Identification of Novel (*R*NAi *de*ficient) genes in *C. elegans*

A DISSERTATION PRESENTED

BY

CHUN-CHIEH G. CHEN

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

September 26, 2006

Interdisciplinary Graduate Program

COPYRIGHT INFORMATION

Approval Page

Identification of Novel (*R*NAi *de*ficient) genes in *C. elegans*

A DISSERTATION PRESENTED

BY

CHUN-CHIEH GISELA CHEN

Approved as to style and content by:

Nicholas Rhind, Chair of Committee

William Theurkauf, Member of Committee

Phillip Zamore, Member of Committee

Kirsten Hagstrom, Member of Committee

Gregory Hannon, Member of Committee

Craig C. Mello, Dissertation Mentor

Anthony Carruthers, Dean of the Graduate School of Biomedical Sciences

ACKNOWLEDGEMENTS

It has been a wonderful six-year time in Worcester and I would like to thank everyone who made my graduate studies exciting and memorable. I would like to thank my advisor, Craig Mello, for the opportunity to join his lab and work on such a fascinating project. I am also greatly appreciative of Craig's comments and encouragements during my graduate studies.

In addition, I would like to say thank you to the past and present members of the Mello lab for making the lab a fun place to work. These people include Tuba Bas, Pedro Batista, Yanxia Bei, Daniel Brownell, Daniel Chaves, Julie Claycomb, Darryl Conte Jr., Thomas Duchaine, Ahmed Elewa, Alla Grishok, Weifeng Gu, Takao Ishidate, Jennifer Keys, Soyoung Kim, Na Li, Lisa Maduzia, Maral Markisian, Jennifer McCullough, Kuniaki Nakamura, KaMing Pang, TaeHo Shin, Masaki Shirayama, Martin Simard, Martha Soto, Chris Trzepacz, Bill Tsai, Yingdee Unhavaithaya, Jessica Vasale and Erbay Yigit. Hiroaki and Darryl had helped me to get started in the lab; Daniel and Pedro have been wonderful baymates; and I can't thank enough for the friendship and help of Lisa and Julie and a very especial thank you to Julie for helping me with my thesis writing. I would also like to express my gratitude to Jean and Darla for their excellent administrative assistance.

I would also like to express my gratitude to the members of my thesis committee, Bill Theurkauf, Phil Zamore and Nick Rhind who had made very helpful comments

during the formation of this thesis and Kirsten Hagstrom and Greg Hannon for their participation in the final thesis defense.

I am appreciative of my classmates, Gen, Julie and Jess for their friendship in the beginning of the graduate school when I first came to the United States, of the Taiwanese Student and Scholar Association for their support and friendship, and of the Seven Hills Symphony at UMass for providing me with a chance to play music in my leisure time.

I would like to thank my family, my father Jen-Po Chen, Yen-Chu Ting and my two sisters Chun-Wei Chen and Chun-Ti Chen. My parents were supportive of me ever since the first day I made the decision to pursue my graduate studies and had encouraged me during the process. My dearest sisters were living near me and together we had spent the most wonderful time together. They also helped me with labeling the plates for the genetic screens on a Friday night before we went out. Finally, I would like to thank Ying, who used to be my labmate and is my husband-to-be. I thank him for his support, comments on my work, and the endless love from him expressed in the form of cooking the most delicious meals everyday.

Abstract

RNA interference or RNAi was first discovered as an experimental approach that induces potent sequence-specific gene silencing. Remarkably, subsequent studies on dissecting the molecular mechanism of the RNAi pathway reveal that RNAi is conserved in most eukaryotes. In addition, genes and mechanisms related to RNAi are employed to elicit the regulation of endogenous gene expression that controls a variety of important biological processes. To investigate the mechanism of RNAi in the nematode *C. elegans*, we performed genetic screens in search of *R*NAi *de*ficient mutants (*rde*). Here I report the summary of the genetic screens in search of *rde* mutants as well as the identification of two novel genes required for the RNAi pathway, *rde-3* and *rde-8*. In addition, we demonstrate that some of the *rde* genes, when mutated, render the animals developmentally defective, suggesting that these *rde* genes also function in developmental gene regulation. This work presents novel insights on the components of the RNAi pathway and the requirement of these components in the regulation of endogenous gene expression.

TABLE OF CONTENTS

Chapter I

Introduction

RNA Interference and Gene Silencing Pathways Utilizing Small RNAs in *C. elegans*

Summary

The ribonucleic acid (RNA) molecules play important roles in diverse biological processes such as messengers and adaptors to translate the genetic code from DNA to protein (mRNAs and tRNAs), and structural and catalytic units of protein synthesis (rRNAs). Recently, mounting evidence has suggested yet another fundamental function for RNA as a regulator of gene expression in the organism. Such regulation can exist within the RNA itself through structural changes or by the utilization of a second RNA molecule of a wide range of sizes. These second RNAs, very often complementary to the regulated target, control gene expression at the level of transcription, translation, and chromatin structure. One type of these conserved regulatory mechanisms employs small RNAs as a sequence-specific guide to elicit gene silencing. Many aspects of the mechanisms of the small RNA generation as well as the function of the small RNAs were revealed through the study of a silencing phenomenon termed RNA interference or RNAi. RNAi was first discovered in *C. elegans* as an experimental tool to down-regulate the expression of specific genes. Subsequent studies demonstrated that RNAi is triggered by double-stranded RNAs and that this double-stranded structure gives rise to the generation of the small RNAs. In this chapter, I present both a summary of our current knowledge and a discussion of future avenues to increase our understanding of these diverse yet molecularly conserved small-RNA pathways. In particular, I will focus on RNAi and related pathways in the nematode *Caenorhabditis elegans*, which is the model system used for the study of RNAi in this thesis.

The mechanism of RNAi in *C. elegans*

In the nematode *C. elegans*, the use of antisense RNAs to block gene expression, presumably through physical interference with the translational machinery, was reported with the surprising observation that the sense strand functioned to the same extent as the antisense strand, with both resulting in a penetrant loss-of-function phenotype (Guo and Kemphues, 1995). This use of antisense RNAs in *C. elegans* to produce a mutant phenotype in a sequence-specific manner was later termed RNA interference or RNAi (Rocheleau et al., 1997). The surprising observation that the sense RNAs also functioned to block gene expression in RNAi argued against the mechanism of RNAi being the physical hybridization between the introduced RNAs and the complementary mRNAs. Subsequent work by Fire *et al.* demonstrated that during RNAi (1) double-stranded RNAs (dsRNAs) could induce a more efficient gene silencing than either the gel-purified sense or antisense single-stranded RNA along, indicating that dsRNAs were the trigger of RNAi; (2) a relatively small amount of dsRNAs can cause the phenotype in an animal and the whole brood of the next generation, suggesting that RNAi is a heritable and most likely amplified process; (3) the level of the target mRNA is drastically decreased (Fire et al., 1998). Since first reported in *C. elegans*, dsRNAs have been shown to induce gene silencing in a variety of organisms. The discovery of RNAi in *C. elegans* opened a new page in the RNA biology where small RNAs play an important role in gene regulation. This section details our current understanding of the mechanism of RNAi in *C. elegans* and discusses those aspects that remain to be elucidated.

The delivery of trigger double-stranded RNAs into *C. elegans*

Introduction of dsRNAs into the cells is the first step in RNAi (Figure I-1). In *C. elegans*, RNAi can be administered by injection of dsRNAs into any part of the body of the worm, by soaking the worms in dsRNAs-containing solutions, by feeding the worms with bacteria expressing dsRNAs, or by expressing dsRNAs in an animal from a transgene (Tabara et al., 1998; Tabara et al., 1999; Timmons et al., 2001). Interestingly, the RNA signal can spread from one tissue to another in the worms, which is a special feature shared by a subset of organisms, including planaria and plants (Newmark et al., 2003; Vaucheret et al., 2001). The transport of the RNAi signal between the cells has been referred to as systemic RNAi or RNA spreading (Tijsterman et al., 2004; Winston et al., 2002). Although several genes required for the transport of dsRNAs have been identified, the detailed mechanism of how dsRNAs are transported and how the RNAi signal is transmitted in different tissues in the worm is not well understood.

sid-1 was the first mutant isolated from a genetic screen that specifically sought mutants that are systemic RNA*i* *d*eficient (*sid*) (Winston et al., 2002). The *sid-1* mutant is resistant to RNAi except when dsRNAs are provided in the target tissue by transgene expression or direct injection, suggesting that *sid-1* functions to transport an RNAi signal. The *sid-1* gene encodes a transmembrane protein and facilitates import of long but not small dsRNAs in both *C. elegans* and when ectopically expressed in Drosophila S2 cell lines, suggesting that SID-1 is involved in transport of dsRNAs (Feinberg and Hunter, 2003). However, several important questions remain about SID-1 function, including

Figure I-1

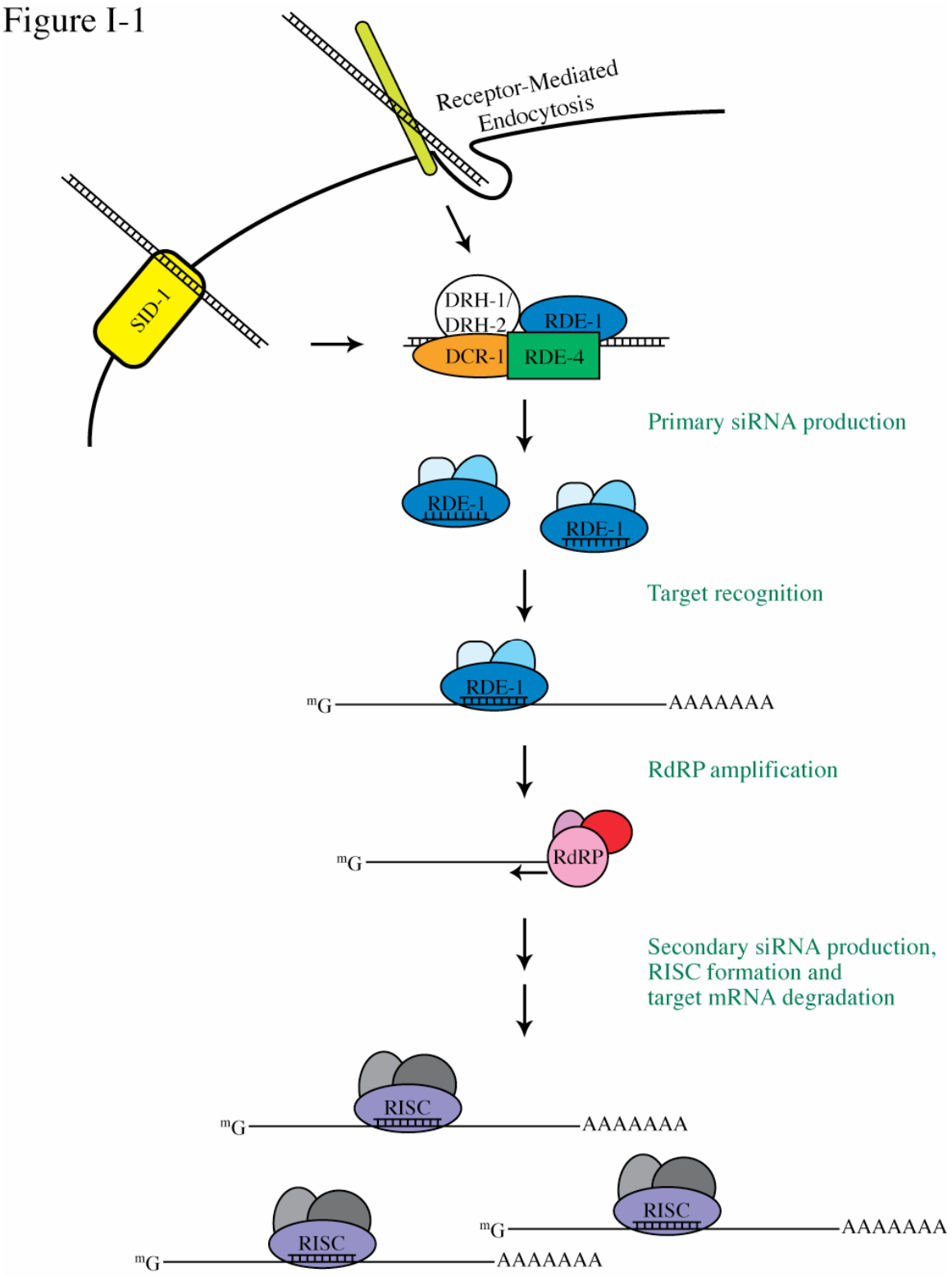

Receptor-Mediated Endocytosis
SID-1
DRH-1/ DRH-2
RDE-1
DCR-1
RDE-4
Primary siRNA production
RDE-1
RDE-1
Target recognition
RDE-1
mG
AAAAAAA
RdRP amplification
mG
RdRP
Secondary siRNA production, RISC formation and target mRNA degradation
RISC
mG
AAAAAAA
RISC
mG
AAAAAAA
RISC
mG
AAAAAAA

Figure I-1 A model for the uptake of dsRNAs into the cell and the RNAi pathway inside the cell in *C. elegans*. See text for details.

how does SID-1 recognize long dsRNAs; what is the mechanism of dsRNA transport; how are the dsRNAs passed onto the RNAi machinery after import; and if there are any cofactors that assist SID-1. It is worth noting that, as revealed by RNAi-based genome-wide screen for genes required for RNAi or the uptake of dsRNAs into the S2 cells, dsRNAs can be internalized independently of SID-1 in Drosophila S2 cells via the Scavenger Receptor-mediated endocytosis pathway (Saleh et al., 2006; Ulvila et al., 2006). Moreover, RNAi knockdown of several gene orthologs identified from one of the screens, including four components of the intracellular vesicle transport, two lipid-modifying enzymes and four genes of unknown function, also affect RNAi in *C. elegans* (Saleh et al., 2006). It remains unclear whether the SID-1 pathway works collaboratively with or independently of the endocytosis pathway. Interestingly, although the SID-1 protein preferentially recognizes long dsRNAs for import in *C. elegans*, over-expression of a putative human *sid-1* homolog in the pancreatic ductal adenocarcinoma cell line, PANC1, results in improved transport of siRNA duplexes and subsequent siRNA-induced gene silencing (Duxbury et al., 2005). As RNAi has been considered a possible method for gene therapy, this finding is of special interest in improving the delivery of nucleic acid of therapeutic potentials.

Another class of genes that were suggested to be involved in transport of RNAi signal includes several *R*NA *s*preading *d*efective (*rsd*) mutants (Tijsterman et al., 2004). These mutants were resistant to RNAi by feeding but sensitive to RNAi by injection and displayed different requirements for tissue specific RNAi. Several *rsd* genes have been

cloned, but their molecular function is unknown, impeding our understanding of if and how these genes contribute to systemic RNAi.

Recently, several members in the ATP-binding cassette (ABC) transporter gene superfamily have been reported to be required for efficient RNAi (Sundaram et al., 2006). ABC transporters are transmembrane proteins that use ATP to translocate small molecules across the membrane and are involved in the export of drugs or toxins from the cell, organelle biogenesis and viral infection. Different sets of transporters are expressed in various tissues (for example, the germline, the intestine, or the muscle) and mutations in an ABC transporter result in a weak RNAi-deficient phenotype in the tissues where that transporter is expressed. The weak RNAi defect could be attributed to the redundant function of several transporters in RNAi or that these ABC transporter genes function indirectly in RNAi. These transporters could translocate RNAi signals in addition to the initial digested dsRNAs or could regulate the cellular metabolism, such as the level of ATPs, for RNAi to function in response to the outer environmental cues. Although all of the components and pathways described above are involved in some aspects of the transport of RNAi signals, further work is required to understand how these pathways integrate to function in systemic RNAi.

The initiation of RNAi

Two genes required for the initiation of RNAi in *C. elegans*, *rde-1* and *rde-4*, were identified from genetic screens in search of *R*NAi *de*ficient (*rde*) mutants (Tabara et al., 1999). Both *rde-1* and *rde-4* are essential for RNAi. In addition, *rde-1* and *rde-4* were

genetically shown to function in an upstream step to initiate an RNAi signal. This signal is required for the inheritance of RNAi in the next generation, but not in a downstream step in which the execution of RNAi occurs in the next generation. Furthermore, once RNAi is initiated, the activity of *rde-1* and *rde-4* are no longer required in the downstream execution step (Grishok et al., 2000). These results suggest that *rde-1* and *rde-4* function at an early step in the RNAi pathway. *rde-1* is a member of the Argonaute protein family, which contains two signature domains, PAZ and PIWI (Tabara et al., 1999). Members of this family were subsequently found to be essential components of the RNA induced silencing complex (RISC), a small RNA-directed nuclease complex that performs endonucleolytic cleavages of the target RNAs (Hammond et al., 2001). *rde-4* encodes a protein with two dsRNA binding motifs near the amino terminus of the protein and was shown accordingly to interact with dsRNAs (Tabara et al., 2002).

In a mechanistically similar silencing pathway in plants, post-transcriptional gene silencing (PTGS), a species of small antisense RNAs complementary to the target RNA was detected by Northern blotting in correlation with ongoing silencing (Hamilton and Baulcombe, 1999). This small RNA species was hypothesized to serve as the sequence-specificity determinant in targeting an mRNA. Examination of RNA species in *C. elegans* in which RNAi was induced by injection of radiolabeled dsRNAs also revealed the presence of small RNAs around 25nt long, suggesting that the accumulation and perhaps the requirement of small RNAs is conserved in these mechanistically related pathways (Parrish et al., 2000). Biochemical characterization of RNAi in Drosophila cell or embryonic lysates further demonstrated that the small RNAs are processed directly

from the initial long dsRNAs (Hammond et al., 2000; Zamore et al., 2000). The enzyme that generates these small RNAs during RNAi, known as small interfering RNAs (siRNAs), from the trigger RNAs was first identified in Drosophila S2 cell lysate as the endonuclease Dicer (Bernstein et al., 2001). Dicer is a large multi-domain protein with conserved helicase motifs at its amino terminus followed by a PAZ domain, two RNase III-like domains and a double-stranded RNA binding motif.

Interestingly, RDE-1 and RDE-4 form a complex with the only homolog of Dicer in *C. elegans*, DCR-1 *in vivo* (Tabara et al., 2002). In addition, two very similar helicase proteins homologous to the helicase domains of DCR-1 were also found in this complex and were named DRH-1 and DRH-2 for Dicer related helicases (Duchaine et al., 2006; Tabara et al., 2002). The analysis of *drh-1* and *drh-2* mutants has not been performed, due to a lack of mutant strains. However, reducing the expression of *drh-1* and *drh-2* by RNAi, referred to as *drh-1*(RNAi) and *drh-2*(RNAi), greatly compromised the ability of RNAi to target a second reporter gene, suggesting that *drh-1* and *drh-2* are also required for the proper function of the RNAi pathway (Tabara et al., 2002). The immuno-complex of RDE-4 interacts with dsRNAs corresponding only to the trigger region, suggesting that the initiation complex may make the first encounter with the exogenously provided trigger RNA. The composition of the initiation complex predicts a model for its action (Figure I-1), in which RDE-4 binds to long trigger dsRNAs that are subsequently processed by DCR-1 into siRNAs (Figure I-1). Because only single-stranded siRNAs are detected in the RDE-1 complex, The DRH-1 and DRH-2 proteins may assist the unwinding of double-stranded siRNAs into single-stranded siRNAs that associate with

the RDE-1 protein (Yigit et al., 2006). The single-stranded siRNA-containing RDE-1 complex then searches for its complementary mRNA target and perhaps direct the first round of endonucleolytic cleavage of the target mRNA. However, consistent with the observation that *rde-1* and *rde-4* only function in the upstream steps of the RNAi pathway, the siRNA-containing RDE-1 complex is not responsible for the degradation of most targeted mRNAs (Sijen et al., 2001). In contrast, the initial round of the mRNA degradation by the RDE-1 complex appears to function in triggering the amplification of the RNAi signal for a downstream effector step where destruction of most targeted mRNAs happens.

Unlike *rde-1*, *rde-4*, *drh-1*(RNAi) and *drh-2*(RNAi) animals, which do not exhibit additional phenotypic defects under laboratory growth conditions, *dcr-1* mutants exhibit developmental defects (Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). *dcr-1* homozygotes display a lethal phenotype due to a failure in processing another class of developmentally essential RNAs termed microRNAs. Recently, a more detailed biochemical examination of DCR-1 protein interactors demonstrated that DCR-1 is present in several distinct protein complexes which participate in different small RNA mediated silencing pathways (Duchaine et al., 2006). This suggests that these different small RNA pathways are insulated from each other as their respective trigger dsRNAs are channeled into distinct DCR-1 subcomplexes.

Amplification of the RNAi signal

A remarkable feature of RNAi in *C. elegans* is that only a very small amount of dsRNA is necessary to silence the expression of a particular gene in an animal (Fire et al., 1998; Kennerdell and Carthew, 1998). For example, it was estimated that ~100 molecules of trigger RNA is sufficient to trigger silencing of ~5000 molecules of target mRNA in each cell for the *unc-22* gene in *C. elegans* (Sijen et al., 2001). In another instance which dsRNAs targeting the gene frizzled were injected into fly embryos, ~30 molecules would be distributed into each cell to cause silencing of the target mRNA which were thought to be present at a higher level (Kennerdell and Carthew, 1998). In addition, *in C. elegans*, the silencing effect can be transmitted to the entire brood of its progeny. The huge difference in the quantity of trigger RNA versus target mRNA suggests that the initial trigger signal may be amplified to elicit silencing. Interestingly, in some silencing phenomena that are mechanistically similar to RNAi, a requirement for members in the RNA dependent RNA Polymerase (RdRP) family has been reported (Cogoni and Macino, 1999; Dalmay et al., 2000; Mourrain et al., 2000; Sijen et al., 2001; Smardon et al., 2000). These phenomena include quelling in Neurospora, post-transcriptional gene silencing (PTGS) in plants, and RNAi in *C. elegans*. Other organisms, including Drosophila and mammals, lack members of this family of RdRPs in their genomes and RdRP activity has not been detected in their RNAi pathways (Schwarz et al., 2002; Stein et al., 2003).

The RdRP activity was predicted to generate double-stranded RNAs that then enter the RNAi pathway. Interestingly, in *C. elegans*, although RNAi is initiated by dsRNAs, the participation of an RdRP homolog is essential (Sijen et al., 2001). The

finding that RdRPs play an essential role in RNAi suggests that the initial exposure of worms to dsRNA may be insufficient to trigger and/or maintain a potent RNAi response and that an amplification process after initiating RNAi is required for potent gene silencing. In light of this, our model is that the initiation complex processes the original trigger dsRNAs into siRNAs, and along with RDE-1, searches for the target mRNA and performs the first round of cleavage (Figure I-1). This step, in turn, triggers an amplification event in which RNAi is potentiated by the generation of more siRNAs to produce additional RISCs through the action of RdRPs.

Analysis of the accumulation of antisense siRNAs in *C. elegans* undergoing RNAi revealed the presence of siRNAs whose sequences correspond to the region immediately upstream of the double-stranded trigger, indicating that these siRNAs are not the result of *dcr-1* directly processing dsRNA triggers (Fire et al., 1998). Additionally, examination of the strand polarity of siRNAs demonstrated that only antisense siRNAs are present. The accumulation of antisense siRNAs also requires the presence of a target RNA, which suggests that mRNAs could be required as a template for RdRPs to generate more siRNAs (Grishok and Mello, unpublished). As these siRNAs most likely arise from a secondary source other than the originally introduced trigger dsRNAs, they are termed secondary siRNAs. In contrast, primary siRNAs are processed from trigger dsRNAs directly introduced to the worms.

In order to test whether the secondary siRNAs are capable of eliciting silencing, a reporter assay, termed a transitive RNAi assay, was developed in *C. elegans* (Sijen et al., 2001). In this assay, two reporter genes are expressed in an animal with one gene sharing

homologous sequences to the trigger dsRNAs. The second target gene bears sequences homologous to part of the first target gene that is not present in the trigger dsRNAs. When exposed to the trigger dsRNA, the silencing of the second target gene is only possible by the secondary siRNAs amplified from the mRNAs that are expressed by the first target gene. Wildtype *C. elegans* worms are capable of eliciting transitive RNAi, suggesting that secondary siRNAs are produced and are functional in eliciting silencing.

There are four RdRP-like genes in the genome of *C. elegans*. Among those, *ego-1* and *rrf-1* have been shown to be required for germline RNAi and somatic RNAi, respectively. Since the *ego-1* mutant also exhibits a sterile phenotype that impedes practical biochemical and genetic analysis, the requirement for RdRPs in RNAi was assessed using the *rrf-1* mutant and the somatic RNAi pathway (Sijen et al., 2001). The *rrf-1* gene is not required for the formation of the RDE-1 complex. However, the *rrf-1* mutant fails to exhibit transitive RNAi and does not accumulate secondary siRNAs. Because the initiation step is not affected in *rrf-1* mutants, this suggests that RRF-1 acts after the initiation step and perhaps uses target mRNAs as a template to generate secondary siRNAs.

Other proteins that could act in secondary siRNA generation include some of the *mut*ator (*mut*) genes. The *mut* genes are defective in transposon silencing in the germline of *C. elegans*, and many of those were also identified from Rde screens (Ketting et al., 1999; Tabara et al., 1999). These include *rde-2/mut-8*, *rde-3/mut-2*, *rde-5/mut-15*, *rde-6/mut-16*, *mut-7* and *mut-14* (Chen et al., 2005; Tijsterman et al., 2002a; Tops et al., 2005; Vastenhouw et al., 2003). Isolation of the *rde* mutants that are defective in these

germline silencing pathways indicated that the components of the RNAi pathway are shared with other silencing pathways. The *mut* genes are collectively named so because of their mutator phenotype in the mutants. However, what the function of individual genes is and how these genes function in the RNAi pathway as well as other silencing pathways remain mostly unknown.

Several independent experimental results suggest that the mutators function downstream of *rde-1* and *rde-4*. First, *rde-2* and *mut-7* have been shown to be required genetically for the execution of the RNAi signal initiated by *rde-1* and *rde-4* (Grishok et al., 2000). Second, it was observed by Tijsterman *et al*. that gene silencing triggered in *C. elegans* by ~25nt antisense RNAs appeared to bypass the requirement for *rde-1* and *rde-4*, but not for *mut-7* and *mut-14* (Tijsterman et al., 2002a). Third, genetic analysis of a *mut-16* allele which exhibits a temperature-sensitive Rde phenotype indicates that *mut-16* functions at the effector step of the RNAi pathway (See chapter II in this thesis). Finally, in an *in vitro* binding experiment using 2′-*O*-methyl oligos to detect the interaction between RDE-1 and the primary siRNAs, mutations in the *rde-3* gene do not affect this interaction (Chen and Mello, unpublished). Taken together, these results suggest that most, if not all, of the *mut* genes function downstream of the initiation step in the RNAi pathway.

The *rde-2* and *mut-7* genes are only required for germline RNAi. It is not clear whether homologous genes or functionally equivalent genes to *rde-2* and *mut-7* function in somatic RNAi or whether the activity of *rde-2* and *mut-7* is unique to the germline RNAi. *rde-2* and *mut-7* encode a *C. elegans*-specific protein and a protein with a 3′ to 5′

exonuclease domain homologous to that of the *E. coli* RNase D, respectively (Ketting et al., 1999). In *Arabidopsis thaliana*, deletion of a gene, *WEX*, which encodes a protein with an RNase D-like domain, also renders the mutant plant defective in PTGS, suggesting that the requirement for the RNase D-like gene is conserved (Glazov et al., 2003). RDE-2 and MUT-7 have been shown to form a protein complex *in vivo* (Tops et al., 2005). The MUT-7 protein complex, which is around 250 kDa in size in the absence of exogeneously provided dsRNAs, exhibits a mobility shift to ~350 kDa in an gel filtration analysis after the animals are exposed to dsRNAs. This mobility shift is independent of the presence of a target gene and dependent on *rde-1* and *rde-4*, implicating a modification of the MUT-7 complex after dsRNA processing but before target cleavage (Tops et al., 2005). However, the unidentified nature of the modification impedes the understanding of its functional significance. Consistent with the observation that the MUT-7 complex may function before target cleavage, the analysis of siRNA accumulation in the *rde-2* and *mut-7* mutants demonstrates a defect in the accumulation of secondary siRNAs in both mutants, suggesting that the MUT-7 complex is required for secondary siRNA accumulation or stability.

The *rde-3* gene, which is required for efficient RNAi in all tissues, also exhibits a defect in the accumulation of antisense secondary siRNAs (Chen et al., 2005). In addition, *rde-3* is required for the fertility and the viability of the animal at high temperatures, although it is not clear what is the cause of these defects and whether it is related to a deficiency in a small RNA-mediated gene silencing pathway. *rde-3* encodes a member of the Polymerase β nucleotidyltransferase superfamily (Aravind and Koonin,

1999). Members of this family include conventional poly(A) polymerases which function to add poly(A) chains to the 3′ end of a mRNA for its stability, 2′-5′ oligoadenylate synthetase (OAS) which is required to activate RNase L in the double-stranded RNA induced interferon response against viral invasion in mammals, and yeast Trf4p, an unconventional poly(A) polymerase which adds poly(A) tails to mis-modified tRNAs resulting in accelerated degradation by the nuclear exosome (Aravind and Koonin, 1999). Two different lesions identified in the *rde-3* mutants alter conserved amino acid residues required for the activity of this polymerase family, suggesting that its activity is important for its function (Chen et al., 2005). Interestingly, the Rdp1 complex in *S. pombe* contains a homolog of *rde-3*, cid12. cid12 mutant fails to accumulate siRNAs, suggesting that this protein family plays an important role for the production or the stability of siRNAs (Motamedi et al., 2004). However, the exact function of RDE-3 in the RNAi pathway remains to be investigated. In addition, it is not clear whether *rde-3* and cid12 function similarly in the RNAi pathway.

Unlike *rde-2*, *mut-7* and *rde-3*, it was observed in the *mut-14* mutant that the accumulation of the secondary siRNAs does not appear to be affected (Tijsterman et al., 2002a). The MUT-14 protein contains a conserved DEAD box motif followed by a helicase C domain, suggesting that MUT-14 may function as a helicase downstream of secondary siRNA production. It also also suggests that distinct mutator genes are required for different steps in the RNAi pathway. However, the exact function of the mutator genes in *C. elegans* remains to be elucidated.

What remains unclear is how, mechanistically, dsRNAs are synthesized by RdRPs, and why only antisense and not both strands of siRNAs accumulate during secondary siRNA production. The biochemical activity of cellular RdRPs, which are encoded in the genome of some eukaryotes was characterized first in plants and subsequently in fungi (Schiebel et al., 1993). These cellular RdRPs appear to be distinct from the viral RdRPs at the primary sequence level. The RdRP from tomato leaves performs primer-dependent or primer-independent RNA synthesis using both DNA and RNA oligo templates. QDE-1, the RdRP homolog in *Neurospora crassa* that is required for quelling, also displays two modes of polymerase activity on single-stranded RNA templates to produce long dsRNAs or small, approximately 9-21-mer complementary RNAs, along the RNA template (Makeyev and Bamford, 2002). Additionally, *S. pombe* Rdp1 has been shown to generate complete long dsRNAs of 500nt using RNA templates. In *C. elegans*, the secondary siRNAs are derived from the trigger region and the area immediately 5' of the trigger region (Sijen et al., 2001). Unlike plants, siRNAs are not detected from the 3' region next to the trigger in worms. The abundance of siRNAs generated from the 5' upstream region diminishes as a function of distance from the trigger. Secondary siRNAs could be produced from the DCR-1 processing of dsRNAs generated by RdRP using the target mRNA as a template, and followed by asymmetric accumulation of the antisense strand by a yet-to-be-identified mechanism. In this case, the dsRNA synthesis could be a primer dependent or independent event. Alternatively, secondary siRNAs could be synthesize by RdRP using the target mRNA as template and modified to bear a 5′ monophosphate, in which case only antisense siRNAs would be

produced. In this model, DCR-1 activity could be dispensable and additional endonucleases would process the RdRP product into small RNAs.

The degradation of the target mRNA

Several initial observations suggested that mRNA is the target during RNAi in *C. elegans*. First, dsRNAs targeting promoter or intron regions do not cause silencing (Fire et al., 1998). Second, the DNA sequence of the endogenous targeted gene locus is not modified (Montgomery et al., 1998). Third, targeting a gene transcribed in a polycistronic unit operon does not interfere with another gene in the same operon, suggesting that RNAi functions after transcription and splicing (Montgomery et al., 1998). Finally, the steady state level of cytoplasmic mRNAs of the targeted genes is greatly diminished, implying that dsRNAs cause degradation of the target mRNA (Fire et al., 1998; Montgomery et al., 1998).

In vitro studies performed in lysate systems in Drosophila embryos or S2 cells were subsequently established to recapitulate RNAi *in vitro.* These studies demonstrated that the addition of double stranded RNAs into the lysates results in sequence-specific, RNA-dependent mRNA degradation of the target mRNAs by a nuclease complex, named the RNA-induced silencing complex (RISC) (Hammond et al., 2000; Tuschl et al., 1999). Similar results were observed in the lysates prepared from mammalian cell lines when siRNAs were transfected (Elbashir et al., 2001a). The composition of the RISC was further analyzed to identify the nuclease activity, nicknamed “slicer”, responsible for the mRNA degradation. The RISC contains several protein components which, together with

small RNAs, direct the sequence–specific cleavage of targeted mRNAs. Among those protein components is Ago2, a member in the Argonaute protein family, of which RDE-1 is also a member (Hammond et al., 2001). Proteins in this family are characterized by two conserved domains: the PAZ (Piwi/Argonaute/Zwille) domain resides near the amino-terminus and the PIWI domain resides near the carboxyl-terminus of the protein (Cerutti et al., 2000).

