RNA Silencing Pathways in *Schizosaccharomyces pombe* and *Drosophila melanogaster*: A Dissertation

Alla A. Sigova
*University of Massachusetts Medical School*

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

Part of the Amino Acids, Peptides, and Proteins Commons, Animal Experimentation and Research Commons, Fungi Commons, and the Nucleic Acids, Nucleotides, and Nucleosides Commons

Repository Citation
https://escholarship.umassmed.edu/gsbs_diss/225

This material is brought to you by eScholarship@UMMS. It has been accepted for inclusion in GSBS Dissertations and Theses by an authorized administrator of eScholarship@UMMS. For more information, please contact Lisa.Palmer@umassmed.edu.
RNA SILENCING PATHWAYS IN

_Schizosaccharomyces pombe_ and _Drosophila melanogaster_

A Dissertation Presented

By

ALLA A. SIGOVA

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

November 3, 2006

Biochemistry and Molecular Pharmacology
COPYRIGHT INFORMATION

The chapters of this dissertation have appeared in whole or part in separate publications:


*These authors contributed equally
RNA SILENCING PATHWAYS IN
_Schizosaccharomyces pombe_ and _Drosophila melanogaster_

A Dissertation Presented

By

Alla A. Sigova

Approved as to style and content by

__________________________
Phillip Zamore, PhD., Thesis Advisor

__________________________
Celia Schiffer, PhD., Member of Committee

__________________________
Nicholas Rhind, PhD., Member of Committee

__________________________
David Lambright, PhD., Member of Committee

__________________________
Traci Hall, PhD., Member of Committee

__________________________
Reid Gilmore, PhD., Chair of Committee

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

Department of Biochemistry and Molecular Pharmacology

November 3rd, 2006
Acknowledgments

First of all, I would like to thank Dr. Phillip Zamore, my thesis advisor, for the opportunity to do science in his lab, for creating an excellent environment for efficient work and study, for his inspirational enthusiasm about science and life in general, for his strong belief in my success as a scientist even at difficult times and, most of all, for his teaching me the importance of doing good science by setting an example of adherence to the highest standards of professionalism and integrity.

I am really grateful to members of Zamore lab, current and former, for stimulating discussions, generous sharing of reagents and expertise, support and fruitful collaborations. I will always remember Juanita, Ben, Gyorgy, Guiliang, Dianne, and Tingting, whose company I enjoyed from the very day I joined the lab. I was lucky to have the best possible benchmates (Chris, Brad, and Chengjian). Despite the hard work, we could always find time for a good laugh and curious conversations.

I am thankful to the great people who worked with me on the Drosophila project (Vasia, Chengjian and Herve) for the most successful collaboration and for the possibility to share my excitement about science with you. Vasia Vagin performed experiments for Chapter III Fig. 1A, Fig. 5, and Fig. 6A and 6B. I performed experiments for Fig. 3. Chengjian Li’s results are in Fig. 6C and 6D. Fig. 2, Fig. 3E, and Fig. 4 are the result of a joint effort of Vasia Vagin and myself. Herve Seitz designed probes for microarrays and helped to analyze microarray data.

I would also like to give thanks to Yuki, Klaus, Jennifer, Megha, Wee, Gwen and Alicia. I was really fortunate to work with you.
I am really grateful to Nick Rhind and the Rhind lab for helping me to discover the great world of fission yeast. I would also like to acknowledge my committee members (Reid Gilmore, Celia Schiffer, Nick Rhind, David Lambright, and Traci Hall) for their suggestions and faithful service even during the busiest times. I am really indebted to Tiffanie Covello, our secretary, for her incredible help with organizing both big and small events of the lab life. Things always run smoothly when Tiffanie is in charge.

I would like to thank all the people in the Biochemistry Department for making it such a special place to be a graduate student.

I am the luckiest person… I have the most wonderful family. Thank you for all your love and support, for great sacrifices you made for me, for bearing with me through successes and failures, moments of happiness as well as sadness. My dear parents and sister, my husband and son, I am really blessed that you are the way you are. To my lifelong friends, for always being there for me, in both good and bad times, and for giving me the opportunity to do the same for you.
RNA silencing is an evolutionary conserved sequence-specific mechanism of regulation of gene expression. RNA interference (RNAi), a type of RNA silencing in animals, is based on recognition and endonucleolytic cleavage of target mRNA complimentary in sequence to 21-nucleotide (nt) small RNA guides, called small interfering RNAs (siRNAs). Another class of 21-nt small RNAs, called micro RNAs (miRNAs), is endogenously encoded in eukaryotic genomes. Both production of siRNAs from long double-stranded RNA (dsRNA) and biogenesis of miRNAs from hairpin structures are governed by the ribonuclease III enzyme Dicer. Although produced as duplex molecules, siRNAs and miRNAs are assembled into effector complex, called the RNA-induced silencing complex (RISC), as single-strands. A member of the Argonaute family of small RNA-binding proteins lies at the core of all known RNA silencing effector complexes. Plants and animals contain multiple Argonaute paralogs. In addition to endonucleolytic cleavage, Argonaute proteins can direct translational repression/destabilization of mRNA or transcriptional silencing of DNA sequences by the siRNA-directed production of silent heterochromatin.

The *Schizosaccharomyces pombe* genome encodes only one of each of the three major classes of proteins implicated in RNA silencing: Dicer (Dcr1), RNA-dependent RNA polymerase (RdRP; Rdp1), and Argonaute (Ago1). These three proteins are required for silencing at centromeres and for the initiation of transcriptionally silent heterochromatin at the mating-type locus. That only one Dicer, RdRP and Argonaute is expressed in *S. pombe* might reflect the extreme specialization of RNA silencing pathways regulating targets only at the transcriptional level in this organism. We decided
to test if classical RNAi can be induced in *S. pombe*. We introduced a dsRNA hairpin corresponding to a GFP transgene. GFP silencing triggered by dsRNA reflected a change in the steady-state concentration of GFP mRNA, but not in the rate of GFP transcription. RNAi in *S. pombe* required *dcr1*, *rdp1*, and *ago1*, but did not require *chp1*, *tas3*, or *swi6*, genes required for transcriptional silencing. We concluded that the RNAi machinery in *S. pombe* could direct both transcriptional and posttranscriptional silencing using a single Dicer, RdRP, and Argonaute protein. Our findings suggest that, in spite of specialization in distinct siRNA-directed silencing pathways, these three proteins fulfill a common biochemical function.

In *Drosophila*, miRNA and RNAi pathways are both genetically and biochemically distinct. Dicer-2 (Dcr-2) generates siRNAs, whereas the Dicer-1 (Dcr-1)/Loquacious complex produces miRNAs. Argonaute proteins can be divided by sequence similarity into two classes: in flies, the Ago subfamily includes Argonaute1 (Ago1) and Argonaute2 (Ago2), whereas the Piwi subfamily includes Aubergine, Piwi and Argonaute 3. siRNAs and miRNAs direct posttranscriptional gene silencing through effector complexes containing Ago1 or Ago2. The third class of small RNAs, called repeat-associated small interfering RNAs (rasiRNAs), is produced endogenously in the *Drosophila* germ line. rasiRNAs mediate silencing of endogenous selfish genetic elements such as retrotransposons and repetitive sequences to ensure genomic stability.

We examined the genetic requirements for biogenesis of rasiRNAs in both male and female germ line of *Drosophila* and silencing of 8 different selfish elements, including tree LTR retrotransposons, two non-LTR retrotransposons, and three repetitive sequences. We find that biogenesis of rasiRNAs is different from that of miRNAs and
siRNAs. rasiRNA production appears not to require Dicer-1 or Dicer-2. rasiRNAs lack the 2',3' hydroxy termini characteristic of animal siRNA and miRNA. While siRNAs derive from both the sense and antisense strands of their dsRNA precursors, rasiRNAs accumulate in antisense polarity to their corresponding target mRNAs. Unlike siRNAs and miRNAs, rasiRNAs function through the Piwi, rather than the Ago, Argonaute protein subfamily. We find that rasiRNAs silence their target RNAs posttranscriptionally: mutations that abrogate rasiRNA function dramatically increase the steady-state mRNA level of rasiRNA targets, but do not alter their rate of transcription, measured by nuclear run-on assay.

Our data suggest that rasiRNAs protect the fly germ line through a silencing mechanism distinct from both the miRNA and RNAi pathways.
TABLE OF CONTENTS

Title i
Copyright information ii
Signature page iii
Acknowledgments iv
Abstract vi
Table of contents ix
List of figures xiii
List of tables xvii

CHAPTER I: INTRODUCTION

Discovery of RNA interference 1
Small interfering RNAs 4
RISC assembly 14
miRNAs: a historical perspective 16
miRNA-directed mRNA destabilization and translational repression 18
miRNA biogenesis 20
Pre-miRNA processing 21
Endogenous plant small RNAs 25
siRNA-mediated heterochromatin formation 26
Small RNA clusters 29
RNase III enzymes in RNA silencing 30
CHAPTER II: A SINGLE ARGONAUTE PROTEIN MEDIATES BOTH TRANSCRIPTIONAL AND POSTTRANSCRIPTIONAL SILENCING IN SCHIZOSACCHAROMYCES POMBE

Summary

Introduction

Results

Discussion

Materials and methods

S. pombe strain construction

Construction of the silencing trigger

RNA analysis

Nuclear run-on analysis

FACS analysis

Chromatin immunoprecipitation (ChIP)

RNase protection assay

RNA isolation and detection by Northern blot

Acknowledgments

CHAPTER III: A DISTINCT SMALL RNA PATHWAY SILENCES SELFISH GENETIC ELEMENTS IN THE GERM LINE

Summary
CHAPTER IV:  MEASURING THE RATES OF TRANSCRIPTIONAL ELONGATION IN THE FEMALE DROSOPHILA MELANOGASTER GERM LINE BY NUCLEAR RUN-ON

Summary

Introduction
Results and discussion  143

Nuclear run-on measures relative rates of transcription  143

Measuring the rates of repetitive element transcription  146

Repetitive elements are transcribed by RNA pol II  153

Repetitive elements in D. melanogaster ovaries are silenced posttranscriptionally  153

Conclusions  161

Materials and methods  162

Acknowledgments  165

Appendix to Chapter IV  166

CHAPTER V: GENERAL CONCLUSIONS  171

REFERENCES  179
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure I-1</th>
<th>RNA-dependent RNA polymerase (RdRP) is required for both initial production of dsRNA trigger and amplification of silencing</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure I-2</td>
<td>RNAi and miRNA pathways in Drosophila</td>
<td>10</td>
</tr>
<tr>
<td>Figure I-3</td>
<td>RNAi and miRNA pathways in mammalian cells</td>
<td>24</td>
</tr>
<tr>
<td>Figure I-4</td>
<td>piRNA pathways in mammalian cells</td>
<td>31</td>
</tr>
<tr>
<td>Figure II-1</td>
<td>Experimental strategy</td>
<td>51</td>
</tr>
<tr>
<td>Figure II-2</td>
<td>The GFP hairpin triggers adh1:gfp silencing</td>
<td>54</td>
</tr>
<tr>
<td>Figure II-3</td>
<td>GFP silencing reflects a reduction in the steady-state level of adh1:gfp mRNA</td>
<td>57</td>
</tr>
<tr>
<td>Figure II-4</td>
<td>GFP silencing triggered by the GFP hairpin is posttranscriptional</td>
<td>60</td>
</tr>
<tr>
<td>Figure II-5</td>
<td>GFP silencing depends on the RNAi machinery but not the transcriptional silencing proteins Chp1, Tas3, or Swi6</td>
<td>63</td>
</tr>
<tr>
<td>Supplementary Figure II-1</td>
<td>Expression of GFP hairpin measured by RNase protection assay</td>
<td>70</td>
</tr>
<tr>
<td>Supplementary Figure II-2</td>
<td>Nuclear run-on analysis reflects changes in transcriptional rates</td>
<td>72</td>
</tr>
<tr>
<td>Supplementary Figure II-3</td>
<td>The nuclear run-on assay is linear and can readily detect a 25% change in transcriptional rates</td>
<td>74</td>
</tr>
<tr>
<td>Supplementary Figure II-4</td>
<td>Chromatin immunoprecipitation analysis of</td>
<td></td>
</tr>
</tbody>
</table>
histone 3 lysine 9 methylation 76

**Supplementary Figure II-5** Detection of GFP-specific small RNAs by Northern hybridization in wild-type and dcr1Δ strains 78

**Supplementary Figure II-6** Detection of GFP-specific small RNAs by Northern hybridization in wild-type, ago1Δ, dcr1Δ, and rdp1Δ strains 80

**Figure III-1** Sense and antisense siRNAs accumulate during RNAi in vivo 96

**Figure III-2** rasiRNAs derived from the endogenous silencing trigger Su(Ste) 99

**Figure III-3** RNA expression from selfish genetic elements 102

**Figure III-4** roo rasiRNAs in ovaries arise mainly from the antisense strand and require the putative helicase Armi for their accumulation 105

**Figure III-5** Northern analysis of mitotic recombinant dcr-1Q1147X homozygous mutant ovaries and nonrecombinant controls 108

**Figure III-6** roo rasiRNAs are modified at their 3’ terminus and associate with Piwi subfamily Argonaute proteins 111

**Supplementary Figure III-1** The structure of the white exon 3 hairpin trigger RNA 115

**Supplementary Figure III-2** Characterization of Su(Ste) rasiRNAs isolated from testes 117

**Supplementary Figure III-3** High resolution tiling microarray analysis across the roo LTR and Fourier-transform analysis of the combined sense and antisense roo rasiRNA data for wild-type ovaries 119
Supplementary Figure III-4 Northern analysis of mitotic recombinant

dcr-1^{Q1147X} homozygous mutant ovaries and

nonrecombinant controls 121

Supplementary Figure III-5 Northern analysis of enzymatic and chemical

probing of total small RNA from fly testes 123

Supplementary Figure III-6 Western blotting analysis of tagged and endogenous

Piwi and Aub proteins and Northern blotting

analysis of small RNAs immunoprecipitated

with them 125

Supplementary Figure III-7 Schematic representations of the selfish genetic

elements examined in this study 127

Figure IV-1 Schematic of nuclear run-on analysis 144

Figure IV-2 Nuclear run-on analysis can detect small changes in

transcriptional rates 147

Figure IV-3 Position of primers used to prepare T7 RNA polymerase PCR

templates to synthesize the antisense RNA probes used to detect

repetitive elements 149

Figure IV-4 Nuclear run-on likely measures the rate of transcription without

substantial re-initiation by RNA polymerase (pol) II 151

Figure IV-5 Nuclear run-on can detect the rates of repetitive element

transcription 154

Figure IV-6 Repetitive element transcripts are largely RNA pol II products 156

Figure IV-7 Transcriptional rates of the retrotransposon gypsy and the
repetitive sequence mst40 in homozygous mutant armi ovaries

measured by nuclear run-on

Figure IV-A1 Transcriptional rates of the retrotransposons I-element, roo, and HeT-A in homozygous mutant armi ovaries measured by nuclear run-on

Figure IV-A2 Transcriptional rates of the Stellate in homozygous mutant spn-E testes measured by nuclear run-on
LIST OF TABLES

Table II-1. Strains used in the study 83

Table III-1. Probes and primers 134

Table IV-1. Primers used to generate T7 RNA polymerase transcription templates for production of antisense RNA probes 164
CHAPTER I

INTRODUCTION

Discovery of RNA interference

RNA silencing is a homology-dependent repression of gene expression mediated by RNA mainly by means of posttranscriptional destruction of sequences homologous to the RNA, inhibition of their translation, or reduction in transcriptional rates. RNA silencing was discovered by plant scientists in 1990, several years after the first transgenic plants were produced (van der Krol et al., 1990, Napoli et al., 1990). In various types of experiments with transgenic *Petunia hybrida*, introduction of additional copies of genes encoding the flower pigmentation pathway enzymes chalcone synthase or dihydroflavonol-4-reductase led to silencing of endogenous copies of these genes as well as transgenic copies manifested by a partial or complete loss of flower pigmentation in these plants. The loss of pigmentation was caused by the decrease in the steady-state level of mRNA for these enzymes. The silencing was dubbed “co-suppression” to reflect the coordinate nature of the effect on both the transgene and the endogenous gene. As similar phenomena were discovered in other plants, it has become clear that at least two different mechanisms operate in various cases of co-suppression. If homology is between coding regions of interacting genes present in either sense or antisense orientation, the silencing is posttranscriptional and leads to *de novo* methylation of homologous DNA. Posttranscriptional gene silencing (PTGS) is not transmitted to the progeny. Instead, it is reset in every generation and fully reversible when transgenes, “triggers” of silencing, are lost upon outcrossing (Flavell, 1994, Napoli et al., 1990, Depicker and Montagu, 1997,
Kunz et al., 1996). Infecting plants with modified RNA viruses carrying fragments of the transgenes homologous to the sequences in plant genome induces a subtype of PTGS called virus-induced gene silencing (VIGS) (Lindbo et al., 1993, Ruiz et al., 1998, Angell and Baulcombe, 1997). Although infections by viruses are localized, they always induce global silencing of the integrated transgenes in all organs of the plant (Voinnet and Baulcombe, 1997). In grafting procedures, silencing spreads from silenced stocks to non-silenced scions and less efficiently in reverse directions (Palauqui et al., 1997). These observations led to realization that silencing in plants is systemic.

Homology in the promotor region of interacting genes induces their transcriptional inhibition, which correlates with methylation of the promotor sequences. Transcriptional silencing is meiotically inherited independent of the trigger of silencing (Park et al., 1996, Jones et al., 2001).

PTGS in plants is mechanistically related to another posttranscriptional silencing phenomenon discovered in filamentous fungi Neurospora crassa termed 'quelling'. When Neurospora is transformed with different portions of the carotenogenic albino-3 (al-3) and albino-1 (al-1) genes, the presence of the exogenous sequences (which are randomly integrated in ectopic locations) provokes a severe impairment in the expression of the endogenous al-1 or al-3 genes and recovery of transformants showing an albino phenotype. Similar to plants, silencing is caused by reduction in the abundance of transcripts from the duplicated sequences and fully reversible when the exogenous genes are lost (Cogoni et al., 1996, Romano and Macino, 1992).

Studies in plants and fungi paved the way for the central breakthrough made by scientists working on nematode Caenorhabditis elegans. When trying to inhibit the
function of par-1 gene using antisense oligonucleotides, they made a surprising discovery: injection of both antisense and control sense RNA into the gonad of wild-type worms produced phenotypes resembling loss-of-function mutants of par-1 (Guo and Kemphues, 1995). Fire and colleagues in similar experiments with unc-22 gene determined that interference was induced by residual amounts of contaminating double-stranded RNA (dsRNA) (Fire et al., 1998). In fact, injection of either purified sense or antisense transcripts did not cause twitching phenotype of loss-of-function unc-22 mutants. Instead, dsRNA homologous to unc-22 gene was much more potent in triggering highly specific destruction of the endogenous unc-22 mRNA (Montgomery et al., 1998). Thus, in plants and animals, dsRNA induces posttranscriptional silencing of homologous sequences by means of destruction of their mRNA in a process termed RNA interference (RNAi). RNAi has quickly become a valuable tool for studying gene function in an increasing number of organisms. In worms, in addition to injection, dsRNA can be introduced into cells by soaking and feeding worms with bacteria expressing dsRNA. dsRNA can be expressed as a long hairpin or produced by bidirectional transcription. In plants, mRNA of a single-copy overexpressed gene is recognized as “aberrant” and used as a template for both primed and unprimed synthesis of complimentary RNA by RNA-dependent RNA polymerase (RdRP) resulting in dsRNA production (Horiuchi et al., 2001, Schiebel et al., 1993, Makeyev and Bamford, 2002, Motamedi et al., 2004, Tang et al., 2003). In nematodes and planarians, similar to plants, silencing is systemic, e.g. it spreads to silence targeted genes throughout the animal (Newmark et al., 2003, Timmons and Fire, 1998). In C. elegans, silencing signal moves from one cell to another by passive diffusion through transmembrane channels formed by protein Sid-1 (Winston et al., 2002,
Feinberg and Hunter, 2003). dsRNA is the most probable candidate on the role of messenger in this organism because it can be transported from cell-to-cell more efficiently than shorter RNA. Non-cell autonomous nature of RNAi in worms and systemic silencing in plants are manifestations of a major role of RNA silencing in protection of these organisms against viral infections. Plant viruses have evolved to evade RNA silencing by expressing diverse set of viral suppressors of RNAi, which are essential for their virulence (Roth et al., 2004). Understanding the mechanisms of RNAi is being used to engineer crops with improved disease resistance. Although no worm viruses have been found yet, flock house virus and vesicular stomatitis virus can replicate in *C. elegans* and are silenced by RNAi (Lu et al., 2005a, Wilkins et al., 2005, Schott et al., 2005). RNAi is thought to be a major defense mechanism protecting *Drosophila melanogaster* against viral infections (Wang et al., 2006, Zambon et al., 2006). RNAi and related phenomena operate in four out of six kingdoms of living organisms: protista, fungi, plantae, and animalia. In addition to protecting against foreign invaders, RNA silencing plays a major role in regulating gene expression in development and disease. Our current understanding of the biological role and biochemical mechanisms underlying RNA silencing pathways is summarized in this chapter.

**Small interfering RNAs**

In 1999 Hamilton and colleagues carried out analyses of four classes of PTGS in plants: transgene-induced PTGS of an endogenous gene, PTGS induced by transgenes, which did not have homology to the endogenous gene, systemic PTGS of plants expressing a GFP transgene, which was initiated by infiltrating a single leaf with
Agrobacterium tumefaciens containing GFP sequences, and VIGS of plants infected with potato virus X (Hamilton and Baulcombe, 1999). In all four classes of PTGS ~ 25 nucleotides (nt) antisense RNAs complimentary to targeted mRNAs (or the virus) were detected. The authors suggested that small RNAs were both the specificity determinants for RNA silencing and the mobile signal by which silencing spreads from the initial site of silencing (such as the Agrobacterium infiltrated leaf) to the rest of the plant (Palauqui et al., 1997, Voinnet and Baulcombe, 1997, Voinnet et al., 1998). Biochemical experiments conducted in animal cell extracts that could recapitulate RNAi demonstrated that the small RNAs—now called small interfering RNAs (siRNAs)—derive directly from the dsRNA that triggered silencing and function as guides for protein complexes that degrade the targeted mRNA. In solanaceous plants, siRNAs produced at the initial site are believed to move over long distances through the vasculature system, and then spread from cell to cell through the plasmodesmata of recipient tissues (Palauqui et al., 1997, Voinnet and Baulcombe, 1997, Voinnet et al., 1998). In other plants, extensive movement results from reiterated short-distance silencing events. For example, in Arabidopsis thaliana and Nicotiana benthamiana, siRNAs derived from original dsRNA (“primary” siRNAs) spread to 10-15 neighboring cells in target-independent manner by trafficking through the plasmodesmal channels connecting plant cells (Himber et al., 2003, Schwach et al., 2005, Dunoyer et al., 2005). If mRNA complimentary to primary siRNAs is present, it is converted to new dsRNA by one of the plant RdRPs called RdR6 (previously known as SDE1). “Secondary” siRNAs, which are produced from new dsRNA, spread farther and so on. Thus, in addition to producing the dsRNA trigger of silencing, RdRPs are required for amplification of the silencing signal (Fig. 1).
RNA-dependent RNA polymerase (RdRP) is required for both the initial production of dsRNA trigger and amplification of silencing. An “aberrant” single-stranded RNA is a substrate for RdRP. dsRNA is processed by Dicer into primary siRNAs. RdRP can also use cleaved target mRNA as a template for synthesis of dsRNA and, consequently, secondary siRNA production.

**Figure 1.** RNA-dependent RNA polymerase (RdRP) is required for both the initial production of dsRNA trigger and amplification of silencing. An “aberrant” single-stranded RNA is a substrate for RdRP. dsRNA is processed by Dicer into primary siRNAs. RdRP can also use cleaved target mRNA as a template for synthesis of dsRNA and, consequently, secondary siRNA production.
Cell-free systems recapitulating in vitro the conversion of dsRNA into siRNAs, the assembly of siRNAs into functional protein complexes, and the destruction of target RNAs were developed first in Drosophila, either from syncitial blastoderm embryos or cultured Drosophila S2 cells and later from cultured human HeLa cells (Tuschl et al., 1999, Hammond et al., 2000, Billy et al., 2001). During the RNAi reaction, both strands of the dsRNA are processed to RNA segments ~21-23 nt (Zamore et al., 2000, Elbashir et al., 2001c) by ribonuclease (RNase) III-like enzyme Dicer and consequently bear the hallmarks features of RNase III cleavage products: they are double-stranded, have 2-nt long overhangs at their 3’ ends, 5’ monophosphate required for entry of an siRNA into the RNAi pathway, and 2’ and 3’ hydroxyls at the 3’ end (Bernstein et al., 2001, Elbashir et al., 2001c, Nykanen et al., 2001, Schwarz et al., 2002). 5’ phosphates are actively maintained in vitro by an unidentified kinase in an ATP-dependent fashion. Processing of dsRNA to siRNAs by Dicer in vitro does not require the presence of targeted mRNA or its translation (Zamore et al., 2000). After their production from long dsRNA by Dicer, siRNAs are incorporated into a functional protein-RNA complex, the RNA-induced silencing complex (RISC), through which they direct the endonucleolytic cleavage of the target 10 nt upstream of the 5’-most nucleotide of the siRNA (Elbashir et al., 2001c, Tuschl et al., 1999, Zamore et al., 2000, Nykanen et al., 2001, Schwarz et al., 2002, Elbashir et al., 2001a, Martinez and Tuschl, 2004). For example, introduction of 21-nt siRNA duplexes in Drosophila embryo lysates or cultured mammalian cells leads to sequence-specific mRNA degradation (Elbashir et al., 2001a, Elbashir et al., 2001c, Elbashir et al., 2001b, Schwarz et al., 2002). Furthermore, siRNAs inherited by the progeny of silenced worms can elicit silencing of the targeted genes (Grishok et al.,
In order to bind target mRNA, two strands of an siRNA duplex have to be separated. In vitro in *Drosophila* embryo lysate, this process requires ATP (Nykanen et al., 2001, Tomari et al., 2004a). Single-stranded siRNA can bypass the requirement for siRNA strand separation in vitro in *Drosophila* embryo lysate or HeLa cell S100 extracts and in vivo in cultured human cells (Schwarz et al., 2002, Hammond et al., 2000, Martinez et al., 2002). Purified human RISC contains only single-stranded siRNAs, suggesting that of two strands of an siRNA duplex only one strand is assembled into RISC and the complimentary strand is degraded (Martinez et al., 2002). Which of the two siRNA strands assembles into RISC is not random. The strand that forms RISC is termed the guide and the strand disfavored for RISC assembly is called the passenger. If the guide strand happens to be complimentary to the target, it will cause mRNA destruction. If the complimentary strand happens to be the passenger strand, no cleavage will be observed. This explains why some synthetic siRNAs duplexes are not active; they load the wrong strand into RISC. Statistical and biochemical studies revealed that the decision which strand of an siRNA duplex gets incorporated into RISC is made based on relative thermodynamic stability of the ends of the 19-nt double-stranded region of the siRNA (Khvorova et al., 2003, Schwarz et al., 2003). Thermodynamic stability of an siRNA can be calculated using “nearest neighbor” analysis, which can predict relative stability at different positions of an RNA helix by measuring the free energy of base-pairing interactions between the flanking nucleotides (Xia et al., 1998). The siRNA strand with the 5′ end corresponding to the less stable end of the duplex is preferentially selected and such an siRNA is called “asymmetric”. siRNAs, both strand of which have equal chances to be incorporated into RISC, are called “symmetric”. Both asymmetry and symmetry are
functional terms because they are specified by underlying RNA-protein interactions. In *Drosophila*, dsRNA is processed into siRNAs by Dicer-2 (Dcr-2), one of the two Dicers encoded in the genome, in an ATP-dependent fashion (Fig. 2) (Nykanen et al., 2001).

