{S}]} + [\mathrm{ES}] + \frac{[\mathrm{E}][\mathrm{I}]}{K_2}$$
$$[\mathrm{E_T}] = \frac{K_1[\mathrm{ES}]}{[\mathrm{S}]} + [\mathrm{ES}] + \frac{K_1[\mathrm{ES}][\mathrm{I}]}{[\mathrm{S}]K_2}$$
$$[\mathrm{E_T}] = [\mathrm{ES}]\left(1 + \frac{K_1}{[\mathrm{S}]} + \frac{K_1[\mathrm{I}]}{[\mathrm{S}]K_2}\right)$$
$$[\mathrm{ES}] = \frac{[\mathrm{E_T}]}{1 + \frac{K_1}{[\mathrm{S}]} + \frac{K_1[\mathrm{I}]}{[\mathrm{S}]K_2}}$$
$$[\mathrm{ES}] = \frac{[\mathrm{E_T}][\mathrm{S}]}{[\mathrm{S}] + K_1\left(1 + \frac{[\mathrm{I}]}{K_2}\right)}$$

Therefore,

$$v = \frac{k_3[\mathrm{E_T}][\mathrm{S}]}{[\mathrm{S}] + K_1\left(1 + \frac{[\mathrm{I}]}{K_2}\right)}$$

And that

$$v = \frac{V_{max}[\mathrm{S}]}{[\mathrm{S}] + K_I}$$

where

$$K_I = K_1\left(1 + \frac{[\mathrm{I}]}{K_2}\right)$$

Competitive inhibitors increase the K_M of the enzyme:substrate reaction denoted by K_1. The final apparent K_M is represented by K_I. V_{max} remains unchanged.

NON-COMPETITIVE INHIBITION

In non-competitive inhibition the kinetic scheme is described as follow,

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

$$E + I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EI$$

$$ES + I \underset{k_{-4}}{\overset{k_4}{\rightleftharpoons}} ESI$$

Likewise, the reaction rate for enzyme: substrate is given by

$$v = k_3[ES]$$

Dividing both sides of the equation by E_T, we get

$$\frac{v}{E_T} = \frac{k_3[ES]}{[E] + [ES] + [EI] + [ESI]}$$

We can also define the following variables

$$K_1 = \frac{k_{-1}}{k_1} \qquad [ES] = \frac{[E][S]}{K_1}$$

$$K_2 = \frac{k_{-2}}{k_2} \qquad [EI] = \frac{[E][I]}{K_2}$$

$$K_4 = \frac{k_{-4}}{k_4} \qquad [ESI] = \frac{[ES][I]}{K_4} = \frac{[E][S][I]}{K_1 K_4}$$

The rate equation is now

$$\frac{v}{E_T} = \frac{k_3 \dfrac{[E][S]}{K_1}}{[E] + \dfrac{[E][S]}{K_1} + \dfrac{[E][I]}{K_2} + \dfrac{[E][S][I]}{K_1 K_4}}$$

With rearrangement,

$$\frac{v}{E_T} = \frac{k_3 \frac{[S]}{K_1}}{1 + \frac{[S]}{K_1} + \frac{[I]}{K_2} + \frac{[S][I]}{K_1 K_4}}$$

$$\frac{v}{E_T} = \frac{k_3[S]}{K_1 + [S] + \frac{K_1[I]}{K_2} + \frac{[S][I]}{K_4}}$$

$$\frac{v}{E_T} = \frac{k_3[S]}{[S]\left(1 + \frac{[I]}{K_4}\right) + K_1\left(1 + \frac{[I]}{K_2}\right)}$$

Assuming that the dissociation rate constants of inhibitor to free enzyme, E and the enzyme:substrate complex, ES are the same such that $K_2 = K_4 = K_I$, we can rewrite the equation as

$$\frac{v}{E_T} = \frac{k_3[S]}{[S]\left(1 + \frac{[I]}{K_4}\right) + K_1\left(1 + \frac{[I]}{K_2}\right)}$$

$$\frac{v}{E_T} = \frac{k_3[S]}{[S]\left(1 + \frac{[I]}{K_I}\right) + K_1\left(1 + \frac{[I]}{K_I}\right)}$$

$$v = \frac{k_3[E_T][S]}{[S]\left(1 + \frac{[I]}{K_I}\right) + K_1\left(1 + \frac{[I]}{K_I}\right)}$$

$$v = \frac{V_{max}[S]}{[S]\left(1 + \frac{[I]}{K_I}\right) + K_1\left(1 + \frac{[I]}{K_I}\right)}$$

$$v = \frac{V_{max} \Big/ \left(1 + \frac{[I]}{K_I}\right)[S]}{[S] + K_1}$$

Defining

$$V_{max(app)} = V_{max} \Big/ \left(1 + \frac{[I]}{K_I} \right)$$

We arrive at the final rate equation,

$$v = \frac{V_{max(app)}[S]}{[S] + K_1}$$

Therefore, in non-competitive inhibition, $K_M = K_1$ for enzyme:substrate reaction remains unchanged but the maximum velocity is reduced to $V_{max(app)}$.

UNCOMPETITIVE INHIBITION

In uncompetitive inhibition, the model is such that

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

$$ES + I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} ESI$$

Similarly, the velocity equation for enzyme:substrate reaction is given by

$$v = k_3[ES]$$

With the following variables define as

$$K_1 = \frac{k_{-1}}{k_1} \qquad [ES] = \frac{[E][S]}{K_1}$$

$$K_2 = \frac{k_{-2}}{k_2} \qquad [ESI] = \frac{[ES][I]}{K_2} = \frac{[E][S][I]}{K_1 K_2}$$

We can express the velocity equation as

$$\frac{v}{E_T} = \frac{k_3[ES]}{[E] + [ES] + [ESI]}$$

$$\frac{v}{E_T} = \frac{k_3 \dfrac{[E][S]}{K_1}}{[E] + \dfrac{[E][S]}{K_1} + \dfrac{[E][S][I]}{K_1 K_2}}$$

With rearrangement,

$$\frac{v}{E_T} = \frac{k_3[S]}{K_1 + [S] + \dfrac{[S][I]}{K_2}}$$

$$\frac{v}{E_T} = \frac{k_3[S]}{K_1 + [S]\left(1 + \dfrac{[I]}{K_2}\right)}$$

Redefining K_2 as K_I, we have

$$v = \frac{k_3[\mathrm{E_T}][\mathrm{S}]}{K_1 + [\mathrm{S}]\left(1 + \frac{[\mathrm{I}]}{K_I}\right)}$$

$$v = \frac{V_{max}[\mathrm{S}]}{K_1 + [\mathrm{S}]\left(1 + \frac{[\mathrm{I}]}{K_I}\right)}$$

$$v = \frac{V_{max} \Big/ \left(1 + \frac{[\mathrm{I}]}{K_I}\right)[\mathrm{S}]}{K_1 \Big/ \left(1 + \frac{[\mathrm{I}]}{K_I}\right) + [\mathrm{S}]}$$

With

$$V_{max(app)} = V_{max} \Big/ \left(1 + \frac{[\mathrm{I}]}{K_I}\right)$$

$$K_{1(app)} = K_1 \Big/ \left(1 + \frac{[\mathrm{I}]}{K_I}\right)$$

The final rate equation is

$$v = \frac{V_{max(app)}[\mathrm{S}]}{K_{1(app)} + [\mathrm{S}]}$$

Therefore, in uncompetitive inhibition, both V_{max} and $K_M = K_1$ is changed in the presence of inhibitors and are reflected as $V_{max(app)}$ and $K_{1(app)}$.

QUADRATIC EQUATIONS FOR COMPETITIONS

Give the following scheme,

$$A + B \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} AB,\ K_{D(AB)} = \frac{k_{-1}}{k_1}$$

$$A + X \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} AX,\ K_{D(AX)} = \frac{k_{-2}}{k_2}$$

Where A, B and X represent protein, labeled canonical target and unlabeled competitive target. $K_{D(AB)}$ is the equilibrium dissociation constant for the protein for its target whereas $K_{D(AX)}$ is the equilibrium dissociation constant for the protein for the competitor target.

We also know that

$$K_{D(AB)} = \frac{k_{-1}}{k_1} = \frac{[A][B]}{[AB]}$$

$$K_{D(AX)} = \frac{k_{-2}}{k_2} = \frac{[A][X]}{[AX]}$$

According to the law of mass action,

$$[A_T] = [A] + [AB] + [AX]$$

$$[AX] = [A_T] - [A] - [AB]$$

From

$$K_{D(AX)} = \frac{k_{-2}}{k_2} = \frac{[A][X]}{[AX]}$$

We get

$$K_{D(AX)} = \frac{[A][X]}{[A_T] - [A] - [AB]}$$

Rearranging the equation,

$$K_{D(AX)} = \frac{[A][X]}{[A_T] - [A] - [AB]}$$

$$[A_T]K_{D(AX)} - [A]K_{D(AX)} - [AB]K_{D(AX)} = [A][X]$$

$$-[A][X] - [A]K_{D(AX)} + [A_T]K_{D(AX)} - [AB]K_{D(AX)} = 0$$

$$-\left([X] + K_{D(AX)}\right)[A] = [AB]K_{D(AX)} - [A_T]K_{D(AX)}$$

$$[A] = \frac{[AB]K_{D(AX)} - [A_T]K_{D(AX)}}{-[X] - K_{D(AX)}}$$

$$[A] = \frac{[A_T]K_{D(AX)} - [AB]K_{D(AX)}}{[X] + K_{D(AX)}}$$

Substitute the expression for [A] into the equation that defines $K_{D(AB)}$, From

$$K_{D(AB)} = \frac{[A][B]}{[AB]}$$

$$K_{D(AB)} = \frac{\frac{[A_T]K_{D(AX)} - [AB]K_{D(AX)}}{[X] + K_{D(AX)}}[B]}{[AB]}$$

$$K_{D(AB)} = \frac{[A_T][B]K_{D(AX)} - [AB][B]K_{D(AX)}}{[X][AB] + [AB]K_{D(AX)}}$$

$$[X][AB]K_{D(AB)} + [AB]K_{D(AB)}K_{D(AX)} = [A_T][B]K_{D(AX)} - [B][AB]K_{D(AX)}$$

$$[X][AB]K_{D(AB)} + [AB]K_{D(AB)}K_{D(AX)} + [B][AB]K_{D(AX)} = [A_T][B]K_{D(AX)}$$

$$\left([X]K_{D(AB)} + K_{D(AB)}K_{D(AX)} + [B]K_{D(AX)}\right)[AB] = [A_T][B]K_{D(AX)}$$

Expressing in terms of [AB],

$$[AB] = \frac{[A_T][B]K_{D(AX)}}{[X]K_{D(AB)} + K_{D(AB)}K_{D(AX)} + [B]K_{D(AX)}}$$

$$[AB] = \frac{[A_T][B]}{[X]\frac{K_{D(AB)}}{K_{D(AX)}} + K_{D(AB)} + [B]}$$

$$\frac{[AB]}{[A_T]} = \frac{[B]}{[X]\frac{K_{D(AB)}}{K_{D(AX)}} + K_{D(AB)} + [B]}$$

$$\frac{[AB]}{[A_T]} = \frac{[B]}{\left(1 + \frac{[X]}{K_{D(AX)}}\right)K_{D(AB)} + [B]}$$

At subnanomolar $K_{D(AB)}$ and $K_{D(AX)}$, we have to account for free [B] and free [X] and the equation has to be recast as follow,

$$\frac{[AB]}{[A_T]} = \frac{[B_T] - [AB]}{\left(1 + \frac{[X_T] - [AX]}{K_{D(AX)}}\right)K_{D(AB)} + [B_T] - [AB]}$$

$$[AB] = \frac{[A_T]\left([B_T] - [AB]\right)}{\left(1 + \frac{[X_T] - [AX]}{K_{D(AX)}}\right)K_{D(AB)} + [B_T] - [AB]}$$

To simplify the equation, we assume that $[X_T] >> [AX]$, i.e. $K_{D(AX)}$ is >> $K_{D(AB)}$ {Lin and Riggs, 1972, J Mol Biol, 72, 671-90},

$$[AB] = \frac{[A_T]\left([B_T] - [AB]\right)}{\left(1 + \frac{[X]}{K_{D(AX)}}\right)K_{D(AB)} + [B_T] - [AB]}$$

If we let

$$\Theta = \frac{[AB]}{[B_T]}$$

Where Θ represents the fraction of protein:labeled RNA [AB] complex to the total labeled target $[B_T]$.