Elegant structural analyses of the Argonaute family PAZ and PIWI domains further elucidated the mechanism of siRNA binding, mRNA recognition by siRNAs and the "slicer" activity. The PAZ domain contains a novel OB fold that is topologically distinct but structurally similar to the previously characterized OB fold shown to interact with single-stranded nucleic acids (Lingel et al., 2003; Song et al., 2003; Yan et al., 2003). *In vitro* binding experiments further demonstrated that the PAZ domain interacts with siRNA duplexes at the 3' two-nucleotide overhang, indicating that the PAZ domain is a siRNA-binding domain. A PAZ domain has also been found in the Dicer protein family.

The analysis of the structure of the archeabacterial Argonaute protein has revealed the identity of "slicer" (Parker et al., 2004; Song et al., 2004). The structure demonstrated that the PIWI domain resembles the fold of the RNase H (Ribonuclease H) family and together with the N-terminal domain and the middle domain between PAZ and PIWI, forms a crescent-shaped base with PAZ domain held above the base. A conserved catalytic motif "DDE" or "DDH" found in both RNase H and PIWI domains is present in the center of a positively charged groove formed by these four domains. Structural

Analysis of the PIWI domain bound to RNAs provided further information about the RNA-protein interaction, which sheds light on the mechanism of target RNA recognition and cleavage (Ma et al., 2005; Parker et al., 2005; Yuan et al., 2005). The 5' phosphate of the guide RNA is anchored in a metal binding pocket. The first base of the guide RNA is not paired with the target RNA where as the rest of the RNA duplex assumes an A form helix (Parker et al., 2005). This RNA duplex is positioned in the PIWI pocket so that the scissile phosphate of the target RNA is adjacent to the putative catalytic site of the PIWI domain. The hallmarks of RISC cleavage resemble that of RNase H (Martinez and Tuschl, 2004; Schwarz et al., 2004), and mutations in the catalytic residues of Ago2 abolish slicer activity, further confirming that Argonaute is the slicer (Liu et al., 2004).

There are 27 Argonaute genes in the *C. elegans* genome (Yigit et al., 2006). Among those, *rde-1* was demonstrated to be essential only for the initiation of RNAi, raising the possibility that other Argonaute proteins may participate in the downstream steps of RNAi. Indeed, an extensive analysis of the Argonaute gene falimy in *C. elegans* revealed the requirement of several other argonaute genes, such as *ppw-1*, F58G1.1, *sago-1*, *sago-2*, for efficient RNAi (Tijsterman et al., 2002b; Yigit et al., 2006). Mutants with deletion of a single argonaute gene in this category show weak or no RNAi defect. However, a mutant with all these genes deleted, named *MAGO* (for multiple argonautes), exhibits an accumulative deficiency in both germline and somatic RNAi.

The RNAi defect of *MAGO* mutant appears to occur at a downstream step. It has been demonstrated that the initiation phase of the RNAi pathway is not affected in the *MAGO* mutant (Yigit et al., 2006). In addition, immunoprecipitation experiments showed

that these Argonautes interact with secondary antisense siRNAs, hence the name "secondary Argonautes" (Yigit et al., 2006). Interestingly, overexpression of a single secondary argonaute, such as *ppw-1*, *sago-1*, *sago-2*, not only rescues the *rde* defect of the *MAGO* mutant but also render the strain hypersensitive to RNAi. In correlation with the enhanced RNAi phenotype, the level of secondary antisense siRNAs is increased in the rescued background (Yigit et al., 2006). These results suggest that whereas only the *rde-1* gene functions in the upstream step of the RNAi pathway, several Argonautes function interchangeably at the effector step for the degradation of mRNAs.

Curiously, sequence alignment analysis of the secondary Argonautes revealed that the conserved DDH residues for the nuclease activity of the PIWI domain are missing in these proteins, raising the question of how mRNA degradation is elicited. It is possible that another cryptic catalytic site is present in the PIWI domain of these proteins. However, unlike in *Drosophila* and human, RISC activity in *C. elegans* is poorly characterized mainly because of the inability to recapitulate RISC cleavage *in vitro*. Therefore, the slicer activity of the Argonautes cannot be readily examined. Alternatively, the secondary Argonautes might bring the mRNAs to close proximity of other nucleases that execute mRNA cleavage and degradation.

One candidate for this nuclease is the Tudor Staphylococcal Nuclease (or Tudor-SN), which is a component of the RISC in Drosophila and human (Caudy et al., 2003). The homolog of Tudor-SN in *C. elegans*, TSN-1, associates with another RISC component, VIG-1. However, the interaction between TSN-1 and the Argonautes in *C. elegans* has not been characterized. In addition, although TSN-1(RNAi) worms show a

partial defect in microRNA pathway, the role of TSN-1 in RNAi remains to be elucidated. Generation and characterization of a TSN-1 mutant should aid in addressing these questions.

Small RNA Mediated Gene Silencing Pathways

Research on the mechanism of RNAi has shed light on the mechanisms of additional endogenous pathways that utilize small RNAs and members in the Argonaute protein family to control gene expression in an organism (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). Cloning and profiling of small RNAs in distinct developmental stages or tissues in a variety of organisms have revealed the existence of abundant small RNAs (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). Some small RNAs are transcribed as single or polycistronic gene units using their own promoters or are encoded in the intron of other transcripts. The resulting transcripts often form imperfectly paired, but highly double-stranded hairpin structures that are processed by Dicer protein family members to release mature small RNAs. Others are likely to be produced by Dicer processing from perfectly paired double-stranded RNAs (Ambros and Lee, 2004). These perfectly paired dsRNAs are the result of bi-directional transcription or the transcription of inversely repeated sequences, which are often found in heterochomatic regions, viral sequences, transgenes, or transposons. Interestingly, the repeat associated small RNAs (rasiRNAs) which protect genome stability in the Drosophila germline, appear to function in a pathway distinct from that of the microRNA or siRNAs (Vagin et al., 2006). In general, the small RNAs, accompanied by distinct

effector proteins, elicit a variety of gene regulatory modes including DNA elimination in Tetrahymena, chromatin modification, transcriptional repression, RNA degradation and translational repression. In the following sections, distinct small RNA mediated gene silencing pathways in *C. elegans* as well as other organisms will be described.

MicroRNAs

MicroRNAs are genes encoding RNA precursors that are processed to release small RNAs of ~22nt in length (Figure I-2). The first microRNA, *lin-4*, was identified in *C. elegans* in 1993, and plays an important role by temporally regulating the protein level of LIN-14 in the heterochronic pathway, a pathway which controls developmental timing in worms (Lee et al., 1993; Wightman et al., 1993). Subsequently, another microRNA, *let-7*, was also demonstrated to function in the heterochronic pathway. Interestingly, the sequence and expression pattern of *let-7* are both conserved in a wide range of animals (Reinhart et al., 2000). The search for the importance and generality of microRNAs culminated with several studies that identified hundreds of small RNAs in a variety of organisms including *C. elegans*, Drosophila, and humans (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). To date, microRNAs have been demonstrated to be involved in development, cell death, cancer, and even memory formation.

The biogenesis of microRNAs is regulated both temporally and spatially (Ambros, 2004). The microRNA genes are transcribed by RNA polymerase II into

Figure I-2

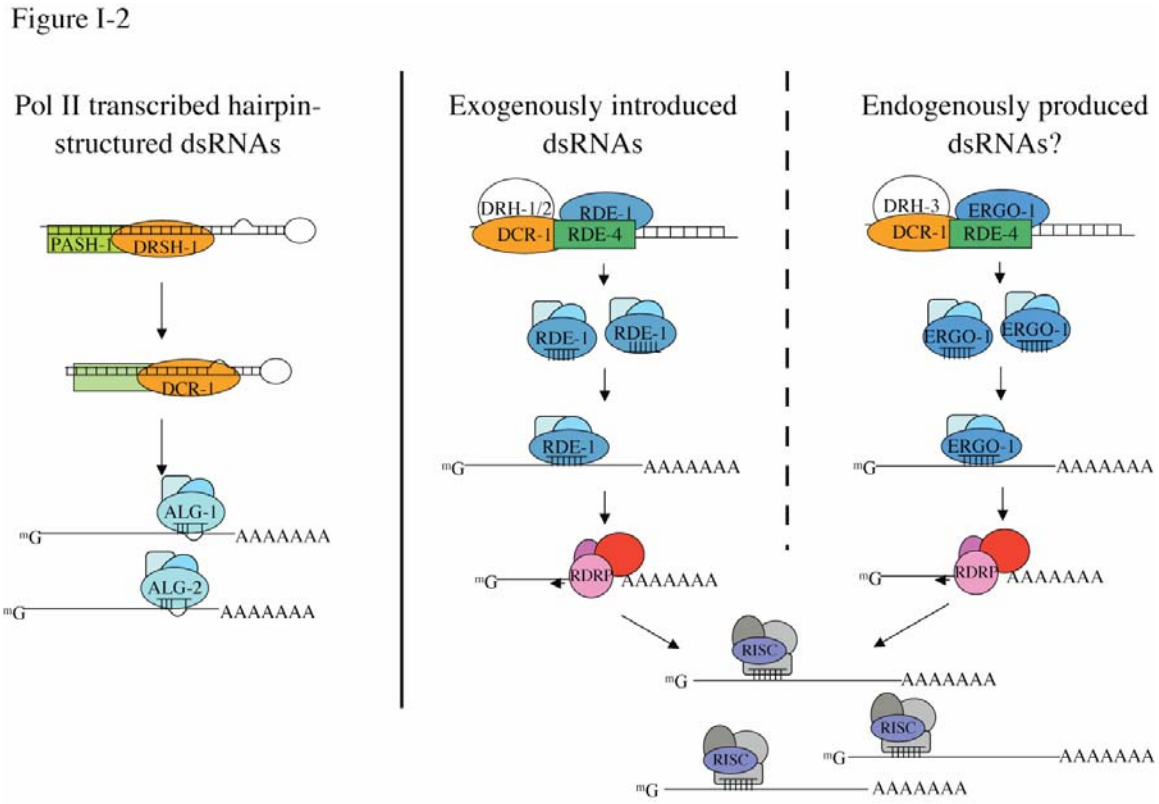

Pol II transcribed hairpin-structured dsRNAs
PASH-1
DRSH-1
DCR-1
ALG-1
mG
AAAAAAA
ALG-2
mG
AAAAAAA
Exogenously introduced dsRNAs
DRH-1/2
RDE-1
DCR-1
RDE-4
RDE-1
RDE-1
RDE-1
mG
AAAAAAA
RDRP
mG
AAAAAAA
Endogenously produced dsRNAs?
DRH-3
ERGO-1
DCR-1
RDE-4
ERGO-1
ERGO-1
ERGO-1
mG
AAAAAAA
RDRP
mG
AAAAAAA
RISC
mG
AAAAAAA
RISC
mG
AAAAAAA
RISC
mG
AAAAAAA

Figure I-2 A model for the small RNA pathways in *C. elegans*. See text for details.

long and highly structured transcripts termed primary miRNAs (pri-miRNAs) (Lee et al., 2004). Mechanisms that control PolII transcriptional regulation have also been observed to control the expression of *let-7* in *C. elegans* (Johnson et al., 2003). The pri-miRNAs are capped and polyadenylated, and can encode many miRNAs (polycistronic) or single miRNAs (monocistronic). The pri-miRNAs are processed by Drosha, an RNase III enzyme similar to Dicer, into stem-looped pre-miRNAs of ~60 to 70nt in length, termed pre-miRNAs in the nucleus (Han et al., 2004; Lee et al., 2004). The Drosha complex also contains the dsRNA binding protein Pasha/DGCR8. The *C. elegans* Drosha homolog, *drsh-1,* and *Pasha* homolog, *pash-1,* appear to exhibit heterochronic defects similar to *dcr-1,* suggesting that the microRNAs pathway is impaired (Denli et al., 2004). The pre-miRNAs are recognized by a RanGTP-dependent dsRNA binding protein Exportin 5 for export into the cytoplasm in HeLa cells (Bohnsack et al., 2004; Lund et al., 2004; Yi et al., 2003). Whether worms use the same export mechanism for the pre-microRNAs remains to be elucidated.

Once released into the cytoplasm the pre-miRNAs are processed by Dicer to produce mature miRNAs, which are single-stranded RNAs of 21nt in length (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). In *C. elegans,* two Argonaute protein family members ALG-1 and ALG-2 interact with DCR-1 and microRNAs (Grishok et al., 2001). The *dcr-1* mutant and the *alg-1* and *alg-2* double mutant exhibit severe heterochronic defects and accumulate mature microRNAs at lower levels (Grishok et al., 2001; Ketting et al., 2001; Knight and Bass, 2001). Both ALG-1 and ALG-2 proteins contain the “DDH” motif that is important for the cleavage activity

of the Argonaute proteins (Yigit et al., 2006). However, it is not known whether ALG-1 and ALG-2 elicit cleavage activity when a perfectly complementary target RNA is present. Inhibition of the activity of ALG-1 and ALG-2 by RNAi does not affect RNAi targeting a reporter gene (Grishok et al., 2001). In addition, expression of ALG-1 protein in the *rde-1* or *mago* mutant background does not rescue the RNAi defect, suggesting that ALG-1 and ALG-2 functions mainly in the microRNA pathway (Yigit et al., 2006).

Like the RNAi initiation complex in *C. elegans*, double stranded RNA binding proteins have been reported to partner with Dicer and Agonautes in the RISC in Drosophila and human (Forstemann et al., 2005; Jiang et al., 2005; Saito et al., 2005). These dsRNA-binding proteins may assist Dicer processing or small RNA loading into the RISC for both the RNAi and microRNA pathways. However, the requirement of a dsRNA binding protein in the microRNA pathway functioning with ALG-1, ALG-2 and DCR-1 has not yet been reported and remains to be investigated.

Most animal microRNAs recongnize their targets through an impefect base-pairing of the miRNA to the the 3' untranslated regions of the target mRNAs (Lee et al., 1993; Wightman et al., 1993). Neverthless, the complementarity between the 5' region (2 to 8nt) of a microRNA, termed "seed" region, and its target appears to be important for the target recognition (Lai, 2002; Lewis et al., 2003; Lim et al., 2003). The imperfect pairing of the miRNA and target mRNA prevents RISC cleavage and in turn is thought to cause the inhibition of translation (Doench et al., 2003; Hutvagner and Zamore, 2002; Zeng et al., 2002; Zeng et al., 2003). It is worth noting that in plants, most microRNAs are 100% complementary to the targets and function like siRNAs (Baulcombe, 2004). In

contrast, most animal microRNAs pairs with the target mRNAs imperfectly and generally function to repress translation.

In *C. elegans,* the function of a microRNA was first characterized in the studies of the *lin-4* microRNA (Olsen and Ambros, 1999). It was observed that the expression of *lin-4* repressed the protein level produced by *lin-14*, which functions in the same developmental pathway as *lin-4* (Wightman et al., 1993). Interestingly, the sequence of the *lin-4* small RNA was found to be partially complementary to that of the *lin-14* mRNA, suggesting that the physical interaction between *lin-4* and *lin-14* mRNA negatively regulates the expression of LIN-14 protein. *lin-4* and *lin-14* are present in the same polysome fraction, indicating that translation is properly initiated in the presence of the *lin-4/lin-14* interaction. However, protein synthesis is not observed, suggesting that the effect of *lin-4* inhibition happens after translational initiation (Olsen and Ambros, 1999).

Recently, Ago1 and Ago2 were reported to associate with a component of P bodies, GW182 (Liu et al., 2005; Rehwinkel et al., 2005; Sen and Blau, 2005). Immunostaining demonstrated that Ago2 localizes to the polar bodies (P bodies). P bodies are translationally inert cytoplasmic foci that contain, among other things, untranslated RNAs. Dissociation of the Argonaute proteins from P bodies disrupts translational repression, suggesting that the localization of the RISC complexes in the P bodies is important for translational repression. Both ALG-1 and ALG-2 have also been shown to localize to the P bodies in *C. elegans* through the interaction with a GW182 homolog, AIN-1, suggesting that this localization is conserved (Ding et al., 2005).

In addition to translational control, a recent report demonstrated that the *let-7* microRNA in *C. elegans* also appears to cause destabilization of its target mRNAs (Bagga et al., 2005). This mRNA degradation mechanism is independent of the RNAi pathway. Instead, two exonucleases, *xrn-1* and Y48B6A.3, are required for the destruction of the target. Similar downregulation of transcript levels have also been observed in human cells, suggesting that microRNAs regulate gene expression through at least two distinct mechanisms (Lim et al., 2005). It is possible that the combination of translational repression and destabilizing mRNA function together in the microRNA regulation. However, it is not clear how these two regulatory modes intersect to achieve the final inhibition of gene expression.

The small RNAs used as a defense mechanism

All cells encounter hostile foreign nucleic acids. Invasion of these molecules could lead to detrimental consequences as they may try to 1) take control over in the host to produce biological materials for their own propagation (for example, a virus), 2) insert into a part of the host genome which may contain essential genetic information (for example transposable element) or 3) overexpress materials that may or may not cause toxicity to the cell (i.e. transgenes). It was first described in plants that introduction of a transgene induces silencing of the transgene as well as the endogenous gene homologous to the transgene(Baulcombe, 2004). This phenomenon is conserved in a variety of organisms and is given a number of names including transgene silencing, cosuppression, post-transcriptional gene silencing (PTGS), and, in *Neurospora crassa*, quelling. In

addition, the cell has also developed defense mechanisms to silence viruses (virus-induced gene silencing) and transposable elements (tranposon silencing).

In *C. elegans*, repetitive transgenes are often silenced or cosuppressed in the germline of the animal. In addition, transposable elements are only active in the somatic tissues but are silenced in the germline. Characterization of the phenotypes of *rde* mutants revealed that some of them are defective in transposon silencing and cosuppression (Ketting et al., 1999; Tabara et al., 1999). These mutants include *rde-2/mut-8, mut-7, rde-3/mut-2, mut-16/rde-6* and *mut-14*. However, genes that act at the initiation step of RNAi, *rde-1* and *rde-4*, are not required for transposon silencing and cosuppression and, likewise, not all *mut* genes function in RNAi. Genetic analysis has placed *rde-2* and *mut-7* downstream of *rde-1* and *rde-4*, suggesting that RNAi and these two silencing pathways may be initiated by different players but share a common set of components in some downstream steps (Grishok et al., 2000).

Molecular analysis of RNA intermediates that accumulate during transposon silencing demonstrated that, *in vivo*, dsRNAs are predominantly produced from the terminal inverted repeats (TIR) which flank the coding sequence of the transposase Tc1 (Sijen and Plasterk, 2003). The accumulation of TIR dsRNAs is present in several mutants defective only in transposon silencing, suggesting that transposon mobilization in these strains is due to a defect downstream of dsRNA production. To examine if the dsRNAs are processed into small RNAs as in RNAi, the presence of siRNAs corresponding to the Tc1 transposase and TIR in wildtype and various mutant backgrounds was analyzed. Whereas only antisense siRNAs accumulate during RNAi,

siRNAs of both orientations targeting the TIR and the coding region of the transposase were detected. The TIR-derived siRNAs can be a direct product of Dicer whereas the siRNAs derived from the transposase coding region can be generated through the Dicer-mediated processing of dsRNAs produced via the action of RdRPs or via low level of read-through transcription from two Tc1 sequences in inverted orientations. Consistent with the genetic analysis, the accumulation of transposon-derived siRNAs was observed in wildtype worms as well as *rde-1* and *rde-4* mutants, but not in mutants that exhibit transposon desilencing.

To further characterize the silencing ability of transposon-derived siRNAs, a germline-expressed green fluorescent protein (*gfp*) reporter gene was fused with TIR sequences and the GFP signal was analyzed in different genetic backgrounds. The GFP signal was suppressed in the wildtype worms but de-repressed in the *mut* animals. In addition, several observations also suggest that transposon silencing functions as RNAi at the post-transcriptional level. First, biochemical fractionation experiments demonstrated that the siRNAs are present in the S100 cytosolic polysome fractions. Second, *gfp* reporter assays showed that TIR siRNAs can only silence GFP when TIR is fused to the *gfp* coding region but not the 3' untranslated region. Third, the ratio between the levels of spliced mRNA to that of un-spliced mRNA increases in *mut* background as a result of the increased level of cytoplasmic mRNA, which is the target in post-transcriptional gene silencing.

The presence of dsRNAs and siRNAs derived from a *gfp* gene undergoing cosuppression in the germline of *C. elegans* was also examined. Although dsRNAs

formed by pre-mRNAs of the cosuppressed *gfp* gene can be detected, siRNAs, presumably produced by processing of the dsRNAs by DCR-1, remain undetectable (Robert et al., 2005). Reduction of the DCR-1 protein activity by RNAi feeding caused a mild decrease in *gfp* silencing in an animal undergoing cosuppression, suggesting that DCR-1 might be involved in the process but that the siRNAs are below detection. Nevertheless, the mRNA level of the cosuppressed *gfp* gene is drastically reduced in wildtype animals and is partially recovered in mutant backgrounds previously shown to affect cosuppression, *rde-2* and *mut-16*. This indicates that mRNA is the target of the silencing caused by cosuppression.

Several observations suggest that cosuppression functions by transcriptional repression. First, a genome-wide RNAi screen in search for genes required for cosuppression identified several genes that are involved in chromatin structure and transcription. Second, full recovery of a cosuppressed gene requires a period longer than that of the recovery of most RNAi treatments, which is at least three generations. Finally, the ratio between the levels of spliced mRNA to that of un-spliced mRNA remain unchanged in the de-cosuppressed animals, indicating that the levels of both spliced and unspliced RNAs increase as a consequence of transcriptional re-activation.

In summary, the present data indicate that although the RNAi, transgene silencing and cosuppression pathways share several components, they differ in how the silencing is initiated and how the silencing is executed (Figure I-2). It is not clear which genes are involved in recognizing dsRNAs derived from transposons and cosuppressed transgenes. Identification of these players will help us to understand how these distinct complexes

recognize dsRNA intermediates in each pathway. It is unlikely that it is due to the subcellular localization of the initiation complex as the RDE-1/RDE-4 complex also responds to promoter driven RNAi in which dsRNAs are produced in the nucleus. While RNAi and transposon silencing elicit silencing post-transcriptionally, cosuppression seems to function transcriptionally, perhaps by inducing the formation of heterochromatin. However, unlike heterochromatin silencing in *S. pombe*, siRNAs are not detected in animals undergoing cosuppression. It is possible that cosuppression is initiated by an RNAi like mechanism but that the siRNAs may not be required for the subsequent maintenance phase of the silencing. Alternatively, other pathways may function in parallel with RNAi for silencing such that the involvement of the RNAi pathway is only partial and thus the siRNAs are not easily detected.

Members of the Argonaute gene family are ideal candidates to orchestrate the convergence and divergence of these silencing pathways. RNAi-based genome-wide screens have identified *csr-1* and C14B1.7 functioning in cosuppression and *ppw-2* functioning for both cossupression and transposon silencing (Robert et al., 2005; Vastenhouw et al., 2003). It is possible that these pathways also function in a "two-step" mechanism like the RNAi pathway. The fact that *rde-1* is not required for cosuppression or transposon silencing suggests that some of these Argonautes are required to initiate these silencing pathways (Dernburg et al., 2000). In addition, whether the secondary Argonautes are shared between these pathways also remains to be elucidated.

<u>More endogenous small RNAs</u>

The cloning of small RNAs in *C. elegans* has revealed several classes of endogenous small RNAs other than microRNAs (Ambros et al., 2003). The first class of these small RNAs is named tiny noncoding RNAs (tncRNAs). The tncRNAs are encoded in the intergenic regions of the genome. In addition, the timing of the expression of some tncRNAs is developmentally regulated. Although most of the tnc RNAs require DCR-1 processing, tncRNAs do not appear to be encoded within long hairpin structure like the microRNAs. Therefore, they define a distinct class of small RNAs. The function of the tncRNAs remains unclear. The developmentally regulated expression of tncRNAs suggests that they could function to regulate gene expression during development. However, it is unclear what the targets of tncRNAs are. The second class of the newly identified small RNAs represents small antisense RNAs perfectly complementary to around 500 coding sequences of the *C. elegans* genome, suggesting that these RNAs are endogenously produced siRNAs (endo-siRNAs). Some of the endo-siRNAs are originated from the sequences of transposons, supporting the finding that one of the functions of RNAi is to silence transposable elements. Lastly, 41 distinct (but partially overlapping in some cases) small RNAs were found at a particular locus on Chromosome X. This locus does not encode previously reported cDNAs or predicted coding sequences. A few of the X-cluster sequences are predicted to form hairpin structures that may be sufficiently long for Dicer-mediated processing. However, it is not clear what the precursors of these small RNAs are.

The first insight on how these classes of small RNAs are produced came from a DCR-1-associated complex distinct from those that are involved in the microRNA or the

exogenous RNAi pathway (Duchaine et al., 2006). This complex contains, in addition to DCR-1, several proteins whose functions inhibit the exogenous RNAi pathway, including ERI-1 (Enhancer of RNAi), ERI-3, ERI-5 and RRF-3. The *eri-1* gene encodes an exonuclease with a SAP domain that degrades small RNAs (Kennedy et al., 2004). *eri-3* encodes a novel protein. *eri-5* encodes a protein with two tudor domains and *rrf-3* is a homolog of the RdRP gene family (Simmer et al., 2002). Thus, the tncRNAs and the X cluster RNAs appear to require distinct DCR-1 complexes for processing than that of the microRNAs.

The *eri* and *rrf-3* mutants are hypersensitive to exogenous RNAi (exo-RNAi). It seems that more than one mechanism contribute to the enhancement of RNAi in these mutants. First, the *eri-1* gene encodes an exonuclease which appears to degrades siRNAs. Therefore, in the absence of *eri-1*, more exo-siRNAs accumulate and function in the exo-RNAi pathway. In addition, the *eri* pathway seems to compete with the exogenous RNAi pathway for some shared and limiting components. Northern blot analysis on the expression of the endogenous small RNAs in the *MAGO* mutant, in which multiple secondary Argonautes are defective, demonstrated that the level of these small RNAs is drastically reduced. This suggests that these two pathways share some of the secondary Argonautes. Furthermore, overexpression of a single secondary Argonuate in the *MAGO* mutant renders the animal hypersensitive to RNAi, suggesting that the secondary Argonautes are one of the limiting factors in these pathways.

In addition to the secondary Argonautes, a single Argonaute, ERGO-1, appears to be essential for the production of some endogenous small RNAs. In addition, *ergo-1*

mutant is hypersensitive to exo-RNAi. It is possible that ERGO-1 functions strictly in the endogenous siRNA pathway along with the secondary Argonautes. Alternatively, ERGO-1 could function like RDE-1, upstream in the endo-RNAi pathway, to initiate the amplification of the secondary endogenous RNAs by *rrf-3*.

Another homolog of the DRH protein family, DRH-3, is involved in the production of endogenous small RNAs. In addition, *drh-3* is required for germline RNAi and fertility (Duchaine et al., 2006). Unlike its homologs DRH-1 and DRH-2, the DRH-3 protein does not interact with RDE-4. The exact function of *drh-3* in both the exo-RNAi and endo-RNAi pathways requires further elucidation. In addition to the genes mentioned above, several players in the exo-RNAi pathway are also necessary for the production or stability of mature endogenous small RNAs. These genes include *rde-2*, *rde-3*, *rde-4*, *mut-7*, and *mut-14* (Lee et al., 2006).

It is worth noting that different endogenous small RNAs appear to exhibit distinct requirements for the genes mentioned above, suggesting that distinct pathways which share some common components are employed for the generation or stability of these small RNAs. Furthermore, it seems likely that additional small RNAs remain to be identified. It is not surprising that a complex network of interconnecting pathways is required for the biogenesis and the function these small RNAs (Figure 1-2). Further studies are required to understand the whole picture.

Conclusion

Since RNAi was first discovered in *C. elegans* eight years ago, our understanding of the mechanisms of RNAi as well as RNAi-like gene silencing pathways has advanced at an incredibly rapid pace. We are beginning to appreciate that RNAi embodies a conserved and important gene regulation mechanism. In addition, due to its sequence-specific nature, the use of RNAi could offer a putative sequence-specific approach for gene therapy. However, despite the remarkable progress, as presented in this chapter, many unknowns still remain to be elucidated. Many new genes, and even small RNAs in these related pathways remain to be identified, and the molecular mechanisms of the small RNA-mediated pathways are still unclear. In the past, research on the RNAi pathway in *C. elegans* has enriched our knowledge about this topic. The studies presented in this thesis will continue to shed light on our understanding of RNAi.

Chapter II

Identification of *RNAi deficient* (*rde*) mutants

Chun-Chieh G. Chen, Jennifer McCullough, Darryl Conte Jr., Pedro Batista, Na Li[1], Weifeng Gu, and Craig Mello[1]

Program in Molecular Medicine and [1]Howard Hughes Medical Institute, University of Massachusetts Medical School, Worcester, Massachusetts 01605

Chun-Chieh G. Chen led and performed the genetic screens and analysis of the mutant candidates with Jennifer McCullough summarized in Table II-1. She isolated 2 *rde-1* alleles for the biochemical analysis performed by Pedro Batista in Figure II-2 and 11 alleles of *rrf-1* analyzed in Figure II-3 by Darryl Conte Jr. and Na Li. She isolated and analyzed the *rde-5* alleles in Figure II-4. She designed the assay and assisted Jennifer McCullough in analyzing the temperature sensitive alleles of *mut-16* and *rrf-1* in Figure-5 and 6. She also assisted Na Li to sequence the *rde-1* and *mut-16* alleles. She led and performed the genetic screen with Martin Simard and Bill Tsai and isolated the temperature sensitive *drh-3(ne4197)* allele. She assisted Weifeng Gu to clone and analyse the *drh-3(ne4197).*

Summary

RNA interference was first discovered in *C. elegans* to experimentally downregulate gene expression. Subsequent studies have shown that similar mechanisms mediate a variety of important gene regulatory programs. To investigate the molecular mechanism of RNAi and related pathways in *C. elegans*, we conducted genetic screens to identify mutants that are *R*NAi *de*ficient (*rde*). Here we present the result of the screens. We have isolated 76 alleles of mutants that are deficient in RNAi. We categorized the mutants by complementation test and identified many new alleles of previously known *rde* mutants and 3 *rde* mutant alleles that defined a new *rde* locus, *rde-8*. Several of these *rde* mutants also exhibit additional defects during development. These defects suggest that an additional small RNA pathway could also be affected. In summary, our result report new players in the RNAi and related pathways and provide new reagents for further analysis of the RNAi pathway in *C. elegans*.

Introduction

The experimental delivery of dsRNAs into a variety of eukaryotic organisms triggers a potent and sequence-specific gene-silencing phenomenon termed RNAi (Fire et al., 1998; Rocheleau et al., 1997). First discovered in the nematode *Caonorhabditis elegans*, RNAi has been shown underlie a conserved mechanism that regulates gene expression. RNAi begins with the introduction of dsRNAs into the cells by a mechanism termed systemic RNAi (Winston et al., 2002). These dsRNAs are subsequently processed into small RNAs of ~21 to 27 nucleotides in length known as small interfering RNAs (siRNAs) (Bernstein et al., 2001; Elbashir et al., 2001b). The siRNAs then direct the sequence-specific search and degradation of the target mRNAs cognate to the dsRNAs (Hammond et al., 2000; Tuschl et al., 1999).

To investigate the molecular mechanism of RNAi, our laboratory utilizes the nematode *C. elegans*. *C. elegans* is a well-established genetic organism and the first insights into the genes required for RNAi have come from genetic screens in search of mutants that are *R*NAi *de*ficient (*rde*) (Tabara et al., 1999). As described in Tabara *et al*., two classes of mutants were found to be required for RNAi (Tabara et al., 1999). The first class includes *rde-1* and *rde-4*, which are essential for RNAi but display no other phenotype under laboratory growth conditions. The second class includes *rde-2*, *rde-3* and *mut-7* which exhibit defects in both RNAi and germline gene silencing phenomena such as transposon silencing and cosuppression. Interestingly, *rde-2* and *mut-7* are only required for RNAi targeting genes expressed in the germline but not soma, suggesting that distinct genes function in germline and somatic RNAi (Ketting et al., 1999). Further

genetic analysis suggests that *rde-1* and *rde-4* act at an upstream step in initiating the RNAi signal whereas *rde-2* and *mut-7* function downstream of *rde-1* and *rde-4* in the execution step (Grishok et al., 2000). This result suggests that RNAi and germline gene silencing are initiated separately. However, at least some components are utilized by both pathways to elicit gene silencing in the downstream steps.

As discovered by our laboratory and others, the microRNA pathway employs RNAi related genes and mechanisms in *C. elegans* to elicit gene silencing (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). This finding further depicts the convergence and divergence of these small RNA mediated gene-silencing pathways. First discovered in *C. elegans,* the miRNAs are small RNAs essential for the development and differentiation of most metazoans (Lee et al., 1993; Wightman et al., 1993). Mutants that fail in the production of mature miRNAs exhibit developmental defects. The central player in both RNAi and miRNA pathways in *C. elegans* is DCR-1, a protein with RNase III-like enzymatic activity. Dicer functions in the RNAi pathway and the microRNA pathway to process long dsRNAs and miRNA stem-looped precursors into siRNAs and microRNAs, respectively. In contrast, members in the argonaute protein family, which includes RDE-1 and ALG-1/2, are required for distinct pathways. For example, RDE-1 is only required for RNAi while ALG-1/2 is only required for the microRNA pathway. Presumably, different argonaute proteins insulate these small RNA pathways from interfering with each other.

Because RNAi shares components with a variety of small-RNA-mediated processes, some of which are essential for the development of the animal, lesions in these

genes will affect the viability of the corresponding mutant and prevent the isolation of such mutants using our previous screening strategy. In a continuation of our effort to identify these new genes, especially genes that function in both RNAi and development, I have conducted several genetic screens that aim at isolating not only novel *rde* mutants but also temperature sensitive lethal *rde* mutants.