Neither recognition nor processing of dsRNA by Dcr-2 in vitro requires any auxiliary proteins; recombinant Dcr-2 directs efficient production of siRNAs (Liu et al., 2003). However, Dcr-2 function in vivo requires the dsRNA-binding (dsRB) protein R2D2. r2d2 mutant embryos are completely defective for the dsRNA-initiated RNAi response. R2D2 contains two dsRNA binding domains (dsRBD) and is homologous to *C. elegans* protein RDE-4. RDE-4 binds dsRNA in vitro and is required for processing of long dsRNA by *C. elegans* Dicer (DCR-1) in vivo (Tabara et al., 2002, Parrish and Fire, 2001, Parker et al., 2006). In contrast, R2D2 is dispensable for dsRNA processing activity of Dcr-2 in vivo (Liu et al., 2006). R2D2 is unstable in the absence of Dcr-2, which suggests that dsRNA processing in vivo is probably directed by the Dcr-2/R2D2 complex rather than Dcr-2 alone (Liu et al., 2003, Liu et al., 2006). The major function of R2D2, however, is in assisting Dcr-2 in loading of siRNAs into RISC, a function of Dcr-2 both genetically and biochemically different from dsRNA processing (Lee et al., 2004, Pham et al., 2004). siRNA strand selection for incorporation into RISC is independent from the dsRNA processing polarity of Dcr-2, which suggests that newly generated siRNAs are released from Dcr-2 before they re-enter the Dcr-2-dependent RISC assembly pathway (Preall et al., 2006). Neither Dcr-2 nor R2D2 can bind an siRNA alone (Liu et al., 2003, Tomari et al., 2004b). R2D2 recognizes 5′ phosphates at the ends of an siRNA, and in asymmetric siRNA, binds preferentially to the end with more double-stranded RNA character. Dcr-2 binds to the available end of the siRNA in a coordinate fashion with R2D2. Both dsRBDs
Figure 2. RNAi and miRNA pathways in *Drosophila*.
of R2D2 are required for this association. Phosphates at the 5´ ends do not affect Dcr-2 binding; suggesting R2D2 is responsible for specific recognition of Dicer products. Thus, R2D2 binding to an siRNA has two aspects. First, it recognizes the identity of the siRNA and second, it is the primary protein sensor of siRNA thermodynamic stability. Binding of Dcr-2/R2D2 to the siRNA promotes its loading into siRISC.

Several lines of evidence indicate Dicer is also required for RISC assembly in human cells. In mammalian cells, long dsRNA induce non-specific interferon (INF) response mainly by activating two enzymes 2´-5´-oligoadenylate synthetase required for activation of sequence-non-specific RNase L and protein kinase R, the active form of which phosphorylates the eukaryotic translation initiation factor 2, leading to general inhibition of protein synthesis and consequently cell death. However, if dsRNA is introduced in mouse oocytes, early embryos or embryonic stem cells, which are refractory to INF response activation, similar to worms and flies, it leads to specific inhibition of gene expression (Svoboda et al., 2000, Yang et al., 2001, Billy et al., 2001, Wianny and Zernicka-Goetz, 2000). Human RISC works inefficiently when programmed with siRNAs, which suggests that there is a direct coupling of Dicer processing activity with loading of siRNAs into RISC (Maniataki and Mourelatos, 2005, Martinez and Tuschl, 2004). Recombinant human Dicer binds siRNAs in an ATP-independent fashion (Pellino et al., 2005). Dicer complex assembled on an siRNA in vivo is ~250-300 kDa, which suggests that it contains additional components. Assembly of the Dicer complex is ATP-dependent. It contains only double-stranded siRNAs, which suggests that it is distinct from RISC. In crosslinking experiments, human Dicer binds preferentially to the
3’ end of the guide strand, suggesting siRNA asymmetry might be recognized differently in human cells than in *Drosophila*.

Active siRISC purified from *Drosophila* S2 cells is ~ 500 kDa. It consists of several proteins including Argonaute 2 (Ago2), Vasa Intrinsic Gene (VIG), *Drosophila* homolog of human Fragile X Retardation Protein (dFXR) and Tudor staphylococcal nuclease (TSN) (Caudy et al., 2003, Caudy et al., 2002, Hammond et al., 2001, Ishizuka et al., 2002). Active RISC purified from *Drosophila* embryo extracts is smaller, <230 kDa (Nykänen et al., 2001). The smallest complex retaining endonuclease activity of RISC was purified from HeLa cells. It is <160 kDa. Similar to *Drosophila* RISC, it contains Argonaute proteins, human Ago1 (hAgo1) and/or hAgo2 (Martinez et al., 2002). Argonaute proteins comprise two highly conserved domains: the PAZ domain (first identified in the *Drosophila* protein Piwi and the plant proteins Argonaute and Zwille) and the PIWI domain (Cerutti et al., 2000). The crystal structure of archaean Argonaute proteins revealed that the PIWI domain is similar to RNase H, strongly implicating Argonautes as the sole elements possessing endonucleolytic “slicer” activity (Song et al., 2004, Parker et al., 2004). Indeed, recombinant hAgo2 and an siRNA form minimal active RISC (Rivas et al., 2005). Although the exact role of auxiliary RISC components is unknown, they, perhaps, enhance or modify the intrinsic activity of Argonaute proteins.

Several known characteristics of mRNA cleavage by RISC are consistent with an RNase H-like enzyme. Both *Drosophila* and human RISCs are Mg$^{2+}$-dependent endonucleases producing a 3’ cleavage product with 5’ phosphate and a 5’ cleavage product with 3’ hydroxyl (Martinez and Tuschl, 2004, Schwarz et al., 2004). As a true enzyme, RISC performs multiple rounds of catalysis without consuming its guide siRNA. Human RISC
cleaves ~10 target molecules and *Drosophila* one more than 50 (Haley and Zamore, 2004, Hutvagner and Zamore, 2002). While RISC does not require ATP for target recognition or cleavage, product release is greatly facilitated by ATP, suggesting that an ATP-dependent RNA helicase acts to dissociate the cleavage products from the siRNA guide (Haley and Zamore, 2004). Alternatively, ATP may promote a conformational change in RISC itself, releasing the cleaved target (Rivas et al., 2005). Following endonucleolytic cleavage by RISC, cleavage products are actively degraded from the 3´end by the exosome and from the 5´end by the 5´-to-3´ exonuclease, XRN1. Neither process requires prior decapping or deadenylation (Orban and Izaurralde, 2005).

Although siRNAs typically pair fully with their target RNAs, the 5´, central and 3´ regions of an siRNA make distinct contributions to binding and catalysis. The first position of a guide siRNA (G1) does not contribute much to the binding energy of RISC with target mRNA. Nucleotides G2 through G8 provide most of the energy for target binding. This region of the small RNA is called the seed sequence. Central and 3´regions, however, must be base-paired with the target in order to achieve efficient catalysis. In general, target cleavage requires base pairing of nucleotides 2 to12, which corresponds to one turn of A-form helix. RISC is remarkably tolerant to mismatches. Up to nine contiguous mismatches at the 3´end and up to five at the 5´end can be tolerated to achieve substantial target cleavage (Haley and Zamore, 2004, Martinez and Tuschl, 2004). This suggests that siRNAs can potentially cause off-target effects if present at concentrations higher then those of their targets, although regulation by siRNAs with so little pairing with their target RNAs is unlikely to proceed through an endonucleolytic mechanism in vivo.
RISC assembly

RISC assembly, the process by which a small RNA is loaded into an Argonaute protein, has been most extensively studied for Drosophila Ago2. Assembly of Drosophila Ago2-RISC in vitro is a highly ordered process. Two similar RISC assembly pathways have been proposed. First begins with formation of “complex B”, which is the immediate precursors of the RISC loading complex (RLC), which requires Dcr-2, its dsRNA-binding partner protein R2D2, and is stimulated by ATP. Complex B can be assembled in the absence of ATP but only on a pre-phosphorylated siRNA (Tomari et al., 2004a). RLC contains both double-stranded and single-stranded siRNA species but RISC contains only single-stranded siRNAs, suggesting separation of two strands of an siRNA is initiated in this complex. Strand separation requires Ago2, since double-stranded siRNAs accumulate in the RLC formed in lysates prepared from ago2 mutant embryos.

As the assembly proceeds, Dcr-2 and then R2D2 are exchanged with Ago2, placing the 5’ end of the guide strand of an siRNA in the phosphate-binding pocket of PIWI domain of Ago2 (Matranga et al., 2005, Rand et al., 2005). The single-stranded RNA-binding PAZ domain of Ago2 is hypothesized to bind the 3’ end of the guide strand, displacing R2D2. For an siRNA duplex, the resulting configuration of siRNA on Ago2 is essentially the same as an Ago2:siRNA complex bound to a target RNA. In vitro and cell culture data suggest that endonucleolytic cleavage of the passenger strand initiates its displacement from Ago2, generating a functional RISC in which the guide strand is ready to bind an mRNA target. Passenger strand cleavage greatly facilitates loading of siRNA into RISC but it is not absolutely required for RISC assembly (Matranga et al., 2005). In worms and humans, Dicers seem to also participate in RISC loading. Both Dicers have been found in
complexes with Argonaute proteins Rde-1 and hAgo2, respectively (Tabara et al., 2002, Chendrimada et al., 2005).

Another group identified two complexes named “R1” and “R2” in the order of their formation, which immediately precede complex R3 (RISC) (Pham et al., 2004). Complex R3 is very large ~ 80S. It contains only single-stranded siRNAs, requires ATP for formation and co-fractionates with known RISC components: VIG, TSN, Ago2 and dFXR. Complex R1 is indistinguishable from Dcr-2/R2D2 heterodimer bound to an siRNA (Pham and Sontheimer, 2005). In addition to Dcr-2/R2D2, complex R2 contains Ago2, suggesting R2 is similar to RLC. RLC formation, however, requires ATP. R2 forms in the absence of ATP and contains only double-stranded siRNAs. Dicer-1 (Dcr-1), the second Drosophila Dicer, is also required for complex R2 formation and dcr-1 mutant embryos assemble only complex R1. This suggests that Dcr-1 and Dcr-2 are components of a common complex. Indeed, when Drosophila embryo lysate is programmed with exogenous siRNAs, both Dicers and R2D2 co-sediment together in an 80S complex. In the absence of siRNAs, however, the 80S complex contains the same components except for Dcr-2/R2D2. Because Dcr-1 in flies is required for biogenesis of endogenous small RNAs called micro-RNAs (miRNAs), the 80S complex probably represents holo-RISC pre-assembled on miRNAs. Holo-RISC might interact with either Dcr-2/R2D2 or RLC only when siRNAs are present. Association with ribosomes could explain large size of holo-RISC. Drosophila RISC components are found in ribosomal fraction during purification (Hammond et al., 2001, Caudy et al., 2003, Caudy et al., 2002). Both Ago2 and dFXR are in complexes containing ribosomal proteins, L5 and L11 (Ishizuka et al., 2002). Association with polyribosomes might indicate the requirement for active
translation for efficient RNAi. Although translation is not required for RNAi in
*Drosophila* embryo lysate, untranslated mRNAs are refractory to RNAi in vivo in
*Drosophila* oocytes (Kennerdell et al., 2002). siRNAs co-sediment with actively
translating polyribosomes in *Trypanosoma brucei* (Djikeng et al., 2003). Perhaps,
translation increases target accessibility, which affects RISC mediated catalysis, at least
in human cells (Brown et al., 2005).

An alternative association of RISC with cytoplasmic processing bodies (P-bodies)
has been recently proposed. P-bodies, also referred to as GW or Dcp bodies, contain
GW182 protein and might play a role in RNAi (Jakymiw et al., 2005, Sen and Blau,
2005, Rehwinkel et al., 2005). Target mRNAs and RISC components are co-localized in
P-bodies, which suggests that P-bodies might be the *bona fide* site for RISC-mediated
target cleavage.

**miRNAs: a historical perspective**

If a small RNA molecule is imperfectly complimentary to the target, instead of
endonucleolytic cleavage, it induces translational repression and/or destabilization of the
mRNA. Such a mode of silencing is characteristic of animal miRNAs. Only a few animal
miRNAs pair extensively with their targets to promote endonucleolytic cleavage. In
contrast, all plant miRNAs direct endonucleolytic cleavage of their target mRNAs.
miRNAs are close relatives of siRNAs. They have the same length of 21-24 nt and differ
from siRNAs only in details of their biogenesis. miRNAs regulate expression of
endogenous genes and may even function to restrict the host range and tissue-tropism of
viruses infecting mammals (Lecellier et al., 2005, Jopling et al., 2005).
The first miRNAs (formerly small temporal (stRNAs)), lin-4 and let-7, were identified in a genetic screen for mutations that disrupt the timing of development in *C. elegans*. *C. elegans* development is regulated by a cascade of heterochronic genes, the repression or activation of which in tissues throughout the animal coordinates the succession of cell fates within lineages. At each stage, a subset of these regulators acts as a switch between alternative choices of cell fates. If a heterochronic gene is mutant, development is either precocious, skipping stage-specific events, or retarded, reiterating stage-specific events in subsequent stages, with respect to other unaffected events (Ambros and Horvitz, 1984). One of the identified heterochronic genes, *lin-4*, instead of protein, encodes a 21-nt single-stranded miRNA (Lee et al., 1993). *lin-4* temporally negatively regulate the expression level of LIN-14 and, in part, LIN-28 through imperfect basepairing with multiple repetitive complementary sites in *lin-14* and a single site in *lin-28* 3′-untranslated regions (UTR). These interactions signal the transition from the first to later larval patterns of cell lineage and differentiation (Moss et al., 1997, Feinbaum and Ambros, 1999, Seggerson et al., 2002). Animals carrying a *lin-4* loss-of-function mutation display reiterations of early fates at inappropriately late developmental stages. Although *lin-4* is found only in nematodes, its target LIN-28 seems to be expressed and developmentally regulated in many animal species (Moss and Tang, 2003). *lin-4* is an example of a multifunctional miRNA. In addition to regulating developmental timing of larval stages, it also regulates life span in the adult in a *lin-14*-dependent fashion. *lin-4* mutants have shortened life span and accelerated tissue aging. Overexpression of *lin-4* extends their life span (Boehm and Slack, 2005). Another 21-nt miRNA, *let-7*, negatively regulates LIN-41 during L3 and later stages of development by means of imperfect
basepairing interactions with two complementary sequence motifs in the lin-41 3’ UTR. This interaction signals the transition to the adult stage, since LIN-41 downregulates the expression of adult transcriptional factor LIN-29 (Slack et al., 2000, Vella et al., 2004). In let-7 null mutants, larval fates are reiterated in adult worms (Reinhart et al., 2000). let-7 sequence is highly conserved in all three clades of bilaterian animals: the deuterostomes, lophotrochozoans, and ecdysozoans. In diverse species, let-7 is expressed in characteristic differential developmental patterns; it tends to be expressed at later developmental stages, possibly regulating the transition to adult (Pasquinelli et al., 2000).

**miRNA-directed mRNA destabilization and translational repression**

Initial studies demonstrated that lin-4 directs repression of lin-14 mRNA post-transcriptionally and suggested that repression occurred through a block to translation of LIN-14 protein (Wightman et al., 1993, Olsen and Ambros, 1999). Surprisingly, repression of LIN-14 protein synthesis by lin-4 did not appear to alter the association of either lin-14 mRNA with polysomes or the length of the lin-14 poly(A) tail, indicating that LIN-14 translation was repressed after the initiation of protein synthesis. Similar conclusions were reached for a second lin-4-regulated gene, lin-28 (Moss et al., 1997, Seggerson et al., 2002). More recent re-examination of the mechanism by which the miRNA let-7 regulates expression of lin-41 and by which lin-4 regulates lin-14 and lin-28 suggested that these miRNAs act mainly by decreasing the stability of their mRNA targets, rather than by altering their rates of translation (Bagga et al., 2005). Apparently conflicting data has also been reported for human miRNA regulatory mechanisms. In cultured human cells, endogenous let-7 appeared to repress translation of a reporter
construct by decreasing the rate of translational initiation (Pillai et al., 2005, Humphreys et al., 2005). In contrast, siRNAs engineered to bind imperfectly to multiple sites in the 3’ untranslated region of a reporter mRNA appear to repress protein synthesis without altering the rate of translational initiation (Doench et al., 2003, Petersen et al., 2006, Doench and Sharp, 2004).

miRNAs can sequester their mRNA targets in P-bodies, and both miRNA-directed mRNA destruction and translational repression may reflect sequestration of the mRNA in P-bodies (Liu et al., 2005b, Pillai et al., 2005, Andrei et al., 2005, Sheth and Parker, 2006, Teixeira et al., 2005). P-bodies contain a pool of translationally repressed mRNAs (Teixeira et al., 2005, Coller and Parker, 2005). Mammalian Argonaute proteins Ago1 and Ago2 directly associate with P-/GW-bodies component GW182. Depletion of GW182 impairs small RNA-directed silencing of a reporter mRNA (Jakymiw et al., 2005, Behm-Ansman et al., 2006, Liu et al., 2005a). Mutations in hAgo1 and hAgo2, which prevent them from localizing to P-/GW-bodies, disrupt translation inhibition of a target mRNA when both proteins are tethered to the 3’ UTR of the target through a bacteriophage RNA-binding protein (Liu et al., 2005a). Yet, RCK/p54, an ATP-dependent DEAD box helicase and a general repressor of translation found in P-bodies, interacts with hAgo1 and hAgo2, and this interaction as well as miRNA-directed translational repression of target mRNAs is independent of P-body integrity. Thus, it remains possible that miRNA-directed translational repression does not require P-bodies per se, but rather that the re-localization of miRNA-repressed mRNAs to P-bodies is a consequence of translation repression (Chu and Rana, 2006). Adding to the plethora of proposed mechanisms for miRNA-directed regulation of mRNA expression, Drosophila,
zebrafish and mammalian cells can also promote mRNA degradation by triggering poly(A) tail shortening and by promoting removal of the 5´ cap (Wu et al., 2006, Behm-Ansmant et al., 2006, Humphreys et al., 2005, Giraldez et al., 2006). A more detailed molecular understanding of all these proposed mechanisms for miRNA function awaits the development of robust in vitro systems that recapitulate both the loading of miRNAs into RISC and the subsequent repression of target mRNA expression by these in vitro generated, miRNA-programmed Argonaute protein complexes.

**miRNA biogenesis**

Although miRNAs accumulate as 21-23-nt single-stranded molecules, they originate as much longer primary miRNAs (pri-miRNAs), hundreds of which are encoded in plant and animal genomes (Lau et al., 2001, Lagos-Quintana et al., 2001, Lee and Ambros, 2001). pri-miRNAs are transcribed by RNA polymerase (pol) II, and thus contain a 5´cap and a poly (A) tail (Cai et al., 2004, Lee et al., 2002). miRNAs are also encoded in genomes of some animal viruses and transcribed by either RNA pol II or III (Cai et al., 2006, Pfeffer et al., 2004, Pfeffer et al., 2005, Sullivan et al., 2005). Within the pri-miRNA, the mature miRNA resides in one arm of a ~70 nt stem-loop structure. A pri-miRNA may contain a single miRNA—with thousands of bases transcribed, apparently to yield a single 21 nt RNA guide—or several miRNAs (polycistronic pri-miRNA). The majority of worm and human miRNAs reside in their own pri-miRNA transcripts, whereas more than half of *Drosophila* miRNAs are polycistronic (Bartel, 2004, Aravin et al., 2003). Pri-miRNA must undergo additional processing steps in order to become mature miRNA. In the nucleus, a complex of proteins including the RNase III enzyme
Drosha and its dsRNA-binding protein partner, Pasha in flies or DGCR8 in mammals recognize the stem-loop of pri-miRNA, cleaves it about one helical turn up from the base of the stem, liberating a ~60 nt precursor of miRNA (pre-miRNA) (**Fig. 2**) (Denli et al., 2004, Gregory et al., 2004, Han et al., 2004a). Like all RNase III enzymes, Drosha cleavage produces 2-nt, 3’ overhanging ends (Lee et al., 2003). Pasha/DGCR8 binds pri-miRNA, restricting Drosha to pri-miRNA processing (Lee et al., 2003). Pasha comprises an amino-terminal WW domain and two carboxy-terminal dsRBDs. In vivo, Drosha and Pasha/DGCR8 are components of ~ 400-650 kDa complex named “Microprocessor”. A higher molecular mass (~ 15 kDa higher than in the 400-650 kDa complex) Drosha has been purified from mammalian cells as a part of much bigger (>700 kDa) but inactive complex. Among the identified components of that complex are Ewing’s sarcoma gene protein, the DEAD/H box polypeptide 17, a putative RNA helicase, and heterogeneous nuclear ribonucleoprotein M (Gregory et al., 2004). In plants, HYL1 (*hyponastic leaves*) is the probable functional homolog of Pasha/DGCR8. HYL1 contains a nuclear localization signal and two dsRBDs (Lu and Fedoroff, 2000). *hyl1* mutants have developmental defects, which are the result of decreased levels of mature miRNAs and, consequentially, accumulation of their target mRNAs (Vazquez et al., 2004a, Han et al., 2004b).

Pre-miRNA processing

After Microprocessor releases a pre-miRNA, Exportin 5, a member of karyopherin β family of nucleocytoplasmic transport factors, transports it to the cytoplasm (Yi et al., 2003, Lund et al., 2004). Exportin 5 mediates transport of both protein and RNA cargo. It
recognizes a characteristic minihelix motif in RNA the structural cis-acting export element that comprises a double-stranded stem (>14 nt) with a base-paired 5’ end and a 3–8 nt protruding 3’ end and, in an RanGTP-dependent manner, transports it to the cytoplasm where hydrolysis of Ran-GTP to Ran-GDP induces release of the pre-miRNA cargo (Brownawell and Macara, 2002, Gwizdek et al., 2004).

In plants, nuclear Dicer-like-1 (DCL1) enzyme processes both pri-miRNAs and pre-miRNAs, which are transported to the cytoplasm as mature miRNAs (Papp et al., 2003, Kurihara and Watanabe, 2004). The plant ortholog of Exportin 5, HASTY, is required for stabilization of pre-miRNAs and miRNAs in the nucleus (Park et al., 2005, Zeng and Cullen, 2004).

In the cytoplasm, a pre-miRNA is recognized by Dicer, which cleaves the loop off the pre-miRNA generating the second end of the miRNA/miRNA* duplex molecule containing 2-nt 3’ overhang. Biochemical data suggest that processing by Drosha increases accuracy and production of mature miRNAs by Dicer (Lee et al., 2003). In Drosophila, processing of pre-miRNAs requires Dcr-1 and its dsRNA binding partner Loquacious (Loqs) (Fig. 2) (Forstemann et al., 2005, Saito et al., 2005, Jiang et al., 2005). Loqs comprises three dsRBDs (two canonical dsRBDs at the N-terminus, and one non-canonical dsRBD at the C-terminus). Loqs not only stimulates pre-miRNA processing activity of Dcr-1 in vitro, it makes it specific. Indeed, Dcr-1 alone efficiently processes long dsRNA and pre-miRNA into ~ 21-23-nt small RNAs. Loqs restricts Dcr-1 to pre-miRNA processing, inhibiting this non-specific effect (Saito et al., 2005). Loqs is absolutely required for pre-miRNA processing in vivo; loqs mutant flies contain highly elevated levels of pre-miRNAs and display desilencing of a miRNA-regulated reporter
transgene (Forstemann et al., 2005).

Mammals appear to use two Loqs homologs, transactivating response RNA-binding protein (TRBP) and PACT (Fig. 3) (Maniataki and Mourelatos, 2005, Lee et al., 2006a, Gregory et al., 2005). Dicer binds TRBP in vitro and, likely, in vivo, since either of the proteins is unstable in the absence of the other (Chendrimada et al., 2005). TRBP and PACT contain three dsRBDs and use the third dsRBD for interaction with Dicer (Lee et al., 2006a). Both proteins are not required for pre-miRNA processing by Dicer. TRBP and PACT have partially redundant functions in RISC assembly. TRBP is required at the downstream step of loading of miRNAs and siRNAs into human RISC, since depletion of the protein in human cells leads to decrease in silencing of a reporter gene but does not change miRNA abundance (Doi et al., 2003, Chendrimada et al., 2005). Depletion of PACT, on the other hand, decreases the abundance of mature miRNAs, but has really little or no effect on the silencing of the reporter.

Similarly to siRNA, miRNA/miRNA* duplexes have 2-nt 3´ overhangs. In contrast to siRNA duplexes, they contain mismatches between two strands usually in the middle of the duplex.

In plants, both siRNA and miRNA/miRNA* duplexes are methylated at the 2´ hydroxyl of the 3´ nucleotide by methyltransferase HEN1 most likely to protect the 3´ends of small RNAs from uridylation (Yang et al., 2006, Park et al., 2002, Yu et al., 2005). In hen1 mutants, the size of small RNAs increases due to addition of one to five uracil residues to the 3´ ends (Li et al., 2005).

Current data suggest that, similarly to the RNAi pathway, the strand with the less stably base-paired 5´ end is preferentially selected to become the miRNA, whereas the
CHAPTER I, Figure 3

Figure 3. RNAi and miRNA pathways in mammalian cells.
strand with the more stably base-paired 5’ end, the miRNA* strand, is degraded (Khvorova et al., 2003, Schwarz et al., 2003). How the thermodynamic asymmetry of miRNA/miRNA* duplexes is recognized remains to be elucidated.

Realization that miRNA/miRNA* duplexes are true intermediates in the miRNA biogenesis pathway came as a result of studies of viral suppressors of RNA silencing in plants. Viral suppressors, such as P19 of *Cymbidium ringspot virus*, have very low affinity for 21-nt single-stranded RNA, but bind siRNA duplexes very efficiently blocking the progression of RNAi (Silhavy et al., 2002, Lakatos et al., 2004). Surprisingly, the P19 of *Tomato bushy stunt virus* was found to co-immunoprecipitate with both miRNA171 and miRNA171* in infected *Arabidopsis* (Dunoyer et al., 2004).

**Endogenous plant small RNAs**

Small RNAs are produced not only from exogenous dsRNA, but from endogenous dsRNA as well. In plants, three classes of endogenous small RNAs have been found.

*Trans*-acting siRNAs (ta-siRNAs) derive from long non-coding single-stranded RNAs that are cleaved by certain miRNAs. Truncated RNA is transformed to dsRNA by the RDR6 and processed into 21-nt siRNAs by one of the plant Dicer like proteins, DCL4 (Allen et al., 2005, Gasciolli et al., 2005, Vazquez et al., 2004b, Xie et al., 2005, Peragine et al., 2004, Yoshikawa et al., 2005). Both strands of the formed siRNAs direct cleavage of endogenous mRNAs *in trans*, and antisense strands of the duplex can mediate cleavage of their own precursor mRNA *in cis*, establishing negative feedback mechanism of regulation of their production.

The second class of endogenous siRNAs is 24-nt species, which are produced from
dsRNA formed as a result of overlapping transcription of a pair of neighboring genes by DCL2 in the RDR6-dependent manner. Initial cleavage in cis of the mRNA of only one of the pair of the genes sets a phase for the formation of the additional dsRNA and generation of 21-nt siRNAs by DCL1 and further cleavage of the mRNA (Borsani et al., 2005).

The third class of endogenous siRNAs combines those corresponding to repetitive sequences in centromeric and pericentromeric regions as well as transposons and retroelements (Xie et al., 2004). Production of these 24-nt primary siRNAs requires plant specific RNA polymerase IV, DCL3 and RDR2 to direct DNA methylation and heterochromatin formation (Chan et al., 2004, Herr et al., 2005, Kanno et al., 2005, Onodera et al., 2005).

**siRNA-mediated heterochromatin formation**

Mechanisms directing heterochromatin formation mediated by endogenous siRNAs are best characterized in fission yeast, *Schizosaccharomyces pombe*. In this organism, heterochromatin is associated with centromeres, telomeres, and the silent region inside the mating-type locus. Heterochromatin formation at tandem and inverted arrays of centromeric repeats that surround the unique central core is required for cohesion and proper segregation of chromosomes (Huang and Moazed, 2006). At the K-region of the mating type locus, which is homologous to centromERIC repeats, heterochromatin suppresses recombination, promoting long-distance chromosome interactions, which are essential for non-random selection of donor loci during mating-type switching (Grewal and Klar, 1997, Jia et al., 2004).
Accumulation of siRNAs corresponding to centromeric repeats requires *S. pombe* Dicer, Dcr1. The substrate for Dcr1 might be either bi-directionally transcribed by RNA pol II centromeric repeats or dsRNA synthesized on nascent pol II transcripts by *S. pombe* RdRP, Rdp1. Binding of RISC-like RNA-mediated Initiation of Transcriptional Silencing (RITS) complex, composed of siRNAs, Argonaute protein Ago1, chromodomain protein Chp1 and protein of unknown function TAS3, to centromeric regions leads to methylation of lysine 9 of histone 3 (H3 K9) by the methyltransferase Clr4, recruitment of the chromodomain protein Swi6 and heterochromatin formation (Verdel et al., 2004). It is not known if RITS binds centromeric DNA or nascent transcripts. Both sense and antisense strands of siRNAs are present in RITS, suggesting DNA binding (Cam et al., 2005). RITS components, however, can be immunoprecipitated with RNA corresponding in sequence to centromeric repeats, suggesting binding to nascent transcripts (Motamedi et al., 2004). RITS might be recruited to nascent transcripts in order to cleave them via the slicing activity of Ago1 or to localize another complex called RNA-directed RNA polymerase complex (RDRC). RDRC binds to both RITS complex and centromeric RNA (Irvine et al., 2006, Motamedi et al., 2004). RDRC has an RdRP activity. It consists of Rdp1 and two conserved proteins, Hrr1, putative RNA helicase, and Cid12, a member of the poly (A) polymerase. siRNA accumulation in *S. pombe* depends not only on Dcr1 but also on components of RDRC, which suggests that only dsRNA generated by Rdp1 can be utilized by Dcr1. Rdp1 does not use siRNAs to prime the synthesis of dsRNA, because tethering of RITS downstream of euchromatic gene, although dependent on Rdp1, is sufficient to initiate the silencing and heterochromatin formation (Buhler et al., 2006). Clr4 is also required
for siRNA accumulation. In *clr4Δ* mutants, the association of RITS with Rdpl and with centromeric RNA is abolished. If RITS indeed recruits RDRC to RNA templates, the failure to localize RITS to centromeric RNA would result in a failure to recruit RDRC to initiate dsRNA synthesis and siRNA production (Motamedi et al., 2004).

