THE FATE OF EDITED RNA IN *CAENORHABDITIS ELEGANS*

by

Sabine Hellwig

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

The University of Utah

August 2007

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a dissertation submitted by

Sabine Hellwig

This dissertation has been read by each member of the following supervisory committee
and by majority vote has been found to be satisfactory.

21 June 2007

6-21-07
_____
Dana Carroll

6·21·07
_____
Katharine S. Ullman

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the dissertation of _____ Sabine Hellwig _____ in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

Date   L. Bass
Chair: Supervisory Committee

Approved for the Major Department

Dana Carroll
Chair/Dean

Approved for the Graduate Council

ABSTRACT

In contrast to DNA, which functions to store genetic information, RNA
effectuates genetic information. The versatility of RNA is evident in the multiple
cellular functions RNA fulfills. RNA molecules are structural elements, catalysts,
adaptors, messengers and regulators. Accordingly, RNA shows great diversity in
nucleotide composition and structure. This dissertation explores how sequence and
structural features of RNA influence the fate of RNA in nematode cells and impact the
entire organism.

First, the *C. elegans* noncoding transcript *rncs-1* is characterized. Throughout
development, *rncs-1* is expressed in intestine and hypodermis. Upon starvation,
transcription of *rncs-1* is upregulated by an ELT-2-dependent mechanism. A feature of
the *rncs-1* transcript is its extended double-stranded structure. Double-stranded RNA is
a substrate for the RNase III enzyme Dicer, which cleaves long RNA duplexes into
~23-nucleotide products that subsequently silence cognate messages. Surprisingly,
rncs-1 RNA is resistant to Dicer cleavage *in vitro*, and additional studies demonstrate
that secondary structural elements in *rncs-1* occlude access of Dicer to the double-
stranded region. While not a cleavage substrate for Dicer *in vitro*, *rncs-1* inhibits Dicer
activity. Significantly, overexpression of *rncs-1* modulates expression of endogenous
Dicer silencing targets *in vivo*. This result represents a previously uncharacterized
cellular function of noncoding RNAs.

Second, the effect of RNA editing on the subcellular localization of RNA is examined. Deamination of adenosine to inosine within double-stranded RNA is catalyzed by adenosine deaminases that act on RNA (ADARs). The majority of cellular inosine is found in noncoding sequences, but the function of inosine in noncoding regions of RNA is unknown. In contradiction to existing models for nuclear retention of ADAR-edited RNA, research presented here confirms that endogenous editing substrates of *C. elegans* are exported to the cytoplasm.

Finally, in continuation of previous behavioral studies that revealed chemotaxis defects of *C. elegans* ADAR-mutants, the response of editing-deficient worms to dauer-inducing pheromone is assayed. Dauer pheromone controls the entry of *C. elegans* into a specialized, highly resistant alternative larval stage called dauer. Nematodes lacking ADARs are found to show wildtype dauer formation, indicating that ADAR activity is dispensable for response to dauer pheromone.

To my mother,
Erika Hellwig

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

CHAPTER 1

INTRODUCTION

The paradox of complexity and genomic protein-coding content

Biological complexity can be defined as a combination of metabolic versatility and the number of cells and different cell types of an organism. During the course of evolution, protein families have expanded, contributing to cell differentiation and environmental adaptation in higher organisms (1). However, paradoxical to the differences in complexity, the absolute number of protein-coding genes is surprisingly constant among eukaryotes, varying only ~3-fold between yeast (~6,500 genes), *C. elegans* (~20,000 genes), zebrafish (~19,000 genes) and human (~22,000 genes) (2).

To achieve complexity, organisms employ sophisticated mechanisms that control gene expression and diversify gene products (3,4). *Cis*-acting elements in the genome modulate chromatin architecture and regulate transcription of a gene. At the mRNA level, alternative splicing events and RNA editing alter coding content of the transcript, and regulatory elements in untranslated regions control localization, stability and translation of messages. Posttranslational modifications, dictated by amino acid sequence, diversify function at the protein level. In addition, higher organisms draw on regulatory noncoding RNAs (ncRNAs) to modulate expression and function of proteins *in cis* and *in trans*.

The work described in this dissertation focused on two aspects of diversifying gene regulation at the RNA level in the nematode *Caenorhabditis elegans*: regulatory ncRNA and RNA editing. We characterize a noncoding transcript and demonstrate its potential to control gene expression through modulation of the RNA interference pathway. Further, we study subcellular localization of cellular RNAs that are edited by adenosine deaminases that act on RNA (ADARs). Lastly, we explore the physiological consequences of ADAR editing by assaying behavior of wildtype and ADAR-mutant *C. elegans*.

Regulatory noncoding RNAs

Comparison of genome sizes in lower and higher eukaryotic organisms revealed that genomes have expanded in the course of evolution, ranging from 12Mb in yeast, ~34Mb in amoebae, ~100Mb in nematodes to ~3000Mb in humans (2,4). In contrast, the total size of translated sequences varies less than 2-fold between amoebae (21Mb) and humans (32Mb). Thus, while ~70% of the yeast genome is transcribed and translated, this percentage is reduced to ~25% in *C. elegans* and <1.5% in humans (5).

The majority of noncoding genome content is transcribed (60-70% in humans, (6)), and much of it is transcribed from both strands. In humans an estimated 22-25% of transcription clusters form sense-antisense pairs, and many of the antisense transcripts are ncRNAs (7). It is intriguing to consider that the human transcriptome could potentially be larger than the genome. Large scale sequencing of cDNAs and genome tiling approaches have identified tens of thousands of transcripts that do not map to coding regions of the mammalian genome (6,8-11). We have only begun to assign functions to noncoding transcripts, and examples are described in the following paragraphs. The

myriad of mechanisms by which ncRNAs can influence expression and function of proteins and other ncRNAs suggests that much, if not all, of the noncoding transcriptional output of a genome may be meaningful.

The genomic organization of ncRNAs is intricate. Some ncRNAs are produced from well-defined genes, and some are posttranscriptionally processed by splicing and polyadenylation. Other noncoding transcripts map to intronic regions, and still others originate from intergenic loci. Many ncRNAs fulfill housekeeping functions. Among those are ribosomal, transfer, splicesomal and telomerase RNAs, as well as the ncRNA 7SL, a component of the signal recognition particle. These ncRNAs are often constitutively expressed and are essential for viability of the cell.

In addition to these infrastructural RNAs, many regulatory ncRNAs have been described. Much research has been dedicated to identify and characterize micro-RNAs (miRNAs), a family of 21-25-nucleotide RNAs that regulate gene expression at the post-transcriptional level (12). The focus of the work presented in this dissertation will be the function of large ncRNAs, sometimes referred to as macro-RNAs. In contrast to miRNAs, large ncRNAs are evolving rapidly and are poorly conserved even in closely related organisms (13). The regulatory function of large ncRNAs is carried out through various mechanisms that modulate gene expression at a transcriptional or post-transcriptional level (14-17), and prominent examples are summarized in Table 1.1.

The noncoding RNA *Xist* is essential for dosage compensation in mammals, where transcription from a random X is silenced in XX females (18). Through long-range *cis*-action, *Xist* initiates silencing, most likely by recruitment of chromatin remodeling factors that establish irreversible transcriptional repression. By a similar

Table 1.1. **Functions of selected large noncoding RNAs**

RNA	Organism	Size	Function and mechanism
Transcriptional repression by sense ncRNA genes			
Xist (X-inactive-specific transcript	Mammals	15-19kb	Dosage compensation by transcriptional silencing of a single X chromosome by coating of inactive X and recruitment of silencing factors
roX1, roX2 (RNA on the X)	Drosophila	3.7kb, 0.5-1.2kb	Dosage compensation by hyperactivation of the single X in males by exclusive expression from the male X and binding of the male X as part of an RNA-protein complex
H19	Mammals	2.3kb	Imprinting of the *Igf2* (insulin-like growth factor) locus by competition with *Igf2* for engagement of a shared enhancer element
Transcriptional repression by antisense ncRNAs			
Tsix	Mammals	>40kb	Antisense to Xist, negatively regulates Xist expression in undifferentiated embryonic stem cells
Air (Antisense Igf2r RNA)	Mammals	108kb	Imprinting of the *Igf2r* (insulin-like growth factor type 2 receptor) locus by promoter occlusion or RNAi degradation of *Igf2r* followed by recruitment of chromatin silencing factors
DISC2 (disrupted in schizophrenia 2)	Human		Negative regulation of *DISC1*
UBE3A-ATS (ube3a antisense)	Human	450kb	Imprinting of the PWS/AS (Prader-Willi/Angelman syndrome) locus
Transcriptional regulation by intergenic ncRNA			
SGR1	Yeast		Serine-induced transcription represses expression of adjacent SER3 gene by transcriptional interference
Xite (X inactivation intergenic transcription element)	Mammals		Enhancer of *Tsix* expression in dosage compensation
iab transcripts (infra-abdominal)	Drosophila		Regulation at the *Bithorax* locus
Post-transcriptional regulation by ncRNAs			
SRA (steroid receptor RNA activator)	Human	700nt	Co-activator of several steroid hormone receptors, required for transcriptional activation by nuclear hormone receptors
HSR1 (heat-shock RNA1)	Human	600nt	Assists in mobilization of HSF1 during heat stress
NRON (noncoding repressor of NFAT)	Mammals	2-4kb	Negative regulator of transcription factor NFAT by preventing nuclear localization
7SK	Mammals	331nt	General repression of transcription elongation by association with elongation factor P-TEFb
B2	Mouse	178nt	General repression of transcription by binding to the core of RNA Polymerase II in response to heat-shock

mechanism, which includes coating of an X chromosome, the ncRNAs *roX1* and *roX2* mediate dosage compensation in *Drosophila*, but with an opposite outcome, the upregulation of gene expression from the male X (19).

Some natural antisense transcripts are involved in transcriptional repression of their sense counterparts. The mechanisms by which this is accomplished are largely unclear, but may involve transcriptional interference or formation of dsRNA, leading to masking of regulatory elements in the sense transcripts, RNA editing or degradation by the RNA interference (RNAi) pathway. The ~108kb antisense noncoding *Air* transcript is essential for genomic imprinting of the *Igf2r* locus (20). *Air* partially overlaps with the *Igf2r* transcript, and presumably initiates silencing by promoter occlusion or through RNAi. The silenced state is subsequently maintained and spread along the genomic locus by recruitment of chromatin modifying proteins. In the human nervous system, antisense *DISC2* modulates expression of its sense counterpart *DISC1*, which is essential for cell polarity and neuronal migration, and disruption of the *DISC2* locus is implicated in schizophrenia and bipolar disorder (21).

Transcription of ncRNA from intergenic loci has also been linked to transcriptional regulation. However, the negative effect appears to be largely dependent on transcriptional activity per se, but not on the ncRNA product. In yeast, serine-dependent transcription of the intergenic *SRG1* ncRNA represses expression of the adjacent *SER3*, a serine biosynthetic gene (22). In the case of *SRG1*, transcriptional interference appears to be the mechanism of negative regulation. Alternatively, intergenic transcription could promote expression from nearby loci by establishing an active chromatin state through the action of RNA Polymerase-II-associated histone-modifying

enzymes. This mechanism has been suggested for the *Xite* ncRNA, whose transcription positively regulates expression of *Tsix*.

Multiple ncRNAs regulate RNA Polymerase II (Pol II) transcription at the posttranscriptional level (23). Some ncRNAs, such as 7SK and B2, have a general negative effect on gene expression by binding canonical components of the Pol II complex (24-26). In case of SRA, the ncRNA is an essential coactivator of several steroid hormone receptors (27), and the *HSR-1* transcript, in complex with translation elongation factor eEF1A, promotes trimerization and activation of HSP-1 (28). The research described in Chapter 2 suggests yet another mechanism for modulation of gene expression by noncoding RNA that involves competitive inhibtion of the RNAi silencing pathway.

Changing RNA sequence and structure by adenosine deamination

Genetic information can be diversified by co- or posttranscriptional alteration of RNA nucleotide sequence, a phenomenon known as RNA editing. RNA editing occurs by nucleotide insertion or deletion or by enzymatic modification of bases. For the latter, two types of enzymatic activities have been characterized. Cytidine deaminases catalyze the conversion of cytidine (C) to uridine (U), while ADARs deaminate adenosine (A) to inosine (I).

The model enzyme of C-to-U deaminases is the mammalian APOBEC-1 (apoB editing catalytic subunit 1) (29). As part of a multi-subunit editosome, it catalyzes editing of apolipoprotein B mRNA involved in lipid metabolism. Other members of the cytidine deaminase family play roles in innate and adaptive immunity (30). Activation-induced deaminase is required for class switch recombination and somatic hypermutation

during maturation of B lymphocytes, resulting in production of diversified antibodies. Further, APOBEC-3G mediates cellular antiviral defense by altering viral nucleotide sequences during the reverse transcription phase of retroviral infection.

The more common form of RNA editing by base modification is A-to-I editing by ADARs (31,32). ADAR activity was discovered in *Xenopus* embryos (33,34), but members of the ADAR family were subsequently found to be widespread among metazoans (35). ADARs share a common domain structure that includes a variable number of double-stranded RNA binding motifs (dsRBMs) at the amino terminus and a conserved catalytic domain at the carboxyl terminus. *C. elegans*, the focus of studies presented in this work, encodes two ADAR family members (36,37). The *adr-1* gene gives rise to several splice variants and encodes a protein with two conserved dsRBMs and a catalytic domain that lacks several key residues. It is unlikely that ADR-1 has deamination activity by itself. The second nematode ADAR, ADR-2, contains a single dsRBM and a catalytically active deaminase domain.

Deamination occurs at the C6 position of adenine through nucleophilic attack of water (38). ADARs are metalloenzymes that require zinc for activity. In addition, the recent solution of the crystal structure of the human ADAR2 catalytic domain revealed the presence of a molecule of inositol hexakisphosphate, a cofactor essential for activity (39). In contrast to APOBEC-1, ADARs do not require accessory protein factors for catalysis.

Because the first known ADAR substrates were components of neurotransmitter and ion channels that function in the nervous system, A-to-I editing is often thought of as a neuronal enzymatic activity. The expression patterns of *Drosophila* ADAR and *C.*

elegans ADR-1 support a function in the nervous system (37,40), but ADARs have been cloned and purified from a variety of tissues, including bovine liver and thymus (41,42), *Xenopus* eggs (43) and chicken lungs (44). In Chapter 2 of this work, expression of a *C. elegans* ADAR substrate is shown to be largely intestinal and hypodermal. As this substrate is heavily edited, we can deduce that ADARs are also active in these tissues of the nematode. Within the cell, ADARs typically localize to the nucleus. However, a splice variant of human ADAR1 contains nuclear export as well as import signals and can shuttle between the nucleus and cytoplasm (45-47).

ADARs exclusively deaminate A residues in double-stranded RNA (dsRNA). The specificity of ADARs is largely dictated by the structure of the RNA substrate, but also by the sequence context of adenosines within the substrate. Long, perfectly duplexed RNAs are promiscuously deaminated at 50-60% of A residues, while short duplexes that contain mismatches, loops and bulges are selectively edited at only a few sites. *In vitro* studies and observations made for *in vivo* substrates, indicate that ADARs have preferences for selection of A residues within a dsRNA. Position relative to the helix terminus (48) and the 5' and 3' nearest neighbors (49,50) are deciding factors, but also the greater sequence context, including the nucleotide opposite a potential editing site, is important (51-54).

For RNA substrates, editing by ADARs has two molecular consequences. First, the primary nucleotide sequence is altered. Most cellular enzymes, including the translation and splicing machineries, interpret inosine as guanosine, as it preferentially base-pairs with cytidine. Second, the secondary structure of the substrate is changed, since deamination changes AU base-pairs to IU mismatches. Thus, extensive editing of

RNA can shift the conformational equilibrium from a double-stranded to a largely single-stranded state. As ADARs do not bind single-stranded RNA (33,55), the model is that progressive editing of a substrate lowers affinity of the enzyme and eventually results in termination of the reaction.

In vivo, ADAR editing can occur whenever RNA is sufficiently base-paired. Such double-stranded regions can include exons, introns, untranslated regions or noncoding transcripts. Prominent examples for editing events in exons include the mRNAs of mammalian glutamate and serotonin receptors (56-59), *Drosophila* calcium, sodium, potassium and chloride channels (60-65) and squid potassium channels (40). So far, all reported ADAR editing in coding regions is facilitated by base-pairing of exons with exon complementary sequences in introns. However, recent bioinformatic analyses of transcriptomes revealed that antisense transcription, occurring from the opposite strand of a protein-coding gene, is common (7,66-68). Thus, formation of sense-antisense pairs could potentially promote ADAR editing, as has been described for Xenopus basic fibroblast growth factor mRNA (69) . The result of A-to-I editing in coding regions is the alteration of codons. For example, ADAR editing of the glutamate receptor subunits gluR-B, gluR-5 and gluR-6 changes a CAG codon, encoding glutamine, to CIG, which is interpreted as CGG and translated to arginine by the ribosome.

Adenosine deamination events in introns have the potential to remove branch sites, create 5' donor sites and create or destroy 3' acceptor sites; however, direct alteration of splicing events by ADARs has been demonstrated for only a few examples. In mammalian ADAR2 pre-mRNA, creation of an alternative 3' splice site by the ADAR2 enzyme results in production of a truncated, nonfunctional protein, suggesting a

possible autoregulatory mechanism (70). In addition, a splicing branch site is edited in human tyrosine phosphatase SHP-1 (71).

The first known ADAR substrates were serendipitously discovered on account of discrepancies between genomic and mRNA sequences. More recently, large-scale biochemical screens (72,73) and bioinformatic approaches (61,74-77) have identified tens of thousands of ADAR editing sites in various transcriptomes. A common observation of these studies was that inosine is most prevalent in noncoding regions of transcripts, and that editing in exons may, in fact, be an exception to the rule. The most abundant source of noncoding dsRNA that is edited by ADARs are repetitive elements that are interspersed throughout the genome (78). Approximately half of the human genome is comprised of repetitive elements, with primate-specific *Alu* elements representing the most prevalent type of repeat (79). The great majority of human ADAR editing sites maps to *Alu* sequences in 3'UTR, 5'UTRs and introns.

The consequences of adenosine deamination in coding regions are obvious, as alteration of codons can lead to synthesis of different protein isoforms. In contrast, while noncoding regions of messenger RNAs and noncoding transcripts contain the bulk of cellular inosine, the function of noncoding inosine largely remains a mystery. Suggested functions include regulation of translation, stability (80,81) or nuclear localization (82) of edited transcripts, but none of these possibilities has been confirmed for endogenous ADAR substrates.

Of particular interest is the prospect that ADAR editing can interfere with other dsRNA-mediated cellular pathways, notably the RNA interference (RNAi) pathway (83,84). RNAi is initiated by cleavage of a long dsRNA trigger into 19-23bp small

interfering RNAs (siRNAs) by the RNase III enzyme Dicer. siRNAs subsequently serve as guides for gene silencing by message degradation and possibly by chromatin-based mechanisms (85). The first indication that ADAR activity antagonizes RNAi came from studies in our laboratory, which indicated that A-to-I editing of transgene-derived dsRNA prevented silencing of transgenes in the *C. elegans* soma (86). Moreover, the chemotactic defect of *C. elegans* ADAR mutants was rescued in RNAi defective backgrounds, substantiating a link between the editing and interference pathways in the nematode (87).

Two possible models for the antagonistic effect of ADARs on RNAi exist, and both may be realized in the cell. First, because dsRNA binding proteins (dsRBPs) are sequence-nonspecific (88), ADARs can compete with Dicer or other dsRBPs of the interference pathway for the binding of dsRNA substrates. In human cell culture, the cytoplasmic splice variant of ADAR1 was shown to sequester transfected siRNAs, and cells lacking ADAR1 were more susceptible to RNAi (88). The second model for inhibition of RNAi by ADARs suggests that promiscuous deamination of dsRNA increases the single-stranded character of the substrate. Such destabilized duplexes are less likely to be good binding partners for Dicer. *In vitro* studies in *Drosophila* extracts showed that editing reduces production of siRNAs from dsRNA and diminishes degradation of the cognate message (88). However, complete inhibition of siRNA formation required extensive editing of ~50% of A residues, while RNAi occured, albeit less efficiently, at lower editing levels.

<u>*C. elegans* ADAR substrates</u>

In C. elegans, all currently known *in vivo* ADAR substrates were identified in biochemical screens of polyadenylated RNA samples performed in our laboratory (72,73). The screening strategy employed a protocol for specific chemical cleavage of RNA 3' of inosine residues by T1 ribonuclease (89), followed by reverse transcription and differential display of the cleavage products. T1-dependent products were cloned and sequenced and subsequently mapped to genomic sequences. A total of 10 nematode ADAR substrates were identified in 2 independent screens. Table 1.2 summarizes our current knowledge of *C. elegans* editing substrates.

Inverted repeats are common in *C. elegans*, comprising ~3-6% of the 100Mb genome (90,91). Similar to the frequent editing observed in human repetitive elements, half of the identified *C. elegans* ADAR substrates also contained inverted repeat elements, while the remaining substrates contained intramolecular dsRNA regions that were not repeat-derived. With the exception of *pop-1* and *unc-64*, whose gene products are essential for embryonic development and normal locomotion, respectively, little is known about expression or function of editing substrates.

Endogenous ADAR substrates in *C. elegans* provide tools to study the effects of A-to-I editing in noncoding regions on the cellular fate of RNAs. Of particular interest to us are the substrates 36A and *rncs-1*, which are the most heavily edited nematode transcripts, with 20-25% of adenosines deaminated to inosine (Figure 1.1). A major caveat in investigating the function of inosine in these two transcripts was that their cellular fate and function in the wildtype nematode was unknown at the outset of our study. We address this issue in two chapters of this work. In Chapter 2, we characterize

Table 1.2. **Cellular ADAR substrates in** *C. elegans*

Substrate name[a]	Gene name	Editing	Repeat	Anatomic expression	Function or inferred function[b]
1TC	W10C8.2 *(pop-1)*	3'UTR	--	Hypodermal seam cells, gonadal precursors, developing vulva	Transcription factor, component of the Wnt signaling pathway
9A	ZC239.6	3'UTR	IR-3	Unknown	Polymerase delta-interacting protein related (inferred), potassium channel tetramerization (inferred)
16G	Y6D11A.1	3'UTR	--	Unknown	3'-5' exoribonuclease (inferred)
Syntaxin	F56A8.7a *(unc-64)*	3'UTR/ intron	IR-5	Nervous system, secretory cells in the vulva, excretory gland cells	Syntaxin, plasma membrane receptor for intracellular vesicles
Laminin-γ	C54D1.5 *(lam-2)*	3'UTR	--	Unknown	Extracellular matrix glycoprotein
SSADH	F45H10.1 *(alh-7)*	3'UTR	--	Intestine, renal glands, hypodermis	Succinic semialdehyde dehydrogenase
12A	D2024.3 *(elo-3)*	3'UTR	IR-2	Pharynx, body wall muscle, hypodermis	Polyunsaturated fatty acid elongase
36A	C35E7.6	3'UTR	IR-4	Unknown	Unknown
	M05B5.3	5'UTR	IR-5	Unknown	Unknown
52G	F55A9.2 *(rncs-1)*	Non-coding	--	Intestine, hypodermis[b]	Putative stress signal, modulation of Dicer activity[c]

[a]   Substrate names are listed according to references 72 and 73.

[b]   Functions are assigned based on orthology group assignments and domain conservation as listed by *Wormbase.*

[c]   Expression and function of *rncs-1* are characterized in Chapter 2.

Figure 1.1. **The *C. elegans* ADAR substrates 36A and *rncs-1*.** Diagrams show the extended double-stranded structures of the substrates 36A (A) and *rncs-1* (B). The 36A substrate is found in the 3'UTR of the C35E7.6 gene. The structures are drawn to scale, and scale bars are shown. Dots indicate single-nucleotide mismatches and blue asterisks represent GU mismatches. Colored vertical bars indicate the position of edited adenosine residues. The color-coding and bar height is indicative of the frequency of editing at a particular site (red: > 70%, blue: 40-70%, green: <40%). Editing levels were inferred from references 72 and 73 and/or determined in experiments described in subsequent chapters.

A ORF (1203bp)

A_n

0 20 40 60 80 100 nt

B

A_n

0   20   40   60   80   100 nt
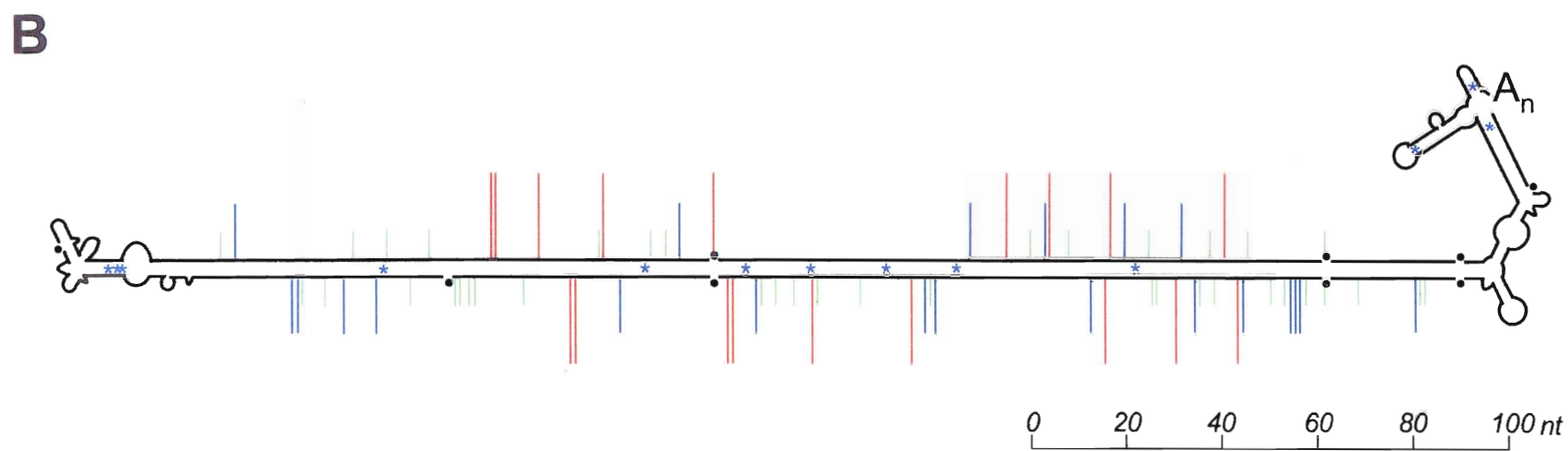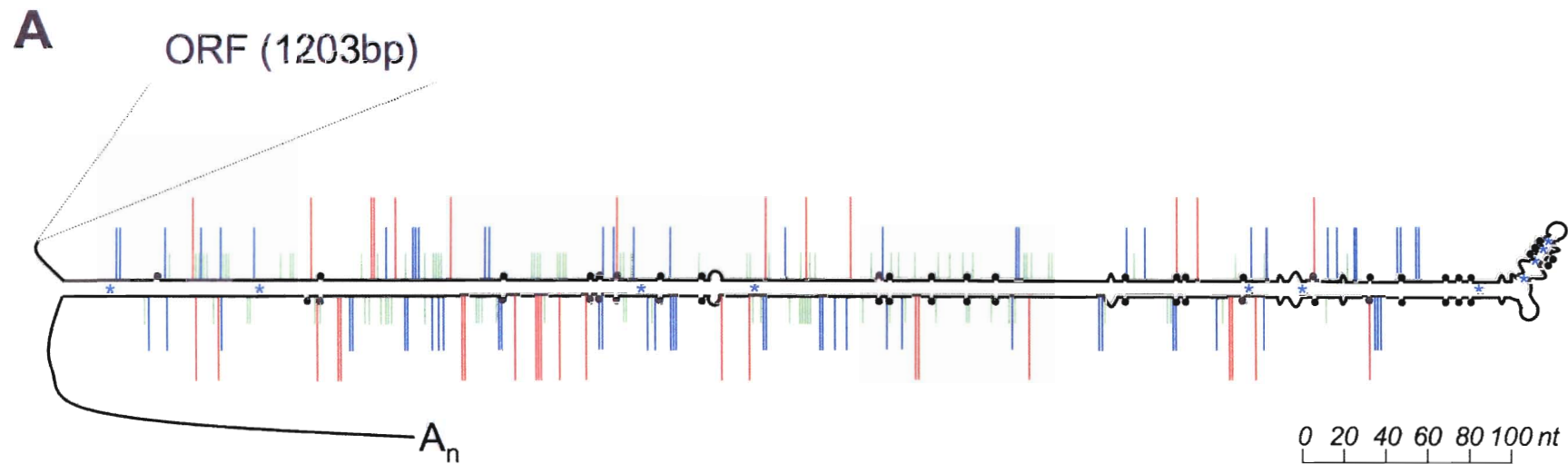
expression and transcriptional regulation of the ADAR substrate *rncs-1* in detail. We demonstrate its upregulation in response to stress and show its capability to modulate the RNAi pathway by competitive inhibition of Dicer *in vitro* and *in vivo*. In Chapter 3, we use a cell fractionation approach to determine the subcellular localization of 36A and *rncs-1*. We find that 36A, and likely *rncs-1*, are exported to the cytoplasm in spite of hyperediting by ADARs. In Chapter 4, we discuss the consequences of ADAR editing for *C. elegans* behavior. We show that worms lacking ADAR activity, while deficient in chemotaxis (37), have a wildtype response to dauer-inducing pheromone. Suggestions for future research are considered in Chapter 5.

References

1. Vogel, C. and Chothia, C. (2006) Protein family expansions and biological complexity. *PLoS Comput Biol*, **2**, e48.

2. Hubbard, T.J., Aken, B.L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T. *et al.* (2007) Ensembl 2007. *Nucleic Acids Res*, **35**, D610-617.

3. Hahn, M.W. and Wray, G.A. (2002) The g-value paradox. *Evol Dev*, **4**, 73-75.

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

5. Huttenhofer, A., Schattner, P. and Polacek, N. (2005) Noncoding RNAs: hope or hype? *Trends Genet*, **21**, 289-297.

6. Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C. *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559-1563.

7. Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z. and Rowley, J.D. (2004) Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res*, **32**, 4812-4820.

8. Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X., Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S. *et al.* (2004) Global

identification of human transcribed sequences with genome tiling arrays. *Science*, **306**, 2242-2246.

9. Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., Long, J., Stern, D., Tammana, H., Helt, G. *et al.* (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science*, **308**, 1149-1154.

10. Jongeneel, C.V., Delorenzi, M., Iseli, C., Zhou, D., Haudenschild, C.D., Khrebtukova, I., Kuznetsov, D., Stevenson, B.J., Strausberg, R.L., Simpson, A.J. *et al.* (2005) An atlas of human gene expression from massively parallel signature sequencing (MPSS). *Genome Res*, **15**, 1007-1014.

11. Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermueller, J., Hofacker, I.L. *et al.* (2007) RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science*.

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

13. Pang, K.C., Frith, M.C. and Mattick, J.S. (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet*, **22**, 1-5.

14. Mattick, J.S. and Makunin, I.V. (2006) Noncoding RNA. *Hum Mol Genet*, **15 Spec No 1**, R17-29.

15. Mehler, M.F. and Mattick, J.S. (2006) Noncoding RNAs in the nervous system. *J Physiol*, **575**, 333-341.

16. Morey, C. and Avner, P. (2004) Employment opportunities for noncoding RNAs. *FEBS Lett*, **567**, 27-34.

17. Prasanth, K.V. and Spector, D.L. (2007) Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev*, **21**, 11-42.

18. Ng, K., Pullirsch, D., Leeb, M. and Wutz, A. (2007) Xist and the order of silencing. *EMBO Rep*, **8**, 34-39.

19. Kelley, R.L. (2004) Path to equality strewn with roX. *Dev Biol*, **269**, 18-25.

20. Sleutels, F., Zwart, R. and Barlow, D.P. (2002) The noncoding Air RNA is required for silencing autosomal imprinted genes. *Nature*, **415**, 810-813.

21.   Millar, J.K., James, R., Brandon, N.J. and Thomson, P.A. (2004) DISC1 and DISC2: discovering and dissecting molecular mechanisms underlying psychiatric illness. *Ann Med*, **36**, 367-378.

22.   Martens, J.A., Wu, P.Y. and Winston, F. (2005) Regulation of an intergenic transcript controls adjacent gene transcription in Saccharomyces cerevisiae. *Genes Dev*, **19**, 2695-2704.

23.   Goodrich, J.A. and Kugel, J.F. (2006) Noncoding-RNA regulators of RNA polymerase II transcription. *Nat Rev Mol Cell Biol*, **7**, 612-616.

24.   Allen, T.A., Von Kaenel, S., Goodrich, J.A. and Kugel, J.F. (2004) The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nat Struct Mol Biol*, **11**, 816-821.

25.   Nguyen, V.T., Kiss, T., Michels, A.A. and Bensaude, O. (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*, **414**, 322-325.

26.   Yang, Z., Zhu, Q., Luo, K. and Zhou, Q. (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature*, **414**, 317-322.

27.   Zhao, X., Patton, J.R., Davis, S.L., Florence, B., Ames, S.J. and Spanjaard, R.A. (2004) Regulation of nuclear receptor activity by a pseudouridine synthase through posttranscriptional modification of steroid receptor RNA activator. *Mol Cell*, **15**, 549-558.

28.   Shamovsky, I., Ivannikov, M., Kandel, E.S., Gershon, D. and Nudler, E. (2006) RNA-mediated response to heat shock in mammalian cells. *Nature*, **440**, 556-560.

29.   Wedekind, J.E., Dance, G.S., Sowden, M.P. and Smith, H.C. (2003) Messenger RNA editing in mammals: new members of the APOBEC family seeking roles in the family business. *Trends Genet*, **19**, 207-216.

30.   Turelli, P. and Trono, D. (2005) Editing at the crossroad of innate and adaptive immunity. *Science*, **307**, 1061-1065.

31.   Bass, B.L. (2002) RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem*, **71**, 817-846.

32.   Valente, L. and Nishikura, K. (2005) ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. *Prog Nucleic Acid Res Mol Biol*, **79**, 299-338.

33.   Bass, B.L. and Weintraub, H. (1987) A developmentally regulated activity that unwinds RNA duplexes. *Cell*, **48**, 607-613.

34.   Rebagliati, M.R. and Melton, D.A. (1987) Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity. *Cell*, **48**, 599-605.

35.   Keegan, L.P., Leroy, A., Sproul, D. and O'Connell, M.A. (2004) Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. *Genome Biol*, **5**, 209.

36.   Hough, R.F., Lingam, A.T. and Bass, B.L. (1999) Caenorhabditis elegans mRNAs that encode a protein similar to ADARs derive from an operon containing six genes. *Nucleic Acids Res*, **27**, 3424-3432.

37.   Tonkin, L.A., Saccomanno, L., Morse, D.P., Brodigan, T., Krause, M. and Bass, B.L. (2002) RNA editing by ADARs is important for normal behavior in Caenorhabditis elegans. *EMBO J*, **21**, 6025-6035.

38.   Polson, A.G., Crain, P.F., Pomerantz, S.C., McCloskey, J.A. and Bass, B.L. (1991) The mechanism of adenosine to inosine conversion by the double-stranded RNA unwinding/modifying activity: a high-performance liquid chromatography-mass spectrometry analysis. *Biochemistry*, **30**, 11507-11514.

39.   Macbeth, M.R., Schubert, H.L., Vandemark, A.P., Lingam, A.T., Hill, C.P. and Bass, B.L. (2005) Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. *Science*, **309**, 1534-1539.

