

SAMIR BOUASKER

**PARTICIPATION DE L'ACTIVITÉ  
ENDONUCLÉASIQUE DES PROTÉINES  
ARGONAUTES ALG-1 ET ALG-2 DANS LA  
MATURATION DES miARN CHEZ *C. ELEGANS***

Thèse présentée

à la Faculté des études supérieures et postdoctorales de l'Université Laval  
dans le cadre du programme de doctorat en biologie Cellulaire et Moléculaire  
pour l'obtention du grade de Philosophiae Doctor (Ph.D)

DÉPARTEMENT DE BIOLOGIE MOLÉCULAIRE,  
BIOCHIMIE MÉDICALE ET PATHOLOGIE  
FACULTÉ DE MÉDECINE  
UNIVERSITÉ LAVAL  
QUÉBEC

2012

© Samir Bouasker, 2012

## Résumé

Au sein de l'ensemble des protéines Argonautes codées par les génomes des organismes métazoaires, certains membres de cette famille de protéines ont conservé des acides aminés importants pour l'activité endonucléasique. La signification fonctionnelle de ces résidus (composés des résidus DDH), pour les Argonautes spécifiques des miARN chez les animaux, est encore inconnue puisque le ciblage des ARNm par les miARN ne provoque pas de clivage site spécifique comme c'est le cas pour les siARN. *In vitro*, nous avons mis en évidence, chez le nématode *C. elegans*, que les protéines Argonautes spécifiques aux miARN, ALG-1 et ALG-2, conservant ce motif, possèdent une activité de clivage similaire à celle impliquée dans la voie des siARN. Nous démontrons également que les protéines Argonautes ALG-1 et ALG-2 ont la capacité de lier et d'utiliser comme substrats différents duplexes de courts ARN, similaires aux duplexes de miARN produits par l'enzyme DCR-1. Les Argonautes ALG-1 et ALG-2 sont capables de coupure endonucléolytique sur ces duplexes, lorsque le degré de complémentarité entre les deux brins le permet, et de séparer les deux brins d'un duplex de courts ARN contenant des mésappariements. *In vivo*, l'activité endonucléasique de ALG-1 ou ALG-2 est essentielle pour la voie des miARN, et la perte de cette activité conduit à une accumulation des précurseurs de miARN tronqués et une altération de la formation de miRISC fonctionnel. Prises dans leur ensemble, nos données indiquent que l'activité endonucléasique des protéines Argonaute ALG-1 ou ALG-2 contribue à la maturation des miARN chez le nématode *C. elegans*.

## Abstract

Within the Argonaute protein family, encoded by the genomes of metazoan organisms, some members have conserved amino acids important for the endonuclease activity. The functional significance of these residues (DDH), for the miRNAs-specific Argonaute proteins in animals, is still unknown, because mRNA targeting by miRNAs does not trigger cleavage at any specific site. We have demonstrated, in the nematode *C. elegans*, that the miRNA-specific Argonaute protein, ALG-1 and ALG-2, have *in vitro* cleavage activity, similar to what is observed for siRNA pathway. We also demonstrate that the Argonaute proteins ALG-1 and ALG-2 have the ability to bind and use different RNA substrates as short RNA duplexes, similar to miRNA duplexes produced by the DCR-1 enzyme. The Argonautes ALG-1 and ALG-2 are capable of endonucleolytic cleavage of these duplexes, when the degree of complementarity between the two strands allows it, and are able to separate the two strands of short RNA duplex containing mismatches. *In vivo*, the endonuclease activity of ALG-1 or ALG-2 is essential for the miRNA pathway, and loss of this activity leads to an accumulation of truncated miRNA precursors and impairs the formation of functional miRISC. Altogether, our data indicate that the endonuclease activity of the Argonaute ALG-1 and ALG-2 contributes to the maturation of miRNAs in the nematode *C. elegans*.

## Avant-Propos

Dans le cadre de mon doctorat, j'ai eu la chance de travailler dans le domaine de la recherche sur la biogenèse des miARN et j'ai contribué à plusieurs projets dont certains ont mené à la publication d'articles. La production de protéines Argonaute recombinante m'a permis notamment de travailler sur l'interaction entre la protéine ribosomale RACK-1 et la protéine Argonaute ALG-1, publiée en 2011. Des essais mis au point sur la protéine PRG-1 seront décrits dans cette thèse et font partie d'une collaboration avec le Dr Miska, dont le manuscrit est publié dans le journal «Science». Certains travaux portant sur la caractérisation de mutants d'ALG-1 du groupe de Dr Ambros font actuellement l'objet d'une étude et ne seront pas décrits. Enfin, mon projet de recherche principal est publié dans le journal «Nucleic Acids Research» et se trouve en Annexe 1.

Avant de commencer, je tiens à remercier le Dr Martin Simard de m'avoir accueilli dans un laboratoire stimulant et sur un sujet de recherche aussi captivant. Je tiens également à remercier les différentes personnes du laboratoire, notamment Alejandro et Irfan pour les nombreuses et longues discussions scientifiques ainsi que Gabriel, Nellie et Johannie pour leur support. L'expérience que j'ai acquise durant ces années a indéniablement renforcé mon désir de continuer une aventure en recherche.

J'aimerais aussi remercier particulièrement les proches qui m'ont soutenu durant cette thèse. Tout d'abord Nellie, qui a partagé les plus durs moments avec moi et qui a été sans aucun doute ma plus grande réussite durant ce doctorat. Ensuite, les amis de «galère» Samuel, Laure, Sophie et Bertrand avec qui j'ai vécu des moments inoubliables. Enfin, une pensée pour toute ma famille et notamment mon père, qui n'a pas pu voir le terme de ce projet, ainsi que ma mère, Ralia, Lounès et Yazid. Et évidemment, un remerciement au Tout-Puissant, sans qui je ne serais rien.