Heterochromatin formation in *S. pombe* correlates with both transcriptional and posttranscriptional silencing of underlying repetitive sequences mediated by a single Argonaute protein (Volpe et al., 2002, Yamada et al., 2005, Noma et al., 2004).

*Drosophila* Argonaute proteins Aubergine (Aub) and Piwi and *C. elegans* Alg-1 are required for transcriptional silencing and heterochromatin formation (Pal-Bhadra et al., 2004, Grishok et al., 2005).

In mammalian cells, similar to plants, siRNAs directed against the promoter of the reporter gene induce its transcriptional silencing, DNA methylalation and/or formation of silenced histone modifications (H3 K9 and H3 K27 methylation) (Morris et al., 2004, Weinberg et al., 2006, Janowski et al., 2006, Ting et al., 2005). TGS requires hAgo1, hAgo2, TRBP and, in some cases, correlates with binding of the human homolog of the *Drosophila* Enhancer of Zeste (E(Z)) histone methyltransferase component of the Polycomb group (PcG) of transcriptional repressors, EZH2, connecting RNAi and Polycomb-based silencing (Kuzmichev et al., 2002, Kim et al., 2006, Janowski et al., 2006). Thus, RNAi machinery might direct transcriptional silencing through heterochromatin formation and recruitment of PcG.

Current evidence suggests a role for RNAi machinery in maintaining nuclear organization. For example, in *Drosophila*, Dcr-2, Piwi and Ago1 are not required for establishment of PcG-based silencing but are indispensable for its maintenance, by
stabilizing clustered PcG target elements at specific nuclear bodies (Grimaud et al., 2006). Similarly, in *S. pombe*, Ago1, Rdpl and Dcr1 are not required for heterochromatin formation at telomeres, but are required for their nuclear clustering (Sugiyama et al., 2005, Hansen et al., 2006, Kanoh et al., 2005).

**Small RNA clusters**

Endogenous small RNAs found in worms are as abundant as endogenous plant siRNAs. The first class of endogenous small RNAs, ~22-nt single-stranded siRNAs, is antisense to mRNAs encoded in the genome of *C. elegans* (Ambros et al., 2003). In worms, siRNAs processed from exogenous dsRNA also accumulate primarily in antisense polarity to their corresponding mRNAs in a process termed target-dependent accumulation (Tijsterman et al., 2002b).

The second class of endogenous small RNAs in worms, ~20-22-nt tiny noncoding RNAs (tncRNAs), derives from sequences outside the protein coding regions. Like miRNAs, they do not have perfectly matched targets, but, unlike precursors of miRNAs, they cannot fold into hairpins. On Chromosome X of *C. elegans*, tncRNA genes form a cluster, which consists of 41 distinct sequences that are oriented in the same direction. Accumulation of endogenous small RNAs is dependent on factors required for exogenous RNAi in worms, including DCR-1 and RDE-4, and some components, including RdRP RRF-3 and the putative 3´-5´ exonuclease ERI-1, that are specific to endogenous RNAi (Lee et al., 2006b).

Recently, endogenous ~24-30-nt RNAs have been found to originate from genomic clusters in mammalian genomes. These small RNAs are named PIWI-interacting RNAs
pain RNAs) for their specific associated with proteins of the Piwi subfamily of Argonaute proteins, expression of which is restricted to the germ line (Fig. 4) (Aravin et al., 2006, Girard et al., 2006, Grivna et al., 2006, Lau et al., 2006, Watanabe et al., 2006).

Importantly, almost all piRNA genes in a given cluster are in the same orientation. This extreme strand bias suggests that piRNAs may be processed from long primary transcripts. In some instances, neighboring clusters seem to diverge from the same location, which suggests that they may be transcribed from a central promoter. Currently, there is no evidence of how piRNAs are generated or what are the possible functional interactions they are involved in. Two mouse piRNA-associated proteins, MIWI and MILI, are required for normal spermatogenesis and expressed at different but overlapping stages (Kuramochi-Miyagawa et al., 2004, Deng and Lin, 2002). It is possible that their role in spermatogenesis is mediated by piRNAs.

**RNase III enzymes in RNA silencing**

RNase III enzymes can be divided into three major classes. Class I enzymes contain a catalytic RNase III domain and a dsRBD. An example of such enzyme is *Escherichia coli* (E. coli) RNase III. Although bacterial RNase III orthologs are the structurally simplest family members, they are able to cleave their cellular substrates in a highly site-specific manner, which is determined by the specific RNA structural and sequence elements, also referred to as reactivity epitopes (Pertzev and Nicholson, 2006). Substrates that exhibit strict double-helical structure undergo coordinate cleavage of both strands, while substrates with internal loops or bulges usually are cleaved at a single site, which is contained within or near the internal loop or bulge (Nicholson, 1999, Amarasinghe et al.,
Figure 4. piRNA pathways in mammalian cells.
Discovery of bulge-helix-bulge motif, which can permit binding of *E. coli* RNase III, but inhibits cleavage, suggests the possible role for RNase III as a dsRNA binding protein (Blaszczyk et al., 2004, Calin-Jageman and Nicholson, 2003, Dasgupta et al., 1998, Gan et al., 2005).

Since bacterial RNase III orthologs function as homodimers, the holoenzyme contains two dsRBDs. Recently solved crystal structure of catalytically active bacterial RNAse III and its substrate is highly symmetric. It supports the proposed single catalytic center model that suggests a single RNA cleavage event occurs on each strand of the RNA within each cleavage site (Zhang et al., 2004). Overall 4 RNA binding motifs (RBM) in each monomer contribute to protein-RNA interactions. While RBMs 1 and 2 in the dsRBD are responsible for dsRNA recognition and binding, RBMs 3 and 4 in the RNase III domain are responsible for substrate recognition and scissile-bond selection.

Dimerization of RNase III is essential for enzyme function in that residues from one subunit (RBM 3 residues) are involved in the selection of the scissile bond, while those from the partner subunit (strictly conserved residues) are involved in the cleavage chemistry (Gan et al., 2006). The RNase III catalysed dsRNA cleavage is Mg\(^{2+}\)-dependent process. Although only one Mg\(^{2+}\) ion is coordinated in each active site in the RNase III-product structure, biochemical data suggest two-metal-ion catalysis (Sun et al., 2005). The exact answer of how many Mg\(^{2+}\) ions participate in catalysis waits the time when RNase III-substrate complex structure is solved. Exhaustive cleavage of polymeric dsRNA produces duplex products of ~12-15 nt, which corresponds to slightly greater than one turn of the A-helix (11 base pairs) (Amarasinghe et al., 2001). 21-25-nt size-selected RNAs obtained by limited digestion by recombinant RNAse III induce efficient
target cleavage in both *Drosophila* S2 and cultural mammalian cells, suggesting that size is the only difference between RNase III and Dicer products (Yang et al., 2002).

Bacterial RNase III enzymes are required for processing of ~5500-nt primary ribosomal RNA (rRNA) precursors to the immediate precursors of mature 16S, 23S and 5S rRNAs, maturation of tRNAs and mRNAs, and separation of mRNAs from co-transcribed tRNAs (Conrad and Rauhut, 2002). In *E. coli*, the degradation of mRNA is generally triggered by endoribonucleolytic cleavage, and the resulting intermediate products are further degraded by endo- and 3´ to 5´ exoribonucleases (Jain, 2002). RNase III plays an important role in mRNA decay particularly through cleavage of a stem loop upstream of the initiation codon in *pnp* transcripts encoding polynucleotide phosphorylase, which together with RNase II are two major exonucleases in *E.coli*. In RNase III deletion mutant, *pnp* mRNA levels are increased more then 11 fold (Portier et al., 1987). RNase III is responsible for the cleavage of small non-coding regulatory RNA (sRNA) RyhB upon its binding to the start codon region of *sodB* mRNA encoding the iron superoxide dismutase. When *E.coli* adjusts to decreasing concentrations of iron, base pairing between RyhB and *sodB*, in a process requiring binding by RNA chaperone Hfq, leads to translational inhibition of mRNA followed by endonucleolytic cleavage by the RNase E endonuclease at the cryptic site in the coding region of *sodB* and its further destruction (Afonyushkin et al., 2005). Although RNase III plays an important role in RNA turnover in *E.coli*, bacteria survive in the absence of RNase III, just slow their growth, suggesting that functional redundancy exists between RNase III and other nucleases.

Between 50 and 100 sRNAs have been found in *E.coli* (Gottesman, 2004). A
majority of them regulates mRNA stability or translation. Although a few RNA regulators are antisense to their target mRNAs (cis-acting), majority of sRNA are encoded far from their targets from single-gene operons. Base pairing of such sRNAs with their targets does not require a high degree of complementarity and, therefore, provides a possibility for sRNAs to affect multiple transcripts (Tjaden et al., 2006). sRNAs are different from miRNAs in their size (some of them are several hundreds nucleotides long).

It is not clear how extensive the role of RNase III is in sRNA-mediated gene regulation. The mechanism of regulation seems to be different from the one in eukaryotes in that RNase III is required at the effector step of small RNA-mediated regulation, not at the production step. It definitely does not require Argonaute proteins (no Argonaute protein is encoded in eubacterial genomes). Interestingly, although Argonaute proteins are encoded in archaenal genomes, RNase III homologs have not been found there. What role Argonaute homologs play in archaea is unknown.

Class II RNase III enzymes comprise Drosha and its homologs. Droshas are 130-160 kDa nuclear proteins, which contain two RNase III catalytic domains, a dsRBD and a long N-terminal segment (Filippov et al., 2000). Mouse and human Drosha enzymes are 99% identical in amino acid sequence for the two catalytic domains and dsRBD. Both of them contain a proline-rich region (PRR) and a serine-and arginine-rich (RS) domain in N-terminal segment (Fortin et al., 2002, Wu et al., 2000). In mammalian cells, inhibition of Drosha expression with antisense oligonucleotides leads to accumulation of 32 S and 12 S pre-rRNAs, suggesting that in addition to its role in pre-miRNA biogenesis the protein is required for pre-rRNA processing (Wu et al., 2000). Although Drosha
orthologs have been identified in worms, flies, mice and human, no Drosha is encoded in plant or fission yeast genomes. Class II RNases have not been studied as extensively as class I and III enzymes. There are no crystallographic data on any class II enzyme. Because Drosha alone does not possess specific nuclease activity towards its substrates, all the current data on Drosha’s catalytic activity were obtained with the purified Microprocessor complex. Site-directed mutagenesis studies of human Drosha suggest that similar to class I enzymes, class II enzymes contain a single processing center with two catalytic sites. In contrast to class I enzymes, catalytic center forms as a result of intramolecular dimerization of RNase III domains of Drosha. Each catalytic site in the processing center cuts one strand of dsRNA at nearby sites in coordinate fashion, generating a short ~ 2-nt 3’ overhang (Han et al., 2004a). Deletion of the P-rich region and most of the RS-region does not affect catalytic activity of the enzyme but it might contain other important features like a nuclear localization signal. The middle region of the protein, including a part of the RS-rich region and the RNAse III domains, is required for Drosha’s binding to DGCR8, but the P-rich region is dispensable for this interaction. The middle region of the protein, the RNAse III domains, and the dsRBD are required for pri-miRNA processing. Although Drosha and DGCR8 are the only essential subunits of the Microprocessor complex, it is not clear whether the complex forms as a result of heterodimerization or heterotetramerization of its components. Current data suggest that selection of strand of the miRNA/miRNA* duplex for specific incorporation into miRISC does not depend on the position of the overhang (5’ or 3’) on the pre-miRNA after Drosha cleavage (Han et al., 2006). An average pri-miRNA consists of a stem of 3 helical turns surrounded by single-stranded RNA segments at both ends (the basal segments on one
side and the terminal loop on the other side). The terminal loop of the pri-miRNA is not essential for processing. Instead, Microprocessor recognizes the single-stranded character of the basal segments of the pri-miRNA, not their sequence, and thereby measures the ~11-nt distance from the junction of the basal segments and the dsRNA stem in order to make a cut (Zeng and Cullen, 2005, Han et al., 2006). Manipulating the length of the outer stem affects cleavage site selection, suggesting the existence of a molecular device that measures the distance from the junction of single-stranded and double-stranded regions. Because Drosha alone does not have any RNA binding activity, but DGCR8 does, it has been proposed that DGCR8 is the likely candidate on the role of molecular anchor, which recognizes the substrate and binds to it forming “pre-cleavage complex”. It is not clear whether Drosha binding to DGCR8 itself (Drosha/DGCR8 interaction does not require bridging RNA) is sufficient to create active center ~11 nt from the juncture of single-stranded and double-stranded regions, or Drosha transitively binds RNA in a context of pre-cleavage complex to determine the cleavage site.

Substrates of the Microprocessor complex are not restricted to hairpin-like structures. Long dsRNA with large internal loops separated by ~3 helical turns of dsRNA can be recognized as well (Han et al., 2006). It would be of great interest to determine if Drosha processes such substrates in vivo.

The third class of RNase III enzymes consists of Dicer-like proteins, which are approximately 200 kDa multidomain proteins. In higher eukaryotes, Dicers comprise an N-terminal DExD/H helicase domain, the DUF283 domain, C-terminal dsRBD, and characteristic for all Dicers PAZ and tandem RNase III domains. Different organisms contain various number of Dicer paralogs, which probably reflects the evolution of
specialized RNA silencing pathways. *S. pombe* encodes one Dcr1 required for accumulation of ~21-23-nt siRNAs (Motamedi et al., 2004). *Neurospora* contains two Dicers (Dcl-1 and Dcl-2), which perform redundant functions in processing of dsRNA into ~25-nt siRNAs in quelling. Dcl-1 is also required for genetically distinct pathway termed meiotic silencing by unpaired DNA (Catalanotto et al., 2004, Nakayashiki et al., 2006). *Tetrahymena thermophila* genome encodes three Dicers, all producing endogenous small RNAs; DCL1 generates ~27-31-nt scan RNAs (scnRNAs) required for developmentally programmed DNA-elimination, DCR2 is the most likely producer of ~23–24-nt RNAs expressed throughout the whole cell cycle of this ciliated protozoan, and DCR1 likely produces ~30-35-nt RNAs, which accumulate only under conditions of starvation (Lee and Collins, 2006). *Drosophila* has two biologically distinct Dicers; Dcr-1 is required for miRNA biogenesis and Dcr-2 for processing of dsRNA (Lee et al., 2004). *Arabidopsis* genome encodes four DCL proteins; DCL1 is required for miRNA biogenesis, DCL2 produces ~22-nt siRNAs during viral defense and PTGS induced with dsRNA triggers, DCL3 produces longer ~24-nt siRNAs required for heterochromatin formation, and DCL4 produces ~21-23-nt trans-acting siRNAs during the vegetative phase change as well as ~21-nt siRNAs during the PTGS induced with inverted repeat triggers and sense transgenes (Xie et al., 2005, Park et al., 2002, Margis et al., 2006). Worms and vertebrates contain only one Dicer, which has dual role in processing of dsRNA and biogenesis of miRNAs.

Similar to class II enzymes, Dicer functions as a monomer containing a single processing center formed through intramolecular dimerization of the two RNase III domains of the same molecule. The center contains two independent catalytic sites, each
cutting one RNA strand of the duplex and generating products with 2-nt 3’ overhangs (Zhang et al., 2004). Biochemical data suggest that the enzyme processes dsRNA progressively from the ends (Zamore et al., 2000, Ketting et al., 2001, Zhang et al., 2002, Vagin et al., 2006). There is no strict requirement, however, for the end availability in order to achieve cleavage. Dicer can initiate cleavage in the middle of the substrate with much slower kinetics if ends are chemically changed to tetra-loops or to RNA/DNA sequences (Zhang et al., 2002). Binding of dsRNA by Dicer does not require Mg$^{2+}$ ions but cleavage is Mg$^{2+}$-dependent. Recombinant or endogenous human Dicer does not require ATP for dsRNA processing, but ATP is required for siRNA production in Drosophila embryo lysate and C. elegans extracts. Energy is probably required for dsRNA unwinding by Dicer’s helicase domain to move it along or for product release (Zhang et al., 2002, Ketting et al., 2001, Nykanen et al., 2001). This hypothesis is supported by the requirement for functional DExH helicase domain for efficient dsRNA processing by Drosophila Dcr-2. Mutation in ATP binding site of the domain blocks this activity (Lee et al., 2004).

Giardia intestinalis Dicer comprises PAZ domain and tandem RNase III domains, but lacks the rest of the domains present in eukaryotic Dicers. Crystal structure of Giardia Dicer reveals that there are two metal ions located in the catalytic site of each RNase III domain, which implicates two-metal ion mechanism of catalysis in the hydrolysis of each strand of the dsRNA (Macrae et al., 2006). The PAZ domain of Argonaute proteins specifically recognizes a 3’ overhang in an siRNA (Ma et al., 2004, Lingel et al., 2004, Song et al., 2003, Yan et al., 2003). Superposition of PAZ domains of Giardia Dicer and hAgo1 reveals that two have the same overall fold and binding pocket
for 3’ overhang but *Giardia* PAZ domain has an extended loop, which is unique for Dicer sequences. The structure suggests the mechanism for measuring the length of the siRNA. There is exactly 25-nt of A-form RNA helix distance between the 3’ overhang binding pocket and the catalytic center of the enzyme. *Giardia* processes long dsRNA into ~25-nt siRNAs, suggesting PAZ domain might serve as a molecular ruler for Dicers to measure and cleave at certain distances from the end of the dsRNA. Some Dicers, for example DCL4 of *Arabidopsis*, Dcr1 of *S. pombe* or Dcr-2 of *Drosophila*, have highly diverged PAZ domains or do not have them at all. It would be of great interest to determine how the correct size of an siRNA is measured by these Dicers.

**Argonaute proteins**

Crystal structures and mutagenesis analysis of Argonaute proteins reveal the basis for their specific ability to bind and mediate cleavage of RNA. Highly conserved at the structural level PAZ domains are ~110 amino acids long and present only in Argonaute and Dicer proteins. The central domain of PAZ is a left-handed twisting barrel of 5 or 6 β-strands containing two α-helices at the N-terminus. It resembles an oligonucleotide-binding (OB) fold present in some nucleic acid binding proteins (Arcus, 2002). This region and a unique β-hairpin/α-helix subdomain found only in PAZ domains form a cleft lined with aromatic and basic residues required for RNA binding (Song et al., 2004, Yan et al., 2003, Lingel et al., 2003). Interestingly, PAZ domains do not rearrange upon nucleic acid binding. A 2-nt 3’ overhang is positioned deep inside the cleft, which, due to steric constrains, cannot accommodate a terminal phosphate or paired nucleotides (Ma et al., 2004, Lingel et al., 2004, Collins and Cheng, 2005). Outside the binding pocket, A-
form RNA makes a $90^\circ$ turn, and phosphate groups of the strand with the buried 3’ end form electrostatic interactions along the cleft.

Three structural domains of the full-length *Pyrococcus furiosus* Argonaute (PfAgo), N-terminal, middle and Piwi, form a crescent-shaped base (Song et al., 2004). PfAgo PAZ domain has no similarity at the nucleic sequence level with canonical PAZ domains. In the crystal structure, however, PAZ domain contains aromatic residues equivalent to those in eukaryotic proteins required for 2-nt 3´ overhang binding. The PAZ domain is positioned in the middle of the crescent above the PIWI domain.

*Achaeoglobus fulgidus* PIWI (AfPiwi) is smaller than PfAgo and structurally corresponds to the middle and PIWI domains of PfAgo (domains A and B in structure) (Parker et al., 2004). The structure of AfPiwi with 21-nt duplex siRNA reveals that PIWI domain contains preformed basic binding pocket for the 5´ phosphate of the guide strand of the duplex organized by the C-terminal carboxylate and a divalent cation, which is the most conserved structural feature of the PIWI domain (Ma et al., 2005). In accordance with biochemical data, the 5´ nucleotide of the guide strand of an siRNA does not contribute to mRNA binding. It is buried inside the binding pocket unwound from the 3´ end of the passenger strand allowing 2-nt overhang to fold back along the protein surface (Parker et al., 2005). The phosphate groups of nucleotides 2 to 5 of the guide strand form extensive interactions with AfPiwi and few contacts with the passenger strand, strongly indicating a conserved mode of the 5´ end of RNA binding within the Argonaute family. The remaining base-paired nucleotides assume an A-form helix, accommodated within a channel in the PIWI domain. Extension of the A-form helix positions the proposed mRNA cleavage site opposite the putative catalytic residues.
associated with the RNase H fold of domain B of the AfPiwi protein, suggesting that catalytic cleavage site is measured from the 5’ end of the guide strand of the siRNA. In the structure of Ma and colleagues, only nucleotides 2 to 5 make base-pairing interaction with the complimentary strand; the rest of the guide strand is disorganized, which can be explained by the necessity to dislodge 3’ end of the siRNA from the 3’ overhang binding pocket of PAZ domain in order to interact with the target (Ma et al., 2005). Partial unwinding of the siRNA duplex in AfPiwi−RNA complex indicates that binding to the PIWI domain might contribute to siRNA or miRNA/miRNA* duplex unwinding during RISC assembly and/or participate in strand selection.

The ability of Argonaute proteins to cleave is determined not only by the presence of conserved residues required for the binding of the 5’ end of the guide strand but also by catalytic residues (Ma et al., 2005, Parker et al., 2004, Liu et al., 2004). The Piwi domains of PfAgo and AfPiwi contain a fold similar to that of RNase H, an enzyme, which cleaves the RNA strand of DNA–RNA hybrids. RNase H and related enzymes, contain three highly conserved carboxylates, Asp-Asp-Glu (DDE motif) (Nowotny et al., 2005). DDE residues in RNase H coordinate two metal ions in the active site of the enzyme required for nucleophile activation and stabilization of the transition state of the catalysis. Positioning of the metal ions is determined by carboxylates of the enzyme and both the scissile phosphate and 2’ hydroxyl of the backbone of the substrate, so that catalysis can take place only when metal ions are separated by ideal distance and form appropriate geometry. In the catalytic center of PfAgo, metal ions are coordinated by two aspartates and a histidine, instead. If the catalytic residues in hAgo2 are mutated to those present in Argonaute proteins which do not possess cleavage activity, mRNA cleavage is
abolished (Rivas et al., 2005, Liu et al., 2004). It is not possible, however, to make non-cleaving Argonaute a cleaver by just introducing catalytic residues present in cleavage-competent homologs, suggesting that the extent of base-pairing interactions between small RNA and its target might either position or exclude target from the active site.

**Interconnection of silencing pathways**

A central role for Argonaute proteins in RNA silencing pathways was first appreciated with the discovery of mutations compromising RNAi and related phenomena, normal development and causing stem cell defects (Cox et al., 1998, Harris and Macdonald, 2001, Morel et al., 2002, Tabara et al., 1999). Multiple Argonaute paralogs present in every given organism (*S. pombe* is the exception, it has only one Argonaute) can be further subdivided into two subfamilies, Ago and Piwi.

RNAi and miRNA pathways can be completely separated from each other. For example, in *Drosophila*, RNAi pathway requires Dcr-2, R2D2 and Ago2, and miRNA pathway requires Dcr-1, Loqs and Ago1 (Fig. 2). Although both recombinant Ago1 and Ago2 direct siRNA-mediated target cleavage, Ago1 immunopurified from *Drosophila* S2 cells contains only miRNAs, not siRNAs (Miyoshi et al., 2005). This suggests the existence of a well-ordered programmed mechanism distinguishing an siRNA from a miRNA to be loaded into corresponding Argonaute protein. Indeed, genetic studies indicate that only Dcr-1 is required for miRNA production, and biochemical evidence suggests that Ago1, but not Ago2, is associated with Dcr-1/Loqs and required for stabilization of mature miRNAs (Saito et al., 2005, Okamura et al., 2004, Lee et al., 2004). Argonaute proteins regulate distinct biochemical pathways not only in *Drosophila*. 
In *C. elegans*, the Argonaute protein Rde-1 is required for RNAi but not for miRNA function, whereas the Argonaute proteins Alg-1 and Alg-2 are required for miRNA function but not RNAi. Separation of silencing pathways suggests that once a small RNA is loaded in either siRISC or miRISC, only the degree of complementarity between the small RNA and target mRNA determines the outcome of interaction between them. In *Drosophila* and human cells, both siRNAs and miRNAs direct target cleavage if perfectly complimentary to their targets and translationally repress them if they are not (Hutvagner and Zamore, 2002, Zeng et al., 2003, Doench et al., 2003). For example, in *Drosophila*, miR-2b, perfectly complimentary to its target, directs cleavage by Ago1, the Argonaute protein of miRNA pathway (Okamura et al., 2004, Miyoshi et al., 2005). Because *Drosophila* miRNAs are imperfectly complimentary to their identified targets, it is assumed that, similar to *C. elegans* and human miRNAs, they direct translational repression of mRNAs. This suggests Ago1 can direct both target cleavage and translational repression. Are other Argonaute proteins multifunctional? In mice and humans, miRNAs are associated with all four proteins of Ago subfamily of Argonaute proteins, but Ago2 is the only protein responsible for target cleavage by both miRNAs and siRNAs (Fig. 3) (Meister et al., 2004, Liu et al., 2004). If hAgo2 is tethered to the 3´ UTR of a reporter gene, it causes translational inhibition of the reporter mRNA without its degradation (Pillai et al., 2004, Pillai et al., 2005). Unfortunately, it is not clear in this case whether inhibition is directly caused by hAgo2 or by another Argonaute protein interacting with it.