40.   Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000) dADAR, a Drosophila double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *Rna*, **6**, 1004-1018.

41.   Kim, U., Garner, T.L., Sanford, T., Speicher, D., Murray, J.M. and Nishikura, K. (1994) Purification and characterization of double-stranded RNA adenosine deaminase from bovine nuclear extracts. *J Biol Chem*, **269**, 13480-13489.

42.   O'Connell, M.A. and Keller, W. (1994) Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. *Proc Natl Acad Sci U S A*, **91**, 10596-10600.

43.   Hough, R.F. and Bass, B.L. (1994) Purification of the Xenopus laevis double-stranded RNA adenosine deaminase. *J Biol Chem*, **269**, 9933-9939.

44.   Herbert, A., Lowenhaupt, K., Spitzner, J. and Rich, A. (1995) Chicken double-stranded RNA adenosine deaminase has apparent specificity for Z-DNA. *Proc Natl Acad Sci U S A*, **92**, 7550-7554.

45. Eckmann, C.R., Neunteufl, A., Pfaffstetter, L. and Jantsch, M.F. (2001) The human but not the Xenopus RNA-editing enzyme ADAR1 has an atypical nuclear localization signal and displays the characteristics of a shuttling protein. *Mol Biol Cell*, **12**, 1911-1924.

46. Nie, Y., Zhao, Q., Su, Y. and Yang, J.H. (2004) Subcellular distribution of ADAR1 isoforms is synergistically determined by three nuclear discrimination signals and a regulatory motif. *J Biol Chem*, **279**, 13249-13255.

47. Poulsen, H., Nilsson, J., Damgaard, C.K., Egebjerg, J. and Kjems, J. (2001) CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domain. *Mol Cell Biol*, **21**, 7862-7871.

48. Lehmann, K.A. and Bass, B.L. (1999) The importance of internal loops within RNA substrates of ADAR1. *J Mol Biol*, **291**, 1-13.

49. Lehmann, K.A. and Bass, B.L. (2000) Double-stranded RNA adenosine deaminases ADAR1 and ADAR2 have overlapping specificities. *Biochemistry*, **39**, 12875-12884.

50. Polson, A.G. and Bass, B.L. (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J*, **13**, 5701-5711.

51. Dawson, T.R., Sansam, C.L. and Emeson, R.B. (2004) Structure and sequence determinants required for the RNA editing of ADAR2 substrates. *J Biol Chem*, **279**, 4941-4951.

52. Kallman, A.M., Sahlin, M. and Ohman, M. (2003) ADAR2 A-->I editing: site selectivity and editing efficiency are separate events. *Nucleic Acids Res*, **31**, 4874-4881.

53. Wong, S.K., Sato, S. and Lazinski, D.W. (2001) Substrate recognition by ADAR1 and ADAR2. *RNA*, **7**, 846-858.

54. Yi-Brunozzi, H.Y., Stephens, O.M. and Beal, P.A. (2001) Conformational changes that occur during an RNA-editing adenosine deamination reaction. *J Biol Chem*, **276**, 37827-37833.

55. Wagner, R.W., Smith, J.E., Cooperman, B.S. and Nishikura, K. (1989) A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells and Xenopus eggs. *Proc Natl Acad Sci U S A*, **86**, 2647-2651.

56.   Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E. and Emeson, R.B. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature*, **387**, 303-308.

57.   Kohler, M., Burnashev, N., Sakmann, B. and Seeburg, P.H. (1993) Determinants of Ca2+ permeability in both TM1 and TM2 of high affinity kainate receptor channels: diversity by RNA editing. *Neuron*, **10**, 491-500.

58.   Lomeli, H., Mosbacher, J., Melcher, T., Hoger, T., Geiger, J.R., Kuner, T., Monyer, H., Higuchi, M., Bach, A. and Seeburg, P.H. (1994) Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science*, **266**, 1709-1713.

59.   Sommer, B., Kohler, M., Sprengel, R. and Seeburg, P.H. (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell*, **67**, 11-19.

60.   Hanrahan, C.J., Palladino, M.J., Ganetzky, B. and Reenan, R.A. (2000) RNA editing of the Drosophila para Na(+) channel transcript. Evolutionary conservation and developmental regulation. *Genetics*, **155**, 1149-1160.

61.   Hoopengardner, B., Bhalla, T., Staber, C. and Reenan, R. (2003) Nervous system targets of RNA editing identified by comparative genomics. *Science*, **301**, 832-836.

62.   Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000) A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. *Cell*, **102**, 437-449.

63.   Reenan, R.A., Hanrahan, C.J. and Barry, G. (2000) The mle(napts) RNA helicase mutation in drosophila results in a splicing catastrophe of the para Na+ channel transcript in a region of RNA editing. *Neuron*, **25**, 139-149.

64.   Semenov, E.P. and Pak, W.L. (1999) Diversification of Drosophila chloride channel gene by multiple posttranscriptional mRNA modifications. *J Neurochem*, **72**, 66-72.

65.   Smith, L.A., Peixoto, A.A. and Hall, J.C. (1998) RNA editing in the Drosophila DMCA1A calcium-channel alpha 1 subunit transcript. *J Neurogenet*, **12**, 227-240.

66.   Dahary, D., Elroy-Stein, O. and Sorek, R. (2005) Naturally occurring antisense: transcriptional leakage or real overlap? *Genome Res*, **15**, 364-368.

67.   Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J. *et al.* (2005) Antisense transcription in the mammalian transcriptome. *Science*, **309**, 1564-1566.

68. Munroe, S.H. and Zhu, J. (2006) Overlapping transcripts, double-stranded RNA and antisense regulation: a genomic perspective. *Cell Mol Life Sci*, **63**, 2102-2118.

69. Saccomanno, L. and Bass, B.L. (1999) A minor fraction of basic fibroblast growth factor mRNA is deaminated in Xenopus stage VI and matured oocytes. *RNA*, **5**, 39-48.

70. Rueter, S.M., Dawson, T.R. and Emeson, R.B. (1999) Regulation of alternative splicing by RNA editing. *Nature*, **399**, 75-80.

71. Beghini, A., Ripamonti, C.B., Peterlongo, P., Roversi, G., Cairoli, R., Morra, E. and Larizza, L. (2000) RNA hyperediting and alternative splicing of hematopoietic cell phosphatase (PTPN6) gene in acute myeloid leukemia. *Hum Mol Genet*, **9**, 2297-2304.

72. Morse, D.P., Aruscavage, P.J. and Bass, B.L. (2002) RNA hairpins in noncoding regions of human brain and Caenorhabditis elegans mRNA are edited by adenosine deaminases that act on RNA. *Proc Natl Acad Sci U S A*, **99**, 7906-7911.

73. Morse, D.P. and Bass, B.L. (1999) Long RNA hairpins that contain inosine are present in Caenorhabditis elegans poly(A)+ RNA. *Proc Natl Acad Sci U S A*, **96**, 6048-6053.

74. Athanasiadis, A., Rich, A. and Maas, S. (2004) Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol*, **2**, e391.

75. Blow, M., Futreal, P.A., Wooster, R. and Stratton, M.R. (2004) A survey of RNA editing in human brain. *Genome Res*, **14**, 2379-2387.

76. Kim, D.D., Kim, T.T., Walsh, T., Kobayashi, Y., Matise, T.C., Buyske, S. and Gabriel, A. (2004) Widespread RNA editing of embedded alu elements in the human transcriptome. *Genome Res*, **14**, 1719-1725.

77. Levanon, E.Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z.Y., Shoshan, A., Pollock, S.R., Sztybel, D. *et al.* (2004) Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol*, **22**, 1001-1005.

78. Levanon, K., Eisenberg, E., Rechavi, G. and Levanon, E.Y. (2005) Letter from the editor: Adenosine-to-inosine RNA editing in Alu repeats in the human genome. *EMBO Rep*, **6**, 831-835.

79.   Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

80.   Scadden, A.D. and Smith, C.W. (1997) A ribonuclease specific for inosine-containing RNA: a potential role in antiviral defence? *Embo J*, **16**, 2140-2149.

81.   Scadden, A.D. and Smith, C.W. (2001) Specific cleavage of hyper-edited dsRNAs. *EMBO J*, **20**, 4243-4252.

82.   Zhang, Z. and Carmichael, G.G. (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell*, **106**, 465-475.

83.   Bass, B.L. (2006) How does RNA editing affect dsRNA-mediated gene silencing? *Cold Spring Harb Symp Quant Biol*, **71**, 285-292.

84.   Nishikura, K. (2006) Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nat Rev Mol Cell Biol*, **7**, 919-931.

85.   Grishok, A. (2005) RNAi mechanisms in Caenorhabditis elegans. *FEBS Lett*, **579**, 5932-5939.

86.   Knight, S.W. and Bass, B.L. (2002) The role of RNA editing by ADARs in RNAi. *Mol Cell*, **10**, 809-817.

87.   Tonkin, L.A. and Bass, B.L. (2003) Mutations in RNAi rescue aberrant chemotaxis of ADAR mutants. *Science*, **302**, 1725.

88.   Saunders, L.R. and Barber, G.N. (2003) The dsRNA binding protein family: critical roles, diverse cellular functions. *Faseb J*, **17**, 961-983.

89.   Morse, D.P. and Bass, B.L. (1997) Detection of inosine in messenger RNA by inosine-specific cleavage. *Biochemistry*, **36**, 8429-8434.

90.   Devine, S.E., Chissoe, S.L., Eby, Y., Wilson, R.K. and Boeke, J.D. (1997) A transposon-based strategy for sequencing repetitive DNA in eukaryotic genomes. *Genome Res*, **7**, 551-563.

91.   Sulston, J.E. and Brenner, S. (1974) The DNA of Caenorhabditis elegans. *Genetics*, **77**, 95-104.

CHAPTER 2

CHARACTERIZATION OF THE *C. ELEGANS*

NONCODING RNA *RNCS-1*

Introduction

With a growing number of metazoan genome sequences completed, it is apparent

that an increase in organismal complexity cannot be explained by a concurrent increase in

the number of protein coding sequences (1). For example, at current count the genome of

the ~1000-cell nematode *Caenorhabditis elegans* contains ~20,000 protein-coding genes,

an only slightly smaller number than the estimated 22,000 genes predicted in humans (2).

Certainly, posttranscriptional processes, such as RNA splicing and RNA editing, account

for some of this discrepancy. Another solution to this paradox lies in the nonprotein-

coding transcripts produced from an organism's genome. While the protein-coding

sequences in *C. elegans* encompass 26Mb, an additional 33Mb of transcripts are

noncoding. In contrast, the 34Mb of human coding sequence are dwarfed by the 1600Mb

of noncoding human transcription products (3). Noncoding RNA (ncRNA) is implicated

in a variety of cellular functions that include dosage compensation, chromatin remodeling

and epigenetic imprinting, modulation of transcription and translation, regulation of RNA

splicing and editing, and stress response (4). However, in light of the immense repertoire

of noncoding transcripts found in complex organisms, our knowledge of ncRNA function

and biochemistry is merely anecdotal.

In recent years, large-scale approaches using experimental RNomics have identified noncoding RNAs that do not belong to familiar classes of structural and functional RNAs (5,6). While such studies caused the number of noncoding RNAs to surge from a handful of known transcripts to thousands of candidates, experimental design usually restricted searches to a limited size range or to binding partners of known RNA binding proteins. In addition, reverse transcription techniques employed during the cloning of noncoding transcripts present a bias against longer and highly structured ncRNAs, and such transcripts are likely to be underrepresented in the current catalog of ncRNA candidates.

Previous studies performed in our laboratory used a biochemical approach to identify polyadenylated *C. elegans* and human transcripts that are substrates of adenosine deaminases that act on RNA (ADARs), editing enzymes that introduce adenosine to inosine changes in double-stranded regions of RNA (7,8). A major discovery of this work was that extended double-stranded regions are found in noncoding regions of mature messenger RNAs as well as in completely noncoding transcripts. One of the *C. elegans* ADAR substrates identified was the product of the F55A4.9 gene (Figure 2.1). This highly structured 800-nucleotide (nt) transcript, from here on referred to as *rncs-1* (RNA noncoding, starvation up-regulated), is spliced and polyadenylated, but does not contain a functional open reading frame and is not found on polysomes (8).

Intrigued by the unique nature of the *rncs-1* transcript, we set out to characterize expression, regulation and function of this RNA. We find that *rncs-1* is expressed abundantly throughout development in hypodermal and intestinal tissues, with increased expression observed in males. Transcription of *rncs-1*, at least partially under control of

Figure 2.1. **Genomic location and secondary structure of *rncs-1*.** (A) Diagram of nucleotides 1009000 to 1014000 on the X chromosome, showing the genomic location of the 847-nucleotide *rncs-1* gene and part of its immediate neighbors. Grey boxes, coding exons; white boxes and arrows, noncoding exons; connecting lines, introns. The 815bp region deleted in the *rncs-1(tm1632)* mutant is indicated. Diamonds represent the locations of the (A/T)GATA(A/G) motifs recognized by GATA-type transcription factors, with white diamonds indicating the extended TGATAAGA consensus frequently found in *C. elegans* intestinal genes. The genomic regions used as promoter sequence in the *P_{rncs-1}:GFP* reporter (promoter) and as rescue fragment in the *rncs-1* rescue and overexpressing lines (rescue) are indicated. (B) Secondary structure of the mature *rncs-1* RNA as predicted by *mfold* and biochemical methods (nuclease probing, data not shown) Aasterisks, GU base pairs; dots, single nucleotide mismatches and bulges; grey arrowhead, exon-exon junction
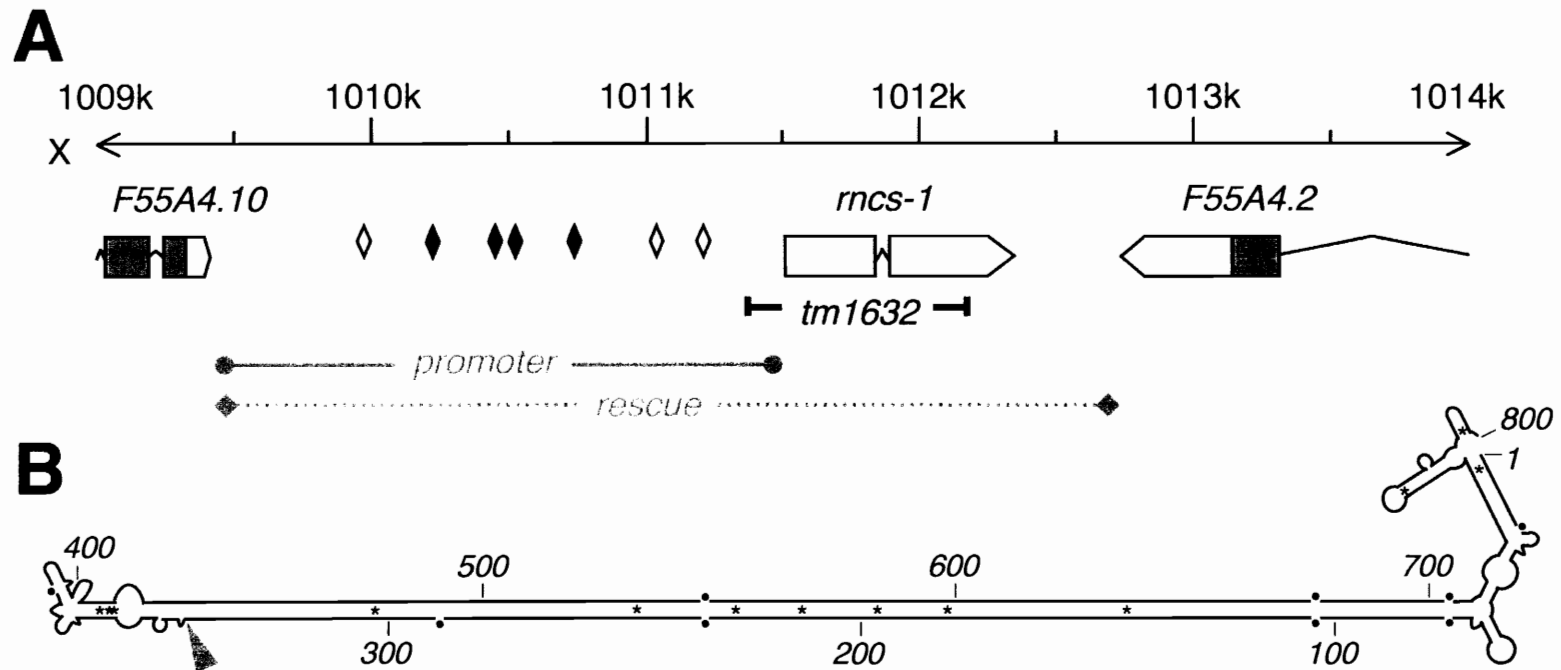
the intestinal GATA factor ELT-2, is modulated in response to environmental food supply, suggesting that this noncoding RNA may play a role in a starvation stress response. Ectopic overexpression of *rncs-1* leads to increase in frequency of male offspring, a phenomenon previously reported to occur in response to heat stress. While highly double-stranded in nature, *rncs-1* RNA is not a substrate for the RNase III enzyme Dicer *in vitro*. We show that branched terminal structures protect the double-stranded region from cleavage by Dicer, a finding that may have implications for other endogenous double-stranded RNAs (dsRNAs). Moreover, *rncs-1* RNA inhibits Dicer activity *in vitro*, raising the possibility that this is a cellular function of *rncs-1*. In support of this model, we find that several genes negatively regulated by Dicer *in vivo* show mRNA levels that vary in correlation with changes in *rncs-1* expression.

Experimental procedures

Strains and worm culture. The following *C. elegans* strains were maintained using standard culture techniques and conditions (9): *N2, BB15 rncs-1(tm1632) X, BB16 uuEx6 [pPrncs-1:GFP], BB17 rncs-1(tm1632) uuEx7 [rncs-1(+) pPrncs-1:GFP], BB32 uuEx7 [rncs-1(+) pPrncs-1:GFP], BB18 uuEx8 [rncs-1(+) pRF4], JM124 elt-4(ca16), elt-7(tm840), BB1 dcr-1(ok247); unc-32(e189)III* (10), *BB13 rde-4(ne299) III*. The *rncs-1(tm1632)* mutant was received from the National Bioresource Project for the Nematode, Japan, and backcrossed six times to generate the *BB15* strain.

To grow synchronized, developmentally staged cultures, eggs were isolated by sodium hypochlorite treatment of gravid hermaphrodites, hatched overnight in M9 buffer, transferred to NGM agar plates seeded with *E. coli* OP50 and incubated at 20°C for 6-8 hrs (L1), 18-20 hrs (L2), 28-30 hrs (L3), 40-42 hrs (L4), or 50-54 hrs (young adults).

Arrested L1 samples were isolated after allowing embryos to hatch in M9 for 16-24 hrs without subsequent introduction to food.

For starvation time courses, worms were grown to L4 in well-fed liquid culture or on plates and subsequently starved in M9 buffer. In experiments that reintroduced starved cultures to food, worms were placed onto NGM agar plates seeded with *E. coli* OP50.

Males were isolated by crossing wildtype hermaphrodites to males to generate ~50% male progeny. Male offspring were separated from hermaphrodite siblings by sorting to ~99% purity at the L4 or young adult stage.

Dauer larvae were isolated by growing liquid culture for 7 to 10 days at 20°C to high population density and food exhaustion. Dauers were purified from nondauers by treatment with 1% SDS as published (11). Pheromone-containing plates were prepared with partially purified dauer pheromone as described in detail in Chapter 4. Sufficient pheromone extract was added to NGM agar prepared without peptone to induce ~100% dauer formation in N2 at 20°C; plates were seeded with *E. coli* OP50. Several gravid N2 hermaphrodites were placed on a pheromone plate and allowed to lay eggs for 8 hrs before being removed. After 3 days progeny that entered dauer were harvested.

Transgenics. The P_{rncs-1}:*GFP* reporter was constructed by amplifying ~2000bp of the putative upstream promoter and regulatory sequences of genomic *rncs-1* ("promoter", Figure 2.1A) using primers 52GpromUHindIII (AGTATCAAGCTTATTTAAATAATA TACTTGTGTGC) and 52GpromLKpnI (TCTTGCGGTACCTGAAACATTTTATAGTT TTAGATGAA) followed by ligation into the HindIII and KpnI restriction sites of pPD49.26. This plasmid was injected into N2 at 5ng/μL to generate transgenic lines as

described (12). To establish lines overexpressing *rncs-1*, a ~3400bp genomic fragment

("rescue" Figure 2.1A) was amplified using primers 52GRSQU2 (CTGTAGTATCCATG

TTATTTAAATAATATACTTGTGTGC) and 52GRSQL2 (TTTCTGAGAAAATATAA

ATACGCATTGATAATTAAACC) and injected at 5ng/μL into *rncs-1(tm1632)* along

with the P_{rncs-1}:*GFP* reporter to generate BB17 or into N2 along with pRF4 to generate

BB18. All microinjected DNA contained excess nonspecific DNA (1kb ladder,

Invitrogen) to yield a final DNA concentration of 100ng/μL.

RNA isolation, northern analysis, and quantitative RT-PCR. For all experiments

that quantify RNA levels in nonintregrated transgenic lines, ~99% pure transgenic

population samples were provided by picking individuals carrying the injection marker.

Similarly, *dcr-1(ok247); unc-32(e189)* homozygous populations were isolated by picking

coiling progeny as described (13).

Total RNA was isolated from worm pellets using Trizol (Invitrogen). This RNA

was used directly for northern analyses. For use in RT-PCR, total RNA was further

purified by treatment with TURBO DNase (Ambion) followed by RNeasy column

chromatography (Qiagen). RNA used in northern analyses to detect siRNAs was size

selected for species <200nt using the *mir*Vana™ miRNA Isolation Kit (Ambion).

Northern analyses for large RNAs were performed using standard protocols for

1% formaldehyde agarose gel electrophoresis and nylon membrane blotting. The *rncs-1*

probe was synthesized using PCR to nucleotides 323-442 of the *rncs-1* cDNA, with a T7

promoter sequence linked to the antisense primer (see Chapter 3 for primer sequences).

Similarly, a template for the GFP probe was amplified using primers gfp5'northern

(CTCCAATTGGCGATGGCCCTGTC) and gfpT7 (TAATACGACTCACTATAGGGA

GATAGTTCATCCATGCCATGTGTAATCC). RNA northern probes were synthesized by *in vitro* transcription using the Strip-EZ® T7 Kit (Ambion). The region of *rncs-1* RNA detected by this probe is not subject to RNA editing *in vivo* (8). To detect 18S ribosomal RNA, 18S-complementary oligonucleotides (TAAGTTTCGCGCCTGCTGCC TTCCTTGGACGTGGTAGCCG, AGGAGAGCTGGAATTACCGCGGCTGCTGGCA CCAGACTTG, AGCGACGGGCGGTGTGTACAAAGGGCAGGGACGTAATCAA) were ^{32}P-labeled using T4 polynucleotide kinase (New England Biolabs). ULTRAhyb™ hybridization buffer (Ambion) was used in all northern experiments to detect *rncs-1* or 18S. Total polyadenylated RNA was detected by dot blotting serial dilutions of total RNA onto nitrocellulose followed by detection with a ^{32}P-labeled $(dT)_{30}$ oligonucleotide as published (14).

Northern analyses to detect small RNAs were performed employing standard techniques for 15% polyacrylamide gel electrophoresis and blotting. For hybridization, a protocol developed by Nelson Lau was adapted (15). In brief, blots were prehybridized for 2 hrs at 42°C in 25mL of PreHyb/Hyb solution (5X SSC, 20mM Na_2HPO_4 pH7.2, 7% SDS, 2X Denhardt's Solution; 1mg of denatured herring sperm DNA added before use). The prehybridization buffer was discarded and a fresh 25mL aliquot of PreHyb/Hyb solution, containing denatured herring sperm DNA and labeled probes was added to the blot. The blot was hybridized at 42°C over night. The hybridization solution was disposed and the blot was washed two times for 10 min at 42°C in 40mL of nonstringent wash solution (3X SSC, 25mM NaH_2PO_4 pH7.5, 5% SDS, 10X Denhardt's Solution), followed by two washes for 30 min at 42°C in 40mL of nonstringent wash solution supplemented with 20μL of 100mM ATP. The blot was washed once with 80mL of

stringent wash solution (1X SSC, 1% SDS) and exposed on a PhosphorImager screen. All solutions were prewarmed to 42°C.

To detect small RNA species with F53A9.2 sequence a mixture of labeled 50-nucleotide sense-oriented DNA oligos that span the entire F53A9.2 cDNA was used. Labeled antisense DNA oligos were used to detect *let-7* and U6 RNAs.

To synthesize cDNA samples used in quantitative RT-PCR analyses, 1μg aliquots of DNase treated total RNA (see above) were reverse transcribed using SuperScript™ II reverse transcriptase and oligo-dT primers (Invitrogen). Samples were subsequently treated with RNaseH (New England Biolabs) for 30 min and diluted fourfold with ddH$_2$O. 5μL of cDNA samples were analyzed per quantitative PCR in a LightCycler® 2.0 instrument using the LightCycler® FastStart DNA Master^{PLUS} SYBR Green I Kit. Primer pairs were designed to span at least one exon-exon junction and produce products of 150 to 300bp. All primer pairs used for quantitative RT-PCR are listed in Table 2.1.

Cloning of rncs-1. Because amplification of the full-length mature transcript from cDNA was unsuccessful, a two-step approach to cloning was taken. First, the *rncs-1* gene was amplified from N2 genomic DNA by PCR using primers SH21 (TGGAGCTCA TTGAGCAAGATGGGGAT) and SH22 (CATTCTGCAGTTTATTGATTCAACATTT CAAAAAC), which contain SacI and PstI restriction sites, respectively. The PCR product was cloned into the SacI and PstI sites of pGEM3zf(+). To eliminate the intron in a second cloning step, exons 1 and 2 were amplified individually from the genomic clone. The exon 1 primers SH45 (TCGAATTCATTGAGCAAGATGGGGATGT) and SH46 (AAGAACTTATTCTGCCTTTCAGCATCGACCACG) linked an EcoRI site to the 5' end of the product and the most 5' nucleotides of exon 2, including an XmnI site,

Table 2.1. **Primer sequences used for quantitative PCR analysis.**

mRNA	5' primer	3' primer	Product size (bp)
let-413	let-413LCU CCGCA GGCAC TCAAA ACATG CACAC AATCA G	let-413LCL ATGTT CTCGA AACTG GCGAT GGGGA CCG	315
eps-8	eps-8LCU CCGGA AGACG TGACT CGTTG GTTAC AGG	eps-8LCL ATTGC TTTCA ATTCA TCGCC AGTGT GAGTT C	218
elt-2	elt-2LCU ATGTG TCACA TGCCG ACTTG TATCC CG	elt-2LCL TAAGA ATCTC CGTCG ACCGC TTCCA AATC	265
elt-4	elt-4LCU CTTAG ATGCT TCTCA TCGGA AACGG CTT	elt-4LCL TTCTT CATTG GGATG GGGCG GGT	156
elt-7	elt-7LCU TGCTGCTCACACTGCTCAACA ACTACAACCA	elt-7LCL CGCCG CTTTC GAGTA GTTGG CTTCT GC	156
gpd-3	gpd3 For GGAGG AGCCA AGAAG GTC	gpd3 Rev AAGTG GAGCA AGGCA GTT	144
F53A9.2	F53A9.2LCU GAGGA TACAA TCCAT ATGGA GCCTA CGGG	F53A9.2LCL GTGGT GACCA TGGTG ATGCC CTCC	119
B0222.3	B0222.3LCU TACGT TATCA AGACT GTTGG AACAA AAATG TCA	B0222.3LCL ATGAC ATGAA AATAT TCCGG AACGT AGACC A	217
C30F12.6	C30F12.6LCU CTTGA TGTCT GCTAG ATTCC GTGCT GC	C30F12.6LCL CCAGCCATTGATGTTATTAATG GAGATTTTG	162
C06G8.1	C06G8.1LCU ACCAG GACAA CGTGC CCCAA TCG	C06G8.1LCL ATGAG TATAA GCGAA TTGAT CCTTT GTTGG C	235
T07C5.1	T07C5.1LCU GATGC AGCAT ATCAT TGACT ACAAG TTCG	T07C5.1LCL TTGTC CATGA TGGCC TCGCT G	190
F35D11.3	F35D11.3LCU TGCTG GGACT GGACC AATTG ATAAT CG	F35D11.3LCL ACAAT TGGAT GCTCT TCCTC TTCAC CG	258
mab-3	mab-3LCU CTTGT GGAAC ACCGT CGAAA TCTTA TGGC	mab-3LCL ATTGT TGATT CATTG GGATC GGCGG A	325
wrt-10	wrt-10LCU CTCCA GCACC AACAA CTACC CCAAC TGTG	wrt-10LCL AAGCT CTACG GATTC CAGTC GATTG TCTTG	243

to the 3' end of the product. Exon 2 was amplified using primers SH47 (CAGAATAAGT TCAACCCGAT) and SH48 (CTAAGCTTTTTATTGATTCAACATTT), which added a HindIII recognition sequence to the 3' end of the exon. The exon 1 product was subsequently digested with EcoRI and XmnI, the exon 2 product with XmnI and HindIII. The *rncs-1* cDNA clone pGEM52G was established by three-way-ligation into the EcoRI and HindIII sites of pGEM3zf(+), with exon-exon linkage mediated by joining of the XmnI sites.

In vitro transcription. Full-length *rncs-1* RNA was transcribed from HindIII-linearized plasmid pGEM52G using T7 RNA polymerase. Transcription templates for *rncs-1* derivatives were generated by PCR amplification of the following segments of *rncs-1* cDNA with T7 promoter sequences linked to the 5' primers: nucleotides (nt) 1-342 (dervative 2, 5'strand), nt 438-800 (derivative 2, 3' strand), nt 71-393 (derivative 3, 5' strand), nt 394-709 (derivative 3, 3' strand), nt 71-342 (derivative 4, 5' strand), nt 438-709 (derivative 4, 3' strand). Transcription template for the 3' strand of derivative 5 was generated by amplifying nt 71-342 with a T7 promoter site on the 3' primer. Note that derivative 4 is an intermolecular duplex, and primers were modified to replace the tetraloop of nt 392-395 with two GC basepais in the RNA derivative. Primer sequences used for the synthesis of transcription templates are listed in Table 2.2.

Transcription templates for the 300bp *unc-22* dsRNA were produced by PCR amplification of part of exon 6 of unc-22 from genomic DNA using primers unc-22300bpU (GGAGTCCCGCCGATCGAGTAGAACTGA) and unc-22300bpL (GGCGGCCTTGTGTATTTGTCACGCG) with T7 promoter sites linked to the 5' primer for sense template or the 3' primer for the antisense template.

Table 2.2. **Primer sequences used for synthesis of transcription templates for *rncs-1* derivatives.**

Derivative	Strand	5' primer	3' primer
2	5'	52GET7 *TAATA CGACT CACTA TA*ATT GAGCA AGATG GGGAT GTAGT CA	52GB GGCTT TCAGC ATCGA CCACG GCTT
	3'	52GCT7 *TAATA CGACT CACTA TA*GGC TTTCA GCATC GACCA	52GF TTTAT TGATT CAACA TTTCA AAAAC T
3	5'	52GAT7 *TAATA CGACT CACTA TA*GGC ATCGA GCTGA AAGAG A	52GG CCCCT TGAGT TCCAT CTTTG TC
	3'	52GHT7 *TAATA CGACT CACTA TA*GGC CTCAG AGACA ATCAG G	52GD GGCAT TGAGC TGAAA GAGAG AGCA
4	5'	52GAT7 *TAATA CGACT CACTA TA*GGC ATCGA GCTGA AAGAG A	52GB GGCTT TCAGC ATCGA CCACG GCTT
	3'	52GCT7 *TAATA CGACT CACTA TA*GGC TTTCA GCATC GACCA	52GD GGCAT TGAGC TGAAA GAGAG AGCA
5	5'	52GAT7 *TAATA CGACT CACTA TA*GGC ATCGA GCTGA AAGAG A	52GB GGCTT TCAGC ATCGA CCACG GCTT
	3'	52GA GGCAT CGAGC TGAAA GAGAG AGCA	52BT7 *TAATA CGACT CACTA TA*GGC TTTCA GCATC GACCA CGGCT T

Internally ^{32}P-labeled RNA was transcribed with T7 RNA polymerase as previously described (16), and purified from a 4% denaturing polyacrylamide gel. Intermolecular and intramolecular duplex RNAs were annealed as described (17), and purified from a 5% native gel. Unlabeled *rncs-1* RNA was synthesized similarly, except that components of the transcription reaction were adjusted to improve RNA yield (18).

In vitro Dicer activity assay. Embryo extract preparation and Dicer activity assays were performed as previously described (17) with the following modifications: For the competition assay, glycerol in the extract was reduced to 20% and the extract was further cleared by centrifugation at 100,000g for 60 min. All Dicer activity assays consisted of 20μL reactions containing 50μg of total protein and 10nM labeled RNA substrates. Reactions containing unlabeled *rncs-1* as an inhibitor were preincubated for 5 min at 20°C with unlabeled *rncs-1* before adding labeled 300bp *unc-22* dsRNA.

Results

The rncs-1 transcript is expressed in intestine and hypodermis in all stages of C. elegans development and is enriched in males. C. elegans development proceeds from the embryo through four larval stages (L1 to L4) into adult worms, completing a lifecycle in ~3 days. Postembryonic development is dependent on the presence of food, and C. elegans arrest at the beginning of L1 under starved conditions (L1 arrest). Further, lack of food supply and high culture density in early development prompts entry into an alternative larval form called dauer. To study *rncs-1* expression, we isolated total RNA from synchronized cultures at each developmental stage, as well as from arrested L1s, and dauer larvae purified from starved and crowded liquid culture. Northern blot analysis revealed an abundant RNA species of 800 nt (Figure 2.2A). This transcript was present

at constant levels in embryos, larval stages and adults. Interestingly, we noticed an appreciable increase in *rncs-1* RNA in arrested L1 worms and dauers.

To investigate the spatial expression of *rncs-1,* we constructed P_{rncs-1}:*GFP*, a reporter with ~2kb of putative *rncs-1* regulatory and promoter sequences driving expression of green fluorescent protein (Figure 2.1A). The reporter was microinjected into the gonad of wildtype (N2) worms. Six independent lines carrying the P_{rncs-1}:*GFP* extrachromosomal array were isolated, and the expression pattern was identical in all lines. Expression of GFP initiated early in embryogenesis, with fluorescence becoming evident in the early gastrula (data not shown). Gastrulation in *C. elegans* is initiated by the ingression of the daughters of the E founder cell, which subsequently divide and differentiate to give rise to the midgut section of the intestinal tract (19). We observed robust expression of P_{rncs-1}:*GFP* in the E lineage starting at the 28-cell stage and continuing into adulthood (Figure 2.2B-D). By the comma stage of embryogenesis (Figure 2.2B), additional GFP fluorescence was visible in the embryo periphery, in cells that give rise to hypodermal tissue. In L1 larva and subsequent stages, we observed strong expression of GFP in hypodermal cells, including the Hyp 7 syncytium and cells of the head and tail hypodermis. In addition, adult hermaphrodites displayed fluorescence in the vulval epithelium. Reporter expression was notably absent in germline, seam cells, nervous system, and pharynx.

Our developmental northern blot indicated that *rncs-1* RNA increases during starvation and in dauer larvae (Figure 2.2A). Similarly, the P_{rncs-1}:*GFP* reporter was continuously expressed throughout prolonged starvation periods (data not shown). While

fluorescence intensity was enhanced under starved conditions, the spatial expression pattern was unchanged.