$$\Theta = \frac{[A_T] - \frac{[A_T][AB]}{[B_T]}}{\left(1 + \frac{[X]}{K_{D(AX)}}\right) K_{D(AB)} + [B_T] - [AB]}$$

$$\Theta = \frac{[A_T]\left(1 - \Theta\right)}{\left(1 + \frac{[X]}{K_{D(AX)}}\right) K_{D(AB)} + [B_T]\left(1 - \Theta\right)}$$

Rearranging the equation,

$$K_{D(AB)}\Theta + \frac{K_{D(AB)}[X]\Theta}{K_{D(AX)}} + [B_T]\Theta - [B_T]\Theta^2 = [A_T] - [A_T]\Theta$$

$$K_{D(AB)}\Theta + \frac{K_{D(AB)}[X]\Theta}{K_{D(AX)}} + [B_T]\Theta - [B_T]\Theta^2 - [A_T] + [A_T]\Theta = 0$$

$$K_{D(AB)}\Theta + \frac{K_{D(AB)}[X]\Theta}{K_{D(AX)}} + [B_T]\Theta - [B_T]\Theta^2 - [A_T] + [A_T]\Theta = 0$$

$$[B_T]\Theta^2 - K_{D(AB)}\Theta - \frac{K_{D(AB)}[X]\Theta}{K_{D(AX)}} - [B_T]\Theta - [A_T]\Theta + [A_T] = 0$$

$$[B_T]\Theta^2 - \left(K_{D(AB)} + \frac{K_{D(AB)}[X]}{K_{D(AX)}} + [B_T] + [A_T]\right)\Theta + [A_T] = 0$$

Again, Θ can be solved using the quadratic solution,

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Such that,

$$a = [B_T]$$
$$b = -\left(K_{D(AB)} + \frac{K_{D(AB)}[X]}{K_{D(AX)}} + [B_T] + [A_T] \right)$$
$$c = [A_T]$$

Therefore,

$$\Theta = \frac{K_{D(AB)} + \frac{K_{D(AB)}[X]}{K_{D(AX)}} + [B_T] + [A_T] - \sqrt{\left(K_{D(AB)} + \frac{K_{D(AB)}[X]}{K_{D(AX)}} + [B_T] + [A_T] \right)^2 - 4[A_T]}}{2[B_T]}$$

From this earlier equation,

$$\Theta = \frac{[A_T]\left(1-\Theta\right)}{\left(1+\frac{[X]}{K_{D(AX)}}\right)K_{D(AB)} + [B_T]\left(1-\Theta\right)}$$

When $X = X_{1/2}$ such that it is the concentration of unlabeled competitor target RNA that causes half maximal binding between protein A and labeled RNA B, i.e. $\Theta = ½$, and we can rewrite the above equation as

$$\frac{1}{2} = \frac{[A_T]\left(1-\frac{1}{2}\right)}{\left(1+\frac{[X_{\frac{1}{2}}]}{K_{D(AX)}}\right)K_{D(AB)} + [B_T]\left(1-\frac{1}{2}\right)}$$

With arrangement of the equation, we can calculate $K_{D(AX)}$, the equilibrium dissociation of unlabeled competitor RNA to protein A given that we know the equilibrium dissociation of labeled target RNA to protein A $K_{D(AB)}$.

$$\frac{1}{2} = \frac{\frac{[A_T]}{2}}{\left(1+\frac{[X_{\frac{1}{2}}]}{K_{D(AX)}}\right)K_{D(AB)} + \frac{[B_T]}{2}}$$

$$K_{D(AB)} + \frac{K_{D(AB)}[X_{\frac{1}{2}}]}{K_{D(AX)}} + \frac{[B_T]}{2} - [A_T] = 0$$

$$\frac{2K_{D(AB)}K_{D(AX)} + 2K_{D(AB)}[X_{\frac{1}{2}}] + [B_T]K_{D(AX)} - 2K_{D(AX)}[A_T]}{2K_{D(AX)}} = 0$$

$$2K_{D(AB)}K_{D(AX)} + 2K_{D(AB)}[X_{\frac{1}{2}}] + [B_T]K_{D(AX)} - 2K_{D(AX)}[A_T] = 0$$

$$\left(2K_{D(AB)} + [B_T] - 2[A_T]\right)K_{D(AX)} = -2K_{D(AB)}[X_{\frac{1}{2}}]$$

$$K_{D(AX)} = \frac{2K_{D(AB)}[X_{\frac{1}{2}}]}{2[A_T] - 2K_{D(AB)} - [B_T]}$$

Alternatively, we can also express the equation in terms of $X_{1/2}$, also known as the IC_{50},

$$[X_{\frac{1}{2}}] = \frac{\left(2K_{D(AB)} + [B_T] - 2[A_T]\right)K_{D(AX)}}{-2K_{D(AB)}}$$

$$IC_{50} = [X_{\frac{1}{2}}] = \frac{\left(2[A_T] - 2K_{D(AB)} - [B_T]\right)K_{D(AX)}}{2K_{D(AB)}}$$

DISSOCIATION RATE CONSTANT

For the following scheme,

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES$$

Assuming that after we drive assembly of ES complex, we allow it to decay with time, we can express the dissociation of ES complex with time into free E and S as

$$\frac{d[ES]}{dt} = -k_{-1}[ES] + k_1[E][S]$$

For the dissociation experiment, we can prevent the formation of new ES from free E and S by adding competitor to quench either free component. We also ensure that the competitor will not affect how ES dissociates, we can simplify the above equation to

$$\frac{d[ES]}{dt} = -k_{-1}[ES]$$

Rearranging the equation,

$$\frac{1}{[ES]} d[ES] = -k_{-1} dt$$

Taking the integral from time, t = 0 to t = x, such that

$$\int_0^x \frac{1}{[ES]} d[ES] = \int_0^x -k_{-1}\, dt$$

We get

$$\ln[ES]_0^x = [-k_{-1}t]_0^x$$

Expanding and simplifying the equation,

$$\ln[ES]_x - \ln[ES]_0 = [-k_{-1}t]_x - [-k_{-1}t]_0$$

$$\ln[ES]_x - \ln[ES]_0 = [-k_{-1}t]_x$$

$$\ln\frac{[ES]_x}{[ES]_0} = -k_{-1}x$$

$$\ln\frac{[ES]_x}{[ES]_0} = -k_{-1}x\ln e$$

By comparison,

$$\frac{[ES]_x}{[ES]_0} = e^{-k_{-1}x}$$

This means that at t = x, the fraction of [ES] complex relative to the starting total [ES] complex at t = 0, is described by an single exponential curve with a dissociation rate constant k_{-1}.

In the event that complex formed at time zero is not equal to 1 and product dissociation can never be zero at infinite time, the dissociation curve can be normalized and fitted by the following equation,

$$\frac{[ES]_x}{[ES]_0} = \left(f_{\max} - f_{\min}\right)e^{-k_{-1}x} + f_{\min}$$

DISSOCIATION HALF-TIME

We can derive the half-time ($x_{½}$) for a dissociation event. Using the above example, it is the time required for half the ES complex to dissociate into free E and free S. From

$$\frac{[ES]_x}{[ES]_0} = e^{-k_{-1}x}$$

When $[ES]_x = ½[ES]_0$, $x = x_{½}$

$$\frac{1}{2} = e^{-k_{-1}x_{\frac{1}{2}}}$$

Solving for $x_{½}$,

$$\ln\frac{1}{2} = \ln e^{-k_{-1}x_{\frac{1}{2}}}$$

$$\ln\frac{1}{2} = -k_{-1}x_{\frac{1}{2}} \ln e$$

$$\ln\frac{1}{2} = -k_{-1}x_{\frac{1}{2}}$$

$$x_{\frac{1}{2}} = \frac{\ln\frac{1}{2}}{-k_{-1}}$$

$$x_{\frac{1}{2}} = \frac{0.693}{-k_{-1}}$$

Appendix II: *Drosophila* Argonaute2 kinetics and thermodynamics

Scenario 1: Perfect target of *let-7* siRNA guide

K_M = 1 nM (Purified) and 25 nM (Lysate)
k_{cat} = 0.06 s^{-1} (Purified) and 0.02–0.06 s^{-1} (Lysate)
K_D = 4 pM (Purified)
k_{off} = 8.8 x 10^{-5} s^{-1} (Purified)
k_{on} = 2.0 x 10^7 $M^{-1}s^{-1}$ or 6.0 x 10^7 $M^{-1}s^{-1}$ (Purified; calculated from K_D and k_{off} and calculated from k_{cat}/K_M)

Since

$$K_M = \frac{k_{cat} + k_{off}}{k_{on}}$$

Based on k_{cat}, k_{on} and k_{off} measurement, K_M calculated for purified dAgo2 is given by

$$K_M = \frac{k_{cat} + k_{off}}{k_{on}}$$

$$K_M = \frac{0.06 + 8.8 \times 10^{-5}}{2.0 \times 10^7} \text{ or } \frac{0.06 + 8.8 \times 10^{-5}}{6.0 \times 10^7}$$

$$K_M = 3 \text{ nM or } 1 \text{ nM}$$

This agrees with the K_M obtained for purified dAgo2 in Michaelis-Menten kinetics.

Scenario 2: Target of *let-7* siRNA guide with g4g5 mismatch

K_M = 2415 nM (Lysate) and predicted is 96.6 nM for purified (based on perfect target above)
k_{cat} = 0.2–0.6 s^{-1} (Lysate)
K_D = 2.3 nM (Purified)
k_{off} = 3.6 x 10^{-3} s^{-1} (Purified)
k_{on} = 1.6 x 10^6 $M^{-1}s^{-1}$ (Purified; calculated from K_D and k_{off})

Based on k_{cat}, k_{on} and k_{off} measurement, K_M calculated for purified dAgo2 is given by

$$K_M = \frac{k_{cat} + k_{off}}{k_{on}}$$

$$K_M = \frac{0.2 + 3.6 \times 10^{-3}}{1.6 \times 10^6} \text{ or } \frac{0.6 + 3.6 \times 10^{-3}}{1.6 \times 10^6}$$

$$K_M = 127 \text{ nM or } 377 \text{ nM}$$

Scenario 3: Target of *let-7* siRNA guide with g10g11 mismatch

K_M = 8 nM (Lysate for g10g11 mismatch therefore an overestimate) and predicted is 0.32 nM for purified (based on perfect target above)
k_{cat} = 2.0 x 10^{-3} s^{-1} or 6.0 x 10^{-3} s^{-1} (Lysate)
K_D = 4 pM (Purified)
k_{off} = 1.1 x 10^{-3} s^{-1} (Purified)
k_{on} = 2.8 x 10^8 $M^{-1}s^{-1}$ (Purified; calculated from K_D and k_{off})

Based on k_{cat}, k_{on} and k_{off} measurement, K_M calculated for purified dAgo2 is given by

$$K_M = \frac{k_{cat} + k_{off}}{k_{on}}$$

$$K_M = \frac{2.0\times10^{-3} + 1.1\times10^{-3}}{2.8\times10^{8}} \text{ or } \frac{6.0\times10^{-3} + 1.1\times10^{-3}}{2.8\times10^{8}}$$

$$K_M = 0.01 \text{ nM or } 0.03 \text{ nM}$$

Scenario 4: Target of *let-7* siRNA guide with g15g16 mismatch

K_M = 20 nM (Lysate) and predicted is 0.8 nM for purified (based on perfect target above)
k_{cat} = 0.01 s^{-1} (Lysate)
K_D = 1 nM (Purified; based on g15g16 mismatches) or 85 pM (Purified; based on g2–g10 base pairing)
k_{off} = 2.6 x 10^{-2} s^{-1} (Purified; based on g2–g10 base pairing)
k_{on} = 2.6 x 10^{7} $M^{-1}s^{-1}$ (Purified; calculated from K_D and k_{off}) or 3.1 x 10^{8} $M^{-1}s^{-1}$

Based on k_{cat}, k_{on} and k_{off} measurement, K_M calculated for purified dAgo2 is given by

$$K_M = \frac{k_{cat} + k_{off}}{k_{on}}$$

$$K_M = \frac{0.01 + 2.6 \times 10^{-2}}{2.6 \times 10^{7}} \text{ or } \frac{0.01 + 2.6 \times 10^{-2}}{3.1 \times 10^{8}}$$

$$K_M = 1.4 \text{ nM or } 0.12 \text{ nM}$$

Kinetic and thermodynamic parameters for 4 different scenarios.