In organisms with a single Dicer, there is a greater interconnection between miRNA and RNAi pathways. The type of silencing in this case seems to be determined mainly by
the identities of Argonaute proteins. When human cells are transfected with siRNAs, all Ago subfamily proteins associate equally well with both endogenous miRNAs and siRNAs (Liu et al., 2004). Since hAgo2 is the only cleaver, endogenous miRNAs bound to it likely regulate their targets by means of cleavage, and miRNAs bound to other Argonautes mediate translational repression of their targets. There is a precedent of the mouse miRNA196 directing cleavage of mRNA corresponding to the developmental regulator HoxB8 in embryos (Yekta et al., 2004). Ago2-deficient mice display severe developmental abnormalities; supporting the view that cleavage competency is required for miRNA-mediated regulation of at least some targets (Liu et al., 2004). Even in Drosophila, separation between the miRNA and siRNA pathways found in vitro seems not to be absolute in vivo. For example, Dcr-1 and its partner Loqs are required not only for miRNA biogenesis, but also for efficient silencing of the endogenous gene by long double-stranded hairpin in vivo, probably due to their role in siRISC assembly (Lee et al., 2004, Forstemann et al., 2005). Although Ago1 is not required for target cleavage by siRNAs in vitro, ago1 mutations impair the siRNA-triggered RNAi in Drosophila embryos and dsRNA-driven silencing of the endogenous gene in adult eyes (Williams and Rubin, 2002, Meyer et al., 2006). This suggests that Ago1 and Ago2 might have partially overlapping functions. In Drosophila S2 cells, Ago1 and Ago2 regulate partially overlapping subsets of miRNA targets (Rehwinkel et al., 2006). Ago2 complexes purified from Drosophila embryos contain mature miRNAs, and those from S2 cells contain both pre- and mature miRNAs, suggesting that both proteins might be components of the same complex (Rehwinkel et al., 2006, Meyer et al., 2006).
CHAPTER II

A SINGLE ARGONAUTE PROTEIN MEDIATES BOTH TRANSCRIPTIONAL AND POSTTRANSCRIPTIONAL SILENCING IN SCHIZOSACCHAROMYCES POMBE

Summary

The Schizosaccharomyces pombe genome encodes only one of each of the three major classes of proteins implicated in RNA silencing: Dicer (Dcr1), RNA-dependent RNA polymerase (RdRP; Rdp1), and Argonaute (Ago1). These three proteins are required for silencing at centromeres and for the initiation of transcriptionally silent heterochromatin at the mating-type locus. Here, we show that the introduction of a dsRNA hairpin corresponding to a GFP transgene triggers classical RNAi in S. pombe. That is, GFP silencing triggered by dsRNA reflects a change in the steady-state concentration of GFP mRNA, but not in the rate of GFP transcription. RNAi in S. pombe requires dcr1, rdp1, and ago1, but does not require chp1, tas3, or swi6, genes required for transcriptional silencing. Thus, the RNAi machinery in S. pombe can direct both transcriptional and post-transcriptional silencing using a single Dicer, RdRP, and Argonaute protein. Our findings suggest that these three proteins fulfill a common biochemical function in distinct siRNA-directed silencing pathways.
Introduction

In many eukaryotic cells, exogenous long double-stranded RNA (dsRNA) triggers the specific degradation of cellular mRNAs with corresponding sequences, a phenomenon termed RNA interference (RNAi) (Fire et al., 1998). The dsRNA is cleaved by the multi-domain ribonuclease III enzyme, Dicer, into a population of 21-27 nt dsRNAs termed small interfering RNAs (siRNAs). (Bernstein et al., 2001) siRNAs are the specificity determinants of the RNAi pathway (Hamilton and Baulcombe, 1999, Hammond et al., 2000, Zamore et al., 2000). siRNAs are assembled into a protein-RNA complex, the RNA induced silencing complex (RISC), which directs cleavage of the target RNA (Zamore et al., 2000, Hammond et al., 2000, Nykanen et al., 2001). Among the various protein factors required for RNAi, members of the Argonaute family of proteins are universally associated with siRNA-directed gene silencing (Cogoni and Macino, 1997, Tabara et al., 1999, Cerutti et al., 2000, Fagard et al., 2000, Catalanotto et al., 2002, Kennerdell et al., 2002, Mochizuki et al., 2002, Pal-Bhadra et al., 2002, Tabara et al., 2002, Tijsterman et al., 2002a, Williams and Rubin, 2002, Shi et al., 2004). Plants and animals contain multiple Argonaute paralogs; the Drosophila genome encodes at least five distinct Argonaute proteins; humans have eight; and worms, 27! The remarkable diversity of Argonaute proteins suggests that each Argonaute paralog plays a specific role in RNA silencing. The Ago2 protein is a core-component of biochemically purified RISC from both Drosophila and mammals (Hammond et al., 2001, Hutvagner and Zamore, 2002, Martinez et al., 2002). In fact, among the human Argonaute proteins, only Ago2 can mediate microRNA- (miRNA-) or siRNA-directed endonucleolytic cleavage of target RNA (Liu et al., 2004, Meister et al., 2004). The recent three-dimensional structure of an archeal Argonaute protein, together with experiments
evaluating the importance of predicted catalytic residues in human Ago2, suggests that human Ago2 itself is the small RNA-guided endoribonuclease that cleaves target RNA (Liu et al., 2004; Song et al., 2004). Other human Argonaute proteins may be specialized to direct translational repression of mRNA or transcriptional silencing of DNA sequences by the small RNA-directed production of silent heterochromatin. Such functional specialization might also extend to the animal miRNA pathway, in which small RNAs typically direct translational repression, but not destruction, of their target mRNAs (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). In *C. elegans*, the Argonaute protein Rde-1 is required for RNAi but not for miRNA function, whereas the Argonaute proteins Alg-1 and Alg-2 are required for miRNA function but not RNAi (Tabara et al., 1999; Grishok et al., 2001). In *Drosophila*, Ago2 function is restricted to the siRNA-directed PTGS pathway, whereas miRNA function is associated with Ago1 (Okamura et al., 2004). In *Arabidopsis thaliana*, AGO1 is required for miRNA function and PTGS, whereas AGO4 acts in a silencing pathway that targets chromatin, rather than mRNA (Fagard et al., 2000; Morel et al., 2002; Zilberman et al., 2003; Vaucheret et al., 2004).

In contrast to the multiplicity of Argonaute proteins in higher organisms, the sequenced fungal genomes appear to encode one (*Neurospora crassa* and *Schizosaccharomyces pombe*) or no Argonaute proteins (*Saccharomyces cerevisiae*). The *Neurospora* Argonaute protein QDE-2 is required for quelling, an RNAi-like phenomenon (Cogoni and Macino, 1997) and is a component of an siRNA-containing complex (Catalanotto et al., 2002). In contrast, Ago1 is required for the silencing of transcription at centromeres in the fission yeast *S. pombe* (Volpe et al., 2002) and initiation of silent heterochromatin at the mating type locus (Hall et al., 2002). *S. pombe*
Ago1 is a component of the RNA-induced Initiation of Transcriptional Silencing (RITS) complex (Verdel et al., 2004). In addition to Ago1, the RITS contains siRNAs derived from centromeric sequences, as well as the chromodomain protein Chp1 and the protein Tas3, whose function is unknown.

Current evidence suggests that the RITS uses the sequence information in its siRNA component to direct the methylation of histone H3 at lysine 9 (K9) bound to centromeric DNA (Verdel et al., 2004). Histone H3 K9 methylation, in turn, triggers the formation of transcriptionally silent heterochromatin, a process dependent on the histone methyltransferase Clr4 and a second chromodomain protein, Swi6 (Elgin and Grewal, 2003). The initiation of silent heterochromatin requires not only the RITS complex, but also Dcr1, the S. pombe homolog of Dicer, and a putative RNA-dependent RNA polymerase, Rdp1 (Hall et al., 2002). By analogy to silencing in higher organisms, Dcr1 is presumed to generate siRNA from dsRNA derived from centromeric DNA. Rdp1 has been postulated either to generate this dsRNA itself or to amplify the production of siRNA, but the precise function of RNA-dependent RNA polymerases in siRNA-directed silencing is not yet understood.

In S. pombe, as in other eukaryotes, silencing can be triggered by the introduction of exogenous long dsRNA transcribed from transgenic DNA. Such dsRNA can silence both endogenous and exogenous genes, although it has not been determined if this silencing is transcriptional—like the endogenous silencing of centromeric DNA—or post-transcriptional, as is observed in the RNAi pathway in animals and plants (Raponi and Arndt, 2003). Grewal, Martienssen and colleagues have shown that the production of heterochromatin at the S. pombe centromere contains features characteristic of both transcriptional and post-transcriptional silencing (Volpe et al., 2002). Forward
transcription at this locus appears to be repressed by the formation of heterochromatin, yet reverse transcription through the same locus is not, suggesting that reverse transcripts are degraded post-transcriptionally. To address the question if PTGS exists in *S. pombe* in the absence of TGS, we developed a simplified silencing system based on green fluorescent protein (GFP) expression. Here, we show that the introduction of a dsRNA hairpin corresponding to a GFP transgene triggers classical RNAi in *S. pombe*: introduction of GFP dsRNA causes a change in the steady-state concentration of GFP mRNA, but not the rate of GFP transcription. RNAi in *S. pombe* requires Ago1, Dcr1, and Rdp1, but does not require Chp1, Tas3, or Swi6, which are required for transcriptional silencing. Our data suggest that Argonaute, Dicer, and RdRP play common biochemical roles in functionally distinct silencing pathways.
Results

To test if dsRNA triggers RNAi in *S. pombe*, we engineered a strain containing the enhanced GFP (gfp) protein coding sequence fused in-frame to the endogenous adh1 gene. PCR analysis (data not shown) verified that the strain contained the adh1:gfp fusion gene integrated at the adh1 locus on chromosome 3. The fusion mRNA encoded by this locus was the only source of adh1 mRNA in the strain. The strain also contained the kanamycin resistance gene (aph, encoding the enzyme, aminoglycoside-3´-phosphotransferase) downstream from adh1:gfp, with 628 bp separating the start of aph open reading frame from the end of the gfp open reading frame.

To trigger silencing of adh1:gfp, we engineered a plasmid-borne GFP hairpin (Fig. 1). The 760 bp gfp open reading frame (ORF) was cloned as an inverted repeat, with the sense and antisense arms of the repeat separated by a 67-base pairs (bp) spacer containing the first intron of the rad9 gene. The intron was included, because intron-containing hairpin RNAs induce PTGS in plants with nearly 100% efficiency (Smith et al., 2000). The construct used here, when spliced, is presumed to leave a loop of 14 unpaired nucleotides (nt). Transcription of the GFP hairpin was under the control of the thiamine-repressible nmt1 promoter; in the absence of thiamine, nmt1 is among the strongest promoters in *S. pombe* (Moreno et al., 1991). The plasmid also contained the ura4 gene, to permit selection for retention of the plasmid in the absence of uracil. (The reporter strain contains an ura4 mutation and thus cannot synthesize uracil.) Expression of the GFP hairpin in the ura4− strain had no effect on cell morphology or growth rate (data not shown).

When the plasmid encoding the GFP hairpin was introduced into the adh1:gfp target strain, the GFP fluorescence intensity was reduced by more than 2-fold (adh1:gfp +
**Figure 1.** Experimental strategy. The coding sequence of GFP was fused in frame with the *adh1* locus on chromosome 3. The kanamycin resistance gene (*aph*) was inserted adjacent to the *adh1:gfp* transgene as an insertion marker under the control of the *Ashbya gossypii* translation elongation factor 1α gene promoter. Red arrows indicate the position of *adh1:gfp*-specific primers used in the quantitative RT-PCR assays in Fig. 3. The silencing trigger was expressed from an episomal plasmid encoding a hairpin transcript corresponding to GFP expressed from the *nmt1* promoter. The GFP dsRNA hairpin contained within the loop sequences the 67-bp intron 1 from *rad9* to facilitate hairpin expression.
hairpin; Figs. 2 and 5). Silencing was observed only in the presence of the GFP hairpin plasmid (Supplementary Fig. 1). No silencing was observed in a adh1:gfp strain transformed with the same plasmid lacking the hairpin (Fig. 2A, empty vector), nor was silencing triggered by the same plasmid expressing only sense or antisense GFP transcript (Fig. 2B). Furthermore, when the silenced strain was grown in the presence of uracil, the plasmid expressing the GFP hairpin was lost (adh:gfp – hairpin), as expected, and GFP fluorescence was restored to that of the original adh1:gfp strain (Fig. 2A). The loss of silencing in non-selective conditions argues against epigenetic (i.e., heritable) GFP silencing. Thus, silencing in our GFP/GFP hairpin system is distinct from that observed previously at both the centromeres and the mating-type locus, where dsRNA is proposed to trigger assembly of epigenetically heritable, repressive chromatin structures (Volpe et al., 2002).

We could imagine two possible mechanisms for the silencing of adh1:gfp by the GFP hairpin. Silencing might reflect either unstable transcriptional repression of the locus or bona fide posttranscriptional gene silencing, i.e., RNAi. To distinguish between these possibilities, we measured both the steady-state level (Fig. 3) and the rate of nuclear transcription (Fig. 4) of the adh1:gfp mRNA. To measure steady-state mRNA abundance, we performed quantitative RT-PCR using primers that spanned the fusion site in the adh1:gfp transgene (Fig. 1); actin (act1) mRNAs levels were measured as a control. In the presence of the GFP hairpin, the level of gfp mRNA was reduced more than two-fold. The steady-state level of aph mRNA was unchanged, demonstrating the specificity of the silencing triggered by the GFP hairpin.

In contrast to its effect on steady-state mRNA levels, the GFP hairpin caused no observable change in the transcription of the adh1:gfp mRNA. Transcription was
**CHAPTER II, Figure 2**

**A**

- adh1:gfp - hairpin
- adh1:gfp + hairpin
- adh1:gfp + empty vector
- adh1:gfp only
- hairpin only

**log fluorescence intensity**

**number of cells**

**B**

- adh1:gfp + sense
- adh1:gfp + antisense
- adh1:gfp + hairpin
- adh1:gfp only
**Figure 2.** The GFP hairpin triggers *adh1::gfp* silencing. (A) Representative FACS data demonstrating that GFP silencing occurs only in the presence of the GFP hairpin. Silencing of GFP expression, measured by fluorescence intensity, was not observed for the empty vector alone, and silencing was lost when the silenced strain was grown in the presence of uracil to cause loss of the hairpin-encoding plasmid. Ten thousand cells were analyzed for each genotype. See also Fig. 5. (B) Representative FACS data demonstrating that GFP silencing is triggered by the GFP hairpin but not by either sense or antisense GFP transcripts.
assessed by the nuclear run-on method. Briefly, yeast cells were permeabilized with detergent to permit the introduction of α-32P-UTP to label transcripts from elongating RNA polymerase II (RNA Pol II) (Volpe et al., 2002). This technique measures the density of RNA Pol II on a gene at the time of lysis, a reflection of the transcriptional rate of the gene in vivo. To distinguish transcription of the adh1:gfp mRNA from that of the GFP hairpin, we measured the rate of transcription for the adh1:gfp fusion mRNA using a probe for adh1. The rate of transcription of the act1 locus provided a normalization control. The GFP hairpin did not detectably decrease the rate of transcription of adh1:gfp (Fig. 4B). Control experiments demonstrate that the nuclear run-on assay can detect changes in the rate of transcription at least as small as 25% (Supplementary Fig. 2 and 3). Thus, the nuclear run-on data suggest that the reduction of steady-state level of gfp mRNA triggered by the hairpin was posttranscriptional.

To test our hypothesis by an independent method, we performed quantitative chromatin immunoprecipitation (ChIP) using an antibody to RNA Pol II. Like the nuclear run-on method, this method measures the relative density of RNA Pol II on a DNA sequence, a reflection of its transcriptional rate. Again, we could detect no decrease in the association of RNA Pol II with adh1:gfp in the presence of the GFP hairpin, relative to its association in the absence of the hairpin or in the presence of the plasmid lacking the hairpin sequences (empty vector) (Fig. 4C). We conclude that silencing of the adh1:gfp locus by the GFP hairpin occurs posttranscriptionally.

Silencing at the centromere locus can be accompanied by lysine 9 methylation of histone 3 (H3 K9) and the spread of silencing across at least 3 kb of adjacent noncentromeric sequence; the spread of silencing correlates with the coating of the adjacent DNA with the Swi6 protein (Partridge et al., 2000). In contrast, silencing of
CHAPTER II, Figure 3

The figure shows a bar graph comparing the relative mRNA steady-state level of GFP hairpin and empty vector under different conditions.

- **GFP hairpin**: The graph indicates a significant increase in mRNA level (+) for the mutant compared to the control (−).
- **Empty vector**: The mRNA level is slightly increased (+) compared to the control (−).

The y-axis represents the relative mRNA steady-state level, ranging from 0 to 1.4. The x-axis categorizes the experimental conditions as either 'adh1:gfp' or 'aph'.
**Figure 3.** GFP silencing reflects a reduction in the steady-state level of *adh1:gfp* mRNA. The steady-state level of mRNA was measured for *gfp* and *aph* relative to the expression level of the *act1* mRNA. For each strain three independent cultures were assayed; each assay was analyzed by quantitative RT-PCR in triplicate. The data are presented as the average of the three independent trials (using the average value for the triplicate quantitative RT-PCR reactions) ± the standard deviation of the three independent trials.
adh1:gfp by the GFP hairpin was not accompanied by silencing of the adjacent aph locus (Fig. 4), whose promoter is less than 600 bp downstream from the end of the gfp open reading frame (Fig. 1). Introduction of GFP hairpin did not increase H3 K9 methylation levels at the adh1:gfp locus (Supplementary Fig. 4A). Thus, in every respect, silencing of adh1:gfp by the GFP hairpin appears to be entirely posttranscriptional. We conclude that in S. pombe, as in animals, plants, protozoa, and basal eukaryotes, dsRNA triggers RNAi.

H3 K9 methylation levels were slightly reduced at the control cenromere locus (Supplementary Fig. 4B), which suggests that a limiting component, possibly Clr4, is recruited from the locus in the presence of GFP hairpin.

Dicer orthologs are required for posttranscriptional gene silencing in Arabidopsis thaliana, Drosophila, C. elegans, Dictyostelium discoideum, and Neurospora (Bernstein et al., 2001, Knight and Bass, 2001, Martens et al., 2002, Catalanotto et al., 2004). Similarly, genes encoding putative RdRP enzymes are required for PTGS in plants, worms, Neurospora, and Dictyostelium (Cogoni and Macino, 1999, Dalmay et al., 2000, Smardon et al., 2000, Martens et al., 2002). The S. pombe dcr1 and rdp1 genes are required for the initiation of silent heterochromatin at centromeres and the mating type locus (Hall et al., 2002, Volpe et al., 2002). In fact, the Rdpl protein is physically associated with centromeric repeats, suggesting it plays a direct role in generating the dsRNA silencing trigger (Volpe et al., 2002).

Small RNAs corresponding to GFP sequences accumulate in strains containing the GFP hairpin, but not the empty vector (Supplementary Fig. 5). No small RNAs were detected in either the dcr1Δ strain (Supplementary Fig. 5 and 6) or the rdp1Δ strain (Supplementary Fig. 6). Small RNA production required only the GFP hairpin,
CHAPTER II, Figure 4

A

\[ Pp-luc \quad aph \quad adh1 \quad act1 \]

- hairpin only
- \(adh1:gfp\) only
- \(adh1:gfp\) + empty vector
- \(adh1:gfp\) + hairpin

B

<table>
<thead>
<tr>
<th></th>
<th>Relative Transcriptional Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>hairpin only</td>
<td>0</td>
</tr>
<tr>
<td>(adh1:gfp) only</td>
<td>1.0</td>
</tr>
<tr>
<td>(adh1:gfp) + empty vector</td>
<td>0.5</td>
</tr>
<tr>
<td>(adh1:gfp) + hairpin</td>
<td>1.0</td>
</tr>
</tbody>
</table>

C

<table>
<thead>
<tr>
<th></th>
<th>Relative RNA Pol II-bound DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(adh1:gfp) only</td>
<td>0</td>
</tr>
<tr>
<td>(adh1:gfp) + empty vector</td>
<td>1.0</td>
</tr>
<tr>
<td>(adh1:gfp) + hairpin</td>
<td>1.5</td>
</tr>
</tbody>
</table>
**Figure 4.** GFP silencing triggered by the GFP hairpin is posttranscriptional. (A) Representative nuclear run-on experiment. (B) The average value ± standard deviation for three independent run-on experiments is presented. For each strain, the rates of transcription for both *aph* and *adh1* were measured and normalized to the rate of transcription of *act1*. Firefly luciferase (*Pp* luc) served as a negative control. No decrease in transcription of *adh1:gfp* was observed when GFP was silenced by introduction of the GFP hairpin-expressing plasmid. The rate of *aph* transcription was essentially constant in all *aph*-containing genotypes. (C) RNA Pol II chromatin immunoprecipitation (ChIP). The density of RNA Pol II at the *adh1:gfp* locus was not decreased by the introduction of the GFP hairpin, reinforcing the view that silencing triggered by this hairpin is post-transcriptional. Data are the average of two independent trials. The data for *adh1:gfp* and those for *aph* were separately normalized to their respective values in the absence of the empty or hairpin vector.
not the \textit{adh1:gfp} target RNA, suggesting that the RdRP plays a direct role in the production of the small RNAs from the hairpin trigger itself or their subsequent stability, rather than in the generation of secondary siRNAs templated from the target RNA. Small RNAs could be detected in the \textit{ago1Δ} strain (Supplementary Fig. 6), consistent with a role for Ago1 in the function, but not the production, of small RNAs.

Next, we measured GFP fluorescence in the \textit{dcr1Δ, rdp1Δ,} and \textit{ago1Δ} strains containing both the \textit{adh1:gfp} locus and the GFP hairpin. GFP was not silenced in the absence of Dcr1 or Rdp1 (Fig. 5), consistent with the requirement for these enzymes in small RNA production. Thus, both Dicer and RdRP are required for posttranscriptional silencing of GFP in \textit{S. pombe}, just as they are in \textit{C. elegans}, the classical model for RNAi. Unlike plants and \textit{Drosophila}, \textit{S. pombe} encodes a single \textit{dicer} gene. Our data, together with previous studies of silent heterochromatin in \textit{S. pombe}, demonstrate that a single Dicer protein can support both TGS and PTGS. The functional specializations of Dicer proteins in plants and flies are unlikely to reflect an inherent biochemical limitation of the ancestral Dicer protein (Xie et al., 2004, Lee et al., 2004). Unlike \textit{C. elegans}, which contains four RdRP genes, \textit{S. pombe} encodes a single RdRP homolog, demonstrating that a single RdRP protein can mediate both transcriptional and posttranscriptional gene silencing. The simplest explanation is that the \textit{S. pombe} RdRP protein supplies a common biochemical function to both pathways. In \textit{C. elegans} both somatic and germline RNAi triggered by dsRNA requires an RdRP (Smardon et al., 2000). The finding that in \textit{S. pombe} RNAi triggered by dsRNA requires a functional RdRP suggests the mechanism of \textit{S. pombe} silencing is more closely related to RNAi in worms than PTGS in plants, where the RdRP homolog SDE1 is required only for silencing triggered by sense RNA-producing transgenes, not dsRNA (Dalmay et al.,
relative fluorescence Intensity

adh1:gfp only
adh1:gfp + empty vector
adh1:gfp + GFP hairpin

wild type
mutant

CHAPTER II, Figure 5
**Figure 5.** GFP silencing depends on the RNAi machinery but not the transcriptional silencing proteins Chp1, Tas3, or Swi6. The geometric mean of fluorescence intensity was determined for each strain from experiments like that in Fig. 2, and normalized to the geometric mean for the wild-type *adh1:GFP* strain. The data are the average ± standard deviation for three trials.
We note that it remains possible that one of the RdRP genes distinct from SDE1 could be required for dsRNA-triggered silencing in plants.

In contrast to the Dicer and RdRP genes, Argonaute proteins form a highly ramified family in all higher eukaryotes whose genomes have been fully sequenced. For example, the *C. elegans* genome encodes 27 Argonaute proteins. Current evidence suggests that Argonaute proteins are functionally specialized in higher organisms. For example, the *C. elegans* Argonaute protein, Rde-1, is required for RNAi, but not for miRNA function. Conversely, the Argonaute proteins Alg-1/Alg-2 are required for miRNA function in worms. In contrast, miRNAs associate with at least four distinct Argonaute proteins in humans, only one of which, Ago2, can direct target mRNA cleavage (Liu et al., 2004, Meister et al., 2004). The *Drosophila* genome encodes at least five Argonaute proteins. In *Drosophila*, miRNAs associate with the Argonaute protein Ago1, whereas siRNAs associate with Ago2; both Ago1 and Ago2 can direct small RNA-guided target RNA cleavage (Okamura et al., 2004). Both the *Drosophila* Argonaute protein-encoding genes, *piwi* and *aubergine (aub)* are required for H3 Lys9 methylation and correct localization of two heterochromatic proteins, HP1 and HP2, all processes associated with TGS (Pal-Bhadra et al., 2004). Yet, mutations in *aub* abrogate the RNAi-based silencing of the *Stellate* locus, RNAi triggered by exogenous dsRNA, and RISC assembly (Aravin et al., 2001, Kennerdell et al., 2002, Tomari et al., 2004a). Furthermore, mutations in *piwi* partially block PTGS of *Adh* transgenes (Pal-Bhadra et al., 2002). Because Argonaute proteins have been implicated both in the production of the siRNA silencing trigger and the execution of target RNA, a partial resolution to this apparent paradox would be if Argonaute proteins were required at multiple steps in both the TGS and RNAi pathways (Tabara et al., 2002, Hammond et al., 2001, Hutvagner and
In this view, discrete Argonaute proteins would be biochemically specialized for functions common to TGS and PTGS, and others dedicated to functions unique to each process. Alternatively, the diversity of Argonaute proteins might simply reflect their specialized patterns of spatial or temporal expression or their intracellular localization.

The genome of *S. pombe* encodes a single Argonaute protein, Ago1, which is required for transcriptional silencing (Hall et al., 2002, Volpe et al., 2002). Is it also required for RNAi in *S. pombe*? We introduced the GFP hairpin into an ago1Δ strain bearing the *adh1:gfp* transgene. Posttranscriptional silencing of *adh1:gfp* by the GFP hairpin required ago1 (Fig. 5). Thus, a single Argonaute protein mediates both RNAi and TGS is *S. pombe*.

Ago1, Rdpl and Dcr1 are all required for both TGS and PTGS in *S. pombe*. Does a common complex mediate these two distinct silencing pathways? Ago1 is a component of the RITS complex, which also contains Chp1, Tas3, and siRNA (Verdel et al., 2004). To test if the RITS complex mediates RNAi in *S. pombe*, we asked if either Chp1 or Tas3 is required for silencing of the *adh1:gfp* transgene by the GFP hairpin. We also asked if the Swi6 or Clr4 proteins, which are not components of the RITS, but are required for RITS-initiated production of silent heterochromatin, are required for posttranscriptional silencing of the *adh1:gfp* transgene by the GFP hairpin. Deletion of *chp1*, *tas3*, or *swi6* had no effect on GFP silencing by the GFP hairpin (Fig. 5). Thus, the requirements for TGS and PTGS in *S. pombe* are genetically distinct. In contrast, deletion of *clr4* had an effect on posttranscriptional silencing of the *adh:gfp* locus. This finding is consistent with the decrease of H3 K9 methylation at the centromeric locus (Supplementary Fig. 4B) and a previous report that siRNAs do not accumulate to normal levels in the *clr4*
mutant strain (Motamedi et al., 2004); it is possible that Clr4 functions in siRNA biogenesis in both TGS and RNAi. That RNAi in *S. pombe* does not require Chp1, Tas3, or Swi6 suggests that a complex distinct from RITS mediates siRNA-directed target mRNA degradation. Alternatively, once siRNAs are produced by the action of Dcr1 and Rdpl, only Ago1 itself may be required for PTGS in *S. pombe*. 
Discussion

In this study, we have demonstrated that a dsRNA derived from a hairpin transcript can trigger posttranscriptional silencing of a corresponding mRNA in *S. pombe*. Schramke and Allshire previously demonstrated that a similar hairpin transcript, corresponding to the *ura4* locus could trigger transcriptional silencing (Schramke and Allshire, 2003). In both studies, silencing triggered by a hairpin transcript required the RNAi machinery—Dcr1, Rdp1, and Ago1. Transcriptional silencing, unlike post-transcriptional silencing, required components of the transcriptional silencing apparatus—Chp1, Tas3, or Swi6. Robust silencing by both pathways requires the chromodomain protein Clr4, which appears to play a role in siRNA biogenesis or stability. Why does the GFP hairpin construct presented here trigger exclusively post-transcriptional silencing, whereas the previously studied *ura4* hairpin triggered transcriptional silencing? One possible explanation is that the GFP hairpin used here included an efficiently spliced intron between the two arms of the hairpin. We presume that splicing of the intron promotes the accumulation of GFP dsRNA in the cytoplasm. In contrast, the *ura4* hairpin construct of Schramke and Allshire contained an unspliced spacer sequence between the hairpin arms. Thus, the *ura4* hairpin may be localized largely to the nucleus. A difference in subcellular localization might explain the different results obtained by the two studies. Alternatively, silencing of *ura4* by the *ura4*-specific hairpin might comprise a mixture of transcriptional and post-transcriptional silencing. In this case, transcriptional silencing might not occur at the *adh1* locus, even if the GFP hairpin-derived siRNAs trigger histone modification, perhaps because the gene is strongly expressed or is in a region of the genome otherwise refractory to

1 Retracted (Allshire, 2005)
heterochromatin formation. Nonetheless, our data, together with those of Schramke and Allshire, clearly show that at least two distinct silencing responses can be initiated by a common RNAi machinery, without resorting to specialized forms of Dicer, RdRP, or Argonaute proteins. The demonstration that fission yeast contain a functional RNAi pathway now provides a simplified, genetically tractable model in which to study how the nature of the silencing trigger or of the silencing target determines the silencing pathway evoked—posttranscriptional or transcriptional.
<table>
<thead>
<tr>
<th>Sample</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNase</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>adh1:gfp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>hairpin</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>vector</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>antisense gfp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>sense gfp</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

- **full-length antisense act1 probe**
- **protected antisense act1 probe**
- **full-length sense gfp probe**
- **protected sense gfp probe**
Supplementary Figure 1. Expression of GFP hairpin measured by RNase protection assay. Antisense act1 and sense gfp probes were used simultaneously to detect act1 mRNA and antisense region in the GFP hairpin. 1- probes only, no RNA, no RNases; 2 to 10- RNases were added; 2- probes plus Saccharomyces cerevisiae total RNA (no target RNA); 3- probes plus wild-type S. pombe total RNA; 4- probes plus total RNA isolated from wild-type strain of S. pombe expressing the GFP hairpin; 5- probes plus total RNA isolated from S. pombe strain containing adh1:gfp transgene; 6- probes plus total RNA isolated from S. pombe strain containing adh1:gfp transgene and empty vector; 7- probes plus total RNA isolated from S. pombe strain containing adh1:gfp transgene and the GFP hairpin; 8- probes plus total RNA isolated from S. pombe strain containing adh1:gfp transgene and antisense strand of the GFP hairpin; 9- probes plus total RNA from S. pombe strain containing adh1:gfp transgene and sense strand of the GFP hairpin; 10- probes plus wild-type S. pombe total RNA mixed with exogenous antisense gfp and sense act1 transcripts.
CHAPTER II, Supplementary Figure 2