The result of the screens can be summarized by the isolation of four categories of *rde* mutants which include: 1) new alleles of known genes involved in RNAi, including *rde-1*, *rde-3*, *rde-5/mut-15*, *rde-6/mut-16*, *sid-1* and *rrf-1*; 2) new alleles that defined a new *rde* locus, *rde-8*; 3) two temperature sensitive alleles of *rde-6/mut-16* and *rrf-1*; and 4) a temperature-sensitive lethal allele of *drh-3*. The following chapter discusses the strategy and the outcome of the screens and describes the further characterization of some mutant alleles. The cloning and characterization of *rde-3* and *rde-8* will be reported separately in Chapters III and IV, respectively.

Material and methods

C. elegans genetics and culture

C. elegans genetics and cultures were done as described in Brenner (1974) (Brenner, 1974). The Bristol strain N2 was used as the standard wild-type strain. Worm strains used for the screens and mutant alleles used in the complementation tests are listed by chromosome as follows, LGI*: lin-11(n566), rrf-1(ne734), rde-6/mut-16(ne322), drh-3(tm1217)*; LGV: *sid-1(ne302), rde-1(ne300), rde-5/mut-15(ne321)*; LGX: *pes-10::gfp*.

RNAi tests were performed as described in Fire *et al.*, 1998, Tabara *et al.*, 1999 and Timmons *et al.*, 2001(Fire et al., 1998; Tabara et al., 1999; Timmons et al., 2001).

Plasmid construction for *let-2::gfp* bacterial food

Primers CMo 4455 GAT TGC CAG GAA GCA CCA GGG AGC and CMo 4456 TTA ATG ACG TCC GTC GGT G were used to amplify a 731 bp fragment at the 3' end of the coding sequence of *let-2* from a wild-type cDNA library. The PCR fragment was then cloned into PCR-Blunt II-TOPO using Zero Blunt TOPO PCR Cloning Kit (Invitrogen). The *let-2* fragment was sub-cloned into L4440 (with two T7 promoters at two opposite orientations) using EcoRV and KpnI (New England Biolabs; NEB), and L4417 (L4440 with *gfp* sequence in between the two T7 promoters) using SpeI and SmaI (NEB). Both L4440 and L4417 were obtained from the Fire Lab plasmid kit.

Mutagenesis

WM147, a strain with Egl (egg laying defective) phenotype caused by the *lin-11* mutation and an integrated transgene expressing *gfp*, was used as the starting strain. The Egl phenotype results in the accumulation of embryos inside the mother. Once hatched, the progeny are trapped inside of the cuticle of the mother and appears as "a bag of worms". The "worm bag" phenotype facilitates the screening process such that is easily distinguished under microscope and allows the recovery of clones of worms. 1.5 million worms at larval stage 4 (L4) were incubated in solution containing 50 mM EMS or 0.5 mM ENU for 4 hrs, then plated out onto 8 15cm nematode growth medium (NGM) plates

with 200,000 animals on each plate. These worms were grown for 24 hrs to produce F1 progeny. The mutagenized F1 population was harvested in synchrony by hypochlorite treatment. Approximately 20 million F1 worms were recovered and plated out as 200,000 animals per 150 mm NGM plate (~100 plates in total). Each NGM plate was treated independently to obtain L1 larvae of the F2 and F3 generation. 200,000 F3 larvae in 2mL of concentrated *let-2::gfp* bacterial food were plated onto NGM plates containing 50 μg/mL ampicillin and 80 μg/mL IPTG and allowed to grow until "worm bags" (i.e. the egg laying defective mothers were full of eggs/hatched larvae) formed. 10 worm bags were picked from each 150mm NGM plates and transferred to a 35 mm NGM plateworm bag. Approximately 1,000 worm bags were isolated which represented at least 100 independent mutations. These *rde* candidates were further analyzed as described in the results section.

Complementation test

Complementation tests were performed using males from the known mutants *rde-1(ne300)*, *rrf-1(ne734)* and *sid-1(ne302)* and Egl hermaphrodites from the candidate mutants. F1 wildtype cross progeny were tested on *let-2::gfp* food for RNAi resistance.

Sequencing of *rde-1*

Sequencing of *rde-1* was done as described in Tabara *et al*., 1999.

Sequencing of *rrf-1*

CMo 3492 GTC AGA GGA CAC ATC AAC and CMo 3493 CTC ATT CAT CTG TTT GGC were used to amplify the 5’ genomic region of *rrf-1* locus. The PCR product was sequenced with the following primers: CMo 3496 TGA AGG TTA ATG AGA AC, CMo 3497 ATA TTT CTA GGT TCT TCG, CMo 3498 ATT TCA ATC ATT GGC AGG, CMo 3499 CAT GTG TTC ATT TTG AAG, CMo 3500 GTC TCG TCT GAG GTC TCG, CMo 3501 ACT GAA TGA GCA ACA TTG, CMo 3502 GAG TTC TAC GGA AAT TCG, CMo 3503 CAA TGC TTC ACT CAG AGC, CMo 3504 GTC ATC GCT ATT GAA CCC, CMo 3525 ACG ATC CAT CAA TTC CTC, CMo 3526 CCA AAT TGC ATC GCT TGG, CMo 3527 TAT ATT CGA TGC TGT CGG, CMo 3528 CTA TGC TCT CCT TTT CTG, CMo 3529 TTC GGT GGG AAT ATC TGC, CMo 3530 TGG AAC TCA ATT GTG AAG, CMo 3531 ATA GCT ACA GCT AAA CAC, CMo 3532 TGA ATT CAA GTA CTC ACC, CMo 3533 GAG TTA TCG AAT TCC GG and CMo 3534 CAG TGA GCA AGT ATC GTG. CMo 3494 GCA GAT TCG ATT CCG CGG and CMo 3495 GTA TTG TGC AAC AAC ACC were used to amplify the 3’ genomic region of *rrf-1* locus. The following primers were used for sequencing this PCR product: CMo 3504, CMo 3505 TCA GGA GTT TCC ACG TCG, CMo 3506 GTC CCC GGA GAT GTT CGG, CMo 3507 CTC GAA CAG TCA TCT TCA C, CMo 3508 GAA CTG GAT GAG ACT ATC, CMo 3509 TTC TAT GCC ATT ATC AGG, CMo 3510 TGA TTG ATC ACG ATG CGG, CMo 3511 AGA TCA GAT ATG GTA AGG, CMo 3512 GTT CAG AAA AAT GTC GAC, CMo 3513 AAA CGA TAA GGC GAA ACC, CMo 3514 CTG ACC AAT TCC ATC ATC, CMo 3515 TTT AAC TGA AAC CCG AGG, CMo 3516 GAT AGC AAA GAG ATG CC, CMo 3517 TAC ATA

CCG GGT AAC TCG, CMo 3518 GTC AAC TTT GAT GAA TCG, CMo 3519 TGC CAA TTG TCG TGT TGC, CMo 3520 GTC ATA TGC AAT CCA AGC, CMo 3521 TCA TCT TGG TCT TTA TCC G, CMo 3522 GTC AGG AAT TCG CTC AGG, CMo 3523 CTC TCT ATC ATA TGG TAC, CMo 3524 CGT ATT GTA AAC GTC CAG and CMo 3525.

Sequencing of */mut-16*

CMo 6409 CAC GCT CAC CAC GAT CTC C and CMo 6414 GAA GCG ACT GTT CAT TAG were used to PCR-amplify the 5' genomic region of *mut-16*. The PCR product was then sequenced with the following primers: CMo 6391 GCA CTG GAA ATT GAA GAA C, CMo 6392 CGG AAA CTG ATC GTG AAG, CMo 6393 CCG TTC TGG ATA CTC CTC, CMo 6394 GTC ACG AAA AGC TCT GAA C, CMo 6395 CTC CTT GCC AAA TCA CCA C, CMo 6396 CGA GGA ACA CTG AAG AAG, CMo 6397 GCT GGG GGC ATT TCT AAC A. CMo 6398 GGA GAT CGT GGT GAG CGT G and CMo 6408 GAC TCC CGG TTC TTC TAC were used to PCR-amplify the 3' genomic region of *mut-16*. The PCR product was then sequenced with the following primers: CMo 6398, CMo 6399 CAA ACA CCA ACG ATA TTC CG, CMo 6400 GAC GCA ATG TTC AAC TTC C, CMo 6401 GCC TCC GTG CAT TTC TGT AC, CMo 6402 GTC CAA CAC AGT CAG AAA C, CMo 6403 CGC TCA AAT GGA CCC AAA TC, CMo 6404 GGC ACC ACA GTA TCC ACC A, CMo 6405 CCT CCA GAA TAC GAT CCA C, CMo 6406 CGA AGC ATG AAC TTG ACA G, CMo 6407 GAA CAG AAG TGA GCA GAC.

Sequencing of *rde-3*

CMo 7711 GAT TGT ATT GTT TCT TTT GTC TTA TAC CGG G and CMo 7712 GGA TCG ATT TCT GTC ATC GAT GCA C were used to PCR-amplify the 5' genomic region of *rde-3*. The PCR product was then sequenced with the following primers: CMo 7713 CTC CCT GTC GAA GGC TCC G, CMo 7714 CGC GAT GGT TTA GGC ACT TC, CMo 7715 CCT GTA ATG GTC ATG GAT CGG TAG, CMo 7716 GAA GAA GAT CCA ATT GGT TAA TTT GTA TTC, CMo 7717 CCA ATA CTG CCA AAT TTG CAA GAA ATC, CMo 7718 CAT CCC AAG CAA TGA AAT TAT CCT CTG. CMo 7719 CCT GTT GAA TCC TCC ATC TTT TGG ATC and CMo 7720 GCA CAT TTG TGC ATT TCC TTC CAG were used to PCR-amplify the 3' genomic region of *rde-3*. The PCR product was then sequenced with the following primers: CMo 7721 CAC AGA TGA CGA ACG ACA AGC ATC, CMo 7722 GAA CGA GCT GGA TTC CAT GTT CCC, CMo 7723 CCG GTT ACT GTA CTT CCA GTT GGC, CMo 7724 CAT TTA ATT CAA AGC AAC CTG TTT CG, CMo 7725 GAG TTC CTT ATT TTC AGC CAA AAG AGG.

2´-*O*-methyl oligo binding assay

The 2´-*O*-methyl oligo binding experiments were performed as described in Yigit *et al.* (Yigit et al., 2006)

Results

Strategy of large-scale *rde* screen

In the initial screens performed by Tabara *et al.*, 1999, *rde* mutants were selected by feeding of bacteria expressing dsRNAs targeting a germline-expressed essential gene. The *rde* mutants were isolated because they were resistant to the lethality produced by the bacterial food. In addition to *rde-1*, *rde-2*, *rde-3* and *rde-4*, three more *rde* loci were identified in this type of screens, *rde-5*, rde-6 and *rde-7*. Both *rde-5* and *rde-6* are also of the mutator class and are allelic to *mut-15* and *mut-16*, respectively (Vastenhouw et al., 2003). *rde-7* is allelic to *sid-1* (Winston et al., 2002). Furthermore, in another small-scale genetic screen in search of *rde* mutants resistant to promoter driven *unc-22* RNAi (somatic RNAi), a new *rde* locus, *rde-9*, was identified. Cloning of *rde-9* revealed that *rde-9* encodes the same gene as *rrf-1* (Darryl Conte and Craig C. Mello, unpublished)(Sijen et al., 2001).

We reasoned that at least two classes of mutants were omitted from these genetic screens: those that are only resistant to somatic RNAi and those that are required for both RNAi and development. In an attempt to recover these mutants, I first constructed a bacterial strain that simultaneously expressed a chimeric dsRNA targeting *let-2*, an essential somatic gene, and *gfp* (Figure II-1a). When fed to transgenic worms expressing *gfp* in the somatic tissues under the *pes-10* promoter, this strain causes gfp silencing and lethality at larval stage 4 (L4) caused by bursting phenotype.

For one representative screen, 1.5 million worms (representing 3 million haploid genomes) were mutagenized using chemical mutagens ethyl methanesulphonate (EMS)

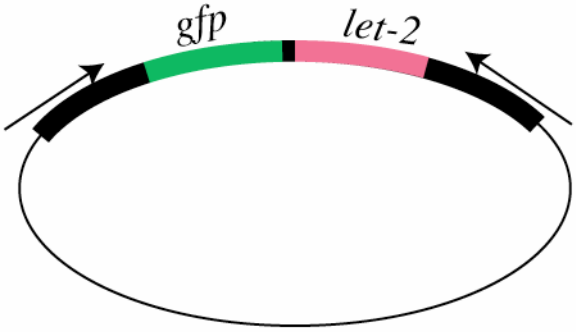
or ethylnitrosourea (ENU). The mutants were then plated in independent pools and allowed to homozygoze the mutation by self-fertilization at the F2 generation. The F3 animals were selected on *let-2/gfp* RNAi food in the F3 generation for viable animals (Figure II-1b). These animals would show up as worm bags due to the Egl phenotype caused by the *lin-11* mutation. Viable animals were clonally isolated for further analysis to select for mutants that were strongly resistant to RNAi as described below.

Individual mutants were first tested for RNAi targeting pos-1 (germline RNAi), let-2 (somatic RNAi) and gfp (germline and somatic RNAi) by feeding and scored for sensitivity to RNAi. Mutants that were only resistant to somatic RNAi were tested for complementation with rrf-1, which is required for only somatic RNAi, whereas mutants that are resistant to all RNAi by feeding were then tested for somatic RNAi by injection. Mutants that were strongly resistant to RNAi by injection were then subjected to complementation tests with candidate RNAi loci including *rde-1* and *sid-1*. Lastly, mutants that complemented *rde-1* and *sid-1* were analyzed by further mapping.

The results after the primary categorization are summarized in Table II-1. 17 alleles of *rde-1*, 22 alleles of *sid-1*, and 27 alleles of *rrf-1* were isolated. The high rate of recovery of the alleles of these mutants is likely due to the larger size of these genes which provides more sequence to be targeted in the mutagenesis. In addition, we also found 3 *rde-3* alleles, 2 *rde-5* alleles and 1 *rde-6* allele. Three alleles that define a new *rde* locus, *rde-8*, were also isolated from this screen. The isolation of the additional alleles of the *rde-3* locus facilitated the cloning of the *rde-3* gene (see Chapter III). Characterization of some of these mutants will be described in separate sections in this

Figure II-1

A.



B.

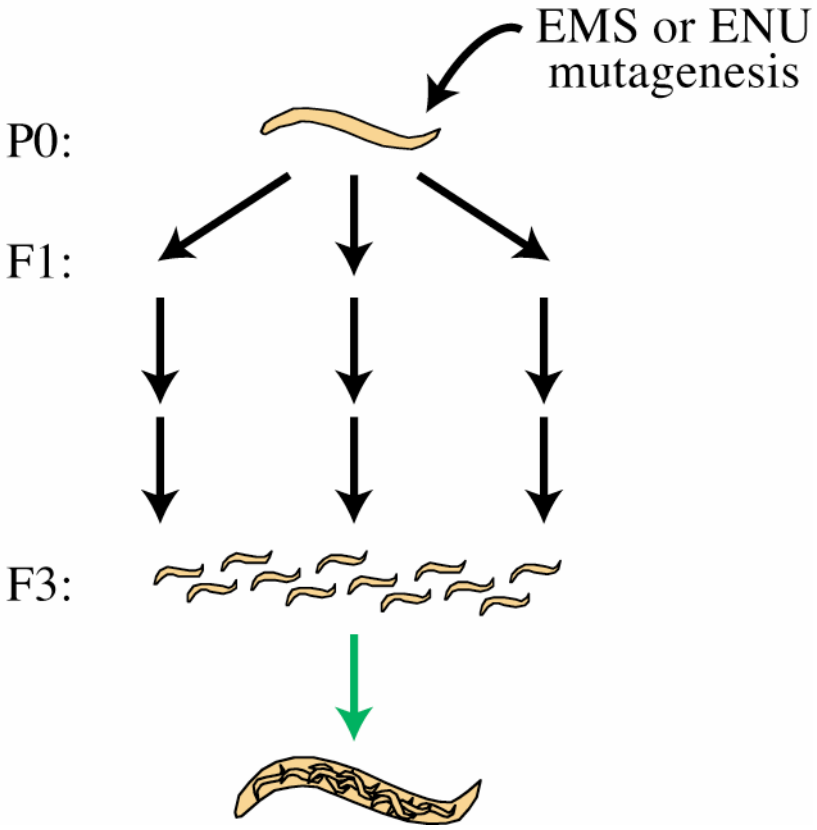


Figure II-1 Genetic screens in search of *rde* mutants

A) The schematic presentation of the bacterial feeding strain used for the screens. A 3' fragment of the *let-2* gene coding sequence of about 700bp were cloned in to the L4417 Vector from the Fire Lab Plasmid kit which contains two T7 promoters oriented at opposite orientations and the coding sequence of the *gfp* gene. The arrows indicate the orientation of the T7 promoters.

B) The schematic presentation of the *rde* genetic screens. See text for details.

chapter. Identification and characterization of *rde-8* will be presented in Chapter IV.

Characterization of two *rde-1* alleles

Sequencing analysis of the new *rde-1* alleles showed that four alleles *ne4085*, *ne4086*, *ne4087* and *ne4088* encode missense mutations in conserved residues of G283E, I357L, L457W and S697F, respectively (Figure II-2b). Examination of the presence of the mutant proteins by Western blotting reveals that only *ne4085* and *ne4086* produce stable proteins (Duchaine T, personal communication). To test if these mutations affect the association between RDE-1 and primary siRNAs, we performed 2´-*O*-methyl binding experiments with protein lysates prepared from these mutants. In this assay, the interaction between RDE-1 and the trigger-derived single stranded RNAs (i.e. primary siRNAs) can by captured by an oligo matrix with sequences homologous to the trigger dsRNAs. Both mutant proteins interact with siRNAs, demonstrating that the mutations do not abolish the protein/RNA interaction (Figure II-2a). Instead, these mutant proteins could affect other steps such as the endonucleolytic activity of RDE-1 or the interaction with other proteins required for initiating the downstream event.

Identification of the *rrf-1* alleles

We identified 27 alleles of *rrf-1* by phenotypic analysis, complementation test and sequencing of the *rrf-1* genomic locus. As shown in Figure 3a, 13 alleles bear early stop codon mutations whereas *ne3298* encodes an allele with a small deletion. Interestingly, all but one missense mutation occurred in the middle-third of the protein, where the

Table II-1

rde **mutant**	**RNAi sensitivity**[a]		**Additional phenotypes**[b]	**Number of new alleles isolated**	**Gene description**
	Germline	**Soma**			
rde-1	R	R	Not observed	27	Argonaute
rde-2	R	S	Him, Mut	0	Novel
rde-3	R	R	Him, Mut	3	Chapter III
rde-4	R	R	Not observed	0	dsRNA binding
rde-5	R	R	Him, Mut	2	*mut-15*/ Novel
rde-6	R	R	Him, Mut	1	*mut-16*/ QN rich domains
rde-7	R	R	Not observed	22	*sid-1*
rde-8	R	R	Not observed	3	Chapter IV
rde-9	R	R	Not observed	27	*rrf-1*/ RdRP

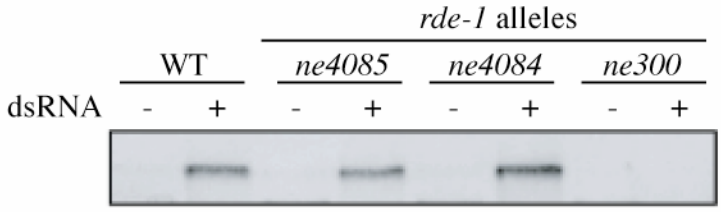
Table II-1 Summary of the *rde* mutants isolated from the genetic screens

[a] The sensitivity of the *rde* mutants were tested by feeding using RNAi targeting the *pos-1* gene (germline), and the *unc-22* gene (soma).

[b] *Him*: high incidence of males; *Mut*: mutator.

Figure II-2

A.



B.

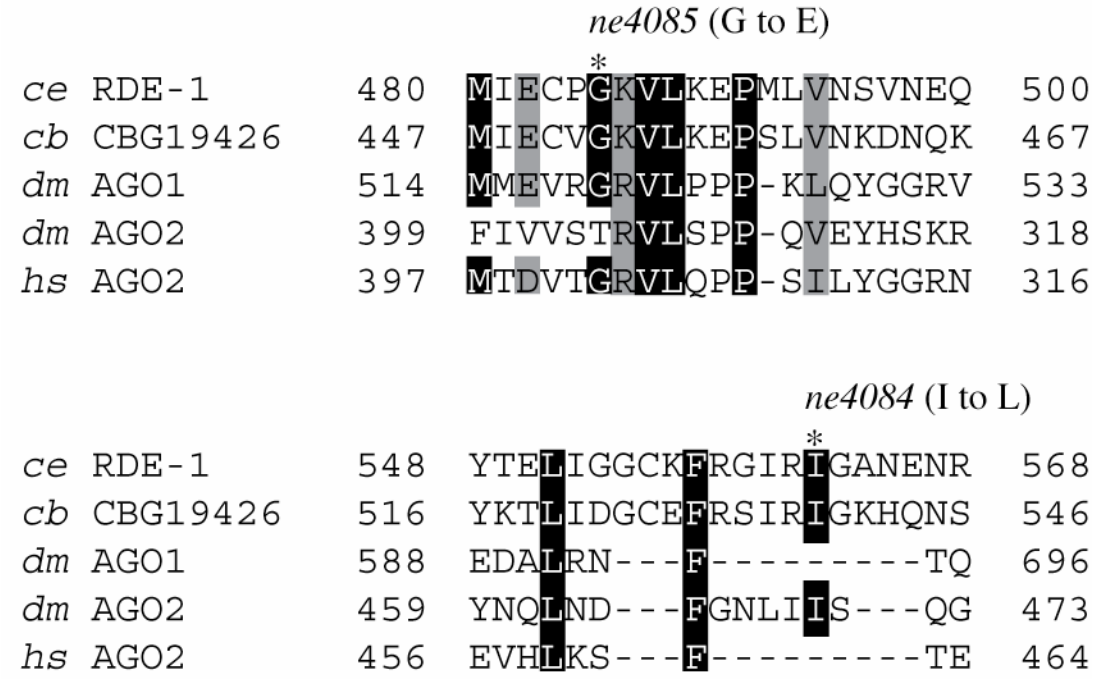


Figure II-2 Characterization of two *rde-1* alleles

A) 2´-O-Methyl experiments on two alleles of the *rde-1* mutant that bear missense mutations were performed. RDE-1 proteins bound to the sense 2´-O-Methyl sequence columns were detected by Western blot using anti-RDE-1 antibody. The wildtype worms and these two alleles express stable mutant proteins while an allele of the *rde-1* mutant that contains an early-stop-codon mutation does not express any RDE-1 protein. Both mutant proteins interact with the sense 2´-O-Methyl matrix.

B) The alignment of the Argonaute proteins including RDE-1 in *C. elegans* (*ce*), CBG19426 in *C. briggsae* (*cb*), Drosophila AGO1 and AGO2 (*dm*) and human AGO2 (*hs*). The mutations in these *rde-1* alleles are marked by an asterik.

function of the RdRPs. Identification of these residues that are essential for *rrf-1* activity will provide further insights when they are used for biochemical analysis (see discussion).

Identification of *rde-5/mut-15* alleles that exhibit temperature sensitive sterility

The sequencing of two additional *rde-5/mut-15* alleles, *ne3358* and *ne3359*, identified single nucleotide changes that introduce an early stop codon and splice acceptor mutation, respectively (Figure II-4a). The early stop codon mutation in *ne3359* truncates more than two-thirds of the protein and most likely generates a null phenotype.

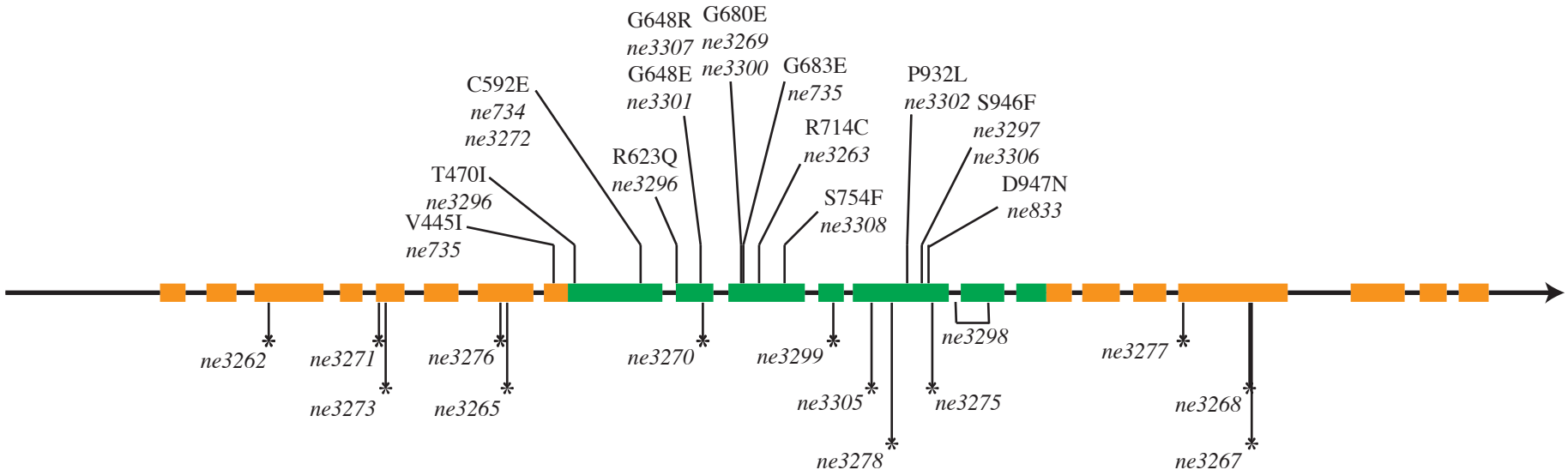
Interestingly, we observed that *ne3359* exhibit a strongly penetrant sterile phenotype at 25°C, in which the oocytes in the proximal region of the gonad "stack up" and become tightly packed (Figure II-4b). This phenotype is reminiscent to that of the *fem* mutants in which the germline is *fem*inized and does not produce sperm, so we examined whether this defect can be rescue by providing external sperm via wildtype males. The result showed that *rde-5/mut-15* hermaphrodites mated with wildtype males produced wildtype progeny, indicating that the sterile phenotype is due to a sperm defect in *rde-5/mut-15* hermaphrodites.

Characterization of two temperature sensitive *rde* mutants

Two mutants exhibited temperature sensitive Rde phenotype. When exposed to RNAi food at permissive temperature of 15°C, both mutants were 100 percent sensitive to RNAi. When the temperature rises above 20°C or increases to a non-permissive

Figure II-3

A.



B.

Figure II-3 Mutations in the *rrf-1* locus

A) The rrf-1 locus is shown. All the non-sense mutations are marked below the gene whereas all the missense mutations are marked above the gene. All missense mutations are within the highly conserved C-terminal region. The exons and introns are represented by boxes and lines, repectively. The green boxes markes the "HS" (high similarity) region (See discussion).

B) Three missense mutations fall in the "HS" region. The alignment of the C-terminal portion of the "HS" region is shown as well as the mutations.

Figure II-4

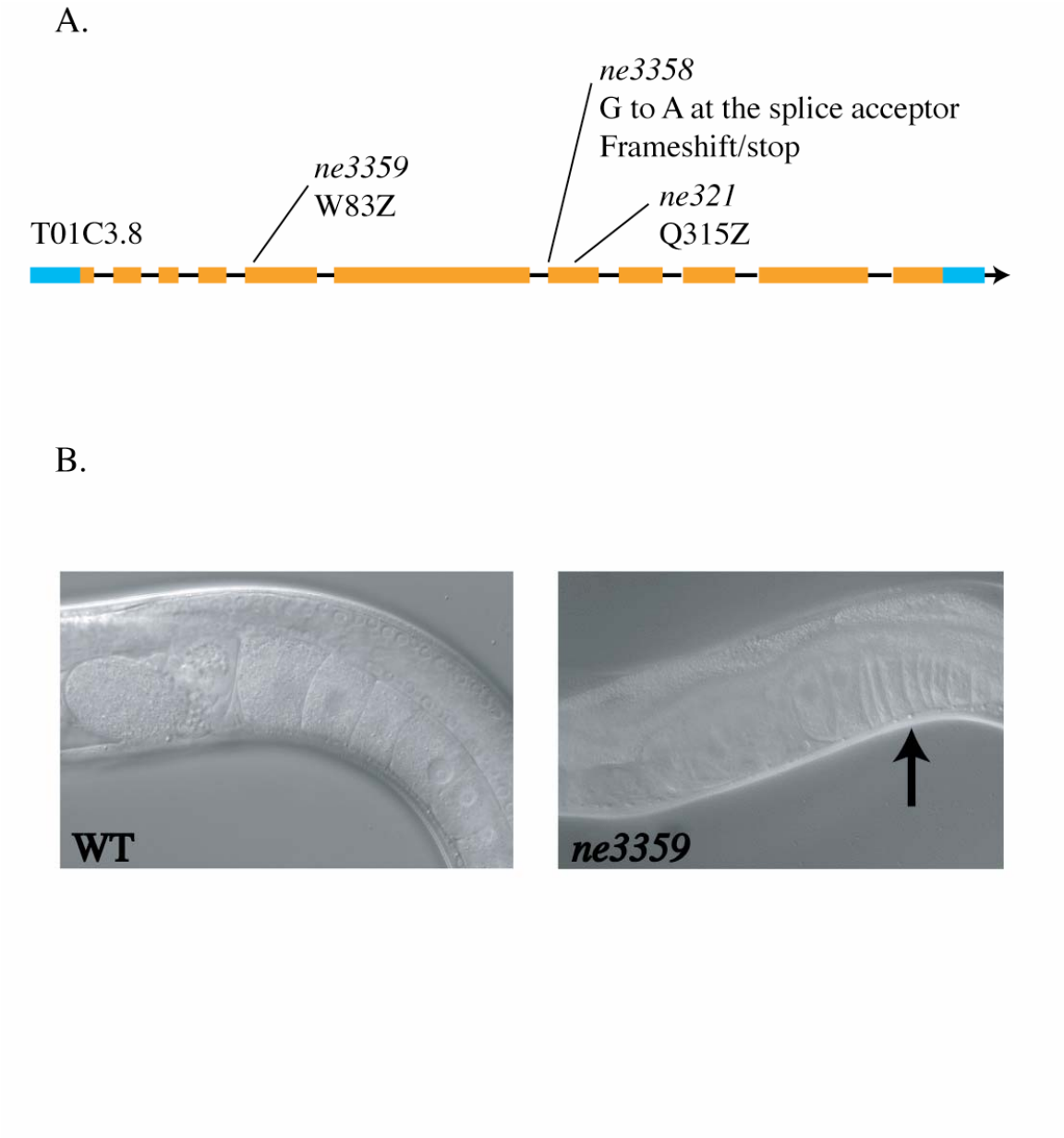

A.
ne3358
G to A at the splice acceptor
Frameshift/stop
ne3359
W83Z
ne321
Q315Z
T01C3.8
B.
WT
ne3359

Figure II-4a Characterization of the *mut-15/rde-5* locus

A) The *mut-15/rde-5* gene and mutations identified are shown. Two mutants bear early-stop-codon mutations while the third one contains a G to A change at the splicer acceptor of exon 7. This mutation causes a frameshift in the subsequent exon which introduces several early stop codons downstream. Exons and introns are represented by boxes and lines, respectively. Blue boxes represent untranslated regions whereas orange boxes represent coding regions.

B) The temperature sensitive sterility of the *mut-15/rde-5 (ne3359)* mutant

In the germline of the wildtype hermaphrodites, the oocytes are produced from the distal end of the gonad and move toward the proximal gonad where the spermatheca locates. Once the oocytes move through the spermatheca, the sperm fertilize the oocyte and produce embryos at the proximal site. In the *mut-15/rde-5 (ne3359)* mutant, fertilization fails due to a sperm defect and as a result the oocyte accumulates in the gonad and exhibits a “stacked oocyte” phenotype which is marked by the black arrow.

temperature (for viability) of 25°C, both mutants show 100 percent resistance to RNAi. Complementation tests and sequencing analysis showed that *ne3310* is an allele of *mut-16* and *ne3263* is an allele of *rrf-1*.

mut-16 encodes a novel protein a with glutamine/asparagines (Q/N)-rich domain-containing protein. In addition to functioning in RNAi, *mut-16* is also required for transposon silencing and cosuppression. The mutation in the *mut-16* (*ne3310*) mutant was identified to encode a D94N change in a conserved residue (Figure II-5). The *rrf-1* gene encodes a RNA dependent RNA Polymerase (RdRP). In order to test where in the RNAi pathway these two mutants function genetically, we performed a temperature shift assay. In the temperature shift-up experiments, we exposed the worms to a bacterial food containing *let-2* dsRNAs at the permissive temperature to initiate RNAi for approximately 24 hours, and shifted the worms up to the non-permissive temperature to assay for their resistance to RNAi (Figure II-6). We reasoned that once RNAi is initiated at the permissive temperature, mutants that are required for the initiation would not be affected by the temperature shift. In contrast, mutants involved in the effector step would not have a functional activity at the non-permissive temperature. Thus, RNAi will be abolished at the higher temperature. Therefore, mutants acting at the initiation step will render sensitivity to RNAi whereas mutants acting at the effector step will be resistant to RNAi in this assay. As shown in Table II-2, both *rrf-1 (ne3263)* and *rde-6 (3310)* exhibit resistance to RNAi at non- permissive temperature, suggesting that neither mutant is required for initiation, but rather function at a downstream effector step. In the reciprocal shift-down experiments, RNAi is initiated at the non-permissive temperature. The

mutants defective in the initiation is likely to render resistance because RNAi can not be initiated whereas the mutants defective in the effector step are not affected. Both *mut-16* and *rrf-1* are sensitive to RNAi in the shift-down experiments, indicating that the initiation of RNAi is not affected in these mutants. Taken together, these experiments suggest that *mut-16* and *rrf-1* act in an effetor step downstream of RNAi.