	Perfect Target	g4g5 mismatch	g10g11 mismatch	g15g16 mismatch
K_M (nM); Lysate	25	2415	8	20
K_M (nM); Purified dAgo2	1	96.6 (predicted)	0.32	0.8
k_{cat} (s^{-1})	0.06	0.2	2.0×10^{-3}	0.01
K_D (nM)	4×10^{-3}	2.3	4×10^{-3}	1
k_{off} (s^{-1})	8.8×10^{-5}	3.6×10^{-3}	1.1×10^{-3}	2.6×10^{-2}
k_{on} ($M^{-1}s^{-1}$)	5.0×10^{7}	1.6×10^{6}	2.8×10^{8}	2.6×10^{7}
Calculated K_M (nM)	1	127	0.01	1.4

Appendix III: *Drosophila* Ago2 has at least one rate-limiting step after seed pairing

Given the following reaction scheme,

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ES \xrightarrow{k_{cat}} E + P$$

At steady state,

$$\frac{dES}{dt} = 0$$

$$k_{on}[E][S] - k_{off}[ES] - k_{cat}[ES] = 0$$

Rearranging the equation to get [ES],

$$[ES] = \frac{k_{on}[E][S]}{k_{off} + k_{cat}}$$

The rate of the reaction is given by

$$k_{cat}[ES]$$

Which when we substitute for [ES] gives,

$$\frac{k_{on}k_{cat}[E][S]}{k_{off} + k_{cat}}$$

$$\frac{k_{on}k_{cat}}{k_{off} + k_{cat}}[E][S]$$

$$k_{obs}[E][S]$$

where

$$k_{obs} = \frac{k_{on}k_{cat}}{k_{off} + k_{cat}}$$

Units of $k_{on} = M^{-1}s^{-1}$; $k_{cat} = s^{-1}$; $k_{off} = s^{-1}$

Therefore the units for k_{obs} is

$$\frac{(M^{-1}s^{-1}) \bullet s^{-1}}{s^{-1} + s^{-1}}$$
$$= M^{-1}s^{-1}$$

Scenario 1: $k_{off} << k_{cat}$

From

$$k_{obs} = \frac{k_{on}k_{cat}}{k_{off} + k_{cat}}$$

We can simplify the equation to

$$k_{obs} = k_{on}$$

This means that when the reaction is very efficient, i.e. k_{cat} is so large that the formation of product occurs much faster than the dissociation of the ES complex i.e. k_{off} is so small. In this case, the rate of the reaction is limited by how fast the substrate binds the enzyme indicated by k_{on}.

Scenario 2: $k_{off} >> k_{cat}$

From

$$k_{obs} = \frac{k_{on}k_{cat}}{k_{off} + k_{cat}}$$

The equation can be recast as

$$\begin{aligned} k_{obs} &= \frac{k_{on}k_{cat}}{k_{off}} \\ &= \frac{k_{on}}{k_{off}}k_{cat} \\ &= \frac{k_{cat}}{K_D} \end{aligned}$$

Therefore, in this case, the reaction rate is defined by how fast the enzyme and substrate make successful contact, the stability of the E:S complex (dissociation constant: K_D) and the catalytic rate (k_{cat}).

To calculate the rate of the second step, k_{cat}, we rearrange the equation as follows,

$$\begin{aligned} k_{obs} &= \frac{k_{on}k_{cat}}{k_{off} + k_{cat}} \\ k_{obs}\left(k_{off} + k_{cat}\right) &= k_{on}k_{cat} \\ k_{obs}k_{off} + k_{obs}k_{cat} &= k_{on}k_{cat} \\ k_{on}k_{cat} - k_{obs}k_{cat} &= k_{obs}k_{off} \\ k_{cat}\left(k_{on} - k_{obs}\right) &= k_{obs}k_{off} \\ k_{cat} &= \frac{k_{obs}k_{off}}{k_{on} - k_{obs}} \end{aligned}$$

The unit for k_{cat} is,

$$\frac{M^{-1}s^{-1} \bullet s^{-1}}{M^{-1}s^{-1} - M^{-1}s^{-1}}$$

$$= s^{-1}$$

For fly Ago2, using binding data for seed-only and a perfect target as follow, we can calculate a predicted value for k_{cat},

$k_{on(seed)}$: 2.1 x 10^8 $M^{-1}s^{-1}$
$k_{off(seed)}$: 0.045 s^{-1}
$k_{obs} = k_{on(fully\ paired)}$: 2.4 x 10^7 $M^{-1}s^{-1}$ or 5.9 x 10^7 $M^{-1}s^{-1}$

Therefore, putting all the variables to calculate k_{cat},

$$k_{cat} = \frac{2.4 \times 10^7 M^{-1}s^{-1} \bullet 0.045 s^{-1}}{2.1 \times 10^8 M^{-1}s^{-1} - 2.4 \times 10^7 M^{-1}s^{-1}}$$

$$= 0.0058 s^{-1}$$

Alternatively,

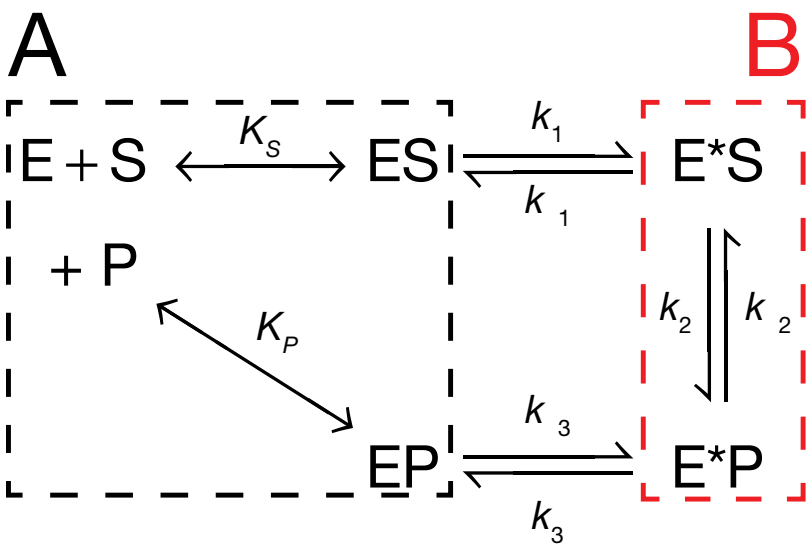
$$k_{cat} = \frac{5.9 \times 10^7 M^{-1}s^{-1} \bullet 0.045 s^{-1}}{2.1 \times 10^8 M^{-1}s^{-1} - 5.9 \times 10^7 M^{-1}s^{-1}}$$

$$= 0.018 s^{-1}$$

The k_{cat} value derived from target cleavage by purified fly Ago2 is 0.06 s^{-1} (Figure 3.3E; Fly Ago2 in embryo lysate, k_{cat} = 0.018–0.06 s^{-1}). As comparison, the k_{cat} value calculated from binding data (above) is smaller than or similar to the catalytic rate measured by target cleavage. This suggests that base pairing beyond the seed or structural rearrangement of fly Ago2 into an active enzyme is rate limiting.

Appendix IV: Kinetic Model of RNAi by *Drosophila* Argonaute2

Given the following reaction model,



E: enzyme (dAgo2); E*: activated enzyme; S: substrate (mRNA target) and P: product (cleaved mRNA target).

To simplify the derivation of the rate equation, we assume segments A and B are in rapid equilibrium with dissociation equilibrium constant K_S and K_P. The conversion among segments A and B are at steady states with rate constants k_1, k_{-1}, k_2, k_{-2}, k_3, k_{-3}.

For each segment, we can define the fraction of enzyme that is free (E) and the fraction that is in complex with substrate (ES or E*S) or product (E*P).

For segment A,

The fraction of enzyme:substrate (ES) complex f_{A1} is given by

$$\frac{[ES]}{[E]+[ES]+[EP]} = \frac{\frac{[E][S]}{K_S}}{[E]+\frac{[E][S]}{K_S}+\frac{[E][P]}{K_P}} = \frac{\frac{[S]}{K_S}}{1+\frac{[S]}{K_S}+\frac{[P]}{K_P}} = \frac{\frac{[S]}{K_S}}{D_A}$$

And the fraction of enzyme:product (EP) complex f_{A2} is given by

$$\frac{[EP]}{[E]+[ES]+[EP]} = \frac{\frac{[E][P]}{K_P}}{[E]+\frac{[E][S]}{K_S}+\frac{[E][P]}{K_P}} = \frac{\frac{[P]}{K_P}}{1+\frac{[S]}{K_S}+\frac{[P]}{K_P}} = \frac{\frac{[P]}{K_P}}{D_A}$$

For segment B.

The fraction of activated enzyme:substrate (E*S) complex, f_{B1} is given by

$$\frac{[E^*S]}{[E^*S]+[E^*P]} = \frac{\frac{k_{-2}[E^*P]}{k_2}}{\frac{k_{-2}[E^*P]}{k_2}+[E^*P]} = \frac{\frac{k_{-2}[E^*P]}{k_2}}{\frac{k_{-2}[E^*P]}{k_2}+\frac{k_2[E^*P]}{k_2}} = \frac{k_{-2}}{k_{-2}+k_2}$$

And the fraction of activated enzyme:product (E*P) complex, f_{B2} is given by

$$\frac{[E^*P]}{[E^*S]+[E^*P]} = \frac{\frac{k_2[E^*S]}{k_{-2}}}{\frac{k_2[E^*S]}{k_{-2}}+[E^*S]} = \frac{\frac{k_2[E^*S]}{k_{-2}}}{\frac{k_2[E^*S]}{k_{-2}}+\frac{k_{-2}}{k_{-2}}} = \frac{k_2}{k_{-2}+k_2}$$

With these definitions, we can simply the reaction scheme to the following,

$$A \underset{k_{-1}{}^*f_{B1} + f_{B2}k_3}{\overset{k_1{}^*f_{A1} + f_{A2}k_3}{\rightleftharpoons}} B$$

We can define the ratio of A to E_T as

$$\frac{A}{E_T} = \frac{f_{B1}k_{-1}+f_{B2}k_3}{E_T}$$

And the ratio of B to E_T as

$$\frac{B}{E_T} = \frac{f_{A1}k_1 + f_{A2}k_{-3}}{E_T}$$

where

$$E_T = f_{A1}k_1 + f_{A2}k_{-3} + f_{B1}k_{-1} + f_{B2}k_3$$

Defining the rate limiting step,

$$v = k_2[E^*S]$$

Therefore,

$$\frac{v}{E_T} = \frac{k_2[E^*S]}{E_T}$$
$$\frac{v}{k_2E_T} = \frac{[E^*S]}{E_T}$$

This is equivalent to

$$\frac{v}{k_2E_T} = \frac{f_{B1}{}^*B}{A+B}$$

Substituting for A and B, we arrive at

$$\frac{v}{k_2E_T} = \frac{f_{B1}{}^* f_{A1}k_1 + f_{A2}k_{-3}}{f_{A1}k_1 + f_{A2}k_{-3} + f_{B1}k_{-1} + f_{B2}k_3}$$
$$\frac{v}{k_2E_T} = \frac{f_{B1}f_{A1}k_1 + f_{B1}f_{A2}k_{-3}}{f_{A1}k_1 + f_{A2}k_{-3} + f_{B1}k_{-1} + f_{B2}k_3}$$

Expanding the terms further,

$$\frac{v}{k_2 E_T} = \frac{\frac{k_{-2}}{k_{-2}+k_2}\frac{\frac{[S]}{K_S}}{D_A}k_1 + \frac{k_{-2}}{k_{-2}+k_2}\frac{\frac{[P]}{K_P}}{D_A}k_{-3}}{\frac{\frac{[S]}{K_S}}{D_A}k_1 + \frac{\frac{[P]}{K_P}}{D_A}k_{-3} + \frac{k_{-2}}{k_{-2}+k_2}k_{-1} + \frac{k_2}{k_{-2}+k_2}k_3}$$

Simplifying the equation,

$$\frac{v}{k_2 E_T} = \frac{\frac{k_{-2}}{k_{-2}+k_2}\frac{[S]}{K_S}k_1 + \frac{k_{-2}}{k_{-2}+k_2}\frac{[P]}{K_P}k_{-3}}{\frac{[S]}{K_S}k_1 + \frac{[P]}{K_P}k_{-3} + \frac{D_A k_{-2}}{k_{-2}+k_2}k_{-1} + \frac{D_A k_2}{k_{-2}+k_2}k_3}$$