A

<table>
<thead>
<tr>
<th>nmt1</th>
<th>Pp Luc</th>
<th>act1</th>
<th>adh1</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ thiamine</td>
<td>+ thiamine</td>
<td>+ thiamine</td>
<td>+ thiamine</td>
</tr>
<tr>
<td>- thiamine</td>
<td>- thiamine</td>
<td>- thiamine</td>
<td>- thiamine</td>
</tr>
</tbody>
</table>

B

Graph showing relative transcription for nmt1 and adh1 under + thiamine and - thiamine conditions.
**Supplementary Figure 2.** Nuclear run-on analysis reflects changes in transcriptional rates. Wild-type cells were grown to mid-log phase in minimal media or minimal media supplemented with 2 μM thiamine. After nuclear run-on, RNA was isolated and hybridized to membranes containing RNA probes for firefly luciferase, *nmt1*, *aph*, and *act1*. (A) Representative nuclear run-on experiment. (B) Quantification of the results in A. The rate of transcription for *nmt1* and *adhl* was normalized to the rate of transcription for *act1*. 
Relative transcriptional rate vs. relative concentration of apc nuclear transcripts.
Supplementary Figure 3. The nuclear run-on assay is linear and can readily detect a 25% change in transcriptional rates. The linearity of the nuclear run-on assay was confirmed by a serial dilution analysis. $^{32}$P-radiolabeled RNA purified from cells expressing aph (yFS313) was mixed with RNA of the same specific activity purified from wild type cells that lack the aph gene (yFS106). RNA was mixed at 3:1, 1:1, and 1:3 ratios, then hybridized to membranes containing RNA probes for firefly luciferase, aph, and act1. The rate of transcription for aph was normalized to the rate of transcription for act1 and reported relative to that of the undiluted RNA probe purified from the undiluted aph-expressing cells.
relative H3 K9 methylation levels

A

- adh1:gfp locus

vector only + -
hairpin - +

B

centromeric locus

vector only + -
hairpin - +
**Supplementary Figure 4.** Chromatin immunoprecipitation analysis of histone 3 lysine 9 (H3 K9) methylation. The density of methylated H3 K9 at the *adh1:gfp* locus was not increased by the introduction of the GFP hairpin, reinforcing the view that silencing triggered by this hairpin is posttranscriptional. Data are the average of three independent trials. The data for (A) *adh1:gfp* and those for (B) centromeric locus were normalized separately to their respective values in the absence of hairpin vector. Absolute level of H3 K9 methylation in the presence of empty vector is by a factor of ~130 lower at the *adh1:gfp* locus then at the centromeric locus.
CHAPTER II, Supplementary Figure 5

Image of a gel showing the expression of small RNAs in wild-type and dcr1Δ strains. The gel is labeled with the expression of adh1:gfp and indicates bands at 30, 25, 23, 21, and 19. There is a loading control at the bottom. The gel shows an empty vector and high-throughput (hp) conditions.
**Supplementary Figure 5.** Detection of GFP-specific small RNAs by Northern hybridization in wild-type and *dcr1Δ* strains. Ø, no episomal plasmid; hp, GFP hairpin-expressing plasmid. The corresponding ethidium bromide-stained loading control is shown in the lower panel. The small increase in small RNAs in the presence of both the GFP hairpin and the *adh1:gfp* target, relative to the hairpin alone, can be explained by a small difference in the amount of total RNA loaded in the two lanes. The electrophoretic mobility of 5'-32P-radiolabeled RNA size markers (nts) is indicated at left.
Supplementary Figure 6. Detection of GFP-specific small RNAs by Northern hybridization in wild-type, ago1Δ, dcr1Δ, and rdp1Δ strains. Ø, no episomal plasmid; hp, GFP hairpin-expressing plasmid. The corresponding ethidium bromide-stained loading control is shown in the lower panel. The electrophoretic mobility of 5’-32P-radiolabeled RNA size markers (nts) is indicated at left.
Materials and Methods

*S. pombe* strain construction

Fission yeast were grown and manipulated as described previously (Moreno et al., 1991). Unless otherwise stated, all strains (Table 1) were grown at 30°C in EMM2 media supplemented with histidine, leucine, adenine, and uracil as appropriate. Deletion and fusion strains were constructed by PCR-based cassette mutagenesis as described previously using the following oligonucleotides: 5’-CCT TCT CTA CTC TTC CCG ACG TCT ACC GTC TCA TGC ATG AGA ACA AGA TTG CCG GCC GTA TCG TCT TGG ACC TTT CCA AGC GGA TCC CCG GGT TAA TTA A-3’ and 5’-AAA GAT TCA TGA GAT GAA CAG ATT TGA AAC CAA GCT GCA ACC AAG CAC ACA AAC TGA GAT TGA CAA CGG CCG TAC CAT CGG AAT TCG AGC TCG TTT AAA C-3’ for *adh1::gfp*; 5’-GTT TGG TAT ATA TAA GCT TCC AAC CGC CAA AGC GAA TTG TCT TCA GCC AAC TCG TCC TTT ATG ATT CAG AGT GAG TAG GGC GCG CCA CTT CTA AAT AAG C-3’ and 5’-TAA GGA AGT AAA AGT TGT GGG CAA TCC AGT AGT CAA TCG TAT ATC TAT TTC ATT AC T TAT TGC ATG CAA TCC ATC AAA CAG AAT TCG AGC TCG TTT AAA C-3’ for *ago1::kanMX*; 5’-TTA TTT TTA CAA AAG GTT CCA AAT TGA TTT GTT GTC AGG GTT ATA ATT AAA TCC AGA TTT TGT ATG GCA ATG TAA AGC AGG CGC GCC ACT TCT AAA TAA G-3’ and 5’-TTC TAT GTT CCA TTG TCG ATA ATG AGT ACA ATT ACT ATT AGC CTT TTA CGA GCA ATT AAA GCT TCG AAT GTA TAT TCG ATG AAT TCG AGC TCG TTT AAA C-3’ for *rdp1::kanMX*; 5’-TAG CTT AGG ATT CAT TAT TTT TTA AGA GAC AAA TTT CTC GTC AAT TGA ATG AAA CCT TCC GCC TTT ATT TTC TTT TTG AGG CGC GCC ACT TCT AAA TAA G-3’ and 5’-GCT TTG GAG ACC CAA ATT GAA AGT TTG AAA AGT TAC AAG
CHAPTER II, Table 1. Strains used in the study.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>yFS106</td>
<td>\textit{leu1-32 h}^+ \textit{ura4-294}</td>
</tr>
<tr>
<td>yFS312</td>
<td>\textit{leu1-32 h}^+ \textit{ura4-294 pAS1}</td>
</tr>
<tr>
<td>yFS313</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-D18 his7-366 ade6-210 adh1:gfp}</td>
</tr>
<tr>
<td>yFS314</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-D18 his7-366 ade6-210 adh1:gfp pAS1}</td>
</tr>
<tr>
<td>yFS315</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-? ade6-210 adh1:gfp pRIP2}</td>
</tr>
<tr>
<td>yFS316</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-? ade6-210 ago1::kanMX adh1:gfp pAS1}</td>
</tr>
<tr>
<td>yFS317</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-? ade6-210 rdp1::kanMX adh1:gfp pAS1}</td>
</tr>
<tr>
<td>yFS318</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-? dcr1::kanMX adh1:gfp pAS1}</td>
</tr>
<tr>
<td>yFS319</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-D18 his7-366 ade6-210 chp1::his7 adh1:gfp pAS1}</td>
</tr>
<tr>
<td>yFS327</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-D18 ade6-210 tas3::kanMX adh1:gfp pAS1}</td>
</tr>
<tr>
<td>yFS328</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-? his7-366 ade6-210 arg3D swi6::arg3 adh1:gfp pAS1}</td>
</tr>
<tr>
<td>yFS329</td>
<td>\textit{h}^+ \textit{leu1-32 ura4-? his7-366 ade6-210 clr4::LEU2 adh1:gfp pAS1}</td>
</tr>
</tbody>
</table>

A question mark indicates it is not known if the strain contains either the \textit{ura4-294} or \textit{ura4-D18} allele of \textit{ura4}. 
Construction of the silencing trigger

Full-length eGFP sequence was amplified using the following primers: 5′-GGG
AAT TCC ATA TGG AGT AAA GGA GAA GAA CTT TTC-3′, 5′-CAC TAG CTA
GCCT CTAT CTT TAC CAA TTA GTT TCA ATG TTT AGT AAG GTT TGA AAA
AAG TTC CAA CAC ACC TGA TTT TAT TTG TAT AGT TCA TCC ATG CCA TGT
G-3′, 5′-GGC GCG GAT CCG AGT AAA GGA GAA CTT TTC-3′, and 5′-CAC
TAG CTA GCT ATT TGT ATA GTT CAT CCA TGC CAT GTG-3′. PCR fragments
were directly subcloned into pCR4-TOPO (Invitrogen) to produce pCR4-TOPO-B and
pCR4-TOPO-N constructs. Both constructs were digested with Nhe I and Not I
restriction enzymes, restriction fragments gel purified, and ligated to produce pCR4-
TOPO-B+N. This construct was digested with Nde I and BamHI and the insert
subcloned into the corresponding restriction sites of pRIP2 to produce pAS1 (Maundrell,
1993). A second plasmid bearing an ars1 sequence was also prepared by subcloning the
EcoRI fragment of pREP2 into corresponding site in pAS1 to produce pAS2. When
transformed, both pAS1 and pAS2 produced episomal transformants that gave similar
results in our experiments. Results for the pAS1 construct only are presented here.

RNA analysis

Cells were harvested from liquid cultures at mid-expontial phase (OD600 = 0.1-
0.4). Total RNA was extracted from cells using the hot phenol method (Lyne et al.,
2003). Purified RNA was treated with DNase (RQ1, Promega) and analyzed by
Quantitative RT-PCR in a DNA Engine OPTICON2 (MJ Research) using the QuantiTect SYBR Green PCR Kit (Qiagen) according to manufacture’s instruction. Analysis was performed using Opticon Monitor (MJ Research), Excel (Microsoft), and IgorPro 5.0 (Wavemetrics) software. PCR primers were 5´-GAT TGC CGG CCG TAT CGT CTT-3´ and 5´-GCC CAT TAA CAT CAC CAT CTA-3´ for adh1:gfp; 5´-CTG AAA CAT GGC AAA GGT AGC-3 and 5´-GGG ATC GCA GTG GTG AGT AAC-3´ for aph; 5´-ACT TTG CTA CGT CGC TTT GGA C-3´ and 5´-CGT TTC CGA TAG TGA TAA CTT G-3´ for act1; 5´-TTTGTGGGATAAACAAGTTCTCAA-3´ and 5´-CGCAGCATATTGGTCTTT-3´ for centromeric outermost region (otr). Relative steady-state mRNA levels were determined from the threshold cycle for amplification using the 2-ΔΔCT method (Livak and Schmittgen, 2001). Control experiments measuring the change in ΔC_T with template dilution demonstrated that the efficiency of amplification of the target gene (adh1:gfp or aph) and the control (act1) was approximately equal.

**Nuclear run-on analysis**

RNA transcriptional elongation was analyzed by nuclear run-on essentially as described (Volpe et al., 2002). 250 ml cultures were grown to mid-log phase (OD_{600} = ~ 0.3 in YEA at 30ºC. Cell cultures were chilled on ice and cells were harvested by spinning for 5 min at 2,860 g at 4ºC. Cells were washed with 5 ml cold TMN buffer (10 mM Tris-HCl pH 7.4, 100 mM NaCl, 5 mM MgCl_2) and resuspended in 950 µl of cold H_2O. Cells were permeabilized by incubating them for 20 min on ice with 0.5% sarkosyl. At the end of incubation, cells were collected by centrifugation for 1 min at 1,830 g at 4ºC, supernatant was removed, and cells were resuspended in 300 µl of transcription
buffer (50 mM Tris-HCl pH 7.9, 5 mM MgCl$_2$, 1 mM MnCl$_2$, 100 mM KCl, 2 mM DTT, 1 mM rATP, 0.5 mM rCTP, 0.5 mM rGTP, 400 µCi α-UTP (6,000 Ci/mmol, MP Biomedicals), 100 U RNasin (Promega), 10 mM phosphocreatine phosphate, 12 µg/ml phosphocreatine kinase) and incubated for 10 min at 25°C. Reaction was stopped with 1 ml of ice-cold TMN buffer containing 1 mM cold UTP. Nuclei were pelleted by centrifugation at 1,830 g for 1 min and supernatant was removed. RNA was isolated with hot-phenol method as described above and purified from unincorporated nucleotides by passing through G-25 spin column (Roche). Incorporation efficiency was determined by calculating the trichloroacetic acid precipitable counts. Full-length RNA probes were transcribed with T7 RNA polymerase from PCR templates prepared with the following oligonucleotides: 5’-gcg taa tac gac cta tag GAA AAA CTC ATC GAG CAT CAA ATG-3’ and 5’-GGG TAA GGA AAA GAC TCA CG-3’ for aph; 5’-gcg taa tac gac cta tag ggC TTG GAA AGG TCC AAG ACG ATA C-3’ and 5’-GAC TAT TCC TGA CAA GCA GTT G-3’ for adh1; 5’-gcg taa tac gac cta tag GAA GCA CTT ACG GTA AAC GAT AC-3’ and 5’-GGA AGA AGA AAT CGC AGC GTT G-3’ for act1; 5’-gcg taa tac gac cta tag GGT TGG CAG GCT CGA CAT AAC GAT AC-3’ and 5’-GTC TAC TAA CAA GAT CAC TTTC-3’ for nmt1, where the T7 promoter sequence is in lowercase. *Pp* luc RNA was transcribed from a PCR-generated DNA template amplified from the pGL-2 Control vector (Promega) using 5’-gcg taa tac gac cta tag GAG AGG AAT TCA TTA TC-3’ and 5’-GGA GAG ATA CGC CCT GGT TCC TG-3’ primers. 5 µg of each RNA probe for firefly luciferase *Photinus pyralis* (*Pp*-luc), *aph*, *adh1*, *nmt1* and *act1* was denatured in 500 µl of ice-cold 10 mM NaOH and 1 mM EDTA, transferred to Hybond-XL membrane (Amersham) using a Bio-Dot SF Microfiltration Apparatus (Bio-Rad) according to manufacturer’s instructions, and immobilized on the membrane by UV
irradiation (200 μjoules/cm; Stratalinker, Stratagene, LaJolla, CA, USA). The membrane was pre-hybridized in Church buffer (Church and Gilbert, 1984) for 1 h at 65°C and hybridization carried out overnight at 65°C. After hybridization, membranes were washed twice with 2x SSC/0.2% (w/v) sodium dodecyl sulfate (SDS) for 30 min at 65°C and analyzed by phosphorimagery (Fuji, Tokyo, Japan).

**FACS analysis**

Log-phase cells were harvested, washed in PBS (10 mM Na₂HPO₄, 1.7 mM KH₂PO₄, pH 7.4, 137 mM NaCl and 2.7 mM KCl) and resuspended in 500 μl PBS. Fluorescence was analyzed using a FACScan Flow Cytometer (Becton-Dickinson). Constant settings were maintained for all experiments. Data were acquired from ten thousand cells in all experiments and analyzed with Cell Quest (Becton-Dickinson) and FACSPress 1.3 (Ray Hicks) software.

**Chromatin immunoprecipitation (ChIP)**

The density of RNA polymerase II on the *adh1: GFP* and *aph* genes and H3 K9 methylation was measured essentially as described (Takahashi et al., 2000). Briefly, 50 ml cultures were grown until OD₆₀₀ = 0.8. Protein was crosslinked to DNA by adding formaldehyde to a final concentration of 1% and incubating the cultures for 10–15 minutes at room temperature with gentle shaking. Crosslinking was quenched by adding glycine to a final concentration of 0.125 M. Cells were harvested by centrifugation, washed 3 times with ice-cold PBS, frozen in liquid nitrogen, and stored at –80°C until use. Cell pellets were resuspended in 200 μl of lysis buffer (50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 1 mM EDTA, 0.1% (w/v) sodium deoxycholate, 1% (w/v) Triton-X
100, 0.1% (w/v) SDS) containing for each 50 ml one tablet of Complete, EDTA-free Protease Inhibitor Cocktail (Roche) and 500 µl of Protease Inhibitor Cocktail For Fungal and Yeast Cells (Sigma), than vortexted with an equal volume of silica beads for 20 minutes at 4°C. The volume of the cell suspension was then adjusted by adding lysis buffer to a final volume of 650 µl. After sonicating to shear the chromosomal DNA to 250 to 500 bp, cell suspensions were centrifuged at 16,000 g for 10 min at 4°C and the supernatants divided into 3 equal portions. The volume of each portion was adjusted to 500 µl, anti-RNA polymerase II (Pol II) antibody (Covance), anti-H3 K9 antibody (Upstate), anti-HA antibody (Covance), or no antibody was added, and the samples incubated on a rotating wheel for 4 h at 4°C. Fifty µl salmon sperm DNA/protein A agarose beads (Upstate) was added and the incubation continued for 1.5-2 h at 4°C. Beads were pelleted by centrifugation at 500 g for 2 min, supernatants discarded, then the beads were washed once for 5 min with 1 ml of each of the following buffers: buffer 1 (lysis buffer with 1% sodium deoxycholate), buffer 2 (lysis buffer with 1% sodium deoxycholate and 1M NaCl), buffer 3 (50 mM Tris-HCl (pH 8.0), 0.25 M LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% sodium deoxycholate), and twice with 10 mM Tris-HCl (pH 7.6) containing 10 mM EDTA. Samples were treated with proteinase K for 4 h at 42°C and cross-linking reversed by overnight incubation at 65°C. After phenol-chloroform extraction and ethanol precipitation, the DNA was resuspended in 20 µl of water. Five percent of the total DNA was used in quantitative PCR analysis as described above (‘RNA analysis’). The difference in threshold cycles obtained when amplifying Pol II immunoprecipitated DNA and control DNA precipitated using non-specific (HA) or no antibody was at least 10 cycles. Control experiments demonstrated that < 2-fold
differences in the density of RNA Pol II on a gene were readily detectable with our protocol. The difference in threshold cycles obtained when amplifying H3 K9 immunoprecipitated centromeric DNA from wild type and control clr4Δ mutant strain was at least 4 cycles.

**RNase protection assay**

To distinguish the GFP hairpin transcripts from mRNA of the adh1:gfp fusion, we measured the rate of transcription for the GFP hairpin using sense probe for gfp. RNA probes were transcribed with T7 RNA polymerase from PCR templates prepared with the following oligonucleotides: 5′-gcgtaatacagtacactatacGAGTAAGGGAGAAGAAC
TTTTC-3′ and 5′-TTGGTCGCTTCGGATTGTTTACATAACCGGACATAATCA
TAGGTCCTCTGGGTTAGTTTCCGTATG-3′ for gfp, and 5′-gcgtaatacagtacactatacctagGAAGCACTTACGGTAAACGATAC
TTTTC-3′ and 5′-TTGGTCGCTTCCGGATTGTTTACATAACCGGACATAATCA
TAGGTCCTCTGAGGTAAAGGAGAAGAAC
TTTTC-3′ for act1, where T7 promoter region is in lowercase. Each probe contained 50 nt of Pp-luc sequence to distinguish full-length probes from protected products. The 180-nt sense gfp probe was synthesized using 40 µM cold UTP and 400 µCi of \(^{32}\)P α-UTP (6,000 Ci/mmol, MP Biomedicals, Irvine, CA, USA). The concentration of cold UTP during the synthesis of the 375-nt antisense act1 probe was 500 µM. Probes were purified from 5% polyacrylamide gel by electroelution using ElutaTube™ Protein, DNA and RNA Extraction and Dialysis Kit (Fermentas) according to manufacturer’s instructions and treated with DNase (RQ1, Promega). ~7x10⁴ cpm of each probe was used immediately for hybridization.
Cells were harvested from liquid cultures at mid-exponential phase (OD<sub>600</sub> = 0.1-0.4). Total RNA was extracted from cells using MasterPure<sup>TM</sup> Yeast RNA Purification Kit (Epicentre). Purified RNA was treated with DNase (RQ1, Promega); 25 µg was used for hybridization. Sample RNA and labeled probes were mixed and ethanol coprecipitated. Pellets were washed with 70% ethanol, dried and dissolved in 20 µl of hybridization buffer (40 mM Pipes pH 6.4, 1 mM EDTA, 0.4 M NaCl and 50% formamide). The mixture was heated at 100°C for 1 minute and hybridization was conducted at 46°C overnight. 300 µl of RNase digestion mix [20 µg/ml RNase A (Roche), 10 U/ml RNase T1 (Ambion), 10 U/ml RNase One (Promega), 10 mM Tris-HCl pH 7.5, 5 mM EDTA pH 8.0, and 0.3 M NaCl] were added and incubation was continued for 1 hr at 30°C followed by 1 hr at 37°C. The digestion was stopped by adding 5 µl of proteinase K (10 mg/ml) and 15 µl of 10% SDS and incubating for 15 min at 37°C. After phenol/chloroform extraction and ethanol precipitation, samples were resuspended in 10 µL of formamide loading buffer (98% w/v formamide, 10 mM EDTA pH 8.0, 0.025% w/v xylene cyanol and 0.025% w/v bromphenol blue) and analysed on 5% polyacrylamide gel.

**RNA isolation and detection by Northern blot**

Total RNA was purified using MasterPure Yeast RNA Purification Kit (Epicentre) and analyzed as previously described (Hamilton and Baulcombe, 1999). 50 µg of total RNA was resolved by 15% denaturing polyacrylamide/urea gel electrophoresis (National Diagnostics, Atlanta, GA, USA). 5´-<sup>32</sup>P-radiolabeled RNA oligonucleotides were used as size markers. After electrophoresis, the polyacrylamide gel
was transferred to Hybond N+ (Amersham-Pharmacia, Little Chalfont, UK) in 0.5x TBE by semi-dry transfer (Transblot SD, Bio-Rad) at 400 mA for 1h. The RNA was cross-linked to the membrane by UV irradiation (1200 µjoules/cm; Stratalinker, Stratagene, La Jolla, CA, USA) and pre-hybridized in 15ml of hybridization buffer (125 mM sodium phosphate buffer pH 7.2, 250 mM NaCl, 7% SDS and 50% deionized formamide) for 1-2 hr at 42°C. Hybridization buffer was replaced and freshly denatured probe was used for hybridization.

Full-length sense RNA probe for GFP was transcribed with T7 RNA polymerase from PCR template prepared with the following oligonucleotides: 5´-gcgtaatacgactcaatatagGAGTAAAGGAGAAGAACTTTTC-3´ and 5´-GTATAGTTCATCCATGCCA TG-3´, where T7 promotor region is in lowercase. Purified RNA was treated with DNase (RQ1, Promega), phenol/chloroform extracted and ethanol precipitated. Dried pellet was resuspended in 50 µl H2O. In order to generate fragments of ~50 nt, probe was incubated for 1h at 60°C in hydrolyzation buffer (60 mM Na2CO3 and 40 mM NaHCO3). Reaction was stopped by adding 3 µl of 3 M NaOAc (pH 5.0) and 5 µl of 10% acetic acid. Probe was purified from free nucleotides by passing through Sephadex G-25 spin column, denatured for 2 min at 95°C, and hybridization was carried out at 42°C overnight. After hybridization, membranes were washed twice with 2x SSC/0.2% (w/v) for 30 min at 42°C. Membranes were analyzed by phosphorimagery (Fuji, Tokyo, Japan). To strip probes, membranes were boiled in 10 mM Tris/HCl, pH 8.5 for 1 min, then re-exposed to confirm probe removal.
Acknowledgments

We thank members of the Zamore lab for support and for comments on the manuscript, Tamal Raha and Michael Green for help with Pol II ChIP, Janet Partridge and Robin Allshire for the \textit{chp}\Delta, \textit{clr4}\Delta, and \textit{swi6}\Delta strains, and Shiv Grewal and Danesh Moazed for the \textit{tas3}\Delta strain.
CHAPTER III

A DISTINCT SMALL RNA PATHWAY SILENCES SELFISH GENETIC ELEMENTS IN THE GERM LINE

Summary

In the *Drosophila* germ line, repeat-associated small interfering RNAs (rasiRNAs) ensure genomic stability by silencing endogenous selfish genetic elements such as retrotransposons and repetitive sequences. While small interfering RNAs (siRNAs) derive from both the sense and antisense strands of their double-stranded RNA precursors, rasiRNAs arise mainly from the antisense strand. rasiRNA production appears not to require Dicer-1, which makes microRNAs, or Dicer-2, which makes siRNAs, and rasiRNAs lack the 2′,3′ hydroxy termini characteristic of animal siRNA and miRNA. Unlike siRNAs and miRNAs, rasiRNAs function through the Piwi, rather than the Ago, Argonaute protein subfamily. Our data suggest that rasiRNAs protect the fly germ line through a silencing mechanism distinct from both the miRNA and RNAi pathways.
Introduction

In plants and animals, RNA silencing pathways defend against viruses (Galiana-Arnoux et al., 2006, Wang et al., 2006, Zambon et al., 2006) regulate endogenous gene expression (Du and Zamore, 2005), and protect the genome against selfish genetic elements such as retrotransposons and repetitive sequences (Kalmykova et al., 2005).

Common to all RNA silencing pathways are RNAs 19–30 nt long that specify the target RNAs to be repressed. In RNA interference (RNAi) (Fire et al., 1998), 21-23 nt small interfering RNAs (siRNAs) are produced from long exogenous double-stranded RNA (dsRNA). In contrast, ~22 nt microRNAs (miRNAs) are endonucleolytically processed from endogenous RNA Pol II transcripts. Dicer RNase III enzymes produce both siRNAs and miRNAs. In flies, Dicer-2 (Dcr-2) generates siRNAs, whereas the Dicer-1 (Dcr-1)/Loquacious (Loqs) complex produces miRNAs (Lee et al., 2004, Hoa et al., 2003, Forstemann et al., 2005, Saito et al., 2005, Jiang et al., 2005). After their production, small silencing RNAs bind Argonaute proteins to form the functional RNA silencing effector complexes believed to mediate all RNA silencing processes.
Results and Discussion

Phased sense and antisense siRNAs in vivo

In *Drosophila*, processive dicing of long dsRNA and the accumulation of sense and antisense siRNAs without reference to the orientation of the target mRNA are hallmarks of RNAi in vitro (Schwarz et al., 2003, Zamore et al., 2000) and in vivo (Fig. 1). We prepared total small RNA from the heads of adult males expressing a dsRNA hairpin (Supplementary Fig. 1A) that silences the *white* gene via the RNAi pathway (Lee and Carthew, 2003). *white* silencing requires Dcr-2 (Lee et al., 2004), R2D2 (Forstemann et al., 2005), and Ago2. siRNAs were detected with a microarray containing T_M (melting temperature)-normalized, probes, 22 nt long, for all sense and antisense siRNAs that theoretically can be produced by dicing the *white* exon 3 hairpin (Fig. 1A). Both sense and antisense *white* siRNAs were detected in wild-type flies but not in dcr-2^{L811fsX} homozygous mutant flies. The Dcr-2-dependent siRNAs were produced with a periodicity of ~22 nt (Fig. 1B), consistent with the phased processing of the dsRNA hairpin from the end formed by the 6-nt loop predicted to remain after splicing of its intron-containing primary transcript (Supplementary Fig. 1B).

*Su*(Ste) rasiRNAs

*Drosophila* repeat-associated small interfering RNAs (rasiRNAs) can be distinguished from siRNAs by their longer length, 24 to 29 nt (Aravin et al., 2003, Aravin et al., 2001). rasiRNAs have been proposed to be diced from long dsRNA triggers (Kalmykova et al., 2005, Tritto et al., 2003), such as the ~50 copies of the bidirectionally transcribed *Suppressor of Stellate* [Su(Ste)] locus on the Y-chromosome (Tritto et al., 2003) (Supplementary Fig. 2B) that in testes silence the ~200 copies of the protein-coding gene *Stellate* (*Ste*) found on the X-chromosome.
A

sense and antisense siRNA in wild-type

microarray signal (normalized)

1 0.8 0.6 0.4 0.2 0

probe position in white exon 3

B

amplitude

2 4 6

oscillation period (nt)
**Figure 1.** Sense and antisense siRNAs accumulate during RNAi in vivo. (A) Microarray analysis of the siRNAs derived from the *white* exon 3 hairpin RNAi trigger. (B) Fourier-transform analysis of the Dcr-2-dependent siRNAs in (A).
Microarray analysis of total small RNA isolated from fly testes revealed that *Su(Ste)* rasiRNAs detectably accumulate only from the antisense strand (Fig. 2A), with little or no phasing (Supplementary Fig. 2A). As expected, *Su(Ste)* rasiRNAs were not detected in testes from males lacking the *Su(Ste)* loci (*cry*¹Y) (Fig. 2A). *Su(Ste)* rasiRNA were also absent from *armitage* (*armi*) mutant testes (Fig. 2A), which fail to silence *Ste* and do not support RNAi in vitro (Tomari et al., 2004a). *armi* encodes a non-DEAD box helicase (Cook et al., 2004) homologous to the *Arabidopsis thaliana* protein SDE3, which is required for RNA silencing triggered by transgenes and some viruses (Dalmay et al., 2001), and depletion by RNAi of the mammalian Armi homolog Mov10 blocks siRNA-directed RNAi in cultured human cells (Meister et al., 2005). Normal accumulation of *Su(Ste)* rasiRNA and robust *Ste* silencing also requires the putative helicase Spindle-E (Spn-E) a member of the DExH family of adenosine triphosphatases (ATPases) (Aravin et al., 2001, Stapleton et al., 2001) (Fig. 2B and Supplementary Fig. 2B).