Isolation of a temperature sensitive lethal allele of *drh-3*

We reasoned that one way to recover temperature-sensitive lethal *rde* mutants was to test for the RNAi resistance at an intermediate temperature. This intermediate temperature will, in theory, allow the mutants to tolerate the lethal effect caused by the mutation thus allowing the mutants to survive during the test. Because the analysis of the temperature sensitive mutants showed that at 20°C, these mutants already exhibit a strong resistance to RNAi, it suggested that such a condition is attainable. Therefore, by performing the screen at this temperature, we can first identify mutants that are potentially temperature sensitive *rde* and then test for viability by further shifting the temperature to non-permissive state.

The mutant *rde-x* was isolated from a pilot screen using the rationale described above. *rde-x* was resistant to *pos-1* and *let-2* RNAi at 20°C by feeding. In addition, when shifted to non-permissive temperature, *rde-x* exhibits 100% sterility of adult animals. Further genetic mapping placed *rde-x* on LGI. Complementation test with candidate genes showed that *rde-x* failed to complement *drh-3*. Finally, sequencing of the drh-3

Figure II-5

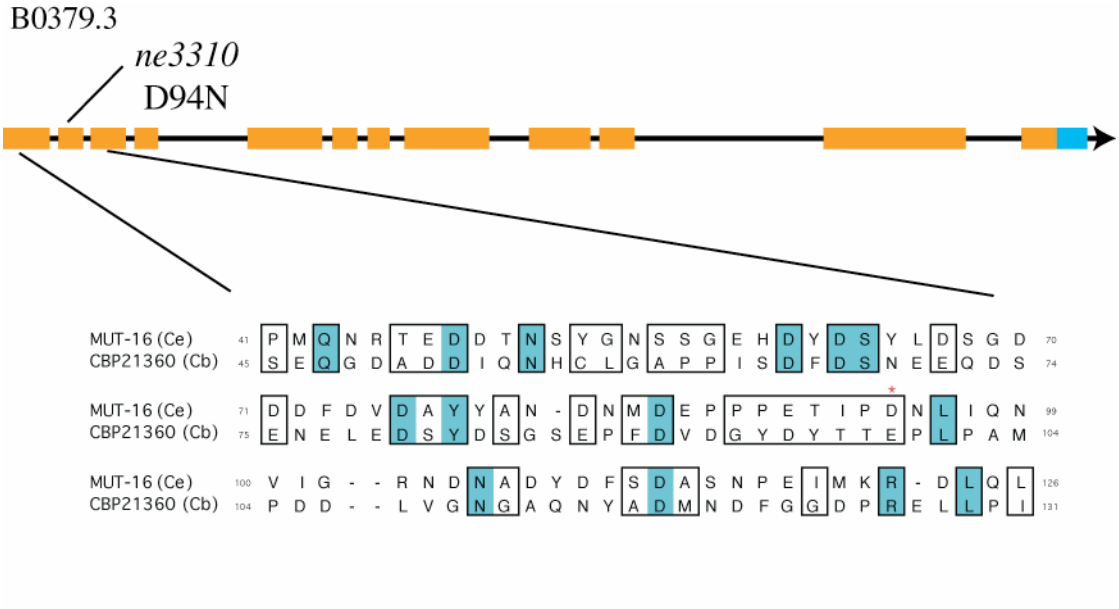

B0379.3
ne3310
D94N
MUT-16 (Ce) 41 P M Q N R T E D D T N S Y G N S S G E H D Y D S Y L D S G D 70
CBP21360 (Cb) 45 S E Q G D A D D I Q N H C L G A P P I S D F D S N E E Q D S 74
MUT-16 (Ce) 71 D D F D V D A Y Y A N - D N M D E P P P E T I P D N L I Q N 99
CBP21360 (Cb) 75 E N E L E D S Y D S G S E P F D V D G Y D Y T T E P L P A M 104
MUT-16 (Ce) 100 V I G - - R N D N A D Y D F S D A S N P E I M K R - D L Q L 126
CBP21360 (Cb) 104 P D D - - L V G N G A Q N Y A D M N D F G G D P R E L L P I 131

Figure II-5 The mutation in the temperature sensitive mut-16/rde-6 mutant allele

The mutation of the ne3310 allele is located in the exon 2 of the gene. A protein sequence alignment of mut-16/rde-6 and its homolog in *C. briggsae* is shown. The red asterik marks the mutated residue which is conserved in both proteins. Exons and introns are represented by boxes and lines, respectively. Blue boxes represent untranslated regions whereas orange boxes represent coding regions.

Figure II-6

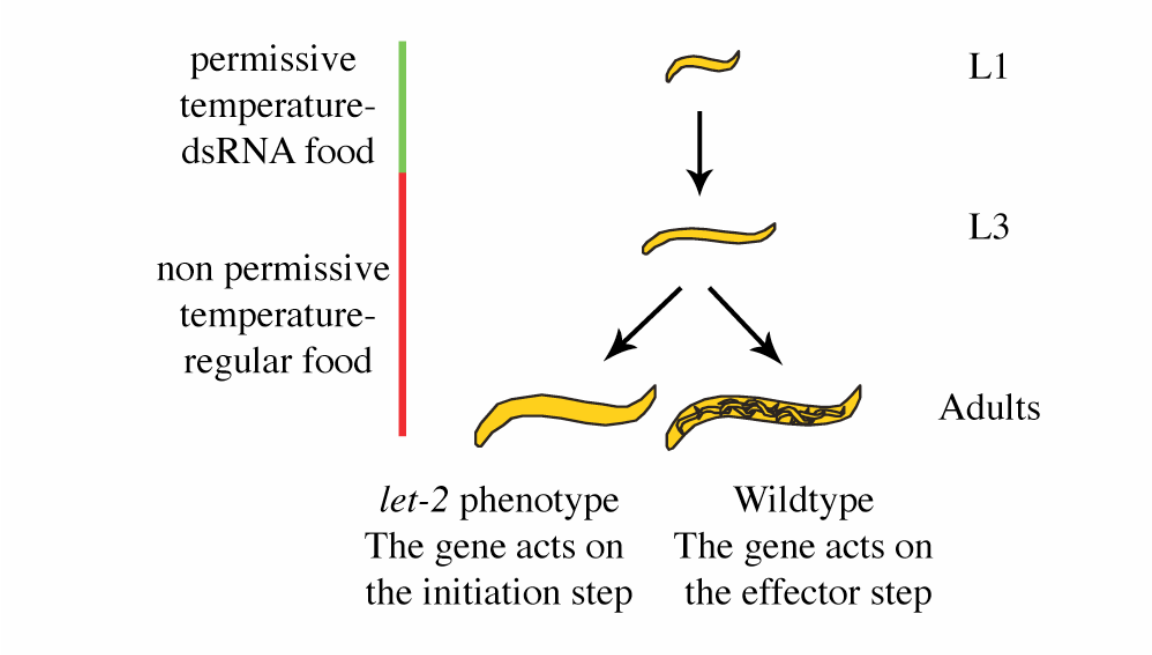


Table II-2

	RNAi Resistance (Viability)		
	Wildtype	*rrf-1(ne3263)*	*mut-16(ne3310)*
15°C → 25°C	5%	96%	90%
25°C→ 15°C	2%	10%	15%
15°C →15°C	0%	0%	0%

Number of worms tested/strain=90

Figure II-6 Schematic diagram of the temperature sensitive RNAi assay

See text for detail.

Table II-2 The *mut-16* gene and the *rrf-1* function at the downstream step of RNAi

The temperatures indicated are the initial temperatures in which the worms are exposed to dsRNAs by feeding followed by the temperatures in which the worms are moved to regular plate. Both *mut-16* and *rrf-1* are resistant RNAi even after the worms are exposed to RNA.

locus revealed a glycine to aspartic acid residue change in a glycine residue that is conserved throughout the DRH gene family, suggesting *rde-x* is *drh-3*.

drh-3 encodes a conserved helicase protein that associated with DCR-1. Analysis of a deletion mutant of *drh-3* suggested that *drh-3* is required for both germline RNAi and fertility. Therefore, isolation of this missense allele of drh-3 further validates the rationale and strategy of the temperature sensitive rde screen.

Discussion

Analysis of the *rde* mutants

Analysis of the missense mutations in *rde-1* mutant alleles revealed several important residues that are critical for the function of RDE-1. For example, a previously described allele of *rde-1*, *ne219*, encodes a Glu to Lys change. Subsequent NMR structure analysis of the PAZ domain in the *Drosophila* Ago2 protein revealed that the corresponding residue is important for stabilizing the fold of the PAZ domain(Lingel et al., 2003). In the screens described here, we have isolated two missense alleles of the *rde-1* gene which do not affect the stability of the RDE-1 protein or the interaction between RDE-1 and the primary siRNAs. Both mutations are found outside of the PAZ or PIWI domain. The mutations in these two alleles do not affect the interaction of RDE-1, as shown in Figure II-2 with siRNAs but rather could affect downstream events such as the endonucleolytic activity of the PIWI domain or interaction between RDE-1 and other protein components in the RDE-1 complex to create an siRNA-containing but non-

functional RDE-1 complex. Further biochemical analysis of the functional defects of these mutant proteins will shed light on our understanding of how the RDE-1 complex functions.

We have also isolated 12 alleles encoding missense mutations in the *rrf-1* gene which encodes a RdRP homolog in the RNAi pathway. Sequence analysis of cellular RdRP homologs in fungi, worms and plants demonstrated that the carboxyl terminus of these proteins is more conserved that the amino terminus. Interestingly, all the missense mutations we have isolated are found within this conserved region. Sequence comparison of the cellular RdRPs to the motifs in the viral RdRPs that catalyze nucleotide transfer also revealed a "high similarity" (HS) region with conserved aspartates that appears to resemble the catalytic aspartate in the viral RdRPs(Makeyev and Bamford, 2002). Three *rrf-1* alleles *ne3297*, *ne3306* and *ne833* possess mutations in the HS region, with *ne833* bearing a mutation precisely in a conserved aspartatic acid (Figure II-b).

One allele of *rrf-1*, *ne3263*, exhibits a temperature sensitive RNAi deficient phenotype in which the mutant worms are sensitive to RNAi at lower temperature but strongly resistant to RNAi when shifted to higher temperature. The isolation of this allele allowed for the genetic analysis of the requirement of the *rrf-1* gene in the RNAi pathway, which suggests that *rrf-1* functions at the effector step. The molecular lesion in *ne3263* is a residue that is also conserved in *ego-1*, a *C. elegans* RdRP homolog of *rrf-1*. While ego-1 plays a role in germline RNAi, *rrf-1* is required for somatic RNAi (Smardon et al., 2000). Overexpression of the EGO-1 protein in the soma under a muscle specific promoter, *myo-3*, rescues the RNAi defect in *rrf-1* mutants, suggesting that these two

proteins are functionally redundant. *ego-1* is also required for proper germline development of *C. elegans*(Qiao et al., 1995). Mutation in the *ego-1* gene causes a sterile phenotype which impedes analysis of the function of *ego-1* in RNAi. Because the *ne3263* lesion is also conserved in ego-1, we question whether this mutant protein, when expressed under *ego-1* promoter, can rescue the activity of EGO-1 at permissive temperature but abolish EGO-1 function at non-permissive temperature in an *ego-1* null mutant. Generation of such a synthetic temperature sensitive *ego-1* strain will greatly advance further biochemical analysis of *ego-1* in both RNAi and development.

Biochemical studies on the activity of the cellular RdRP in fungi, QDE-1, have demonstrated that QDE-1 exhibits two modes of polymerase activity (Makeyev and Bamford, 2002). The full-length protein synthesizes long dsRNAs from a single stranded RNA template and the catalytically active carboxyl terminus produces complementary small RNAs of ~9 to 21 nucleotides scattered along the template, suggesting that more than one activity could be employed for the silencing pathways. It is unclear what the biochemical activity of RdRPs is in *C. elegans*. The isolation of an array of missense rrf-1 alleles provides a toolbox for the biochemical analysis of the catalytic activity of this enzyme, as well as the other RdRP homologs, which will shed light on the origins and synthesis of RNA intermediates during RNAi.

Identification of *rde* mutants from genetic screens

We have conducted a large-scale *rde* screen that allowed us to recover additional mutant alleles of all previously isolated mutants except for *rde-4*. Some of these novel

alleles provide us with information about residues that are important for the gene function. In addition, two of the mutant alleles which exhibit a temperature sensitive *rde* phenotype allowed us to demonstrated that *rrf-1* and *mut-16/rde-6* function in the effector step of the RNAi pathway. Furthermore, we have also observed a temperature lethal phenotype in a presumptive null allele of *rde-5/mut-15*. Future analysis of these alleles will provide us with further insight on the mechanism of RNAi as well as developmental pathways essential for the viability of the animal.

From these screens, we have also identified a novel *rde* locus, *rde-8*. *rde-8* was not isolated in the recent genome-wide RNAi-based screens, searching for genes required for RNAi, indicating that these two methods complement each other to identify novel genes in the RNAi pathway. The isolation of *rde-8* also suggests that genetic screens in search of mutants that affect RNAi but not viability are not yet saturated. However, our results also present a challenge for this type of screens: mutant alleles of large-sized genes, for example *rde-1, rrf-1* and *sid-1*, are a bigger target for mutagenesis and will represent most of the mutant population, thus leading to a time-consuming search or a loss of other novel mutants. One solution to this problem is to introduce transgenes expressing these large RNAi genes so that an extra wildtype copy can rescue the mutant phenotype when the endogenous copy is mutated and allow for the isolation of novel genes. Although identification of genes required for RNAi can be conducted in a variety of approaches such as biochemical purification, reverse genetics, and RNAi-based screens, the isolation of mutants deficient in RNAi from genetic screens allows both the

identification of the genes and the generation of useful mutant alleles that can be used for future analyses.

Improving the strategy for temperature sensitive *rde* screen

The recovery of two temperature sensitive alleles from our initial large-scale *rde* screen has provided a basis for designing future screens to identify temperature sensitive lethal *rde* mutants. We observed that at a semi-permissive temperature of 20°C, both of the temperature sensitive *rde* mutants are resistant to RNAi. In other independently conducted genetic screens in our laboratory in search of temperature sensitive embryonic lethal mutants, we noticed that at the same semi-permissive temperature, most mutants could produce viable progeny. Therefore, we hypothesize that at this semi-permissive temperature, mutations that affect both RNAi and viability can be identified as viable *rde* mutants. Our preliminary screen allowed us to recover an allele of *drh-3* mutant which exhibits RNAi resistance and a sterile phenotype at the non-permissive temperature but is viable at the permissive temperature, which validates the rationale of the screen. Although further modifications are needed to improve the efficiency of the screen, this strategy will provide a starting point to identify genes that are involved in both RNAi and developmentally essential pathways.

Acknowledgement

We would like to thank Shi-Chang (Bill) Tsai and Martin Simard who participated in the temperature sensitive *rde* screen for their assistance. Some of the nematode strains

were provided by *Caenorhabditis* Genetics Center, funded by National Institutes of Health National Center for Research Resources.

Chapter III

A member of the polymerase ß nucleotidyltransferase superfamily is required for RNA interference in *C. elegans*

Chun-Chieh G. Chen, Martin J. Simard, Hiroaki Tabara[1], Daniel R. Brownell[2], Jennifer A. McCollough, and Craig C. Mello[2,3]

Program in Molecular Medicine and [2]Howard Hughes Medical Institute, University of Massachusetts Medical School, Worcester, Massachusetts 01605

[1]Graduate School of Medicine, HMRO, Kyoto University, Kyoto, 606-8501 Japan

This chapter was published in Current Biology, Vol. 15, 378-383 (2005).

Chun-Chieh G. Chen cloned the *rde-3* gene and assisted Daniel R. Brownell in sequencing and identifying lesions in the *rde-3* alleles. She performed the analyses in Figure III-2 and Figure III-3.

Summary

RNA interference (RNAi) is an ancient, highly conserved mechanism, in which small RNA molecules (siRNAs) guide the sequence-specific silencing of gene expression (Mello and Conte, 2004). Several protein components of the silencing machinery have been identified, including helicases, RNase-related proteins, double and single-stranded RNA binding proteins and RNA-dependent RNA polymerase-related proteins (Meister and Tuschl, 2004). Work on these factors has led to the revelation that RNAi mechanisms intersect with cellular pathways required for development and fertility (Ambros, 2004; Baulcombe, 2004). Despite rapid progress in understanding key steps in the RNAi pathway, it is clear that many factors required for both RNAi and related developmental mechanisms have not yet been identified. Here we report the characterization of the *C. elegans* gene *rde-3.* Genetic analysis of presumptive null alleles indicates that *rde-3* is required for siRNA accumulation and for efficient RNAi in all tissues, and is essential for fertility and viability at high temperatures. RDE-3 contains conserved domains found in the polymerase ß nucleotidyltransferase superfamily, which includes conventional poly (A) polymerases, 2'-5' oligoadenylate synthetase (OAS) and yeast Trf4p (Aravind and Koonin, 1999). These findings implicate a new enzymatic modality in RNAi and suggest possible models for the role of RDE-3 in the RNAi mechanism.

Results and discussion

Molecular identification of *rde-3*

The *rde-3* locus was originally defined by a single allele (Tabara et al., 1999). To further characterize this locus we first conducted additional genetic screens to identify new alleles, and then used genetic mapping to define a small interval containing the gene (Figure 1A). We then examined the nucleotide sequence of candidate genes within this region for lesions in *rde-3*. In each allele, we found a mutation predicted to disrupt the expression or to change the amino-acid sequence of the gene K04F10.6. In addition, we found a lesion in K04F10.6 in the previously described mutant *mut-2* (*r459*) (Collins and Anderson, 1994), suggesting that *mut-2* and *rde-3* are allelic. Finally, we found that a PCR fragment containing only K04F10.6 rescues the somatic RNAi defect of *ne3364*, confirming that K04F10.6 is *rde-3* (Figure 1).

rde-3 encodes a conserved protein in the polymerase β nucleotidyltransferase superfamily. RDE-3, and other members in this family, including conventional poly (A) polymerases (PAP), 2'-5' oligoadenylate synthetase (OAS), *C. elegans* GLD-2 and budding yeast Trf4p, all contain a nucleotidyltransferase 2 (NTP transferase 2) domain (Aravind and Koonin, 1999). In addition, RDE-3 and a subset of family members including GLD-2 and Trf4p contain an additional domain of unknown function, named the PAP/25A (Poly A Polymerase/ 2'-5' Oligo Adenylate Synthetase) associated domain

Figure III-1

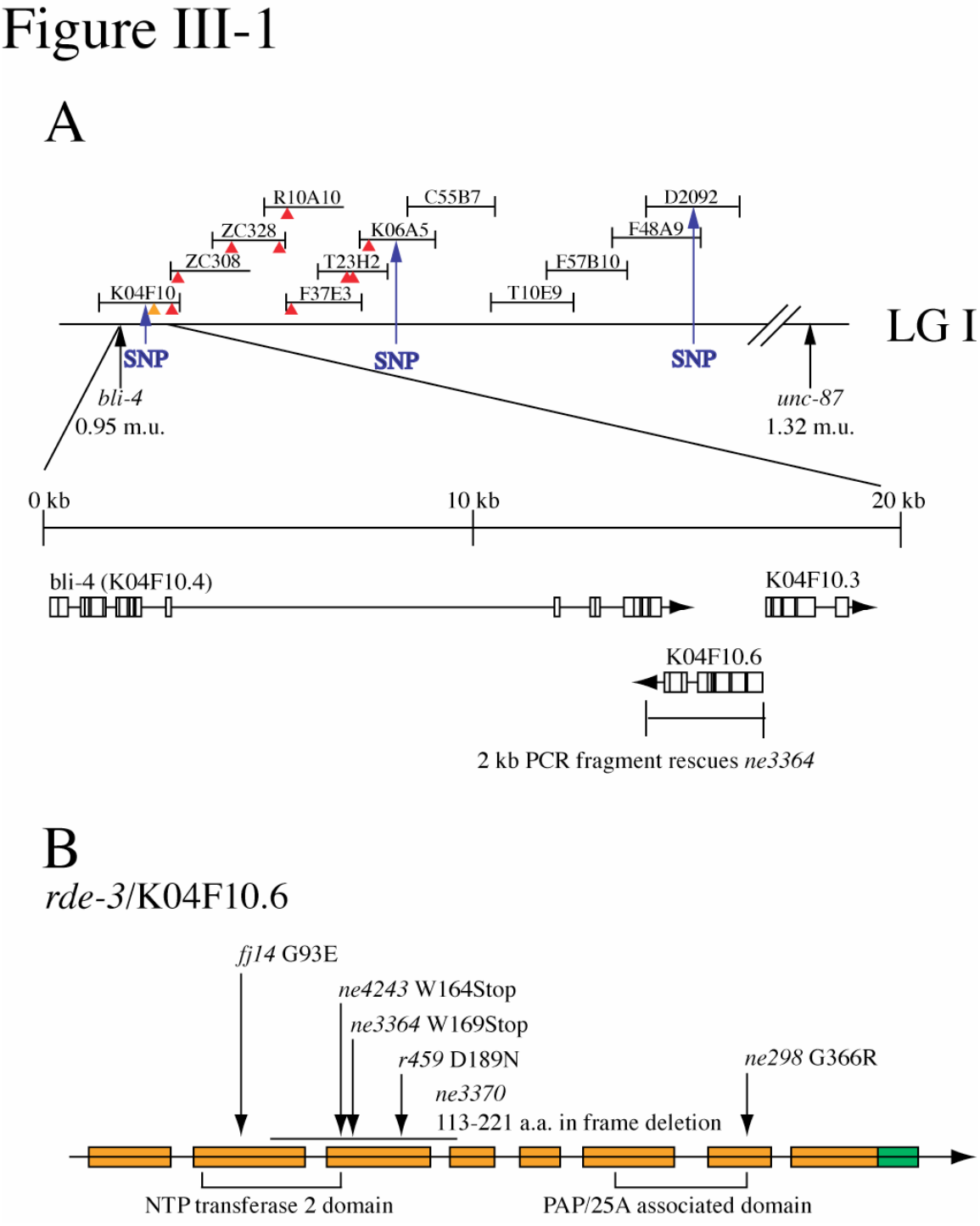

A
R10A10
C55B7
D2092
ZC328
K06A5
F48A9
ZC308
T23H2
F57B10
K04F10
F37E3
T10E9
LG I
SNP
SNP
SNP
bli-4
0.95 m.u.
unc-87
1.32 m.u.
0 kb
10 kb
20 kb
bli-4 (K04F10.4)
K04F10.3
K04F10.6
2 kb PCR fragment rescues ne3364
B
rde-3/K04F10.6
fj14 G93E
ne4243 W164Stop
ne3364 W169Stop
r459 D189N
ne298 G366R
ne3370
113-221 a.a. in frame deletion
NTP transferase 2 domain
PAP/25A associated domain

Figure III-1 Molecular identification of *rde-3*.

(A) Schematic diagram of the *rde-3* genetic interval. *rde-3* was mapped genetically to an interval of approximately 0.37 m.u., very close and to the right of *bli-4.* Single Nucletide Polymorphisms (blue arrows) were used to narrow the interval to 7 cosmids. Candidate genes shown as red triangles were sequenced. Mutations were found in K04F10.6 (orange triangle). A single-gene K04F10.6 PCR fragment was used for rescue experiments. (forward primer: GAT TGT ATT GTT TCT TTT GTC TTA TAC CGG G, reverse primer: CGT TGG AGA AAC GAA GAA TGT GCA TAG) (B) Line and block diagram of the intron/exon structure of *rde-3.* Reverse-transcriptase PCR was used to confirm the exon structure predicted by wormbase. RDE-3 contains two conserved domains including an NTP transferase 2 domain at its N-terminus and PAP/25A (Poly A polymerase/2'-5' OligoAdenylate) associated domain at its C-terminus. The molecular lesions in six *rde-3* alleles are indicated above the diagram.

(Aravind and Koonin, 1999). The *rde-3* alleles *ne298*, *r459* and *fj14* each contains point mutations predicted to alter conserved amino acids. The *ne3370* allele contains an in-frame deletion of 423bp (Figure 1B). *ne4243* and *ne3364* each contain nonsense mutations predicted to truncate the protein after 164 and 169 amino acids respectively, and are thus likely to represent null mutations. Two of the lesions in *rde-3* are predicted to alter conserved residues within the nucleotidyltransferase domain, suggesting that polymerase activity may be important for RDE-3 function (See discussion below).

rde-3 is required for fertility and viability as well as siRNA accumulation

The previous characterization of *ne298* showed that *rde-3* had additional phenotypes including a Him (High incidence of males) phenotype which reflects an increase in the non-disjunction of the X chromosome, and a Mut (Mutator) phenotype resulting from the activation of endogenous transposon (Tabara et al., 1999). The new alleles of *rde-3,* including the presumptive null alleles, exhibit phenotypes similar to those of *rde-3(ne298),* suggesting that all of these mutants represent strong loss-of-function alleles. *rde-3* mutants do not exhibit the specific developmental defects expected from a loss in the function of the miRNAs *lin-4* or *let-7* (data not shown), suggesting that, as with other mutator-class *rde* mutants, its developmental functions in fertility and viability may reflect a role in other as yet unknown mechanisms.

We next examined the sensitivity of these alleles to RNAi targeting genes expressed in different tissues: *pos-1* (germline) and *unc-22* (muscle). The three alleles

Figure III-2

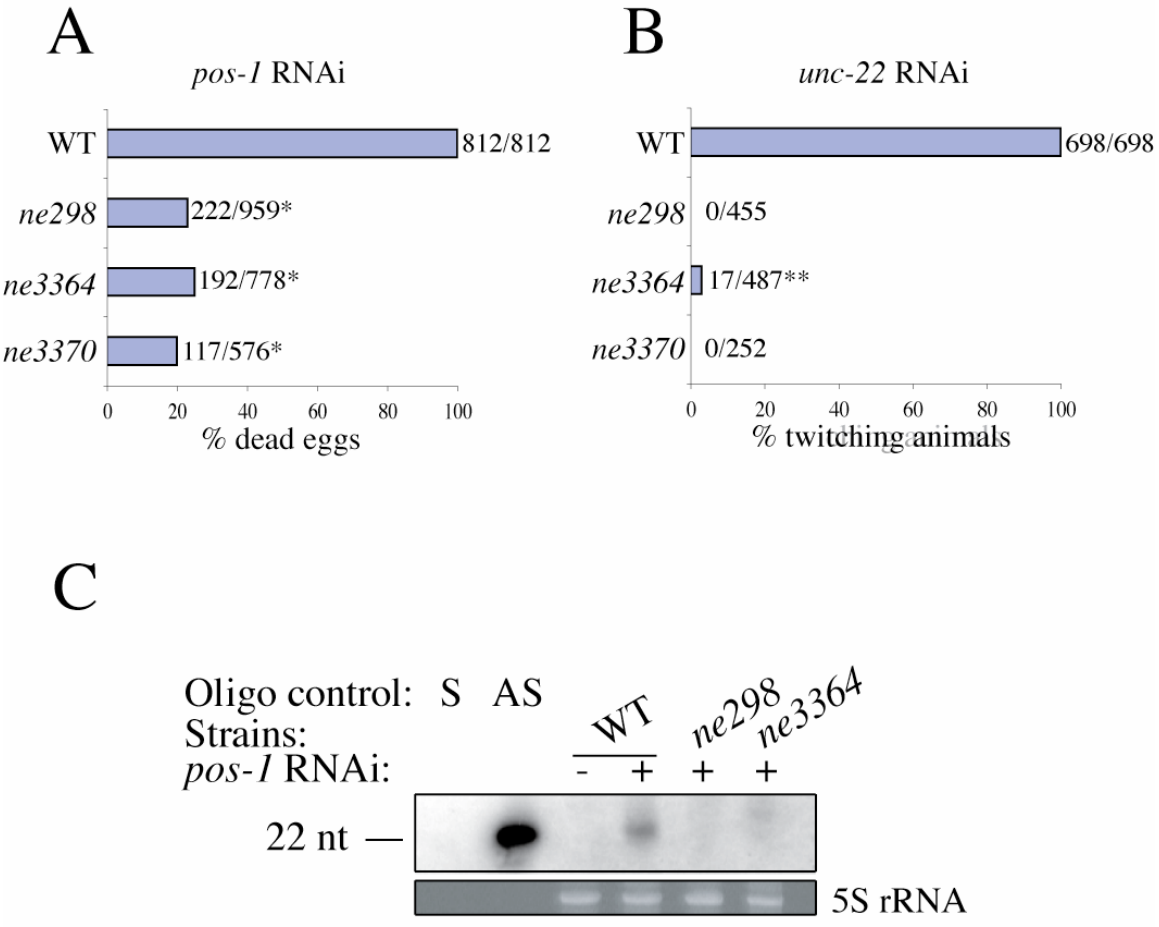

A
pos-1 RNAi
WT
ne298
ne3364
ne3370
812/812
222/959*
192/778*
117/576*
0
20
40
60
80
100
% dead eggs
B
unc-22 RNAi
WT
ne298
ne3364
ne3370
698/698
0/455
17/487**
0/252
0
20
40
60
80
100
% twitching animals
C
Oligo control: S AS
Strains:
pos-1 RNAi:
WT
ne298
ne3364
- + + +
22 nt
5S rRNA

Figure III-2 *rde-3* is defective in RNAi.

(A-B) Graphical representations of the sensitivity of three *rde-3* mutant strains to RNAi induced by microinjection of: (A) double-stranded *pos-1* RNA (1mg/mL) or (B) double-stranded *unc-22* (1mg/mL). Young adult animals were injected with dsRNA and were transferred individually to fresh plates every 12hrs. Phenotypes were scored for the progeny produced between 24 to 48 hrs post-injection. The number of sensitive animals/ over the total number of animals scored is presented next to each bar. *The data for the three *rde-3* strains includes a background of approximately 20% inviable embryos that arise spontaneously due to the incompletely penetrant embryonic lethal phenotype of *rde-3*. **These progeny were very weakly twitching in comparison to the wild-type.

(C) *rde-3* is defective in siRNA accumulation. Northern blot analysis of siRNAs in wild-type and *rde-3* mutant animals. Total RNAs were extracted using TRI Reagent (MRC) from embryos. Small RNAs were further purified using MirVana (Ambion). 10μg of RNAs were separated in 15% UREA-PAGE and transfer to nitrocellulose membrane. Northern analysis was performed using ^{32}P labeled sense ribo-probe covering the whole *pos-1* coding region (Grishok et al., 2001). Ethidium bromide staining of 5S rRNAs was used for loading control. Sense and antisense *pos-1* RNA oligos of 22 nt were included as hybridization controls.

tested (*ne298*, *ne3364* and *ne3370*) showed similar levels of resistance to RNAi (Figure 2A and B). Note that the dead embryos observed in Fig. 2a includes a background of approximately 20% inviable embryos that arise spontaneously due to the incompletely penetrant lethal phenotype associated with *rde-3* alleles (Table III-1). We found that the presumptive null allele, *ne3364*, like *ne298* (Tabara et al., 1999), is sensitive to RNAi induced by a transgene that simultaneously drives the expression of sense and antisense *unc-22* RNA. In these transgene experiments, two independent transgenic lines were analyzed and both exhibited a penetrance of the *unc-22* twitching phenotype that was indistinguishable from that induced by the same transgenes when expressed in wild-type strains (Table III-2). These findings indicate that RDE-3 is essential for RNAi induced by feeding but is not required for RNAi induced by transgene driven expression of dsRNA. RDE-3 has one close homolog and at least 10 more distantly related family members in *C. elegans*. Therefore, it is possible that the ability of *rde-3* mutants to respond to transgene-driven RNAi reflects compensation in this pathway from one or more of these homologs. Alternatively, it is possible that RDE-3 function is required for RNAi amplification, and that RDE-3 is thus less critical when very large amounts of dsRNA are delivered by transgene-driven expression directly in target tissues (For more on these possibilities see the discussion below and the model in Figure 4B).

In *C. elegans*, antisense siRNAs accumulate in animals undergoing RNAi. Some mutants defective in RNAi, such as *rde-1* and *mut-7*, fail to accumulate siRNAs while other mutants like *mut-14* show no defect in siRNA accumulation (Tijsterman et al., 2002a). To further understand why RNAi is defective in *rde-3*, we examined the level of

Table III-1 Embryonic lethality of different rde-3 alleles

	Percentage of dead eggs (S.D.)
Wildtype (n=3)	0.5% (±0.5)
ne298 (n=7)	26% (±5.86)
ne3364 (n=7)	31.9% (±16.01)
ne3370 (n=6)	32.7% (±12.52)

Individual L4 animals were cultured at 25℃. Adults were transferred to a new plate every 12 hrs and dead eggs and total number of progeny were scored over 24 to 48 hrs

Table III-2 PCR product rescue of rde-3 (ne3364) and sensitivity of rde-3 (ne3364) to transgene driven RNAi

	F2 lines with unc animals	F2 lines sensitive to *let-2* RNAi by feeding
Control Injection	2/2	0/2
Control Injection + K04F10.6 PCR fragment	7/7	6/7

A mixture of three plasmids were injected in the control experiment: pPD125.18 [myo-3 promoter::*unc-22* sense], pPD125.29 [myo-3 promoter::*unc-22* antisense], pRF4 [*rol-6*(*su1006gf*)]. The phenotypes were scored after independent F2 inherited lines are generated

siRNA in *rde-3* mutants. As shown in Figure 2C, wild-type animals accumulate antisense *pos-1* siRNAs when exposed to RNAi targeting *pos-1*. However, siRNAs are not detected in either *ne298* or *ne3364* worms. This result suggests that RDE-3 activity is required upstream of the accumulation of siRNA during RNAi.

rde-3 is a member of a functionally diverse multi-gene family

The *rde-3* family includes several *C. elegans* members, as well as numerous homologs in other organisms. These homologs have been implicated in a variety of cellular mechanisms, including polyadenylation of transcripts in the nucleus (PAP) and cytoplasm (GLD-2) (Wang et al., 2002). The OAS members of this family are up-regulated by the interferon anti-viral response in mammalian cells (Hovanessian et al., 1977; Roberts et al., 1976). Viral dsRNA is thought to bind to OAS and to activate the synthesis of 2'-5' adenyl oligomers, which in turn activate RNase L (Dong and Silverman, 1995). RNase L then degrades cytoplasmic RNAs of both the host and virus, in a non-sequence-specific manner, preventing further viral amplification. The budding yeast protein Trf4p has been implicated in the polyadenylation and recruitment of improperly processed tRNAs to the exosome (Kadaba et al., 2004), and has also been implicated in DNA synthesis and chromosome cohesion (Wang et al., 2000b).