In male worms carrying the P_{rncs-1}:*GFP* transgene we noticed enhanced expression of the reporter. To examine expression of the endogenous *rncs-1* locus in males, we prepared total RNA from populations of wildtype, well-fed male and hermaphrodite worms, and quantified *rncs-1* by northern blot (Figure 2.2E,F). We observed an ~2.5-fold increase in *rncs-1* expression in males when compared to hermaphrodites. In contrast, differences in mRNA levels were not observed for the mRNAs of *let-413* and *eps-8*, two intestinal genes involved in midgut formation and maintenance (20,21) (Figure 2.2F).

rncs-1 transcription is regulated by food supply. To determine whether the increase in *rncs-1* levels observed in arrested L1 and dauer larvae was a direct response to lack of food, we removed food from cultures of synchronized L4 worms, and harvested worms for RNA isolation over the course of 2 days. We observed an ~3-fold induction of *rncs-1* levels over the unstarved control in as little as 1 hr of starvation (Figure 2.3A,B). After 10 hrs without food, *rncs-1* RNA levels reached a maximum of ~6-fold enrichment, and this abundance was maintained for the 48-hr starvation period. After 48 hrs of starvation, worms were reintroduced to food and within 6 hrs *rncs-1* RNA levels decreased to baseline. To address whether *C. elegans* mRNA levels generally increase as a result of starvation, we subjected the total RNA samples isolated during the starvation time course experiments to northern analyses using an oligo-dT probe (data not shown, see Experimental Procedures). Unlike *rncs-1*, the general population of polyadenylated RNA did not increase in the absence of food (Figure 2.3B, dashed line).

Figure 2.2. **Temporal and spatial expression of *rncs-1*.** (A) Northern blot of *rncs-1* in 20μg total RNA samples isolated from wildtype embryos, larval stages (L1-L4), young adults, arrested L1s, and dauer larvae purified from exhausted liquid culture (see Experimental Procedures). The 800-nucleotide *rncs-1* RNA was present at constant levels throughout development. In starvation-arrested L1 worms and in dauer larvae, transcript levels were increased approximately 5 to10-fold over baseline levels. The blot was rehybridized with a probe for 18S ribosomal RNA (rRNA) as a loading control. Data are representative of multiple northern blots (n=2-4, depending on stage). (B-D) GFP expression under the control of the *rncs-1* upstream regulatory and promoter sequences. (B) In the comma stage embryo, strong fluorescence was evident in the E cell lineage of the developing midgut (dmid). Additional GFP expression was observed in the embryo periphery, in cells of the developing hypodermis (dh). (C, D) In adults, the midgut (mid), but not other sections of the intestinal tract, showed strong expression of the P_{rncs-1}:*GFP* reporter. Fluorescence was also observed in cells of the head hypodermis (hh), the tail hypodermis (th), the Hyp 7 syncytium (Hyp 7), and the vulval epithelium (ve). Visible GFP expression was absent in the germline, seam cells (sc), and nervous system. Diagrams in the bottom right corner indicate the approximate focal plane. (E) Northern blot for *rncs-1* in total RNA samples prepared from hermaphrodite (H) and male (M) populations of wildtype worms. The blot was rehybridized to a probe for 18S ribosomal RNA to control for loading. (F) Quantification of relative RNA levels in four independent samples of males (grey bars) and hermaphrodites (white bars). *rncs-1* amounts were determined by northern blot, while *let-413* and *eps-8* levels were quantified by real-time quantitative RT-PCR analyses. Error bars, the standard error of the mean, S.E.M.; *, $P < 0.05$, t-test.
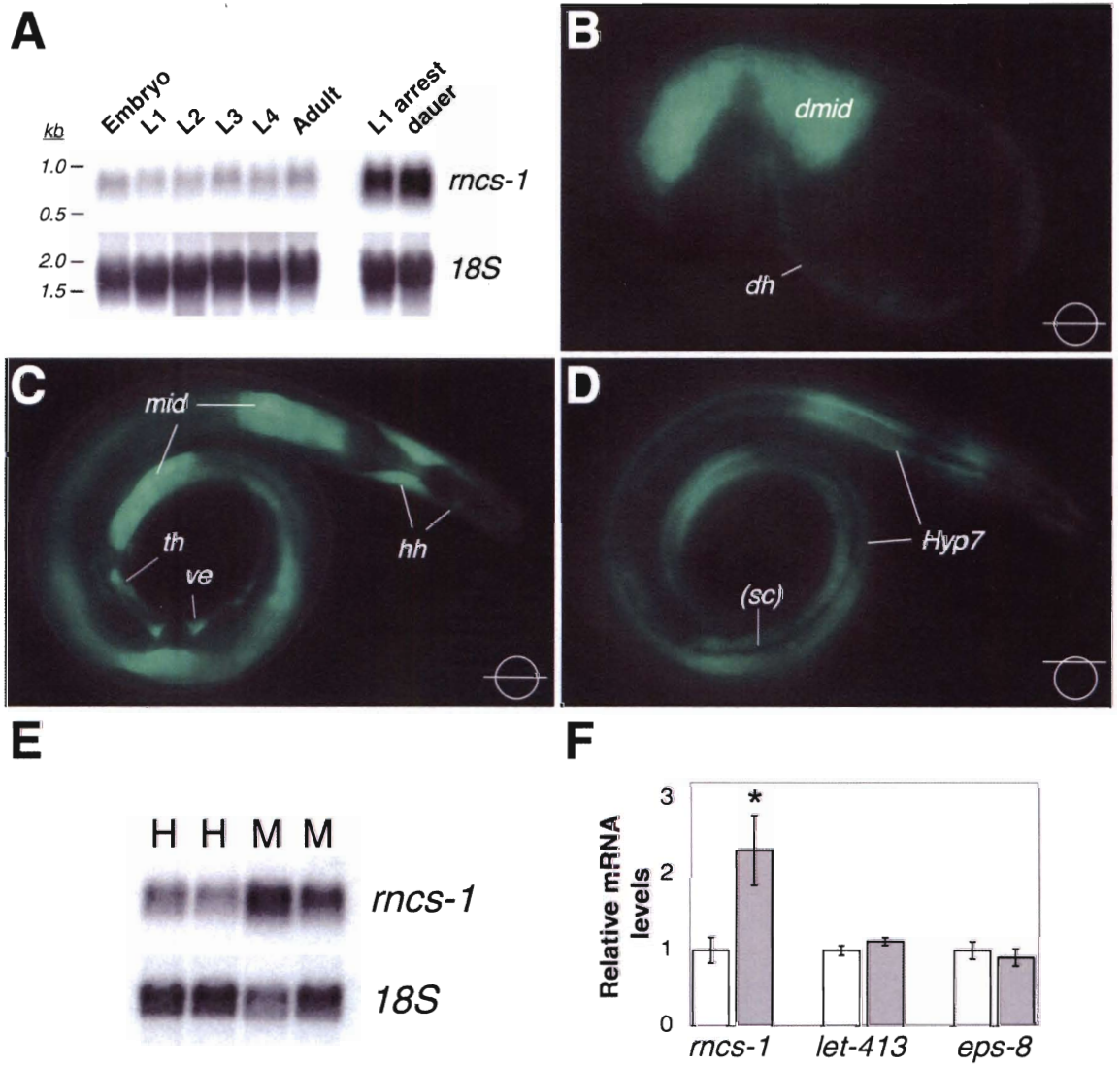
Instead, we observed a slight decrease in polyadenylated transcript levels to ~80% of unstarved levels. Similar to *rncs-1*, the overall polyadenylated RNA population returned to baseline levels within 6 hrs after reintroduction of food.

The *C. elegans* dauer larva is an enduring, non-aging and non-feeding stage of the nematode lifecycle (22). Entry into dauer is mediated by sensory integration of two environmental cues: the lack of food and the abundance of dauer pheromone (daumone). Daumone is a soluble fatty acid derivative secreted by the worm throughout its lifecycle (23). High levels of daumone in the growth medium, indicative of high population density, can prompt entry into dauer despite the presence of sufficient food (24). We took advantage of the different pathways to dauer to investigate whether *rncs-1* transcript levels were elevated due to the absence of food in the growth medium, or because of lack of food in the worm's digestive tract. When RNA was isolated from dauer larvae purified from exhausted liquid culture, where starvation conditions as well as high pheromone levels existed, we observed increased levels of *rncs-1* when compared to L3 worms (Figure 2.3C). However, in samples where dauer larva formation was initiated on agar plates containing high levels of daumone but also sufficient food, no increase was detected. This result indicated that regulation of *rncs-1* occurs in response to environmental food signals rather than actual food uptake.

To investigate whether the increase of *rncs-1* RNA during starvation was due to posttranscriptional stabilization of the transcript, or transcriptional up-regulation, wildtype animals carrying the P_{rncs-1}:*GFP* reporter was subjected to a starvation time course. In RNA samples isolated at various time points of starvation and after re-feeding, we compared the changes in GFP mRNA levels to changes in endogenous *rncs-1*

transcript levels by northern analyses. Consistent with the enhanced fluorescence observed in starved worms carrying the P_{rncs-1}:GFP transgene, we found that induction of GFP under control of the *rncs-1* promoter parallels the regulation of genomic *rncs-1* (Figure 2.3D). It is therefore likely that regulation of *rncs-1* in response to food supply is at the level of transcription, via upstream regulatory and promoter sequences.

Upregulation of rncs-1 during starvation is dependent on the GATA factor ELT-2. To investigate which factors are involved in transcriptional regulation of *rncs-1*, we searched the upstream regulatory sequence of *rncs-1* for recognition sites of known transcription factors. We identified seven sequence motifs containing the (A/T)GATA(A/G) consensus recognized by the GATA family of transcription factors (Figure 2.1A), including three extended GATA consensus sites frequently found in *C. elegans* intestinal genes (25). *C. elegans* encodes three GATA-type transcription factors that, like *rncs-1*, are expressed in the postembryonic intestine: ELT-2, ELT-4 and ELT-7 (26). Deficiency in *elt-2* causes lethality and a gut obstruction at the L1 stage (27), while worms with null mutations in either or both of the *elt-4* and *elt-7* genes are essentially wildtype (28). Therefore, ELT-2 is believed to be the major GATA factor in the intestine.

Because of the lethality of deleting the *elt-2* gene, we decreased ELT-2 activity by RNA interference (RNAi). Synchronized cultures of wildtype *C. elegans* were grown on *E. coli* OP50 until the L2 stage and subsequently transferred to feed on bacteria expressing dsRNA containing GFP (dsGFP) or *elt-2* (dsELT-2) sequence. After 24 hrs of dsRNA feeding, half of the population was harvested, while the remainder was starved for 3 hrs. By northern analyses we determined that *rncs-1* levels under fed conditions

Figure 2.3. **Regulation of *rncs-1* transcription in response to environmental food supply.** (A) Northern blot for *rncs-1* in 15μg of total RNA isolated from wildtype L4 larvae removed from food for the indicated amount of time (1 to 48 hrs) and subsequently reintroduced to food for several hours (48 hrs starved, 1 to 6 hrs fed). The blot was rehybridized to a probe for 18S rRNA to control for loading. (B) Quantified northern results for two independent cohorts of worms subjected to the starvation/feeding time course. Black diamonds/solid line, *rncs-1* RNA; white diamonds/dashed line, total polyadylated RNA detected by an oligo-dT northern probe; black arrowhead, time of food addition; error bars, scatter. Transcript levels were normalized to 18S rRNA and are shown relative to the unstarved 0 hr sample. (C) Comparison of *rncs-1* levels in well-fed L3 larvae (L3), dauer larvae isolated from starved and crowded liquid culture (dau), and dauer larvae grown in the presence of food on dauer pheromone rich plates (dau (ph)). (D) Northern blot of GFP mRNA and *rncs-1* transcript in RNA samples from wildtype worms carrying the P_{rncs-1}:*GFP* transgene that were subjected to a starvation/feeding time course.

A

0	1	3	6	10	24	48	48	48	48	Hrs starved
							1	3	6	Hrs fed

rncs-1

18S

B

C

L3	dau	dau (ph)	
-	+	+	pheromone
+	-	+	food

rncs-1

18S

D

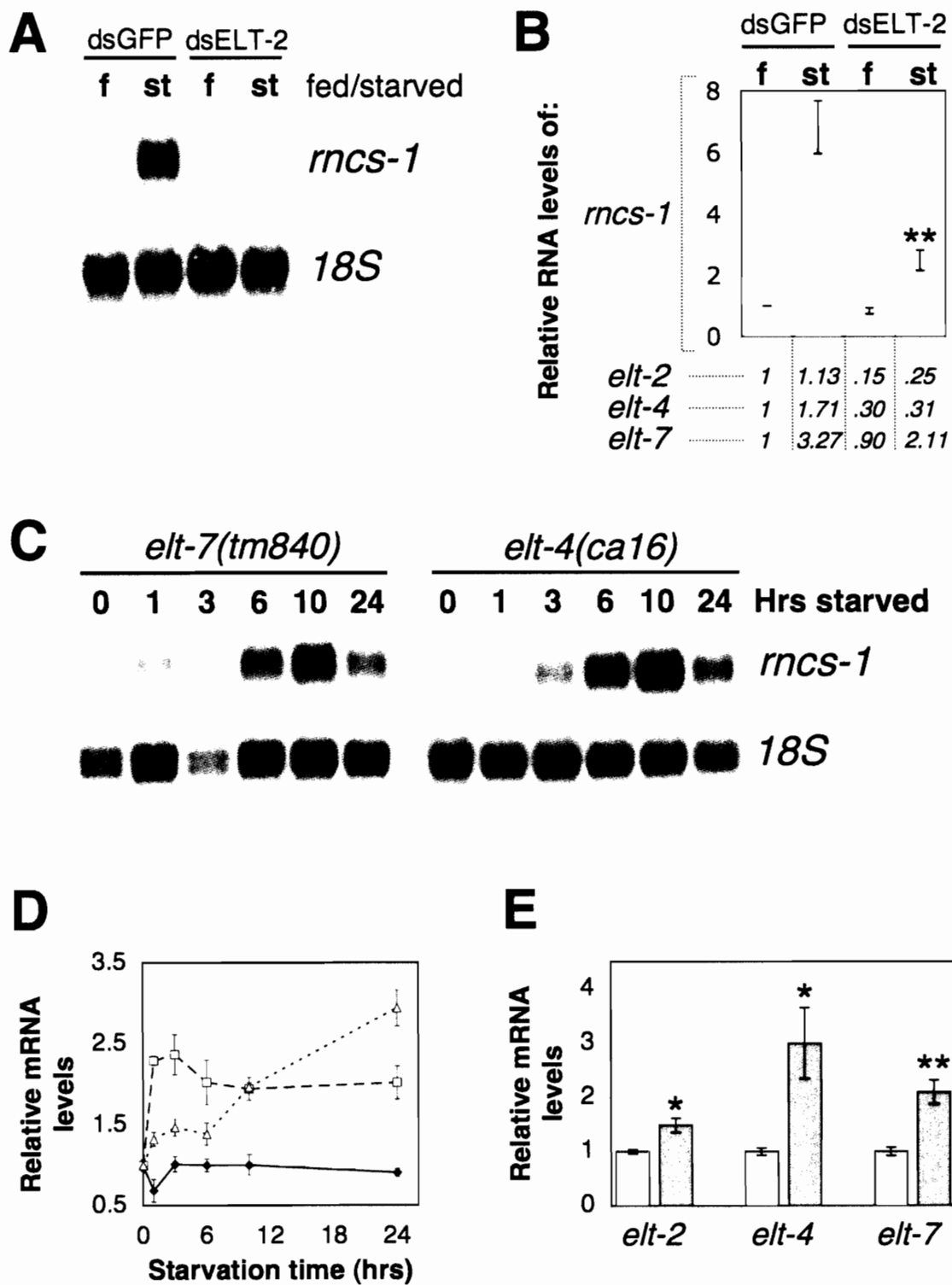| 0	3	6	10	24	24	Hrs starved
					6	Hrs fed

P_{rncs-1}:*GFP*

rncs-1
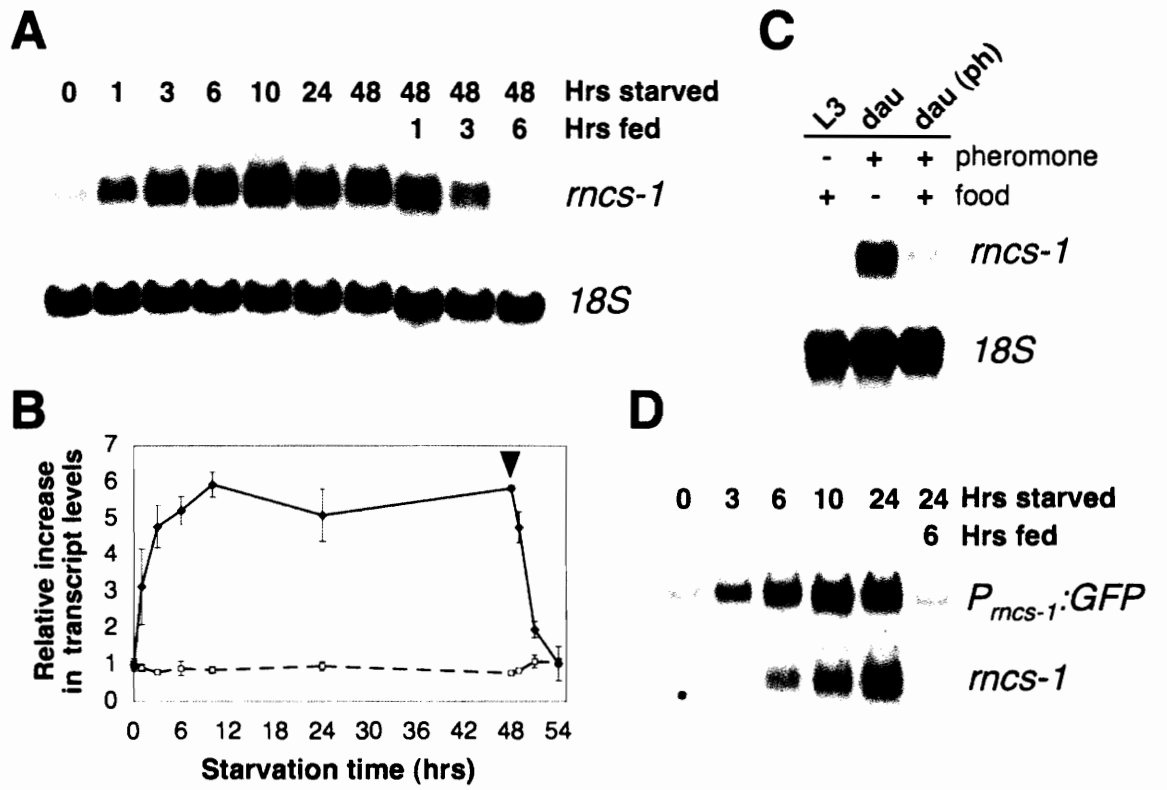
were only modestly reduced by *elt-2* RNAi to ~80% of control levels (Figure 2.4A,B). This weak reduction was not surprising, as transient RNAi against *elt-2* only causes a mild, if any, reduction in expression of other ELT-2 regulated genes (29). Feeding of dsELT-2 reduced levels of *elt-2* mRNA to an average of 15% in fed samples when compared to the fed dsGFP control (tabulation, Figure 2.4B), and this level may be sufficient for ELT-2 function under nonstarved conditions.

Importantly, starved animals in which *elt-2* was reduced by RNAi showed a significantly weaker induction of *rncs-1* (~2.5-fold compared to ~7-fold). Because the *elt-4* and *elt-2* mRNAs share high sequence conservation, targeting of *elt-2* by RNAi also reduced the mRNA levels of *elt-4* (tabulation, Figure 2.4B). To test the possibility that *elt-4* or *elt-7* were also involved in *rncs-1* regulation, we examined *rncs-1* induction during starvation in *elt-4(ca16)* and *elt-7(tm840)* mutant worms (Figure 2.4C). No difference in *rncs-1* induction was observed in either single mutant as well as the *elt-4;elt-7* double mutant (data not shown) when compared to wildtype.

Interestingly, we observed that *elt-4* and *elt-7* mRNA, but not *elt-2* mRNA levels increased upon starvation (Figure 2.4D). For this reason we are cautious in interpreting the dsELT-2 results. Although deletion of *elt-4* in worms wildtype for *elt-2* had no effect on *rncs-1* induction, we do not exclude the possibility that the reduction of *rncs-1* levels during starvation reported in Figures 4.2A and B required reduced expression of both *elt-2* and *elt-4*. It is also unclear whether the effect of *elt-2* on *rncs-1* induction is direct or mediated by a separate *elt-2*-dependent transcription factor.

In addition, we found that *elt-2*, *elt-4* and *elt-7* mRNA levels are enriched in males (Figure 2.4E), correlating with the male enrichment observed for *rncs-1*.

Figure 2.4. ***ELT-2*-dependence of *rncs-1* induction during starvation.** (A) Northern blot of *rncs-1* in young adult wildtype worms fed with dsRNA against GFP (dsGFP) or *elt-2* (dsELT-2). f, worms harvested well fed; st, worms harvested after 3 hrs of starvation. The blot was rehybridized for 18S rRNA to control for loading. (B) In three independent experiments *rncs-1* levels were quantified in by northern analyses (bar graph; error bars, S.E.M.); *elt-2*, *elt-4* and *elt-7* mRNA levels were quantified by real-time quantitative RT-PCR (table below graph; numbers indicate mean mRNA levels relative to the dsGFP fed control). (C) Northern blot of *rncs-1* in RNA samples from *elt-7(tm840)* and *elt-4(ca16)* mutant worms subjected to a starvation time course. (D) Relative mRNA levels of *elt-2* (black diamonds, solid line), *elt-7* (white squares, dashed line) and *elt-4* (white triangles, dotted line) as quantified by real-time RT-PCR in wildtype worms during a starvation time course. An average of two independent experiments is plotted relative to the unstarved expression level of each message. Error bars, S.E.M. (E) Relative mRNA levels of *elt-2*, *elt-4* and *elt-7* in wildtype adult hermaphrodites (white bars; average of three independent samples) and males (grey bars; four samples) as quantified by real-time RT-PCR. Error bars, S.E.M.; *, $P < 0.05$; **, $P < 0.01$; t-test.

A dsGFP dsELT-2
f st f st fed/starved

rncs-1

18S

B dsGFP dsELT-2
f st f st

Relative RNA levels of:

rncs-1

	f	st	f	st
elt-2	1	1.13	.15	.25
elt-4	1	1.71	.30	.31
elt-7	1	3.27	.90	2.11

C elt-7(tm840) elt-4(ca16)
0 1 3 6 10 24 0 1 3 6 10 24 Hrs starved

rncs-1

18S

D

Relative mRNA levels

Starvation time (hrs)

E

Relative mRNA levels

elt-2 elt-4 elt-7

The GATA factor ELT-2 has recently been implicated in regulating intestinal innate immunity genes during bacterial and fungal infections of *C. elegans* (29,30). The *elt-2*-dependent induction of *rncs-1* during food deprivation hints at an involvement of ELT-2 in the starvation stress response pathway.

Ectopic overexpression of rncs-1 RNA causes increased spontaneous male frequency. To study the physiological consequences of loss of *rncs-1* RNA we obtained a strain carrying a deletion in the *rncs-1* locus. The *tm1632* deletion (Figure 2.1A) removes 815bp of DNA, including ~150bp of promoter sequence, and inserts 232bp of sequence not found elsewhere in the *C. elegans* genome. Northern analyses using probes complementary to either the 5' or 3' half of the *rncs-1* RNA detected no transcript of any size in *rncs-1(tm1632)* homozygous worms (data not shown), indicating that *tm1632* is a null allele. We also constructed a rescue line by introducing a genomic fragment (Figure 2.1A), into *rncs-1(tm1632)* mutants along with the P_{rncs-1}:*GFP* injection marker. Expression of *rncs-1* RNA in the rescue line was analyzed by northern blot, and found to be 10 to 15-fold enhanced over levels observed in unstarved wildtype worms. We used this line to investigate effects of overexpressing *rncs-1*, and also crossed the P_{rncs-1}:*rncs-1*, P_{rncs-1}:*GFP* array into a wildtype background.

While *rncs-1(tm1632)* animals appeared healthy, in examining *rncs-1* rescue/overexpressing worms we noticed an unusually high occurrence of males. This prompted us to score male frequency in self-progeny of wildtype, *rncs-1(tm1632)*, *rncs-1* rescue/overexpressing lines, as well as control lines carrying the P_{rncs-1}:*GFP* reporter. Groups of 500 to 1000 L4 hermaphrodite larvae were transferred to a separate plate and allowed to grow to the first day of adulthood. Progeny were isolated by hypochlorite

treatment and synchronized by hatching without food over night. The number of

hermaphrodites and males was counted at the adult stage. Table 2.3 summarizes the data

from several independent experiments. Our assay revealed no discernible difference in

spontaneous male frequency in *rncs-1(tm1632)* mutants and wildtype worms (0.041 and

0.042% respectively). However, when *rncs-1* was overexpressed in the mutant or

wildtype background, spontaneous male frequency was signifcantly increased (8 to 15-

fold; P values <0.0001).

An increase in spontaneous male progeny of lesser magnitude (~3-fold) was also

observed when GFP was expressed under the *rncs-1* promoter in N2 or *rncs-1* mutants.

We previously reported that dsRNA synthesized from transgenes leads to transgene

silencing in ADAR mutant *C. elegans* (31). After crossing the P_{rncs-1}:*GFP* transgene into

worms carrying mutations in both *C. elegans* ADAR genes, *adr-1* and *adr-2*, we

observed silencing of GFP (data not shown). This indicates that dsRNA is generated

from the P_{rncs-1}:*GFP* transgene. Possibly, the increase in male frequency by *rncs-1*

overexpression is mediated by the double-stranded character of the transcript, and dsRNA

from the P_{rncs-1}:*GFP* extrachromosomal array can mimic this effect, albeit to a lesser

degree.

We observed ~0.04% males among self-progeny of wildtype hermaphrodites, a

value less than the ~0.2% previously reported (32). This difference may derive from the

fact that we only scored progeny of day 1 adults (isolated by hypochlorite treatment),

while the cited report scored progeny over the worm's lifespan. It is possible that the

frequency of meiotic nondisjunction that leads to spontaneous male progeny increases

with age of the parent.

Table 2.3. **Frequency of spontaneous male offspring from day 1 adult hermaphrodites**

Strain	Number of hermaphrodites	Number of males	Male frequency (%)	P values (χ^2)	
N2	11802	5	**0.042**		
N2 + P_{rncs-1}:rncs-1 + P_{rncs-1}:GFP	6846	43	**0.624**	<0.0001	0.026
N2 + P_{rncs-1}:GFP	8312	11	**0.132**	<0.0001	
rncs-1(tm1632)	9654	4	**0.041**		
rncs-1(tm1632) + P_{rncs-1}:rncs-1 + P_{rncs-1}:GFP	9646	33	**0.341**	<0.0001	0.114
rncs-1(tm1632) + P_{rncs-1}:GFP	4370	5	**0.114**	0.017	

Terminal branched structures protect rncs-1 RNA from processing by Dicer.

Long dsRNA, when introduced into *C. elegans*, is cleaved by the RNase III enzyme

Dicer into ~23-nucleotide small interfering RNAs (siRNAs). The *rncs-1* secondary

structure contains an almost perfectly double-stranded helix of ~300bp (Figure 2.1B),

suggesting that *rncs-1* RNA may be a cellular substrate for Dicer. While *C. elegans*

Dicer has not been purified in an active form, Dicer activity can be assayed by incubating

^{32}P-labeled dsRNA in wildtype embryo extracts (33). We used this assay to test whether

siRNAs are generated from *rncs-1* RNA *in vitro*. We cloned the full-length mature *rncs-1*

transcript and synthesized ^{32}P-labeled *rncs-1* RNA by *in vitro* transcription. Using

nuclease mapping techniques we verified that the resulting transcription product adopted

the secondary structure predicted by *mfold* (34) (Figure 2.5A substrate 1, mapping data

not shown). Surprisingly, when *rncs-1* RNA was incubated in wildtype embryo extract,

no small RNA species were observed (Figure 2.5B, lane 1). The absence of detectable

siRNAs persisted over all ranges of RNA concentration, extract concentration, and

reaction time tested (data not shown).

Previous studies using recombinant human Dicer suggested that cleavage of long

dsRNA is more efficient if initiated from its termini (35). We speculated that the

branched structural elements flanking the central RNA helix of *rncs-1* hindered efficient

processing. To test this hypothesis, we synthesized *rncs-1* derivatives lacking one or

both of the terminal structures (Figure 2.5A, substrates 2-4). A substrate consisting of the

rncs-1 helix with all mismatches repaired served as a positive control (Figure 2.5A,

substrate 5). Incubation of the *rncs-1* derivatives in embryo extract produced siRNAs

when at least one helix terminus was blunt-ended (Figure 2.5B, lanes 2-5). This
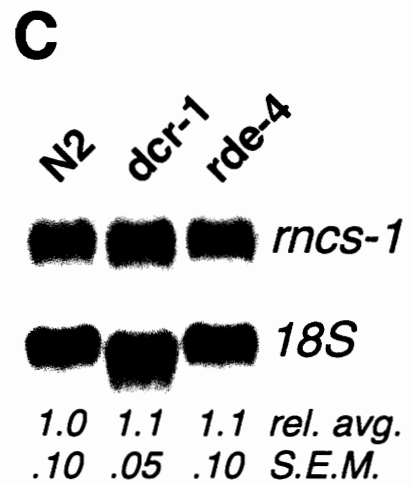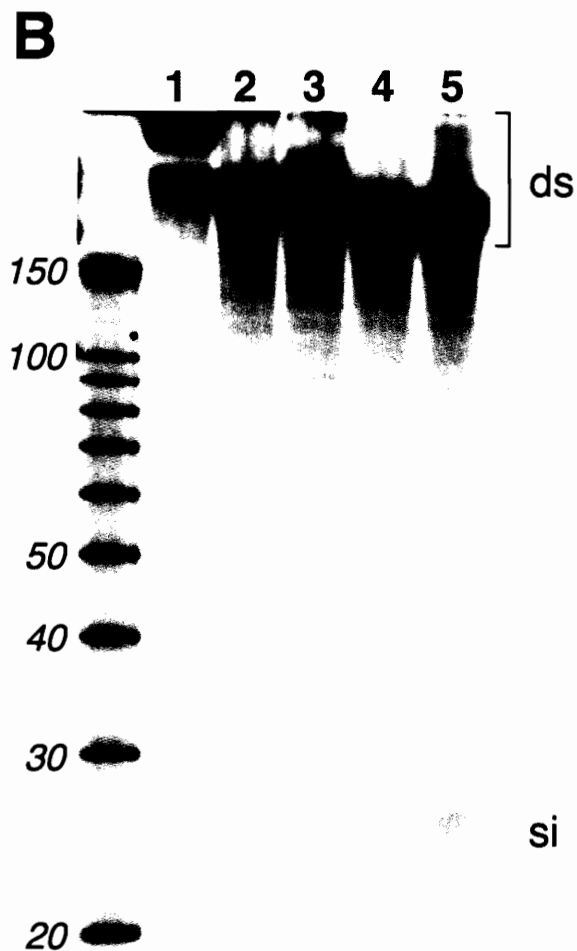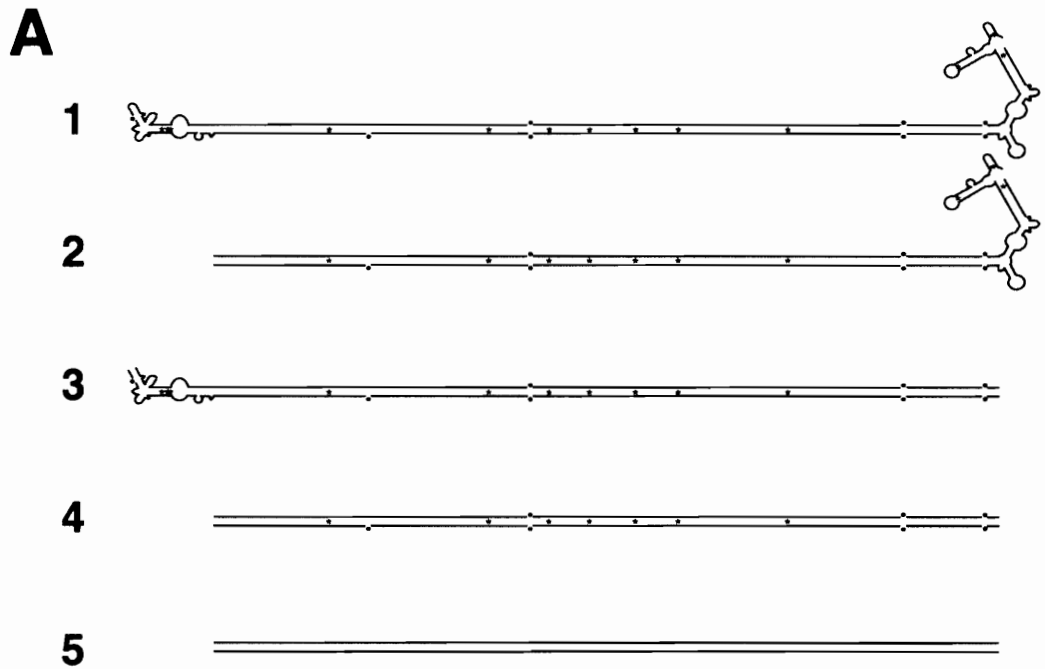
indicated that the terminal branched structures of *rncs-1* protect it from cleavage by Dicer *in vitro*.

To investigate whether *C. elegans* Dicer also fails to cleave full-length *rncs-1* RNA *in vivo*, we used northern analyses to determine *rncs-1* transcript levels in *dcr-1(ok247)* and *rde-4(ne299)* mutant worms, strains incabable of siRNA production (10,36). We observed that the amount of *rncs-1* RNA is not altered in either mutant when compared to wildtype (Figure 2.5C). Further, northern blots failed to detect any RNA species other than the full-length *rncs-1* transcript, suggesting that this RNA is stable *in vivo*.

rncs-1 RNA inhibits dsRNA cleavage by Dicer in vitro. *C. elegans* siRNA production involves at least two dsRNA binding proteins. The domain structure of Dicer includes a dsRNA binding motif (dsRBM) at the C-terminus (37). In addition, the small accessory protein RDE-4, containing two dsRBMs, is essential for Dicer-mediated cleavage of dsRNA (17,38). dsRBMs generally bind to RNA in a sequence-nonspecific manner (39,40). This property is thought to allow different cellular dsRNA pathways to intersect (31,41), and several recent studies have focused on how various dsRNA binding proteins can compete for substrates (42,43). Conversely, and relevant to *rncs-1*, it is plausible that different dsRNA substrates may compete with each other for dsRBM-containing proteins.