The accumulation in vivo of only antisense rasiRNAs from *Su(Ste)* implies that sense *Su(Ste)* rasiRNA are not produced or are selectively destroyed. Either process would make *Ste* silencing mechanistically different from RNAi. In support of this view, mutations in the central components of the *Drosophila* RNAi pathway—*dcr-2*, *r2d2*, and *ago2*—did not diminish *Su(Ste)* rasiRNA accumulation (Fig 2B). Deletion of the *Su(Ste)* silencing trigger (*cry*¹Y) caused a factor of ~65 increase in *Ste* mRNA (Fig. 3A), but null or strong hypomorphic mutations in the three key RNAi proteins did not (Fig. 3B).

Fly Argonaute proteins can be subdivided into Ago (Ago1 and Ago2) and Piwi [Aubergine (Aub), Piwi, and Ago3] subfamilies. Unlike *ago1* and *ago2*, the *aub*, *piwi*, and *ago3* mRNAs are enriched at the germ line (Williams and Rubin, 2002, Cox et al., 1998, Harris and Macdonald, 2001). Aub is required for *Ste* silencing (Aravin et al., 2001) and *Su(Ste)* rasiRNA accumulation (Aravin et al., 2004). In *aub¹/HN2/aub²QC42* transheterozygous mutants, *Su(Ste)* rasiRNA were not detected by microarray
CHAPTER III, Figure 2

A

![Graph showing microarray signal (normalized) for probe position in Su(Ste) transcript.](image)

B

<table>
<thead>
<tr>
<th>Gene</th>
<th>WT</th>
<th>+/-</th>
<th>-/-</th>
<th>+/-</th>
<th>-/-</th>
<th>+/-</th>
<th>-/-</th>
<th>+/-</th>
<th>-/-</th>
<th>+/-</th>
<th>-/-</th>
</tr>
</thead>
<tbody>
<tr>
<td>Su(Ste) rasiRNA</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-8</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2S rRNA</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Legend:**
- loqs
- armi
- dcr-2
- spn-E
- ago2
- piwi
- aub
- r2d2
- ago2
**Figure 2.** rasiRNAs derived from the endogenous silencing trigger *Su(Ste).* (A) Microarray analysis of *Su(Ste)* small RNAs in wild-type, *cry^1Y*, and *armi* mutant testes. (B) Northern analysis of the most abundant *Su(Ste)* rasiRNA in testes mutant for RNA silencing genes.
(Supplementary Fig. 2B) or Northern analysis (Fig. 2B), and Su(Ste)-triggered silencing of Ste mRNA was lost completely (Fig. 3B). Even aub^{HN2}/+ heterozygotes accumulated less of the most abundant Su(Ste) rasiRNA than wild-type (Fig. 2B). That the Ago subfamily protein Ago2 is not required for Ste silencing, whereas the Piwi subfamily protein Aub is essential for it, supports the view that Ste is silenced by a pathway distinct from RNAi. Intriguingly, Su(Ste) rasiRNAs hyperaccumulated in piwi mutant testes, where Ste is silenced normally (Figs. 2B and 3B and Supplementary Fig. 2B).

Mutations in aub also cause an increase in sense, but not antisense Su(Ste) RNA (Aravin et al., 2001); our results suggest that antisense Su(Ste) rasiRNAs can silence both Ste mRNA and sense Su(Ste) RNA, but that no Su(Ste) rasiRNA exist that can target the antisense Su(Ste) transcript. Our finding that Su(Ste) rasiRNAs are predominantly or exclusively antisense is essentially in agreement with the results of small RNA cloning experiments, in which four of five Su(Ste) rasiRNAs sequenced were in the antisense orientation (Aravin et al., 2003), but is at odds with earlier reports detecting both sense and antisense Su(Ste) rasiRNAs by non-quantitative Northern hybridization (Aravin et al., 2001).

A third RNA silencing pathway in flies

Is germline RNA silencing of selfish genetic elements generally distinct from the RNAi and miRNA pathways? We examined the expression of a panel of germline expressed selfish genetic elements—three long terminal repeat (LTR)-containing retrotransposons (roo, mdg1, and gypsy), two non-LTR retrotransposons (I-element and HeT-A, a component of the Drosophila telomere), and a repetitive locus (mst40)—in mutants defective for eight RNA silencing proteins. All selfish genetic elements tested behaved like Ste: loss of the RNAi proteins Dcr-2, R2D2, or Ago2 had little or no effect on retrotransposon or repetitive element silencing (Fig. 3, C and D). Instead, silencing required the putative helicases Spn-E and Armi plus one or both of the Piwi subfamily
CHAPTER III, Figure 3

A

change in steady-state RNA

Stellate mRNA (target)  Su(Ste) sense RNA (trigger locus)

wild-type  cry'/Y

B

change in steady-state RNA

dcr-2  r2d2  spn-E  armi  ago2  piwi  aub

C

change in steady-state RNA

dcr-2  r2d2  loqs  spn-E  armi  ago2  piwi  aub

D

change in steady-state HeT-A mRNA

dcr-2  r2d2  loqs  armi  ago2  aub

E

change in steady-state RNA

roo  Het-A  mdg1  mast40
Figure 3. (A to D) RNA expression from selfish genetic elements was measured in homozygous mutants relative to heterozygotes for Ste silencing in testes [(A) and (B)] and for the repeated locus mst40, the LTR-retrotransposons roo, mdg1, and gypsy, and the non-LTR-retrotransposons I-element, and HeT-A in ovaries [(C) and (D)]. (E) RNA expression from selfish genetic elements in dcr-1<sup>Q1147X</sup> null mutant clones generated by mitotic recombination in the ovary.
Argonaute proteins, Aub or Piwi. Silencing did not require Loqs, the dsRNA-binding protein required to produce miRNAs (Fig. 3, C and D).

The null allele dcr-1^{Q1147X} is homozygous lethal, making it impossible to procure dcr-1 mutant ovaries from dcr-1^{Q1147X} dcr-1^{Q1147X} adult females (Lee et al., 2004). Therefore, we generated clones of dcr-1^{Q1147X} dcr-1^{Q1147X} cells in the ovary by mitotic recombination in flies heterozygous for the dominant female-sterile mutation ovo^{D1} (Sahut-Barnola and Pauli, 1999). We measured RNA levels, relative to rp49 mRNA, for three retrotransposons (roo, HeT-A, and mdg1) and one repetitive sequence (mst40) in dcr-1 dcr-1 recombinant ovary clones and in ovo^{D1} TM3 and dcr-1 ovo^{D1} nonrecombinant ovaries. The ovo^{D1} mutation blocks oogenesis at stage 4, after the onset of HeT-A (Savitsky et al., 2006) and roo rasiRNA production. Retrotransposon or repetitive sequence transcript abundance was unaltered or decreased in dcr-1 dcr-1 relative to ovo^{D1} TM3 and dcr-1 ovo^{D1} controls (Fig. 3E). We conclude that Dcr-1 is dispensable for silencing these selfish genetic elements in the Drosophila female germ line.

roo is the most abundant LTR-retrotransposon in flies. We analyzed roo silencing in the female germ line with the use of microarrays containing 30-nt probes, tiled at 5-nt resolution, for all ~18,000 possible roo rasiRNAs (Fig. 4, A and B); we corroborated the data at 1-nt resolution for those rasiRNAs derived from LTR sequences (Supplementary Fig. 3A). As observed for Su(Ste) but not for white RNAi, roo rasiRNAs were nonhomogeneously distributed along the roo sequence and accumulated primarily from the antisense strand (Fig. 4A and Supplementary Fig. 3A). In fact, the most abundant sense rasiRNA peak (asterisk in Fig. 4, A and B) corresponded to a set of probes containing 16 contiguous uracil residues, which suggests that these probes nonspecifically detected fragments of the mRNA polyadenylate [poly(A)] tail. Most of the remaining sense peaks were unaltered in armi mutant ovaries, in which roo expression is increased; this result implies that they do not contribute to roo silencing.
**CHAPTER III, Figure 4**

**A**

antisense roo rasiRNA in wild-type

microarray signal (normalized)

probe position in roo

**B**

antisense roo rasiRNA in armi/armi

microarray signal (normalized)

probe position in roo

**C**

<table>
<thead>
<tr>
<th>dcr-2</th>
<th>r2d2</th>
<th>loqs</th>
<th>armi</th>
<th>spin-E</th>
<th>ago2</th>
<th>piwi</th>
<th>aub</th>
</tr>
</thead>
<tbody>
<tr>
<td>+/--/-</td>
<td>+/--/-</td>
<td>+/--/-</td>
<td>+/--/-</td>
<td>+/--/-</td>
<td>+/--/-</td>
<td>+/--/-</td>
<td>+/--/-</td>
</tr>
</tbody>
</table>

#1 roo rasiRNA

pre-miR-8

miR-8

pre-miR-311

miR-311

2S rRNA
Figure 4. roo rasiRNAs in ovaries arise mainly from the antisense strand and require the putative helicase Armi for their accumulation. (A and B) Microarray analysis of roo rasiRNAs in (A) wild-type and (B) armi mutant ovaries. The asterisk marks a peak arising from non-specific hybridization to poly(A)-containing RNAs. (C) Northern analysis for roo rasiRNA [peak 1 in (A)] in ovaries heterozygous (+/−) or homozygous mutant (−/−) for components of the RNA silencing machinery. The open red arrowhead highlights pre-miR-311 accumulation in loqs^{00791}; the solid red arrowhead highlights loss of mature miR-311. miR-311 is expressed predominantly in the germ line.
(Figs. 3C and 4A). We detected no phasing in the distribution of roo rasiRNAs (Supplementary Fig. 3B).

As for Su(Ste), wild-type accumulation of antisense roo rasiRNA required the putative helicases Armi and Spn-E and the Piwi subfamily Argonaute proteins Piwi and Aub, but not the RNAi proteins Dcr-2, R2D2, and Ago2 (Fig. 4, B and C). Moreover, accumulation of roo rasiRNA was not measurably altered in loqs<sup>010791</sup>, an allele that strongly disrupts miRNA production in the female germ line (Fig. 4C).

**Are roo rasiRNAs not made by dicing?**

Loss of Dcr-2 or Dcr-1 did not increase retrotransposon or repetitive element expression, which suggests that neither enzyme acts in rasiRNA-directed silencing. Moreover, loss of Dcr-2 had no detectable effect on Su(Ste) rasiRNA in testes or roo rasiRNA in ovaries (Figs. 2B and 4C). We measured the amount of roo rasiRNA and miR-311 in dcr-1/dcr-1 ovary clones generated by mitotic recombination. Comparison of recombinant (dcr-1/dcr-1) and nonrecombinant (ovo<sup>D1</sup>/TM3 and dcr-1/ovo<sup>D1</sup>) ovaries by Northern analysis revealed that roo rasiRNA accumulation was unperturbed by the null dcr-1<sup>Q1147X</sup> mutation (Fig. 5A and Supplementary Fig. 4). Pre-miR-311 increased and miR-311 declined by a factor of ~3 in the dcr-1/dcr-1 clones (Fig. 5B and Supplementary Fig. 4), consistent with about two-thirds of the tissue corresponding to mitotic dcr-1/dcr-1 recombinant cells. Yet, although most of the tissue lacked dcr-1 function, we observed improved, rather than diminished, silencing for the four selfish genetic elements examined (Fig 3E). Moreover, the dsRNA-binding protein Loqs, which acts with Dcr-1 to produce miRNAs, was also dispensable for roo rasiRNA production and selfish genetic element silencing (Fig. 3, C and D, and Fig. 4C). Although we cannot exclude the possibility that dcr-1 and dcr-2 can fully substitute for each other in the production of rasiRNA in the ovary, previous biochemical evidence suggests that none of the three RNase III enzymes in flies—Dcr-1, Dcr-2, and Drosha—can cleave long...
A miR-311—2S rRNA—pre-miR-311—roo rasiRNAs—ovo D1/TM3

B

relative small RNA abundance

roo rasiRNA  miR-311  pre-miR-311
**Figure 5.** (A) Northern analysis of mitotic recombinant $dcr-1^{Q1147X}$ homozygous mutant ovaries and non-recombinant controls. *roo* rasiRNA were detected using a mixture of five hybridization probes. Arrowheads are as in Fig. 4C. (B) Quantification of the data in (A) normalized to the 2S rRNA loading control.
dsRNA into small RNAs 24 to 30 nt long (Saito et al., 2005, Bernstein et al., 2001, Liu et al., 2003).

Animal siRNA and miRNA contain 5’ phosphate and 2’,3’ hydroxy termini (Elbashir et al., 2001c, Hutvagner et al., 2001). We used enzymatic and chemical probing to infer the terminal structure of roo and Su(Ste) rasiRNAs. RNA from ovaries or testes was treated with calf intestinal phosphatase (CIP) or CIP followed by polynucleotide kinase plus ATP. CIP treatment caused roo (Fig. 6A) and Su(Ste) (Supplementary Fig. 5) rasiRNA to migrate more slowly in polyacrylamide gel electrophoresis, consistent with the loss of one or more terminal phosphate groups. Subsequent incubation with polynucleotide kinase and ATP restored the original gel mobility of the rasiRNAs, indicating that they contained a single 5’ or 3’ phosphate before CIP treatment. The roo rasiRNA served as a substrate for ligation of a 23-nt 5’ RNA adapter by T4 RNA ligase, a process that requires a 5’ phosphate; pre-treatment with CIP blocked ligation (Fig. 6B), thus establishing that the monophosphate lies at the 5’ end. The rasiRNA must also contain at least one terminal hydroxyl group, because it could be joined by T4 RNA ligase to a preadenylated 17-nt 3’ RNA adapter (Fig. 6B). Notably, the 3’ ligation reaction was less efficient for the roo rasiRNA than for a miRNA in the same reaction (22% versus 50% conversion to ligated product).

RNA from ovaries or testes was reacted with NaIO₄, then subject to β-elimination, to determine whether the rasiRNA had either a single 2’ or 3’ terminal hydroxy group or had terminal hydroxy groups at both the 2’ and 3’ positions, as do animal siRNA and miRNA. Only RNAs containing both 2’ and 3’ hydroxy groups react with NaIO₄; β-elimination shortens NaIO₄-reacted RNA by one nucleotide, leaving a 3’ monophosphate terminus, which adds one negative charge. Consequently, NaIO₄-reacted, β-eliminated RNAs migrate faster in polyacrylamide gel electrophoresis than does the original unreacted RNA. Both roo (Fig. 6A) and Su(Ste) (Supplementary Fig. 5) rasiRNA lack either a 2’ or a 3’ hydroxyl groups, because they failed to react with NaIO₄;
Figure 6. *roo* rasiRNAs are modified at their 3´ terminus and associate with Piwi subfamily Argonaute proteins. (A) Chemical and enzymatic probing of *roo* rasiRNA structure. *roo* rasiRNA (peak 1 in Fig. 4A) were detected by Northern hybridization. The membrane was then stripped and reprobed for miR-8. (B) *roo* rasiRNA can serve as a 3´ or a 5´ substrate for T4 RNA ligase. Solid arrowheads, 5´ ligation products; open arrowheads, 3´ ligation products. (C and D) *roo* rasiRNA associate with myc-tagged Piwi (C) and GFP-tagged Aub protein (D), but not with Ago1. I, input; S, supernatant; B, bound.
miRNAs in the same samples reacted with NaIO₄. Together, our results show that rasiRNA contain one modified and one unmodified hydroxyl. Because T4 RNA ligase can make both 3´–5´ and 2´–5´ bonds (Yang et al., 2006), we cannot currently determine the blocked position. Some plant small silencing RNAs contain a 2´-O-methyl modification at their 3´ terminus (Yang et al., 2006).

**rasiRNA bind Piwi and Aub**

*Drosophila* and mammalian siRNA and miRNA function through members of the Ago subfamily of Argonaute proteins, but *Su(Ste)* and *roo* rasiRNAs require at least one member of the Piwi subfamily for their function and accumulation. To determine whether *roo* rasiRNAs physically associate with Piwi and Aub, we prepared ovary lysate from wild-type flies or transgenic flies expressing either myc-tagged Piwi or green fluorescent protein (GFP)-tagged Aub protein (Harris and Macdonald, 2001, Cox et al., 2000); immunoprecipitated them with monoclonal antibodies (mAb) to myc, GFP, or Ago1; and then analyzed the supernatant and antibody-bound small RNAs by Northern blotting (*Fig. 6, C and D, and Supplementary Fig. 6*). We analyzed six different *roo* rasiRNAs. All were associated with Piwi but not Ago1, the *Drosophila* Argonaute protein typically associated with miRNAs (Okamura et al., 2004); miR-8 (*Fig. 6D and Supplementary Fig. 6C*), miR-311 and *bantam* immunoprecipitated with Ago1 mAb. No rasiRNAs immunoprecipitated with the myc mAb when we used lysate from flies lacking the myc-Piwi transgene (*Supplementary Fig. 6B*).

Although *aub* mutant ovaries silence *roo* mRNA normally, they showed reduced accumulation of *roo* rasiRNA relative to *aub/+* heterozygotes (*Fig. 4C*), which suggests that *roo* rasiRNAs associate with both Piwi and Aub. We analyzed the supernatant and antibody-bound small RNAs after GFP mAb immunoprecipitation of ovary lysate from GFP-Aub transgenic flies and flies lacking the transgene. *roo* rasiRNA was recovered only when the immunoprecipitation was performed with the GFP mAb in ovary lysate.
from GFP-Aub transgenic flies (Fig. 6D and Supplementary Fig. 6D). The simplest interpretation of our data is that roo rasiRNA physically associate with both Piwi and Aub, although it remains possible that the roo rasiRNAs are loaded only into Piwi and that Aub associates with Piwi in a stable complex. The association of roo rasiRNA with both Piwi and Aub suggests that piwi and aub are partially redundant, as does the modest reduction in roo silencing—by a factor of 3.3 ± 0.6 (average ± SD, n = 3)—in piwi but not in aub mutants (Fig. 3C). Alternatively, roo silencing might proceed through Piwi alone, but the two proteins could function in the same pathway to silence selfish genetic elements.

Our data suggest that in flies, rasiRNAs are produced by a mechanism that requires neither Dcr-1 nor Dcr-2, yet the patterns of rasiRNAs that direct roo and Ste silencing are as stereotyped as the distinctive siRNA population generated from the white hairpin by Dcr-2 (Fig. 1A) or the unique miRNA species made from each pre-miRNA by Dcr-1. A key challenge for the future will be to determine what enzyme makes rasiRNAs and what sequence or structural features of the unknown rasiRNA precursor lead to the accumulation of a stereotyped pattern of predominantly antisense rasiRNAs.
CHAPTER III, Supplementary Figure 1

A

GUC
CAG

white exon 3

white exon 3

5’
3’

B

Dicer-2

~22 nt ~22 nt ~22 nt ~22 nt
**Supplementary Figure 1.** (A) The structure of the *white* exon 3 hairpin trigger RNA. (B) A model describing the data in Fig. 1, which suggest that Dicer-2 initiates dsRNA processing from the loop end of the *white* exon 3 hairpin, then moves in ~22 nt steps to create a phased series of siRNAs.
CHAPTER III, Supplementary Figure 2

A

![Graph showing amplitude vs. oscillation period (nt)]

B

![Diagram showing probe position in Su(Ste) transcript]

antisense rasiRNA in wild-type
antisense rasiRNA in spn-E
antisense rasiRNA in wild-type
antisense rasiRNA in piwi
antisense rasiRNA in wild-type
antisense rasiRNA in aub
Supplementary Figure 2. Characterization of Su(Ste) rasiRNAs isolated from testes. (A) The microarray data for sense and antisense Su(Ste) rasiRNAs in wild-type were examined for phasing using Fourier-transform analysis. No oscillations corresponding to the length of either a typical siRNA (21 to 23 nt) or rasiRNA (24 to 30 nt) were detected. The amplitude scale here is the same as that in Fig. 1B. (B) Microarray analysis of Su(Ste) antisense rasiRNA in testes from wild-type and spn-E, piwi, or aub mutant flies. Above the array data, the schematic indicates the position of the rasiRNAs within a single Su(Ste) locus.
A

1 nt tiling

sense and antisense

roo rasiRNA in wild-type

probe position in roo retrotransposon

B

amplitude

oscillation period (nt)
Supplementary Figure 3. (A) High resolution tiling microarray analysis across the roo LTR. The absolute peak heights in this experiment have been normalized so that they can be compared directly with those in Fig. 4A and 4B. (B) Fourier-transform analysis of the combined sense and antisense roo rasiRNA data for wild-type ovaries. No oscillations corresponding to the length of either a typical siRNA (21 to 23 nt) or rasiRNA (24 to 30 nt) were detected. The amplitude scale here is the same as that in Fig. 1B and Supplementary Fig. 2.
CHAPTER III, Supplementary Figure 4

miR-311—

2S rRNA—

pre-

miR-311—

roo

rasiRNAs—

dcr-1/dcr-1

ovoD1/TM3

dcr-1/ovoD1

miR-311—

2S rRNA—
**Supplementary Figure 4.** Northern analysis of mitotic recombinant *dcr-IQ1147X* homozygous mutant ovaries and non-recombinant controls. *roo* rasiRNA were detected using a mixture of hybridization probes for five different rasiRNA peaks. The open red arrowhead highlights pre-miR-311 accumulation in the *dcr-1* mutant tissue; the solid red arrowhead highlights loss of mature miR-311. The data here correspond to an experiment independent from that in Fig. 5. A third independent experiment that produced essentially identical results is not shown.
CHAPTER III, Supplementary Figure 5

Superscript RNA—miR-311

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Su(Ste) rasiRNA</th>
<th>miR-311</th>
</tr>
</thead>
<tbody>
<tr>
<td>no treatment</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>CIP</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>CIP + PNK</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>β-elimination</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>RNase</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
<tr>
<td>DNase</td>
<td>![Image]</td>
<td>![Image]</td>
</tr>
</tbody>
</table>
Supplementary Figure 5. Northern analysis of enzymatic and chemical probing of total small RNA from fly testes. The major Su(Ste) rasiRNA contains a 5´ monophosphate terminus, but lacks the 2´, 3´ hydroxy terminus characteristic of siRNA and miRNA. The major Su(Ste) rasiRNA was detected by Northern hybridization. The membrane was then stripped and re-probed for miR-311. As expected, miR-311 contains a 5´ phosphate and a 2´, 3´ hydroxy terminus.
CHAPTER III, Supplementary Figure 6

A

anti-Piwi anti-myc
myc-Piwi Piwi

anti-Aub anti-GFP
GFP-Aub Aub

B

I.P.
anti-myc anti-GFP anti-Ago1

#1 roo rasiRNA
wild-type ovaries

C

anti-Ago1 I.P.

WT myc-Piwi

anti-myc anti-GFP anti-Ago1

#4 roo rasiRNA
myc-Piwi transgenic ovaries
#5 roo rasiRNA
#6 roo rasiRNA
wild-type ovaries
#4 roo rasiRNA
Supplementary Figure 6. (A) Transgenic myc-tagged and endogenous Piwi or transgenic GFP-tagged and endogenous Aub proteins were resolved by high resolution SDS-PAGE and measured by quantitative Western blotting. Representative data from the analysis is shown and expression values are given in the text. The blots were probed first with anti-Piwi or anti-Aub, then stripped and reprobed with myc mAb or GFP mAb. (B) roo rasiRNA were not detected by Northern analysis in the immunoprecipitates when mAbs to myc, GFP, or Ago1 were used with ovary lysate prepared from wild-type flies carrying neither a myc-Piwi- nor a GFP-Aub-expressing transgene. (C) miR-8 was detected associated with Ago1 in both wild-type (WT) and myc-Piwi-expressing transgenic ovary lysate. (D) Three roo rasiRNAs identified by microarray analysis (Fig. 4A) co-immunoprecipitated with myc-Piwi in ovary lysate isolated from myc-Piwi-expressing transgenic flies but not wild-type ovary lysate. I.P., immunoprecipitation; I, input; S, supernatant; B, bound.
**Supplementary Figure 7.** Schematic representations of the selfish genetic elements examined in this study, showing the position of the 5´ ends of the quantitative RT-PCR primers (open arrowheads). All elements are drawn at the same scale. The primers for heterochromatic *Stellate* also amplify RNA from the euchromatic *Stellate* loci and thus monitor total *Stellate* expression.
Materials and Methods

Fly stocks

The following fly stocks were used: eyFlp; FRT82B dcr-1^Q1147X/TM3 Sb (Hatfield et al., 2005, Lee et al., 2004), hsFlp; P[neoFRT]82B P[ovo^{D1-18}] 3R/TM3 Sb, w^{1118}; loqs^{00791}/CyO (Forstemann et al., 2005), y w eyFlp; FRT42D dcr-2^L8116X (Lee et al., 2004, Pham et al., 2004), y w;r2d2/CyO (Liu et al., 2003), y w;ago2^{414}/TM6B Tb (Okamura et al., 2004), w;ago2^{51B}/TM6B Tb (Xu et al., 2004), w;piwi^{2}/CyO (Cox et al., 1998), aub^{H321} cn bw/CyO, and aub^{O242} cn bw/CyO (Schupbach and Wieschaus, 1991), w; spn^{E100.37}/TM3 Sb (kind gift of R. Lehmann), armi^{72.1}/TM3 Sb P[hs-hid] (Cook et al., 2004) and P[nos-Gal4-VP16]/TM3 Sb (Reinke et al., 2000), P[UAS-GFP-aub]/TM3 Sb (Harris and Macdonald, 2001) and P[hs-myc-piwi]/CyO (Cox et al., 2000), FRT42D dcr-2^L8116X/CyO; P[w-IR]/P[w-IR] and P[w-IR]/TM6B Tb (Kim et al., 2005, Lee and Carthew, 2003) and cry^{1}Y (Belloni et al., 2002, Schmidt et al., 1999).

Tiling microarrays

The expression profile for all known miRNAs was analyzed together with all possible rasiRNAs derived from Su(Ste) or roo sequence using custom synthesized microarrays containing 22 nt, T_M^-normalized probes every 1 nt or 30 nt, T_M^-normalized probes every 5 nt. Microarray fabrication, hybridization and data acquisition were performed at LC Sciences (Houston, TX USA). 20 µg ovary, testis, head total small RNA was isolated from wild-type or mutant flies using the mirVana kit (Ambion, Austin, TX, USA). The presence of antisense and the absence of sense rasiRNA peaks for Su(Ste) and roo were confirmed by Northern hybridization. All microarray data is available at http://www.ncbi.nlm.nih.gov/geo/ using Gene Expression Omnibus accession number GSE4932. Fig. 1A displays the average of three independent samples for wild-type; Fig. 2A shows the average of seven independent samples.
**Microarray data analysis**

Raw data was processed as follows: background (defined as the average of signal readings on "BKG0" spots, which were spotted with chemical linker, but no probe) was subtracted from each intensity value. Intensities above $\exp(5) \approx 150$ were considered significant, as described in reference (Lu et al., 2005b), and intensities below 30,000 were considered non-saturated, as established with a titration of synthetic 21 nt RNAs added to the sample. Control RNAs (ribosomal RNAs, snRNAs, snoRNAs, tRNAs and synthetic RNAs added to the sample) were used for normalization, dividing each ‘mutant’ reading by the geometric average of all mutant/wt ratios for significantly, non-saturated, control RNAs. Irreproducible peaks, i.e. those that appeared in only one of a series of replicate experiments, and single-probe peaks, i.e. those peaks where the probe signal was both higher than the average signal and more than five times greater than the signal from its immediate neighbors, were removed from the datasets. For 5-nt tiling resolution, we imposed an additional criterion before removing peaks from the data set: the difference in signal intensity between the single peak and its neighbors was required to be greater than 500. When available, color-reversed datasets were averaged. Data were scaled to the highest peak in the wild-type genotype.