Like GLD-2, OAS and Trf4p, RDE-3 lacks any recognizable RNA binding motif. GLD-2 is thought to interact with target mRNAs through its association with the KH-domain RNA-binding protein GLD-3 (Wang et al., 2002). Mammalian OAS family

Figure III-3

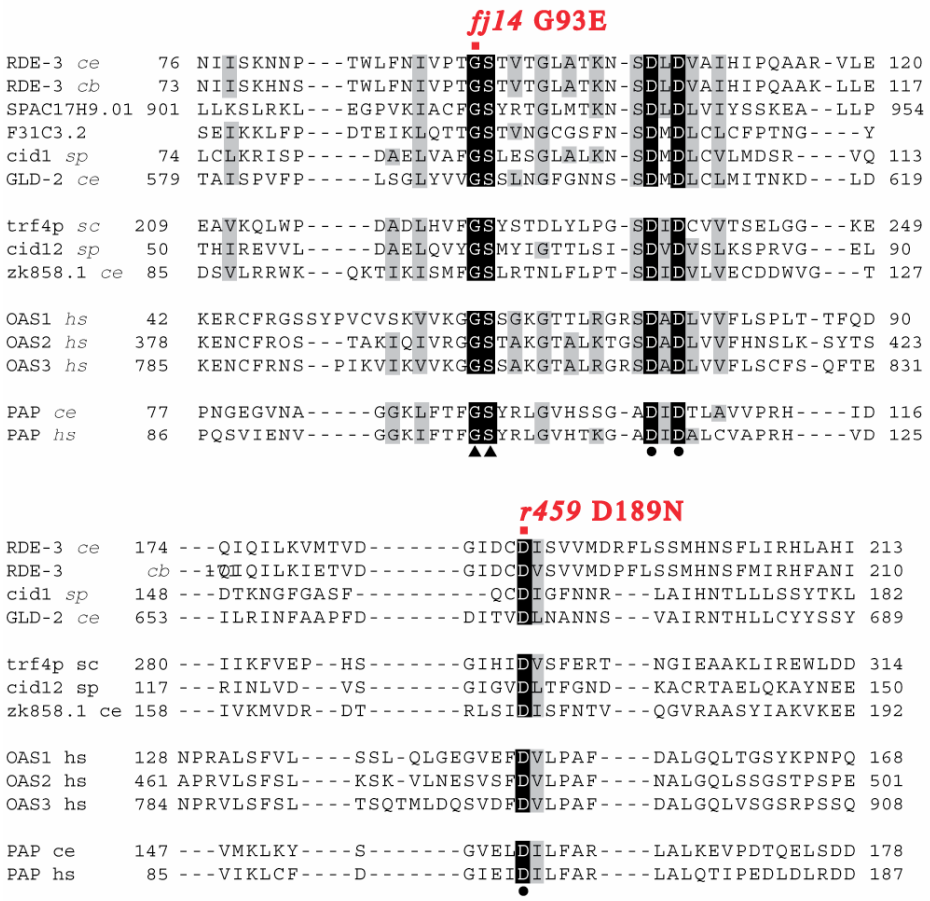

fj14 G93E
RDE-3 ce 76 NIISKNNP---TWLFNIVPTGSTVTGLATKN-SDLDVAIHIPQAAR-VLE 120
RDE-3 cb 73 NIISKHNS---TWLFNIVPTGSTVTGLATKN-SDLDVAIHIPQAAK-LLE 117
SPAC17H9.01 901 LLKSLRKL---EGPVKIACFGSYRTGLMTKN-SDLDLVIYSSKEA--LLP 954
F31C3.2 SEIKKLFP---DTEIKLQTTGSTVNGCGSFN-SDMDLCLCFPTNG----Y
cid1 sp 74 LCLKRISP-----DAELVAFGSLESGLALKN-SDMDLCVLMDSR----VQ 113
GLD-2 ce 579 TAISPVFP-----LSGLYVVGSSLNGFGNNS-SDMDLCLMITNKD---LD 619
trf4p sc 209 EAVKQLWP-----DADLHVFGSYSTDLYLPG-SDIDCVVTSELGG---KE 249
cid12 sp 50 THIREVVL-----DAELQVYGSMYIGTTLSI-SDVDVSLKSPRVG---EL 90
zk858.1 ce 85 DSVLRRWK---QKTIKISMFGSLRTNLFLPT-SDIDVLVECDDWVG---T 127
OAS1 hs 42 KERCFRGSSYPVCVSKVVKGGSSGKGTTLRGRSDADLVVFLSPLT-TFQD 90
OAS2 hs 378 KENCFROS---TAKIQIVRGGSTAKGTALKTGSDADLVVFHNSLK-SYTS 423
OAS3 hs 785 KENCFRNS--PIKVIKVVKGGSSAKGTALRGRSDADLVVFLSCFS-QFTE 831
PAP ce 77 PNGEGVNA-----GGKLFTFGSYRLGVHSSG-ADIDTLAVVPRH----ID 116
PAP hs 86 PQSVIENV-----GGKIFTFGSYRLGVHTKG-ADIDALCVAPRH----VD 125
r459 D189N
RDE-3 ce 174 ---QIQILKVMTVD-------GIDCDISVVMDRFLSSMHNSFLIRHLAHI 213
RDE-3 cb --TQIQILKIETVD-------GIDCDVSVVMDPFLSSMHNSFMIRHFANI 210
cid1 sp 148 ---DTKNGFGASF----------QCDIGFNNR---LAIHNTLLLSSYTKL 182
GLD-2 ce 653 ---ILRINFAAPFD-------DITVDLNANNS---VAIRNTHLLCYYSSY 689
trf4p sc 280 ---IIKFVEP--HS-------GIHIDVSFERT---NGIEAAKLIREWLDD 314
cid12 sp 117 ---RINLVD---VS-------GIGVDLTFGND---KACRTAELQKAYNEE 150
zk858.1 ce 158 ---IVKMVDR--DT-------RLSIDISFNTV---QGVRAASYIAKVKEE 192
OAS1 hs 128 NPRALSFVL----SSL-QLGEGVEFDVLPAF----DALGQLTGSYKPNPQ 168
OAS2 hs 461 APRVLSFSL----KSK-VLNESVSFDVLPAF----NALGQLSSGSTPSPE 501
OAS3 hs 784 NPRVLSFSL----TSQTMLDQSVDFDVLPAF----DALGQLVSGSRPSSQ 908
PAP ce 147 ---VMKLKY----S-------GVELDILFAR----LALKEVPDTQELSDD 178
PAP hs 85 ---VIKLCF----D-------GIEIDILFAR----LALQTIPEDLDLRDD 187

Figure III-3 *rde-3* mutants alter conserved residues in the nucleotydyltransferase domain.

Sequence alignment of RDE-3 with members of the polymerase β nucleotidyltransferase superfamily. Proteins are divided into subgroups based on sequence identity. Residues identical in all proteins are highlighted in black while conservative amino-acid substitutions are highlighted in grey. The conserved GS residues are marked by black triangles. The aspartic acid residues of the predicted catalytic triad are indicated with black circles. The lesions in *fj14* and *r459* are indicated. *ce* (*Caenorhabditis elegans*), *cb* (*Caenorhabditis briggsae*), *hs* (*Homo sapiens*), *sc* (*Saccharomyces cerevisiae*), *sp* (*Schizosaccharomyces pombe*). The sequences were aligned using Clustal W (Thompson et al., 1994). Some divergent sequences among subgroups were aligned manually.

members bind to dsRNA via an undefined RNA binding motif that appears to involve several positively charged residues localized on the surface of the protein (Hartmann et al., 2003).It is not presently known how Trf4p recognizes its targets. Further studies will be required to determine whether and how RDE-3 binds to RNA.

Proteins in this superfamily have two conserved features at the active site of the enzyme. A helical turn which includes a highly conserved glycine-serine motif (GS), followed by a conserved aspartic acid triad (Aravind and Koonin, 1999). Interestingly, *rde-3 (fj14)* contains a glycine to glutamate mutation in the GS motif, and *rde-3 (r459)* contains an aspartic acid to asparagine mutation in the third aspartic acid of the catalytic triad (Figure 3). These findings suggest that the polymerase activity of RDE-3 is likely to be important for its function. These residues are also conserved in a fission yeast homolog, Cid12 (Figure 3). Interestingly, cid12 has recently been shown to function in an RNAi-like mechanism required for chromatin silencing at centromeric repeats (Motamedi et al., 2004). Cid12 associates with the RNA-dependent RNA polymerase Rdp1 and is required for the accumulation of siRNAs involved in the silencing of centromeric chromatin in *S. pombe*. Despite their apparent functional similarity as factors required for siRNA accumulation, RDE-3 and Cid12 do not appear to represent orthologous proteins. The *C. elegans* gene ZK858.1 has a much greater degree of sequence identity with Cid12 (Figure 3). Conceivably, RDE-3 and Cid12/ZK858.1 provide similar functions in distinct branches of the RNAi pathway, for example in post-transcriptional vs. transcriptional silencing.

There are many possible models for how RDE-3 might function in RNAi. In one model, RDE-3 may respond to dsRNA in a manner analogous to OAS, producing oligo-adenylate that, rather than inducing a general non-specific mRNA destruction, instead stimulates nucleases required for siRNA-directed mRNA turnover. Alternatively, RDE-3 could function indirectly to facilitate RNAi by insuring a proper balance of RNA metabolism in the cell. For example, RDE-3 could function like its homolog GLD-2 in the polyadenylation of a subset of mRNAs in the cytoplasm. In the absence of RDE-3, defects in the polyadenylation of these mRNAs could lead to an accumulation of aberrant transcripts that enter the RNAi pathway and titrate limiting factors, reducing the ability of cells to initiate RNAi in response to foreign dsRNA (Figure 4A). Similarly, RDE-3 could function like Trf4p in the turnover of improperly processed tRNAs or of other RNAs normally destined for the exosome. Defects in such a mechanism could, once again (as shown in Figure 4A), lead to inappropriate recognition of these aberrant RNAs and competition for limiting components of the RNAi machinery.

A final model worth considering (Figure 4B) is that RDE-3 is required for the amplification of RNAi in response to low levels of the inducing trigger dsRNA. This model is consistent with our own observation that RDE-3 is not required for RNAi initiated by transgene driven dsRNA (see above), and with the finding from *S. pombe* that the RDE-3 homolog Cid12 interacts with RdRP. In *C. elegans*, the detectable accumulation of siRNAs during RNAi requires RdRP activity (Conte and Mello unpublished; (Sijen et al., 2001)), and is thought to involve an RdRP-dependent synthesis of new dsRNA after an initial round of target mRNA recognition. After ingestion, small

Figure III-4

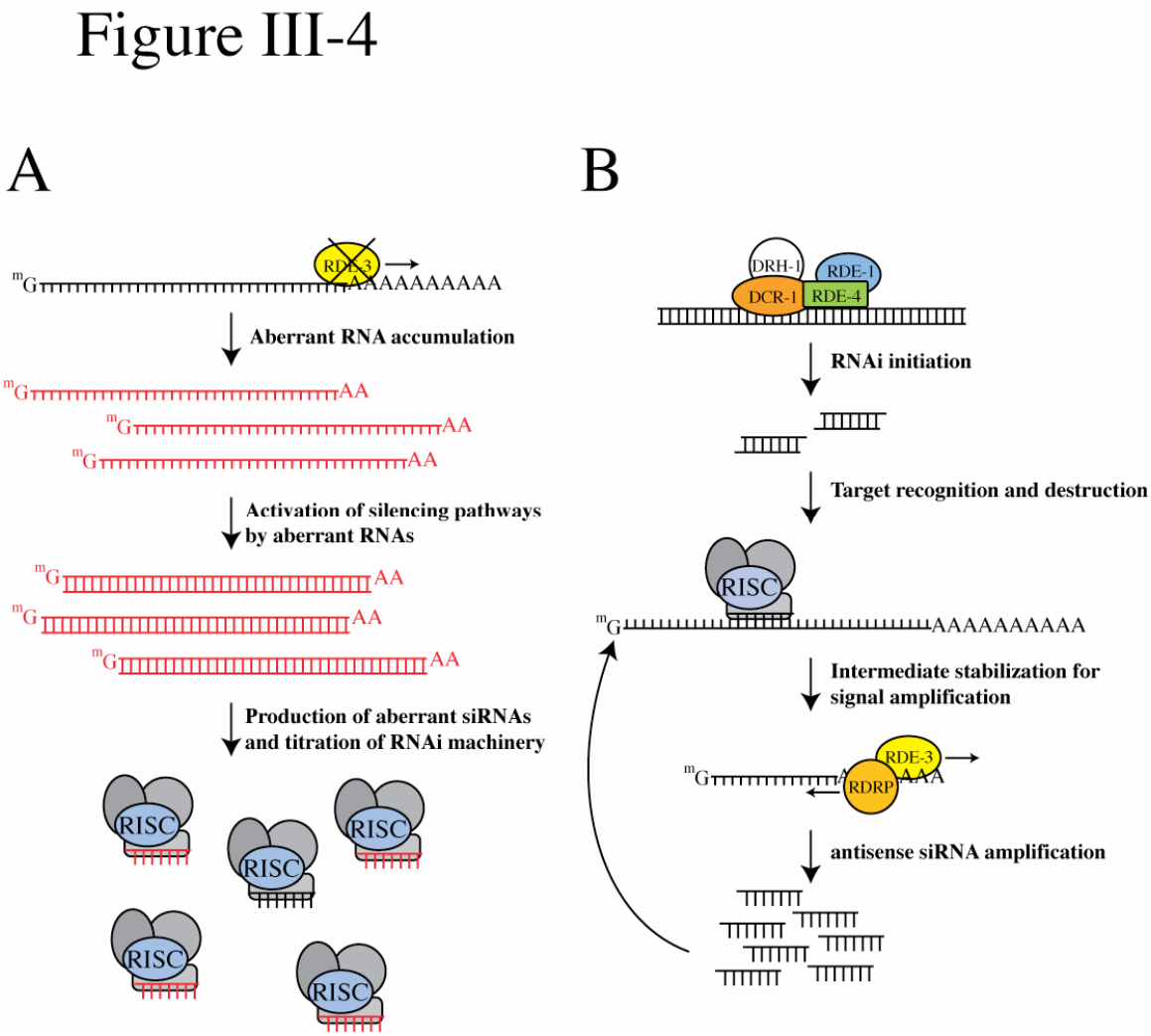

A
RDE-3
AAAAAAAAAA
Aberrant RNA accumulation
AA
Activation of silencing pathways
by aberrant RNAs
Production of aberrant siRNAs
and titration of RNAi machinery
RISC
B
DRH-1
RDE-1
DCR-1
RDE-4
RNAi initiation
Target recognition and destruction
RISC
AAAAAAAAAA
Intermediate stabilization for
signal amplification
RDE-3
RDRP
AAA
antisense siRNA amplification

Figure III-4 Models.

Schematic diagrams illustrating possible roles for RDE-3: (A) as a factor that maintains the balance of RNA metabolism required for efficient RNAi or (B) directly, as a direct component of the RNAi pathway. Model (A) depicts RDE-3 as a poly-A polymerase required for processing a subset of mRNAs in the cell. Loss of RDE-3, indicated by the X could lead to accumulation of 'aberrant' transcripts (red) with short poly-A tails, that can enter the RNAi pathway, titrating RISC, (or other limiting factors) necessary for efficient RNAi. In model (B) proteins that initiate RNAi, including Dicer (DCR-1) which processes long dsRNA into siRNAs, a Dicer-related helicase (DRH-1), an argonaute protein (RDE-1) and the dsRNA binding protein (RDE-4) (Tabara et al., 2002) are shown processing foreign dsRNA into primary siRNA. These extremely low abundance primary siRNAs then target mRNA in an initial round of cleavage mediated by components of the RNA induced silencing complex (RISC). RDE-3 is then proposed to polyadenylate the 5' cleavage product, stabilizing this intermediate and allowing RdRP to amplify the response, generating abundant secondary siRNAs.

quantities of dsRNA entering target tissues may be processed by Dicer to generate low abundance siRNAs. These primary siRNA could then direct a first round of target mRNA cleavage. RDE-3 might then polyadenylate the nascent 3'end of this cleavage product, stabilizing it, and permitting its recognition by RdRP (Figure 4B). In the absence of RDE-3, the initial cleavage product may be too unstable to be recognized efficiently by RdRP, preventing amplification and the consequent accumulation of siRNA. While their precise functions remain to be discovered, the findings that; 1) RDE-3 and Cid12 are required for RNAi-related mechanisms and, 2) that mutations in these factors result in similar defects in siRNA accumulation in these divergent species, strongly suggests that members of this protein family are likely to represent integral components of small-RNA-mediated silencing pathways.

ACKNOWLEDGEMENTS

We would like to thank Danesh Moazed for sharing data prior to publication. Some nematode strains were provided by *Caenorhabditis* Genetics Center, funded by National Institutes of Health National Center for Research Resources. This work has been funded in part by National Institutes of Health (GM58800). MJS is a Canadian Institutes of Health Research postdoctoral fellow; and CCM is a Howard Hughes Medical Institute investigator.

Chapter IV

Characterization of the *rde-8* gene

Chun-Chieh G. Chen, Pedro Batista, Jennifer McCullough, Lina Li, Irina Ostrovisky, and Craig Mello[1]

Program in Molecular Medicine and [1]Howard Hughes Medical Institute, University of Massachusetts Medical School, Worcester, Massachusetts 01605

Chun-Chieh G. Chen performed the analysis and experiments in Figure IV-1, IV-2, IV-3, IV-4, IV-5, IV-6 B, and Figure IV-7. Pedro Batista performed the 2'-*O*-Methyl experiments shown in Figure IV-6 A and C.

Summary

Multiple distinct small RNA-mediate gene silencing pathways regulate fundamental biological processes in a cell. The molecular mechanism that underlies these pathways is known as RNA interference or RNAi. In the RNAi pathway, double-stranded RNAs are processed into small RNAs. The small RNAs, in conjunction with a variety of effector protein complexes, guide these effector complexes to execute gene silencing via mRNA degradation, translational inhibition, or transcriptional silencing. To identify protein components that are required for the biogenesis and the function of these small RNAs, we conducted genetic screens in search of mutants that are *R*NAi *de*fficient (*rde*) in the nematode *C. elegans*. Here we report the cloning and characterization of the *rde-8* gene. *rde-8* encodes a protein with a conserved novel domain, the RDE-8 domain, and a putative nuclear localization sequence. We demonstrate that the *rde-8* gene functions in the downstream effector step of the RNAi pathway and is required for the accumulation of several distinct classes of the small RNAs in the worms. These includes 1) the exogenous siRNAs, 2), the endogenous siRNAs, and 3) the X-cluster RNAs. Our report identifies a novel player in the RNAi pathway and suggests the requirement of a new conserved domain in the RNAi pathway.

Introduction

Small RNAs of approximately 21 to 30nt in length have recently been shown to play roles in a variety of important gene regulatory mechanisms (Zamore and Haley, 2005). Most of these small RNAs are originated from precursors that form double-stranded RNA (dsRNAs) structures and are processed by enzymes of the RNA Endonuclease III-like Dicer protein family into small RNAs (Meister and Tuschl, 2004). These small RNAs are subsequently incorporated into a number of different effector complexes and direct sequence-specific search to find the complementary target. In the heart of the effector complexes are the Argonaute proteins, which belong to a conserved gene family essential for small RNA mediated gene silencing. An Argonaute protein is comprised of two major domains: The PAZ domain interacts with small RNAs and PIWI is a domain with RNase H-like motifs (Lingel et al., 2003; Song et al., 2003; Song et al., 2004). Guided by the small RNAs, the Argonaute-containing effector complexes can execute gene silencing by a various means including RNA degradation, translational inhibition and chromatin state change (Tomari and Zamore, 2005).

In the nematode *C. elegans*, introduction of dsRNAs induced a potent and speficic gene silencing phenomenon, termed RNAi interference or RNAi (Fire et al., 1998). The dsRNAs are transported into an animal via a mechanism termed systemic RNAi (Winston et al., 2002). Once in the cells, dsRNAs are processed into small RNAs by the initiation complex (Tabara et al., 2002). This protein complex contains RDE-4, a double-stranded RNA biding protein, RDE-1, an Argonaute protein, DCR-1, a Dicer family dsRNA endonuclease, and DRH-1 and DRH-2, two Dicer related helicases. The processed small

RNAs are named small interfering RNAs (siRNAs) and, are hypothesized to direct the RDE-1 protein to scan for complementary mRNA targets. After a target has been found, an amplification signal is activated for the production of antisense secondary siRNAs generated by the RNA dependent RNA Polymerase (RdRPs) (Sijen et al., 2001). These siRNAs are then escorted by a set of additional Argonaute proteins to elicit mRNA degradation (Yigit et al., 2006).

In addition to exogenously introduced small RNAs (exo-siRNAs), the *C. elegans* genome produce a variety of endogenously produced small RNAs for the regulation of gene expression. These include miRNAs (microRNAs), tncRNAs (tiny non coding RNAs), endo-siRNAs (endogenous siRNAs) and X-cluster RNAs (X chromosome cluster RNAs) (Ambros et al., 2003; Lau et al., 2001; Lee and Ambros, 2001). While microRNAs mediate gene silencing through translational inhibition, the function of these small RNAs remains largely unknown (Olsen and Ambros, 1999). One exception is a subset of the endo-siRNAs, whose sequences correspond to that of the transposons. Studies have shown that these siRNAs silence the transcripts originated from the tranposon sequences post-transcriptionally, suggesting that RNAi mediates transposon silencing in *C. elegans* (Sijen and Plasterk, 2003). Consistent with this, genetic screens in search of RNAi deficient mutants have identified several mutants that also exhibit defects in transposon silencing (Ketting et al., 1999).

In order to investigate the mechanism of RNAi, we have conducted genetic screens to identify mutants that are *R*NAi *de*ficient (*rde*). Here we report the identificaiton and characterization of *rde-8*. *rde-8* encodes a protein with a conserved

novel domain, the RDE-8 domain, and a putative nuclear localization sequence. We demonstrate that the function of the *rde-8* gene is required downstream of *rde-1*. In addition, *rde-8* is required for the production of secondary siRNAs as well as endogenous small RNAs such as endo-siRNAs and X-cluster RNAs. Our result indicates that *rde-8* functions in the generation of a variety of small RNAs in *C. elegans*.

Material and Methods

C. elegans Culture and strains

C. elegans culture and genetics were as described in Brenner (1974) (Brenner, 1974). The Bristol strain N2 was used as the standard wildtype strain. Alleles, marker mutations, and balancer chromosomes used in this study are listed below. I: C30F12.1 (*tm2615*). II: C29F5.3 (*tm2499*). IV: *unc-5* (*e53*), *rde-8* (*ne3360*, *ne3361*, *ne3309*), *dpy-20* (*e1282*).

C. elegans strains generated in this study are listed as follows: WM148 *ne3361*; *ne*Is14[*rde-8*::GFP:RDE-8, *rol-6*]. WM149 *ne3361*; *ne*Ex13[*rde-8*::GFP:RDE-8DD, *rol-6*]. WM150 *ne3361*; *ne*Ex14[*rde-8*::GFP:RDE-834D, *rol-6*]. WM151 *nee3361*; *ne*Is10.

Mapping of *rde-8*

ne3360 was used for the mapping of *rde-8*. After determining that *rde-8* is located on the right arm of LGIV, a genetic triple of *unc-5 (e53) rde-8 (ne3360) dpy-20 (e1282)* was created to perform three factor mapping and single nucleotide polymorphism

mapping using a Hawaiian isolate of *C. elegans* (Figure IV-1). The results narrowed the genetic interval of *rde-8* down to a region of 6 cosmids. Rescue experiments were then performed for further mapping.

Cosmid rescue

6 cosmids were determined to contain candidates of the *rde-8* gene (Figure IV-1). Cosmid rescue experiments were first performed with a mix of two cosmids each. A DNA mix of 50ng/μL of each cosmid and 100ng/μL of *rol-6* gene as injection marker were used to generate transgenic animals by microinjection of the DNA mix into the germline of the worms. The transgenic worms were then tested for the RNAi resistance by feeding the animals with bacteria expressing *pos-1*, *let-2* or *unc-22* dsRNAs. Transgenic animals containing the mix of two cosmids, ZC477.5 and C34D4 rescued the RNAi defect of the *rde-8* mutant. This procedure was then repeated using a single cosmid of which ZC477.5 rescued the RNAi defect again. Candidate gene approach was then used to select genes on ZC477.5 for PCR amplification. PCR fragment containing a single gene, ZC477.5 rescues the RNAi deficiency of the *rde-8* mutant.

Sequencing of the *rde-8* locus

Two DNA fragments containing the *rde-8* gene from each allele of the *rde-8* mutants were amplified using primers CMo 9781 CCA TTG CTC GTT GTG AAG TAT CCG and CMo 9790 GCA GTG TGA CAA AGG CGC AC. Sequencing of these DNA fragments were then performed using the following primers: CMo 9781 CCA TTG CTC

GTT GTG AAG TAT CCG, CMo 9782 GAA AAT TAC GCA ACG ACG AGG, CMo 9783 GGA AAG TCG ACA CCA ATG ACA G, CMo 9784 CTT TTC AGG TTC GGA TAG CAC, CMo 9785 AAT GCA AAA ATA CCA GAG CTC C, CMo 9786 CGA TAA GCC TAT CAA TTG CAG, CMo 9787 GAA TAG CAA ATC ATG TAT GCA C, CMo 9788 GAT CCT TTC CTT CTG GGC AC, CMo 9789 AAC ACT TGA GAA GTT TTG CAG G, and CMo 9790 GCA GTG TGA CAA AGG CGC AC.

RNA preparation and Northern blotting

TriReagent were used to purify RNAs from the worms. A pellet of about 50,000 to 100,000 adult worms of *C. elegans* was treated following the protocol of the TriReagent for RNA extraction. DNase I treatment were performed to further digest DNAs away from the extracted RNA solutions (Ambion). After DNase I treatment, RNAs were extracted with phenol/chloroform/isoamyl alcohol and recovered by Ethanol precipitation (Molecular Cloning). These RNAs are further fractionated following the protocol of MirVana purification (Ambion) to enrich for fractions that contain small RNAs of around 200bp or smaller. 20μg of RNAs from this preparation were used for gel electrophoresis using a 15% UREA polyacrylamide gel (Sequagel, National Diagnostics). Northern blotting was performed as described in Duchaine *et al*. (Duchaine et al., 2006).

Preparation of Protein lysates, 2'-*O*-Methyl binding experiments and Immuno-precipitation

Preparation of protein lysates and immuno-precipitation were performed as described in Duchaine *et al*. (Duchaine et al., 2006). 2´-*O*-Methyl binding experiments were performed as described in Yigit *et al*. (Yigit et al., 2006).

Results

Cloning of *rde-8*

Three independently isolated alleles of *rde-8* were isolated from a genetic screen in search of *rde* mutants. In order to identify the molecular lesion of *rde-8*, we set out to clone the corresponding gene. We mapped *rde-8* (*ne3360*) to a genetic interval containing six overlapping cosmids (Figure IV-1). Transgenic animals carrying different mixture of cosmid DNAs or single-gene-containing PCR products were then generated to test the animals for the ability to restore the sensitivity to RNAi. In all the combination tested, DNA containing gene ZC477.5 always rescues the RNAi defect of *rde-8*.

Sequencing of ZC477.5 genomic region revealed a G to A substitution in the position 603 of the coding sequence in all three alleles, *ne3360, ne3361* and *ne3309*. This mutation results in the introduction of an early stop codon, which appears to be null mutations (Figure IV-1). Consistent with this, placement of *rde-8* (*ne3360*) over a chromosomal deficiency resulted in worms deficient in RNAi, suggesting that this lesion is most likely a null mutation. In summary, these results indicate that *rde-8* is encoded by ZC477.5 and the *rde-8* mutant is homozygous viable and defective in RNAi.

The RDE-8 protein encodes a 339 amino-acid protein. Sequence analysis revealed a conserved region of 44 amino acids located in the center of the protein. Because this

domain is first identified in the RDE-8 protein, we named it the RDE-8 domain (Figure IV-2). In addition, by using the PSORTII computer program, a putative nuclear localization sequence (NLS) was predicted near the carboxyl terminus, (Figure IV-2). A homolog of *rde-8*, CBP15480, is found in another nematode specie, *C. briggsae* (figure IV-2). In addition, sequence alignment analysis using the RDE-8 domain identified two additional predicted coding sequences, C30F12.1 and C29F5.3, in the *C. elegans* genome. Both the RDE-8 domain and the predicted NLS sequence are also identified in these proteins, suggesting that these two domains are evolutionary conserved.

The *rde-8* domain is required for RNAi

The RDE-8 domain is a small domain of 44 amino acids in length. It contains four highly conserved aspartic acid residues and an alanine residue (Figure IV-2). To investigate if the RDE-8 domain is required for RNAi, we mutated the conserved aspartic acid residues to alanine residues and ask if these mutant proteins can rescue the RNAi defect in the rde-8 mutant. We generated two constructs in which the first two, and the last two conserved glutamate residues were mutated, respectively. We then expressed these mutant proteins by a transgene in *rde-8* (*ne3360*) and tested for the RNAi sensitivity in the *rde-8* mutant by feeding. To exclude the possibility that the mutations generated render the proteins unstable, a green fluorescent (GFP) protein was tagged to the *rde-8* proteins to monitor for the proper expression. As shown in figure IV-3, a wildtype GFP::RDE-8 fusion protein rescued the Rde phenotype of the *rde-8* mutant. In contrast, both of these *rde-8*

mutant constructs fail to rescue the RNAi defect in *rde-8* mutants, suggesting that the RDE-8 domain is important for its function in RNAi.

To test whether these mutant constructs cause a dominant negative effect for RNAi, we have also expressed this protein in the wildtype background (Figure IV-3). These transgenic animals exhibit normal sensitivity to RNAi, suggesting that these mutants do not interfere with the normal RNAi pathway. This result also indicates that the failure of rescuing the RNAi defect in the *rde-8* (*ne3360*) animals is not caused by a dominant negative effect.

The RDE-8 protein is partially localized to the nucleus

Because the RDE-8 protein contains a putative nuclear localization sequence, we generated a transgenic animal expression GFP::RDE-8 fusion protein under its own promoter to monitor the subcellular localization of the RDE-8 proteins. Examination of the localization of the functional GFP::RDE-8 fusion protein demonstrated that fluorescent signals of a fraction of the GFP::RDE-8 protein is found in the nucleus, indicating that a fraction of the RDE-8 protein are present in the nucleus (Figure IV-4). Exposure of the worms to dsRNAs does not change the localization of the GFP::RDE-8 protein suggesting that this localization pattern is not affected by triggering RNAi.