Although full-length *rncs-1* RNA was cleaved by Dicer, we wondered if it could bind to Dicer and/or RDE-4, to sequester these factors and inhibit cleavage of another dsRNA, *in trans*. We added increasing amounts of unlabeled full-length *rncs-1* to embryo extract containing a ^{32}P-labeled 300bp dsRNA substrate of *C. elegans unc-22*

Figure 2.5. **Structure-dependent processing of *rncs-1* and derivatives by Dicer.** (A) Schematic of full-length *rncs-1* (1) and *rncs-1* derivatives (2-5) synthesized by *in vitro* transcription. Substrate 5 is identical to substrate 4, except that GU base pairs and mismatches were repaired. (B) Autoradiogram of the reaction products of 1-hr incubations of the ^{32}P-labeled *rncs-1* derivative substrates in wildtype embryo extract as analyzed by denaturing gel electrophoresis. Successful processing of a dsRNA (ds) substrate by Dicer produces siRNA species (si). Full-length *rncs-1* RNA is resistant to Dicer cleavage (lane 1). Substrates that lacked either one (lanes 2 and 3) or both (lane 4) of the branched structures flanking the central dsRNA helix of *rncs-1* were substrates for Dicer, as was a perfectly double-stranded control (lane 5). RNA Decade Markers (Ambion) were used as size standards. (C) Northern blot of *rncs-1* in RNA samples prepared from wildtype (N2), *dcr-1(ok247)* and *rde-4(ne299)* worms. Numbers below each lane indicate the relative average (rel. avg.) and standard error of the mean (S.E.M.) calculated by northern analyses of three independent samples collected for each strain. The blot was rehybridized to 18S rRNA to control for loading.
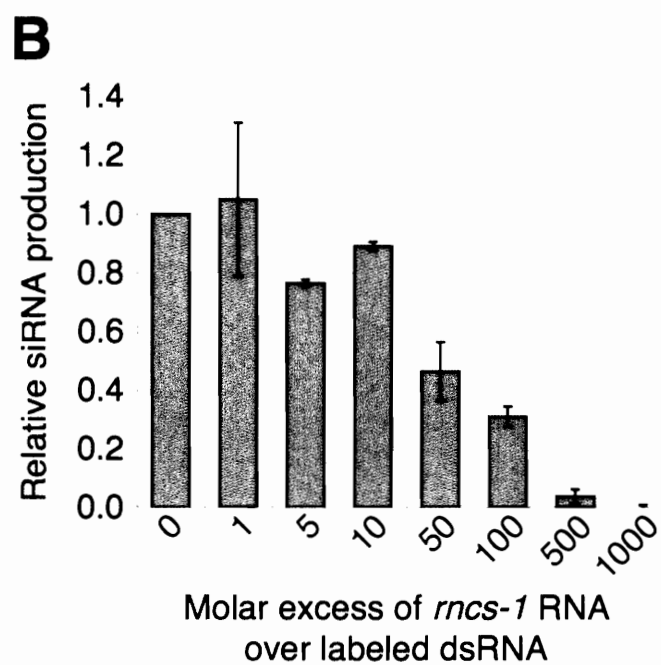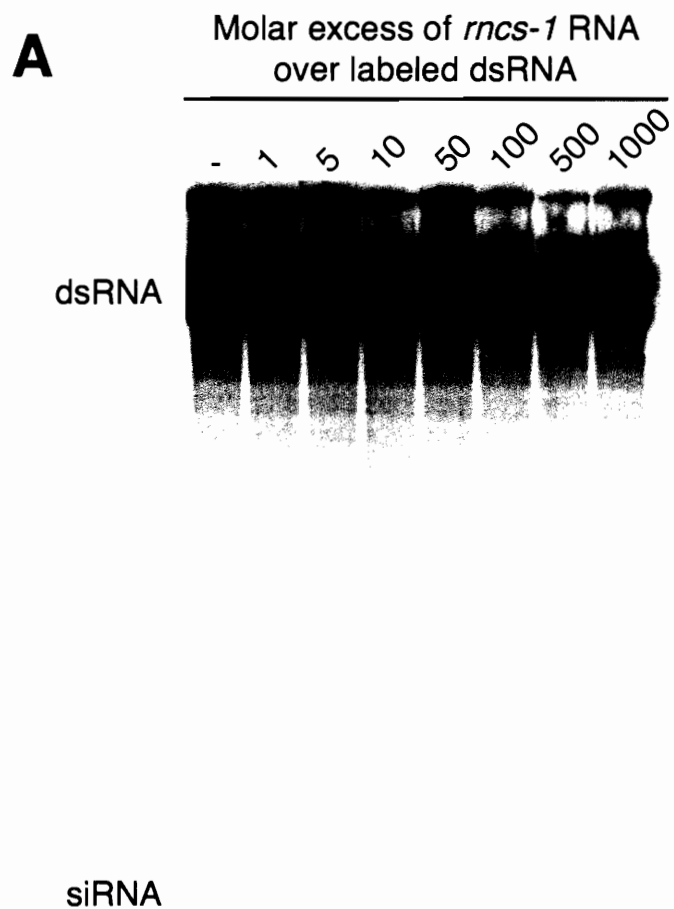
A

1

2

3

4

5

B

1 2 3 4 5

ds

150

100

50

40

30

si

20

C

N2 dcr-1 rde-4

rncs-1

18S

1.0 1.1 1.1 rel. avg.
.10 .05 .10 S.E.M.

sequence. As expected, the *unc-22* dsRNA gave rise to siRNAs when incubated in embryo extract for 30 min (Figure 2.6A, lane 1). However, when extract was preincubated with *rncs-1*, a reduction in ^{32}P-siRNA was observed, beginning at 5 to 10-fold molar excess (Figure 2.6A,B). At 50-fold excess, Dicer cleavage was inhibited by ~50%, and at 500- to 1000-fold excess, siRNA was undetectable. We conclude that *rncs-1* RNA inhibits production of siRNAs *in vitro*, presumably by competing with a dsRNA substrate for binding to Dicer or RDE-4.

Assuming identical binding affinities, true competitive inhibition predicts 50% inhibition at a 1:1 molar ratio of substrate and inhibitor. In our experiments a larger molar excess of *rncs-1* RNA was required to reduce *unc-22* siRNA production by half. This is likely due to the limitations of the experimental system. Cleavage of dsRNA in *C. elegans* embryo extract is inefficient, with siRNA encompassing less than 1% of the radioactive label of the substrate, even after prolonged incubation. To generate amounts of siRNA sufficient for detection, the incubation time must be extended to exceed single turnover of the enzyme. In addition, extracts contain unknown amounts of endogenous dsRNA binding proteins and dsRNAs that may affect the reaction. It is also possible that *rncs-1* RNA has a lower affinity for dsRBMs than the perfectly duplexed *unc-22* dsRNA, thereby reducing its potency as an inhibitor of Dicer binding. For these reasons, we consider the inhibition data to be of a qualitative, rather than a quantitative, nature.

Expression of Dicer-regulated genes depends on rncs-1 levels. To investigate whether *rncs-1* inhibits Dicer activity *in vivo*, we searched for Dicer-regulated genes with altered expression in *rncs-1* mutant and overexpressing lines. Our lab recently used microarray analysis to identify genes whose mRNA levels are Dicer-dependent (13).

Figure 2.6. **Competitive inhibition of Dicer activity by *rncs-1* RNA *in vitro*.** Wildtype embryo extract was preincubated for 5 min with increasing amounts of unlabeled full-length *rncs-1* followed by addition of ^{32}P-labeled 300bp *unc-22* dsRNA substrate (20nM), followed by incubation for 30 min. (A) Autoradiogram of a denaturing gel analyzing the products of a representative competition experiment. Decreased *unc-22* siRNA is observed as the molar excess of *rncs-1* RNA is increased. (B) Quantification of efficiency of siRNA production from *unc-22* dsRNA relative to the reaction in the absence of *rncs-1* transcript. The average of two sets of experiments using independent preparations of embryo extract is shown. Error bars, scatter of data.

A

Molar excess of *rncs-1* RNA
over labeled dsRNA

- 1 5 10 50 100 500 1000

dsRNA

siRNA

B

Molar excess of *rncs-1* RNA
over labeled dsRNA

Within the group of genes negatively regulated by Dicer we searched for candidates with a predicted or confirmed expression pattern similar to that of *rncs-1*. We chose 18 genes and designed primer pairs for quantitative real time RT-PCR analyses. In an initial screen, we compared mRNA expression levels in samples of young adult worms with wildtype, *rncs-1(tm1632)* and *dcr-1(ok247)* genotypes. If *rncs-1* antagonizes Dicer activity *in vivo*, genes normally silenced by Dicer would show a further decrease in mRNA levels in a *rncs-1* deficient background. Of the 18 candidates analyzed, 6 did not reproduce the expected increase in mRNA levels in the Dicer mutant, and were excluded from further analysis (data not shown). Of the remaining 12 genes, 5 genes showed reproducible decrease in mRNA levels in *rncs-1(tm1632)* samples (Figure 2.7A). The reduction in mRNA levels varied in magnitude from ~50%, observed for the F53A9.2 gene, to ~10% for the T07C5.1 mRNA. Seven genes analyzed, including F35D11.3 (Figure 2.7A), showed negative regulation by Dicer that was not further enhanced in the *rncs-1(tm1632)* mutant. F35D11.3 was included in further analyses as a negative control.

We next isolated samples of young adult hermaphrodites overexpressing *rncs-1* RNA in a *rncs-1(tm1632)* or wildtype background, as well as of wildtype worms expressing the P_{rncs-1}:*GFP* transgene. Quantitative RT-PCR analyses verified that Dicer mRNA levels were unchanged in these strains (data not shown). Consistent with the idea that *rncs-1* modulates expression of Dicer-regulated genes, genes with reduced mRNA levels in the *rncs-1* mutant had increased mRNA levels when *rncs-1* was overexpressed (+ lanes, Figure 2.7B). The magnitude of mRNA increase in overexpressing lines correlated with the magnitude of mRNA reduction in *rncs-1(tm1632)* animals. For example, F53A9.2 mRNA showed the greatest decrease in *rncs-1(tm1632)* animals

Figure 2.7. **Dicer-regulated genes respond to *rncs-1* levels.** (A) Relative expression of genes negatively regulated by Dicer in wildtype (grey) and *rncs-1(tm1632)* (white) adult hermaphrodites as quantified by real time RT-PCR. The average of six independent samples for each strain is shown; samples were normalized to *gpd-3* expression. Error bars; S.E.M.; *, P < 0.05, t-test. (B) Graphs: relative expression of the genes shown in (A) for the following strains: wildtype (N2, grey), *rncs-1(tm1632)* (*rncs-1*, white), *rncs-1(tm1632)* overexpressing *rncs-1 (rncs-1(+)*, grey on white hatch), wildtype overexpressing *rncs-1* (N2(+); black on grey hatch) and wildtype with P_{rncs-1}:*GFP* reporter (N2(gfp); black on grey checkerboard). The average of three to six independent samples for each strain is plotted normalized to *let-413* levels. Error bars; S.E.M.; *, P < 0.05, **, P < 0.01, ***, P < 0.001, t-test; grey asterisks, comparison to *rncs-1(tm1632)*; black asterisks, comparison to N2. Tabulation below graphs: Expression of analyzed genes relative to wildtype hermaphrodites in *dcr-1(ok247)* mutants and in samples of wildtype males as quantified by real time RT-PCR. The average of three to four independent samples is listed (asterisks, significance levels as above). Triangles illustrate observed trends for the sensitivity of gene expression to *rncs-1* expression (light grey) and to regulation by Dicer (dark grey).
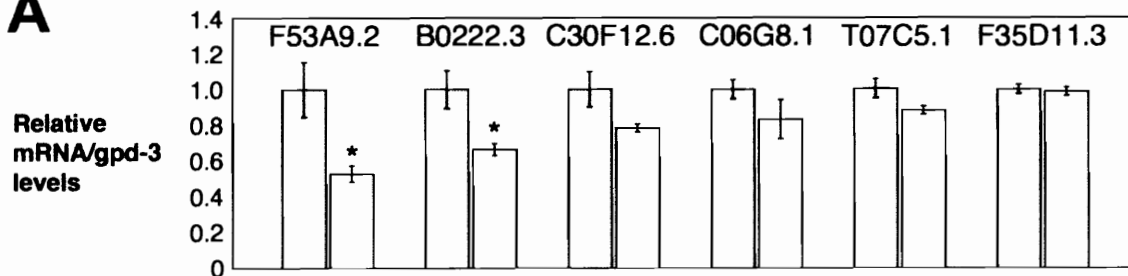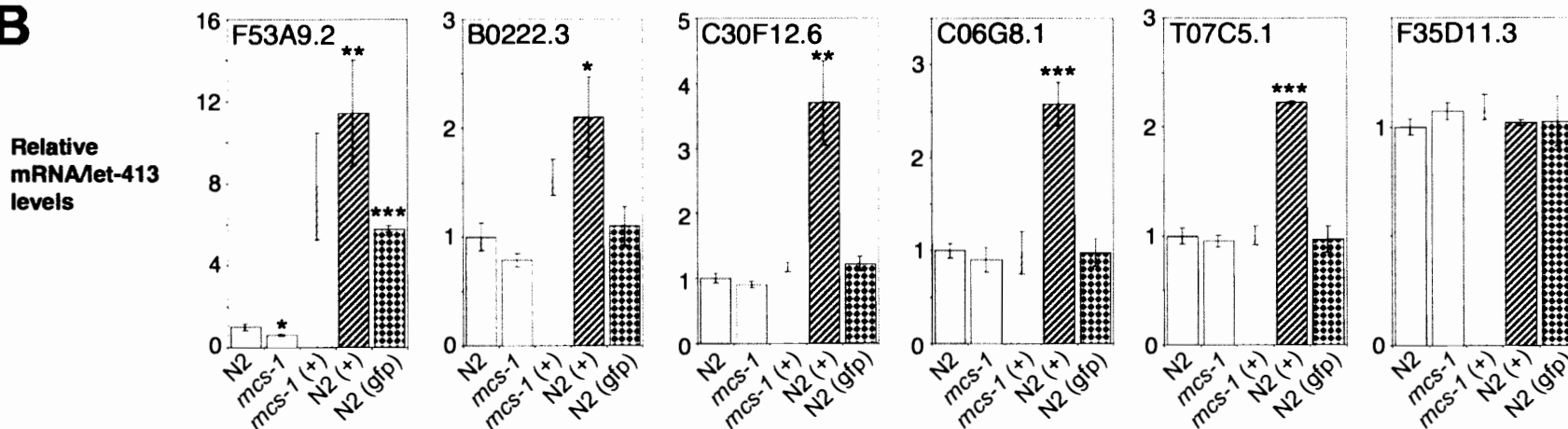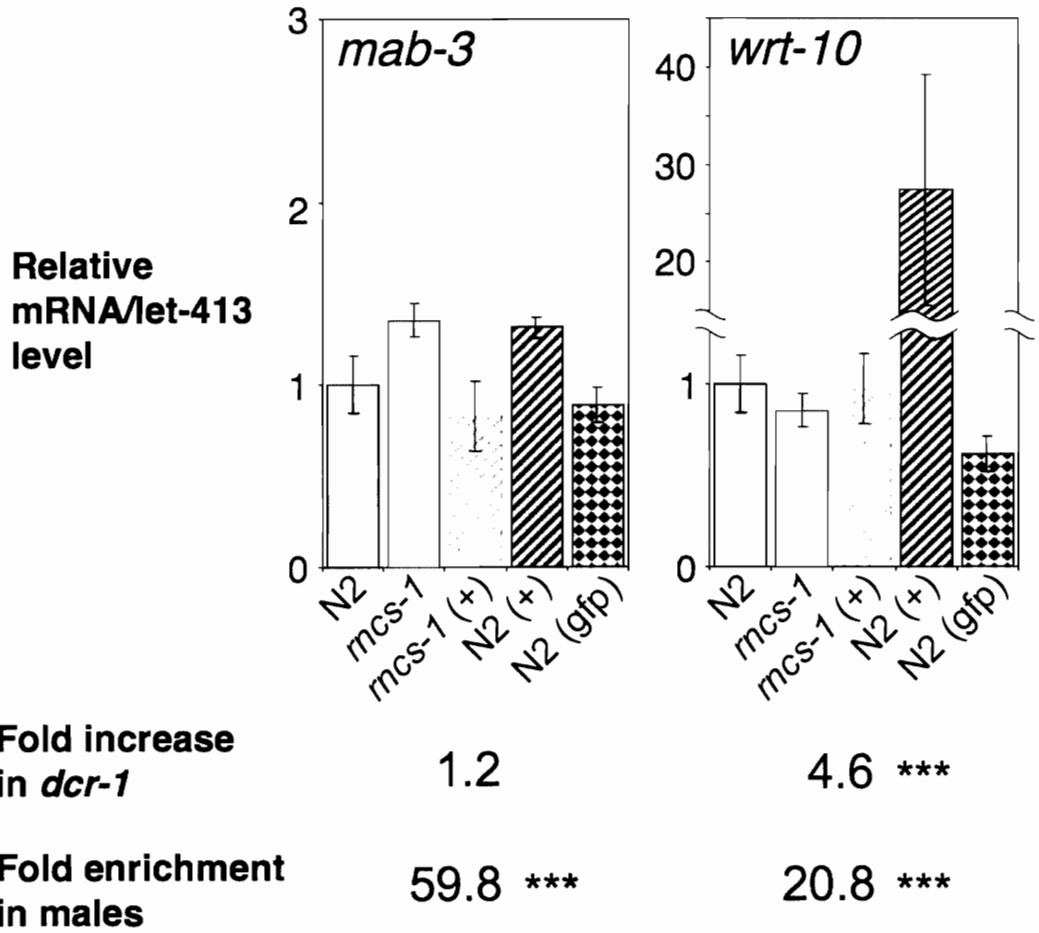
(Figure 2.7A), and the greatest increase in overexpressing strains (+ lanes, Figure 2.7B). Similarly, T07C5.1 mRNA levels decreased only slightly in *rncs-1(tm1632)* animals (Figure 2.7A) and showed only a slight increase (~2-fold) in overexpressing wildtype strains (N2 (+), Figure 2.7B). In addition, sensitivity to *rncs-1* expression also correlated with the magnitude of silencing by Dicer (tabulation, Figure 2.7B). For F53A9.2, but not the other genes, mRNA levels were also increased by the P_{rncs-1}:*GFP* transgene, consistent with the idea that dsRNA produced from this transgene can inhibit Dicer similarly to *rncs-1* overexpression. Of all analyzed genes, F53A9.2 expression showed the greatest sensitivity to Dicer and *rncs-1* levels, implying that even small amounts of dsRNA interfere with its regulation. Because binding of dsRNA by dsRBMs is sequence-nonspecific, any dsRNA, like *rncs-1*, can potentially antagonize Dicer activity.

Interestingly, F53A9.2, C30F12.6 and C06G8.1 were upregulated >5-fold in males (tabulation, Figure 2.7B), while B0222.3, T07G5.1 and the *rncs-1* insensitive gene F35D11.3, showed a slight upregulation. To exclude the possibility that upregulation of Dicer targets in the *rncs-1* overexpressing lines was due to a high percentage of males in our samples, we analyzed expression levels of *mab-3* and *wrt-10*, two intestinal genes previously shown to be enriched in males (44). The *mab-3* mRNA, while ~60-fold male enriched, showed no upregulation when *rncs-1* was overexpressed (Figure 2.8). Interestingly, the mRNA levels of *wrt-10* were greatly increased (~21-fold, Figure 2.8) in response to *rncs-1* overexpression in wildtype, and also showed ~5-fold upregulation in the *dcr-1* mutant (tabulation Figure 2.8). Including *wrt-10*, 4 out of 7 (6 out of 7 if T07C5.1 and F35D11.3 are included) somatic Dicer-regulated genes, showed enrichment

Figure 2.8. **mRNA levels of intestinal male-enriched genes in response to *rncs-1* levels and Dicer.** Graphs: relative expression of *mab-3* and *wrt-10* for the following strains: wildtype (N2, grey), *rncs-1(tm1632)* (*rncs-1*, white), *rncs-1(tm1632)* overexpressing *rncs-1 (rncs-1(+)*, grey on white hatch), wildtype overexpressing *rncs-1* (N2(+); black on grey hatch) and wildtype with P_{rncs-1}:*GFP* reporter (N2(gfp); black on grey checkerboard). The average of three independent samples for each strain is plotted normalized to *let-413* levels. Error bars, S.E.M. Tabulation below graphs: Expression of *mab-3* and *wrt-10* relative to wildtype hermaphrodites in *dcr-1(ok247)* mutants and in samples of wildtype males as quantified by real time RT-PCR. The average of three to four independent samples is listed. ***, P < 0.001, t-test.

Relative mRNA/let-413 level

mab-3

wrt-10

N2   rncs-1   rncs-1 (+)   N2 (+)   N2 (gfp)

Fold increase in *dcr-1*   1.2   4.6 ***

**Fold enrichment in males**   59.8 ***   20.8 ***

in males. While only circumstantial, this correlation suggests *C. elegans* Dicer may function to suppress somatic male-enriched genes in hermaphrodites.

F53A9.2 siRNAs are reduced upon rncs-1 overexpression. Our *in vitro* data suggested that *rncs-1* inhibits production of siRNAs from long dsRNA (Figure 2.6). To explore whether this was the mechanism by which *rncs-1* modulates expression of Dicer-regulated genes *in vivo*, we monitored siRNAs of F53A9.2, the gene most sensitive to *rncs-1* levels. Using sense-oriented DNA oligo probes of F53A9.2 sequence, we detected northern signals for small RNAs of ~19 and ~23 nt (lane 1, Figure 2.9). Occurrence of these siRNAs was dependent on Dicer, as their signals were reduced by ~80% in *dcr-1(ok247)* samples (lane 2, Figure 2.9). Consistent with our model, F53A9.2 siRNAs were reduced in strains overexpressing *rncs-1* (compare lanes 3-5, Figure 2.9). Reduction of siRNAs was less when *rncs-1* was overexpressed in the *rncs-1(tm1632)* mutant compared to overexpression in the wildtype background (~30% compared to ~60% reduction, Figure 2.9B); this correlates with the degree of upregulation of F53A9.2 mRNA in these strains (Figure 2.7B). We did not detect changes in F53A9.2 siRNA production in the *rncs-1(tm1632)* mutant (data not shown), likely because of the limited sensitivity of the northern analysis.

In addition to its role in RNAi, Dicer is required for processing microRNA precursors (pre-miRNAs) into mature miRNAs, such as *let-7* (45). *let-7* miRNA is expressed in seam cells, beginning at L4 and continuing through adulthood (46). *rncs-1* is not expressed in seam cells, so we used northern analyses to confirm that inhibition of Dicer activity by *rncs-1* is tissue specific. As expected, we saw accumulation of pre-*let-7* in worms deficient for Dicer (lane 7, Figure 2.9). However, in contrast to the F53A9.2

Figure 2.9. ***rncs-1* dependence of F53A9.2 siRNA production.** (A) Northern blots of 1μg (lanes 1-2) or 3μg (lanes 3-5) of size-selected RNA from indicated strains probed for small RNAs complementary to F53A9.2 (see Experimental Procedures). The blots were reprobed for *let-7* miRNA and its precursor (pre; lanes 6-10) and for U6 RNA as loading control (panels below lanes 1-5). (B) Three independent northern experiments were quantified for levels of F53A9.2 siRNAs relative to wildtype levels. Error bars, S.E.M.; **, P < 0.01, t-test; strain designations as in Figure 2.8.
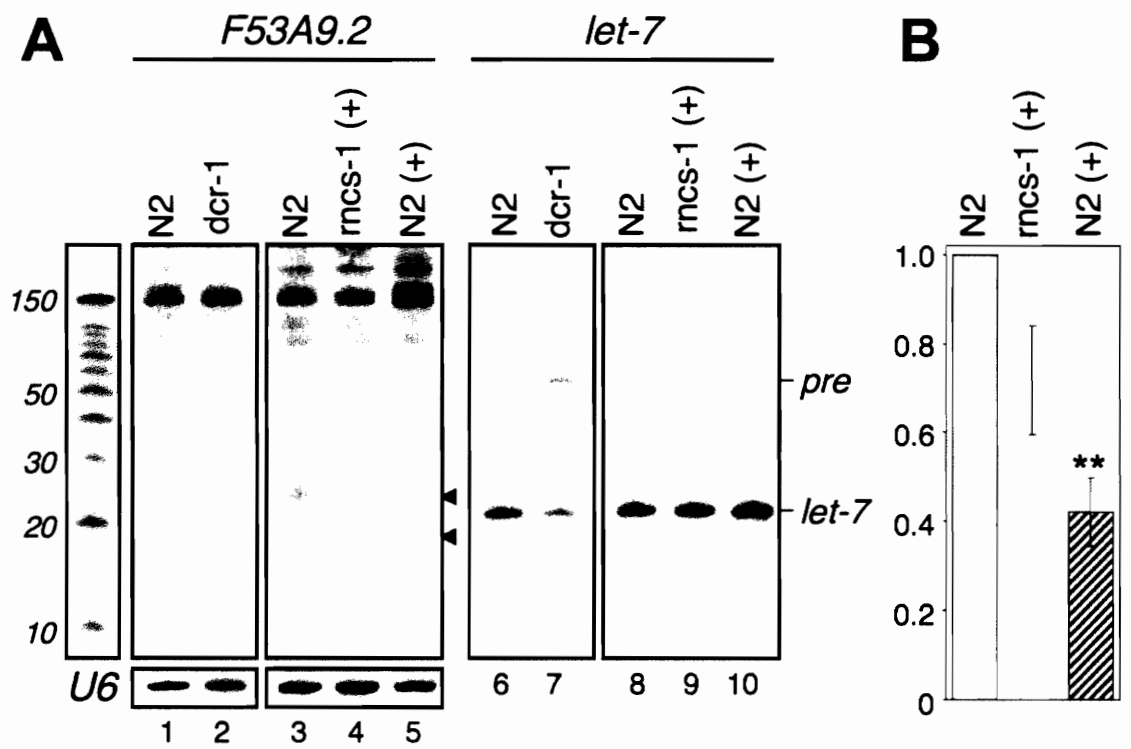
siRNAs, levels of mature *let-7* were not changed upon overexpression of *rncs-1* (lanes 9 and 10, Figure 2.9). Regulation of Dicer activity by *rncs-1*, therefore, is limited to the tissue of *rncs-1* expression or does not interfere with miRNA processing.

Discussion

In the experiments described in this chapter we have characterized *rncs-1*, a ncRNA whose double-stranded structure is unprecedented among characterized ncRNAs. While *rncs-1* is novel among cellular RNAs, as discussed below, its relationship to stress reiterates properties of other ncRNAs. Importantly, we found that changes in *rncs-1* expression lead to corresponding changes in expression of Dicer-regulated genes. In this capacity *rncs-1* provides the first example of a cellular RNA proven to modulate activity of a dsRBP. However, as detailed below, viruses have long been known to employ such strategies. In our analysis of Dicer-regulated genes, we found that siRNAs of F53A9.2 could be directly linked to Dicer activity. While many endogenous siRNAs have been identified (47), few have been shown to be Dicer-dependent, and our study is also noteworthy for this observation.

Is rncs-1 a mediator of stress response in C. elegans? An emerging theme of prokaryotic and eukaryotic noncoding RNAs is their role in stress response pathways. In bacteria, small ncRNAs, such as OxyS, RyhB or DsrA, modulate mRNA stability and translation as part of stress response systems that regulate oxygen and nutrient homeostasis, quorum sensing and temperature shock (48). Heat shock response, probably the best-understood stress pathway in eukaryotes, involves regulatory ncRNAs in a variety of organisms. In *Drosophila*, heat shock induces the ncRNA hsr-omega, which sequesters RNA processing factors in the nucleus, a function fulfilled by repetitive sat III

ncRNAs in human cells (49). Also inducible by heat shock, the small mouse B2 ncRNA acts to repress transcription by RNA polymerase II (50,51), and elimination of the noncoding transcript HSR1 renders mammalian cells thermosensitive (52).

Stress-regulated ncRNAs are largely uncharacterized in *C. elegans*. A large-scale analysis of genes enriched in dauer worms identified the ncRNA *tts-1* (transcribed telomerase-like sequence) (53), and later reports noted significant upregulation of *tts-1* RNA in *daf-2* mutants, during L1 arrest and bacterial infection (54,55). Deletion of *tts-1*, which like *rncs-1* is encoded by the X chromosome, reveals no essential function for this ncRNA (55).

Similarly, while strong induction of *rncs-1* occurred in the absence of a food supply, we did not observe obvious defects in the *rncs-1(tm1632)* deletion strain. The mutant exhibited wildtype behavior in conventional assays of dauer formation, aging or starvation (data not shown). Possibly, *rncs-1* supplies a redundant function during stress, and other ncRNAs, similarly regulated by food supply, can substitute for its function.

In bacteria, transcription of stress-induced ncRNA is orchestrated by the same factors that regulate protein-coding genes during stress, and it is intriguing to consider that ELT-2 acts similarly. While ELT-2 is known to induce innate immunity genes during bacterial infection (29,30), a function for ELT-2 during starvation stress has, to our knowledge, not been suggested. The insulin signaling pathway, mediated by *daf-2* and *daf-16*, upregulates expression of target genes during starvation (56), but *rncs-1* expression and induction is unchanged in *daf-2* or *daf-16* mutants (data not shown). Our data raise the possibility that ELT-2, directly or indirectly, provides an alternative starvation response, and perhaps, by analogy to the bacterial systems, one that involves

induction of ncRNA as well as proteins. Further dissection of the *rncs-1* promoter sequence could be helpful in characterizing the essential starvation response elements and could lead to identification of other genes, possibly encoding ncRNAs, that are regulated by a common mechanism.

Is rncs-1, or dsRNA in general, an indicator of cellular stress? A surprising finding of our work was that overexpression of *rncs-1* increased the frequency of males among hermaphrodite self-progeny. Occurrence of males (XO genotype) within offspring of virgin hermaphrodites (XX) is the result of rare X-chromosome nondisjunction. Spontaneous male frequency is increased by subjecting hermaphrodites to heat stress at the L4 larval stage (57), but the mechanisms underlying this phenomenon are unclear. *rncs-1* is not induced by heat shock, and heat stress increases male frequency in *rncs-1* mutant hermaphrodites to a similar degree as in wildtype (data not shown). Thus, it is unlikely that *rncs-1* upregulation is involved in elevated male frequency during heat shock.

Our current model for the observed increase in X nondisjunction in lines overexpressing *rncs-1* is that accumulation of this noncoding RNA nonspecifically and, in the case of our transgenic lines, nonphysiologically, signals stressed conditions to the organism. Initiation of genetic exchange by switching from asexual to sexual reproduction is a common strategy for adaptation to challenging environments (58). Increasing the frequency of males in *C. elegans* populations is an important prerequisite for sexual reproduction, and it is feasible that heat shock and other stresses induce yet unclear mechanisms to increase X nondisjunction in the hermaphrodite germline. This model would be strongly supported by experimental evidence showing induction of male

frequency under starvation stress, possibly in a *rncs-1*-dependent manner. Transient

starvation of L4 hermaphrodite larvae, using a similar approach employed for male

enrichment by heat shock, did not result in increased male frequency, even in wildtype

(data not shown). We are currently testing other starvation protocols.

Interestingly, we also observed induction of male frequency, albeit of lesser

degree, when GFP mRNA was expressed as a transgene under control of the *rncs-1*

promoter. Because low levels of dsRNA are synthesized from this transgene, it is

possible that dsRNA of any sequence, not just *rncs-1*, can signal stress.

In the mammalian interferon-mediated innate immune response, long dsRNA,

indicative of viral infection, plays a direct role in protein phosphorylation of Protein

Kinase PKR and activation of Oligoadenylate Synthetase OAS, which culminates in cell

death through translation inhibition and RNA degradation (59). This rigorous reaction to

dsRNA is thought to be the reason why gene targeting by RNAi using long trigger

dsRNA is impossible in mammalian tissue culture (60). Nematodes do not have an

interferon response, and long dsRNA works well to trigger RNAi. While it is generally

believed that dsRNA does not induce a stress response in *C. elegans*, our data suggest

that accumulation of dsRNA may, indeed, signal stress in nematodes, perhaps with more

subtle consequences by comparison to mammals.

Does endogenous dsRNA function to fine-tune dsRBP activity? Recent global

analyses of eukaryotic transcriptomes reveal that production of endogenous dsRNA is

common (61). For example, transcription from both strands of the genome, often in the

form of ncRNA, is thought to enable at least 20% of human transcripts to form sense-

antisense pairs (62), and greater than 70% of mouse RNAs overlap at least partially with

transcripts in opposite orientation (63). In *C. elegans*, genomic inverted repeat (IR) structures, formed by paralogous gene pairs, provide a similar source of potential dsRNA (64). An estimated 3.6% of the *C. elegans* genome consists of IRs up to 2000bp in length (65), and while their transcription is largely undocumented, identification of extended ADAR-edited dsRNA structures in the UTRs of various *C. elegans* mRNAs suggests that at least some IRs are part of stable RNA transcripts and provide an additional source of endogenous dsRNA (7,8).

dsRBPs interact with dsRNA in a sequence-independent manner (40), making it possible for one dsRNA to sequester protein binding partners away from other dsRNAs. Along these lines, certain viruses produce dsRNAs to interfere with dsRNA-mediated host defenses. For example, Epstein-Barr virus $EBER_I$ RNA, and human adenovirus VA_I RNA, are virally-produced RNAs that inactivate PKR to evade host cell death (66). In addition, adenovirus VA_I and VA_{II} function as competitive substrates to inhibit host Dicer activity (67).

Here we present evidence that an endogenous dsRNA can function to fine-tune *C. elegans* Dicer activity *in vitro* and *in vivo*, in a manner reminiscent of the adenoviral-associated RNAs. Our data suggest a mechanism by which Dicer-regulated gene expression is controlled in a tissue-specific manner. In tissues with high concentrations of endogenous dsRNA like *rncs-1*, Dicer would be less potent in silencing gene expression, allowing for more robust expression of its targets.

The efficacy by which a competitive binding substrate like *rncs-1* modulates Dicer activity depends on the relative concentration of the dsRNA it is competing with. In our *in vitro* experiments, *rncs-1* only inhibited cleavage of dsRNA substrates when

present at significant molar excess (Figure 2.6). While we do not know the concentration or identity of the endogenous dsRNAs that *rncs-1* competes with, this may explain why *rncs-1* mutants do not show a RNAi-hypersensitive phenotype as observed in animals with mutations in genes for negative regulators of RNAi like *eri-1* or *rrf-3* (68). Conventional methods of RNAi, such as feeding or injection, introduce high concentrations of exogenous dsRNA, which likely overshadow any competition by *rncs-1*. It is also possible that the origin of the dsRNA substrate is important. The efficacy of *rncs-1* as an antagonist of Dicer may require the competing dsRNA Dicer substrate to be transcribed in the nucleus, whereas RNAi by feeding or microinjection introduces dsRNA to the cytoplasm.

It is yet unclear whether the molecular phenotype of fine-tuning endogenous RNAi has any relevance for the proposed role of *rncs-1* as a mediator of stress response in *C. elegans*. Interestingly, mutations in several genes involved in RNAi, transgene silencing and co-suppression—including *mut-7*, *rde-2*, *rde-3*, *mut-14*, *eri-1* and *rrf-3*—display frequent X chromosome nondisjunction (69,70). Of those genes, *rde-2*, *rde-3*, *mut-7*, *mut-14*, *rrf-3* and *eri-1* are required for Dicer-dependent production of endogenous siRNAs to the predicted gene C44B11.6 (47). In addition, ERI-1 and RRF-3 are found in complex with Dicer and are required for production of Dicer-dependent endogenous siRNAs complementary to the somatic gene K02E2.6 (69). While not yet validated by northern analyses, other endogenous siRNAs are likely to have a similar requirement for Dicer cofactors. Future efforts will be directed at characterizing expression of the Dicer-regulated and *rncs-1*-sensitive genes during starvation and other stresses, as well as defining a general role of RNAi in nematode stress pathways.

Endogenous dsRNAs: the sequence versus structure conundrum. Taken together, our data suggest that *rncs-1* functions are largely mediated by its double-stranded character, rather than its sequence. Therefore, it is not surprising that sequence searches have not revealed homologs of *rncs-1*. Further, for unknown reasons, long ncRNAs, including essential transcripts like mammalian *Xist*, are poorly conserved in sequence, even in closely related species (71). Thus, conventional approaches to identify functional analogs of ncRNAs like *rncs-1* are ineffective. We hope that searches based on structure conservation or identification of common promoter elements may prove more successful.