Fourier-transform analysis parses a signal into its periodic components and thus allows comparison of the contribution of each frequency of oscillation within a data set. To permit Fourier transform analysis, deleted peaks were replaced by the average signal intensity of the rest of the dataset to ensure a continuous wave function. Where required, the last data point was removed, so that the number of points was always an even number, a requirement of the fast Fourier transformation algorithm we used. Datasets were centered by subtracting the average from each data point, then fast Fourier transform and its modulus were computed using Igor Pro version 4.06A Carbon (WaveMetrics, Portland, OR, USA). We normalized Fourier-transform data by dividing each value by the average of the values in the interval from 5 nt to 140 nt interval, the
range of values displayed in the figures, to allow direct comparison of different experiments.

**RNA isolation and detection by Northern blot**

Ovaries and testes were manually dissected with forceps into *Drosophila* Ringer’s solution (182 mM KCl, 46 mM NaCl, 3 mM CaCl₂, 10 mM Tris-HCl, pH 7.5). RNA was isolated from testes or ovaries using the mirVana kit (Ambion), and then treated with RQ1 DNase (Promega, Madison, WI, USA). Total RNA was quantified by absorbance at 260 nm, and 1-10 µg of total RNA was resolved by 20% denaturing polyacrylamide/urea gel electrophoresis (National Diagnostics, Atlanta, GA, USA). 5’-32P-radiolabeled RNA oligonucleotides were used as size markers. After electrophoresis, the polyacrylamide gel was transferred to Hybond N+ (Amersham-Pharmacia, Little Chalfont, UK) in 0.5x TBE by semi-dry transfer (Transblot SD, Bio-Rad) at 20 V for 1–2 h. The RNA was cross-linked to the membrane by UV irradiation (1200 µjoules/cm; Stratalinker, Stratagene, La Jolla, CA, USA) and pre-hybridized in Church buffer (Church and Gilbert, 1984) for 1 h at 37 or 65°C. 25 pmol of RNA (Dharmacon, Lafayette, CO, USA) or DNA (IDT, Coralville, IA, USA) probe was 5’-32P-radiolabeled with polynucleotide kinase (New England Biolabs, Beverly, MA, USA) and 330 µCi γ-32P-ATP (7,000 µCi/mmol; New England Nuclear, Boston, MA, USA) and purified using a Sephadex G-25 spin column (Roche, Basel, Switzerland). To detect 2S rRNA, 1/50th of the 32P-radiolabeled probe was diluted with unlabeled 2S rRNA probe. The 32P-radiolabeled probes were hybridized in Church buffer for 4–12 h. For RNA probes, hybridization was at 65°C; for DNA probes, hybridization was at 37°C. After hybridization, membranes were washed twice with 2x SSC/0.1% (w/v) sodium dodecyl sulfate (SDS) and once with 1x SSC/0.1% (w/v) SDS for 30 min. Membranes were analyzed by phosphorimagery (Fuji, Tokyo, Japan). To strip probes, membranes were boiled in 0.1% (w/v) SDS for 5 min, then re-exposed to confirm probe removal. Probe sequences are listed in Table S1.
**Generation of mutant germ line clones**

*dcr-1* mutant female germ line clones were generated by Flp/FRT recombination as described previously (Xu and Rubin, 1993, Chou et al., 1993, Chou and Perrimon, 1992, Chou and Perrimon, 1996). hsFlp; P[neoFRT]82B P[ovo\textsuperscript{D1}] 3R/TM3 Sb males were crossed with eyFlp;FRT82B *dcr-1\textsuperscript{Q1147X}/TM3 Sb* females. After 3 days, parental flies were removed and the larvae heat shocked at 37°C for 2 h daily until eclosion. Two to three days after eclosion, females were collected and fed on yeasted media for 3 days, and then ovaries were manually dissected. Only ovaries containing late stage oocytes were used as *dcr-1/dcr-1* mutant clones for analysis.

**Analysis of rasiRNA and miRNA chemical structure**

15 µg of wild-type total ovary or testis small RNA was used for each reaction. RNase treatment was in 0.2 M NaCl with 1 µl 32 mg/ml RNase A (Sigma-Aldrich, St. Louis, MO, USA) for 2 h at 30°C. DNase treatment was in 1 x RQ1 DNase buffer with 5 U RQ1 DNase (Promega, Madison, WI, USA). Nucleases were inactivated by phenol/chloroform extraction, followed by precipitatin of the RNA in the presence of 20 µg glycogen carrier (Roche Diagnostics GmbH, Mannheim, Germany) at −70°C for 1 h, the precipitate was collected by centrifugation and dissolved in denaturing gel loading buffer. NaIO\textsubscript{4} reaction and β-elimination (Alefelder et al., 1998) was performed by adding 13.5 µl (20 µg) total RNA in water to 4 µl 5x borate buffer (148 mM boric acid, pH 8.6) and 2.5 µl freshly dissolved 200 mM NaIO\textsubscript{4} and incubating for 10 min at room temperature. 2 µl of glycerol was added to quench unreacted NaIO\textsubscript{4} and incubated for an additional 10 min at room temperature. Samples were dried by centrifugation under vacuum for 1 h at room temperature, then dissolved in 50 µl 1x borax buffer (30 mM borax and 30mM boric acid, 50 mM NaOH, pH 9.5) and incubated for 90 min at 45°C. 20 µg of glycogen was added, the RNA was precipitated at −70°C for 1 hour, the precipitate collected by centrifugation and then dissolved in the denaturing gel loading buffer.
buffer. Enzymatic probing of RNA 5’ ends was performed by incubating the RNA with 2 U calf intestinal phosphatase in 1x restriction buffer 3 (CIP; New England Nuclear) for 2 h at 37°C. CIP was inactivated by phenol/ chloroform extraction, precipitated in the presence of 20 µg of glycogen carrier for 1 h at –70°C, the pellet collected by centrifugation, and then dissolved in water. Part of the CIP-treated RNA was then treated with T4 polynucleotide kinase in 10x PNK buffer (New England Nuclear) and 10 µM ATP. The reaction was inactivated by phenol/ chloroform extraction, the RNA precipitated in the presence of 20 µg of glycogen carrier for 1 h at –70°C, the precipitate collected by centrifugation, and then dissolved in denaturing gel loading buffer.

The treated total small RNA was resolved by high-resolution polyacrylamide gel electrophoresis (PAGE) and then analyzed by Northern blotting. A single nylon membrane was probed first to detect the roo or Su(Ste) rasIRNA, then the probe was stripped and the membrane re-probed to detect miR-8 or miR-311 as an internal control.

**Quantitative RT-PCR analysis**

0.5 µg total RNA were used to reverse transcribe target sequences using oligo(dT) primer and Superscript III reverse transcriptase (Invitrogen) according to the manufacturer’s directions. The resulting cDNA was analyzed by quantitative RT-PCR performed in a DNA Engine OPTICON 2 (MJ Research, Bio-Rad, Hercules, CA, USA) or iQ5 (Bio-Rad, Hercules, CA, USA) instrument using the QuantiTect SYBR Green PCR kit (QIAGEN) according to manufacturer’s instructions. Relative steady-state mRNA levels were determined from the threshold cycle for amplification using the \(2^{-\Delta\Delta Ct} \) method (Livak and Schmittgen, 2001). Table S1 lists the PCR primer sequences, and Supplementary Fig. 7 indicates their position on each selfish genetic element. Control experiments measuring the change in \(\Delta Ct \) with template dilution demonstrated that the efficiencies of amplification of the target genes and the control (act5C or rp49) were approximately the same.
**CHAPTER III, Table 1. Probes and primers.**

### Probes for Northern hybridization

<table>
<thead>
<tr>
<th>Small RNA detected</th>
<th>Probe sequence (DNA or RNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major antisense <em>Su(Ste)</em> rasiRNA</td>
<td>5’-CCACGCUGUAACACUUCU-3’</td>
</tr>
<tr>
<td>#1 <em>roo</em> antisense rasiRNA</td>
<td>5’-TGGGCTCCGTCAATCTTTAG-3’</td>
</tr>
<tr>
<td>#1 <em>roo</em> sense rasiRNA</td>
<td>5’-CATAAGATATGAAACGGAGCCCA-3’</td>
</tr>
<tr>
<td>#3 <em>roo</em> antisense rasiRNA</td>
<td>5’-TGAGAGTCTGCTATTTGAGA-3’</td>
</tr>
<tr>
<td>#3 <em>roo</em> sense rasiRNA</td>
<td>5’-TTCTTTGAATAGGTACACTCTCA-3’</td>
</tr>
<tr>
<td>#4 <em>roo</em> antisense rasiRNA</td>
<td>5’-TCGACTCAGTGCCACAATAAAT-3’</td>
</tr>
<tr>
<td>#4 <em>roo</em> sense rasiRNA</td>
<td>5’-ATTTATAGGCCACTGAGTGCA-3’</td>
</tr>
<tr>
<td>#5 <em>roo</em> antisense rasiRNA</td>
<td>5’-CGACGTTGTTAGGCGGAGGA-3’</td>
</tr>
<tr>
<td>#5 <em>roo</em> sense rasiRNA</td>
<td>5’-TCCTCCCGCTAAACACCGTCG-3’</td>
</tr>
<tr>
<td>#6 <em>roo</em> antisense rasiRNA</td>
<td>5’-GCTCGGCTTGGAGAATATTAT-3’</td>
</tr>
<tr>
<td>#6 <em>roo</em> sense rasiRNA</td>
<td>5’-ATAATTCTTCTAGGCGGAGGC-3’</td>
</tr>
<tr>
<td>#7 <em>roo</em> antisense rasiRNA</td>
<td>5’-TCTGAGGCATCCGTGGTAAA-3’</td>
</tr>
<tr>
<td>#7 <em>roo</em> sense rasiRNA</td>
<td>5’-TTTACAAACCGGATGCCTCAGA-3’</td>
</tr>
<tr>
<td>miR-8</td>
<td>5’-GACATCTTTACCTGACAGTATTA-3’</td>
</tr>
<tr>
<td>miR-311</td>
<td>5’-UCAGCGCGGUGAAUGUGCAAUA-3’</td>
</tr>
<tr>
<td><em>bantam</em></td>
<td>5’-CAGCTTTACAATGATCCTCACT-3’</td>
</tr>
<tr>
<td>2S rRNA</td>
<td>5’-TACAACCCTCAACCATATGATGTAACGA-3’</td>
</tr>
</tbody>
</table>
## Primers for quantitative PCR

<table>
<thead>
<tr>
<th>Detects</th>
<th>Sequence (forward primer, reverse primer)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ste</td>
<td>GTCAGAGCGTGGCGTGATTGC, GGCCCGAACATCGCTCCGTCC</td>
</tr>
<tr>
<td>Su(Ste)</td>
<td>TTCCGAAGTCAAGCGCTTCAATG, GGAATCTGTTTAATTGCAACAAC</td>
</tr>
<tr>
<td>roo</td>
<td>CGTCTGCAATGTACTGGCTCT, CGGCAGTCCACTTAACCTTCC</td>
</tr>
<tr>
<td>HeT-A</td>
<td>CGCCGCGGACCATCTTCAGA, CGCCGCAGTCGGTTGGAGT</td>
</tr>
<tr>
<td>I-element</td>
<td>GACCAAATAAAAATAATACGACTTC, AACTAATTGCTGGCTTTGATATG</td>
</tr>
<tr>
<td>mdg1</td>
<td>AACAGAAACGCAGCAACAGC, CGTTCCCCATGTCCGTTGAT</td>
</tr>
<tr>
<td>mst40</td>
<td>TTCCCTAAGTCCCTCGCAATC, GCACAGTTTCGCCGCCCATCA</td>
</tr>
</tbody>
</table>
**Immunoprecipitation**

myc-piwi transgenic flies (Cox et al., 2000) were heat shocked at 37°C for 1 h every 12 h for 3 days before dissecting the ovaries. aub/Cyo; UAS-GFP-aub/nos-Gal4-VP16 and aub/aub; UAS-GFP-aub/nos-Gal4-VP16 transgenic flies (Harris and Macdonald, 2001) were generated by crossing aub/Cyo; UAS-GFP-aub with aub/Cyo; nos-Gal4-VP16 flies. aub/Cyo; UAS-GFP-aub/nos-Gal4-VP16 ovaries composed the majority of the tissue isolated. The nanos (nos) promoter drives expression of GFP-Aub early in spermatogenesis, in the tip of the testes, and is sufficient to rescue the Ste silencing defect in aub mutant males (data not shown). Quantitative western blotting (Supplementary Fig. 6A) established that the ratio of myc-Piwi to endogenous Piwi in the myc-piwi, piwi+/myc-piwi, piwi+ flies used was 0.95 ± 0.15 to 1; the ratio of GFP-Aub to endogenous Aub in the aub/Cyo; UAS-GFP-aub/nos-Gal4-VP16 flies was 3.1 ± 0.8 to 1. Since the myc-Piwi-expressing ovaries contained two wild-type and two transgenic myc-piwi genes and the ovaries expressing GFP-Aub contained one mutant aub gene and no more than one wild-type aub gene, each genotype expressed about twice as much total Piwi or Aub protein as wild-type.

For immunoprecipitation, 30 µl Protein A/G agarose (Calbiochem, San Diego, CA, USA) were incubated with 20 µl anti-Ago1 (1B8) (Okamura et al., 2004), 5 µl anti-GFP (B-2, Santa Cruz Biotechnology, Santa Cruz, CA, USA, or 3E6, Qbiogene, Montreal, Quebec, Canada), or 5 µl anti-myc (9E10, Sigma, St. Louis, Missouri, USA) monoclonal antibody for 2 h at 4°C. Next, the agarose beads were washed twice with lysis buffer (30 mM HEPES-KOH, pH 7.4, 100 mM KOAc, 2 mM Mg(OAc)₂, 5 mM DTT, 0.5% [v/v] NP-40). Subsequently, 50 µl of ovary lysate (20 µg/µl total protein) was added and the mixture agitated gently for 4 h at 4°C. The supernatant (50 µl) and antibody bound material were separated by centrifugation at 13,000 rpm for 2 min at 4°C, and the antibody-bound beads washed four times then resuspended in 50 µl lysis buffer. 2
µl of Input (I), Supernatant (S), or Bound (B) samples were subject to Western Blotting. The remaining samples were treated with Proteinase K (1 mg/ml proteinase K, 100 mM Tris–Cl, pH 7.5, 12.5 mM EDTA pH 8.0, 150 mM NaCl, 1% [w/v] SDS) for 1 h at 65°C, extracted with an equal volume of phenol/chloroform, and then precipitated with 3 volumes of 100% ethanol in the presence of 20 µg glycogen carrier. The precipitate was washed once with 80% ethanol and then dissolved in nuclease-free water. 15% of each RNA sample was resolved in a 20% denaturing polyacrylamide gel and analyzed by Northern hybridization.

Western blotting

Samples reserved for Western blotting were added to 20 µl SDS sample buffer (10 mM Tris-Cl, pH 6.8, 2% [w/v] SDS, 100 mM DTT, 10% [v/v] glycerol), boiled for 10 min, and then resolved by SDS-PAGE and transferred to a PVDF membrane (Immobilon-P, Millipore). To detect Ago1, Aubergine, or Piwi, the membrane was incubated for 2 hours at room temperature with anti-Ago1 (1B8), anti-Aubergine (ab17724, Abcam PLC, Cambridge, UK), or anti-Piwi (ab5207, Abcam PLC, Cambridge, UK) antibody diluted 1:1000 into TBST-milk (25 mM Tris-Cl, pH 7.4, 3.0 mM KCl, 140 mM NaCl, 0.05% [v/v] Tween-20, 5% [w/v] non-fat dry milk). Subsequently, the membrane was washed 3 times with TBST and incubated 1 h at room temperature with either goat anti-mouse (Ago1) or goat anti-rabbit (Aubergine and Piwi) HRP-conjugated secondary antibody diluted 1:1000 into TBST-milk, and then developed with SuperSignal West Dura Extended Duration Substrate (Pierce, Rockford, IL, USA).
Acknowledgments

CHAPTER IV
MEASURING THE RATES OF TRANSCRIPTIONAL ELONGATION IN THE FEMALE DROSOPHILA MELANOGASTER GERM LINE
BY NUCLEAR RUN-ON

Summary

We applied nuclear run-on technique to measure changes in transcriptional rates not only for single-copy genes, but also for multi-copy transcriptional units such as repetitive elements and transposons in the female Drosophila melanogaster germ line. Our data indicate that as little as a ~1.5-fold change in the rate of transcription can be detected by this method. Our nuclear run-on protocol likely measures changes in transcriptional elongation, because rates of transcription decline with time, consistent with a low rate of Pol II re-initiation in the isolated nuclei. We find that the retrotransposon gypsy and the repetitive sequence mst40 are silenced posttranscriptionally in fly ovaries.
Introduction

Eukaryotic genomes exist as chromatin (from the Greek *chroma*, meaning colored), a complex of DNA and histone proteins that facilitates efficient DNA packaging and differential regulation of gene expression. Cytologically, chromatin can be divided into heterochromatin, first identified by its dark staining, and euchromatin, which stains more lightly (Zacharias, 1995). Heitz imagined euchromatin to be genetically active and rich in genes (Zacharias, 1995). By contrast, he proposed heterochromatin to be genetically inert and gene-poor. Nearly eight decades later, his prescient assessment stands essentially uncorrected.

Constitutive heterochromatin (‘α-heterochromatin’) comprises tandem arrays of repetitive DNA sequences (‘satellite’ sequences), punctuated occasionally by insertions of transposable elements. Constitutive heterochromatin is believed to organize specialized structures such as centromeres and telomeres that act to maintain genetic stability and ensure the segregation of chromosomes during mitosis and meiosis (Allshire et al., 1995, Bernard et al., 2001, Nimmo et al., 1998, Pidoux and Allshire, 2005, Huang and Moazed, 2006). Heterochromatic states are epigenetically inherited: the DNA packaging state is maintained after replication and mitosis, irrespective of the underlying DNA sequence (Elgin and Grewal, 2003). At the molecular level, all heterochromatic regions contain lysine 9-methylated histone 3 and the evolutionarily conserved heterochromatic protein 1 (HP1; Swi6 in fission yeast) (Huisinga et al., 2006). Heterochromatin has other unusual properties: it replicates late in S-phase, remains condensed throughout the cell cycle, resists recombination, and silences reporter genes inserted within or nearby, a property that underlies position effect variegation (Henikoff, 1992, Abbott et al., 2005, Gilbert, 2002, Schubeler et al., 2002, Mahtani and Willard, 1998, Puechberty et al., 1999, Copenhaver et al., 1999). In contrast, facultative heterochromatin is transcriptionally active, but can adopt the structural and functional
characteristics of heterochromatin under special circumstances. Examples of facultative heterochromatin include the inactive X chromosome of mammals (which is genetically identical to its active sister), the silenced rRNA genes in nucleolar dominance, a phenomenon in which rRNA genes of only one parent are active in animal and plant hybrids, and the developmentally regulated β-globin locus in animals (Heard, 2005, Forrester et al., 1989, Kim and Dean, 2004, Weintraub et al., 1981, Grummt and Pikaard, 2003, Lawrence and Pikaard, 2004, Lawrence et al., 2004, Lewis and Pikaard, 2001).

Between constitutive heterochromatin and euchromatin lies β-heterochromatin, which is less condensed, largely deficient in tandem sequence arrays, filled with copies of numerous transposable element families, and contains few functional genes (Holmquist et al., 1998). Repetitive sequences account for the majority of all heterochromatic sequences in *Drosophila melanogaster*, whose genome is 30% heterochromatic, suggesting they play a central role in heterochromatin assembly. Sequence comparison of the *D. melanogaster* fourth chromosome, which is largely heterochromatic, with the syntenic *Drosophila virilis* chromosome 6, which is euchromatic, suggests that the difference in chromatin packaging reflects the density and distribution of transposable elements (Riddle and Elgin, 2006, Slawson et al., 2006).

The predominance of repetitive elements in heterochromatin suggests that these selfish genetic elements are silenced transcriptionally. Repetitive element silencing is essential for genome integrity, as their transposition is the major source of genome rearrangements (Kazazian, 2004). In plants, animals, and fungi, RNA silencing has been implicated as a major defense against repetitive element transposition (Nolan et al., 2005, Sijen and Plasterk, 2003, Martienssen and Colot, 2001, Kalmykova et al., 2005, Savitsky et al., 2006). In the male and female germ line of *D. melanogaster*, retrotransposons and repetitive sequences are silenced by the repeat associated small RNA (rasiRNAs) pathway (Vagin et al., 2006), an RNA silencing mechanism distinct from both the RNAi and miRNA pathways. Small silencing RNAs 24–30 nt long, rasiRNAs are not only ~3–7
nt longer than siRNAs and miRNAs, they are also chemically different, in that they lack one of the 3’ terminal hydroxyl groups characteristic of animal siRNAs and miRNAs. Consistent with this chemical difference, rasiRNAs may not be produced by either Dicer-1, which makes *Drosophila* miRNAs, or Dicer-2, which makes siRNAs. rasiRNA-directed silencing of repetitive genetic elements requires the putative helicases Spn-E, and Armitage as well as Piwi or Aubergine, members of the Piwi sub-clade of Argonaute family of proteins.

Supporting the view that rasiRNAs act to silence repetitive sequence by reducing their rate of transcription, mutations in *spn-E*, *aubergine* and *piwi* are reported to cause loss of lysine 9 methylation of histone H3 on silenced genes (Pal-Bhadra et al., 2004). Historically, heterochromatin has been regarded as transcriptionally inert, but new evidence in fission yeast suggests that heterochromatin can be transcribed by RNA polymerase II (pol II), then silenced by the post-transcriptional destruction of the nascent transcript, perhaps at its site of transcription (Buhler et al., 2006, Volpe et al., 2002, Djupedal et al., 2005, Kato et al., 2005).

What mechanism silences repetitive elements in the germ line of any organism is unknown. Mutations in *armi* increase the steady-state concentration of repetitive element mRNA in the fly germ line (Vagin et al., 2006). Here, we examined the transcriptional rates of the silenced and desilenced retrotransposon *gypsy* and the repetitive sequence *male-specific transcript 40* (*mst40*) in the *D. melanogaster* female germ line. Contrary to expectation, we find that loss of *gypsy* or *mst40* silencing caused by loss of the Armi protein is not accompanied by a change in the rate of transcription of either of these selfish genetic elements.
Results and Discussion

Nuclear run-on measures relative rates of transcription

Weiss demonstrated for the first time that isolated eukaryotic nuclei retain the ability to synthesize RNA (Weiss, 1960). Such RNA synthesis results from transcript elongation by RNA polymerases, rather than initiation of transcription after nuclear isolation (Cox, 1976). These experiments lead to the establishment of ‘nuclear run-on’ as the standard method for measuring the relative rates of gene transcription (McKnight and Palmiter, 1979, Swaneck et al., 1979). In a prototypical nuclear run-on experiment, intact nuclei are isolated by centrifugation, and then incubated with exogenous ATP, GTP, CTP and α-32P-UTP, incorporating radiolabeled nucleotide into transcripts initiated prior to cell lysis but elongated in the isolated nuclei. Radiolabeled nuclear RNA is isolated, and then hybridized to strand-specific RNA probes immobilized on a nylon filter. The specificity of hybridization is established by comparing the hybridization signals for genes of interest to those of control genes whose transcription is presumed to be invariant under the conditions compared (Fig. 1).

We sought to apply the nuclear run-on method to germ line tissue from flies. To begin our study, we first tested if nuclear run-on of nuclei isolated from dissected D. melanogaster ovaries accurately reflects changes in transcriptional rates. We compared the apparent transcriptional rates of the RpL32 gene (commonly known as rp49) in nuclei isolated from ovaries of wild-type flies containing two copies of rp49 (2x rp49) to that in mutant Df(3R)L127/TM6; Dp(3;1)B152/Dp(3;1)B152 flies in which one copy of the rp49 locus on the third chromosome was deleted and each X chromosome contained an additional copy of the locus, for a total of three copies of rp49 (hereafter, 3x rp49). As a control, we used the gene act5C. Hybridization to firefly luciferase (Pp Luc) antisense transcript, whose sequence shares little similarity with any gene in D. melanogaster, provided a measure of the non-specific background. The steady-state mRNA level in the
Dissect flies and isolate ovaries

Break cells and purify nuclei

Complete transcription of active genes in presence of ATP, CTP, GTP, & $^{32}$P $\alpha$-UTP

Purify nuclear RNA

Hybridize nuclear RNA to immobilized RNA probes to detect transcription of specific genes or elements
**Figure 1.** Schematic of nuclear run-on analysis.
ovaries of the 3x rp49 flies was greater by a factor of 1.84 ± 0.39 (average ± standard deviation) than in the ovaries of the 2x rp49 flies (Fig. 2A). Transcription of rp49 in the ovaries with three copies of the gene was greater by a factor of 1.75 ± 0.12 (average ± standard deviation for three independent trials) than in the ovaries from flies bearing two copies of the rp49 gene (Fig. 2B). We conclude that the nuclear run-on assay readily detected a ~1.5-fold change in the transcriptional rate of a housekeeping gene in D. melanogaster ovaries.

A central assumption in this assay is that only transcripts initiated by RNA pol II prior to cell lysis are extended during the radiolabeling reaction. If this assumption is justified, the amount of α-32P-UTP incorporated in an elongating transcript should decline with time, as individual molecules of RNA pol II complete transcription but do not initiate new rounds of RNA synthesis in vitro.

Labeled transcripts were isolated and hybridized to strand-specific RNA probes for four different repetitive sequences: two LTR-retrotransposons (roo and gypsy) and two non-LTR-retrotransposons (I-element and HeT-A). Probes for two single-copy genes, act5C and rp49, served as controls. (The positions of the primers used to make PCR templates for RNA probe transcription by T7 RNA polymerase are diagrammed in Fig. 3). If no re-initiation occurs during the run-on reaction, transcription rates should saturate when there are no more unfinished transcripts left to label. We find that after ~40 min, transcriptional rates saturated, consistent with RNA pol II having elongated all available transcripts without re-initiating new rounds of transcription (Fig. 4).

**Measuring the rates of repetitive element transcription**

Hybridization-based nuclear run-on was conceived to measure the relative transcriptional rates of single-copy genes (McKnight and Palmiter, 1979). Because repetitive elements are multi-copy, we were concerned that their high aggregate rates of RNA synthesis might saturate the immobilized RNA probes, preventing our detecting
CHAPTER IV, Figure 2

**A**

**relative steady-state mRNA**

- **2x rp49**
- **3x rp49**

**B**

**relative transcriptional rate**

- **2x rp49**
- **3x rp49**
**Figure 2.** Nuclear run-on analysis can detect small changes in transcriptional rates. (A) Relative steady-state mRNA concentration in ovaries from flies with two or three copies of the *rp49* gene. (B) Nuclear run-on assay in ovaries from flies with two or three copies of the *rp49* gene. The assay readily detected the ~1.5-fold change in the transcription of *rp49* between the wild-type *rp49/rp49* (2x *rp49*) and the Df(3R)L127/TM6; Dp(3;1)B152/Dp(3;1)B152 (3x *rp49*) flies.
**roo**

- ORF: 9,093 bp
- 5' LTR: 7,469 bp
- 3' LTR: 6,231 bp

**gypsy**

- ORF: 7,469 bp
- 5' LTR: 6,231 bp
- 3' LTR: 6,083 bp

**I-element**

- ORF1: 1387 bp
- ORF2: 4,730 bp

**HeT-A**

- ORF: 6,083 bp

**mst40**

- ORF: 1387 bp
Figure 3. Position of primers used to prepare T7 RNA polymerase PCR templates to synthesize the antisense RNA probes used to detect repetitive elements.
CHAPTER IV, Figure 4

---

**roo**

![Graph of roo relative transcription over time (min)]

**rp49**

![Graph of rp49 relative transcription over time (min)]

**I-element**

![Graph of I-element relative transcription over time (min)]

**gypsy**

![Graph of gypsy relative transcription over time (min)]

**act5C**

![Graph of act5C relative transcription over time (min)]

**HeT-A**

![Graph of HeT-A relative transcription over time (min)]
Figure 4. Nuclear run-on likely measures the rate of transcription without substantial re-initiation by RNA polymerase II, because transcriptional rates saturate with time.
changes in transcriptional rates. To exclude this possibility, we performed nuclear run-on using \textit{armi} homozygous mutant ovaries, in which transposon silencing is derepressed. 100, 50, 25, and 10\% of the $^{32}$P-radiolabeled RNA was hybridized to 5 µg of immobilized RNA probes for the transposons \textit{roo}, \textit{gypsy}, \textit{I-element}, and \textit{HeT-A}, the repetitive sequence \textit{mst40}, and the single-copy gene \textit{act5C}. For all six genes, the rate of transcription decrease linearly with dilution, even for \textit{HeT-A} retrotransposon, whose steady-state transcript levels rise \~170 fold in the absence of \textit{Armi} (Vagin et al., 2006) (\textbf{Fig. 5}). We conclude that our experimental conditions can accurately measure repetitive element transcriptional rates despite their high copy number.