The RNAi defect of *rde-8*

To characterize the RNAi deficiency of *rde-8*, we performed several RNAi assays to test the resistance of RNAi introduced via three different approaches for dsRNA

Figure IV-1

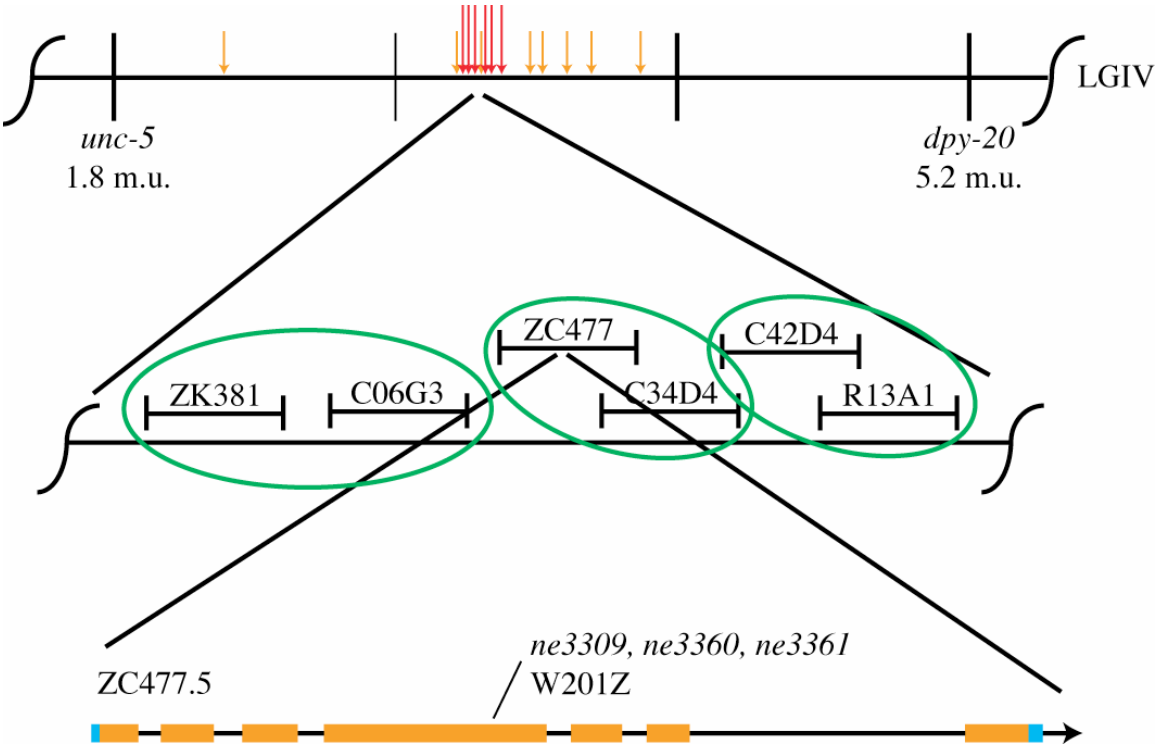

LGIV
unc-5
1.8 m.u.
dpy-20
5.2 m.u.
ZC477
C42D4
ZK381
C06G3
C34D4
R13A1
ne3309, ne3360, ne3361
W201Z
ZC477.5

Figure IV-1 Cloning and Mapping of the rde-8 locus

For detail, please refer to material and methods. The red and orange arrows indicate map positions of single nucleotide polymorphisms used in the mapping. The Green circles represent DNA mixes of cosmids used for rescue experiments. The ZC477.5 locus is shown with exons represented by orange boxes, intron represented by black lines and the UTRs represented by blue boxes.

delivery (Tabara et al., 1999; Timmons et al., 2001). The RNAi response can be triggered by feeding the animal with bacteria expressing dsRNAs, by injecting the animals with dsRNAs or by introducing trasgenes expressing dsRNAs in to an animal. These three methods vary in the amount of dsRNAs that can be efficiently delivered, with feeding being the least efficient form, and transgene-driven RNAi the most efficient. Therefore, a mutant that is fully deficient in RNAi will exhibit strongly resistance to RNAi delivered by any of these methods. However, the mutants that are required for efficient RNAi could only be resistant to RNAi by feeding and injection but sensitive to transgene driven RNAi. In addition, because a transgene can be expressed directly in the cells where an mRNA target is present, the transgene driven RNAi does not require a functional systemic RNAi pathway. Therefore, by using this assay, we can also investigate whether rde-8 is involved in the systemic RNAi. The requirements of the *rde-8* gene in the RNAi pathway were assayed for either germline or somatic RNAi, because these two pathways utilize different genes to elicit silencing.

While the wildtype animals are sensitive to RNAi delivered by feeding, injection and transgene, our results demonstrate that *rde-8* is strongly resistant to both germline and somatic RNAi by feeding (Figure IV-3). In addition, *rde-8* is also resistant to somatic RNAi by injection of double stranded RNAs targeting the *unc-22* gene, which is a gene required for muscle development of *C. elegans* (Figure IV-5). In contrast, the *rde-8* mutant exhibits some degree of sensitivity to germline RNAi when double-stranded RNAs targeting the *src-1* gene were administered. Furthermore, *rde-8* is also sensitive to transgene-triggered RNAi of the *unc-22* gene (6/6 transgenic lines). Taken together, these

Figure IV-2

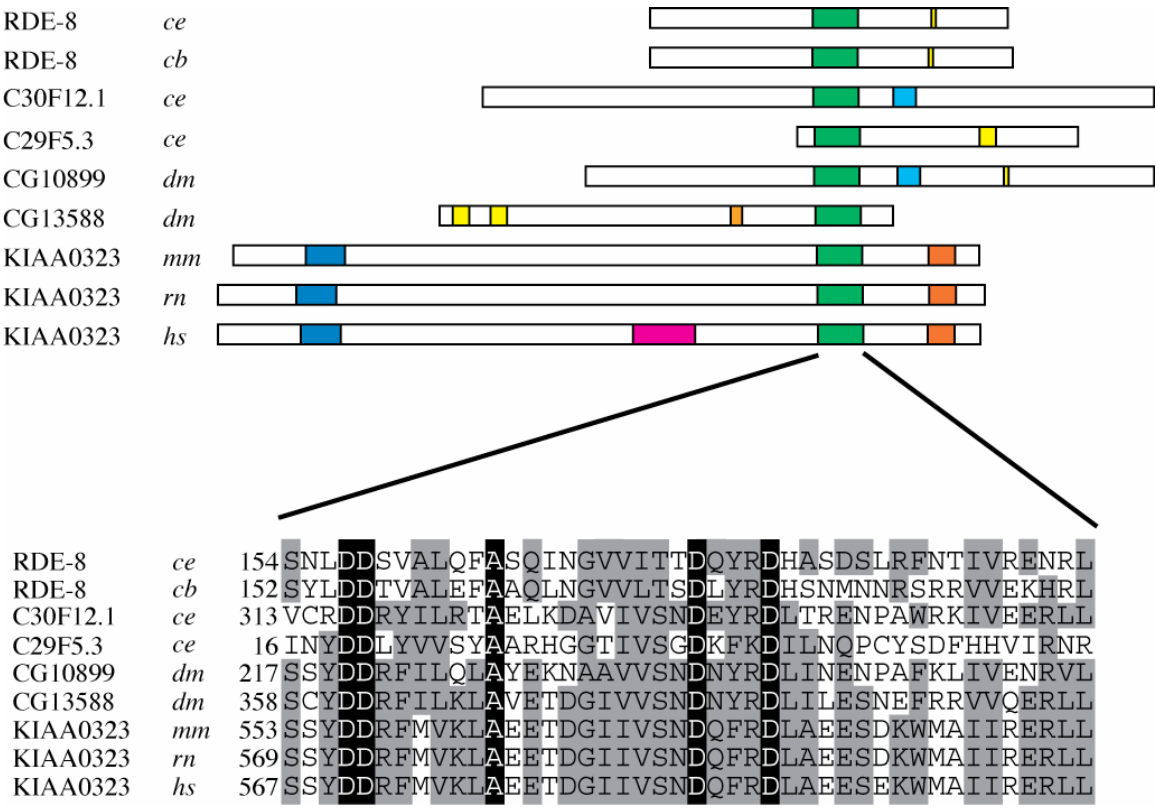

RDE-8 ce
RDE-8 cb
C30F12.1 ce
C29F5.3 ce
CG10899 dm
CG13588 dm
KIAA0323 mm
KIAA0323 rn
KIAA0323 hs
RDE-8 ce 154 SNLDDSVALQFASQINGVVITTDQYRDHASDSLRFNTIVRENRL
RDE-8 cb 152 SYLDDTVALEFAAQLNGVVLTSDLYRDHSNMNNRSRRVVEKHRL
C30F12.1 ce 313 VCRDDRYILRTAELKDAVIVSNDEYRDLTRENPAWRKIVEERLL
C29F5.3 ce 16 INYDDLYVVSYAARHGGTIVSGDKFKDILNQPCYSDFHHVIRNR
CG10899 dm 217 SSYDDRFILQLAYEKNAAVVSNDNYRDLINENPAFKLIVENRVL
CG13588 dm 358 SCYDDRFILKLAVETDGIVVSNDNYRDLILESNEFRRVVQERLL
KIAA0323 mm 553 SSYDDRFMVKLAEETDGIIVSNDQFRDLAEESDKWMAIIRERLL
KIAA0323 rn 569 SSYDDRFMVKLAEETDGIIVSNDQFRDLAEESDKWMAIIRERLL
KIAA0323 hs 567 SSYDDRFMVKLAEETDGIIVSNDQFRDLAEESEKWMAIIRERLL

Figure IV-2 The RDE-8 protein, the RDE-8 domain and homologs containing the RDE-8 domain.

Proteins containing the RDE-8 domain are shown by aligning the RDE-8 domain region. Genes containing this domain can be found in *ce* (*C. elegans*), *cb* (*C. briggsae*), *dm* (*D. melanogaster*), *mm* (mouse), *rn* (rat) and *hs* (human). The sequence alignment of these proteins is further presented. The conserved residues are highlighted in black and the amino acids of similar property are highlighted in grey. The color boxes indicate the following: green box-the RDE-8 homology domain; yellow box-the nuclear localization domain; blue box, the KH domain; orange box-transposase 9 domain; red box-DNA resolvase.

Figure IV-3

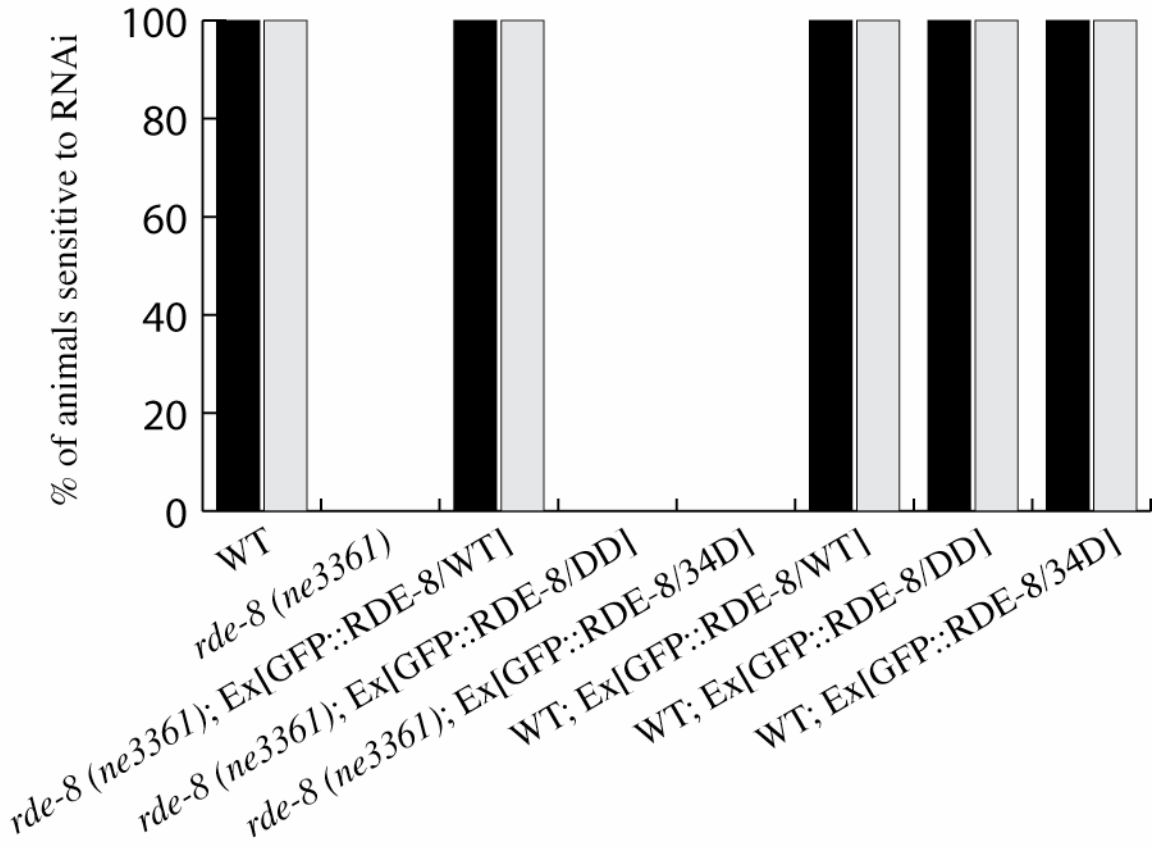

% of animals sensitive to RNAi
100
80
60
40
20
0
WT
rde-8 (ne3361)
rde-8 (ne3361); Ex[GFP::RDE-8/WT]
rde-8 (ne3361); Ex[GFP::RDE-8/DD]
rde-8 (ne3361); Ex[GFP::RDE-8/34D]
WT; Ex[GFP::RDE-8/WT]
WT; Ex[GFP::RDE-8/DD]
WT; Ex[GFP::RDE-8/34D]

Figure IV-3 The RDE-8 domain is required for the function of the *rde-8* gene in RNAi

GFP tagged wildtype or mutant RDE-8 proteins are expressed by transgenes in the *rde-8* mutants or wildtype worms. The transgenic worms (n=20) are then tested for their resistance to germline RNAi (Black columns) or somatic RNAi (Grey columns). Transgenes with mutations in the first two conserved aspartic residues (DD) or last two conserved aspartic acid residues (34D) of the RDE-8 domain does not rescue the RNAi defect of the rde-8 mutant. Expression of these mutant proteins in the wildtype worms does not cause any RNAi defect.

Figure IV-4

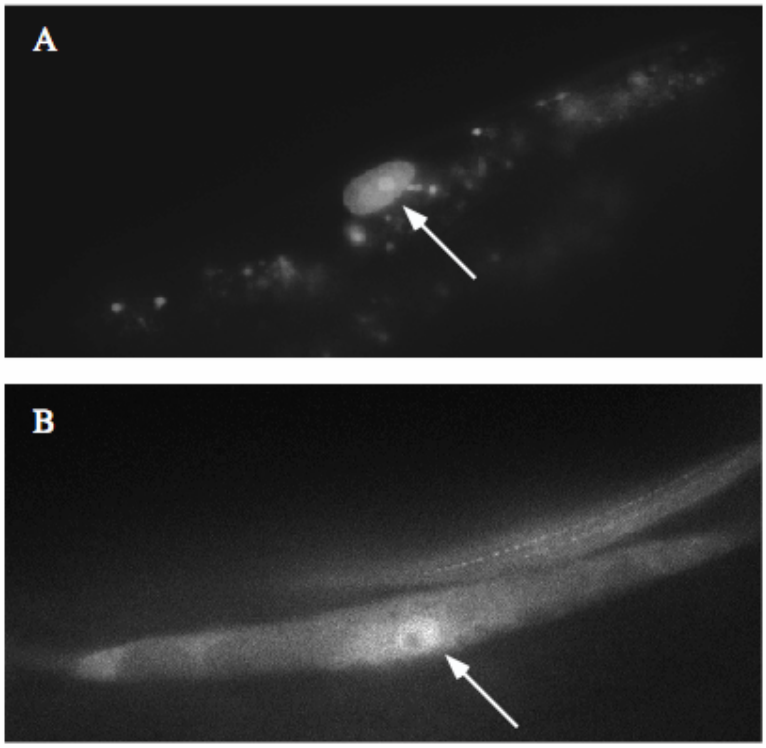


Figure IV-4 The nuclear localization of the GFP::RDE-8 protein

A) The GFP::RDE-8 expression is present in the coelomocyte of the transgenic worms. The GFP::RDE-8 is localized to both cytoplasm and the nucleus (indicated by the white arrow). The specles in the background are resulted from the autofluorescence of the gut granules.

B) Similar pattern of the subcellular localization of the GFP::RDE-8 protein can be observed in the muscle cells. The absence of the fusion protein from the nucleolus can also be observed.

results suggest that *rde-8* is involved in systemic RNAi or that another functionally redundant pathway collaborates with *rde-8* to exert RNAi.

rde-8 functions downstream of RDE-1/siRNA complex formation

To further investigate in which step of RNAi *rde-8* is involved, we test if the trigger dsRNA can be transported into the cells by the 2'-*O*-Methyl oligo binding assay. This assay allows the monitoring of the interaction between RDE-1 and the primary siRNAs of both polarities. The primary siRNAs are siRNAs generated by DCR-1 directly processing the trigger dsRNAs that are transported into the cells via systemic RNAi. Therefore, if rde-8 is required for systemic RNAi, dsRNAs will not be transported into the cells, and the interaction between RDE-1 and siRNA will not be detected. On the other hand, if RDE-8 functions downstream of the formation of the siRNA/RDE-1 complex, the interaction between siRNAs and RDE-1 will not be affected in the *rde-8* mutant. As described in Yigit *et al*., single stranded 2'-*O*-Methyl oligos with sequences complementary to either strand of the trigger dsRNAs can interact with primary siRNAs (Yigit et al., 2006). When the oligos are bound to a matrix, the siRNA/RDE-1 complex can be recovered from protein lysates prepared from the adult animals of *C. elegans*. The presence of RDE-1 which interacts with the primary siRNAs will then be resolved by Western blotting.

Synchronized larval stage 1 (LI) animals were fed on bacteria with or without expression of GFP dsRNAs around 60hrs until reaching adulthood. The worms are then

Figure IV-5

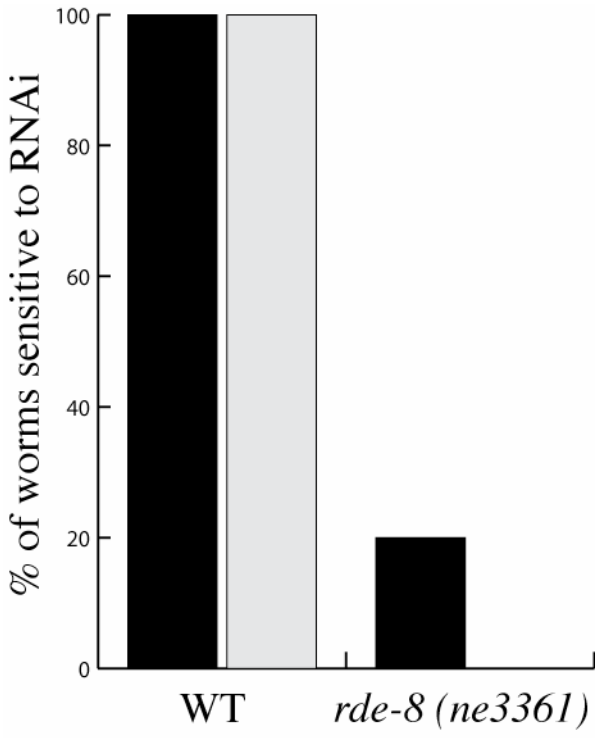


Figure IV-5 The RNAi deficiency of the *rde-8* mutant

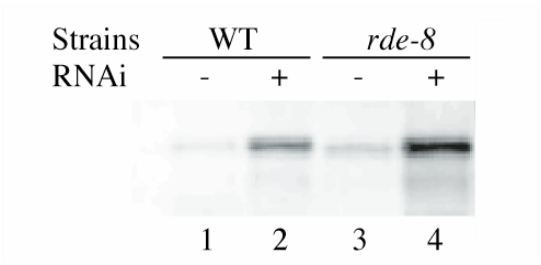
Both wildtype and *rde-8* mutants are injected with 1mg/mL of double stranded RNAs targeting *src-1* (germline RNAi–black columns) or *unc-22* (somatic RNAi–grey columns). The phenotypes caused by the dsRNAs were scored in the progeny. Wildtype worms show 100% sensitivity to *src-1* RNAi (n=959) and *unc-22* RNAi (n=765). The rde-8 mutant is completely resistant to somatic RNAi (n=563) but exhibit some level of sensitivity to the germline RNAi (n=699).

harvested for the preparation of protein lysates. The lysates are incubated with 2'-*O*-Methyl oligos bound to a matrix that allows for recovery of the oligo bound precipitates. The recovered precipitates are then subjected to Western blotting using RDE-1 antibody. As shown in Figure IV-5a, in the lysate prepared from the wildtype animals, the RDE-1 protein is associated with the primary siRNA and can be recovered by the oligo matrix targeting antisense primary siRNAs (Lane 2). This interaction between RDE-1 and the siRNAs is not affected in the lysates prepared from *rde-8* mutants, suggesting that *rde-8* is not required for the initial transport of dsRNAs into the cells (Lane 4).

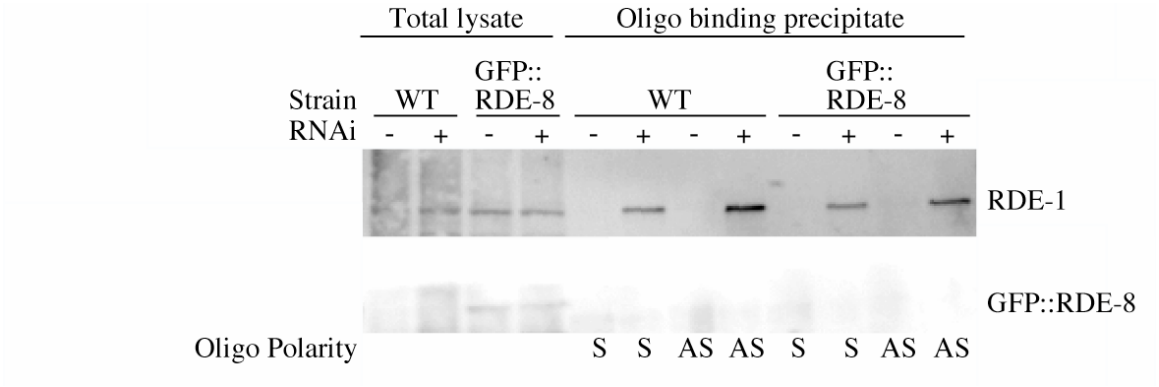
We then ask if the RDE-8 protein interacts with the siRNA/RDE-1 complex and functions at the initiation step of RNAi. We performed immunoprecipitation experiments using anti-GFP antibody to precipitate a GFP::RDE-8-fusion protein that can fully rescue the RNAi defect in *rde-8* mutants (Figure IV-5b). In addition, we also test the ability of GFP::RDE-8 to interact with the 2'-*O*-Methyl oligos (Figure IV-5c). As presented in Figure IV-5b, although both RDE-1 protein and GFP::RDE-8 protein are present in the protein lysate, immunoprecipation of GFP::RDE-8 does not pull down the RDE-1 protein. Furthermore, the presence of the GFP::RDE-8 protein is not detected in 2'-*O*-Methyl oligos targeting either sense or antisense of the primary siRNAs. However, the interaction between the RDE-1 protein and the primary siRNAs of both orientations can be clearly demonstrated in this experiment. Taken together, these experiments show that the *rde-8* gene is not required for the RDE-1/siRNA complex formation. In addition, the GFP::RDE-8 protein was not found to interact the RDE-1/siRNA complex, suggesting that the *rde-8* gene functions downstream of the RDE-1/siRNA complex formation.

Figure IV-6

A.



B.



C.

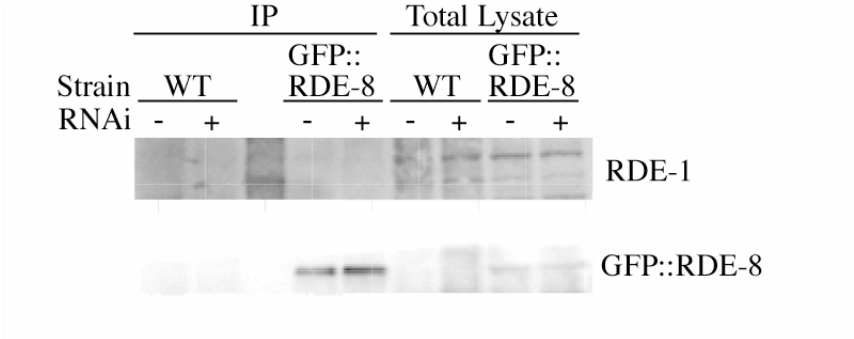


Figure IV-6 The *rde-8* gene acts downstream of the RDE-1/siRNA complex formation

A) 2'-*O*-Methyl binding experiments were performed with lysates prepared from the wildtype and *rde-8* mutant animals. The RDE-1 protein can by pulled down by the oligo matrix targeting the antisense primary siRNAs.

B) immunoprecipitation of the GFP::RDE-8 proteins. Both RDE-1 protein and GFP::RDE-8 protein are expressed in the lysates. The GFP::RDE-8 prteins are precipitated with the anti-GFP antibody. However, the RDE-1 protein cannot be detected after immunoprecipitation.

C) 2'-*O*-Methyl binding experiment using the lysates prepared from the wildtype and GFP::RDE-8 proteins. The RDE-1 protein interacts with the oligo matrix when the lysates are prepared from worms exposed to dsRNAs using oligos of both sense and antisense orientations. The presence of the GFP::RDE-8 protein can not be detected.

rde-8 is required for the accumulation of both exo- and endo-siRNAs

Because the RNAi signal can be initiated in *rde-8*, we next examine whether the secondary antisense siRNAs were produced in the *rde-8* mutant by Northern blotting. In *C. elegans*, siRNAs can be classified into two classes, depending on the origin of the dsRNAs: the exo-siRNAs are produced from dsRNAs that are provided exogenously, whereas the endo-siRNAs are produced from the genome of the worms by a yet-to-be characterized mechanism. We again exposed worms to dsRNAs as described in the previous section. These worms were then harvested and subjected to RNA purification. The total RNAs were then size-selected to enrich for the fractions that contain small RNAs. Northern blotting was performed to detect several class of small RNAs, including 1) the exo-siRNAs targeting the *sel-1* gene, 2) the endo-siRNAs targeting the coding sequence of K02, 3) the X-cluster RNAs and 4) the microRNA *let-7*.

When animals are exposed to bacteria expressing *sel-1* dsRNAs, the *sel-1* antisense siRNAs accumulate in the wildtype animals (Figure IV-7, Panel 1). However, the level of the *sel-1* siRNAs is reduced to 33% in the *rde-8* mutants. The reduction of the *sel-1* siRNAs can be rescued to 70% by the expression of the RDE-8 protein in the *rde-8* mutant, suggesting that *rde-8* is required for the accumulation of antisense secondary exo-siRNAs.

Next, the presence of the endogenously expressed small RNAs originated from the K02E2.6 mRNA and the X clusters is also examined. As shown in figure IV-6 Panel 2 and 3, both classes of small RNAs are present in the wildtype worms but neither of the endo-siRNAs is detected in the *rde-8* mutant (0% of signal and 3% of signal above

background for X cluster and K02E2.6 small RNAs respectively). In addition, the level of these siRNAs can be partially restored by 21% in the *rde-8* rescued animals, suggesting that *rde-8* is also involved in the production or stability of the endo-siRNAs. However, the expression of the let-7 microRNA is not affected in the *rde-8* mutant. We notice that the defect of the accumulation of the endo-siRNAs and the exo-siRNAs is only partially rescued. This is most likely due to the partial rescue of this strain in the somatic tissue but not in the germline, because the expression of transgenes is often silenced in the germline of *C. elegans*. To sum up, our results indicate that the function of *rde-8* is required for the accumulation of both endo-siRNAs and secondary exo-siRNAs but not the microRNAs.

It has been reported that a group of Argonaute proteins, named the secondary Argonatues, interact with the secondary siRNAs (Yigit et al., 2006). This interaction is essential for the accumulation of secondary siRNAs because in a Multiple Argonaute mutant (MAGO), where the secondary Argonautes genes are mutated, secondary siRNAs no longer accumulate. In addition, overexpression of one of the secondary Argonaute proteins leads to a higher level of accumulation of the secondary siRNAs, suggesting that the interaction between the Argonaute proteins and the siRNAs stabilizes the siRNAs.

To test if the decrease in the level of the siRNAs in the *rde-8* mutant is due to a lower expression of the secondary argonaute proteins, we overexpressed a secondary Argonaute protein, SAGO-1, in the *rde-8* mutant. The expression of this protein has previously been shown to stabilize secondary siRNAs. We performed unc-22 RNAi by feeding to test for restoration of RNAi sensitivity in this strain. As previously reported, overexpression of SAGO-1 in the *mago* mutant results in hypersensitivity of the strain to

Figure IV-7

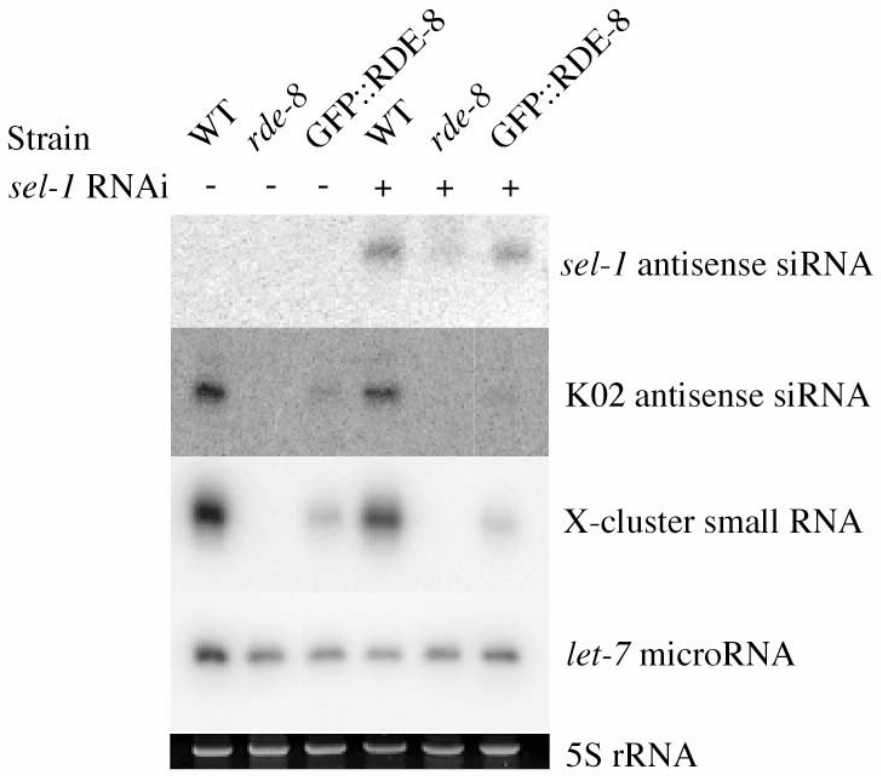

Strain
WT
rde-8
GFP::RDE-8
WT
rde-8
GFP::RDE-8
sel-1 RNAi
- - - + + +
sel-1 antisense siRNA
K02 antisense siRNA
X-cluster small RNA
let-7 microRNA
5S rRNA

Figure IV-7 *rde-8* is required for the accumulation of exo-siRNAs and endo-siRNAs

The expression of different small RNAs in wildtype, *rde-8* mutant and GFP::RDE-8 worms with or without the exposure to dsRNAs was monitored by Northern blotting. The *sel-1* antisense secondary siRNAs can be detected in wildtype worms exposed to *sel-1* dsRNAs but not in the *rde-8* mutant. The presence of the *sel-1* siRNAs can be rescued with the GFP::RDE-8 expression. Similar expression pattern is observed with the K02 endo-siRNAs and the X-cluster small RNAs. The expression of the *let-7* microRNA is not affected in the *rde-8* mutant. The 5S rRNAs were used as loading control.

RNAi (Yigit et al., 2006). This hypersensitivity is due to a higher level of secondary siRNAs in this strain. However, overexpression of SAGO-1 in the *rde-8* mutant does notrender the mutant sensitive to *unc-22* RNAi, suggesting that the RNAi defect in the *rde-8* mutant is independent of the level of the secondary Argonautes (Argonautes).

Discussion

The identification of *rde-8*

Here we report the cloning of the *rde-8* gene. *rde-8* encodes a protein with a conserved domain and a putative nuclear localization sequence. We have named this domain the RDE-8 domain because if was first identified in the *rde-8* gene. The RDE-8 domain is characterized by four conserved glutamic acid residues and one conserved alanine residue. Genes containing the RDE-8 domain can be found in the animal kingdom, suggesting that the RDE-8 domain is evolutionarily conserved. Presently, we do not know about the function of this domain. However, our result demonstrated that the four conserved glutamic acid residues are required for the function of the RDE-8 protein, suggesting that these residues are critical for the gene function of *rde-8*.

We also identified a putative nuclear localization sequence (NLS) motif in the RDE-8 protein. Interestingly, a fraction of the RDE-8 protein localized to the nucleus. However, whether the nuclear localization of the GFP::RDE-8 is directed by the putative NLS is presently unclear. We also tested if the nuclear localization pattern exhibits any change when the worms are exposed to dsRNAs. However, we didn't observe any significant modification in the localization patterns. It is possible that the shuttling of the

GFP::RDE-8 between the nucleus and the cytoplasm is very dynamic and can not be resolved with our current microscopic analysis. Alternatively, it is also possible that two fractions of the RDE-8 proteins function at different locations. Because *rde-8* is required for the accumulation of both exo-siRNAs and endo-siRNAs, it is conceivable that the cytoplasmic RDE-8 proteins are required for the generation of the exo-siRNAs introduced from outside the cell. In contrast, the nuclear RDE-8 proteins maybe required for the production of the endo-siRNAs transcribed from the genome of the *C. elegans*.

The function of *rde-8* in the exo-RNAi pathway and the endo-RNAi pathway

We demonstrated that *rde-8* is required after the primary siRNA are produced. However, in the *rde-8* mutant, the level of the antisense secondary siRNAs is drastically reduced, suggesting that rde-8 functions in the accumulating these siRNAs. *rde-8* could function to assist the activity of the RdRPs to produce secondary siRNAs. Alternatively, *rde-8* could also function in stabilizing these siRNAs. One possibility is that the expression level of the secondary Argonautes that are required for the stability of the secondary siRNAs is reduced, resulting in the decrease of the siRNA level. However, our result indicates that overexpression of a secondary Argonaute protein, SAGO-1, does not rescue the RNAi defect of the *rde-8* mutant, suggesting that this defect is not caused by the loss of secondary Argonaute proteins.

Interestingly, as shown by Northern blotting, there is a very low level of the antisense secondary siRNAs detected in the *rde-8* mutant. It is possible that the mutation of the *rde-8* gene in *rde-8* (*ne3361*) is not null, therefore a very low level of the *rde-8*

activity is accounted for this observation. In addition, the *rde-8* (*ne3361*) mutant is sensitive to RNAi administered with larger quantity of dsRNAs, suggesting that the activity of the *rde-8* gene is not completely abolished. Additional alleles that would severely disrupt the structure of the RDE-8 protein are required to further address this question. Alternatively, considering that the *rde-8* mutant allele is a null allele, the reduced level of the siRNAs would suggest that an additional pathway participates with *rde-8* in producing siRNAs. This pathway may require the activity of other *rde-8* homologs or a different player can function in parallel with the *rde-8* gene for the production of secondary siRNAs.

rde-8 is required for the production of endogenous siRNAs

rde-8 is not required for the production of mature *let-7* microRNAs. Consistent with this, the *rde-8* mutant does not exhibit any developmental defect observed in the *let-7* mutant. However, unlike the low but detectable level of exo-siRNAs in the *rde-8* mutant, none of the K02 endo-siRNAs or X-cluster siRNAs were detected by Nothern blotting. This suggests that the function of *rde-8* is essential for the endo-siRNAs. The discrepancy of the siRNA accumulation between these classes of small RNAs suggests that the biogenesis of the endo-siRNAs and the X-cluster siRNAs are more subceptible to the activity of the *rde-8* gene. At present, we do not know molecularly how *rde-8* exhibits a distinct requirement for the production of the exo-siRNAs, and the endo-siRNAs and the X-cluster siRNAs. Several genes in the mutator class and the Eri class have also been demonstrated to be required for the accumulation of the endogenous siRNAs (Duchaine

et al., 2006; Lee et al., 2006). Therefore, it will be interesting to understand how these different classes of genes function together for the generation of the endogeous siRNAs.