The *C. elegans* genome is rich in long inverted repeat (LIR) elements, with an average of ~125 LIRs occurring per megabase (72). The *rncs-1* gene is among the 20% of LIRs that contain a stem longer than 100bp. In contrast to the majority of LIRs in C. elegans, *rncs-1* is unique in sequence and not repeat-derived.

Inverted DNA repeats with highly base-paired stems are recombinogenic and cause genome instability in eukaryotes (73,74). This could explain the absence of an identifiable *rncs-1* homolog from the genomes of *C. briggsae* and *C. remanei*, two species of nematodes closely related to *C. elegans*. Indeed, in *C. briggsae* and *C. remanei* the region of the X chromosome syntenic with the *rncs-1* neighborhood underwent an inversion, with a putative genomic breakpoint close to the current position of *C. elegans rncs-1*. The fact that *C. elegans* has retained an unstable element in its genome hints at a selective advantage provided by the *rncs-1* gene product.

References

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

2.   Hubbard, T.J.P., Aken, B.L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T. *et al.* (2006) Ensembl 2007. *Nucl. Acids Res.*, gkl996.

3.   Frith, M.C., Pheasant, M. and Mattick, J.S. (2005) The amazing complexity of the human transcriptome. *Eur J Hum Genet*, **13**, 894-897.

4.   Prasanth, K.V. and Spector, D.L. (2007) Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev*, **21**, 11-42.

5.   Huttenhofer, A. and Vogel, J. (2006) Experimental approaches to identify noncoding RNAs. *Nucleic Acids Res*, **34**, 635-646.

6.   Mattick, J.S. and Makunin, I.V. (2006) Noncoding RNA. *Hum Mol Genet*, **15 Spec No 1**, R17-29.

7.   Morse, D.P., Aruscavage, P.J. and Bass, B.L. (2002) RNA hairpins in noncoding regions of human brain and Caenorhabditis elegans mRNA are edited by adenosine deaminases that act on RNA. *Proc Natl Acad Sci U S A*, **99**, 7906-7911.

8.   Morse, D.P. and Bass, B.L. (1999) Long RNA hairpins that contain inosine are present in Caenorhabditis elegans poly(A)+ RNA. *Proc Natl Acad Sci U S A*, **96**, 6048-6053.

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

10.   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.

11.   Wang, J. and Kim, S.K. (2003) Global analysis of dauer gene expression in Caenorhabditis elegans. *Development*, **130**, 1621-1634.

12.   Mello, C. and Fire, A. (1995) DNA transformation. *Methods Cell Biol*, **48**, 451-482.

13.   Welker, N.C., Habig, J.W. and Bass, B.L. (2007) Genes misregulated in C. elegans deficient in Dicer, RDE-4, or RDE-1 are enriched for innate immunity genes. *RNA*.

14.   Harley, C.B. (1987) Hybridization of oligo(dT) to RNA on nitrocellulose. *Gene Anal Tech*, **4**, 17-22.

15.   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.

16. Bass, B.L. and Weintraub, H. (1987) A developmentally regulated activity that unwinds RNA duplexes. *Cell*, **48**, 607-613.

17. Parker, G.S., Eckert, D.M. and Bass, B.L. (2006) RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. *RNA*, **12**, 807-818.

18. Doudna, J.A. (1997) Preparation of homogeneous ribozyme RNA for crystallization. *Methods Mol Biol*, **74**, 365-370.

19. Sulston, J.E., Schierenberg, E., White, J.G. and Thomson, J.N. (1983) The embryonic cell lineage of the nematode Caenorhabditis elegans. *Developmental Biology*, **100**, 64-119.

20. Bossinger, O., Fukushige, T., Claeys, M., Borgonie, G. and McGhee, J.D. (2004) The apical disposition of the Caenorhabditis elegans intestinal terminal web is maintained by LET-413. *Dev Biol*, **268**, 448-456.

21. Croce, A., Cassata, G., Disanza, A., Gagliani, M.C., Tacchetti, C., Malabarba, M.G., Carlier, M.F., Scita, G., Baumeister, R. and Di Fiore, P.P. (2004) A novel actin barbed-end-capping activity in EPS-8 regulates apical morphogenesis in intestinal cells of Caenorhabditis elegans. *Nat Cell Biol*, **6**, 1173-1179.

22. Golden, J.W. and Riddle, D.L. (1984) The Caenorhabditis elegans dauer larva: developmental effects of pheromone, food, and temperature. *Dev Biol*, **102**, 368-378.

23. Jeong, P.Y., Jung, M., Yim, Y.H., Kim, H., Park, M., Hong, E., Lee, W., Kim, Y.H., Kim, K. and Paik, Y.K. (2005) Chemical structure and biological activity of the Caenorhabditis elegans dauer-inducing pheromone. *Nature*, **433**, 541-545.

24. Golden, J.W. and Riddle, D.L. (1982) A pheromone influences larval development in the nematode Caenorhabditis elegans. *Science*, **218**, 578-580.

25. Pauli, F., Liu, Y., Kim, Y.A., Chen, P.J. and Kim, S.K. (2006) Chromosomal clustering and GATA transcriptional regulation of intestine-expressed genes in C. elegans. *Development*, **133**, 287-295.

26. Maduro, M.F. and Rothman, J.H. (2002) Making worm guts: the gene regulatory network of the Caenorhabditis elegans endoderm. *Dev Biol*, **246**, 68-85.

27. Fukushige, T., Hawkins, M.G. and McGhee, J.D. (1998) The GATA-factor elt-2 is essential for formation of the Caenorhabditis elegans intestine. *Dev Biol*, **198**, 286-302.

28. McGhee, J.D., Sleumer, M.C., Bilenky, M., Wong, K., McKay, S.J., Goszczynski, B., Tian, H., Krich, N.D., Khattra, J., Holt, R.A. *et al.* (2007) The ELT-2 GATA-factor and the global regulation of transcription in the C. elegans intestine. *Dev Biol*, **302**, 627-645.

29. Shapira, M., Hamlin, B.J., Rong, J., Chen, K., Ronen, M. and Tan, M.W. (2006) A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. *Proc Natl Acad Sci U S A*, **103**, 14086-14091.

30. Kerry, S., Tekippe, M., Gaddis, N.C. and Aballay, A. (2006) GATA Transcription Factor Required for Immunity to Bacterial and Fungal Pathogens. *PLoS ONE*, **1**, e77.

31. Knight, S.W. and Bass, B.L. (2002) The role of RNA editing by ADARs in RNAi. *Mol Cell*, **10**, 809-817.

32. Hodgkin, J., Horvitz, H.R. and Brenner, S. (1979) Nondisjunction Mutants of the Nematode CAENORHABDITIS ELEGANS. *Genetics*, **91**, 67-94.

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

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

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

36. Parrish, S. and Fire, A. (2001) Distinct roles for RDE-1 and RDE-4 during RNA interference in Caenorhabditis elegans. *RNA*, **7**, 1397-1402.

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

38. 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.

39. Chang, K.-Y. and Ramos, A. (2005) The double-stranded RNA-binding motif, a versatile macromolecular docking platform. *FEBS Journal*, **272**, 2109-2117.

40. Tian, B., Bevilacqua, P.C., Diegelman-Parente, A. and Mathews, M.B. (2004) The Double-Stranded-RNA-Binding Motif: Interference and Much More. *Nature Reviews Molecular Cell Biology*, **5**, 1013-1023.

41. Tonkin, L.A. and Bass, B.L. (2003) Mutations in RNAi rescue aberrant chemotaxis of ADAR mutants. *Science*, **302**, 1725.

42. Nishikura, K. (2006) Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nat Rev Mol Cell Biol*, **7**, 919-931.

43. Bass, B.L. (2007) How does RNA editing affect dsRNA-mediated gene silencing? *Cold Spring Harbor Symposia on Quantitative Biology*, **71**, 285-292.

44. Thoemke, K., Yi, W., Ross, J.M., Kim, S., Reinke, V. and Zarkower, D. (2005) Genome-wide analysis of sex-enriched gene expression during C. elegans larval development. *Dev Biol*, **284**, 500-508.

45. 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.

46. 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.

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

48. Gottesman, S. (2005) Micros for microbes: noncoding regulatory RNAs in bacteria. *Trends Genet*, **21**, 399-404.

49. Jolly, C. and Lakhotia, S.C. (2006) Human sat III and Drosophila hsr omega transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells. *Nucleic Acids Res*, **34**, 5508-5514.

50. Allen, T.A., Von Kaenel, S., Goodrich, J.A. and Kugel, J.F. (2004) The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. *Nat Struct Mol Biol*, **11**, 816-821.

51. Espinoza, C.A., Allen, T.A., Hieb, A.R., Kugel, J.F. and Goodrich, J.A. (2004) B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nat Struct Mol Biol*, **11**, 822-829.

52. Shamovsky, I., Ivannikov, M., Kandel, E.S., Gershon, D. and Nudler, E. (2006) RNA-mediated response to heat shock in mammalian cells. *Nature*, **440**, 556-560.

53.   Jones, S.J., Riddle, D.L., Pouzyrev, A.T., Velculescu, V.E., Hillier, L., Eddy, S.R., Stricklin, S.L., Baillie, D.L., Waterston, R. and Marra, M.A. (2001) Changes in gene expression associated with developmental arrest and longevity in Caenorhabditis elegans. *Genome Res*, **11**, 1346-1352.

54.   Halaschek-Wiener, J., Khattra, J.S., McKay, S., Pouzyrev, A., Stott, J.M., Yang, G.S., Holt, R.A., Jones, S.J., Marra, M.A., Brooks-Wilson, A.R. *et al.* (2005) Analysis of long-lived C. elegans daf-2 mutants using serial analysis of gene expression. *Genome Res*, **15**, 603-615.

55.   O'Rourke, D., Baban, D., Demidova, M., Mott, R. and Hodgkin, J. (2006) Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of C. elegans with M. nematophilum. *Genome Res*, **16**, 1005-1016.

56.   Henderson, S.T. and Johnson, T.E. (2001) daf-16 integrates developmental and environmental inputs to mediate aging in the nematode Caenorhabditis elegans. *Curr Biol*, **11**, 1975-1980.

57.   Hodgkin, J. (1999) In Hope, I. A. (ed.), *C. elegans*. Oxford University Press, Oxford.

58.   de Visser, J.A. and Elena, S.F. (2007) The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat Rev Genet*, **8**, 139-149.

59.   Samuel, C.E. (2001) Antiviral actions of interferons. *Clin Microbiol Rev*, **14**, 778-809.

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

61.   Munroe, S.H. and Zhu, J. (2006) Overlapping transcripts, double-stranded RNA and antisense regulation: a genomic perspective. *Cell Mol Life Sci*, **63**, 2102-2118.

62.   Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z. and Rowley, J.D. (2004) Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res*, **32**, 4812-4820.

63.   Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J. *et al.* (2005) Antisense transcription in the mammalian transcriptome. *Science*, **309**, 1564-1566.

64. Zhao, G., Chang, K.Y., Varley, K. and Stormo, G.D. (2007) Evidence for active maintenance of inverted repeat structures identified by a comparative genomic approach. *PLoS ONE*, **2**, e262.

65. LeBlanc, M.D., Aspeslagh, G., Buggia, N.P. and Dyer, B.D. (2000) An annotated catalog of inverted repeats of Caenorhabditis elegans chromosomes III and X, with observations concerning odd/even biases and conserved motifs. *Genome Res*, **10**, 1381-1392.

66. McKenna, S.A., Kim, I., Liu, C.W. and Puglisi, J.D. (2006) Uncoupling of RNA binding and PKR kinase activation by viral inhibitor RNAs. *J Mol Biol*, **358**, 1270-1285.

67. Andersson, M.G., Haasnoot, P.C., Xu, N., Berenjian, S., Berkhout, B. and Akusjarvi, G. (2005) Suppression of RNA interference by adenovirus virus-associated RNA. *J Virol*, **79**, 9556-9565.

68. Timmons, L. (2004) Endogenous inhibitors of RNA interference in Caenorhabditis elegans. *Bioessays*, **26**, 715-718.

69. 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.

70. Grishok, A. (2005) RNAi mechanisms in Caenorhabditis elegans. *FEBS Lett*, **579**, 5932-5939.

71. Pang, K.C., Frith, M.C. and Mattick, J.S. (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet*, **22**, 1-5.

72. Wang, Y. and Leung, F.C. (2006) Long inverted repeats in eukaryotic genomes: recombinogenic motifs determine genomic plasticity. *FEBS Lett*, **580**, 1277-1284.

73. Gordenin, D.A., Lobachev, K.S., Degtyareva, N.P., Malkova, A.L., Perkins, E. and Resnick, M.A. (1993) Inverted DNA repeats: a source of eukaryotic genomic instability. *Mol Cell Biol*, **13**, 5315-5322.

74. Lobachev, K.S., Stenger, J.E., Kozyreva, O.G., Jurka, J., Gordenin, D.A. and Resnick, M.A. (2000) Inverted Alu repeats unstable in yeast are excluded from the human genome. *EMBO J*, **19**, 3822-3830.

CHAPTER 3

SUBCELLUAR LOCALIZATION OF *C. ELEGANS* ADAR

SUBSTRATES IN WILDTYPE EMBRYO

CELL FRACTIONS

Introduction

It is widely assumed that inosine-containing RNAs are retained in the nucleus. This assumption is based on two lines of evidence. First, dsRNA originating from the pairing of early and late transcripts of mouse polyoma virus is promiscuously edited by host ADARs and subsequently excluded from the cytoplasm (1). Second, synthetic inosine-containing dsRNA and ssRNA is edited and retained in the nucleus of *Xenopus* oocytes after microinjection (2).

Both of these studies originated in the laboratory of Gordon Carmichael, and additional work from his group proposed a model for how nuclear export of inosine-containing RNA is prevented. Crosslinking experiments and affinity chromatography identified a protein complex—consisting of the multifunctional nuclear protein $p54^{nrb}$, the splicing factor PSF, and the nuclear matrix component matrin 3—that cooperatively binds to long inosine-containing RNA in HeLa cell nuclear extract (2). Further analyses indicated that $p54^{nrb}$ and PSF make direct contacts with the RNA, while matrin 3 is associated via protein-protein interactions to confer cooperativity and anchor the

ribonucleoprotein complex to the nuclear matrix. Interestingly, closer examination of the proteins associating with inosine-containing RNA in HeLa nuclear extract revealed an additional protein complex, consisting of the RNA-binding protein vigilin, the Ku70/86 autoantigen, RNA helicase A, the kinase DNA-PKcs and ADAR1 (3). In the nucleus, components in this complex participate in heterchromatic gene silencing, DNA repair and chromosome segregation, suggesting possible new roles for ADARs and deaminated RNA in the nucleus.

While the findings above are intriguing, several caveats need to be considered before generally applying the nuclear retention model to cellular inosine-containing RNAs. First, all of the substrates used were viral or synthetic RNAs. At this point, no cellular promiscuously edited transcripts were found to be associated with the p54[nrb]/PSF/matrin 3 complex. The group of David Spector recently identified a cellular transcript, the 8kb CTN-RNA, which is retained in the nucleus (4). The 3'UTR of CTN-RNA contains inverted repeat sequences that form double-stranded structures and are edited by ADARs, introducing between 7 and 12 inosines to the >4.5kb 3'UTR. Because CTN-RNA colocalized and coimmunoprecipitated with p54[nrb], it was suggested that nuclear retention of the RNA was facilitated by the inosine-binding complex identified by Carmichael, making CTN-RNA the first cellular example for this mechanism of retention. However, no experimental data in support of this assumption—such as showing nuclear retention in the absence of ADARs—were presented (5).

Second, the polyoma viral and synthetic RNAs used by the Carmichael group to demonstrate nuclear retention and identify the putative retention complex were hyperedited to ~50% A-to-I changes. Hyperedited cellular RNAs, such as those

identified in our laboratory (6,7), rarely show editing levels >20%. This is likely due to the fact that editing to >40% requires long perfectly double stranded stretches of dsRNA, while the dsRNA regions occurring as part of cellular transcripts typically contain mismatches, loops and bulges (Figure 1.1).

Third, several ADAR substrates in various organisms, such as the mammalian gluR-B and *C. elegans pop-1* mRNAs, are essential for development and survival and clearly make it into the cytoplasm to be translated (8). The requirement for translating certain edited substrates is in conflict with a model of ubiquitous nuclear retention for inosine-containing RNAs. Carmichael suggests that the answer to this conundrum lies in the kinetics of RNA binding by the p54[nrb]/PSF/matrin 3 complex (2). He proposes that tight binding of inosine-containing RNA requires cooperativity. Assembly of higher ordered cooperative ribonucleocomplexes requires binding of p54[nrb] along an extended stretch of hyperedited RNA. Selectively edited substrates, containing few or widely spaced inosines, would not allow for tight association with the retention complex, allowing their export from the nucleus. It is questionable whether this mechanism could account for the inosine-mediated retention suggested for the CTN-RNA, as very few inosines are found in the 3'UTR of this message. Determinants other than number of inosines may play a role in selecting inosine-containing RNAs for export or retention. Intriguingly, Spector's studies of CTN-RNA offered an alternative mechanism by which messenger RNAs that are retained in the nucleus could be exported for translation. He finds that under stressed condition, part of the CTN-RNA 3'UTR is removed, allowing the mRNA to the exported to the cytoplasm for translation (4). It is possible that a

similar mechanism controls the expression of cellular messenger RNAs with hyperedited structures in their 3'UTR.

In the studies described in this chapter we attempted to establish whether the hyperedited *C. elegans* ADAR substrates 36A and *rncs-1* are retained in the nucleus in the presence of wildtype ADAR activity. We chose a cell fractionation approach that would allow us to examine the subcellular localization of several edited and unedited transcripts in nuclear and cytoplasmic populations of RNA. We were able to confirm that the ADAR substrate 36A, like unedited control messages, was efficiently exported to the cytoplasm and copurified with the polysomal RNA population. In addition, sequence analyses confirmed that the exported 36A population was edited to a degree exceeding editing of the nuclear population. Hyperedited *rncs-1* was found enriched in the cytosol, although limitations of the experimental system did not allow for clear interpretation of this observation. We conclude that the proposed model of inosine-mediated nuclear retention of RNA does not apply to 36A and, possibly, *rncs-1*. Plausible reasons are discussed.

Experimental procedures

Cell fractionation of C. elegans embryos. To separate cellular compartments a differential centrifugation protocol from the lab of Robert Horvitz was adapted (9). Embryos were prepared by hypochlorite treatment and washed extensively (four to five times) with excess M9 buffer, followed by one wash with chilled homogenization buffer (10mM KCl, 1.5mM $MgCl_2$, 1mM EDTA, 1mM EGTA, 1mM DTT, 250mM sucrose; sterile filtered; chilled at 4°C overnight). Embryos were pelleted in a 50mL conical tube at 1,200g. Typically, a 3-5mL pellet of embryos was used.

The following steps were performed on ice and RNase-free glassware was used when possible. Five volumes of homogenization buffer were added to the embryo pellet. Complete protease inhibitors –EDTA (Roche) were added to 1X final concentration, and RNase inhibitor (e.g. RNaseOUT, Invitrogen) was added to a final concentration of 0.5-1.0U/µL. Embryos were gently resuspended by inverting the tube.

Working in small batches (5-6mL), the embryo slurry was transferred to a chilled 15mL Wheaton steel dounce homogenizer, set in a bucket of ice. To break down whole embryos and cells, 13-15 strokes were applied. To verify lysis, an aliquot was examined by microscopy and DAPI staining. The lysate contained whole embryos and cells but also free nuclei. The homogenate was transferred to a 30mL glass COREX tube, set on ice. Lysates of all homogenized batches were pooled. An aliquot of this crude whole cell lysate was taken, flash-frozen in liquid nitrogen and stored at -80°C for analysis.

To eliminate unlysed embryos, the lysate was centrifuged in a Beckman JA25.50 rotor at 100g for 2 min at 4°C. The supernatant was transferred to a fresh COREX tube; the pellet was discarded. The low-speed centrifugation was repeated two more times, increasing the centrifugal force to 200g for 2 min at 4°C followed by 271g (1500rpm) for 3 min at 4°C. The supernatant of the last centrifugation step—largely depleted of whole cells or embryos—was analyzed by microscopy and DAPI staining and had a granular appearance due to released nuclei. As an alternative to the crude cell lysate obtained immediately after homogenization, an aliquot of this cell-free whole cell lysate was on occasion saved for analysis. There was no remarkable difference in protein, DNA or RNA content between both types of whole cell lysate.

As a next step, nuclei were pelleted by centrifugation at 750g (2,500rpm) for 10 min at 4°C. The pellet was reserved, while the supernatant was subjected to a repeated centrifugation step at 750g for 10 min at 4°C. Nuclei pellets were usually combined, resuspended in a minimal amount of homogenization buffer, flash-frozen and saved for analysis.

The supernatant from the previous step was transferred to 16x76mm polyallomer ultracentrifuge tubes (Beckman), dividing the sample and topping with homogenization buffer, if necessary. The sample was centrifuged in a 90Ti rotor (rotor and centrifuge prechilled) at 100,000g for 60 min at 4°C. The supernatant (S100) and pellet (P100, resuspended in a minimal amount of homogenization buffer) were flash-frozen and saved.

Analysis of DNA content of cell fractions. DNA was isolated from the interphase and phenol phase of the Trizol RNA extraction following the manufacturer's instructions (Invitrogen). The DNA was resuspended in HEPES buffer. 1/10,000 of total DNA from lysate, nuclei, S100 and P100 was analyzed by semiquantative PCR using primers SH17 (CCACCGCCGCTTCTTCCTCT) and SH18 (GATGGGCCGGACTCGTCGTA), which amplify a 696bp fragment of genomic *act-1* DNA. Aliquots of the reaction were removed after 15, 20, 25 and 30 cycles and analyzed by agarose gel electrophoresis.

Analysis of protein content of cell fractions by western blot and activity assays. For Western analyses, equal fractions of lysate, nuclei, S100 and P100 were run on a 4-15% Tris-HCl gel (Biorad). Proteins were transfered onto an Immobilon-P (PVDF) membrane (Millipore) in Towbin buffer (25mM Tris-HCl, 192mM glycine, 20% methanol) using an electrophoresis transfer cell (110V, 45 min). The membrane was

blocked by shaking for 1 hr in 3% BSA in TTBS (20mM Tris-HCl pH7.5, 150mM NaCl, 0.1% Tween 20) before adding primary antibody to 1μg/mL and incubating for an additional hour. For detection of Hsp90 a mouse monoclonal antibody (Stressgen, SPA-830D) was used; for detection of histone H3, a goat polyclonal antibody (Santa Cruz Biotechnology, sc-8654). The blot was washed twice for 10 min in TTBS and then incubated with a 1:10,000 dilution of secondary antibody in 3% BSA for 1 hr. Anti-mouse alkaline phosphatase (AP) coupled IgG (Sigma, A4312) and anti-goat AP-coupled IgG (Santa Cruz Biotechnology, sc-2771) were used in the Hsp90 and H3 analyses, respectively. The blot was washed twice for 10 min in TTBS and developed using SIGMA *FAST*™ BCIP/NBT tablets.

For enzyme activity assays ^{32}P-labeled dsRNA (CAT duplex) was prepared by *in vitro* transcription as described in Chapter 2. For deamination assays the buffer of the cell fractions was adjusted to reflect salt and glycerol concentrations suitable for ADAR assays (final: 20mM Tris-HCl pH7.5, 40mM KCl, 5% glycerol, 0.005% NP-40, 0.5mM DTT). Deamination assays were performed as published (10). In each 30μL reaction, 7.5μg of total protein from lysate, nuclei, S100 and P100 and 0.5nM dsRNA substrate were used, and reactions were incubated for 1 hr at 20°C. Dicer activity assays, using 15μg of total protein, were performed as described in Chapter 2.

RNA analysis in cell fractions. RNA from whole cell lysate, nuclei, S100 and P100 was isolated and processed using methods described in Chapter 2. Poly(A)+ RNA was prepared from total RNA using Oligotex® mRNA spin column chromatography (Qiagen). Northern analyses of total and poly(A)+ RNA were performed using standard protocols. ^{32}P-labelled probes were synthesized using the Strip-EZ® RNA T7 system

(Ambion) from cDNA PCR products. Table 3.1 lists the primer sequences used to generate all probe transcription templates.

For RT-PCR analysis cDNA was prepared using random decamer primers (Ambion) and ThermoScript™ reverse transcriptase (Invitrogen). The use of random decamers instead of oligo-dT primers in the reverse transcription allowed for PCR analysis of nonpolyadenylated RNA, such as pre-rRNA, and prevented transcriptional bias against highly structured messages like 36A. cDNA was prepared from 5μg RNA per 20μL reaction; RT reactions were treated with RNaseH (New England Biolabs) and diluted 4-fold prior to PCR analysis. 5μL of cDNA samples were analyzed per quantitative PCR in a LightCycler® 2.0 instrument using the LightCycler® FastStart DNA Master^PLUS SYBR Green I Kit. Primer pairs were designed to span at least one exon-exon junction and produce products of 100 to 350bp. Primer pairs used for RT-PCR are listed in Table 3.1.

Quantification of absolute RNA distribution. Absolute cellular total RNA distribution was calculated from the mass of total RNA isolated from each fraction as determined by spectrophotometry. Similarly, total RNA >200nt (selected by RNeasy® chromatography) and poly(A)+ distribution was determined based on mass, but because only part of the total RNA sample was used in RNeasy® chromatography or poly(A)+ selection, the absolute distribution had to be extrapolated. For example, if 2000μg of total RNA was isolated from nuclei and poly(A)+ preparation from 500μg yielded 5μg of poly(A)+ RNA, the extrapolated absolute amount of poly(A)+ in nuclei would be 2000μg /500μg *5μg = 20μg.

Table 3.1. **Primer sequences used for synthesis of northern probe templates and for quantitative PCR analysis.**

RNA	5' primer	3' primer	Product size (bp)
Primer pairs for synthesis of northern probe templates			
pre-rRNA	ITS1U AACCT TAAGA TGCTC GACTG GCTTC ACGG	ITS1LT7 *TAATA CGACT CACTA TAGGC* CGAAA CCGAA CCACG ATCAT CAAGA	333
uaf-1(ex5)	uaf1ex5U2 ACACT CCTGC TGTCC ACCAG CCAC	uaf1ex5LT7 *TAATA CGACT CACTA TAGGC* CTGAA TTTTG CAAGA TTATT TGGGC TTCCA	235
act-1	act-1U GACGA GGTTG CCGCT CTTGT TGTA	act-1LT7 *TAATA CGACT CACTA TAGGA* AGCGT GAGG GAGAG GACAG	446
36A	36Ap0.u CCTGA GTTTT CACGA AACGA	36Ap0.l2 *TAATA CGACT CACTA TAGGA* TTGCT TGCTC AGTTT	528
mcs-1	52Gp1.u GCCGT GGTCG ATGCT GAAAG GCA	52Gp1.l *TAATA CGACT CACTA TAGGA* AAGGG AATTT CACAC T	138
Primer pairs used in quantitative RT-PCR analysis			
pre-rRNA	ITS1U AACCT TAAGA TGCTC GACTG GCTTC ACGG	ITS1L GCCGA AACCG AACCA CGATC ATCAA GA	315
uaf-1(ex5)	uaf-1ex5LCU TCTGC CCAAC TATTT GACTG AGGAT	uaf-1ex5LCL TGGGC TTCCA GTTGG GAGAT TT	223
gpd-3	gpd3 For GGAGG AGCCA AGAAG GTC	gpd3 Rev AAGTG GAGCA AGGCA GTT	144
act-1	SH17 CCACC GCCGC TTCTT CCTCT	SH18 GATGG GCCGG ACTCG TCGTA	421
uaf-1	uaf-1LCU CTCGT TGGAA ATTCC AAGAC CGTAT GA	uaf-1LCL TTCGT CCGGT CAGAG CAGCC TG	119
36A	36ALCU GGAAG AGTTT TGAAA ATGTG TTATG ATGA	36ALCL GATTT TGTTG GTTTT CTTGT TCTTA GCA	231

Extrapolation was also used to calculate the distribution of RNAs from northern blots and quantitative RT PCR. For instance, *rncs-1* levels in 20μg total RNA per fraction isolated in one experiment were analyzed by northern blot and band intensity was quantified using a Molecular Dynamics PhosphorImager. Measured *rncs-1* band volumes were 440875, 2636356, 162007 for nuclei, S100 and P100, respectively. In this experiment, the absolute amount of total RNA isolated per fraction was 1210μg, 1432μg, 2995μg (nuclei, S100, P100). The theoretical band volumes—measured if the entire RNA samples had been analyzed by northern blot—calculate as 440875/(20μg /1210μg) = 26672911 for nuclear *rncs-1*, 2636356/(20μg /1432μg) = 188763058 for *rncs-1* in S100, and 162007/(20μg /2995μg) = 24260517 for *rncs-1* in P100. Therefore, the percentage of all cellular *rncs-1* cofractioning with the S100 compartment in this experiment was 188763058/(26672911 + 188763058 + 24260517) * 100% = 78.8%. For RNA species that could be quantified by northern analyses and RT-PCR, using results of either method in calculating the absolute cellular distribution led to comparable results.

Determination of editing levels of 36A and rncs-1 by population sequencing. 5μg of total RNA, purified by DNase treatment and RNeasy® chromatography, were used in the cDNA reactions described below. Control reactions, lacking reverse transcriptase enzyme, were included for all RNA samples.

To amplify the 5' strand of the inverted repeat structure in the 3'UTR of 36A, total RNA was reverse transcribed by nested PCR using primer 36ARTL6 (TTTGGCATAGCCTACAATTGACGGA) and ThermoScript™ enzyme. Following manufacturer's instructions (Invitrogen), 20μL reactions were set up and incubated at 53°C for 1 hr before heat-inactivation at 85°C for 10 min. 1μL RNase H (New England

Biolabs) was added to each reaction, and incubation was continued for 20 min at 37°C followed by 5 min at 80°C. cDNA was amplified by nested PCR, using primers 36ARTU0 (GAGAATCCTGTCAAGAAAAATAATGACTACA) and 36ARTL6 for the first PCR and 36ARTU6 (GGAGCAGTAGCCGTAGCAGCCT) and 36ARTL5 (ATTCTCAATTCACAGTACATTTCCAAGATT) for the nested reaction.

To amplify the 3' strand of *rncs-1*, reverse transcription using primer 52GRTL2 (TGATTCAACATTTCAAAAACTTGTATTTTACATCTAAAACTATAAA) was performed as described for 36A, except that the reaction was performed at 60°C. One round of PCR using 52GRTU2 (ATTTTTTCCCGACAAAGATGGAACTCAAGGAT) and 52GRTL2 followed. Because the *rncs-1* PCR products were contaminated with primer-dimer products, the PCR product was purified by agarose gel extraction.

All primers used in the RT-PCR scheme map to regions of the RNA not edited by ADARs. The resulting cDNA sample represents a mixture of cDNAs that mirror the mixture of edited and unedited species found in each RNA sample. Editing levels were determined by visual inspection of the sequencing traces. Overlapping peaks formed by A and G traces indicate A-to-I editing sites. The ratio of G peak height to A peak height approximates the percent of the RNA population edited at a particular site.

Results

Cell fractionation by differential centrifugation successfully separates cellular DNA and protein. To determine the subcellular distribution and editing status of subcellular populations of *C. elegans* ADAR substrates, we separated cellular compartments of embryo cells. Because cell fractionation methods have not been previously used to partition cellular RNA content in *C. elegans*, we adapted a protocol

that had been successfully applied to confirm the subcellular localization of proteins (9).
This method, described in detail in the Experimental Procedures section, utilizes gentle
dounce homogenization that leaves nuclei intact. After removal of unlysed embryos and
cells, differential centrifugation is applied to separate the nuclear compartment, the
soluble cytoplasmic S100 fraction, and the P100 fraction, which contains membranes,
organelles and polysomes.

We verified the success of the cell fractionation by examining subcellular DNA
and protein content. DNA was extracted from the nuclear, S100 and P100 fractions and
analyzed by semiquantitative PCR for the genomic actin 1 (*act-1*) locus (Figure 3.1A).
As expected, we found that isolated nuclei contained the majority of genomic DNA. No
act-1 DNA was amplified from S100 samples, and only a small amount was detected in
P100, presumably resulting from nuclear fragments that co-purified with this fraction.

Western blots confirmed the expected subcellular distribution of proteins (Figure
3.1B). Hsp90, a cytosolic heat shock protein encoded by the *daf-21* gene (11), co-
fractioned almost exclusively with the S100 fraction. In contrast, histone H3 was
predominantly nuclear.

Because western analyses indicated successful partitioning of subcellular protein
populations, we used the isolated cell fractions in enzyme activity assays to determine the
subcellular localization of *C. elegans* ADARs and Dicer. A-to-I editing by ADARs
predominantly occurs in the nucleus in *Xenopus*, *Drosophila* and mammalian cells.
While it is assumed that this is also the case for *C. elegans*, no experimental evidence
exists in support of this claim. In an ADAR activity assay, we monitored editing of *in
vitro* transcribed ^{32}P-labeled ~800bp CAT duplex RNA in wildtype embryo lysate, nuclei

and S100 (Figure 3.1C). Our results confirm that ADAR activity primarily localizes to the nuclear compartment. However, a low level of A-to-I editing was also present in the soluble cytoplasmic fraction, and it remains unclear at this point whether this is due to imperfect separation of cellular compartments during the fractionation or whether the observed activity reflects an actual presence of ADARs in the cytosol.

Consistent with the finding that human Dicer predominantly localizes to the cytoplasm (12), we found that cleavage of CAT duplex RNA into siRNAs was enriched in S100 over lysate, and virtually nondetectable in the nuclear fraction (Figure 3.1D). However, human Dicer colocalizes with the endoplasmic reticulum (12). The P100 fraction, containing most membrane compartments and organelles, was not analyzed for Dicer activity and may contain additional Dicer activity.

mRNAs of housekeeping genes exhibit the expected subcellular distribution. We prepared total RNA from whole cell lysate, nuclei, S100 and P100 and examined the absolute distribution of RNA species using mass determination by spectrophotometry, northern analyses and quantitative RT-PCR (see Experimental Procedures for details on quantification). The majority of total RNA, on average ~60%, co-fractionated with the P100 compartment (Total RNA, Figure 3.2A). Because ribosomal RNAs (rRNAs) constitute the predominant RNA species in the cell, and most rRNAs are expected to be found in the polysomal fraction co-purifying with the P100, this distribution of total RNA was anticipated. After size selection for total RNA >200nt by column chromatography, the P100 contribution to cellular RNA content was increased to ~70% (Total RNA >200nt, Figure 3.2A). Concurrently, the S100 RNA content was reduced from an average of 23% before size selection to 14%, indicating that the S100 is enriched in small

Figure 3.1. **Distribution of genomic DNA and proteins in N2 embryo cell fractions.**
(A) Semiquantitative PCR of genomic *act-1* in 1/10,000 of genomic DNA samples
extracted from lysate and each cell fraction (L, lysate; N, nuclear fraction; S, S100; P,
P100). Aliquots of the PCR were removed after 15, 20, 25 and 30 cycles and analyzed by
agarose gel electrophoresis. The highest concentration of genomic DNA was detected in
the nuclear fraction. No DNA was present in the S100 fraction, while the P100 contained
some genomic material. (B) Western blot of the cytosolic protein Hsp90 (90kDa, white
arrowhead) and histone H3 (15.3kDa, black arrowhead) in equivalent volumes of lysate
and cell fractions. Hsp90 fractionates with the S100; histone H3 is predominantly
nuclear. The two lanes for the nuclear fraction represent the nuclear pellets obtained
from the sequential centrifugation steps at 750g (see Experimental Procedures). Low-
molecular-weight bands observed in the P100 likely represent non-specific detection of
small basic ribosomal proteins or, in case of the histone H3 blot, H3 sedimented with
nuclear debris. (C) *In vitro* editing of CAT duplex RNA in lysate and cell fractions. The
reaction products were treated with nuclease P1, and the resulting 5' nucleotide
monophosphates (pI, pA) were analyzed by thin-layer chromatography. The highest
ADAR activity is measured in the nuclear fraction. Whole cell lysate of *adr-1;adr-2*
double mutant embryos had no detectable ADAR activity. (D) Dicer cleavage of CAT
duplex RNA in lysate and nuclear and S100 fraction (15µg total protein each). The
reaction products were analyzed by PAGE. Dicer activity, evidenced by the appearance
of siRNAs (grey arrowhead), is largely confined to the S100 compartment. Labeled 25bp
DNA ladder (Invitrogen) was used as size standards.