Rearranging the equation and grouping the S and P terms,

$$\frac{v}{k_2 E_T} = \frac{\left(\frac{k_{-2}k_1}{k_{-2}+k_2}\right)\frac{[S]}{K_S} + \left(\frac{k_{-2}k_{-3}}{k_{-2}+k_2}\right)\frac{[P]}{K_P}}{\frac{[S]}{K_S}k_1 + \frac{[P]}{K_P}k_{-3} + \left(1+\frac{[S]}{K_S}+\frac{[P]}{K_P}\right)\left(\frac{k_{-1}k_{-2}}{k_{-2}+k_2}\right) + \left(1+\frac{[S]}{K_S}+\frac{[P]}{K_P}\right)\left(\frac{k_3 k_2}{k_{-2}+k_2}\right)}$$

$$\frac{v}{k_2 E_T} = \frac{\left(\frac{k_{-2}k_1}{k_{-2}+k_2}\right)\frac{[S]}{K_S} + \left(\frac{k_{-2}k_{-3}}{k_{-2}+k_2}\right)\frac{[P]}{K_P}}{\frac{[S]}{K_S}k_1 + \frac{[P]}{K_P}k_{-3} + \frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{[S]}{K_S}\frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{[P]}{K_P}\frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{k_3 k_2}{k_{-2}+k_2} + \frac{[S]}{K_S}\frac{k_3 k_2}{k_{-2}+k_2} + \frac{[P]}{K_P}\frac{k_3 k_2}{k_{-2}+k_2}}$$

$$\frac{v}{k_2 E_T} = \frac{\left(\frac{k_{-2}k_1}{k_{-2}+k_2}\right)\frac{[S]}{K_S} + \left(\frac{k_{-2}k_{-3}}{k_{-2}+k_2}\right)\frac{[P]}{K_P}}{\frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{k_3 k_2}{k_{-2}+k_2} + \frac{[S]}{K_S}k_1 + \frac{[S]}{K_S}\frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{[S]}{K_S}\frac{k_3 k_2}{k_{-2}+k_2} + \frac{[P]}{K_P}k_{-3} + \frac{[P]}{K_P}\frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{[P]}{K_P}\frac{k_3 k_2}{k_{-2}+k_2}}$$

$$\frac{v}{k_2 E_T} = \frac{\left(\frac{k_{-2}k_1}{k_{-2}+k_2}\right)\frac{[S]}{K_S} + \left(\frac{k_{-2}k_{-3}}{k_{-2}+k_2}\right)\frac{[P]}{K_P}}{\left(\frac{k_{-1}k_{-2}+k_3 k_2}{k_{-2}+k_2}\right) + \frac{[S]}{K_S}\left(k_1 + \frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{k_3 k_2}{k_{-2}+k_2}\right) + \frac{[P]}{K_P}\left(k_{-3} + \frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{k_3 k_2}{k_{-2}+k_2}\right)}$$

Further simplification leads to,

$$\frac{v}{k_2 E_T} = \frac{\left(\frac{k_{-2}k_1}{k_{-2}+k_2}\right)\frac{[S]}{K_S} + \left(\frac{k_{-2}k_{-3}}{k_{-2}+k_2}\right)\frac{[P]}{K_P}}{\left(\frac{k_{-1}k_{-2}+k_3k_2}{k_{-2}+k_2}\right) + \frac{[S]}{K_S}\left(\frac{k_1k_{-2}+k_1k_2}{k_{-2}+k_2} + \frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{k_3k_2}{k_{-2}+k_2}\right) + \frac{[P]}{K_P}\left(\frac{k_{-3}k_{-2}+k_{-3}k_2}{k_{-2}+k_2} + \frac{k_{-1}k_{-2}}{k_{-2}+k_2} + \frac{k_3k_2}{k_{-2}+k_2}\right)}$$

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S}k_{-2}k_1 + \frac{[P]}{K_P}k_{-2}k_{-3}}{\left(k_{-1}k_{-2}+k_3k_2\right) + \frac{[S]}{K_S}\left(k_1k_{-2}+k_1k_2+k_{-1}k_{-2}+k_3k_2\right) + \frac{[P]}{K_P}\left(k_{-3}k_{-2}+k_{-3}k_2+k_{-1}k_{-2}+k_3k_2\right)}$$

We can rewrite the equation as

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S}N_S + \frac{[P]}{K_P}N_P}{Const + \frac{[S]}{K_S}Coeff_S + \frac{[P]}{K_P}Coeff_P}$$

$$N_S = k_{-2}k_1$$
$$N_P = k_{-2}k_{-3}$$
$$Const = k_{-1}k_{-2} + k_3k_2$$
$$Coeff_S = k_1k_{-2} + k_1k_2 + k_{-1}k_{-2} + k_3k_2$$
$$Coeff_P = k_{-3}k_{-2} + k_{-3}k_2 + k_{-1}k_{-2} + k_3k_2$$

Scenario 1: When P = 0, as we are measuring reaction at very early timepoints.

The rate equation can be simplify to

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S}N_S}{Const + \frac{[S]}{K_S}Coeff_S}$$

Rearranging the equation,

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{Coeff_S} N_S}{\frac{Const * K_S}{Coeff_S} + [S]}$$

$$v = \frac{k_2 E_T * \frac{[S]}{Coeff_S} N_S}{\frac{Const * K_S}{Coeff_S} + [S]}$$

$$v = \frac{\frac{k_2 E_T N_S}{Coeff_S} [S]}{\frac{Const * K_S}{Coeff_S} + [S]}$$

Therefore,

$$V_{max} = \frac{k_2 E_T N_S}{Coeff_S} = \frac{E_T k_2 k_{-2} k_1}{k_1 k_{-2} + k_1 k_2 + k_{-1} k_{-2} + k_3 k_2}$$

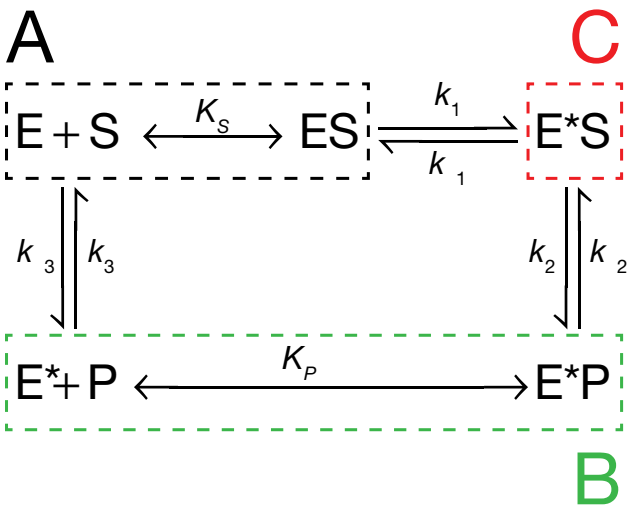
$$k_{cat} = \frac{k_2 N_S}{Coeff_S} = \frac{k_2 k_{-2} k_1}{k_1 k_{-2} + k_1 k_2 + k_{-1} k_{-2} + k_3 k_2}$$

$$K_M = \frac{Const * K_S}{Coeff_S} = \frac{K_S \left(k_{-1} k_{-2} + k_3 k_2 \right)}{k_1 k_{-2} + k_1 k_2 + k_{-1} k_{-2} + k_3 k_2}$$

$$\frac{k_{cat}}{K_M} = \frac{k_2 N_S}{Const * K_S} = \frac{k_2 k_{-2} k_1}{K_S \left(k_{-1} k_{-2} + k_3 k_2 \right)}$$

Appendix V: Kinetic Model of RNAi by *Drosophila* Argonaute2

Given the following reaction model,



E: enzyme (dAgo2); E*: activated enzyme; S: substrate (mRNA target) and P: product (cleaved mRNA target).

To simplify the derivation of the rate equation, we assume segments A and B are in rapid equilibrium with dissociation equilibrium constant K_S and K_P. The conversion among segments A, B and C are at steady states with rate constants k_1, k_{-1}, k_2, k_{-2}, k_3, k_{-3}.

For each segment, we can define the fraction of enzyme that is free (E) and the fraction that is in complex with substrate (ES or E*S) or product (E*P).

For segment A,

The fraction of free enzyme f_{A1} is given by

$$\frac{1}{1+[S]/K_S} = \frac{1}{D_A}$$

And the fraction of ES complex is given by $f_{A2} = 1 - f_{A1}$ is given by

$$\frac{[S]/K_S}{1+[S]/K_S} = \frac{[S]/K_S}{D_A}$$

Likewise, for segment B.

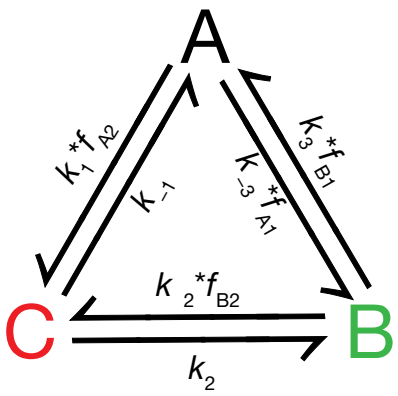
The fraction of free activated enzyme E*, f_{B1} is given by

$$\frac{1}{1+[P]/K_P} = \frac{1}{D_B}$$

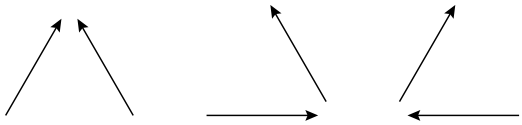
And the fraction of E*P complex is given by $f_{B2} = 1 - f_{B1}$ is given by

$$\frac{[P]/K_P}{1+[P]/K_P} = \frac{[P]/K_P}{D_B}$$

Finally for segment C, 100% are in E*S complex, i.e. $f_C = 1$. With these definitions, we can simply the reaction scheme to the following,

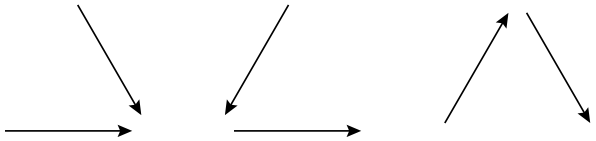


Using the King Altman method, we can define the ratio of A to E_T as



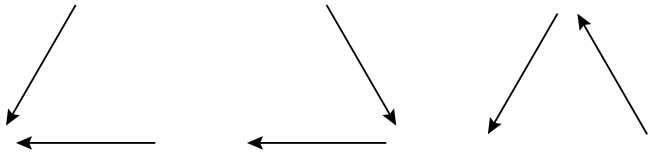
$$\frac{A}{E_T} = \frac{f_{B1}k_{-1}k_3 + f_{B1}k_2k_3 + f_{B2}k_{-1}k_{-2}}{E_T}$$

And the ratio of B to E_T,



$$\frac{B}{E_T} = \frac{f_{A1}k_2k_{-3} + f_{A2}k_1k_2 + f_{A1}k_{-1}k_{-3}}{E_T}$$

And the ratio of C to E_T,



$$\frac{C}{E_T} = \frac{f_{A2}f_{B2}k_1k_{-2} + f_{A1}f_{B2}k_{-2}k_{-3} + f_{A2}f_{B1}k_1k_3}{E_T}$$

Where $E_T = A + B + C$

$$E_T = f_{B1}k_{-1}k_3 + f_{B1}k_2k_3 + f_{B2}k_{-1}k_{-2} + f_{A1}k_2k_{-3} + f_{A2}k_1k_2 + f_{A1}k_{-1}k_{-3} + f_{A2}f_{B2}k_1k_{-2} + f_{A1}f_{B2}k_{-2}k_{-3} + f_{A2}f_{B1}k_1k_3$$

Defining the rate limiting step,

$$v = k_2[E^*S]$$

Therefore,

$$\frac{v}{k_2E_T} = \frac{[E^*S]}{E_T}$$

Substituting [E*S] as C gives

$$\frac{v}{k_2 E_T} = \frac{f_{A2} f_{B2} k_1 k_{-2} + f_{A1} f_{B2} k_{-2} k_{-3} + f_{A2} f_{B1} k_1 k_3}{E_T}$$

$$\frac{v}{k_2 E_T} = \frac{f_{A2} f_{B2} k_1 k_{-2} + f_{A1} f_{B2} k_{-2} k_{-3} + f_{A2} f_{B1} k_1 k_3}{f_{B1} k_{-1} k_3 + f_{B1} k_2 k_3 + f_{B2} k_{-1} k_{-2} + f_{A1} k_2 k_{-3} + f_{A2} k_1 k_2 + f_{A1} k_{-1} k_{-3} + f_{A2} f_{B2} k_1 k_{-2} + f_{A1} f_{B2} k_{-2} k_{-3} + f_{A2} f_{B1} k_1 k_3}$$