Ackowledgement

Some of the nematode strains were provided by *Caenorhabditis* Genetics Center, funded by National Institutes of Health National Center for Research Resources.

Chapter V

Discussion and Conclusion

Here I present a summary of genetic analysis of RNAi in *C. elegans*. Although rapid progress has been made in understanding the mechanism of RNAi, many questions remain to be answered. Genetic screens conducted in our laboratory led to the initial identification of genes that played a role in RNAi. Further analysis of the homologs of the genes required for RNAi has demonstrated that genes and mechanisms related to RNAi play an important role in additional conserved endogenous small RNA pathways. These small RNAs have critical functions in the regulation of a variety of important biological processes, such as developmental timing, tumorgenesis, fertility and memory formation. To enhance our understanding of the RNAi pathway as well as the endogenous small RNA pathways, a genetic approach is presented in this thesis to study RNAi in *C. elegans*. This approach was aimed at isolating new mutant alleles deficient in RNAi using genetic screens. Analysis of these mutants has resulted in the identification of new genes in the RNAi pathway. The study presented in this thesis will further our knowledge about the RNAi and related pathways in *C. elegans* as well as the other organisms that utilize small RNAs for the regulation of gene expression.

Understanding the RNAi pathway in *C. elegans*

In a continuation of a strategy to identify genes that play a role in RNAi and related pathways, I have conducted several genetic screens in search of *R*NAi *de*ficient mutants (*rde*). After a first round of categorization to identify mutants in which we are most interested in, 76 *rde* mutants were isolated. These mutants are strongly resistant to RNAi as assayed by RNAi administered through injection of high concentration

(1mg/mL) of double stranded RNAs. Among the 76 mutants, 69 were alleles of previously identified RNAi genes isolated our laboratory as well as many others: these include 17 *rde-1* alleles, 2 *mut-15/rde-5* alleles, 1 *mut-16/rde-6* alleles, 22 *sid-1/rde-7* alleles, 27 *rrf-1/rde-9* alleles. 3 additional alleles of *rde-3* were recovered, which led to the identification of the *rde-3* gene. In addition, 3 alleles defined a new *rde* locus, *rde-8*, a conserved gene required for both germline and somatic RNAi.

Several genes in the RNAi pathway are shared with the other essential endogenous small RNA pathways. Mutation in these genes results in lethality or sterility of the corresponding mutants that impedes the recovery of these mutants. We therefore designed genetic screens to identify temperature sensitive *rde* mutants that exhibit lethal phenotypes at the non-permissive temperature. From a trial screen, I have isolated a new allele of *drh-3(ne4197)*. *drh-3* is a Dicer Related Helicase which is required for both RNAi and the viability and fertility of the worm (Duchaine et al., 2006). This result presents a new strategy for the further identification of genes that play a role in both RNAi and development of *C. elegans*.

Biochemical approaches as well as reverse genetics have also been employed in search of genes in the RNAi and related pathways in *C. elegans*. Although some of the genes identified through these approaches were also found in our screens, additional new genes were identified from our mutant collections. This suggests that these methods complement each other in the quest for genes required for RNAi. These genetic mutants also readily provide new reagents for further characterization, which will be presented in the following sections.

We have found two *rde-1* alleles that carry two different missense mutations. The mutated residue in one of these alleles, *ne4085*, is also conserved in human Ago2, suggesting that these residues are critical for the function of the Argonaute proteins. Both of the mutated amino acid residues reside in the linker region between the PAZ and the PIWI domain. These mutant proteins do not show a defect in the interaction with the primary siRNAs and the target mRNAs. However, RNAi is completely abolished in these mutants, suggesting that this linker region is important for the function of the RDE-1 protein. It is possible that these mutations interfere with the cleavage activity of the PIWI domain. Alternatively, in these mutants, an interaction between a protein component and the RDE-1 protein could be lost. This loss could impede the release of the cleaved target mRNA, resulting in the blocking of downstream events.

Analysis of the *rrf-1* mutants identified 16 missense mutations in the RRF-1 protein. All of the molecular lesions were found in a conserved and catalytically active region of the RdRP proteins, suggesting that these residues are important for the enzymatic activity of these enzymes. The biochemical activity of the RdRP protein family involved in RNAi has been tested in *Neurospora crassa* and *S. pombe* (Makeyev and Bamford, 2002; Motamedi et al., 2004). In both cases, dsRNAs were efficiently formed. In addition, a truncated but catalytically active fragment of Neurospora RdRP can also synthesize small RNAs. It is not clear what is the biochemical function of *rrf-1* and *ego-1* in *C. elegans*. Characterization of the biochemical activity of the wildtype as

well as the mutant RdRPs identified from these screens will shed light on the mechanism of the production of the small RNAs.

In light of the necessity of the RdRP gene in the RNAi pathway in *C. elegans*, a frequently asked question is why the activity of the RdRPs is essential for RNAi to function. One can imagine that, because *C. elegans* is already highly sensitive to dsRNAs, there should be no need for the RdRPs to generate even more dsRNAs when a large quantity of dsRNAs is provided. However, this is not the case, as the *rrf-1* mutants are resistant to the promoter-driven RNAi, which results in the production of a large quantity of dsRNAs inside the cells.

Although we don't know the true answer to this question, a possible model is proposed. The reasoning that gives rise to this model is based on the following two observations. First, immunoprecipitation analysis of the components of the DCR-1 complex by Duchaine *et al*. demonstrated that the RDE-4 protein is the most abundant protein in the DCR-1 immuno-complex. However, the amount of the RDE-1 protein in the same immuno-complex is very low, almost below detection level. Nevertheless, the interaction between DCR-1 and RDE-1 can be detected readily and reliably by a more sensitive approach such as Western blotting, suggesting that RDE-1 indeed interact with DCR-1. Second, the characterization of the genetic requirements for the production of recently identified endogenous small RNAs by Ambros *et al*. demonstrated a distinct necessity for the *rde-1* and, *rde-4* and *dcr-1* genes (Lee et al., 2006). Whereas *rde-4* and *dcr-1* are essential for the generation of some endo-siRNAs, *rde-1* is dispensable. These results suggest that the DCR-1/RDE-4 complex may be paired with other Argonaute

proteins to process additional dsRNAs. In addition, the amount of the RNAi initiation complex, which includes DCR-1, RDE-4 and RDE-1, is not as abundant as one might expect. As a result, the tiny amount of the initiation complex may not be sufficient for the degradation of the abundant target mRNAs and the need of amplification by the RdRPs, is essential. Based on this model, it is possible that increasing the amount of the initiation complex in the absence of the amplification mechanism could lead to degradation of more target mRNAs which would result in distinguishable silencing. One way to test this hypothesis is to overexpress the RDE-1 proteins in the *rrf-1* mutant and assay for any slight regain of sensitivity to RNAi in this genetic background.

The small RNAs in *C. elegans*

rde-3, *mut-15/rde-5*, and *drh-3* mutants also exhibit developmental defects. *rde-3* homozygous mutant produces around 20 percent of dead embryos and some of the worms are also sterile. A presumptive null allele of *mut-15/rde-5* exhibits sterile phenotype at non-permissive temperature due to a sperm defect. The temperature sensitive *drh-3* allele bearing a missense mutation results in 100% sterility of the animals at non-permissive temperature. All three mutants are also mutators due to defects in transposon silencing in the germline. The desilenced transposons are able to excise themselves out of the genome and re-insert into a new locus, resulting in generation of secondary mutations that can affect the viability of the worms. However, not all mutations are due to transposon silencing defects. In order to exclude the possibility that the lethal phenotypes observed in these mutants are resulted from such secondary mutations, the phenotypic analysis

Figure V-1

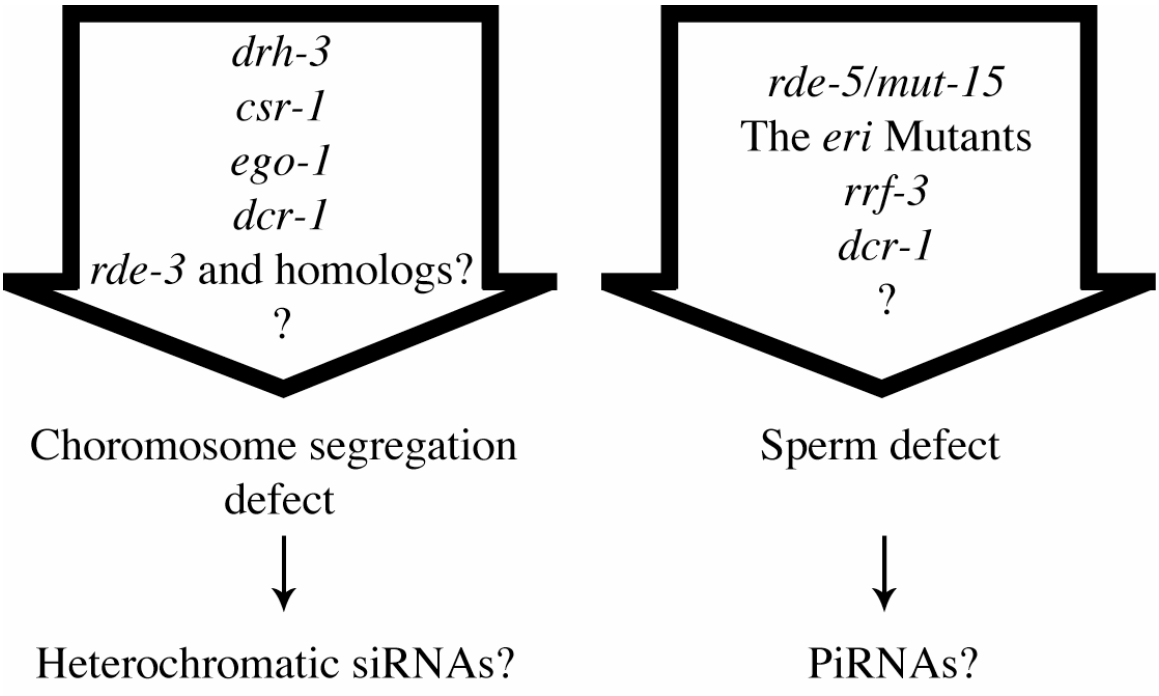

drh-3
csr-1
ego-1
dcr-1
rde-3 and homologs?
?
rde-5/mut-15
The eri Mutants
rrf-3
dcr-1
?
Choromosome segregation defect
Sperm defect
Heterochromatic siRNAs?
PiRNAs?

Figure V-1 A Model for possible small RNA pathways in *C. elegans*

Mutants that exhibit similar phenotypes are grouped in the large arrows. The possible small RNA defects are indicated by the small arrows. See text for details.

were performed immediately after at least 6 outcrossings to get rid of the background mutations. Therefore, it is more likely that these phenotypes are due to a more direct effect of the lesions in these genes.

Currently, we do not know the exact nature that causes lethal phenotypes of these mutants. A mutant allele that bears a deletion in the *drh-3* gene has been reported to exhibit a chromosome segregation phenotype (Duchaine et al., 2006). Similar phenotype has also been observed in a member of the Argonaute gene family, *csr-1* (*C*hromosome *S*egregation and *R*NAi Defective), suggesting that these genes are, directly or indirectly, involved in the proper segregation of the chromosomes (Yigit et al., 2006).

Interestingly, a homolog of *rde-3* in *S. pombe*, cid1, has also been shown to be required for proper chromosome segregation (Wang et al., 2000a). Additionally, another homolog in S. pombe, cid12, has been reported to play a role in the heterochromatin silencing in the centrosomal region and defects in this silencing will also cause a defect in chromosome segregation (Motamedi et al., 2004). This raises the question whether *rde-3* could also be involved in this pathway. The rde-3 mutant does not appear to exhibit a chromosome segregation defect. However, there are at least 9 additional paralogs of *rde-3* in the genome of *C. elegans* and many of which exhibit lethal phenotypes by RNAi. Therefore, it is possible that some of these genes function together with *rde-3* to mediate proper chromosome segregation (Figure V-1).

The *mut-15/rde-5* mutant exhibits a temperature dependent-sterile phenotype. This phenotype is not observed in other mutator class of mutants, suggesting that it is most likely not a general defect caused by tansposon mobilization. Interestingly, another

class of mutants, known as the *eri* mutants (*e*nhancer of *R*NA*i*), also exhibits a temperature dependent sterile phenotype similar to that of *mut-15/rde-5* (Duchaine et al., 2006). The Eri mutants were first isolated from genetic screens in search of mutant that exhibit "enhanced" RNAi phenotype. This phenotype is observed when a mutant is sensitive to RNAi targeting certain genes that are normally not susceptible in the wildtype animals. Several *eri* genes, including *eri-1, eri-3* and *eri-5,* have recently been shown to associate with *dcr-1* (Duchaine et al., 2006). In addition, these *eri* genes are also required for the production of a variety of small RNAs including endo-siRNAs, tncRNAs and X-cluster RNAs, suggesting that these genes function in the endo-RNAi pathway for the generation of endogenous small RNAs. Because both *mut-15/rde-5* and the Eri pathway are involved in the small RNA pathway, it is tantalizing to propose that these genes are required for production of small RNAs that are specifically required for the function of sperm. However, several mutants that do not exhibit the same sterile phenotype as the *eri* and *mut-15/rde-5* mutants are also defective in the production of the endo-siRNAs. These results suggest that the sterile phenotype is not caused by the defect of these small RNAs. Therefore, the identity of these small RNAs remains to be investigated. Recently, a new class of small RNAs of approximately 30nt long has been cloned from rat, mouse and human (Aravin et al., 2006; Girard et al., 2006; Grivna et al., 2006; Lau et al., 2006). These small RNAs associate with Argonaute proteins in the PIWI subfamily and are termed piRNAs (Piwi interacting RNAs). The piRNAs are encoded by clustered genomic loci and accumulate as single stranded RNAs. Although the size and expression pattern of the piRNAs are conserved among rat, mouse and human, the

conservation of the piRNA sequences is not observed. Mutants of several genes in the Piwi subfamily exhibit male sterility, suggesting that the piRNAs may play a role in spermatogenesis (Deng and Lin, 2002; Kuramochi-Miyagawa et al., 2004). Therefore, it is conceivable that the piRNA homologs in *C. elegans* require the function of the *eri* and *mut-15/rde-5* genes for their production. As more genes and small RNAs are identified each day, it will be of great importance to dissect the pathway for the production of each class of small RNAs in more detail.

References

Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350-355.

Ambros, V., and Lee, R. C. (2004). Identification of microRNAs and other tiny noncoding RNAs by cDNA cloning. Methods Mol Biol *265*, 131-158.

Ambros, V., Lee, R. C., Lavanway, A., Williams, P. T., and Jewell, D. (2003). MicroRNAs and other tiny endogenous RNAs in C. elegans. Curr Biol *13*, 807-818.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M. J., Kuramochi-Miyagawa, S., Nakano, T., *et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravind, L., and Koonin, E. V. (1999). DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res *27*, 1609-1618.

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

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

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

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

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.

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

Cerutti, L., Mian, N., and Bateman, A. (2000). Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. Trends Biochem Sci *25*, 481-482.

Chen, C. C., Simard, M. J., Tabara, H., Brownell, D. R., McCollough, J. A., and Mello, C. C. (2005). A member of the polymerase beta nucleotidyltransferase superfamily is required for RNA interference in C. elegans. Curr Biol *15*, 378-383.

Cogoni, C., and Macino, G. (1999). Gene silencing in Neurospora crassa requires a protein homologous to RNA-dependent RNA polymerase. Nature *399*, 166-169.

Collins, J. J., and Anderson, P. (1994). The Tc5 family of transposable elements in Caenorhabditis elegans. Genetics *137*, 771-781.

Dalmay, T., Hamilton, A., Rudd, S., Angell, S., and Baulcombe, D. C. (2000). An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. Cell *101*, 543-553.

Deng, W., and Lin, H. (2002). miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. Dev Cell *2*, 819-830.

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

Dernburg, A. F., Zalevsky, J., Colaiacovo, M. P., and Villeneuve, A. M. (2000). Transgene-mediated cosuppression in the C. elegans germ line. Genes Dev *14*, 1578-1583.

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

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

Dong, B., and Silverman, R. H. (1995). 2-5A-dependent RNase molecules dimerize during activation by 2-5A. J Biol Chem *270*, 4133-4137.

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

Duxbury, M. S., Ashley, S. W., and Whang, E. E. (2005). RNA interference: a mammalian SID-1 homologue enhances siRNA uptake and gene silencing efficacy in human cells. Biochem Biophys Res Commun *331*, 459-463.

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

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

Feinberg, E. H., and Hunter, C. P. (2003). Transport of dsRNA into cells by the transmembrane protein SID-1. Science *301*, 1545-1547.

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

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

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

Glazov, E., Phillips, K., Budziszewski, G. J., Schob, H., Meins, F., Jr., and Levin, J. Z. (2003). A gene encoding an RNase D exonuclease-like protein is required for post-transcriptional silencing in Arabidopsis. Plant J *35*, 342-349.

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

Grishok, A., Tabara, H., and Mello, C. C. (2000). Genetic requirements for inheritance of RNAi in C. elegans. Science *287*, 2494-2497.

Grivna, S. T., Beyret, E., Wang, Z., and Lin, H. (2006). A novel class of small RNAs in mouse spermatogenic cells. Genes Dev *20*, 1709-1714.

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

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

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

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

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

Hartmann, R., Justesen, J., Sarkar, S. N., Sen, G. C., and Yee, V. C. (2003). Crystal structure of the 2'-specific and double-stranded RNA-activated interferon-induced antiviral protein 2'-5'-oligoadenylate synthetase. Mol Cell *12*, 1173-1185.

Hovanessian, A. G., Brown, R. E., and Kerr, I. M. (1977). Synthesis of low molecular weight inhibitor of protein synthesis with enzyme from interferon-treated cells. Nature *268*, 537-540.

Hutvagner, G., McLachlan, J., Pasquinelli, A. E., Balint, E., Tuschl, T., and Zamore, P. D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.

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

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

Johnson, S. M., Lin, S. Y., and Slack, F. J. (2003). The time of appearance of the C. elegans let-7 microRNA is transcriptionally controlled utilizing a temporal regulatory element in its promoter. Dev Biol *259*, 364-379.

Kadaba, S., Krueger, A., Trice, T., Krecic, A. M., Hinnebusch, A. G., and Anderson, J. (2004). Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. Genes Dev *18*, 1227-1240.

Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature *427*, 645-649.

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

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

Ketting, R. F., Haverkamp, T. H., van Luenen, H. G., and Plasterk, R. H. (1999). Mut-7 of C. elegans, required for transposon silencing and RNA interference, is a homolog of Werner syndrome helicase and RNaseD. Cell *99*, 133-141.

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

Kuramochi-Miyagawa, S., Kimura, T., Ijiri, T. W., Isobe, T., Asada, N., Fujita, Y., Ikawa, M., Iwai, N., Okabe, M., Deng, W., *et al.* (2004). Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. Development *131*, 839-849.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853-858.

Lai, E. C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet *30*, 363-364.

Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858-862.

Lau, N. C., Seto, A. G., Kim, J., Kuramochi-Miyagawa, S., Nakano, T., Bartel, D. P., and Kingston, R. E. (2006). Characterization of the piRNA complex from rat testes. Science *313*, 363-367.

Lee, R. C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862-864.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843-854.

Lee, R. C., Hammell, C. M., and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. Rna *12*, 589-597.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H., and Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. Embo J *23*, 4051-4060.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature *433*, 769-773.

Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B., and Bartel, D. P. (2003). The microRNAs of Caenorhabditis elegans. Genes Dev *17*, 991-1008.

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

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

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

Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95-98.

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

Makeyev, E. V., and Bamford, D. H. (2002). Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. Mol Cell *10*, 1417-1427.

Martinez, J., and Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. Genes Dev *18*, 975-980.

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

Mello, C. C., and Conte, D., Jr. (2004). Revealing the world of RNA interference. Nature *431*, 338-342.

Montgomery, M. K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. Proc Natl Acad Sci U S A *95*, 15502-15507.

Motamedi, M. R., Verdel, A., Colmenares, S. U., Gerber, S. A., Gygi, S. P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C., Morel, J. B., Jouette, D., Lacombe, A. M., Nikic, S., Picault, N., *et al.* (2000). Arabidopsis SGS2 and SGS3 genes are required for posttranscriptional gene silencing and natural virus resistance. Cell *101*, 533-542.

Newmark, P. A., Reddien, P. W., Cebria, F., and Sanchez Alvarado, A. (2003). Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. Proc Natl Acad Sci U S A *100 Suppl 1*, 11861-11865.

Olsen, P. H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol *216*, 671-680.

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

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

Parrish, S., Fleenor, J., Xu, S., Mello, C., and Fire, A. (2000). Functional anatomy of a dsRNA trigger: differential requirement for the two trigger strands in RNA interference. Mol Cell *6*, 1077-1087.

Qiao, L., Lissemore, J. L., Shu, P., Smardon, A., Gelber, M. B., and Maine, E. M. (1995). Enhancers of glp-1, a gene required for cell-signaling in Caenorhabditis elegans, define a set of genes required for germline development. Genetics *141*, 551-569.

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

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature *403*, 901-906.

Robert, V. J., Sijen, T., van Wolfswinkel, J., and Plasterk, R. H. (2005). Chromatin and RNAi factors protect the C. elegans germline against repetitive sequences. Genes Dev *19*, 782-787.

Roberts, W. K., Hovanessian, A., Brown, R. E., Clemens, M. J., and Kerr, I. M. (1976). Interferon-mediated protein kinase and low-molecular-weight inhibitor of protein synthesis. Nature *264*, 477-480.

Rocheleau, C. E., Downs, W. D., Lin, R., Wittmann, C., Bei, Y., Cha, Y. H., Ali, M., Priess, J. R., and Mello, C. C. (1997). Wnt signaling and an APC-related gene specify endoderm in early C. elegans embryos. Cell *90*, 707-716.

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

Saleh, M. C., van Rij, R. P., Hekele, A., Gillis, A., Foley, E., O'Farrell, P. H., and Andino, R. (2006). The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. Nat Cell Biol *8*, 793-802.

Schiebel, W., Haas, B., Marinkovic, S., Klanner, A., and Sanger, H. L. (1993). RNA-directed RNA polymerase from tomato leaves. II. Catalytic in vitro properties. J Biol Chem *268*, 11858-11867.

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

Schwarz, D. S., Tomari, Y., and Zamore, P. D. (2004). The RNA-induced silencing complex is a Mg2+-dependent endonuclease. Curr Biol *14*, 787-791.

Sen, G. L., and Blau, H. M. (2005). Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. Nat Cell Biol *7*, 633-636.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K. L., Parrish, S., Timmons, L., Plasterk, R. H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465-476.

Sijen, T., and Plasterk, R. H. (2003). Transposon silencing in the Caenorhabditis elegans germ line by natural RNAi. Nature *426*, 310-314.

Simmer, F., Tijsterman, M., Parrish, S., Koushika, S. P., Nonet, M. L., Fire, A., Ahringer, J., and Plasterk, R. H. (2002). Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. elegans hypersensitive to RNAi. Curr Biol *12*, 1317-1319.

Smardon, A., Spoerke, J. M., Stacey, S. C., Klein, M. E., Mackin, N., and Maine, E. M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in C. elegans. Curr Biol *10*, 169-178.

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

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

Stein, P., Svoboda, P., Anger, M., and Schultz, R. M. (2003). RNAi: mammalian oocytes do it without RNA-dependent RNA polymerase. Rna *9*, 187-192.

Sundaram, P., Echalier, B., Han, W., Hull, D., and Timmons, L. (2006). ATP-binding cassette transporters are required for efficient RNA interference in Caenorhabditis elegans. Mol Biol Cell *17*, 3678-3688.

Tabara, H., Grishok, A., and Mello, C. C. (1998). RNAi in C. elegans: soaking in the genome sequence. Science *282*, 430-431.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell *99*, 123-132.

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

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

Tijsterman, M., Ketting, R. F., Okihara, K. L., Sijen, T., and Plasterk, R. H. (2002a). RNA helicase MUT-14-dependent gene silencing triggered in C. elegans by short antisense RNAs. Science *295*, 694-697.

Tijsterman, M., May, R. C., Simmer, F., Okihara, K. L., and Plasterk, R. H. (2004). Genes required for systemic RNA interference in Caenorhabditis elegans. Curr Biol *14*, 111-116.

Tijsterman, M., Okihara, K. L., Thijssen, K., and Plasterk, R. H. (2002b). PPW-1, a PAZ/PIWI protein required for efficient germline RNAi, is defective in a natural isolate of C. elegans. Curr Biol *12*, 1535-1540.

Timmons, L., Court, D. L., and Fire, A. (2001). Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. Gene *263*, 103-112.

Tomari, Y., and Zamore, P. D. (2005). Perspective: machines for RNAi. Genes Dev *19*, 517-529.

Tops, B. B., Tabara, H., Sijen, T., Simmer, F., Mello, C. C., Plasterk, R. H., and Ketting, R. F. (2005). RDE-2 interacts with MUT-7 to mediate RNA interference in Caenorhabditis elegans. Nucleic Acids Res *33*, 347-355.

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

Ulvila, J., Parikka, M., Kleino, A., Sormunen, R., Ezekowitz, R. A., Kocks, C., and Ramet, M. (2006). Double-stranded RNA is internalized by scavenger receptor-mediated endocytosis in Drosophila S2 cells. J Biol Chem *281*, 14370-14375.

Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P. D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. Science *313*, 320-324.

Vastenhouw, N. L., Fischer, S. E., Robert, V. J., Thijssen, K. L., Fraser, A. G., Kamath, R. S., Ahringer, J., and Plasterk, R. H. (2003). A genome-wide screen identifies 27 genes involved in transposon silencing in C. elegans. Curr Biol *13*, 1311-1316.

Vaucheret, H., Beclin, C., and Fagard, M. (2001). Post-transcriptional gene silencing in plants. J Cell Sci *114*, 3083-3091.

Wang, L., Eckmann, C. R., Kadyk, L. C., Wickens, M., and Kimble, J. (2002). A regulatory cytoplasmic poly(A) polymerase in Caenorhabditis elegans. Nature *419*, 312-316.

Wang, S. W., Toda, T., MacCallum, R., Harris, A. L., and Norbury, C. (2000a). Cid1, a fission yeast protein required for S-M checkpoint control when DNA polymerase delta or epsilon is inactivated. Mol Cell Biol *20*, 3234-3244.

Wang, Z., Castano, I. B., De Las Penas, A., Adams, C., and Christman, M. F. (2000b). Pol kappa: A DNA polymerase required for sister chromatid cohesion. Science *289*, 774-779.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.

Winston, W. M., Molodowitch, C., and Hunter, C. P. (2002). Systemic RNAi in C. elegans requires the putative transmembrane protein SID-1. Science *295*, 2456-2459.

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

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

Yigit, E., Batista, P. J., Bei, Y., Pang, K. M., Chen, C. C., Tolia, N. H., Joshua-Tor, L., Mitani, S., Simard, M. J., and Mello, C. C. (2006). Analysis of the C. elegans Argonaute

Family Reveals that Distinct Argonautes Act Sequentially during RNAi. Cell *127*, 747-757.

Yuan, Y. R., Pei, Y., Ma, J. B., Kuryavyi, V., Zhadina, M., Meister, G., Chen, H. Y., Dauter, Z., Tuschl, T., and Patel, D. J. (2005). Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. Mol Cell *19*, 405-419.

Zamore, P. D., and Haley, B. (2005). Ribo-gnome: the big world of small RNAs. Science *309*, 1519-1524.

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

Zeng, Y., Wagner, E. J., and Cullen, B. R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol Cell *9*, 1327-1333.

Zeng, Y., Yi, R., and Cullen, B. R. (2003). MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci U S A *100*, 9779-9784.

Appendix I

Current Biology, Vol. 15, 378–383, February 22, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.cub.2005.01.009

A Member of the Polymerase β Nucleotidyltransferase Superfamily Is Required for RNA Interference in *C. elegans*

Chun-Chieh G. Chen,[1] Martin J. Simard,[1] Hiroaki Tabara,[3] Daniel R. Brownell,[1,2] Jennifer A. McCollough,[1] and Craig C. Mello[1,2,*]
[1]Program in Molecular Medicine and
[2]Howard Hughes Medical Institute
University of Massachusetts Medical School
Worcester, Massachusetts 01605
[3]Graduate School of Medicine
Horizontal Medical Research Organization (HMRO)
Kyoto University
Kyoto 606-8501
Japan

Summary

RNA interference (RNAi) is an ancient, highly conserved mechanism in which small RNA molecules (siRNAs) guide the sequence-specific silencing of gene expression [1]. Several silencing machinery protein components have been identified, including helicases, RNase-related proteins, double- and single-stranded RNA binding proteins, and RNA-dependent RNA polymerase-related proteins [2]. Work on these factors has led to the revelation that RNAi mechanisms intersect with cellular pathways required for development and fertility [3, 4]. Despite rapid progress in understanding key steps in the RNAi pathway, it is clear that many factors required for both RNAi and related developmental mechanisms have not yet been identified. Here, we report the characterization of the *C. elegans* gene *rde-3*. Genetic analysis of presumptive null alleles indicates that *rde-3* is required for siRNA accumulation and for efficient RNAi in all tissues, and it is essential for fertility and viability at high temperatures. RDE-3 contains conserved domains found in the polymerase β nucleotidyltransferase superfamily, which includes conventional poly(A) polymerases, 2′-5′ oligoadenylate synthetase (OAS), and yeast Trf4p [5]. These findings implicate a new enzymatic modality in RNAi and suggest possible models for the role of RDE-3 in the RNAi mechanism.

Results and Discussion

Molecular Identification of *rde-3*

The *rde-3* locus was originally defined by a single allele [6]. To further characterize this locus, we first conducted additional genetic screens to identify new alleles and then used genetic mapping to define a small interval containing the gene (Figure 1A). We then examined the nucleotide sequence of candidate genes within this region for lesions in *rde-3*. In each allele we found a mutation predicted to disrupt the expression or to change the amino-acid sequence of the gene K04F10.6. In addition, we found a lesion in K04F10.6 in the previously described mutant *mut-2 (r459)* [7], suggesting that *mut-2* and *rde-3* are allelic. Finally, we found that a PCR fragment containing only K04F10.6 rescues the somatic RNAi defect of *ne3364*, confirming that K04F10.6 is *rde-3* (Figure 1).

rde-3 encodes a conserved protein in the polymerase β nucleotidyltransferase superfamily. RDE-3, and other members in this family, including conventional poly(A) polymerases (PAP), 2′-5′ oligoadenylate synthetase (OAS), *C. elegans* GLD-2, and budding yeast Trf4p, all contain a nucleotidyltransferase 2 (NTP transferase 2) domain [5]. In addition, RDE-3 and a subset of family members, including GLD-2 and Trf4p, contain an additional domain, of unknown function, named the PAP/25A-associated domain [5]. The *rde-3* alleles *ne298*, *r459*, and *fj14* each contain point mutations predicted to alter conserved amino acids. The *ne3370* allele contains an in-frame deletion of 423 bp (Figure 1B). *ne4243* and *ne3364* each contain nonsense mutations predicted to truncate the protein after 164 and 169 amino acids, respectively, and are thus likely to represent null mutations. Two of the lesions in *rde-3* are predicted to alter conserved residues within the nucleotidyltransferase domain, suggesting that polymerase activity may be important for RDE-3 function (see discussion below).

rde-3 Is Required for Fertility and Viability as well as siRNA Accumulation

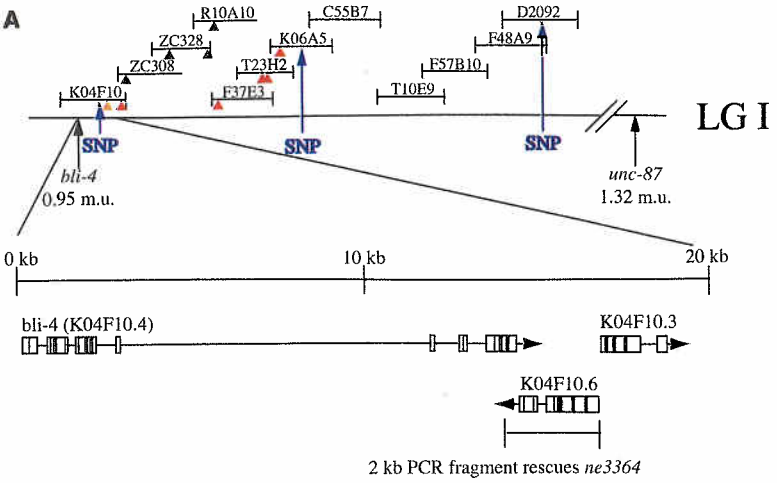
The previous characterization of *ne298* showed that *rde-3* had additional phenotypes, including a Him (High incidence of males) phenotype that reflects an increase in the nondisjunction of the X chromosome and a Mut (Mutator) phenotype resulting from the activation of endogenous transposon families [6]. The new *rde-3* alleles, including the presumptive null alleles, exhibit phenotypes similar to those of *rde-3 (ne298)*, suggesting that all of these mutants represent strong loss-of-function alleles. *rde-3* mutants do not exhibit the specific developmental defects expected from a loss in the function of the miRNAs *lin-4* or *let-7* (data not shown), suggesting that, as with other mutator-class *rde* mutants, its developmental functions in fertility and viability may reflect a role in other as-yet-unknown mechanisms.