A

B

C

D
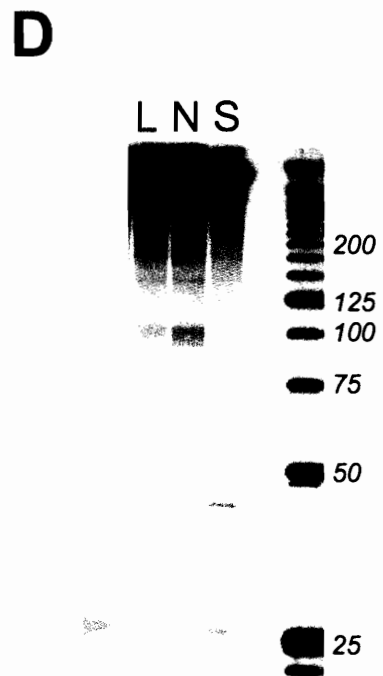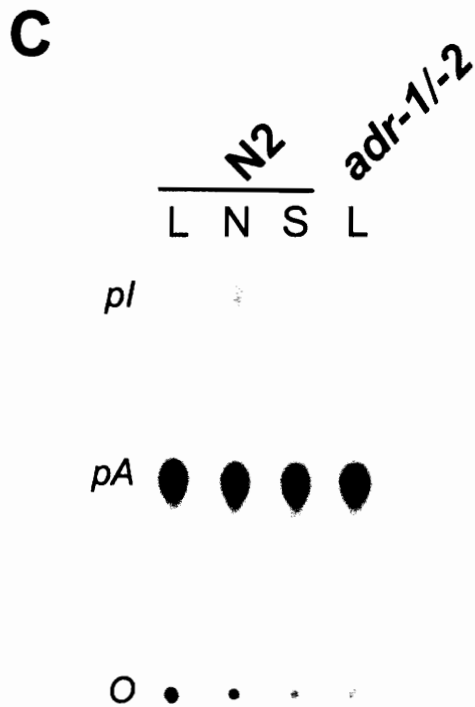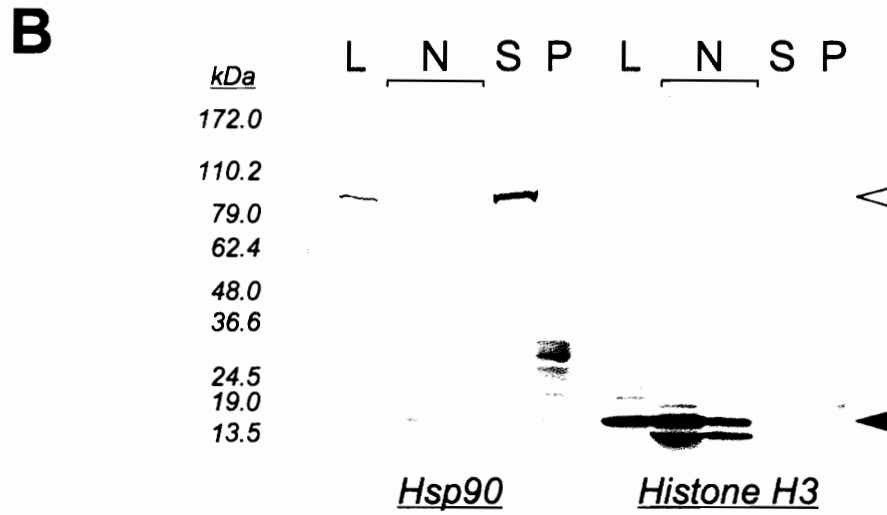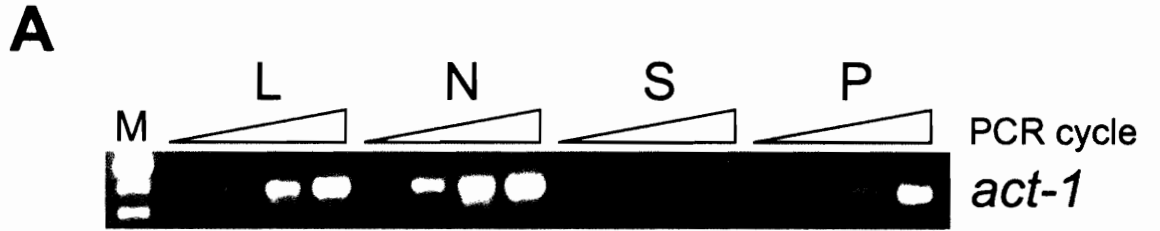
RNA species. This enrichment could reflect the true localization of small RNAs, such as tRNAs, to the soluble cytoplasm. However, because the S100 fraction contains all of the homogenization buffer used in the cell lysis and fractionation, any RNA species that diffused out of the nucleus or other organelles would be found in the S100. This may, indeed, be the case for small RNAs, as is discussed later in this chapter.

When polyadenylated RNA (poly(A)+ RNA) was prepared from cell fractions, ~70% of absolute cellular content co-purified with the P100 compartment and 10% was found in S100 (poly(A)+, Figure 3.2A). 20% of cellular poly(A)+ co-fractioned with the nuclear compartment. This number agrees with published data, indicating that 20% of cellular poly(A)+ is found in the nuclei of HeLa cells (13).

Using northern analyses and quantitative RT-PCR we examined the mRNA distribution of several housekeeping genes. Northern analyses of *act-1* mRNA, encoding actin in pharyngeal and body wall muscle (14), showed no enrichment in nuclear or S100 RNA samples (*act-1*, Figure 3.2B), and quantification of absolute cellular distribution revealed a pattern similar to what was observed for the collective poly(A)+ population, with ~70% of *act-1* localizing to the P100. This distribution is indicative of active translation of *act-1* on polyribosomes, and comparable results were obtained when we quantified the mRNA distribution of *gpd-3* and *uaf-1*, coding for a glyceraldehyde 3-phosphate dehydrogenase and the splicing factor U2AF^{65}, respectively (15,16).

The hyperedited mRNA of ADAR substrate 36A shows subcellular distribution similar to housekeeping genes. We used northern analyses and quantitative RT-PCR to determine the subcellular distribution of the C35E7.6 (36A) mRNA, which contains a ~1700nt 3'UTR that is extensively edited by ADARs (see Chapter 1). Using a probe

Figure 3.2. **Distribution of ADAR substrates and control RNAs in N2 embryo cell fractions.** (A) Absolute distribution of RNAs in cell fractions (grey, nuclei; white, S100; black, P100). The distribution was calculated from mass (total RNA, total RNA >200, poly(A)+), from northern blot (*rncs-1*) or from real-time RT-PCR (all other RNAs) by methods described in the Experimental Procedures. The average distribution determined in at least two independent fractionations is plotted. Numbers indicate values (in %) of absolute cellular RNA found in each fraction. (B) Northern blot of *act-1*, 36A and *rncs-1* in 5μg of poly(A)+ RNA isolated from lysate and cell fractions (L, lysate; N, nuclear; S, S100; P, P100). The blot is representative of two or more northern experiments from independent fractionation experiments. The northern analyses shown were obtained by hybridizations of the same blot.

A
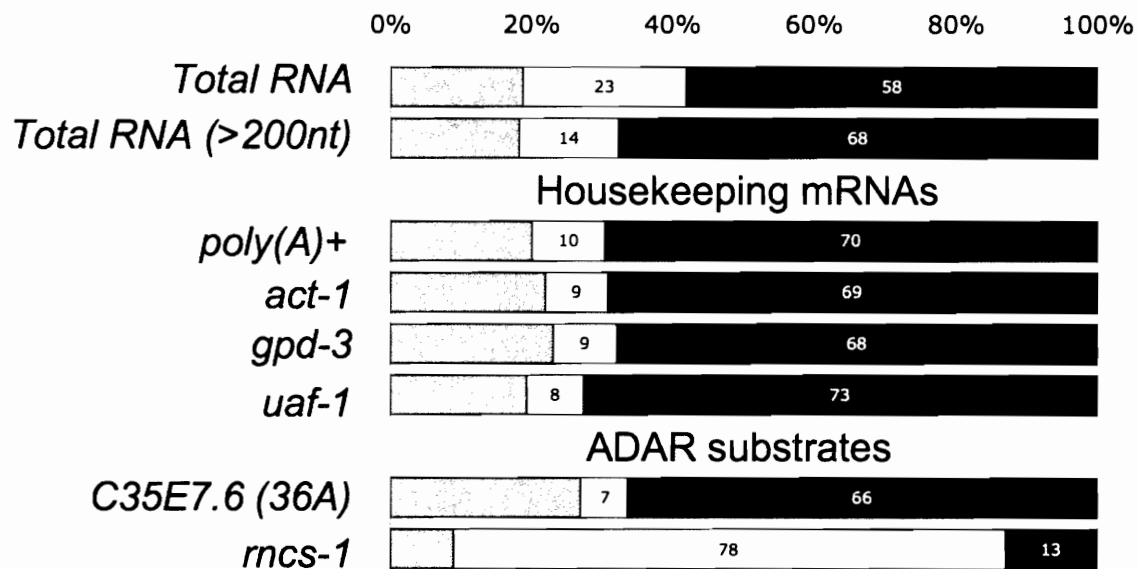
B
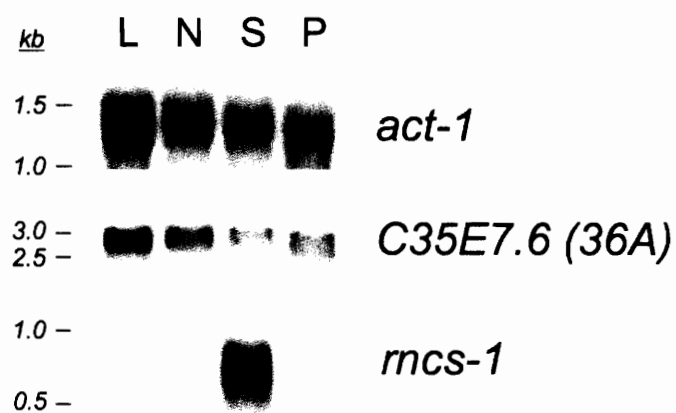
complementary to the unedited coding region of 36A, northern analyses of poly(A)+ RNA isolated from cell fractions detected a single species of ~2.9kb in all samples (36A, Figure 3.2B). This species likely corresponds to the full-length transcript, containing the coding region as well as the 3'UTR. Similar to the relative mRNA levels detected in a northern for *act-1*, no enrichment for 36A in nuclear or S100 fractions was observed. When the absolute cellular distribution of 36A was quantified, the majority of mRNA (66%) localized to the P100 compartment (36A, Figure 3.2A).

Because it was a possibility that the cytoplasmic 36A population consisted of largely unedited species, while heavily edited 36A transcripts were retained in the nucleus, as would be predicted by the Carmichael model of nuclear retention, we sequenced the 36A 3'UTR in cDNA samples isolated from whole cell lysate and cell fractions. We found that the cytoplasmic 36A populations, co-purifying with the S100 and P100 fractions, reproducibly showed editing levels that were increased in comparison to nuclear 36A (Figure 3.3A). For most editing sites analyzed, the editing level in lysate was most similar to that in the P100 fraction, consistent with the finding that the P100 contains the majority of cellular 36A.

Hyperedited rncs-1 is enriched in the S100. Using northern analyses we also examined the subcellular distribution of *rncs-1* RNA, a second ADAR substrate. In contrast to 36A transcript or housekeeping mRNAs, *rncs-1* was greatly enriched in RNA samples isolated from S100 (*rncs-1*, Figure 3.2B), with an absolute amount of >70% of cellular *rncs-1* co-fractioning with this compartment (Figure 3.2A). Less than <15% of *rncs-1* co-purified with the P100 fraction, consistent with earlier findings that *rncs-1* RNA is not associated with polysomes (7). Sequencing of *rncs-1* cDNA samples

revealed that S100 population is greatly enriched for hyperedited species, while the editing levels in nuclear RNA were significantly lower (Figure 3.3B).
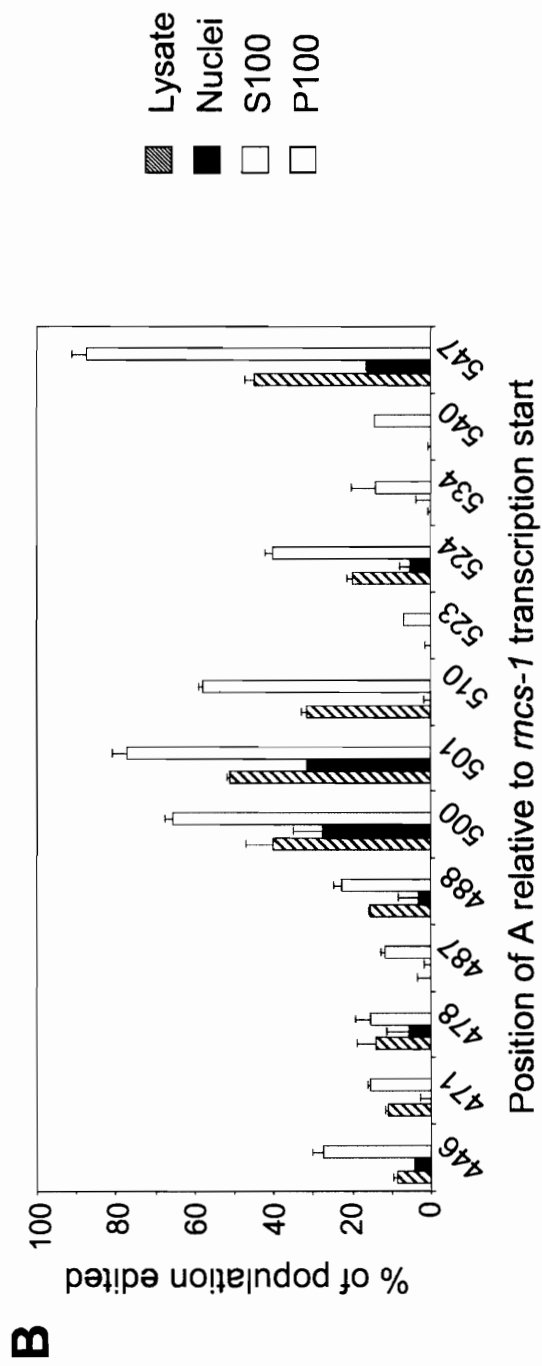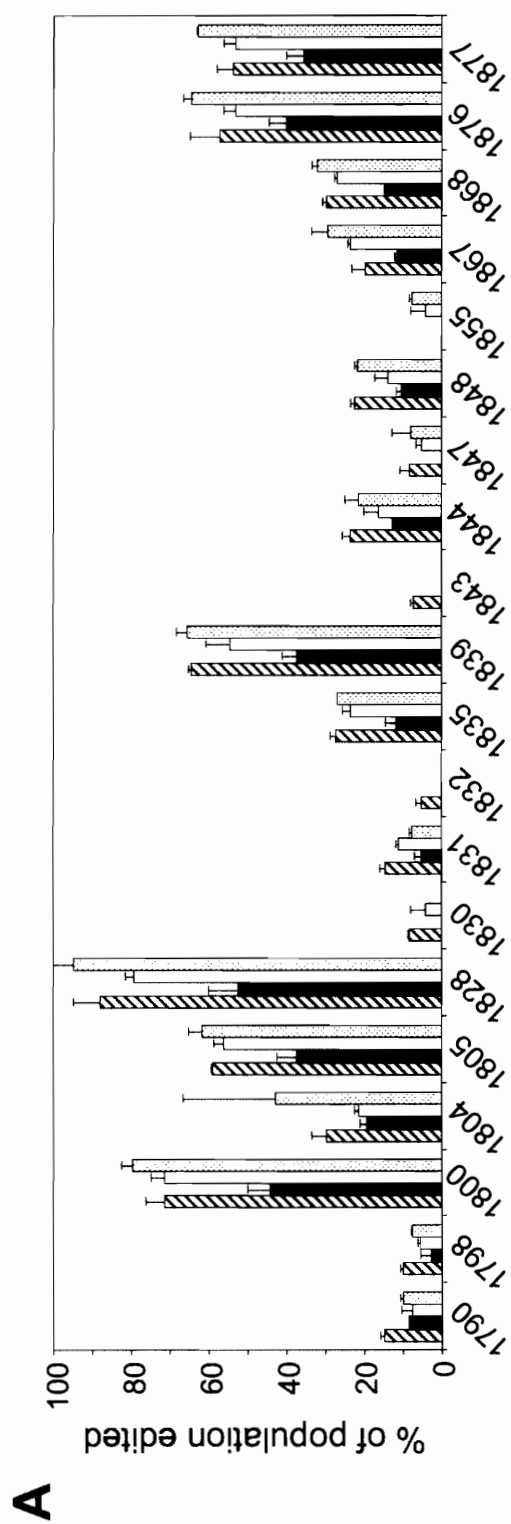
Cell fractionation is a reliable method for subcellular partitioning of large RNA species but allows for diffusion of small RNAs. While our results for the subcellular distribution of hyperedited 36A and *rncs-1* were intriguing, we needed to establish that the cell fractionation was a reliable method for determining the subcellular localization of RNAs. Therefore, we determined the distribution of 3 *C. elegans* RNAs, ranging in size from ~7kb to ~100nt, known to exclusively localize to the nucleus.

The largest nuclear RNA examined was the 7kb ribosomal precursor RNA (pre-rRNA). We used northern analyses and RT-PCR to detect the internal transcribed sequence 1 (ITS1), which is excised from the pre-rRNA transcript prior to nuclear export of the mature ribsomal RNAs (17). We determined that pre-rRNA was greatly enriched in RNA samples from nuclei, with >90% of all cellular pre-rRNA co-purifying with this compartment (pre-rRNA, Figure 3.4).

The *uaf-1* locus, encoding the 1.7kb *uaf-1* mRNA assayed above, also gives rise to a 1.9kb transcript, herein referred to as *uaf-1(ex5)*. This transcript includes an alternatively spliced exon, which contains a premature stop codon and 10 repeats of the U_4CAG/R splice site consensus sequence. This exon causes nuclear retention of the *uaf-1(ex5)* splice variant, presumably through association with splicing factors (18). By northern and RT-PCR analysis we verified that *uaf-1(ex5)* was enriched in the nuclear compartment (*uaf-1(ex5)*, Figure 3.4). However, quantification of the absolute distribution of this transcript revealed that only ~50% of cellular content co-purified with the nuclei.

Figure 3.3. **Editing of ADAR substrates in cell fractions.** RNA in N2 embryo lysate and subpopulations of RNA that cofractioned with nuclei, S100 and P100 were analyzed for A-to-I editing levels in ADAR substrates 36A (A) and *rncs-1* (B). Edited regions were amplified by RT-PCR and sequenced using a population sequencing approach (see Experimental Procedures). Percent editing at any given site represents the percent of the population with an A-to-I change at that site. The mean percentage +/- standard error of the mean determined in two independent cell fractionation experiments is plotted. Editing levels of *rncs-1* in P100 was not analyzed.

A — Position of A relative to 36A transcription start (% of population edited)

B — Position of A relative to *mcs-1* transcription start (% of population edited)

Legend: Lysate, Nuclei, S100, P100

Figure 3.4. **Size dependence of distribution of control nuclear RNAs in N2 embryo cell fractions.** The distribution of RNA species of different sizes with known localization in the nucleus was analyzed. Approximate sizes of the RNA species are indicated on the scale on the right hand scale. Northern blots of 20μg total RNA (pre-rRNA, U6) or 5μg of poly(A)+ RNA (*uaf-1(ex5)*) were probed for pre-rRNA, *uaf-1(ex5)* mRNA and U6 snRNA (L, lysate; N, nuclear; S, S100; P, P100). The absolute cellular distribution of each RNA was calculated from the relative abundance quantified in the northern analyses (see Experimental procedures). The average absolute distribution determined in two cell fractionation experiments is plotted in a bar graph diagram below each northern blot. The absolute percentage of each RNA found in nuclear (light grey), S100 (white) and P100 (dark grey) is indicated by the value number. Black arrowheads indicate approximate sizes of the ADAR substrates 36A and *rncs-1* on the size scale.
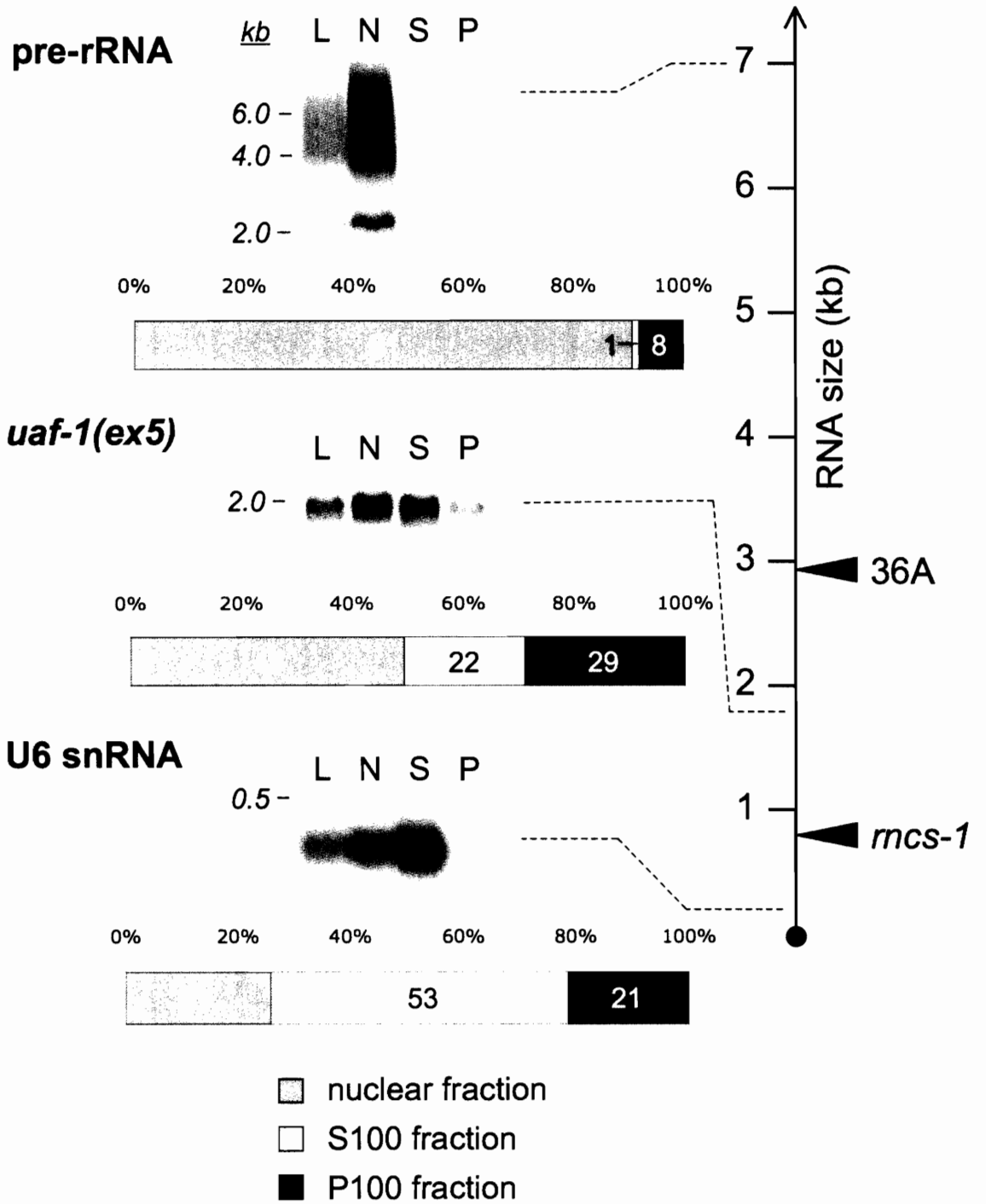
The U6 snRNA is a 102nt spliceosomal RNA. In contrast to other snRNA, biogenesis and function of U6 is entirely confined to the nucleus (19). While we observed some enrichment of U6 concentration in nuclear RNA over lysate RNA by northern blot (compare L and N lanes, U6 snRNA, Figure 3.4), U6 concentration in S100 was found to be the highest. Quantification of absolute distribution showed <30%, instead of the expected ~100%, of U6 remained in the nuclear compartment.

The distribution of the nuclear control RNAs—all of which should be exclusively found in nuclei in an ideal cell fractionation—suggests that the cell fractionation method may only be reliable in ascertaining nuclear retention of RNAs of sufficient size. While large nuclear RNAs, like pre-rRNA, remain associated with the nuclear compartment, very small RNA species, like U6, appear diffusible, as evidenced by the loss of U6 into the S100. Consistent with this idea is the observed enrichment of RNAs <200nt in the S100 fraction noted earlier in this chapter.

While great care was taken to ensure gentle homogenization of embryos, agitation and compression of nuclei during the lysis process could have led to expulsion of small RNAs from the nucleus. In addition, size may not be the only determinant in allowing or preventing loss of nuclear RNA species into the S100 during cell fractionation. RNAs, like U6, are often incorporated into ribonucleoprotein complexes (RNPs). Size of the RNP as well as affinity of an RNA for its associated proteins may influence retention of the RNA in the nuclear compartment during the fractionation.

Given these findings, we are cautious in interpreting the results obtained for the subcellular localization of the *C. elegans* ADAR substrates. Possible conclusions and reservations are discussed below.

Discussion

In the experiments described above, we used a cell fractionation approach to separate subcellular RNA populations of *C. elegans* embryos. Consistent with constitutive translation, the mRNAs of several housekeeping genes largely co-purified with the P100 compartment, presumably due to association of these mRNAs with polyribosomes. Interestingly, and in contradiction to the existing model of nuclear retention of inosine-containing RNAs, the hyperedited ADAR substrates 36A and *rncs-1* were not enriched in the nuclear compartment. Instead, the majority of 36A and *rncs-1* transcripts co-purified with the P100 and S100 fraction, respectively.

Is hyperedited 36A mRNA exported from the nucleus and translated? The relative distribution of 36A in cell fractions assayed by northern and RT-PCR, and the absolute subcellular distribution calculated from these analyses, indicate that the 36A mRNA distribution is comparable to that of housekeeping messages (Figure 3.2). This suggests that the 3'UTR, in spite of hyperediting by ADARs, does not prevent nuclear export of 36A. On the contrary, editing levels were higher in the P100 population than in the nuclear population for all editing sites analyzed (Figure 3.3A).

The study of CTN-RNA by Prasanth et al. suggested that RNAs retained in the nucleus could be released and exported for translation after removal of the 3'UTR (4). The data presented here show no indication that this is the case for 36A. The 36A-specific northern probe was designed to detect the coding region of the message, and would have detected a smaller species resulting from removal of all or part of the 1.7kb edited 3'UTR. However, northern analyses from as much as 5μg of poly(A)+ RNA consistently showed a single species of ~2.9kb, the expected size of the full-length

transcript that includes the 3'UTR. In addition, the RT-PCR primers used to amplify a region of 36A for sequencing span the exon-3'UTR junction. The hyperedited population sequenced in the P100, therefore, clearly corresponds to messages that contain coding sequences and 3'UTR.

While we show that the cellular fractionation approach may not be a suitable method to verify the nuclear retention of all RNAs, we believe our data for 36A is reliable for two reasons. First, according to our findings for nuclear control RNAs, the size of 36A mRNA is sufficient to prevent diffusion during the fractionation (size scale, Figure 3.4). The ~1kb smaller *uaf-1(ex5)* transcript, while also subject to diffusion out of the nucleus, showed significant (~50%) co-purification with the nuclear compartment. Second, erroneous diffusion out of the nucleus during lysis is expected to result in enrichment of RNA in the S100 fraction, as is the case for U6 snRNA and for total RNA species <200nt. However, the majority of 36A was found in the P100 fraction.

Our data indicate that hyperediting of the 36A 3'UTR does not prevent nuclear export of the 36A mRNA. While we observed co-purification with the P100 fraction, we cannot infer association with polyribosomes or translation of 36A. However, experiments using fusion constructs of reporter genes and the 36A 3'UTR as well as polyribosome profiles indicate that hyperediting of the 36A 3'UTR does not prevent translation (Heather Hundley, unpublished data).

Carmichael's studies suggest that the binding of nuclear vigilin to inosine-containing RNA mediates a function of edited dsRNA in heterochromatin silencing (3). Interestingly, he and other groups reported that the majority of cellular vigilin is cytoplasmic and associated with polyribosomes (3,20,21). In *Xenopus*, binding of vigilin

to the 3'UTR of estrogen-inducible vitellogenin mRNA stabilizes the message and prevents cleavage by a polysomal ribonuclease (22,23). It is intriguing to think that cytoplasmic vigilin could, in fact, be facilitating the association of hyperedited mRNAs, such as 36A, with polysomes. *In vitro* selection experiments and mutational analysis of the vigilin binding site in the vitellogenin 3'UTR indicated that RNA secondary structure influences binding of vigilin, with tighter binding observed for flexible single-stranded substrates (24). It is interesting to note that ADAR editing can significantly influence the double-stranded properties of RNA by changing AU basepairs to IU mismatches. Therefore, while hyperediting may not prevent nuclear export of hyperedited 36A message in *C. elegans*, it may impact how 36A associates with cytoplasmic vigilin. This, in turn, could influence mRNA stability, analogous to *Xenopus* vitellogenin, or alter association with polyribosomes and translation efficiency.

Is hyperedited rncs-1 exported to the soluble component of the cytoplasm? Our data suggest that *rncs-1* is found in the soluble cytoplasm. However, because the reliability of the cell fractionation method was size-dependent, and the 800nt *rncs-1* RNA is located at the lower end of the size spectrum (Figure 3.4), the data need to be evaluated cautiously.

Two observations indicate that the enrichment in the S100 reflects the true subcellular localization of *rncs-1*. First, the absolute fraction of *rncs-1* co-purifying with the S100 (78%) was greatest among all RNAs examined. While a large fraction of U6 snRNA was subject to diffusion during the fractionation, the relative comparison of U6 concentration by northern analysis showed an enrichment of U6 in nuclear RNA compared to whole cell lysate (lane N > lane L, U6 snRNA, Figure 3.4). Such

enrichment was never observed for *rncs-1* RNA (lane N < lane L, *rncs-1*, Figure 3.2).

Second, the comparison of editing levels in nuclear and S100 *rncs-1* populations show

clear differences between the compartments, with S100 populations exhibiting an

enrichment in highly edited species (Figure 3.3B). If localization of *rncs-1* was solely

due to random diffusion out of the nucleus during the fractionation, editing levels of

rncs-1 should be roughly equal in both compartments. It is possible that highly edited

rncs-1 species are more prone to diffusion compared to transcripts with fewer edited

sites, however, this would further contradict the model that hyperedited RNAs are

retained in the nucleus by cooperative association with a retention complex. In addition,

we cannot completely exclude the possibility that editing levels of nuclear samples were

artifactually lowered by contamination with genomic DNA. The primers used for RT-

PCR of *rncs-1*—unlike those designed for PCR and sequencing of 36A cDNA—did not

discriminate against unspliced product. While –RT controls were included in all RT-PCR

reaction and showed no product by gel-electrophoresis, small amounts of product may

have been amplified from DNA, mimicking the unedited state of cDNA.

To corroborate the cytoplasmic localization of *rncs-1* suggested by the cell

fractionation method, efforts in our lab and in collaboration with other groups have been

directed at localizing *rncs-1* RNA by in situ hybridization in embryos (Györgyi

Csankovszki & Barbara Meyer, Michael Jäntsch, personal communication). In situ

hybridization, technically difficult in nematodes, is further complicated by the secondary

structure and editing of *rncs-1*. In northern analyses, probes fully complementary to an

unedited copy of *rncs-1* were equally efficient in detecting *rncs-1* in samples from

wildtype and ADAR deficient worms (data not shown). It is unclear, however, whether

this would be true for in situ hybridization. In our in situ hybridization attempts, we used probes complementary to the unedited 5' strand of *rncs-1*, probes complementary to the unedited terminal branched structures or degenerate probes to the 5' strand synthesized from wildtype whole embryo cDNA. The latter should include probes perfectly complementary to any *rncs-1* molecule in the cell, regardless of editing status. As a negative control the probes were also hybridized to *rncs-1* mutant embryos, which show no *rncs-1* transcript by northern analysis.

The majority of in *rncs-1* situ hybridization experiments performed by our lab and in collaborating labs showed cytoplasmic localization of the *rncs-1* probes (data not shown). However, background in *rncs-1(tm1632)* control animals was high, and often indistinguishable from the signal in wildtype worms. In one experiment using degenerate probes, we detected a nuclear signal in several embryos (data not shown), but this was also seen in some *rncs-1* mutant embryos. In summary, in situ hybridization attempts so far have yielded unsatisfactory results. It is possible, that in situ detection of *rncs-1* in *rncs-1* overexpressing lines (see Chapter 2) would prove more successful.

While our current lines of evidence point towards cytoplasmic localization of *rncs-1* RNA, our reservations in regards to the reliability of the methods used to determine subcellular localization, prevent us from making a statement of absolute certainty. In our findings described in Chapter 2, we suggest a molecular function for *rncs-1* as an inhibitor of Dicer activity. Further analysis of the functional and possibly physical interaction of Dicer and *rncs-1* may provide further clues to the subcellular localization of *rncs-1* RNA.

Why are hyperedited RNAs not retained in C. elegans nuclei? The evidence presented in this chapter indicates, at least in the case of 36A, that hyperedited *C.elegans* ADAR substrates are exported from the nucleus. This is in clear contradiction to the current model of nuclear retention of hyperedited dsRNA, based on observations for hyperedited viral and synthetic dsRNAs. What are the possible reasons for these conflicting findings?

First, as discussed earlier in this chapter, editing levels in the cellular ADAR substrates are lower than those in synthetic and viral dsRNAs. On average only ~20-25% of adenosines in the double-stranded regions of 36A and *rncs-1* are changed to inosine. This editing level may be insufficient for nuclear retention. While perfectly base-paired dsRNA is deaminated at 50-60% of its adenosines, internal loops and bulges within ADAR substrates lead to more selective editing of adenosines (25).

Second, it is possible that mRNA processing, such as splicing, capping and polyadenylation, can overcome an inosine-mediated nuclear retention mechanism. While data show that splicing of an adenovirus-derived construct cannot prevent nuclear retention (2), microinjection of such a construct does not truly reflect the complex sequence of processing and packaging steps that nascent cellular transcripts undergo (26).