Replacing for f_{A1}, f_{A2}, f_{B1} and f_{B2},

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]/K_S}{D_A}\frac{[P]/K_P}{D_B} k_1 k_{-2} + \frac{1}{D_A}\frac{[P]/K_P}{D_B} k_{-2} k_{-3} + \frac{[S]/K_S}{D_A}\frac{1}{D_B} k_1 k_3}{\frac{1}{D_B} k_{-1} k_3 + \frac{1}{D_B} k_2 k_3 + \frac{[P]/K_P}{D_B} k_{-1} k_{-2} + \frac{1}{D_A} k_2 k_{-3} + \frac{[S]/K_S}{D_A} k_1 k_2 + \frac{1}{D_A} k_{-1} k_{-3} + \frac{[S]/K_S}{D_A}\frac{[P]/K_P}{D_B} k_1 k_{-2} + \frac{1}{D_A}\frac{[P]/K_P}{D_B} k_{-2} k_{-3} + \frac{[S]/K_S}{D_A}\frac{1}{D_B} k_1 k_3}$$

Simplifying the equation,

$$\frac{v}{k_2 E_T} = \frac{\frac{[S][P]}{K_S K_P} k_1 k_{-2} + \frac{[P]}{K_P} k_{-2} k_{-3} + \frac{[S]}{K_S} k_1 k_3}{D_A k_{-1} k_3 + D_A k_2 k_3 + \frac{D_A [P]}{K_P} k_{-1} k_{-2} + D_B k_2 k_{-3} + \frac{D_B [S]}{K_S} k_1 k_2 + D_B k_{-1} k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P} k_1 k_{-2} + \frac{[P]}{K_P} k_{-2} k_{-3} + \frac{[S]}{K_S} k_1 k_3}$$

Grouping the P and S terms together,

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S} k_1 k_3 + \frac{[P]}{K_P} k_{-2} k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P} k_1 k_{-2}}{D_A k_{-1} k_3 + D_A k_2 k_3 + D_B k_2 k_{-3} + D_B k_{-1} k_{-3} + \frac{D_B [S]}{K_S} k_1 k_2 + \frac{[S]}{K_S} k_1 k_3 + \frac{D_A [P]}{K_P} k_{-1} k_{-2} + \frac{[P]}{K_P} k_{-2} k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P} k_1 k_{-2}}$$

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S} k_1 k_3 + \frac{[P]}{K_P} k_{-2} k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P} k_1 k_{-2}}{\left(D_A k_{-1} k_3 + D_A k_2 k_3 + D_B k_2 k_{-3} + D_B k_{-1} k_{-3}\right) + \frac{[S]}{K_S}\left(D_B k_1 k_2 + k_1 k_3\right) + \frac{[P]}{K_P}\left(D_A k_{-1} k_{-2} + k_{-2} k_{-3}\right) + \frac{[S]}{K_S}\frac{[P]}{K_P} k_1 k_{-2}}$$

Expanding the term D_A and D_B gives,

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S} k_1 k_3 + \frac{[P]}{K_P} k_{-2} k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P} k_1 k_{-2}}{\left(1 + \frac{[S]}{K_S}\right)\left(k_{-1} k_3 + k_2 k_3\right) + \left(1 + \frac{[P]}{K_P}\right)\left(k_2 k_{-3} + k_{-1} k_{-3}\right) + \frac{[S]}{K_S}\left(\left(1 + \frac{[P]}{K_P}\right) k_1 k_2 + k_1 k_3\right) + \frac{[P]}{K_P}\left(\left(1 + \frac{[S]}{K_S}\right) k_{-1} k_{-2} + k_{-2} k_{-3}\right) + \frac{[S]}{K_S}\frac{[P]}{K_P} k_1 k_{-2}}$$

Expanding gives

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S}k_1k_3 + \frac{[P]}{K_P}k_{-2}k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P}k_1k_{-2}}{k_{-1}k_3 + k_2k_3 + \frac{[S]}{K_S}k_{-1}k_3 + \frac{[S]}{K_S}k_2k_3 + k_2k_{-3} + k_{-1}k_{-3} + \frac{[P]}{K_P}k_2k_{-3} + \frac{[P]}{K_P}k_{-1}k_{-3} + \frac{[S]}{K_S}k_1k_2 + \frac{[S]}{K_S}k_1k_3 + \frac{[S]}{K_S}\frac{[P]}{K_P}k_1k_2 + \frac{[P]}{K_P}k_{-1}k_{-2} + \frac{[P]}{K_P}k_{-2}k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P}k_{-1}k_{-2} + \frac{[S]}{K_S}\frac{[P]}{K_P}k_1k_{-2}}$$

Group the S and P terms together,

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S}k_1k_3 + \frac{[P]}{K_P}k_{-2}k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P}k_1k_{-2}}{k_{-1}k_3 + k_2k_3 + k_2k_{-3} + k_{-1}k_{-3} + \frac{[S]}{K_S}k_{-1}k_3 + \frac{[S]}{K_S}k_2k_3 + \frac{[S]}{K_S}k_1k_2 + \frac{[S]}{K_S}k_1k_3 + \frac{[P]}{K_P}k_2k_{-3} + \frac{[P]}{K_P}k_{-1}k_{-3} + \frac{[P]}{K_P}k_{-1}k_{-2} + \frac{[P]}{K_P}k_{-2}k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P}k_1k_2 + \frac{[S]}{K_S}\frac{[P]}{K_P}k_{-1}k_{-2} + \frac{[S]}{K_S}\frac{[P]}{K_P}k_1k_{-2}}$$

And simplifying,

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S}k_1k_3 + \frac{[P]}{K_P}k_{-2}k_{-3} + \frac{[S]}{K_S}\frac{[P]}{K_P}k_1k_{-2}}{\left(k_{-1}k_3 + k_2k_3 + k_2k_{-3} + k_{-1}k_{-3}\right) + \frac{[S]}{K_S}\left(k_{-1}k_3 + k_2k_3 + k_1k_2 + k_1k_3\right) + \frac{[P]}{K_P}\left(k_2k_{-3} + k_{-1}k_{-3} + k_{-1}k_{-2} + k_{-2}k_{-3}\right) + \frac{[S]}{K_S}\frac{[P]}{K_P}\left(k_1k_2 + k_{-1}k_{-2} + k_1k_{-2}\right)}$$

Replacing the terms to give,

$$N_S = k_1k_3$$
$$N_P = k_{-2}k_{-3}$$
$$N_{SP} = k_1k_{-2}$$
$$Const = k_{-1}k_3 + k_2k_3 + k_2k_{-3} + k_{-1}k_{-3}$$
$$Coeff_S = k_{-1}k_3 + k_2k_3 + k_1k_2 + k_1k_3$$
$$Coeff_P = k_2k_{-3} + k_{-1}k_{-3} + k_{-1}k_{-2} + k_{-2}k_{-3}$$
$$Coeff_{SP} = k_1k_2 + k_{-1}k_{-2} + k_1k_{-2}$$

$$\frac{v}{k_2 E_T} = \frac{\frac{[S]}{K_S}N_S + \frac{[P]}{K_P}N_P + \frac{[S]}{K_S}\frac{[P]}{K_P}N_{SP}}{Const + \frac{[S]}{K_S}Coeff_S + \frac{[P]}{K_P}Coeff_P + \frac{[S]}{K_S}\frac{[P]}{K_P}Coeff_{SP}}$$

Scenario 1: When P = 0, as we are measuring reaction at very early time points.

The rate equation can be simplify to

$$\frac{v}{k_2 \mathrm{E_T}} = \frac{\frac{[\mathrm{S}]}{K_\mathrm{S}} N_S}{Const + \frac{[\mathrm{S}]}{K_\mathrm{S}} Coeff_S}$$

Rearranging the equation,

$$\frac{v}{k_2 \mathrm{E_T}} = \frac{\frac{[\mathrm{S}]}{Coeff_S} N_S}{\frac{Const * K_\mathrm{S}}{Coeff_S} + [\mathrm{S}]}$$

$$v = \frac{k_2 \mathrm{E_T} * \frac{[\mathrm{S}]}{Coeff_S} N_S}{\frac{Const * K_\mathrm{S}}{Coeff_S} + [\mathrm{S}]}$$

$$v = \frac{\frac{k_2 \mathrm{E_T} N_S}{Coeff_S} [\mathrm{S}]}{\frac{Const * K_\mathrm{S}}{Coeff_S} + [\mathrm{S}]}$$

Therefore,

$$V_{\max} = \frac{k_2 \mathrm{E_T} N_S}{Coeff_S} = \frac{\mathrm{E_T} k_2 k_1 k_3}{k_{-1} k_3 + k_2 k_3 + k_1 k_2 + k_1 k_3}$$

$$k_{cat} = \frac{k_2 N_S}{Coeff_S} = \frac{k_2 k_1 k_3}{k_{-1} k_3 + k_2 k_3 + k_1 k_2 + k_1 k_3}$$

$$K_M = \frac{Const * K_\mathrm{S}}{Coeff_S} = \frac{K_\mathrm{S}\left(k_{-1} k_3 + k_2 k_3 + k_2 k_{-3} + k_{-1} k_{-3}\right)}{k_{-1} k_3 + k_2 k_3 + k_1 k_2 + k_1 k_3}$$

$$\frac{k_{cat}}{K_M} = \frac{k_2 N_S}{Const * K_\mathrm{S}} = \frac{k_2 k_1 k_3}{K_\mathrm{S}\left(k_{-1} k_3 + k_2 k_3 + k_2 k_{-3} + k_{-1} k_{-3}\right)}$$

Published Manuscripts

Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Structural Biology 387

Structure and function of nematode RNA-binding proteins

Ebru Kaymak, LM Wee and Sean P Ryder

RNA-binding proteins are critical effectors of gene expression. They guide mRNA localization, translation, and stability, and potentially play a role in regulating mRNA synthesis. The structural basis for RNA recognition by RNA-binding proteins is the key to understand how they target specific transcripts for regulation. Compared to other metazoans, nematode genomes contain a significant expansion in several RNA-binding protein families, including Pumilio-FBF (PUF), TTP-like zinc finger (TZF), and Argonaute-like (AGO) proteins. Genetic data suggest that individual members of each family have distinct functions, presumably due to sequence variations that alter RNA-binding specificity or protein interaction partners. In this review, we highlight example structures and identify the variable regions that likely contribute to functional divergence in nematodes.

Address
Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01605, USA

Corresponding author: Ryder, Sean P (Sean.Ryder@umassmed.edu)

Current Opinion in Structural Biology 2010, **20**:305 312

This review comes from a themed issue on
Nucleic acids
Edited by Joseph Puglisi and James Williamson

Available online 24th April 2010

0959 440X/$ see front matter

DOI 10.1016/j.sbi.2010.03.010

Introduction

RNA regulation is pervasive and impacts nearly every aspect of gene expression. RNA molecules function as both regulators and targets in diverse pathways to ensure appropriate decoding of the genome. RNA-binding proteins are central to this form of regulation. They act as effectors of RNA stability and translation efficiency, they guide transcripts to defined locations within a cell, they control the fidelity of gene decoding, and they function as cofactors to promote the activity of functional and structural RNA molecules.

The facile genetics, defined cellular lineage, and ease of observation have made the nematode *Caenorhabditis elegans* a popular model to study RNA regulatory mechanisms. A scan of the *C. elegans* and other nematode genomes reveals a surprising expansion of putative RNA-binding protein relative to other metazoans. For example, the RNA-binding protein Pumilio discovered in flies has two homologs in humans but 11 homologs in *C. elegans* [1,2]. The CCCH-type tandem zinc finger (TZF) family, typified by the mammalian protein tristetraprolin (TTP), has 16 members in worms [3 6]. Finally, there are 27 Argonaute homologs in *C. elegans*, including a clade of worm-specific Argonautes (WAGOs) [7,8].

It is not clear why RNA-binding protein families have expanded in nematodes. Forward and reverse genetic experiments indicate that many play distinct roles in germline development, gametogenesis, and early embryogenesis, where the regulation of maternal RNAs plays a primary role. In this review, we outline representative structures from the PUF, TZF, and AGO families, and highlight data that identify the basis for specialized function in the expanded set of nematode homologs.