We next examined the sensitivity of these alleles to RNAi targeting genes expressed in different tissues: *pos-1* (germline) and *unc-22* (muscle). The three alleles tested (*ne298*, *ne3364*, and *ne3370*) showed similar levels of resistance to RNAi (Figures 2A and 2B). Note that the dead embryos observed in Figure 2A include a background of approximately 20% inviable embryos that arise spontaneously because of the incompletely penetrant lethal phenotype associated with *rde-3* alleles. We found that the presumptive null allele, *ne3364*, like *ne298* [6], is sensitive to RNAi induced by a transgene that simultaneously drives the expression of sense and antisense *unc-22* RNA. Two independent transgenic lines were analyzed in these experiments; the penetrance of the *unc-22* twitching phenotype was indistinguishable from that induced by the same transgenes

*Correspondence: craig.mello@umassmed.edu

T., and Fukui, Y. (2003). DOCK2 regulates Rac activation and cytoskeletal reorganization through interaction with ELMO1. Blood *102*, 2948–2950.

3. Namekata, K., Enokido, Y., Iwasawa, K., and Kimura, H. (2004). MOCA induces membrane spreading by activating Rac1. J. Biol. Chem. *279*, 14331–14337.
4. Yajnik, V., Paulding, C., Sordella, R., McClatchey, A.I., Saito, M., Wahrer, D.C., Reynolds, P., Bell, D.W., Lake, R., van den Heuvel, S., et al. (2003). DOCK4, a GTPase activator, is disrupted during tumorigenesis. Cell *112*, 673–684.
5. Kiyokawa, E., Hashimoto, Y., Kobayashi, S., Sugimura, H., Kurata, T., and Matsuda, M. (1998). Activation of Rac1 by a Crk SH3-binding protein, DOCK180. Genes Dev. *12*, 3331–3336.
6. Nolan, K.M., Barrett, K., Lu, Y., Hu, K.Q., Vincent, S., and Settleman, J. (1998). Myoblast city, the Drosophila homolog of DOCK180/CED-5, is required in a Rac signaling pathway utilized for multiple developmental processes. Genes Dev. *12*, 3337–3342.
7. Cote, J.F., and Vuori, K. (2002). Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. J. Cell Sci. *115*, 4901–4913.
8. Wu, Y.C., and Horvitz, H.R. (1998). C. elegans phagocytosis and cell-migration protein CED-5 is similar to human DOCK180. Nature *392*, 501–504.
9. Erickson, M.R., Galletta, B.J., and Abmayr, S.M. (1997). Drosophila myoblast city encodes a conserved protein that is essential for myoblast fusion, dorsal closure, and cytoskeletal organization. J. Cell Biol. *138*, 589–603.
10. Gumienny, T.L., Brugnera, E., Tosello-Trampont, A.C., Kinchen, J.M., Haney, L.B., Nishiwaki, K., Walk, S.F., Nemergut, M.E., Macara, I.G., Francis, R., et al. (2001). CED-12/ELMO, a novel member of the crkII/Dock180/Rac pathway, is required for phagocytosis and cell migration. Cell *107*, 27–41.
11. Grimsley, C.M., Kinchen, J.M., Tosello-Trampont, A.C., Brugnera, E., Haney, L.B., Lu, M., Chen, Q., Klingele, D., Hengartner, M.O., and Ravichandran, K.S. (2004). Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration. J. Biol. Chem. *279*, 6087–6097.
12. Meller, N., Irani-Tehrani, M., Kiosses, W.B., Del Pozo, M.A., and Schwartz, M.A. (2002). Zizimin1, a novel Cdc42 activator, reveals a new GEF domain for Rho proteins. Nat. Cell Biol. *4*, 639–647.
13. Janardhan, A., Swigut, T., Hill, B., Myers, M.P., and Skowronski, J. (2004). HIV-1 Nef binds the DOCK2-ELMO1 complex to activate Rac and inhibit lymphocyte chemotaxis. PLoS Biol 2(1): e6. DOI: 10.1371/journal.pbio.0020006
14. Albert, M., Kim, J., and Birge, R. (2000). $\alpha v\beta 5$ integrin recruits the CrkII–Dock180–Rac1 complex for phagocytosis of apoptotic cells. Nat. Cell Biol. *2*, 899–905.
15. Fukui, Y., Hashimoto, O., Sanui, T., Oono, T., Koga, H., Abe, M., Inayoshi, A., Noda, M., Oike, M., Shirai, T., et al. (2001). Haematopoietic cell-specific CDM family protein DOCK2 is essential for lymphocyte migration. Nature *412*, 826–831.
16. Hasegawa, H., Kiyokawa, E., Tanaka, S., Nagashima, K., Gotoh, N., Shibuya, M., Kurata, T., and Matsuda, M. (1996). DOCK180, a major CRK-binding protein, alters cell morphology upon translocation to the cell membrane. Mol. Cell. Biol. *16*, 1770–1776.
17. Nishihara, H., Kobayashi, S., Hashimoto, Y., Ohba, F., Mochizuki, N., Kurata, T., Nagashima, K., and Matsuda, M. (1999). Non-adherent cell-specific expression of DOCK2, a member of the human CDM-family proteins. Biochim. Biophys. Acta *1452*, 179–187.
18. Lu, M., Kinchen, J.M., Rossman, K.L., Grimsley, C., deBakker, C., Brugnera, E., Tosello-Trampont, A.C., Haney, L.B., Klingele, D., Sondek, J., et al. (2004). PH domain of ELMO functions in trans to regulate Rac activation via Dock180. Nat. Struct. Mol. Biol. *11*, 756–762. Published online June 11, 2004. 10.1038/nsmb800
19. Wu, Y.C., Tsai, M.C., Cheng, L.C., Chou, C.J., and Weng, N.Y. (2001). C. elegans CED-12 acts in the conserved crkII/DOCK180/Rac pathway to control cell migration and cell corpse engulfment. Dev. Cell *1*, 491–502.
20. Zhou, Z., Caron, E., Hartwieg, E., Hall, A., and Horvitz, H.R. (2001). The C. elegans PH domain protein CED-12 regulates cytoskeletal reorganization via a Rho/Rac GTPase signaling pathway. Dev. Cell *1*, 477–489.
21. Schmidt, A., and Hall, A. (2002). Guanine nucleotide exchange factors for the Rho GTPases: Turning on the switch. Genes Dev. *16*, 1587–1609.



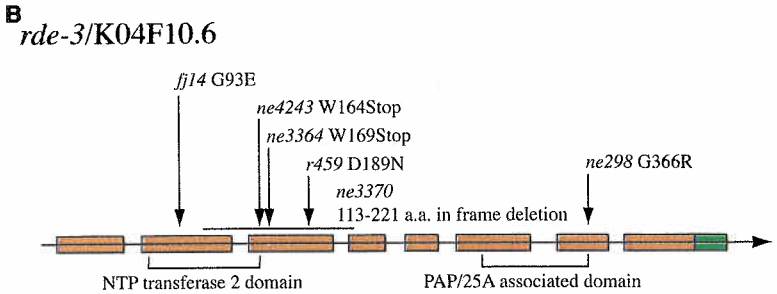


Figure 1. Molecular Identification of *rde-3*

(A) Schematic diagram of the *rde-3* genetic interval. *rde-3* was mapped genetically to an interval of approximately 0.37 map units, very close and to the right of *bli-4*. Single nucleotide polymorphisms (blue arrows) were used to narrow the interval to 7 cosmids. Candidate genes shown as red triangles were sequenced. Mutations were found in K04F10.6 (orange triangle). A single-gene K04F10.6 PCR fragment was used for rescue experiments (forward primer: 5′-GAT TGT ATT GTT TCT TTT GTC TTA TAC CGG G-3′, reverse primer: 5′-CGT TGG AGA AAC GAA GAA TGT GCA TAG-3′).

(B) Line and block diagram of the intron/exon structure of *rde-3*. Reverse-transcriptase PCR was used to confirm the exon structure predicted by Wormbase. RDE-3 contains two conserved domains, including an NTP transferase 2 domain at its N terminus and PAP/25A (poly(A) polymerase/2′-5′ oligoadenylate)-associated domain at its C terminus. The molecular lesions in six *rde-3* alleles are indicated above the diagram.

when expressed in wild-type strains (data not shown). These findings indicate that RDE-3 is essential for RNAi induced by feeding, but is not required for RNAi induced by transgene-driven expression of dsRNA. RDE-3 has one close homolog and at least ten more distantly related family members in *C. elegans*. Therefore, it is possible that the ability of *rde-3* mutants to respond to transgene-driven RNAi reflects compensation in this pathway from one or more of these homologs. Alternatively, it is possible that RDE-3 function is required for RNAi amplification and that RDE-3 is thus less critical when very large amounts of dsRNA are delivered by transgene-driven expression directly in target tissues (for more on these possibilities, see the discussion below and the model in Figure 4B).

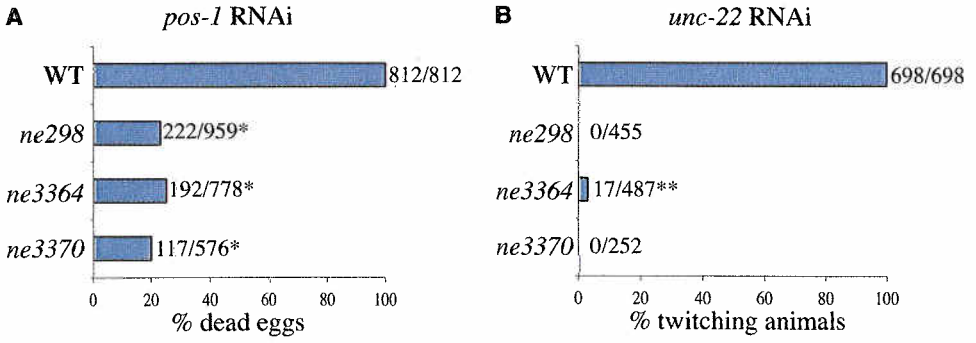
In *C. elegans*, antisense siRNAs accumulate in animals undergoing RNAi. Some mutants defective in RNAi, such as *rde-1* and *mut-7*, fail to accumulate siRNAs, whereas other mutants, such as *mut-14*, show no defect in siRNA accumulation [8]. To further understand why RNAi is defective in *rde-3*, we examined the level of siRNA in *rde-3* mutants. As shown in Figure 2C, wild-type animals accumulate antisense *pos-1* siRNAs when they are exposed to RNAi targeting *pos-1*. However, siRNAs are not detected in either *ne298* or *ne3364* worms. This result suggests that RDE-3 activity is required upstream of the accumulation of siRNA during RNAi.

rde-3 Is a Member of a Functionally Diverse Multigene Family

The *rde-3* family includes several *C. elegans* members, as well as numerous homologs in other organisms. These homologs have been implicated in a variety of cellular mechanisms, including polyadenylation of transcripts in the nucleus (PAP) and cytoplasm (GLD-2) [9]. The OAS members of this family are upregulated by the interferon antiviral response in mammalian cells [10, 11]. Viral dsRNA is thought to bind to OAS and to activate the synthesis of 2′-5′ adenyl oligomers, which in turn activate RNase L [12]. RNase L then degrades cytoplasmic RNAs of both the host and virus in a non-sequence-specific manner, preventing further viral amplification. The budding yeast protein Trf4p has been implicated in the polyadenylation and recruitment of improperly processed tRNAs to the exosome [13], and it has also been implicated in DNA synthesis and chromosome cohesion [14].

Like GLD-2, OAS, and Trf4p, RDE-3 lacks any recognizable RNA binding motif. GLD-2 is thought to interact with target mRNAs through its association with the KH-domain RNA binding protein GLD-3 [9]. Mammalian OAS family members bind to dsRNA via an undefined RNA binding motif that appears to involve several positively charged residues localized on the surface of the protein [15]. It is not presently known how Trf4p recognizes its targets. Further studies will be required to determine whether and how RDE-3 binds to RNA.

Proteins in this superfamily have two conserved features at the active site of the enzyme: a helical turn that includes a highly conserved glycine-serine (GS) motif and a conserved aspartic-acid triad [5]. Interestingly, *rde-3 (fj14)* contains a glycine-to-glutamate mutation in the GS motif, and *rde-3 (r459)* contains an aspartic-acid-to-asparagine mutation in the third aspartic acid of the



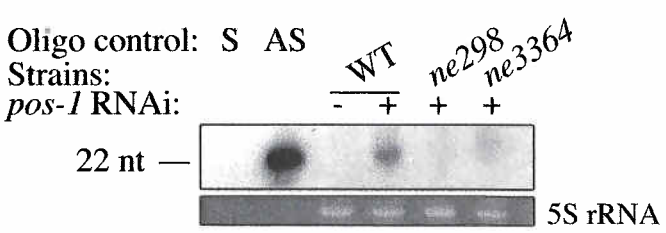


Figure 2. *rde-3* Is Defective in RNAi

(A and B) Graphical representations of the sensitivity of three *rde-3* mutant strains to RNAi induced by microinjection of: (A) double-stranded *pos-1* RNA (1 mg/ml) or (B) double-stranded *unc-22* RNA (1 mg/ml). Young adult animals were injected with dsRNA and were transferred individually to fresh plates every 12 hr. Phenotypes were scored for the progeny produced 24–48 hr after injection. The number of sensitive animals over the total number of animals scored is presented next to each bar. *: The data for the three *rde-3* strains includes a background of approximately 20% inviable embryos that arise spontaneously because of the incompletely penetrant embryonic lethal phenotype of *rde-3*. **: These progeny were very weakly twitching in comparison to the wild-type.

(C) *rde-3* is defective in siRNA accumulation. Northern blot analysis of siRNAs in wild-type and *rde-3* mutant animals is shown. Total RNAs were extracted with TRI Reagent (Molecular Research Center) from embryos. Small RNAs were further purified with MirVana (Ambion). Ten micrograms of RNAs were separated in 15% UREA-PAGE and transferred to nitrocellulose membrane. Northern analysis was performed with ^{32}P-labeled sense riboprobe covering the whole *pos-1* coding region [18]. Ethidium bromide staining of 5S rRNAs was used for loading control. Sense (S) and antisense (AS) *pos-1* RNA oligos of 22 nt were included as hybridization controls.

catalytic triad (Figure 3). These findings suggest that the polymerase activity of RDE-3 is likely to be important for its function. These residues are also conserved in a fission-yeast homolog, Cid12 (Figure 3). Interestingly, Cid12 has recently been shown to function in an RNAi-like mechanism required for chromatin silencing at centromeric repeats [16]. Cid12 associates with the RNA-dependent RNA polymerase Rdp1 and is required for the accumulation of siRNAs involved in the silencing of centromeric chromatin in *S. pombe*. Despite their apparent functional similarity as factors required for siRNA accumulation, RDE-3 and Cid12 do not appear to represent orthologous proteins. The *C. elegans* gene ZK858.1 has a much greater degree of sequence identity with Cid12 (Figure 3). It is conceivable that RDE-3 and Cid12/ZK858.1 provide similar functions in distinct branches of the RNAi pathway—for example, in post-transcriptional versus transcriptional silencing.

There are many possible models for how RDE-3 might function in RNAi. In one model, RDE-3 may respond to dsRNA in a manner analogous to OAS, producing oligoadenylate. However, rather than inducing a general nonspecific mRNA destruction, it may stimulate nucleases required for siRNA-directed mRNA turnover. Alternatively, RDE-3 could function indirectly to facilitate RNAi by insuring a proper balance of RNA metabolism in the cell. For example, RDE-3 could function as its homolog GLD-2 does in the polyadenylation of a subset of mRNAs in the cytoplasm. In the absence of RDE-3, defects in the polyadenylation of these mRNAs could lead to an accumulation of aberrant transcripts that enter the RNAi pathway and titrate limiting factors, reducing the ability of cells to initiate RNAi in response to foreign dsRNA (Figure 4A). Similarly, RDE-3 could function as Trf4p does in the turnover of improperly processed tRNAs or of other RNAs normally destined for the exosome. Defects in such a mechanism could, once again (as shown in Figure 4A), lead to inappropriate recognition of these aberrant RNAs and competition for limiting components of the RNAi machinery.

Another model worth considering (Figure 4B) is that RDE-3 is required for the amplification of RNAi in re-

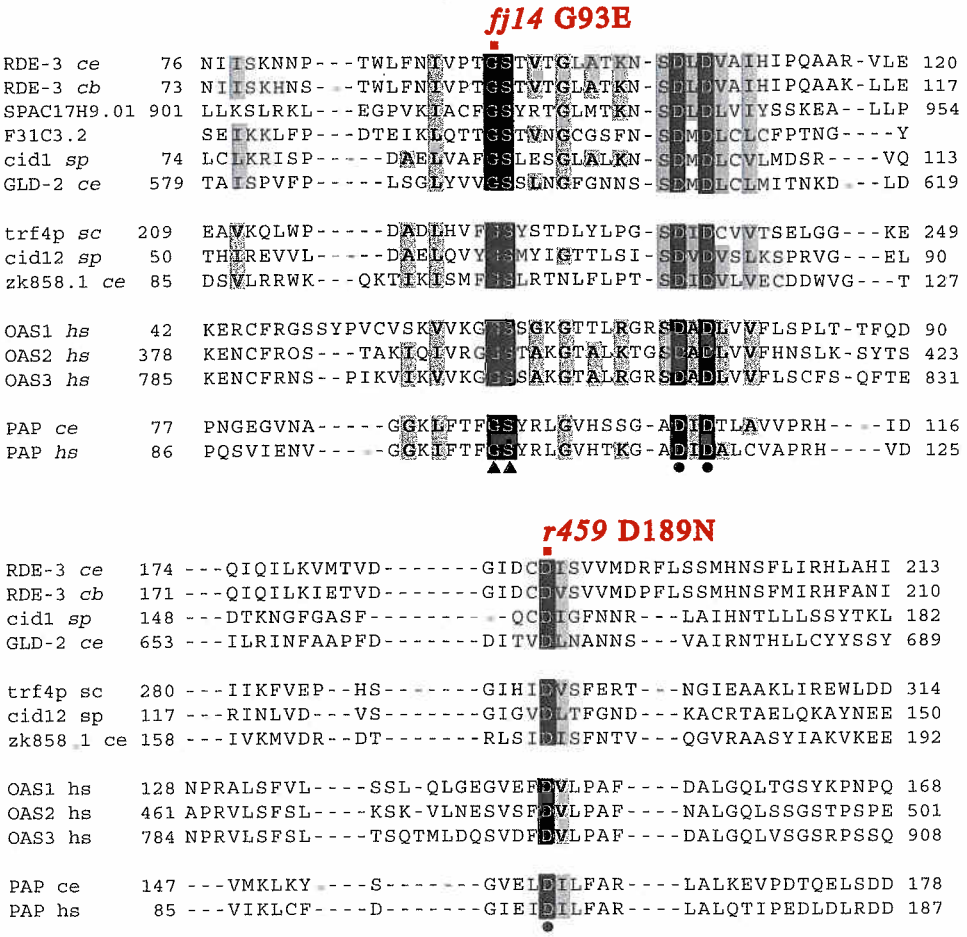


Figure 3. *rde-3* Mutants Alter Conserved Residues in the Nucleotidyltransferase Domain

Sequence alignment of RDE-3 with members of the polymerase β nucleotidyltransferase superfamily. Proteins are divided into subgroups on the basis of sequence identity. Residues identical in all proteins are highlighted in black, and conservative amino-acid substitutions are highlighted in gray. The conserved GS residues are marked by black triangles. The aspartic-acid residues of the predicted catalytic triad are indicated with black circles. The lesions in *fj14* and *r459* are indicated. The following abbreviations were used: *ce*, *Caenorhabditis elegans*; *cb*, *Caenorhabditis briggsae*; *hs*, *Homo sapiens*; *sc*, *Saccharomyces cerevisiae*; and *sp*, *Schizosaccharomyces pombe*. The sequences were aligned with Clustal W [19]. Some divergent sequences among subgroups were aligned manually.

sponse to low levels of the inducing trigger dsRNA. This model is consistent with our own observation that RDE-3 is not required for RNAi initiated by transgene-driven dsRNA (see above) and with the finding from *S. pombe* that the RDE-3 homolog Cid12 interacts with RdRP. In *C. elegans*, the detectable accumulation of siRNAs during RNAi requires RdRP activity (D. Conte and C.C.M., unpublished data; [17]) and is thought to involve an RdRP-dependent synthesis of new dsRNA after an initial round of target mRNA recognition. After ingestion, small quantities of dsRNA entering target tissues may be processed by Dicer to generate low abundance siRNAs. These primary siRNA could then direct a first round of target-mRNA cleavage. RDE-3 might then polyadenylate the nascent 3′ end of this cleavage product, stabilizing it and permitting its recognition by RdRP (Figure 4B). In the absence of RDE-3, the initial cleavage product may be too unstable to be recognized efficiently by RdRP, preventing amplification and the consequent accumulation of siRNA. Although their precise functions remain to be discovered, the findings that (1) RDE-3 and Cid12 are required for RNAi-related mechanisms, and (2) mutations in these factors result in similar defects in siRNA accumulation in these divergent species, strongly suggest that members of this protein family are likely to represent integral components of small-RNA-mediated silencing pathways.

Acknowledgments

We would like to thank Danesh Moazed for sharing data prior to publication. Some nematode strains were provided by the *Caenorhabditis* Genetics Center, funded by the National Institutes of Health National Center for Research Resources. This work has been funded in part by the National Institutes of Health (GM58800). M.J.S. is a Canadian Institutes of Health Research postdoctoral fellow, and C.C.M. is a Howard Hughes Medical Institute investigator.

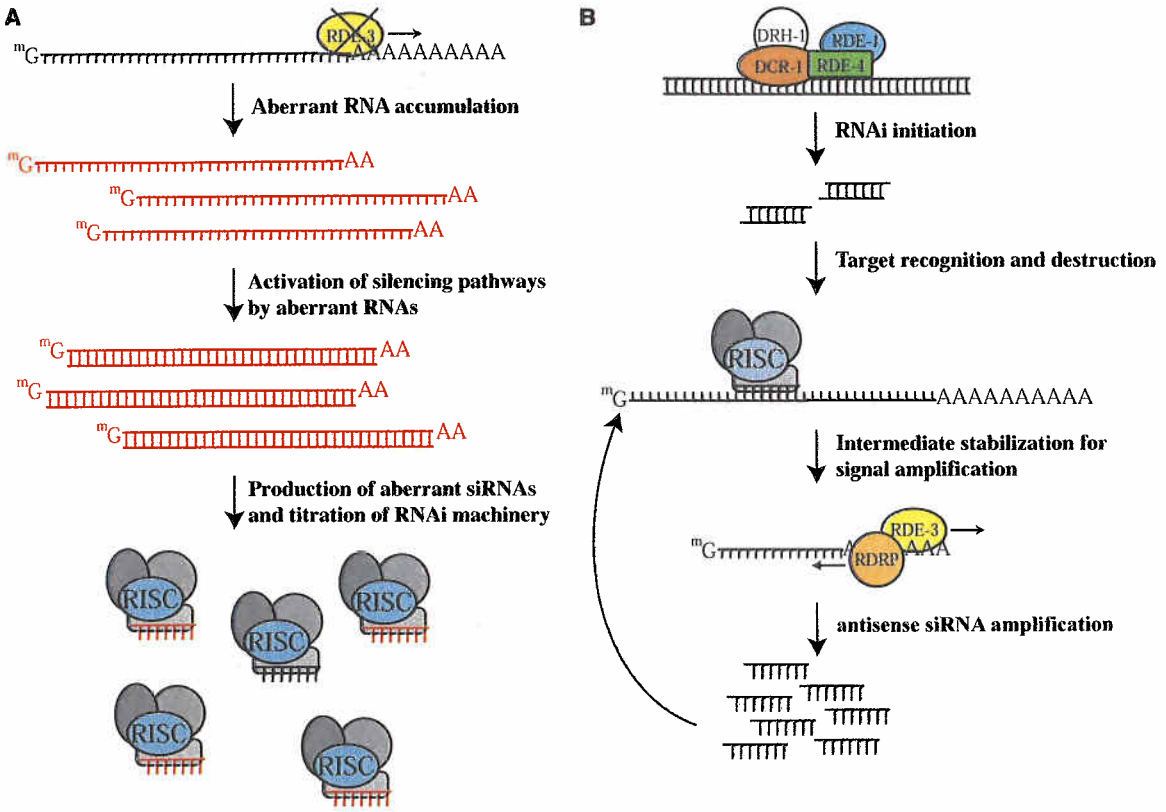


Figure 4. Models

Schematic diagrams illustrating possible roles for RDE-3: (A) as a factor that maintains the RNA metabolism balance required for efficient RNAi or (B) as a direct component of the RNAi pathway. (A) depicts RDE-3 as a poly(A) polymerase required for processing a subset of mRNAs in the cell. RDE-3 loss (indicated by the "X") could lead to accumulation of "aberrant" transcripts (red) with short poly(A) tails that can enter the RNAi pathway, titrating the RNA-induced silencing complex (RISC) or other limiting factors necessary for efficient RNAi. In (B), RNAi-initiating proteins, including Dicer (DCR-1), which processes long dsRNA into siRNAs, a Dicer-related helicase (DRH-1), an argonaute protein (RDE-1), and the dsRNA binding protein (RDE-4) [20], are shown processing foreign dsRNA into primary siRNA. These extremely-low-abundance primary siRNAs then target mRNA in an initial round of cleavage mediated by components of the RISC. RDE-3 is then proposed to polyadenylate the 5′ cleavage product, stabilizing this intermediate and allowing RdRP to amplify the response, generating abundant secondary siRNAs.

Received: December 14, 2004
Revised: January 4, 2005
Accepted: January 6, 2005
Published online: January 13, 2005

References

1. Mello, C.C., and Conte, D., Jr. (2004). Revealing the world of RNA interference. Nature *431*, 338–342.
2. Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343–349.
3. Baulcombe, D. (2004). RNA silencing in plants. Nature *431*, 356–363.
4. Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350–355.
5. Aravind, L., and Koonin, E.V. (1999). DNA polymerase beta-like nucleotidyltransferase superfamily: Identification of three new families, classification and evolutionary history. Nucleic Acids Res. *27*, 1609–1618.
6. Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. Cell *99*, 123–132.
7. Collins, J.J., and Anderson, P. (1994). The Tc5 family of transposable elements in *Caenorhabditis elegans*. Genetics *137*, 771–781.
8. Tijsterman, M., Ketting, R.F., Okihara, K.L., Sijen, T., and Plasterk, R.H. (2002). RNA helicase MUT-14-dependent gene silencing triggered in *C. elegans* by short antisense RNAs. Science *295*, 694–697.
9. Wang, L., Eckmann, C.R., Kadyk, L.C., Wickens, M., and Kimble, J. (2002). A regulatory cytoplasmic poly(A) polymerase in *Caenorhabditis elegans*. Nature *419*, 312–316.
10. Roberts, W.K., Hovanessian, A., Brown, R.E., Clemens, M.J., and Kerr, I.M. (1976). Interferon-mediated protein kinase and low-molecular-weight inhibitor of protein synthesis. Nature *264*, 477–480.
11. Hovanessian, A.G., Brown, R.E., and Kerr, I.M. (1977). Synthesis of low molecular weight inhibitor of protein synthesis with enzyme from interferon-treated cells. Nature *268*, 537–540.
12. Dong, B., and Silverman, R.H. (1995). 2-5A-dependent RNase molecules dimerize during activation by 2-5A. J. Biol. Chem. *270*, 4133–4137.
13. Kadaba, S., Krueger, A., Trice, T., Krecic, A.M., Hinnebusch, A.G., and Anderson, J. (2004). Nuclear surveillance and degradation of hypomodified initiator tRNAMet in *S. cerevisiae*. Genes Dev. *18*, 1227–1240.
14. Wang, Z., Castano, I.B., De Las Penas, A., Adams, C., and Christman, M.F. (2000). Pol kappa: A DNA polymerase required for sister chromatid cohesion. Science *289*, 774–779.
15. Hartmann, R., Justesen, J., Sarkar, S.N., Sen, G.C., and Yee, V.C. (2003). Crystal structure of the 2′-specific and double-stranded

RNA-activated interferon-induced antiviral protein 2′-5′-oligoadenylate synthetase. Mol. Cell *12*, 1173–1185.

16. Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789–802.
17. Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465–476.
18. Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell *106*, 23–34.
19. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. *22*, 4673–4680.
20. Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. Cell *109*, 861–871.

Current Biology, Vol. 15, 384–391, February 22, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.cub.2005.01.046

Vigilins Bind to Promiscuously A-to-I-Edited RNAs and Are Involved in the Formation of Heterochromatin

Qiaoqiao Wang, Zuo Zhang,[1] Katherine Blackwell, and Gordon G. Carmichael*
Department of Genetics and Developmental Biology
University of Connecticut Health Center
Farmington, Connecticut 06030

Summary

The fate of double-stranded RNA (dsRNA) in the cell depends on both its length and location [1]. The expression of dsRNA in the nucleus leads to several distinct consequences. First, the promiscuous deamination of adenosines to inosines by dsRNA-specific adenosine deaminase (ADAR) can lead to the nuclear retention of edited transcripts [2]. Second, dsRNAs might induce heterochromatic gene silencing through an RNAi-related mechanism [3–8]. Is RNA editing also connected to heterochromatin? We report that members of the conserved Vigilin class of proteins have a high affinity for inosine-containing RNAs. In agreement with other work [9–11], we find that these proteins localize to heterochromatin and that mutation or depletion of the *Drosophila* Vigilin, DDP1, leads to altered nuclear morphology and defects in heterochromatin and chromosome segregation. Furthermore, nuclear Vigilin is found in complexes containing not only the editing enzyme ADAR1 but also RNA helicase A and Ku86/70. In the presence of RNA, the Vigilin complex recruits the DNA-PKcs enzyme, which appears to phosphorylate a discrete set of targets, some or all of which are known to participate in chromatin silencing. These results are consistent with a mechanistic link between components of the DNA-repair machinery and RNA-mediated gene silencing.

Results and Discussion

Vigilin and DDP1 Bind Specifically to Inosine-Containing RNAs

We reported previously that nuclear I-RNAs can bind to a complex containing $p54^{nrb}$, PSF, and matrin 3 [2]. This complex prevents the export of promiscuously edited RNAs to the cytoplasm. Our initial identification of $p54^{nrb}$ and PSF as I-RNA binding proteins came from UV crosslinking studies using HeLa-cell nuclear extracts [2] (Figure 1A). Of special interest, however, is the additional band of about 160 kDa that we observed. Although this band is a minor one in crosslinking studies using HeLa-cell nuclear extracts, in similar studies using whole-cell extracts of *Drosophila* S2 cells, a more striking 160 kDa band specific for I-RNA is observed, along with a number of lower-molecular-weight bands (Figure 1A, lanes 3 and 4). Two of the smaller bands are *Drosophila* homologs of $p54^{nrb}$ and PSF (the NonA and NonA-like proteins) and will be described in more detail elsewhere.

We next used affinity chromatography to isolate HeLa-cell I-RNA binding proteins (Figure 1B). Lane 2 shows the proteins bound to an I-RNA affinity column where about 50% of the guanosines in an RNA transcript have been replaced with inosines. Lane 3 shows proteins that bound to a column containing the same RNA without inosines. We excised prominent bands from the gel and identified proteins by mass spectrometry. As expected, the complex of $p54^{nrb}$, PSF, and matrin 3 is retained on the column. The 160 kDa band shown here was not consistently observed before, as a result of proteolysis during the course of some experiments.

The 160 kDa band was identified as Vigilin by mass spectrometry. Vigilin is a ubiquitous and abundant RNA binding protein that may play a role in mRNA stability [12–15]. Vigilins are highly conserved (see Table S1 in the Supplemental Data available with this article online), and all members of this family contain 14 tandem hnRNP K-homology (KH) domains (Figure 1D) that could be involved in protein-nucleic acid and protein-protein interactions.

In parallel experiments using *Drosophila* S2 extracts, a prominent 160 kDa protein was also specifically retained by I-RNA (Figure 1C). The *Drosophila* 160 kDa protein was identified by mass spectrometry as the dodeca-satellite binding protein 1 (DDP1), which is the *Drosophila* homolog of Vigilin. Recombinant Vigilin and DDP1 specifically bind to I-RNA and exist in molecular complexes distinct from the other I-RNA binding factors in HeLa cells or S2 cells (Figure S1). DDP1 was originally identified through its ability to bind specifically to one strand of the centromeric dodecasatellite DNA sequence [10]. Thus, this protein binds to both DNA and RNA. In *Drosophila*, DDP1 associates with pericentric heterochromatin and further colocalizes almost everywhere on polytene chromosomes with the heterochromatin-associated protein HP1 [10]. These results suggested a role for DDP1 in both chromosome segregation and heterochromatin formation [10, 16], and this role has recently been experimentally confirmed [11].

Vigilin Is Associated with Heterochromatin in Mammalian Cells

In COS7 cells, Vigilin exhibits a dual localization to both the cytoplasm and the nucleus (Figure S2). In agreement with results reported by others [9], we found that most of the protein is cytoplasmic and associated with ER and that nuclear Vigilin does not appear to reside in discrete, prominent sites but rather is found generally associated with regions of higher DNA content. Like Vigilin, DDP1 is abundant in the cytoplasm of *Drosophila* S2 cells, where it also appears to decorate the ER. However, within the nucleus DDP1 is concentrated in the region of the chromocenter, which appears as a prominent area of DAPI staining (Figure S2). In this work, we have confirmed the reported association of DDP1 with heterochromatin [10, 11] by the observation that DDP1 and HP1 almost completely colocalize.

*Correspondence: carmichael@nso2.uchc.edu
[1]Present address: Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724.

5097-1