Third, it is not clear that the p54[nrb]/PSF/matrin 3 protein complex implicated in nuclear retention of hyperedited RNA is conserved in *C. elegans*. The *C. elegans* F25B5.7 gene encodes a putative homolog of human p54[nrb] and PSF, but binding of inosine-containing RNAs or association with a possible homolog of matrin 3 remain to be shown. Thus, nuclear retention of hyperedited RNAs may be a mechanism limited to higher eukaryotes.

References

1.   Kumar, M. and Carmichael, G.G. (1997) Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci U S A*, **94**, 3542-3547.

2.   Zhang, Z. and Carmichael, G.G. (2001) The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell*, **106**, 465-475.

3.   Wang, Q., Zhang, Z., Blackwell, K. and Carmichael, G.G. (2005) Vigilins bind to promiscuously A-to-I-edited RNAs and are involved in the formation of heterochromatin. *Curr Biol*, **15**, 384-391.

4.   Prasanth, K.V., Prasanth, S.G., Xuan, Z., Hearn, S., Freier, S.M., Bennett, C.F., Zhang, M.Q. and Spector, D.L. (2005) Regulating gene expression through RNA nuclear retention. *Cell*, **123**, 249-263.

5.   Bass, B.L., Hellwig, S. and Hundley, H.A. (2005) A nuclear RNA is cut out for translation. *Cell*, **123**, 181-183.

6.   Morse, D.P., Aruscavage, P.J. and Bass, B.L. (2002) RNA hairpins in noncoding regions of human brain and Caenorhabditis elegans mRNA are edited by adenosine deaminases that act on RNA. *Proc Natl Acad Sci U S A*, **99**, 7906-7911.

7.   Morse, D.P. and Bass, B.L. (1999) Long RNA hairpins that contain inosine are present in Caenorhabditis elegans poly(A)+ RNA. *Proc Natl Acad Sci U S A*, **96**, 6048-6053.

8.   Bass, B.L. (2002) RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem*, **71**, 817-846.

9.   Chen, F., Hersh, B.M., Conradt, B., Zhou, Z., Riemer, D., Gruenbaum, Y. and Horvitz, H.R. (2000) Translocation of C. elegans CED-4 to nuclear membranes during programmed cell death. *Science*, **287**, 1485-1489.

10.   Macbeth, M.R., Lingam, A.T. and Bass, B.L. (2004) Evidence for auto-inhibition by the N terminus of hADAR2 and activation by dsRNA binding. *RNA*, **10**, 1563-1571.

11.   Birnby, D.A., Link, E.M., Vowels, J.J., Tian, H., Colacurcio, P.L. and Thomas, J.H. (2000) A transmembrane guanylyl cyclase (DAF-11) and Hsp90 (DAF-21) regulate a common set of chemosensory behaviors in caenorhabditis elegans. *Genetics*, **155**, 85-104.

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

13.   Herman, R.C., Williams, J.G. and Penman, S. (1976) Message and non-message sequences adjacent to poly(A) in steady state heterogeneous nuclear RNA of HeLa cells. *Cell*, **7**, 429-437.

14.   Krause, M., Wild, M., Rosenzweig, B. and Hirsh, D. (1989) Wild-type and mutant actin genes in Caenorhabditis elegans. *J Mol Biol*, **208**, 381-392.

15.   Huang, X.Y., Barrios, L.A., Vonkhorporn, P., Honda, S., Albertson, D.G. and Hecht, R.M. (1989) Genomic organization of the glyceraldehyde-3-phosphate dehydrogenase gene family of Caenorhabditis elegans. *J Mol Biol*, **206**, 411-424.

16.   Zorio, D.A., Lea, K. and Blumenthal, T. (1997) Cloning of Caenorhabditis U2AF65: an alternatively spliced RNA containing a novel exon. *Mol Cell Biol*, **17**, 946-953.

17.   Saijou, E., Fujiwara, T., Suzaki, T., Inoue, K. and Sakamoto, H. (2004) RBD-1, a nucleolar RNA-binding protein, is essential for Caenorhabditis elegans early development through 18S ribosomal RNA processing. *Nucleic Acids Res*, **32**, 1028-1036.

18.   MacMorris, M.A., Zorio, D.A. and Blumenthal, T. (1999) An exon that prevents transport of a mature mRNA. *Proc Natl Acad Sci U S A*, **96**, 3813-3818.

19.   Kiss, T. (2004) Biogenesis of small nuclear RNPs. *J Cell Sci*, **117**, 5949-5951.

20.   Klinger, M.H. and Kruse, C. (1996) Immunocytochemical localization of vigilin, a tRNA-binding protein, after cell fractionation and within the exocrine pancreatic cell of the rat. *Ann Anat*, **178**, 331-335.

21.   Lang, B.D. and Fridovich-Keil, J.L. (2000) Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. *Nucleic Acids Res*, **28**, 1576-1584.

22.   Cunningham, K.S., Dodson, R.E., Nagel, M.A., Shapiro, D.J. and Schoenberg, D.R. (2000) Vigilin binding selectively inhibits cleavage of the vitellogenin mRNA 3'-untranslated region by the mRNA endonuclease polysomal ribonuclease 1. *Proc Natl Acad Sci U S A*, **97**, 12498-12502.

23.   Dodson, R.E. and Shapiro, D.J. (1997) Vigilin, a ubiquitous protein with 14 K homology domains, is the estrogen-inducible vitellogenin mRNA 3'-untranslated region-binding protein. *J Biol Chem*, **272**, 12249-12252.

24. Kanamori, H., Dodson, R.E. and Shapiro, D.J. (1998) In vitro genetic analysis of the RNA binding site of vigilin, a multi-KH-domain protein. *Mol Cell Biol*, **18**, 3991-4003.

25. Lehmann, K.A. and Bass, B.L. (1999) The importance of internal loops within RNA substrates of ADAR1. *J Mol Biol*, **291**, 1-13.

26. Aguilera, A. (2005) Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Curr Opin Cell Biol*, **17**, 242-250.

CHAPTER 4

DAUER FORMATION IN *C. ELEGANS* ADAR MUTANTS

Introduction

Mutations in ADAR genes have been isolated in various organisms. These

mutant animals are invaluable tools in studying ADAR biology, and their analysis has

furthered our understanding of the physiological consequences of dysfunctional A-to-I

editing.

In lower organisms ADARs are nonessential. However, viable animals with

ADAR deficiencies display behavioral and neurological defects. *Drosophila* ADAR1

null mutants, while grossly wildtype in morphology and longevity, exhibit

neurobehavioral phenotypes that include locomotion and body posture defects,

excessive grooming habits, defective male mating behavior and a compromised

response to hypoxia (1,2). These defects are aggravated by age, and progressive brain

degeneration is evident in older ADAR mutant flies. Consistent with the findings in

Drosophila, homozygous ADAR deficient *C. elegans* are morphologically normal but

display sensory defects when subjected to chemotaxis assays (3).

In contrast to the dispensable role of A to I editing in flies and nematodes,

ADARs are essential in mammals. Mice homozygous for an ADAR2 deletion become

prone to seizures and die shortly after birth (4), and deficiency of mammalian ADAR1

causes embryonic lethality due to hematopoietic defects and apoptotic disintegration of

the liver and other tissues (5-7). Interestingly, the lethality of the mouse ADAR2

deletion can be reverted by ectopic expression of an AMPA receptor GluR-B mRNA

containing a single A to G change to mimic the edited state of the Q/R site (4). This

raises the possibility that ADAR function evolved to be essential by correcting

otherwise deleterious G to A mutations in a small subset of genes critical for

development.

In addition to the essential function of ADARs in development, mammalian

ADARs have retained the nonessential function of fine-tuning cellular pathways in the

nervous system that is observed in invertebrates. This is evident in a number of human

disorders associated with dysfunctional A to I editing. Misregulated editing of the

AMPA GluR2 mRNA is associated with pathophysiological aspects of amyotrophic

lateral sclerosis, epilepsy, Alzheimer's disease, Huntington's disease, schizophrenia

and forebrain ischemia (8-12). Altered editing of a second ADAR substrate in human

brain, the serotonin receptor subtype 2C, has been implicated in schizophrenia (13) and

may modify the response to pharmacotherapy in patients treated for major depressive

disorders (14-16).

Previous work in our laboratory described the chemosensory defect of ADAR

mutant *C. elegans* (3). Tonkin et al. observed that worms with altered or absent ADAR

activity showed reduced tracking towards volatile chemical attractants. This defect

varied in severity with the type of chemical and its concentration. Both *adr-1* and

adr-2 genes contribute to the ability to detect attractants. However, *adr-2* single

mutants have a more severe chemotaxis defect compared to *adr-1* single mutants,

Several reports indicate that this is the case for the pathways of dauer formation and

consistent with the fact that A to I editing is completely eliminated in an *adr-2* deficient background, while it is only altered in *adr-1* mutants.

In *C. elegans* neurological pathways involved in various processes—such as thermotactic and chemotactic responses, locomotion, egg laying, dauer larva formation, osmotic avoidance and male mating behavior—intersect, and mutants displaying one behavioral defect are often compromised in additional neurological responses (17,18). chemotaxis. Vowels and Thomas showed that dauer formation defective worms carrying mutations in the *daf-11* or *daf-21* genes are abnormal for sensing water-soluble odorants (19), and a screen for mutants with defective amphid neuron morphology identified several genes that share altered chemotaxis and dauer formation phenotypes (20).

The *C. elegans* dauer larva is a facultative, specialized developmental stage, characterized by altered morphology, cessation of feeding, unusual movement and increased resistance to harsh environmental conditions (21). Dauer larva formation is initiated when *C. elegans* encounters adverse conditions during its early development. Two opposing environmental signals—food supply and population density—are integrated by signal transduction pathways (18). Ample food supply antagonizes dauer formation, and dauers are formed in uncrowded culture when food is gradually limited (21). Quorum sensing in *C. elegans* is mediated through dauer pheromone (daumone), a short-chain hydroxylated fatty acid compound (22,23). A high concentration of daumone, which is secreted throughout the worm's life cycle, is indicative of high population density and can prompt dauer formation even in the presence of abundant

food (22,24). In the laboratory, daumone can be partially purified from worm culture medium and is used to identify dauer-defective mutants (25).

Because *C. elegans* ADAR mutants have a chemosensory defect, we wondered whether they were also compromised in dauer formation, as is the case for several other chemotaxis-deficient strains. In the work described below, we exposed wildtype and ADAR-mutant worms to dauer pheromone and monitored efficacy of entry into dauer. While our initial results suggested a function for the *adr-1* gene in dauer formation, subsequent experiments using a different deletion allele and *adr-1* rescue strains could not corroborate these findings. We conclude that A-to-I editing in the *C. elegans* nervous system is not essential for dauer formation in response to daumone. By principle of exclusion, this may ultimately prove helpful in identifying the ADAR editing substrates important for chemotaxis.

Experimental procedures

Strains. Strains used in this study were N2, *BB2 adr-1(gv6)*, *BB3 adr-1(gv42)*, *BB4 adr-1(gv6);adr-2(gv42)* (3), *BB19 adr-1(tm668)*, *BB20 adr-2(ok735)*, *BB21 adr-1(tm668);adr-2(ok735)*, *BB43 adr-1(gv6) uuEx9 [adr-1(+), sur-5::GFP]*. The *adr-1(gv6)* and *adr-2(gv42)* deletion alleles originated in the lab of Mike Krause. The *adr-1(tm668)* deletion allele was received from Dr. Shohei Mitani at the National Bioresource Project for the Nematode, Japan; the *adr-2(ok735)* allele was isolated by Dr. Robert Barstead at the Oklahoma Medical Research foundation, member of the *C. elegans* Gene Knockout Consortium. The *adr-1(tm668)* and *adr-2(ok735)* alleles were backcrossed individually against N2 8 times to generate strains *BB19* and *BB20*, respectively. Crossing of *BB19* and *BB20* yielded the double-mutant strain BB21.

Construction of the *adr-1* rescue strain *BB43* is detailed below. Strains were maintained by standard culture techniques (26).

Construction of the adr-1 rescue strain. A ~6.8kb fragment was PCR amplified from wildtype genomic DNA using primers adr-1PCRRSQU (AACACGGGACTGGA CTCATGAGAAAATGTCT) and adr-1PCRRSQL (CTGTTCCAATTACAGTTGTCA AAATCCATCGTC). This DNA fragment contains ~1.2kb of upstream sequence, the entire *adr-1* gene and ~1.6kb of downstream sequence. The PCR product was microinjected into *adr-1(gv6)* mutant hermaphrodites at a concentration of 5ng/μL along with 5ng/μL of a *sur-5::GFP* marker plasmid and 90ng/μL non-specific DNA (1kb DNA ladder, Invitrogen) to create transgenic lines as described (27).

Partial purification of dauer pheromone from liquid culture medium. To crudely purify daumone a method described by Golden and Riddle was adapted (24). Liquid cultures of worms (2L or more of combined volume), grown for several days, were centrifuged at 1,700g to sediment worms and other particles. The supernatant was transferred to a large-volume glass beaker and reduced by boiling on a heated stir plate in the fume hood until all liquid had evaporated. Using a spatula, the dry residue was loosened from the walls and bottom of the beaker. Small pieces of residue were transferred directly to 50mL plastic conical tubes, while larger pieces were broken up with a mortar and pestle prior to transfer. Conical tubes were filled with solids to no more than half of the tube volume. Pheromone was extracted from residue by addition of 95% ethanol (EtOH) to a final volume of 40mL. Tubes were shaken or vortexed vigorously for 5-10 min before centrifugation at 3,000g. The EtOH extract was decanted and reserved. Extraction of the solids with EtOH was repeated for an

additional four to five times until extract was only slightly colored. The pooled EtOH extract was transferred in batches to a small glass beaker and reduced to solids by boiling on a heated stir plate in the fume hood. The solids were desolved in 10mL of distilled water. The dark brown pheromone solution was sterile filtered through a 0.2µm syringe-top filter and stored at -20°C. The activity of pheromone extract varied between preparations. Typically, between 10 and 40µL pheromone/mL agar were required to induce 50% dauer formation in N2 at 20°C.

Assay for induction of dauer larva formation on pheromone plates. Dauer assays were performed following published protocols (24,25). In brief, indicated amounts of crude pheromone extract were added to NGM agar, prepared without peptone and to a final agar concentration of 2.2%. Petri dishes (35 by 10mm), containing 2mL of agar, were used in the assay. All plates used in an experiment contained extract preparations of the same batch. Plates were seeded with 20µL aliquots of *E. coli* OP50, suspended to 6% (w/w) in M9 buffer supplemented with 50µg/mL streptomycin, and allowed to dry overnight. 8-10 gravid adults were placed on each assay plate and allowed to lay 80-120 eggs before being removed. The plates were incubated at the assay temperature for 3 days until nondauers had reached adulthood. Dauers and nondauer worms were counted using a stereomicroscope by removing individual worms by aspiration. The percentage of the population that had entered the dauer stage was calculated.

Northern and RT-PCR analyses of adr-1. RNA was prepared and northern and RT-PCR analyses were performed by techniques described in Chapter 2. To synthesize a northern probe specific for *adr-1*, a ~2.7kb *adr-1c* cDNA segment was PCR amplified using primers adr-15'LCU (TGGCACTGATCACACCAGCGATTC) and LAT068

(TTACGATACGCCTCAAAGATTTCAG). Random primed oligonucleotide probes, carrying internal ^{32}P-dCTP labels, were synthesized from this PCR product using the Strip-EZ® DNA system (Ambion). Primer pairs used in the quantitative RT-PCR analyses were adr-15'LCU (see above) and adr-15'LCL (CTGTTGTTGCTGCTGCTGCTGTTGTT) for amplification of the 5' fragment of the *adr-1* cDNA and adr-1LCU (CAGTCGATGATGCTGCAACTAAGAAGT) and adr-1LCL (CATTTGCACAGGCATACTCTCCACA) for amplification of the 3' fragment.

Results

Worms carrying the adr-1(gv6) allele, but not those carrying the adr-1(tm668) allele, are dauer formation defective. To test whether *C. elegans* defective in A-to-I editing have defects in dauer formation, we subjected worms with mutations in *adr-1*, *adr-2* or both genes to assays for dauer formation in response to dauer pheromone (24,25). In these assays, embryos are hatched and grown on agar plates containing abundant food and a varying amount of partially purified daumone. Wildtype larvae grown under these conditions will either enter into dauer or develop into adults in a pheromone-concentration-dependent manner, with a higher percentage of the population entering dauer at higher daumone concentrations (Figure 4.1A, N2).

In a first set of experiments, 50% of wildtype worms entered dauer at ~11μL daumone/mL agar (Figure 4.1A). Interestingly, *adr-1(gv6);adr-2(gv42)* double mutants required more than double the concentration of pheromone (~23μL/mL) to induce dauers at the same frequency. In a test of the single mutants, the dauer-defective phenotype segregated with the *adr-1(gv6)* allele, while the *adr-1(gv42)* worms behaved similar to wildtype.

To corroborate these results, we tested a second *adr-1* deletion allele, *adr-1(tm668)*. In contrast to what we observed for the *gv6* allele, *adr-1(tm668)* mutants, as well as *adr-1(tm668);adr-2(ok735)* double mutants, did not display a defect in dauer formation (Figure 4.1B).

The adr-1(gv6) and adr-1(tm668) alleles are similar in mRNA levels and ADAR editing profile. We wondered whether the differences in dauer formation phenotype observed in the *adr-1(gv6)* and *adr-1(tm668)* alleles could be explained by differences in ADR-1 expression or activity caused by these genomic lesions. The *adr-1(gv6)* allele, described in detail in earlier work of our laboratory (3), removes a 1560bp genomic segment, deleting both dsRBMs and resulting in an in-frame stop codon 19 amino acids downstream of the deletion point (Figure 4.2A). In the *adr-1(tm668)* allele, a 876bp deletion, spanning from exon 3 to intron 7, eliminates dsRBM1 and causes an in-frame stop 7 amino acids downstream of the deletion point. Both alleles are predicted to represent null mutations.

A northern blot for *adr-1* revealed truncated messages of the expected sizes in strains carrying either of the *adr-1* mutations (Figure 4.2B). Abundance of the truncated transcripts was reduced in comparison to strains with a wildtype *adr-1* allele. To further quantify reduction in message levels, real-time RT PCR was performed. Using PCR primer pairs for amplification of regions close to the 5' end as well as the 3' end of the message, we verified that message levels are reduced by ~90% in either deletion allele (Figure 4.2C).

Consistent with the data obtained for *adr-1* expression in *adr-1(gv6)* and *adr-1(tm668)*, the editing profile of the endogenous ADAR substrate 36A was similar

Figure 4.1. **Dauer formation in response to pheromone in wildtype and ADAR mutants.** (A) Dauer assay in N2, *adr-1(gv6)*, *adr-1(gv42)* and *adr-1(gv6);adr-2(gv42)* mutants. The *adr-1;adr-2* double mutant displays a dauer-defective phenotype that largely segregates with the *adr-1(gv6)* single mutant. The assay was performed at 17.5°C. Data represent the mean of three replicates per strain and pheromone concentration. Error bars indicate the standard error of the mean (SEM). (B) Repeated dauer formation assay in N2, 2 different *adr-1* deletion strains and 2 different *adr-1;adr-2* double-mutants. Only worms carrying the *gv6* allele, but not those carrying the *tm668* allele are dauer-defective. The assay was performed at 20°C. Data represent the mean +/- SEM of three replicates per strain and pheromone concentration.
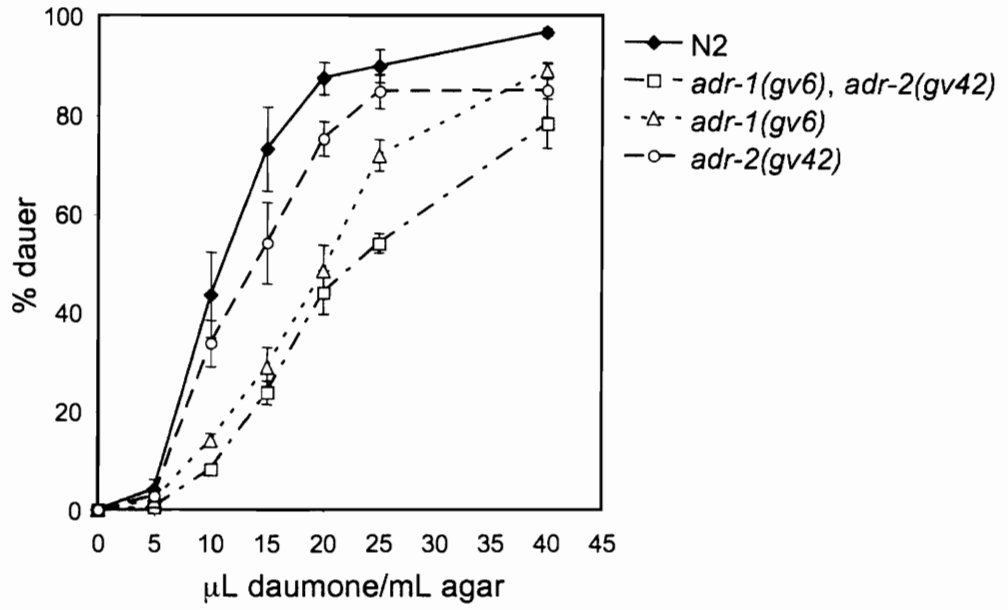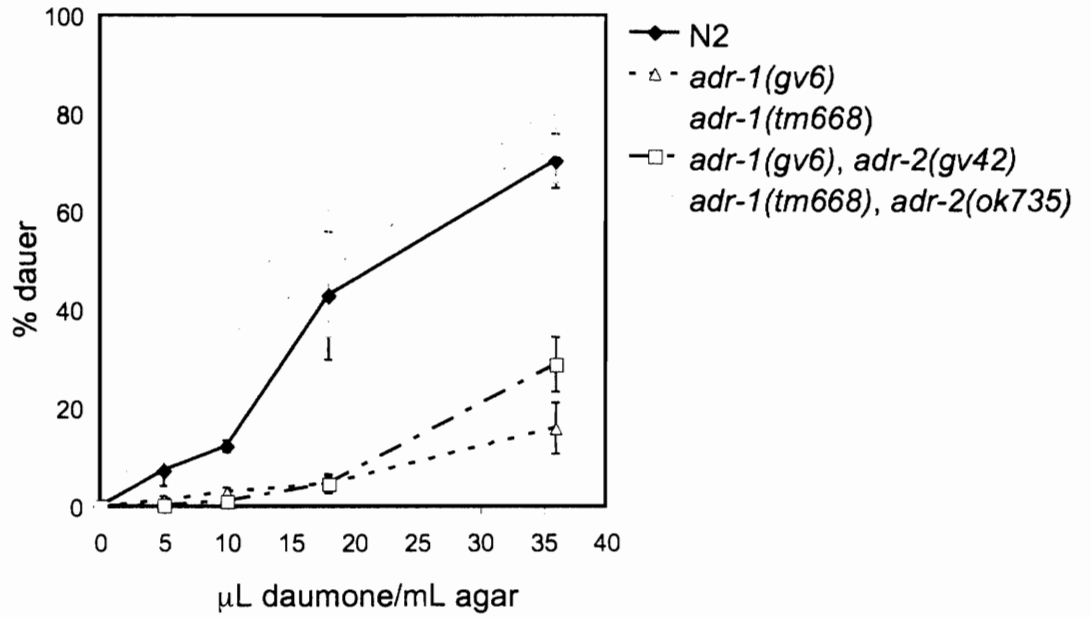
A

B

Figure 4.2. **Genomic lesion and *adr-1* mRNA expression in *adr-1(gv6)* and
adr-1(tm668).** (A) Schematic of the *adr-1* gene. White arrows indicate *adr-1* exons,
grey arrows indicate the relative positions of neighboring genes. The exons coding for
dsRNA binding domains (RBM) and the catalytic domain (CD), the regions eliminated in
the *adr-1* deletion alleles (*gv6*, *tm668*) and the genomic fragment used to construct the
rescue lines (rescue PCR product) are marked. (B) Northern blot of *adr-1* in 20μg of total
RNA prepared from N2, *adr-1(gv6)*, *adr-1(tm668)*, *adr-2(gv42)*, *adr-2(ok735)*,
adr-1(gv6);adr-2(gv42) and *adr-1(tm668);adr-2(ok735)* mixed stage worms. The
migration position of the wildtype mRNA (wt, ~2.9kb), and the truncated transcripts
caused by the *gv6* (Δ gv6, ~1.6kb) and *tm668* (Δ tm668, ~2.2kb) alleles are indicated.
Asterisks mark the very faint, crescent-shaped band of the Δ gv6 transcript. For a better
northern signal of *adr-1(gv6)* in poly(A)+ RNA refer to [3]. (C) Relative *adr-1* mRNA
abundance in N2, *adr-1(gv6)* and *adr-1(tm668)* as quantified by real-time RT-PCR. The
detected amounts of mRNA relative to wildtype levels are plotted. Data represent the
mean +/- SEM of 2 experiments. Two different primer sets to detect a 5' region (5') and
a 3' region (3') of the mRNA were used. Primer pairs mapped to regions outside the
deletions.

A

gv6

rescue PCR product

B

C
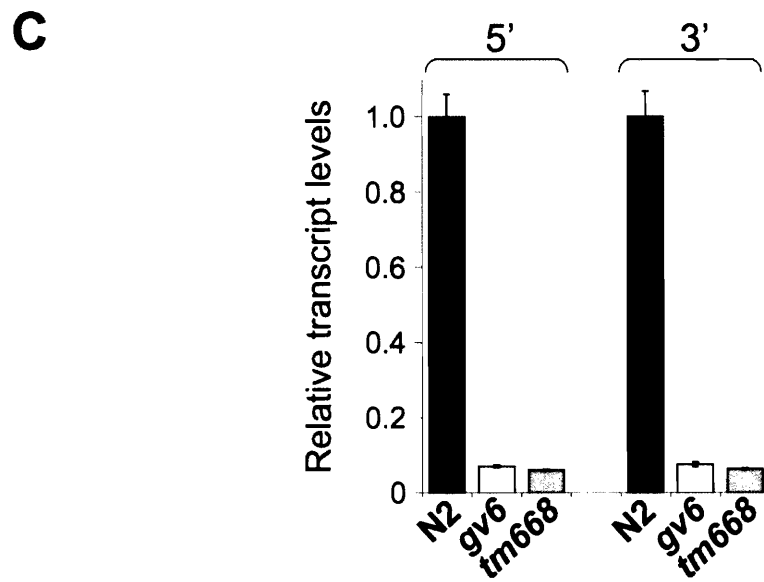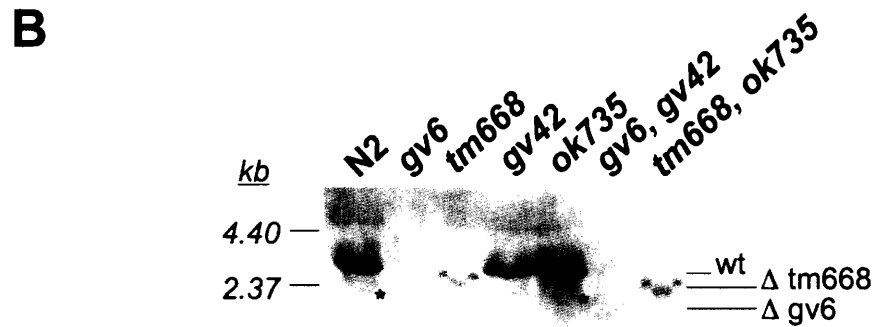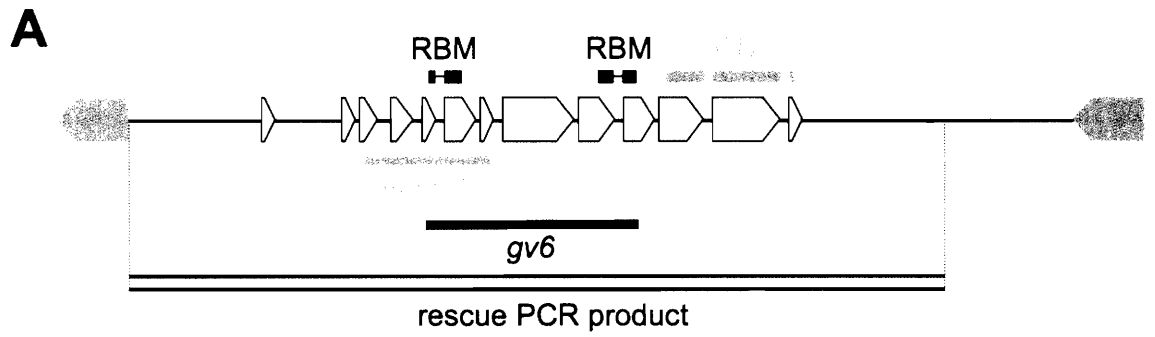
in both alleles. While having no detectable A-to-I editing activity on its own, ADR-1 is able to modulate the activity of the ADR-2 enzyme. In the absence of functional ADR-1, the editing of numerous adenosines is reduced or abolished, while other sites are enhanced. Table 4.1 summarizes a comparison of 48 editing sites in the 36A substrate in N2, *adr-1(gv6)* and *adr-1(668)* worms. While not included in Table 4.1, the identity of modulated editing sites was nearly the same in both alleles (data not shown).

In summary, a comparison of *adr-1* expression and activity in *adr-1(gv6)* and *adr-1(tm668)* revealed no differences that could account for the discrepancy in dauer formation observed in the mutants.

The adr-1(gv6) dauer-defective phenotype is not rescued by wildtype copies of adr-1. In an additional attempt to correlate the defect in dauer formation in *adr-1(gv6)* with the *adr-1* locus, we constructed rescue lines. An extrachromosomal array carrying wildtype copies of the *adr-1* gene was introduced into *adr-1(gv6)* mutants and expression of the wildtype mRNA was verified by northern blot (Figure 4.3A). Introduction of the wildtype copy of *adr-1* into *adr-1(gv6)* also restored the editing of adenosines in 36A to levels typically observed in N2 (Figure 4.3B), indicating that functional ADR-1 is produced in the rescue strains.

However, when a rescue line was subjected to a dauer formation assay, dauer formation was indistinguishable from *adr-1(gv6)* mutants (Figure 4.3C). This indicates that the observed dauer deficiency is not caused by a lesion in the *adr-1* locus.

As a last step, we tested whether *adr-1(gv6)* could have a dominant negative effect on dauer formation. A genomic fragment of the *adr-1* locus amplified from *adr-1(gv6)* mutants was injected into N2 worms. However, transgenic worms carrying

Table 4.1. **A-to-I editing of 36A in *adr-1(gv6)* and *adr-1(tm668)*.** The editing of 48 adenosines was evaluated in comparison to wildtype, and the numbers of reduced, disappeared and enhanced editing sites were scored.

Strain	Unchanged sites	Reduced (disappeared) sites	Enhanced sites	Total sites analyzed
adr-1(gv6)	3	32 (11)	2	48
adr-1(tm668)	1	27 (18)	2	48

Figure 4.3. *adr-1* **expression, 36A editing and dauer formation in** *adr-1* **rescue strains.** (A) Northern blot of *adr-1* in 20μg of total RNA prepared from N2, *adr-1(gv6)*, and 3 rescue lines. A ~2.9kb fragment (wt) representing wildtype mRNA is visible in N2 and rescue lines, while only the ~1.6kb truncated transcript is detected in the *adr-1(gv6)* mutant. The blot was rehybridized to 18S rRNA to control for loading. (B) Sequence trace of selected adenosines of the ADAR substrate 36A in N2 genomic DNA and in cDNA samples of N2, *adr-1(gv6)* and one *adr-1* rescue line. Superimposing peaks of adenosine (green) and guanosine (black) traces are indicative of A-to-I editing. The ratio of peak heights in a double peak approximates the frequency of editing at a particular site. The rescue line displays an editing pattern similar to wildtype. Numbers indicate positions of adenosines relative to the transcription start of 36A. (C) Dauer formation assay at three daumone concentrations in N2, *adr-1(gv6)*, *adr-1(tm668)* and *adr-1* rescue worms. The wildtype copy of *adr-1* in the rescue line does not restore the dauer defect of *adr-1(gv6)*. The assay was performed at 20°C. Data represent the mean + SEM of 3 to 9 replicates, depending on strain and pheromone concentration.
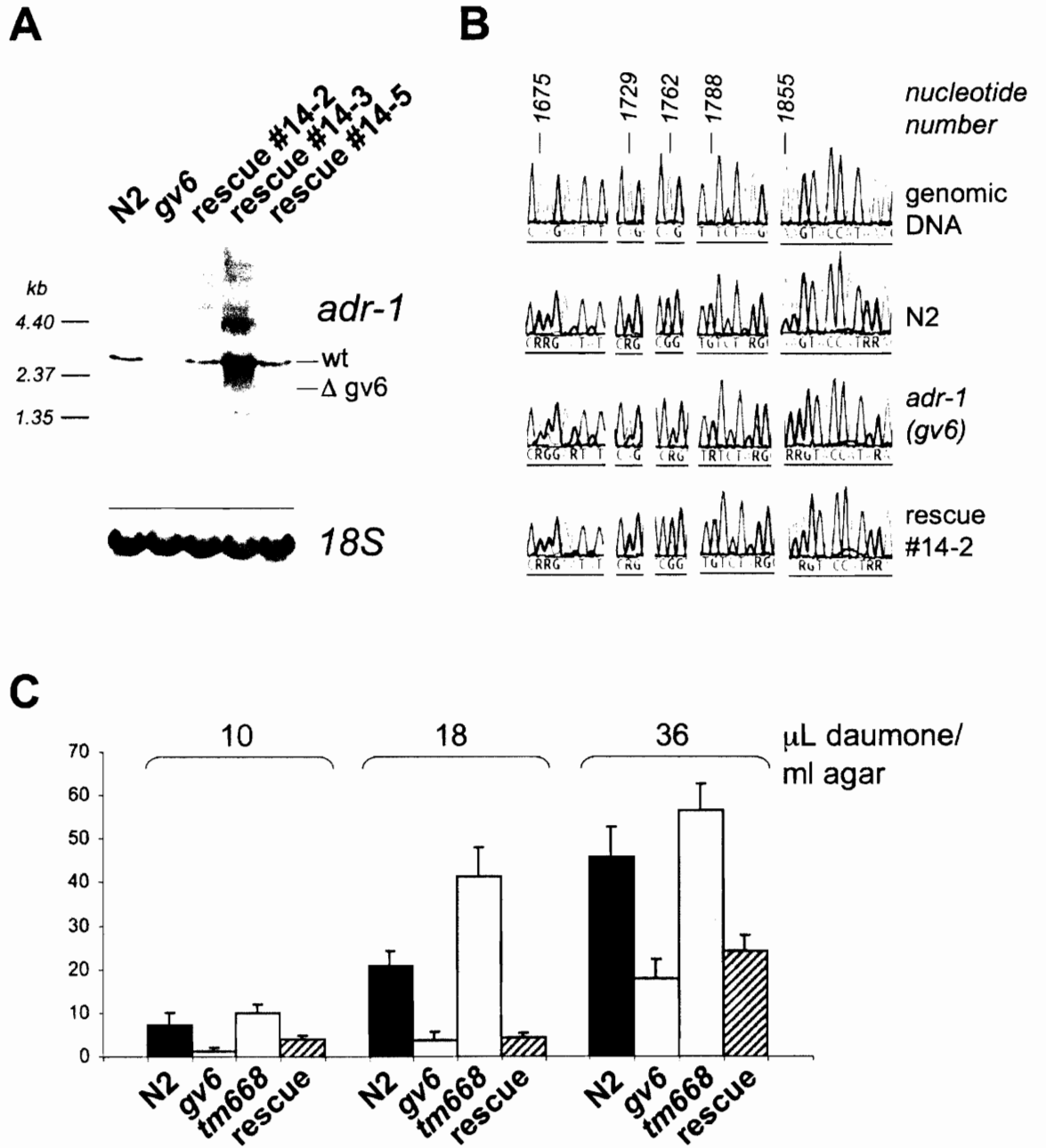
the mutant *adr-1* copy showed a wildtype response to dauer pheromone (data not shown), excluding a dominant negative effect.