\textbf{Repetitive elements are transcribed by RNA pol II}

To establish that our assay detects RNA synthesis, we conducted the run-on in the presence of three RNA chain terminators: 3’-deoxy ATP, 3’-deoxy CTP, and 3’-deoxy GTP. Consistent with RNA synthesis, run-on transcription decreased dramatically in the presence of the RNA chain terminators for all six genes examined, \textit{act5C}, \textit{rp49}, \textit{gypsy}, \textit{roo}, \textit{I-element}, and \textit{HeT-A} (\textbf{Fig. 6}).

In \textit{D. melanogaster}, 10 µg/ml \textit{α}-amanitin inhibits transcription by RNA pol II, but not by RNA pol I or III. To test if RNA pol II predominantly transcribes repetitive elements in \textit{D. melanogaster}, the nuclear run-on assay was conducted on nuclei preincubated with 10 µg/ml \textit{α}-amanitin for 10 min on ice. Hybridization signals were compared with those from mock-preincubated nuclei. \textit{α}-amanitin reduced repetitive element transcription by an amount comparable to the reduction observed for the single-copy control genes \textit{act5C} and \textit{rp49} (\textbf{Fig. 6}). We conclude that, in flies, repetitive elements, like protein-coding genes, are transcribed by RNA pol II.

\textbf{Repetitive elements in \textit{D. melanogaster} ovaries are silenced posttranscriptionally}

In \textit{armi} homozygous mutants, the steady-state RNA concentration increases by a
Figure 5. Nuclear run-on can detect the rates of repetitive element transcription. Dilution analysis detects no saturation of signal when registering repetitive element transcription in nuclear run-on assay.
CHAPTER IV, Figure 6

- $10 \mu g/ml \alpha$-amanitin
- $3'$-dA, $3'$-dC, $3'$-dGTP
- $10 \mu g/ml \alpha$-amanitin

mock treatment (normalized for each element or gene)

relative transcription

$I$-element  gypsy  roo  HeT-A  act5C  rp49
Figure 6. Repetitive element transcripts are largely RNA polymerase II products.
factor of three for the LTR-retrotransposon gypsy and more than 8-fold for mst40 (Vagin et al., 2006). mst40 sequences are located in region 40, at the base of chromosome 2L, close to or within the β-heterochromatin (Russell and Kaiser, 1994). gypsy element insertions cause frequent mutations in D. melanogaster (Peifer and Bender, 1988). Interestingly, studies of reversion of gypsy-induced mutations demonstrated that it is not simply the insertion of gypsy DNA that causes the mutant phenotype, because most gypsy-induced phenotypes can be suppressed by mutations in suppressor of Hairy wing [su(Hw)] (Peifer and Bender, 1988). The 5′ UTR of gypsy contains an insulator sequence that binds Su(Hw) protein and is the only part of gypsy required to block the interaction of enhancers with the promoters they regulate. Insertion of this insulator sequence between an enhancer and a promoter by insertion of a gypsy element uncouples the enhancer from the promoter (Gause et al., 2001).

gypsy element can only be mobilized in male and female progeny of mothers which both contain active copies of gypsy and are homozygous for permissive flamenco (flam) alleles (Sarot et al., 2004, Chalvet et al., 1998). What protein-coding genes, if any, reside in flam is unknown.

To determine the mechanism by which Armi silences mst40 and gypsy, we analyzed ovaries from heterozygous and homozygous armi mutant females by nuclear run-on (Fig. 7). We found no significant change in transcription between armi/+ and armi/armi ovaries for either mst40 or gypsy. We conclude that in the female germ line of D. melanogaster, Armi acts to silencing these two repetitive elements are post-transcriptionally rather than transcriptionally.
**Figure 7.** (A) A nuclear run-on experiment representative of the data used for the quantitative analysis in (B). (B) The transcriptional rates of the retrotransposon *gypsy* and the repetitive sequence *mst40* do not change significantly when silencing is lost in homozygous mutant *armi* ovaries. The average (± standard deviation), relative to the transcriptional rate of *act5C* is presented.
Conclusions

Here, we show that nuclear run-on can be used to measure changes in transcriptional rates in the female *D. melanogaster* germ line. Our data indicate that as little as a ~1.5-fold change in the rate of transcription can be detected by this method. Our nuclear run-on protocol likely measures changes in the density of RNA Pol II, because rates of transcription decline with time, consistent with a low rate of Pol II re-initiation in the isolated nuclei. Our data support the view that the nuclear run-on technique can be applied not only to single-copy genes, but also to multi-copy transcriptional units such as repetitive elements and transposons. Surprisingly, we find that the retrotransposon *gypsy* and the repetitive sequence *mst40* are silenced post-transcriptionally in fly ovaries. Whether repetitive elements in general are silenced post-transcriptionally in the fly germ line remains to be established (see Appendix to Chapter IV).
Materials and Methods

The rate of transcriptional elongation was determined by nuclear run-on analysis essentially as described (So and Rosbash, 1997). Briefly, ovaries from 50 to 100 females were dissected with needles into *Drosophila* Ringer’s solution (182 mM KCl, 46 mM NaCl, 3 mM CaCl₂, 10 mM Tris-HCl, pH 7.5) and stored on ice in a 1.5 ml homogenization tube (Kontes, Vineland, NJ, USA) during isolation. Isolated tissues were centrifuged at 2,000 x g at 4°C, the supernatant discarded, and then homogenized with 20-30 strokes of the pestle in 300 µl homogenization buffer (10 mM HEPES-KOH, pH 7.5, 10 mM KCl, 0.8 M sucrose, 1 mM EDTA, 0.5 mM DTT, and 100 µg/ml yeast tRNA (Ambion, Austin, TX, USA) containing one tablet of Complete Mini EDTA-free Protease Inhibitor Cocktail (Roche, Indianapolis, IN, USA) for each 10 ml). An additional 200 µl of homogenization buffer was used to rinse the tube. The rinse and the homogenate were pooled, and then filtered through a Bio-Spin column (Bio-Rad, Hercules, CA, USA).

The filtrate was overlaid on a 500 µl cushion of 1 M sucrose dissolved in 10 mM HEPES pH 7.5, 10 mM KCl, 10% glycerol, and 1 mM EDTA plus one tablet of Complete Mini EDTA-free Protease Inhibitor Cocktail (Roche) for each 10 ml and centrifuged at 10,000 x g for 10 min at 4°C. After discarding the supernatant, the nuclei were resuspended in 500 µl of Nuclear Resuspension (NR) buffer [40 mM HEPES-KOH, pH 8.0, 25% glycerol, 5 mM magnesium acetate, 1 mM DTT, and 0.1 mM EDTA containing for each 10 ml one tablet of Complete Mini EDTA-free Protease Inhibitor Cocktail (Roche)] and centrifuged at 4°C for 2 min at 8,000 x g.

The pellet was resuspended in 195 µl of NR buffer, and the reaction initiated by adding 50 µl of 5X reaction mixture (5 mM magnesium acetate, 750 mM KCl, 2.5 mM each ATP, CTP and GTP, 10 mM DTT, 1 µl RNasin (Promega), 50 mM creatine phosphate, and 60 µg/ml creatine kinase) and 5 µl of ³²P α-UTP (6,000 Ci/mmol, 40 µCi/µl, MP Biomedicals, Irvine, CA, USA). After incubation at 23°C for 30 min, the
reaction was stopped by adding 25 µl of RNase-free DNase (RQ1, Promega) and incubating for 5 min at 37°C. Proteins were digested by adding 20 µl of 15X proteinase buffer (7.5% SDS, 150 mM EDTA) and 3 µl of 10-mg/ml proteinase K, and then incubating at 37°C for 30 min. After extraction with phenol/chloroform, the RNA was precipitated with two volumes ethanol. The precipitate was recovered by precipitation, washed with 80% (v/v) ethanol, air dried, and resuspended in 50 µl H2O. Unincorporated nucleotides were removed by two consecutive rounds purification using mini Quick Spin Columns for RNA (Roche) according to the manufacturers instructions, and then used for hybridization as described below.

Labeled RNA from isolated nuclei was hybridized to immobilized strand-specific in vitro transcribed RNA probes for firefly luciferase (Photinus pyralis; Pp luc), roo, I-element, HeT-A, mst40, gypsy, rp49 and act5C. RNA probes were transcribed with T7 RNA polymerase from PCR templates prepared with the oligonucleotides reported in Table 1. Pp luciferase was transcribed from the pGL-2 Control vector (Promega) using the primers: 5´-gcg taa tac gac tca cta tag GAG AGG AAT TCA TTA TC-3´ and 5´-GAA GAG ATA GCC CTG GTT CCT G-3´. 5 µg of each in vitro transcribed RNA probe was denatured in 500 µl of ice-cold 10 mM NaOH and 1 mM EDTA, transferred to Hybond-XL membrane (Amersham) using a Bio-Dot SF Microfiltration Apparatus (Bio-Rad) according to manufacturer’s instructions, and immobilized on the membrane by UV irradiation (200 µjoules/cm; Stratalinker, Stratagene, LaJolla, CA, USA). The membrane was pre-hybridized in Church buffer (Church and Gilbert, 1984) for 1 h at 65°C and hybridization carried out overnight at 65°C. After hybridization, membranes were washed twice with 2x SSC/0.1% (w/v) SDS for 30 min at 65°C and analyzed by phosphorimagery (Fuji, Tokyo, Japan).
CHAPTER IV, Table 1. Primers used to generate T7 RNA polymerase transcription templates for production of antisense RNA probes.

<table>
<thead>
<tr>
<th>Element or gene</th>
<th>Primer sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>roo 5’ primer</td>
<td>taatacgactcactattagggagacG</td>
</tr>
<tr>
<td>3’ primer</td>
<td>ccacGGAGGGTTTGATTTAGGGACAGT</td>
</tr>
<tr>
<td>3’ primer</td>
<td>AGCAGAAGCAGCAACAGCAGTAG</td>
</tr>
<tr>
<td>HeT-A 5’ primer</td>
<td>taatacgactcactattagggagaccacGGAGAAGATCGCTGTCTG</td>
</tr>
<tr>
<td>3’ primer</td>
<td>GACACGCGAAAAAGCGAAC</td>
</tr>
<tr>
<td>I-element 5’ primer</td>
<td>taatacgactcactattagggagacT</td>
</tr>
<tr>
<td>3’ primer</td>
<td>CATCAACACAGCCCAATTGAC</td>
</tr>
<tr>
<td>act5C 5’ primer</td>
<td>gcgtaatacgactcactatagggT</td>
</tr>
<tr>
<td>3’ primer</td>
<td>GGTAGGTTGTCTCGTGGAATGC</td>
</tr>
<tr>
<td>3’ primer</td>
<td>GGCCACCGTGAGAAGATGAC</td>
</tr>
<tr>
<td>mst40 5’ primer</td>
<td>taatacgactcactattagggagacAAGATCGCTTTCGATCTAC</td>
</tr>
<tr>
<td>3’ primer</td>
<td>TCAAATCAGACGAAGTTCAAGG</td>
</tr>
<tr>
<td>rp49 5’ primer</td>
<td>gcgtaatacgactcactataaggg</td>
</tr>
<tr>
<td>3’ primer</td>
<td>TTACTCGTTCTCTTGAGAAG</td>
</tr>
<tr>
<td>3’ primer</td>
<td>GACCACCTGCCCAGCATACAG</td>
</tr>
</tbody>
</table>
Acknowledgments

We would like to thank members of Zamore lab for helpful discussions and comments on the manuscript.
Appendix to Chapter IV

Mutations in *armi* and *spn-E* increase the steady-state concentration of repetitive element mRNA in the fly germ line (Vagin et al., 2006). Nuclear run-on technique was used to measure changes in transcriptional rates for repetitive elements and transposons in both male and female *Drosophila melanogaster* germ line. We examined the transcriptional rates of the silenced and desilenced retrotransposons in ovaries and repetitive sequence *Stellate* in testes of *D. melanogaster*. We find that similar to *gypsy* and *mst40*, desilencing of *I-element*, *roo*, *HeT-A* (*Fig. A1*) or *Stellate* (*Fig. A2*) caused by loss of the *Armi* or *Spn-E* protein is not accompanied by a change in the rate of transcription of either of these selfish genetic elements. We conclude that repetitive elements are silenced posttranscriptionally in fly ovaries and testes.
Appendix to CHAPTER IV, Figure 1

A

<table>
<thead>
<tr>
<th>l-elem</th>
<th>roo</th>
<th>HeT-A</th>
<th>Pp Luc</th>
<th>act5C</th>
</tr>
</thead>
<tbody>
<tr>
<td>armi/+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>armi/armi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B

- **armi/+**
- **armi/armi**

Relative transcription

<table>
<thead>
<tr>
<th>l-element</th>
<th>roo</th>
<th>HeT-A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Bars indicate relative transcription levels.
Figure A1. (A) A nuclear run-on experiment representative of the data used for the quantitative analysis in (B). (B) The transcriptional rates of the retrotransposons I-element, roo, and HeT-A do not change significantly when silencing is lost in homozygous mutant armi ovaries. Firefly luciferase (Pp luc) served as a negative control. The average (± standard deviation), relative to the transcriptional rate of act5C is presented.
Appendix to CHAPTER IV, Figure 2

A

Stellate  Pp Luc  act5C

spn-E/+  

spn-E/spn-E

B

\[
\begin{align*}
\text{relative transcriptional rates} \\
\text{spn-E/+} & \quad \text{spn-E/spn-E}
\end{align*}
\]

![Graph showing relative transcriptional rates for Stellate expressing spn-E/+ and spn-E/spn-E genotypes. The graph depicts higher expression in spn-E/+ samples compared to spn-E/spn-E.](image)
**Figure A2.** (A) A nuclear run-on experiment representative of the data used for the quantitative analysis in (B). (B) The transcriptional rates of the *Stellate* do not change significantly when silencing is lost in homozygous mutant *spn-E* testes. Firefly luciferase (*Pp luc*) served as a negative control. The average (± standard deviation), relative to the transcriptional rate of *act5C* is presented.
CHAPTER V

GENERAL CONCLUSIONS

Here, we show that in *S. pombe*, similarly to plants and animals, introduction of dsRNA induces sequence-specific RNAi response. dsRNA was derived from a long hairpin corresponding in sequence to the full-length *gfp* transgene integrated downstream from the *adh1* gene on chromosome 3. The hairpin was expressed episomally from a plasmid under the control of a strong *nmt1* promotor. Using RNAse protection assay we detected the hairpin expression in wild type *S. pombe*. Only dsRNA could trigger GFP silencing analysed by FACS. Neither antisense nor sense strand of GFP hairpin alone was able to induce the response. This feature makes silencing in *S. pombe* different from quelling in *Neurospora*, because quelling can be induced when a transgene with a sequence homologous to an endogenous gene is integrated in the genome in either sense or antisense orientation (Romano and Macino, 1992).

In *S. pombe*, the silencing of GFP was caused by the decrease in the steady-state level of the *adh1:gfp* mRNA. Silencing was specific to *gfp*, because neighboring gene *aph*, situated ~623 bp from the *adh1:gfp* construct was not affected. In silenced *S. pombe*, the steady-state level of the *adh1:gfp* mRNA decreased without changes in its transcription determined by nuclear run-on assay, suggesting that GFP was silenced posttranscriptionally. These findings were corroborated by the demonstration that the occupancy of the *adh1:gfp* locus by RNA Pol II in wild type fission yeast and upon introduction of the GFP hairpin did not change. The posttranscriptional silencing in *S. pombe* depended on conserved proteins Ago1, Rdpl, and Dcr1, which are required for RNAi in other organisms. RNAi is mediated by siRNAs, which are the products of
processing of long dsRNA by Dicer. We detected the accumulation of GFP siRNAs in both wild type and \textit{adh1:gfp} target strains expressing the GFP hairpin, suggesting that the presence of targeted mRNA was not required for production and/or accumulation of siRNAs. This feature makes the RNA silencing in \textit{S. pombe} more similar to RNAi in flies than in worms. In worms, primary siRNAs produced from initial dsRNA trigger in the absence of target mRNA are very low abundant. In the presence of targeted mRNA, RdRP produces additional dsRNAs, which are processed into secondary siRNAs. Due to this amplification step, secondary siRNAs are much more abundant than primary siRNAs (Sijen et al., 2001). We find that in \textit{S. pombe}, siRNA abundance depends on Dcr1 and Rdp1, suggesting that Rdp1 could use the GFP hairpin as a template. Why either sense or antisense strand of the hairpin alone could not induce GFP silencing? We propose that during the initial step of GFP silencing, dsRNAs generated by transcription of the GFP hairpin are recognized and cleaved by Dcr1 to produce primary siRNAs. It is possible that Dcr1 preferentially targets Rdp1-generated dsRNAs compared to the hairpin mRNAs. In this case, low abundant primary siRNAs are incorporated into an unidentified complex that mediates the recruitment of other RNAi components such as Rdp1 and, possibly, Dcr1 to the GFP hairpin transcripts to produce secondary siRNAs. Secondary siRNAs are incorporated into RISC-like complex to direct silencing in trans of the holomologous \textit{adh1:gfp} locus. Similarly, endogenous centromeric \textit{dg} and \textit{dh} repeats are transcribed by RNA Pol II to produce both forward and reverse noncoding RNAs (Volpe et al., 2002). siRNAs are incorporated into RITS complex that mediates the recruitment of heterochromatin assembly factors such as Clr4 and Swi6 to homologous targeted loci, which in turn promotes binding of additional RITS complexes (Sugiyama et al., 2005). Thus, cooperation of components of RITS complex and heterochromatin assembly
factors allows the RNAi machinery to operate in cis to process nascent centromeric transcripts into siRNAs. Components of RITS complex, however, Chp1 and Tas3, as well as heterochromatic protein Swi6 were dispensable for silencing induced with the GFP hairpin, which is possibly the reason why the RNAi machinery could mediate silencing in trans. The GFP silencing did not increase H3 K9 methylation levels at the adh1:gfp locus measured by chromatin immunoprecipitation assay, suggesting that in S. pombe, posttranscriptional silencing could be established in the absence of heterochromatin formation. Introduction of the GFP hairpin decreased H3 K9 methylation levels at the centromeric locus. It is possible that in the presence of the GFP hairpin, the only H3 K9 methyltransferase in S. pombe Clr4 is recruited away from the centromeric locus. This hypothesis is consistent with our finding that deletion of clr4 had an effect on posttranscriptional silencing of the adh:gfp locus and a previous report that siRNAs did not accumulate to normal levels in the clr4 mutant strain (Motamedi et al., 2004). It is possible that Clr4 functions in siRNA biogenesis in both TGS and RNAi. That RNAi in S. pombe does not require Chp1, Tas3 or Swi6 suggests that a complex distinct from RITS mediates siRNA-directed target mRNA degradation in the cytoplasm. Indeed, the RNAi components in fission yeast were found in complexes localized to both nucleus and cytoplasm (Carmichael et al., 2006, Noma et al., 2004). It would be of great interest to determine the function of cytoplasmic RNAi complexes. Thus, RNAi proteins in S. pombe, Ago1, Dcr1, and Rdp1 can function in at least two distinct silencing pathways, suggesting that, in spite of their functional specialization, Argonaute proteins can mediate gene silencing by multiple mechanisms. For example, a single Argonaute protein in S. pombe can regulate mRNA abundance through posttranscriptional destruction of targeted sequences and/or heterochromatin-base transcriptional inhibition. We hypothesize that
some of the multiple Argonaute paralogs in other organisms, are biochemically redundant and their specialized function in gene silencing can be explained by distinct patterns of their spatial or temporal expression or intracellular localization.

In *Drosophila*, heterochromatin comprises about 30% of the genome. The heterochromatin contains mainly satellite sequences, transposons, and few genes, suggesting that repetitive elements are required for heterochromatin formation. On the other hand, in certain tissues, germ line, for example, repetitive elements have to be silenced to ensure uncompromised transmission of genetic information to the next generation. Thus, transcriptional silencing of repetitive elements through heterochromatin formation could be the way to prevent their mobilization and recombination in the germ line.

In *Drosophila*, miRNA and RNAi pathways are both genetically and biochemically distinct. Dcr-2 generates siRNAs, whereas the Dcr-1/Loqs complex produces miRNAs. Two Argonaute proteins in flies, Ago1 and Ago2, belong to the Ago subfamily, whereas the Piwi subfamily includes Aub, Piwi, and Ago 3. siRNAs and miRNAs regulate gene expression by endonucleolytic cleavage or translational repression of target mRNAs through effector complexes containing Ago2 or Ago1, respectively. The third class of small RNAs, rasiRNAs, is produced endogenously in the *Drosophila* germ line and has been proposed to silence retrotransposons by an RNAi-like mechanism (Aravin et al., 2001, Kogan et al., 2003, Savitsky et al., 2006).

To compare the biogenesis of siRNAs produced endogenously in the soma of flies expressing a long hairpin targeting *white* gene from that of rasiRNAs in the germ line, we analyzed the accumulation of both classes of small RNAs by tiling microarrays. We determined that in vivo in wild-type flies, in accordance with previous in vitro data, both
strands of the trigger dsRNA were processed by Dcr-2 and accumulated as siRNAs. Fourier-transform analysis of these siRNAs revealed a pronounced ~22 nt periodicity, demonstrating for the first time that Dcr-2 in flies, similarly to human Dicer, processed dsRNA substrates processively from the ends.

Microarray analysis of Su(Ste) rasiRNAs in fly testes and roo rasiRNAs in ovaries revealed that rasiRNAs were accumulated exclusively in antisense polarity to their respective target mRNAs, with little or no phasing detectable by Fourier-transform analysis. This suggests that biogenesis of rasiRNAs is different from that of siRNAs.

We examined the genetic requirements for biogenesis of rasiRNAs in both male and female germ line of Drosophila and silencing of 8 different selfish elements, including three LTR retrotransposons, two non-LTR retrotransposons, and three repetitive sequences.

Production of rasiRNAs did not require Dcr-2, R2D2, Dcr-1 or Loqs, suggesting that biogenesis of rasiRNAs is distinct from that of siRNAs and miRNAs. We cannot exclude that Dcr-2 and Dcr-1 act redundantly in the production of rasiRNAs like Dicers in Neurospora. In Neurospora, however, both Dicers produce siRNAs of the same size (Catalanotto et al., 2004). rasiRNAs are longer than siRNAs and miRNAs. They are 25-29 nt. The development of dcr-2 dcr-1 double mutants should unambiguously resolve this issue. In addition, our data reveal that rasiRNAs have a blocked hydroxyl group at the 3’ end. This group, possibly a methyl group as in endogenous plant siRNAs and miRNAs, could be added during production steps of rasiRNA biogenesis or later, during maturation steps. Further studies will reveal if this modification is indeed the methyl group, and if Drosophila homolog of HEN1, which is responsible for depositing of methyl group to the 2’ hydroxyl of the 3’ terminal nucleotide in Arabidopsis, plays similar role in the biogenesis of rasiRNAs. It is possible that longer size of rasiRNAs and the modification at the 3’ end facilitate the loading of rasiRNAs into distinct Argonaute proteins. Indeed,
we show that both accumulation of rasiRNAs and silencing of endogenous repeats require Piwi and Aubergine, proteins of Piwi subfamily of Argonaute proteins. Unlike ago1 and ago2, which are ubiquitously expressed, aub, piwi, and ago3 mRNAs are enriched at the posterior pole of the early embryo and subsequently only in the developing germ line (Williams and Rubin, 2002, Cox et al., 1998, Harris and Macdonald, 2001). Although it has been demonstrated that Piwi could be programmed with single-stranded 21-nt siRNA to direct endonucleolytic target cleavage (Saito et al., 2006), we find that Piwi and Aubergine bind only rasiRNAs, whereas Ago1 binds exclusively miRNAs. Since rasiRNA pathway is genetically distinct from both siRNA and miRNA pathway, it is possible that rasiRNA pathway is separated from siRNA and miRNA pathway. An enzyme/complex, which produces rasiRNAs, might load them exclusively into Argonaute proteins of Piwi subfamily. Even if such separation does exist, it is not absolute. rasiRNA accumulation and repetitive element silencing require armi. armi mutant ovaries cannot assemble Ago2-RISC or support RNAi in vitro (Tomari et al., 2004a). Armi encodes a non-DEAD box helicase (Cook et al., 2004) homologous to the Arabidopsis thaliana protein SDE3, which is required for RNA silencing triggered by transgenes and some viruses (Dalmay et al., 2001), and depletion by RNAi of the mammalian Armi homolog, Mov10, blocks siRNA-directed RNAi in cultured human cells (Meister et al., 2005). Normal accumulation of rasiRNA and robust repetitive element silencing also requires the putative helicase, Spn-E, a member of the DExH family of ATPases required for RNAi in activated Drosophila oocytes (Aravin et al., 2001, Curr Biol, 11, 1017–1027, Stapleton et al., 2001, Kennerdell et al., 2002)

Aub and Piwi proteins are 42% identical, their mRNAs are 51% identical, and the two genes lie next to each other on the left arm of chromosome 2, suggesting they are paralogs that arose by gene duplication. (No genetic mutants are available for ago3, which resides in the centromeric heterochromatin of chromosome 3.) There are several lines of evidence supporting the view that Piwi and Aub play partially redundant roles in
rasiRNA pathway. Firstly both proteins can bind the same roo rasiRNA. Although piwi mutants desilence roo only slightly, and aub mutant ovaries still silence roo mRNA normally, they show reduced accumulation of roo rasiRNA relative to aub/+ heterozygotes. Secondly repetitive element silencing in ovaries requires either Piwi or Aub, which probably reflects the differential patterns of expression of these repetitive elements in respect to these Argonaute proteins. In testes, Ste silencing in germ cells requires Aub but not Piwi, which is expressed exclusively in apical somatic hub cells (Saito et al., 2006, Kalmykova et al., 2005). These ideas remain to be tested, and await the development of piwi aub double mutants, a challenging task because the two genes lie next to each other on the left arm of chromosome 2.

Our data suggest that in flies rasiRNAs are produced by a mechanism that requires neither Dcr-1 nor Dcr-2 and function as guides for Argonaute proteins of the Piwi subfamily, silencing expression of selfish genetic elements through a pathway distinct from both the RNAi and miRNA pathways. One possible reason for the mechanistic and genetic differences—including the length and terminal chemical structure of rasiRNAs—between the rasiRNA and the siRNA and miRNA pathways might be that rasiRNAs act to repress gene expression transcriptionally, rather than post-transcriptional as do animal siRNAs and miRNAs. To test this idea, we performed nuclear run-on experiments for selfish genetic elements in wild type and mutant testes and ovaries. The nuclear run-on method measures the density of RNA Pol II synthesizing RNA from a gene, and therefore can detect changes in transcriptional rate between wild type and mutant flies in which selfish genetic element silencing is disrupted. We examined Su(Ste)-triggered silencing of Ste in testes and roo, HeT-A, I-element, gypsy, and mst40 silencing in ovaries from males and females heterozygous or homozygous for mutations in spnE and armi. For all six selfish genetic elements we examined—two examples of repetitive sequence, two non-LTR retrotransposons, and two LTR-retrotransposons, we could detect no change in transcriptional rate in wild type versus
mutant germ line cells. We conclude that although rasiRNA biogenesis and function is distinct from both the RNAi and miRNA pathways, rasiRNA-direct posttranscriptional silencing of selfish genetic elements. Given the remarkable evolutionary conservation of RNA silencing pathways between flies and mammals, it is tempting to speculate that the human germ line, too, contains a third RNA silencing pathway in which long small silencing RNAs, produced without Dicer, act through Piwi subfamily Argonaute proteins to silence selfish genetic elements or other RNA transcripts posttranscriptionally. Such long small RNAs, piRNAs, have been indeed cloned recently from mouse, rat, and human germ line (Lau et al., 2006, Aravin et al., 2006, Girard et al., 2006, Grivna et al., 2006, Watanabe et al., 2006). They bind exclusively to Piwi subfamily of Argonaute proteins, and originate as genomic clusters, which chromosomal position is conserved between all three tested mammals. Within each cluster piRNAs display extreme strand bias, which makes them similar to rasiRNAs. Although only a small portion of piRNAs corresponds to repetitive elements, it remains to be elucidated if yet to be annotated sequences of piRNAs are also rasiRNAs. Alternatively, it is of great interest to determine if rasiRNAs also arise as clusters it trans from their corresponding targets.
REFERENCES


Dasgupta, S., Fernandez, L., Kameyama, L., Inada, T., Nakamura, Y., Pappas, A., and Court, D.L. Genetic uncoupling of the dsRNA-binding and RNA cleavage activities of