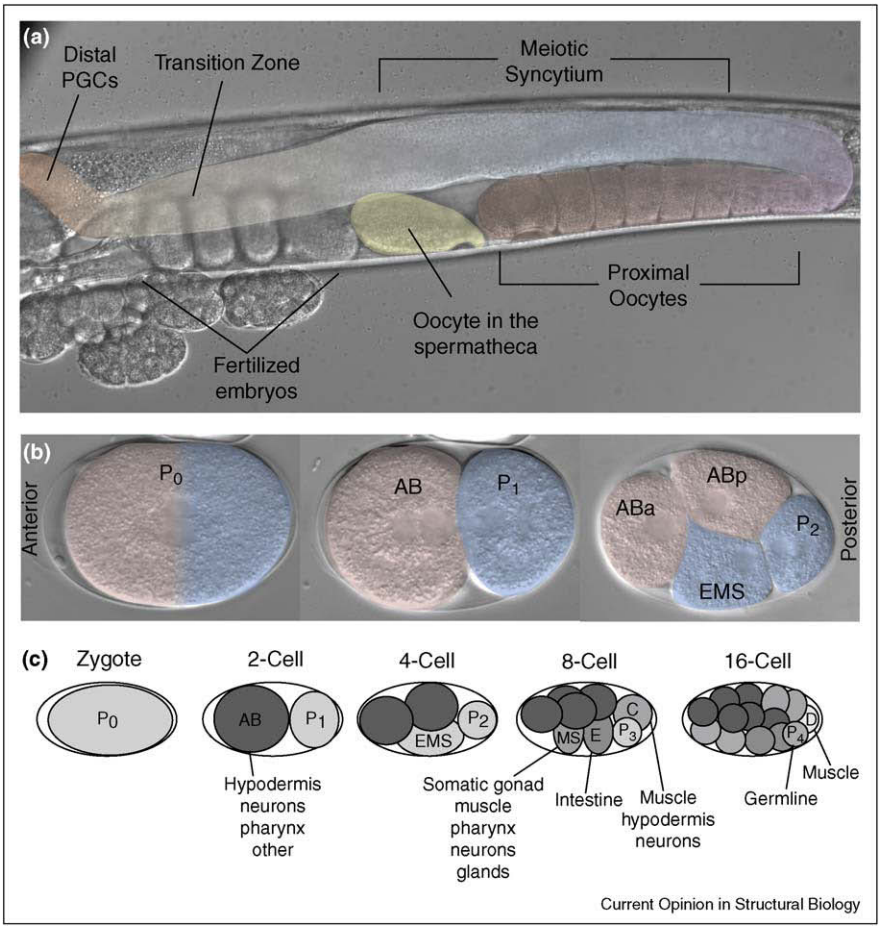
The PUF family

PUF proteins in nematode germline development

The *fem 3* binding factor (FBF) was the first Pumilio homolog identified in *C. elegans* [2]. Pumilio and FBF together comprise the founding members of the PUF family of RNA-binding proteins. FBF is encoded by two nearly identical genes, *fbf 1* and *fbf 2*. Together, they act to maintain the population of progenitor cells in the distal region of the germline and promote the switch from spermatogenesis to oogenesis at the onset of adulthood [2,9] (Figure 1). FBF binds in a sequence-specific fashion to the 3′ untranslated region (UTR) of several messenger RNAs, including *fem 3* and *gld 1* [10]. GLD-1 and FEM-3 promote spermatocyte differentiation, and GLD-1 promotes entry into meiosis [11 13]. FBF represses translation of *gld 1* mRNA in the distal end of the germline, and it represses translation of *gld 1* and *fem 3* mRNA in developing oocytes [2,9].

Nine additional *puf* genes, termed *puf 3* to *puf 12*, are present in the *C. elegans* genome. Most have distinct biological functions defined by phenotypic differences, mRNA target specificity, or expression pattern. Three of these genes *puf 5*, *puf 6*, and *puf 7* are redundantly required for embryonic viability and oocyte maturation [14]. They prevent premature translation of *glp 1* mRNA in oocytes. PUF-8 promotes mitosis in germline progenitor cells, similar to FBF, but binds to RNA with different sequence specificity and as such likely regulates a distinct set of target mRNAs. PUF-9 regulates hunchback-like (*hbl 1*) mRNA in the hypodermis and ventral nerve cord [15]. RNAi screens reveal important roles for PUF-3, PUF-4, PUF-11, and PUF-12 in oogenesis and early embryonic development, but their critical mRNA targets have not been identified [16,17]. In the following

Figure 1



Anatomy of *C. elegans* hermaphrodite reproduction. **(a)** A single gonad arm from a hermaphrodite worm is shown. The gonad is highlighted in false color. The distal arm of the germline contains mitotically dividing progenitor cells (red). There is a transition (orange) from mitosis to meiosis concurrent with a transition from a single celled state to a syncytial region (blue). Meiotic nuclei recellularize, first to form spermatocytes in the L4 larval stage that are stored in the spermatheca (yellow), and then switch to form oocytes (purple) at the onset of adulthood. **(b)** Pattern of the first two cellular divisions after fertilization. The anterior and posterior poles are marked. **(c)** Pattern of division and early lineage of embryogenesis. Several founder cells are established early in embryogenesis that go on to form different tissues in the adult. Adapted with permission from [50].

sections, we review a recently published crystal structure of FBF and highlight biochemical experiments that define differences in RNA recognition in this family [18••].

Biochemical insights into PUF binding specificity

Wickens and co-workers have dissected the RNA-binding properties of several PUF proteins [10,19••,20,21]. The consensus sequence recognized by FBF, termed the FBF binding element (FBE), is 5′-UGURNNAUA-3′ [10]. The FBE is nine nucleotides in length and is partially degenerate at three positions. FBEs are present in the 3′-UTR of *fem 3*, *gld 1*, and numerous other mRNAs regulated by FBF in the germline. Mutation of the FBE in the 3′-UTR of *fem 3* leads to derepression of FEM-3 and failure to switch from spermatogenesis to oogenesis [11].

PUF-8 and PUF-9, on the other hand, recognize an eight nucleotide consensus identical to that bound by human Pum1 (5′-UGUANAUA-3′) termed the Nanos Response Element (NRE) [21,22]. The NRE is similar to the FBE but is a single nucleotide shorter. This difference is critical, as FBF discriminates between these two elements by more than 30-fold. Intriguingly, the specificity of PUF-8 can be converted to that of FBF by swapping a 64-amino acid fragment in the middle of the PUF domain, demonstrating that this region is critical for specificity.

PUF-5 and PUF-6/7 recognize a longer, partially degenerate consensus motif termed the PUF-5 binding element (5BE: 5′-CyCUGUAyyyUGU-3′, where y is a pyrimidine) [20]. PUF-11 binds three sets of RNA targets, 5′-CUGUGAAUA-3′, 5′-CUGUANAAUA-3′ and

5′-NUGUNAAAUA-3′, suggesting multiple modes of RNA recognition through a mechanism that is not immediately apparent [19••]. Clearly, these experiments show that the nematode PUF family has diverged to expand the repertoire of sequences recognized by the PUF domain. Recent crystal structures begin to address the molecular basis for this variance.

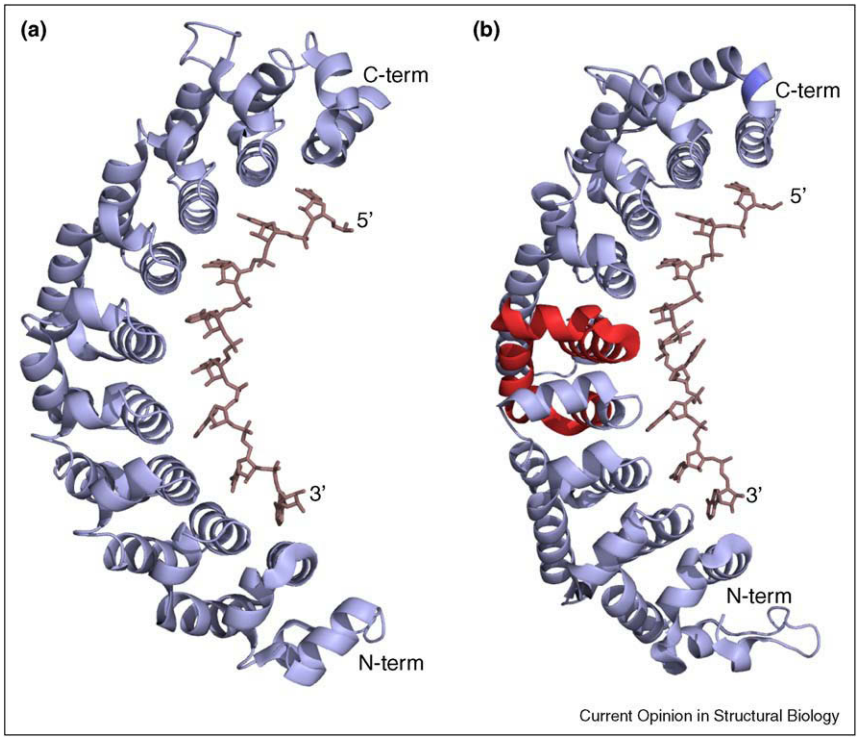
Crystal structures of PUF proteins

The first structures of a PUF domain, including *Drosophila* Pumilio and human Pum1, were determined independently in 2001 [23,24]. The structures revealed an architecture of eight repeat motifs comprising three alpha helices. The repeats pack against one another to form an extended curved structure that vaguely resembles a banana. A subsequent structure of human Pum1 bound to RNA demonstrates that the concave surface comprises the RNA-binding interface, where each repeat recognizes a single nucleotide (Figure 2a) [22]. The amino acids that face the concave surface define the nucleotide specificity at each repeat, which has been reviewed previously [25]. This architecture immediately suggests a model where PUF proteins bind to RNA with modular specificity, such that changing the order of the repeats could modify RNA-binding specificity. Several experiments with chimeric PUF proteins support this model and suggest that this domain is particularly amenable to protein engineering [19••,20,21].

All of the nematode PUF proteins comprise eight repeats, but many bind to a consensus element that contains more than eight nucleotides. To gain insight into the structural basis for recognition of longer elements by this domain, Hall and co-workers crystallized FBF-2 in complex with six different RNA sequences, including four naturally occuring sites [18••]. This study reveals that FBF has an elongated structure with less curvature relative to other PUF domain proteins (Figure 2b). This elongated structure enables a single base to flip out and point away from the protein without affecting interactions with the other eight nucleotides. Thus, a slight variance of the curvature of the overall structure, governed by repeats 4 6, has a profound impact on the RNA-binding specificity.

It is possible that curvature-driven base flipping accounts for the multiple modes of RNA recognition by PUF-11 [19••]. Each mode contains two conserved regions, including a UGU trinucleotide and an element comprising AAAUA. These may represent eight ‘core’ interactions with this protein. Each mode contains an insertion of one

Figure 2



Crystal structures of PUF domain proteins bound to RNA. **(a)** The structure of human Pum1 bound to the NRE. Each PUF repeat specifies a single nucleotide [22]. **(b)** Structure of *C. elegans* FBF bound to the FBE, similar to the NRE but containing an additional nucleotide [18••]. Here, the reduced curvature of the protein enables recognition of a nine nucleotide consensus with an eight repeat PUF domain. One nucleotide flips away from the binding surface of the protein, and eight form interactions with each repeat. The region of the protein that confers FBF like specificity when swapped into PUF 8 is shown in red.

or more nucleotides into a distinct position within this core. If the inserted nucleotides flip out, a model similar to FBF would resolve the multiple modes of binding. The incomplete degeneracy of the inserted nucleotides may be partially explained by differential stacking free energy with neighboring nucleotides. A similar model could be proposed for PUF-5/6/7, where eight nucleotides are specified unambiguously, and five more nucleotides are partially degenerate [20]. More structural work is needed to assess this hypothesis and define the basis for the variance in PUF specificity. It is also important to assess whether conformational flexibility contributes to binding specificity.

TTP-like CCCH tandem zinc finger proteins

TZF proteins in *C. elegans* early embryogenesis

TTP is a mammalian RNA-binding protein that regulates the immune response by promoting the turnover of the mRNA encoding the pro-inflammatory cytokine TNF-alpha [26,27]. TTP is an AU-rich element (ARE) binding protein, which coordinates the stability of mRNAs containing extended repeats of UAUU in their 3′ UTRs. TTP has two $CX_8CX_5CX_3H$ zinc finger motifs. Each motif binds to a single UAUU repeat [28].

There are several TTP paralogs in the *C. elegans* genome, many of which are required for worm fertility. A cascade of TZF proteins, including OMA-1/2, MOE-3, MEX-5/6, MEX-1, POS-1, and PIE-1, guide the progression from the oocyte to embryo. OMA-1/2 and MOE-3 are partially redundant factors that promote oocyte maturation, and inhibit embryonic gene expression before fertilization [4 6,29]. MEX-5 and MEX-6 are required for anterior patterning in the early embryo [5]. They are translated from maternally supplied mRNA shortly after fertilization, and migrate to the anterior of the embryo before the first cellular division. POS-1, PIE-1, and MEX-1 are also translated after fertilization, but accumulate in the posterior of the embryo in a pathway that depends upon MEX-5/6 anterior localization [30,31]. All three proteins are required for posterior patterning and segregation of germline and somatic lineages, but have non-redundant functions [4,29]. In addition to these well-studied examples, there are eight additional TZF genes in the *C. elegans* genome. DCT-13 and possibly Y116A8C.20 promote germline tumor formation in a sensitized genetic background, while CCCH-1, CCCH-2, CCCH-5, F38C2.7, Y116A8C.19, and C35D6.4 have no known function [32].