We conclude that the dauer-defective phenotype observed in *adr-1(gv6)* mutants must be caused by one or more secondary mutations in the genetic background of this strain. Because the dauer-deficiency segregated with the *adr-1(gv6)* mutation through several backcrosses, it is likely that a gene closely linked to the *adr-1* locus is affected.

Discussion

The data collected in the described series of experiments revealed no dauer-deficient phenotype in *C. elegans* mutant for ADARs. This suggests that deficiency in ADARs does not cause a generalized malfunction of the nematode nervous system. Instead, A-to-I editing may be important for only a subset of neurological responses, such as the detection of chemosensory stimulants.

The *C. elegans* nervous system is comprised of 302 neurons, and much research has been directed at understanding the circuitry of this simple neuronal network (17). Laser ablation studies and mutant analyses have furthered our understanding of which neurons are involved in mediating a particular behavioral response. The functions of the 12 amphid neurons, the primary sensory neurons in *C. elegans*, are particularly well characterized. For example, the amphid neurons AWA and AWC are important for volatile chemotaxis, while ADF, ASG and ASI are critical for sensing dauer pheromone (18,28).

The identity of the ADAR substrates responsible for the chemotaxis deficiency of *C. elegans adr*-mutants is currently unknown (3). Refining our understanding of which behavioral circuits are modulated by A-to-I editing and which are not, may

ultimately prove helpful in identifying ADAR substrates with functional consequences in the nervous system. Taking the example of the amphid neurons above, a good candidate substrate could be a gene expressed in AWA and AWC, but not in ADF, ASG or ASI. Similarly, a gene with a known function in chemotaxis is more likely a candidate for editing than a gene vital for dauer formation.

References

1.   Ma, E., Gu, X.Q., Wu, X., Xu, T. and Haddad, G.G. (2001) Mutation in pre-mRNA adenosine deaminase markedly attenuates neuronal tolerance to O2 deprivation in Drosophila melanogaster. *J Clin Invest*, **107**, 685-693.

2.   Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000) A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. *Cell*, **102**, 437-449.

3.   Tonkin, L.A., Saccomanno, L., Morse, D.P., Brodigan, T., Krause, M. and Bass, B.L. (2002) RNA editing by ADARs is important for normal behavior in Caenorhabditis elegans. *EMBO J*, **21**, 6025-6035.

4.   Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R. and Seeburg, P.H. (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature*, **406**, 78-81.

5.   Hartner, J.C., Schmittwolf, C., Kispert, A., Muller, A.M., Higuchi, M. and Seeburg, P.H. (2004) Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *J Biol Chem*, **279**, 4894-4902.

6.   Wang, Q., Khillan, J., Gadue, P. and Nishikura, K. (2000) Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science*, **290**, 1765-1768.

7.   Wang, Q., Miyakoda, M., Yang, W., Khillan, J., Stachura, D.L., Weiss, M.J. and Nishikura, K. (2004) Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *J Biol Chem*, **279**, 4952-4961.

8.   Akbarian, S., Smith, M.A. and Jones, E.G. (1995) Editing for an AMPA receptor subunit RNA in prefrontal cortex and striatum in Alzheimer's disease, Huntington's disease and schizophrenia. *Brain Res*, **699**, 297-304.

9.   Kawahara, Y., Ito, K., Sun, H., Aizawa, H., Kanazawa, I. and Kwak, S. (2004) Glutamate receptors: RNA editing and death of motor neurons. *Nature*, **427**, 801.

10. Kawahara, Y., Kwak, S., Sun, H., Ito, K., Hashida, H., Aizawa, H., Jeong, S.Y. and Kanazawa, I. (2003) Human spinal motoneurons express low relative abundance of GluR2 mRNA: an implication for excitotoxicity in ALS. *J Neurochem*, **85**, 680-689.

11. Peng, P.L., Zhong, X., Tu, W., Soundarapandian, M.M., Molner, P., Zhu, D., Lau, L., Liu, S., Liu, F. and Lu, Y. (2006) ADAR2-dependent RNA editing of AMPA receptor subunit GluR2 determines vulnerability of neurons in forebrain ischemia. *Neuron*, **49**, 719-733.

12. Vollmar, W., Gloger, J., Berger, E., Kortenbruck, G., Kohling, R., Speckmann, E.J. and Musshoff, U. (2004) RNA editing (R/G site) and flip-flop splicing of the AMPA receptor subunit GluR2 in nervous tissue of epilepsy patients. *Neurobiol Dis*, **15**, 371-379.

13. Sodhi, M.S., Burnet, P.W., Makoff, A.J., Kerwin, R.W. and Harrison, P.J. (2001) RNA editing of the 5-HT(2C) receptor is reduced in schizophrenia. *Mol Psychiatry*, **6**, 373-379.

14. Gurevich, I., Tamir, H., Arango, V., Dwork, A.J., Mann, J.J. and Schmauss, C. (2002) Altered editing of serotonin 2C receptor pre-mRNA in the prefrontal cortex of depressed suicide victims. *Neuron*, **34**, 349-356.

15. Niswender, C.M., Herrick-Davis, K., Dilley, G.E., Meltzer, H.Y., Overholser, J.C., Stockmeier, C.A., Emeson, R.B. and Sanders-Bush, E. (2001) RNA editing of the human serotonin 5-HT2C receptor. alterations in suicide and implications for serotonergic pharmacotherapy. *Neuropsychopharmacology*, **24**, 478-491.

16. Yang, W., Wang, Q., Kanes, S.J., Murray, J.M. and Nishikura, K. (2004) Altered RNA editing of serotonin 5-HT2C receptor induced by interferon: implications for depression associated with cytokine therapy. *Brain Res Mol Brain Res*, **124**, 70-78.

17. Hobert, O. (2003) Behavioral plasticity in C. elegans: paradigms, circuits, genes. *J Neurobiol*, **54**, 203-223.

18. Riddle, D.L. and P.S., A. (1997) In Riddle, D., Blumenthal, T, Meyer, BJ, Priess, JR (ed.), *C. elegans II*. Cold Spring Harbor Laboratory Press.

19. Vowels, J.J. and Thomas, J.H. (1994) Multiple chemosensory defects in daf-11 and daf-21 mutants of Caenorhabditis elegans. *Genetics*, **138**, 303-316.

20. Starich, T.A., Herman, R.K., Kari, C.K., Yeh, W.H., Schackwitz, W.S., Schuyler, M.W., Collet, J., Thomas, J.H. and Riddle, D.L. (1995) Mutations affecting the chemosensory neurons of Caenorhabditis elegans. *Genetics*, **139**, 171-188.

21. Cassada, R.C. and Russell, R.L. (1975) The dauerlarva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. *Dev Biol*, **46**, 326-342.

22. Golden, J.W. and Riddle, D.L. (1982) A pheromone influences larval development in the nematode Caenorhabditis elegans. *Science*, **218**, 578-580.

23. Jeong, P.Y., Jung, M., Yim, Y.H., Kim, H., Park, M., Hong, E., Lee, W., Kim, Y.H., Kim, K. and Paik, Y.K. (2005) Chemical structure and biological activity of the Caenorhabditis elegans dauer-inducing pheromone. *Nature*, **433**, 541-545.

24. Golden, J.W. and Riddle, D.L. (1984) The Caenorhabditis elegans dauer larva: developmental effects of pheromone, food, and temperature. *Dev Biol*, **102**, 368-378.

25. Golden, J.W. and Riddle, D.L. (1984) A pheromone-induced developmental switch in Caenorhabditis elegans: Temperature-sensitive mutants reveal a wild-type temperature-dependent process. *Proc Natl Acad Sci U S A*, **81**, 819-823.

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

27. Mello, C. and Fire, A. (1995) DNA transformation. *Methods Cell Biol*, **48**, 451-482.

28. Bargmann, C.I. and Mori, I. (1997) In Riddle, D., Blumenthal, T, Meyer, BJ, Priess, JR (ed.), *C. elegans II*. Cold Spring Harbor Laboratory Press.

CHAPTER 5

PERSPECTIVES

Summary

The work described in the previous chapters explored several aspects of RNA biology in *C. elegans*. In Chapter 2, we characterized the noncoding RNA *rncs-1*, an endogenous substrate of ADARs, and presented evidence that *rncs-1* may function in a starvation stress response and can modulate gene regulation by Dicer *in vitro* and *in vivo*. In Chapter 3, determination of intracellular RNA localization by cell fractionation indicated that hyperedited ADAR substrates are not retained in the nucleus in *C. elegans* embryos. This suggests that a model for inosine-mediated nuclear retention of RNA is not universally applicable. In Chapter 4, we explored the physiological consequences of ADAR editing and demonstrated that ADAR function was not required for dauer formation in response to dauer pheromone.

Among the above, the most interesting finding described is the ability of *rncs-1* to modulate Dicer activity through competition with endogenous dsRNA. This results suggests a new model for how noncoding RNAs can exert a gene-regulatory function. The first section below describes approaches to identify other ncRNAs with properties and functions similar to *rncs-1* and to establish whether ncRNAs like *rncs-1* are molecular links between cellular stress responses and the RNAi pathway. In the second section, experiments are suggested that utilize the ability of *rncs-1* to inhibit Dicer

activity in order to identify endogenous Dicer targets and to delineate the involvement of

cofactors in the silencing pathway. Results of initial experiments are presented. A major

motivation in characterizing *rncs-1* function was to develop a model system that would

allow us study the effects of ADAR editing on the cellular fate and function of editing

substrates. The third section outlines approaches that would shed light on how A-to-I

editing modulates the function of *rncs-1* as a putative stress signal and inhibitor of RNAi.

Elucidating the role of *rncs-1* in stress

Is there a link between rncs-1 induction, starvation stress, X chromosome

nondisjunction and the RNAi pathway? In Chapter 2, we demonstrated that *rncs-1* is

induced in response to starvation. In addition, we showed that ectopic *rncs-1*

overexpression led to an increase in spontaneous male frequency and modulation of gene

silencing by Dicer.

A missing link in the data presented is the demonstration that X chromosome

nondisjunction is induced by starvation stress, possibly in a *rncs-1*-dependent manner.

As discussed in Chapter 2, initial experiments assayed determined male frequency among

progeny of hermaphrodites subjected to transient starvation stress at the L4 larval stage.

No increase in X nondisjunction was observed. To confirm or refute a connection

between male frequency and starvation, other starvation protocols must be tested.

Developmental time point and duration of starvation stress could be varied. After

identifying a protocol that induces increased X nondisjunction in wildtype worms, the

role of *rncs-1* could be tested using the null mutant and overexpressing lines.

A second missing connection is a possible link between starvation and *rncs-1*-

sensitive Dicer targets. Initial experiments indicate that B0222.3 and C30F12.6 are

moderately induced by starvation (~2-5-fold increase after 3 to 6 hrs), but this induction

is largely unchanged in *rncs-1(tm1632)* mutants (data not shown). Starvation induction

of F53A9.2 varied greatly in wildtype samples, ranging from ~30-fold to <2-fold

induction after 3 hrs (data not shown). This magnitude of fluctuation makes it impossible

to compare induction levels between wildtype and mutant strains, and starvation

conditions must be optimized to yield reproducible results within wildtype worms. The

possibility that *rncs-1* function is redundant, as discussed in the following paragraphs,

makes our results difficult to interpret.

Are there functional orthologs of rncs-1 in C. elegans or related species? In spite

of significant starvation-induction of *rncs-1*, which hints at a function for *rncs-1* in stress

response, *rncs-1*-deficient worms behaved largely wildtype in various assays of

starvation survival, dauer formation and aging (data not shown). While additional assays,

for instance for dauer recovery or survival of infection, have not been performed and may

yet reveal defects, the lack of detectable phenotypes in *rncs-1* mutants could indicate that

rncs-1 function is redundant and can be performed by other ncRNAs in its absence. Three

possible approaches to identify other ncRNAs that share properties of *rncs-1* are

described below.

First, in a candidate approach, double mutants of *rncs-1* and other known ncRNAs

could be assayed for synthetic phenotypes. An interesting candidate is the *tts-1* ncRNA.

Like *rncs-1*, *tts-1* is encoded by the X chromosome and upregulated in dauer. Deletion of

tts-1 results in no obvious phenotypic defects. It is unclear whether this enrichment in

dauer is due to lack of food or whether it is induced by dauer pheromone. A starvation

time course experiment, followed by northern or RT-PCR analyses of *tts-1* levels, could

clarify this. In contrast to *rncs-1*, *tts-1* levels are dependent on *daf-2*, a gene important for dauer pheromone, as well as starvation response. It is unclear in which tissues *tts-1* is expressed and whether ELT-2 is involved in its transcriptional regulation. Transcriptional reporter expression and RNAi feeding experiments, as have been described for *rncs-1* in Chapter 2, could reveal further similarities between both ncRNAs.

A second experimental approach could employ tiling array comparison of samples prepared from starved and well fed cultures to identify transcripts upregulated by lack of food. By specifically searching for sequences mapping to genomic regions outside of protein-coding genes, new starvation-enriched ncRNAs could be revealed. In these experiments *rncs-1* would serve as a positive control.

Third, identification of *rncs-1* promoter elements responsible for basic transcription and starvation induction could be helpful in identifying functional orthologs of *rncs-1*. In the experiments described in Chapter 2, we used ~2kb of promoter sequence for expression of *rncs-1* reporter and overexpressing lines. Shortening of this sequence, deletion of sequence stretches and mutation of GATA recognition sequences could identify key regulatory elements. A GFP reporter would be useful for initial screening of various promoter sequences and basal transcription levels and response to starvation could be subsequently quantified by northern analyses. After key promoter sequences have been characterized, BLAST searches of *C. elegans*, *C. briggsae* and *C. remanei* genomes might reveal ncRNAs under transcriptional control similar to that of *rncs-1*.

Using *rncs-1* ncRNA as a tool to study gene silencing by Dicer *in vivo*

Which Dicer-dependent gene regulation pathways are modulated by rncs-1?

Data in Chapter 2 suggested that *rncs-1* inhibits Dicer activity *in vitro* and *in vivo*. Dicer is involved in several distinct pathways of gene silencing (1,2), requiring a variety of cofactors (Figure 5.1). Because *let-7* processing was unchanged when *rncs-1* was overexpressed, inhibition of Dicer by *rncs-1* appears limited to certain tissues or to a particular pathway of Dicer-mediated silencing.

As a next step in elucidating the influence of *rncs-1* on gene silencing, expression of the five Dicer-dependent and *rncs-1*-sensitive genes identified in our study should be analyzed in strains mutant for various genes involved in Dicer-mediated pathways. Initial experiments indicate that Dicer-dependent silencing of all 5 *rncs-1*-sensitive genes is also dependent on *rde-1*, *alg-1* and *hpl-2*, but independent of *rde-4*, *rrf-1*, *rrf-2* and *alg-2* (Figure 5.2 and data not shown). *dcr-1*, *alg-1* and *hpl-2* are reported to be involved in transcriptional silencing of transgenes in the soma (3). Using primer pairs specific for unspliced transcripts, we analyzed pre-mRNA levels for B0222.3 and C30F12.6 and found levels to be increased in *rncs-1* overexpressing lines as well as in *dcr-1*, *alg-1*, *hpl-2* and *rde-1* mutants (Figure 5.3). While this suggests that Dicer-dependent silencing of *rncs-1*-sensitive genes occurs at the transcriptional level, the possibility that pre-mRNAs are degraded cannot be excluded at this point. Future experiments, such as chromatin immunoprecipitation analysis of Dicer target genes, will have to clarify this point. However, silencing at the pre-mRNA level indicates that pathways of RNA interference function in the nucleus, and B0222.3 and C30F12.6 could be the first endogenous genes identified to be subject to a nuclear, Dicer-dependent silencing pathway.

Figure 5.1. **Genes involved in distinct pathways of small-RNA-mediated gene silencing in the *C. elegans* soma.** The diagram outlines proposed pathways of gene silencing that are mediated by miRNAs and siRNAs. Genes shown by genetic and/or biochemical methods to be required for distinct steps in the pathway are indicated.
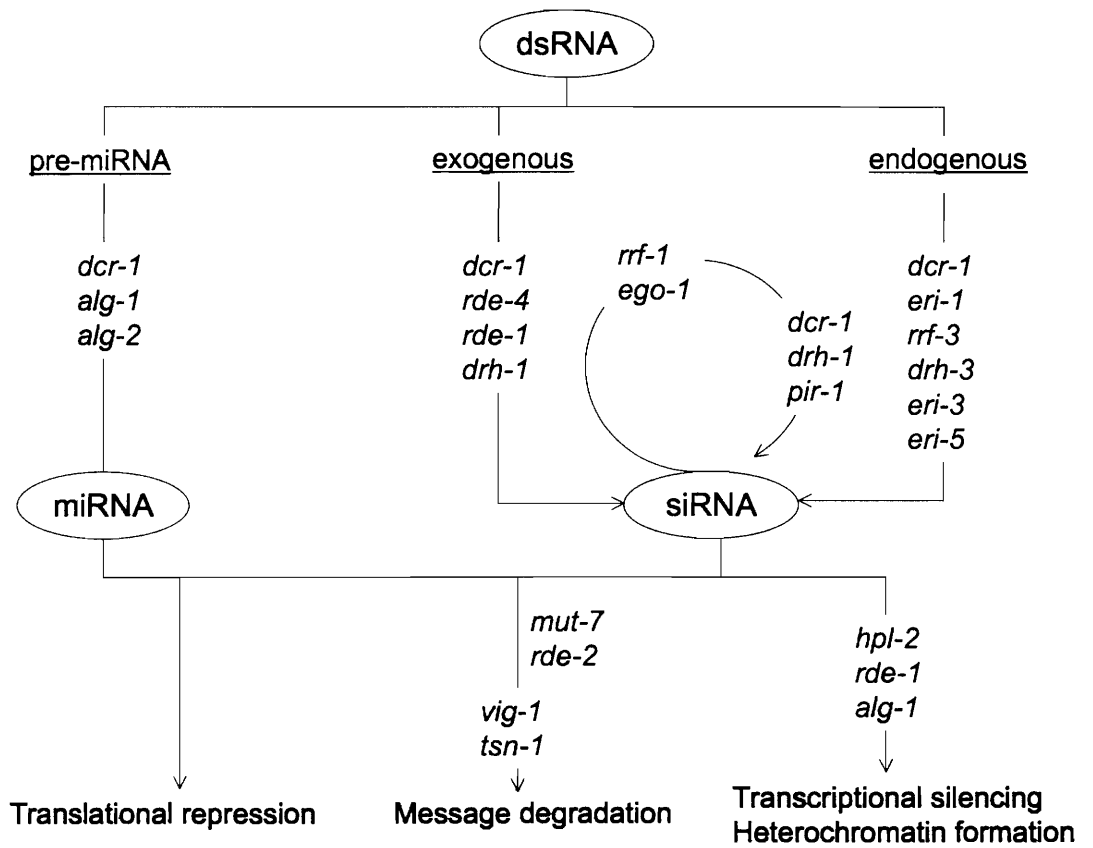
Figure 5.2. **Expression levels of *rncs-1*-sensitive genes F53A9.2, B0222.3 and C30F12.6 in selected mutants for genes implicated in small-RNA-mediated pathways.** mRNA levels were quantified by real-time PCR and normalized to *gpd-3* expression. The mean mRNA level measured in a minimum of three independent samples of young adult RNA is plotted. Error bars represent the standard error of the mean (S.E.M.). Expression levels of all three genes are significantly increased in *dcr-1*, *rde-1*, *alg-1* and *hpl-2* mutants, but remain unchanged in *rde-4*, *rrf-1*, *rrf-2* and *alg-2* deficient worms. **, P < 0.01, ***, P < 0.001.
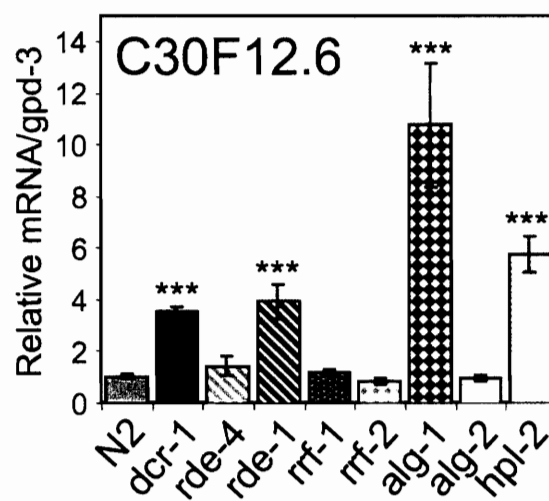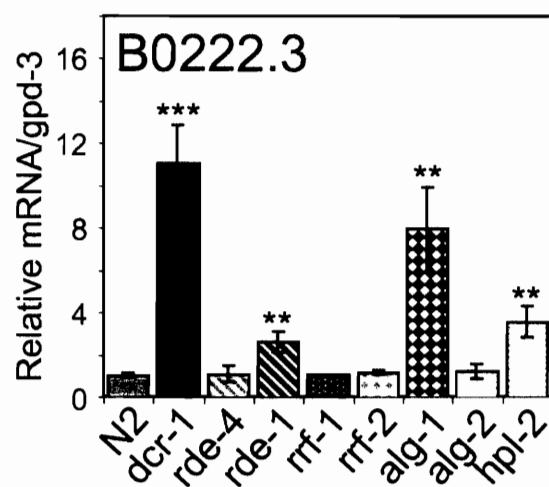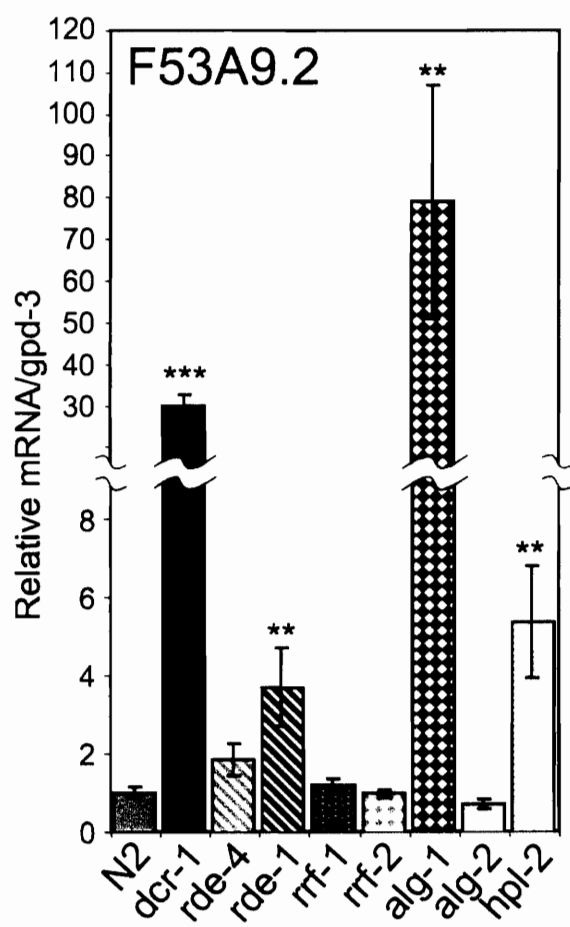
Figure 5.3. **Pre-mRNA levels of B0222.3 and C30F12.6 in *rncs-1(tm1632)*, *rncs-1*
overexpressing lines and RNAi mutants.** (A) RT-PCR products of B0222.3 and
C30F12.6 pre-mRNAs after 20 cycles of amplification using primers that map to intronic
regions. Pre-mRNA levels are visibly increased in lines that overexpress *rncs-1* in a
rncs-1-deficient (rncs-1(+)) or wildtype (N2(+)) background and in *dcr-1*, *alg-1* and
hpl-2 mutants. (B) Quantification of mRNA and pre-mRNA levels in wildtype, *rncs-1*-
deficient and *rncs-1* overexpressing lines. Changes in pre-mRNA levels parallel those of
mRNA. The mean +/- S.E.M. of three independent samples per strain is plotted. (C)
Quantification of mRNA and pre-mRNA levels in wildtype, *rncs-1* overexpressing,
dcr-1, *alg-1*, *hpl-2* and *rde-1* mutant worms, showing changes of pre-mRNA levels that
are similar to changes in mRNA levels. The mean +/- S.E.M. of three independent
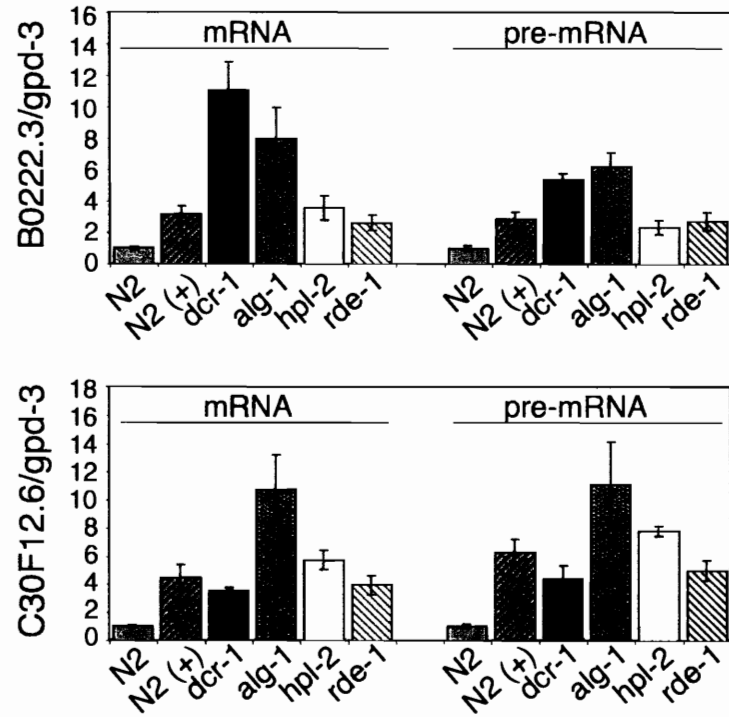samples per strain is plotted.
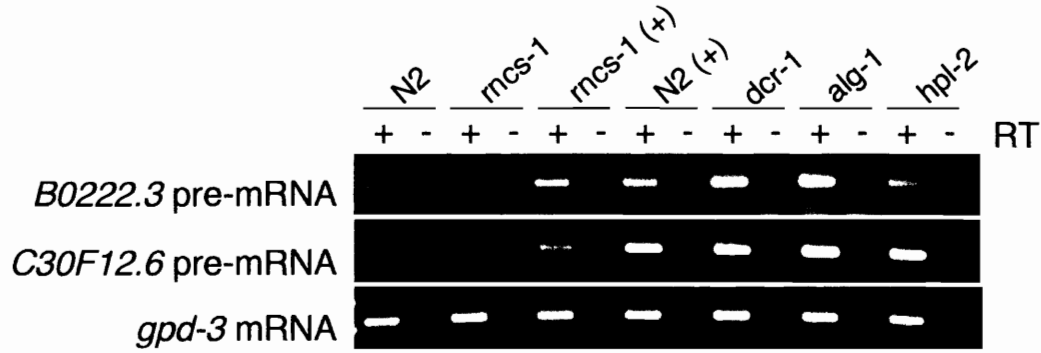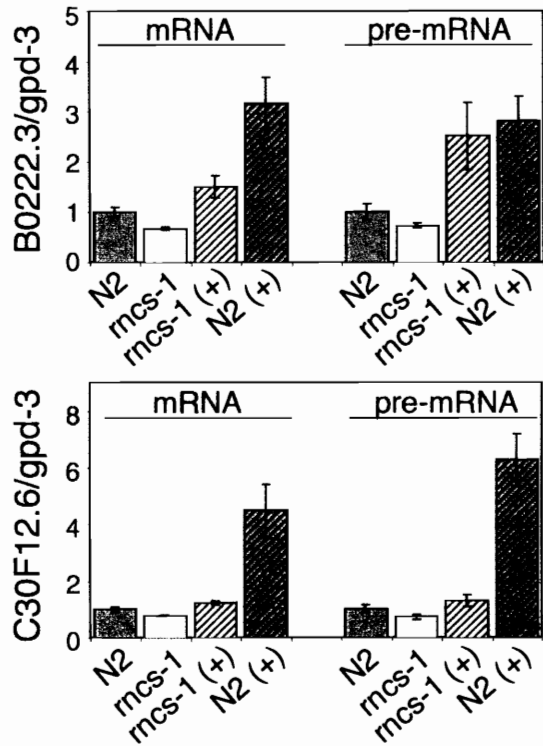
The ability to detect endo-siRNAs complementary to F53A9.2 by northern blot, allows us to identify genes involved in generating these small RNAs *in vivo*. While >1000 putative endo-siRNAs have been cloned, <10 have been validated by northern blot (1,4,5). Analyses of endo-siRNA in Dicer pathway mutants revealed that endo-siRNA generation requires *dcr-1*, *rde-2*, *rde-3*, *rde-4*, *mut-7*, *mut-14*, *rrf-3*, *eri-1*, *eri-3*, *eri-5* and *drh-3*, although dependency on these genes does not appear to be equal for all endo-siRNA tested. Detection of F53A9.2 endo-siRNAs, and possibly endo-siRNAs complementary to other *rncs-1*-sensitive genes, in RNAi pathway mutants could further our understanding of how endogenous siRNAs are generated.

Can we use rncs-1 overexpression to identify genes regulated by Dicer-dependent endo-siRNAs? Deep sequencing of small RNAs in *C. elegans* revealed that most endogenous siRNAs are not direct products of Dicer cleavage (4,5). Moreover, microarray analyses of Dicer-regulated endogenous mRNAs did not show good correlation with sequenced endo-siRNAs (6). Dicer-dependent, endo-siRNA-mediated regulation has only been clearly shown for a handful of genes, including F53A9.2, identified by us. It remains largely unclear how frequently this pathway is employed as a mechanism for expression regulation *in vivo*.

Our data indicated that *rncs-1* overexpression is able to specifically interfere with F53A9.2 endo-siRNA generation by Dicer, coinciding with upregulation of F53A9.2 mRNA. We can use this observation to search for other genes regulated in a Dicer-dependent and *rncs-1*-sensitive manner using a microarray approach. Looking for an overlapping dataset of genes upregulated in *dcr-1*-mutants and *rncs-1* overexpressing lines has several advantages. First, we are examining the consequences of reduced Dicer

activity—caused by genomic mutation or by functional inhibition by *rncs-1*—in different genetic backgrounds. The effects of secondary mutations in either strain will therefore be eliminated. Second, the inhibitory effect of *rncs-1* overexpression appears to be tissue-specific or does not influence miRNA processing. The overlapping dataset of *dcr-1*-dependent and *rncs-1*-sensitive genes is, therefore, expected to be small. Third, depending on the origin and structure of the dsRNA substrate, Dicer may require different combinations of cofactors, such as RDE-4, argonautes, RNA-dependent polymerases or helicases, to produce small RNAs and to silence genes. Beyond a requirement for Dicer-dependence, the microarray experiment described here will reveal regulated genes regardless of dependency on additional cofactors.

Our observation that most mRNAs with *dcr-1*-dependence and *rncs-1*-sensitivity were enriched in the male soma suggested that Dicer may function to suppress somatic male fate in the hermaphrodite. Including samples of wildtype males in the microarray comparison above could substantiate or refute the validity of this model.

The effect of editing by ADARs on *rncs-1* function

How does ADAR editing of rncs-1 influence modulation of endogenous siRNA formation? Results presented in Chapter 2 demonstrated that overexpression of *rncs-1* caused inhibition of endogenous siRNA (endo-siRNA) production and upregulation of the corresponding mRNAs. All experiments were performed in worms wildtype for ADARs. By sequencing we verified that *rncs-1* was edited in overexpressing lines and found that A-to-I editing levels were similar when *rncs-1* was present at wildtype or increased levels (data not shown). While ADAR editing does not influence the induction

of *rncs-1*, it is unclear how hyperediting affects the molecular and physiological consequences of *rncs-1* overexpression.

We proposed that the inhibition of endo-siRNA formation is caused by sequestration of Dicer by *rncs-1* through dsRBM-dsRNA interaction. Because A-to-I editing diminishes the double-stranded character of *rncs-1*, we might expect that lack of ADAR editing can influence its ability to function as a competitive inhibitor of dsRNA cleavage. Future experiments should, therefore, be directed at studying the effects of *rncs-1* overexpression and deficiency in an ADAR-mutant background. Experiments should include quantifying mRNA levels and endo-siRNAs of F53A9.2 and other *rncs-1*-sensitive genes when unedited *rncs-1* is overexpressed. Strains that express wildtype levels of *rncs-1* but are mutant for *adr-1* and/or *adr-2* must be included as controls.

These experiments might indicate that *rncs-1* is a more potent antagonist of Dicer activity in the absence of editing. This is intuitively expected, as unedited *rncs-1* is likely to be a better binding partner of dsRBMs. However, inhibition of endo-siRNA formation by *rncs-1* is based on a model of competition with endogenous dsRNA. ADAR activity is also expected to edit the dsRNA competing with *rncs-1*, and elimination of A-to-I editing will therefore also make the endogenous dsRNA a more efficient binding partner of Dicer. Therefore, it is possible that the competition between *rncs-1* and endogenous dsRNA is only minimally, if at all, perturbed in ADAR-mutants.

How does ADAR editing influence the role of rncs-1 as a stress signal? Induction of *rncs-1* in response to starvation is transcriptional and independent of ADAR editing (Figure 2.3D and data not shown). In Chapter 2 we proposed that overexpression of *rncs-1* or dsRNA of any sequence signals cellular stress and leads to increased

spontaneous male frequency, possibly through inhibition of endogenous RNAi pathways. As described above, elimination of A-to-I enhances the double-stranded character of dsRNA and potentially may alter the ability of *rncs-1* to signal stress.

To test this possibility strains overexpressing *rncs-1* or the P_{rncs-1}:*GFP* transgene, which gives rise to low levels of dsRNA, could be assayed for spontaneous male frequency.

If dsRNA was indeed a cellular stress signal, male frequency will be enhanced further when unedited dsRNA is overexpressed. Alternatively it is possible that the increase of cellular inosine, as a result of edited dsRNA overexpression, could be a molecular marker required for signaling stress through dsRNA. In the latter case, male frequency in dsRNA overexpressing lines will be identical to levels observed in ADAR-mutants with wildtype dsRNA levels.

While only at the early stages of experimentation, initial, and not yet quantified, observations indicate that crossing the P_{rncs-1}:*GFP* transgene into *adr-1;adr-2* mutants increases spontaneous male frequency in comparison to P_{rncs-1}:*GFP* expression in wildtype.

References

1.   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.

2.   Grishok, A. (2005) RNAi mechanisms in Caenorhabditis elegans. *FEBS Lett*, **579**, 5932-5939.

3.   Grishok, A., Sinskey, J.L. and Sharp, P.A. (2005) Transcriptional silencing of a transgene by RNAi in the soma of C. elegans. *Genes Dev*, **19**, 683-696.

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

5.   Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H. and Bartel, D.P. (2006) Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in C. elegans. *Cell*, **127**, 1193-1207.

6.   Welker, N.C., Habig, J.W. and Bass, B.L. (2007) Genes misregulated in C. elegans deficient in Dicer, RDE-4, or RDE-1 are enriched for innate immunity genes. *RNA*.