NMR structure of a TZF family protein

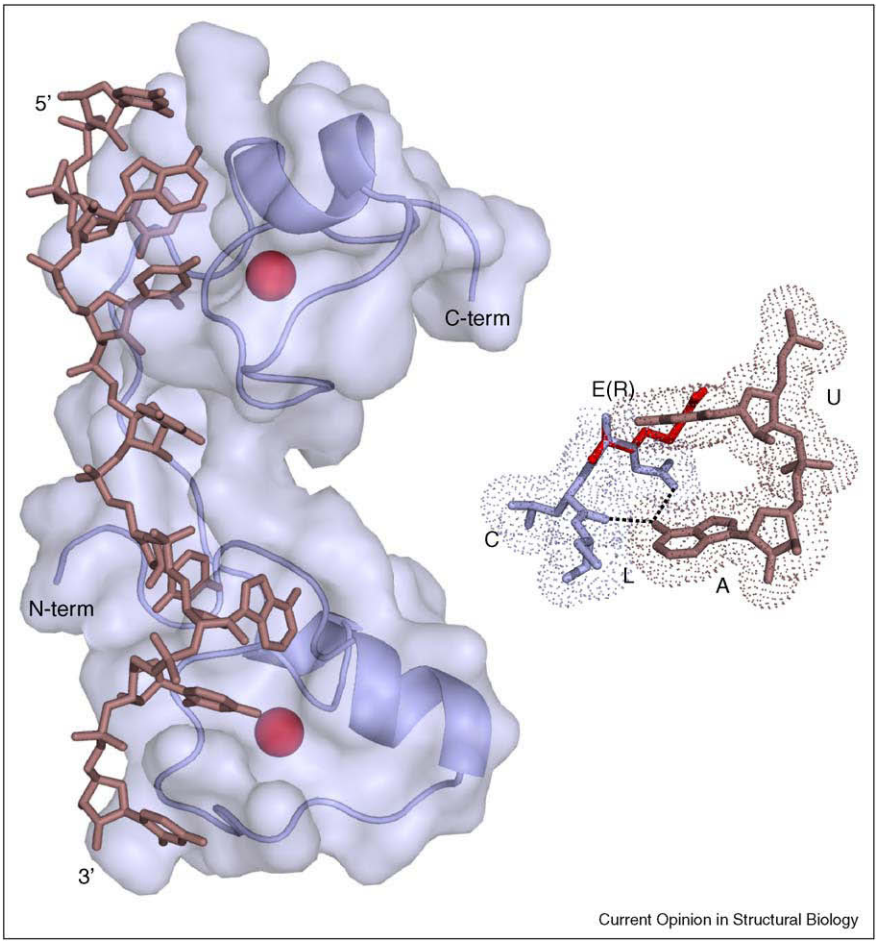
Only one structure of a TZF protein has been determined to date. Wright and co-workers determined the solution structure of the Tis11D bound to the RNA sequence 5′-UUAUUUAUU-3′ (Figure 3) [33]. Tis11D is a mammalian paralog of TTP that regulates mRNA stability in response to growth factors [3]. It binds to RNA with identical specificity to TTP. The structure reveals that each $CX_8CX_5CX_3H$ finger motif independently recognizes the four nucleotide sequence UAUU. A conserved motif with the sequence (R/K)YKTEL lies upstream of the first cysteine of each finger. This region makes numerous contacts with the RNA. These are primarily composed of hydrogen bonds between the protein backbone and the Watson Crick edges of the bases, and van der Waals interactions that specify the shape of the base at each position. In addition, the side chains of two conserved aromatic amino acids form stacking interactions between adjacent RNA bases at two positions within each finger. These amino acids are essential for high affinity binding, and may contribute to specificity through differential stacking propensity. This structure has thus far provided our only glimpse into RNA recognition by this class of RNA-binding proteins, and as such serves as the primary frame of reference for the interpretation of experiments for related factors.

Biochemical insights into nematode TZF binding specificity

In most cases, the RNA-binding activity of nematode TZF proteins has not been investigated in detail. The two exceptions are MEX-5 and POS-1, which bind to RNA but with different specificity compared to TTP, Tis11D, and each other [34,35]. MEX-5 binds with high affinity but relaxed specificity to any uridine rich sequence, including polyuridine. This contrasts with TTP which binds >80-fold more tightly to AREs than polyuridine. POS-1 binds with high affinity to a consensus sequence termed the POS-1 recognition element (PRE: 5′-UA($U_{2–3}$)RD($N_{1–3}$)G-3′, where R is any purine, D is A, G, or U, and N is any base). Compared to TTP binding sequence, the PRE is more degenerate and specifies three purines instead of two.

In Tis11D, three contiguous amino acids in each finger form an adenosine recognition pocket: glutamate, leucine, and the first cysteine of the CCCH motif (Figure 3) [33]. The glutamate side chain accepts a hydrogen bond from the exocyclic amine of the adenosine. The leucine and the cysteine are conserved in both MEX-5 and POS-1, but the glutamate is not. In MEX-5, the analogous amino acids are arginine in the first finger and a lysine in the second. Mutating both to glutamate confers TTP-like specificity to MEX-5, suggesting they are critical specificity determinants [35]. In POS-1, an alanine and a valine occupy the analogous positions. It is not clear how these amino acids contribute to the differences in POS-1 specificity, or how this protein specifies three purines compared to two. Structural data are needed to resolve this problem. One nematode TZF protein, CCCH-1, has two glutamate residues in the analogous position similar to TTP. The rest have basic residues, small hydrophobic residues, or some combination thereof. It is expected that CCCH-1 will bind to RNA with TTP-like specificity, and

Figure 3



NMR structure of human Tis11D bound to RNA [33]. Each zinc finger domain independently recognizes the sequence UAUU through a combination of base specific hydrogen bonding interactions and stacking interactions driven by aromatic side chains. The inset shows recognition of adenosine in the N terminal finger. Three amino acids (blue), glutamate, leucine, and cysteine, come together to form an adenosine (violet) recognition pocket. The exocyclic amine hydrogen bonds with the glutamate side chain and the backbone carbonyl of the leucine. In MEX 5, the glutamate is replaced with an arginine (red), which is proposed to flip away from the adenosine and form nonspecific interactions with the backbone of adjacent nucleotides [35].

that the others will bind to RNA with hybrid specificity, but this has not been experimentally demonstrated.

The worm Argonaute proteins

Biological functions of nematode Argonaute proteins

Argonautes are the primary effectors of small RNA silencing pathways, which have been the subject of intense investigation [36]. Twenty-seven Argonaute genes are annotated in the *C. elegans* genome. These fall into three paralogous groupings: first, similar to *Arabidopsis thaliana* AGO1; second, similar to *Drosophila melanogaster* PIWI; third, WAGOs [37]. Some Argonaute proteins catalyze the cleavage of target RNAs recognized by small RNA guides. Others function in the regulation of mRNA translation in the microRNA pathway. And some are implicated in transcriptional gene silencing through modification of chromatin state.

The large number and apparent diversity of nematode Argonautes suggests a high degree of specialization or functional overlap. Canonical RNA interference triggered by exogenous double strand RNA (dsRNA) is mediated by RDE-1 [7]. Endogenously encoded siRNAs, which are proposed to control cellular homeostasis, are loaded into ERGO-1 [8]. SAGO-1 and SAGO-2 (members of the WAGO clade) are proposed to function in a systemic amplification mechanism, absent in flies and mammals, that leads to silencing of sequences upstream from loci targeted by primary small RNAs [8,38 40]. ALG-1 and ALG-2 load microRNAs required for the temporal regulation of pattern formation during development [41]. The PIWI clade member PRG-1 loads 21U-RNAs and is required for germline maintenance and fertility [42]. More recently, WAGO-1 was shown to repress specific genes, transposons, pseudogenes, and

cryptic loci in conjunction with a class of guide sequences termed the 22G-RNAs [43], and CSR-1 was shown to target euchromatic domains of the genome to enforce appropriate assembly of kinetochores and to facilitate segregation of the holocentric chromosomes [44••,45]. Several additional Argonautes do not have clearly delineated function.

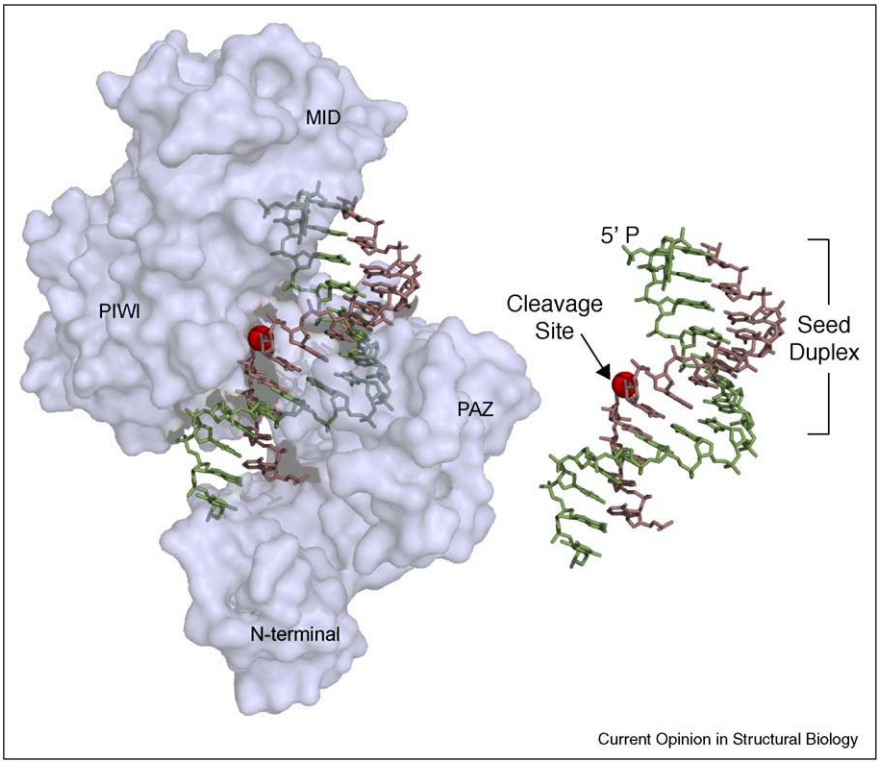
Argonaute structure

Eukaryotic Argonautes consist of four domains: the N-terminal, PAZ, MID, and PIWI domains [37]. A series of recent crystal structures of Argonaute-like RNA endonucleases from hyperthermophilic bacteria begin to define the basis for guide and target recognition as well as the mechanism of site-specific cleavage [46••,47]. As predicted, the structure shows that the 5′ monophosphate of the single stranded guide is lodged between the interface of the MID and PIWI domain while the 3′ end is held by the PAZ domain (Figure 4) [46••]. Target association is proposed to occur in two steps. First, the seed region of the guide pairs to the target [48]. Pairing is limited to the seed because of the 'doubly anchored' conformation of the guide. Once seed pairing is achieved, the helix leading to dissociation of the 3′ end of the guide from the PAZ domain. This remodeling step positions the target adjacent to the metal-coordinated catalytic residues in the PIWI domain required for target cleavage. A minimum of 15 contiguous base pairs is necessary to mediate the remodeling event, which explains why most microRNAs which typically recognize their targets through incomplete pairing do not guide cleavage of their mRNA targets [48].

Implications for nematode-specific Argonautes

The features of the bacterial Argonaute structures shed light on the function of WAGOs. WAGOs involved in systemic silencing lack the residues that coordinate the divalent metal ion required for target cleavage. This suggests that they do no regulate their RNA targets by guide-directed cleavage [8]. However, they do load guide sequences that are at least in principle capable of completely base pairing with their target RNAs. It is not known if these proteins direct a two-step recognition process to bind to their RNA targets. If not, then complementarity between the 3′-end of the guide and the target RNA may be dispensable for function, increasing the number of potential targets as well as increasing the opportunity for off target effects. If so,

Figure 4



Crystal structure of a *Thermus thermophilus* DNA dependent RNA endonuclease related to eukaryotic Argonaute proteins [46••]. The structure shows the protein bound to a DNA guide (green) and a target RNA (violet). The 5′ monophosphate end of the guide is anchored in a cleft between the MID and the PAZ domains. Nucleotides 2 8, which comprise the seed, are exposed on the surface of the protein complex and as such are positioned for substrate recognition. Pairing of the 3′ end of the guide with the target RNA aligns the scissile phosphate (red sphere) with the catalytic residues in the protein. The inset shows the guide target interaction in the absence of protein for clarity.

then cleavage-independent RNA silencing must be possible in a conformation that includes significant pairing between the guide and the target.

In contrast, CSR-1 is capable of guide-directed RNA cleavage in worm extracts, implying but not proving that cleavage activity plays a role in its biological function [44••]. The guide RNAs recognized by CSR-1 harbor a 5′-triphosphate moiety. It is not clear how CSR-1 preferentially accommodates a triphosphate moiety in place of the canonical 5′-monophosphate. The 5′-triphosphate group enhances cleavage activity in extracts relative to the identical sequence with a 5′-monophosphate, suggesting the 5′-triphosphate moeity functions in some aspect of target cleavage [49]. More work is needed to delineate the basis for 5′-end discrimination by CSR-1.

Concluding remarks

The function of RNA-binding proteins is dictated by their structure. For RNA-binding protein families where a common domain has evolved new binding specificity, it is important to understand how structural changes define the basis for novel function. While genetics and biochemical experiments can identify the critical sequence elements, they cannot in most cases address how these elements contribute to novel function in a mechanistic sense. Thus, it is important to continue to put effort into structural studies beyond the first structure in an RNA-binding protein family. Structural experiments can provide key insights needed to understand biological function.

Acknowledgements

The authors would like to thank Brian Farley, John Pagano, Pedro Batista, and members of the Zamore and Mello labs for helpful comments concerning this manuscript, and Dr. Philip Zamore for allowing W.L.M. the opportunity to contribute to this article. S.P.R. is supported by NIH R01 GM081422 and a Basil O'Connor Starter Scholar Award from the March of Dimes. The authors apologize for references omitted due to space limitations.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- •• of outstanding interest

1. Lehmann R, Nusslein Volhard C: **The maternal gene nanos has a central role in posterior pattern formation of the Drosophila embryo**. *Development* 1991, **112**:679 691.

2. Zhang B, Gallegos M, Puoti A, Durkin E, Fields S, Kimble J, Wickens MP: **A conserved RNA binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line**. *Nature* 1997, **390**:477 484.

3. Varnum BC, Ma QF, Chi TH, Fletcher B, Herschman HR: **The TIS11 primary response gene is a member of a gene family that encodes proteins with a highly conserved sequence containing an unusual Cys His repeat**. *Mol Cell Biol* 1991, **11**:1754 1758.

4. Tabara H, Hill RJ, Mello CC, Priess JR, Kohara Y: **Pos 1 encodes a cytoplasmic zinc finger protein essential for germline specification in *C. elegans***. *Development* 1999, **126**:1 11.

5. Schubert CM, Lin R, de Vries CJ, Plasterk RH, Priess JR: **MEX 5 and MEX 6 function to establish soma/germline asymmetry in early *C. elegans* embryos**. *Mol Cell* 2000, **5**:671 682.

6. Detwiler MR, Reuben M, Li X, Rogers E, Lin R: **Two zinc finger proteins, OMA 1 and OMA 2, are redundantly required for oocyte maturation in *C. elegans***. *Dev Cell* 2001, **1**:187 199.

7. Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC: **The rde 1 gene, RNA interference, and transposon silencing in *C. elegans***. *Cell* 1999, **99**:123 132.

8. Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua Tor L, Mitani S, Simard MJ, Mello CC: **Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi**. *Cell* 2006, **127**:747 757.

9. Crittenden SL, Bernstein DS, Bachorik JL, Thompson BE, Gallegos M, Petcherski AG, Moulder G, Barstead R, Wickens M, Kimble J: **A conserved RNA binding protein controls germline stem cells in *Caenorhabditis elegans***. *Nature* 2002, **417**:660 663.

10. Bernstein D, Hook B, Hajarnavis A, Opperman L, Wickens M: **Binding specificity and mRNA targets of a *C. elegans* PUF protein, FBF 1**. *RNA* 2005, **11**:447 458.

11. Ahringer J, Kimble J: **Control of the sperm oocyte switch in *Caenorhabditis elegans* hermaphrodites by the fem 3 3′ untranslated region**. *Nature* 1991, **349**:346 438.

12. Francis R, Barton MK, Kimble J, Schedl T: **gld 1, a tumor suppressor gene required for oocyte development in *Caenorhabditis elegans***. *Genetics* 1995, **139**:579 606.

13. Francis R, Maine E, Schedl T: **Analysis of the multiple roles of gld 1 in germline development: interactions with the sex determination cascade and the glp 1 signaling pathway**. *Genetics* 1995, **139**:607 630.

14. Lublin AL, Evans TC: **The RNA binding proteins PUF 5, PUF 6, and PUF 7 reveal multiple systems for maternal mRNA regulation during *C. elegans* oogenesis**. *Dev Biol* 2007, **303**:635 649.

15. Nolde MJ, Saka N, Reinert KL, Slack FJ: **The *Caenorhabditis elegans* pumilio homolog, puf 9, is required for the 3′UTR mediated repression of the let 7 microRNA target gene, hbl 1**. *Dev Biol* 2007, **305**:551 563.

16. Fraser AG, Kamath RS, Zipperlen P, Martinez Campos M, Sohrmann M, Ahringer J: **Functional genomic analysis of *C. elegans* chromosome I by systematic RNA interference**. *Nature* 2000, **408**:325 330.

17. Sonnichsen B, Koski LB, Walsh A, Marschall P, Neumann B, Brehm M, Alleaume AM, Artelt J, Bettencourt P, Cassin E *et al.*: **Full genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans***. *Nature* 2005, **434**:462 469.

18. •• Wang Y, Opperman L, Wickens M, Hall TM: **Structural basis for specific recognition of multiple mRNA targets by a PUF regulatory protein**. *Proc Natl Acad Sci USA* 2009, **106**:20186 20191.
This article demonstrates the structural basis for differential binding site consensus sequence length between PUF proteins.

19. •• Koh YY, Opperman L, Stumpf C, Mandan A, Keles S, Wickens M: **A single *C. elegans* PUF protein binds RNA in multiple modes**. *RNA* 2009, **15**:1090 1099.
The diversity of sequence recognized by several nematode PUF domain proteins is explored in this research paper.

20. Stumpf CR, Kimble J, Wickens M: **A *Caenorhabditis elegans* PUF protein family with distinct RNA binding specificity**. *RNA* 2008, **14**:1550 1557.

21. Opperman L, Hook B, DeFino M, Bernstein DS, Wickens M: **A single spacer nucleotide determines the specificities of two mRNA regulatory proteins**. *Nat Struct Mol Biol* 2005, **12**:945 951.

22. Wang X, McLachlan J, Zamore PD, Hall TM: **Modular recognition of RNA by a human pumilio homology domain**. *Cell* 2002, **110**:501 512.

23. Wang X, Zamore PD, Hall TM: **Crystal structure of a Pumilio homology domain**. *Mol Cell* 2001, **7**:855 865.

24. Edwards TA, Pyle SE, Wharton RP, Aggarwal AK: **Structure of Pumilio reveals similarity between RNA and peptide binding motifs**. *Cell* 2001, **105**:281 329.

25. Lu G, Dolgner SJ, Hall TM: **Understanding and engineering RNA sequence specificity of PUF proteins**. *Curr Opin Struct Biol* 2009, **19**:110 115.

26. Blackshear PJ: **Tristetraprolin and other CCCH tandem zinc finger proteins in the regulation of mRNA turnover**. *Biochem Soc Trans* 2002, **30**:945 952.

27. Lai WS, Carballo E, Strum JR, Kennington EA, Phillips RS, Blackshear PJ: **Evidence that tristetraprolin binds to AU rich elements and promotes the deadenylation and destabilization of tumor necrosis factor alpha mRNA**. *Mol Cell Biol* 1999, **19**:4311 4323.

28. Brewer BY, Malicka J, Blackshear PJ, Wilson GM: **RNA sequence elements required for high affinity binding by the zinc finger domain of tristetraprolin: conformational changes coupled to the bipartite nature of Au rich MRNA destabilizing motifs**. *J Biol Chem* 2004, **279**:27870 27877.

29. Mello CC, Draper BW, Krause M, Weintraub H, Priess JR: **The pie 1 and mex 1 genes and maternal control of blastomere identity in early *C. elegans* embryos**. *Cell* 1992, **70**:163 176.

30. Cuenca AA, Schetter A, Aceto D, Kemphues K, Seydoux G: **Polarization of the *C. elegans* zygote proceeds via distinct establishment and maintenance phases**. *Development* 2003, **130**:1255 1265.

31. Reese KJ, Dunn MA, Waddle JA, Seydoux G: **Asymmetric segregation of PIE 1 in *C. elegans* is mediated by two complementary mechanisms that act through separate PIE 1 protein domains**. *Mol Cell* 2000, **6**:445 455.

32. Pinkston Gosse J, Kenyon C: **DAF 16/FOXO targets genes that regulate tumor growth in *Caenorhabditis elegans***. *Nat Genet* 2007, **39**:1403 1409.

33. Hudson BP, Martinez Yamout MA, Dyson HJ, Wright PE: **Recognition of the mRNA AU rich element by the zinc finger domain of TIS11d**. *Nat Struct Mol Biol* 2004, **11**:257 264.

34. Farley B, Pagano JM, Ryder S: **RNA target specificity of the embryonic cell fate determinant POS 1**. *RNA* 2008, **14**:2685 2697.

35. Pagano JM, Farley B, McCoig LM, Ryder S: **Molecular basis of RNA recognition by the embryonic polarity determinant MEX 5**. *J Biol Chem* 2007, **282**:8883 8894.

36. Ghildiyal M, Zamore PD: **Small silencing RNAs: an expanding universe**. *Nat Rev Genet* 2009, **10**:94 108.

37. Hutvagner G, Simard MJ: **Argonaute proteins: key players in RNA silencing**. *Nat Rev Mol Cell Biol* 2008, **9**:22 32.

38. Sijen T, Fleenor J, Simmer F, Thijssen KL, Parrish S, Timmons L, Plasterk RH, Fire A: **On the role of RNA amplification in dsRNA triggered gene silencing**. *Cell* 2001, **107**:465 476.

39. Sijen T, Steiner FA, Thijssen KL, Plasterk RH: **Secondary siRNAs result from unprimed RNA synthesis and form a distinct class**. *Science* 2007, **315**:244 247.

40. Pak J, Fire A: **Distinct populations of primary and secondary effectors during RNAi in *C. elegans***. *Science* 2007, **315**:241 244.

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

42. Batista PJ, Ruby JG, Claycomb JM, Chiang R, Fahlgren N, Kasschau KD, Chaves DA, Gu W, Vasale JJ, Duan S *et al.*: **PRG 1 and 21U RNAs interact to form the piRNA complex required for fertility in *C. elegans***. *Mol Cell* 2008, **31**:67 78.

43. Gu W, Shirayama M, Conte DJ, Vasale J, Batista PJ, Claycomb JM, Moresco JJ, Youngman EM, Keys J, Stoltz MJ *et al.*: **Distinct argonaute mediated 22G RNA pathways direct genome surveillance in the *C. elegans* germline**. *Mol Cell* 2009, **36**:231 244.

44. •• Claycomb JM, Batista PJ, Pang KM, Gu W, Vasale JJ, van Wolfswinkel JC, Chaves DA, Shirayama M, Mitani S, Ketting RF *et al.*: **The Argonaute CSR 1 and its 22G RNA cofactors are required for holocentric chromosome segregation**. *Cell* 2009, **139**:123 134.
The biological role of CSR 1, an interesting WAGO with a function in chromosome segregation, is described in this article.

45. van Wolfswinkel JC, Claycomb JM, Batista PJ, Mello CC, Berezikov E, Ketting RF: **CDE 1 affects chromosome segregation through uridylation of CSR 1 bound siRNAs**. *Cell* 2009, **139**:135 148.

46. •• Wang Y, Juranek S, Li H, Sheng G, Wardle GS, Tuschl T, Patel DJ: **Nucleation, propagation and cleavage of target RNAs in ago silencing complexes**. *Nature* 2009, **461**:754 761.
The structural basis for Argonaute guide target recognition and evidence for the two state binding model are presented in this paper.

47. Song JJ, Smith SK, Hannon GJ, Joshua Tor L: **Crystal structure of Argonaute and its implications for RISC slicer activity**. *Science* 2004, **305**:1434 1437.

48. Tomari Y, Zamore PD: **Perspective: machines for RNAi**. *Genes Dev* 2005, **19**:517 529.

49. Aoki K, Moriguchi H, Yoshioka T, Okawa K, Tabara H: **In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans***. *EMBO J* 2007, **26**:5007 5019.

50. Farley B, Ryder S: **Regulation of maternal mRNAs in early development**. *Crit Rev Biochem Mol Biol* 2008, **43**:135 162.