# Table des matières

|  |     |
|--|-----|
| Résumé.....  | i   |
| Abstract.....  | ii  |
| Avant-Propos .....   | iii |
| Table des matières .....   | v   |
| Liste des tableaux.....  | ix  |
| Liste des figures .....  | x   |
| Listes des Abréviations.....   | xii |
| Introduction.....  | 1   |
| Chapitre 1 : Revue de littérature .....                                  | 3   |
| 1. Les miARN .....   | 4   |
| 1.1. Le nématode <i>C. elegans</i> .....                                 | 4   |
| 1.1.1. Généralités .....   | 4   |
| 1.1.2. Développement embryonnaire.....                                   | 5   |
| 1.1.3. Les différents stades larvaires.....                              | 7   |
| 1.1.4. Les miARN et le développement du nématode <i>C. elegans</i> ..... | 8   |
| 1.1.4.1. Le miARN lin-4.....   | 8   |
| 1.1.4.2. Le miARN let-7 .....  | 8   |
| 1.1.4.3. Division des cellules HSC .....                                 | 9   |
| 1.1.4.4. Développement de la vulve.....                                  | 10  |
| 1.1.4.5. Développement neuronal .....                                    | 11  |
| 1.1.4.6. Développement embryonnaire.....                                 | 11  |
| 1.2. Fonction des miARN dans le développement des mammifères .....       | 12  |
| 1.2.1. Régulation du système nerveux .....                               | 12  |
| 1.2.2. Développement du système cardio-vasculaire.....                   | 13  |
| 1.2.3. Développement des organes .....                                   | 14  |
| 1.2.4. Système immunitaire et miARN.....                                 | 15  |
| 1.3. miARN dans les cancers .....  | 16  |
| 1.3.1. miARN suppresseur de tumeurs .....                                | 16  |
| 1.3.2. OncomiRs .....  | 17  |
| 1.3.3. Les miARN et la spécificité cellulaire dans les cancers .....     | 18  |
| 1.4. La biogénèse des miARN .....  | 19  |
| 1.4.1. Organisation génomique et transcription.....                      | 20  |
| 1.4.2. Maturation des miARN.....   | 22  |
| 1.4.2.1. Description des RNases de type III.....                         | 22  |
| 1.4.2.2. Maturation nucléaire par Drosha .....                           | 23  |
| 1.4.2.3. Maturation cytoplasmique par Dicer .....                        | 25  |
| 1.4.2.4. Les cofacteurs de Dicer : le complexe RLC .....                 | 26  |
| 1.4.2.5. Dicer et les protéines Argonautes .....                         | 27  |
| 1.4.3. Mécanisme d'inhibition de la traduction par les miARN .....       | 28  |
| 1.4.3.1. Interaction entre le miARN et sa cible.....                     | 28  |

|  |   |    |
|--|---|----|
| 1.4.3.2.                                 | La machinerie de traduction .....                             | 29 |
| 1.4.3.3.                                 | Inhibition de la traduction médié par les miARN .....         | 30 |
| 1.4.3.4.                                 | Devenir de l'ARNm ciblé .....                                 | 31 |
| 1.4.4.                                   | Stabilité des miARN .....                                     | 33 |
| 2.                                       | Les siARN .....   | 34 |
| 2.1.                                     | Généralités .....   | 34 |
| 2.2.                                     | La voie de l'exo-ARNi .....                                   | 34 |
| 2.2.1.                                   | La transmission systémique de l'ARNdb .....                   | 35 |
| 2.2.2.                                   | Synthèse des siARN primaires .....                            | 35 |
| 2.2.3.                                   | Synthèse des siARN secondaires .....                          | 36 |
| 3.                                       | Les piARN .....   | 38 |
| 3.1.                                     | Découverte et fonction des piARN .....                        | 38 |
| 3.2.                                     | Synthèse des piARN chez les mammifères et la drosophile ..... | 39 |
| 3.3.                                     | Les piARN chez <i>C. elegans</i> .....                        | 40 |
| 4.                                       | Les protéines Argonaute .....                                 | 41 |
| 4.1.                                     | Généralités .....   | 41 |
| 4.1.1.                                   | Classification des protéines Argonaute .....                  | 41 |
| 4.1.2.                                   | Les protéines Argonaute procaryotiques .....                  | 42 |
| 4.2.                                     | Structure et fonction des protéines Argonaute .....           | 44 |
| 4.2.1.                                   | Le domaine PAZ .....  | 44 |
| 4.2.2.                                   | Le domaine Mid .....  | 46 |
| 4.2.3.                                   | Le domaine PIWI .....   | 47 |
| 4.2.3.1.                                 | Caractérisation de l'activité RISC .....                      | 48 |
| 4.2.3.2.                                 | Caractérisation du domaine PIWI .....                         | 48 |
| 4.2.4.                                   | Structures cristallines de complexes Argonaute-ARN .....      | 49 |
| 4.2.4.1.                                 | Avant propos .....  | 50 |
| 4.2.4.2.                                 | Tracing Argonaute binding .....                               | 50 |
| 5.                                       | Contexte et hypothèse de recherche .....                      | 55 |
| 6.                                       | Objectifs de recherche .....                                  | 56 |
| Chapitre 2 : Matériels et méthodes ..... |   | 57 |
| 1.                                       | Biologie moléculaire .....                                    | 57 |
| 1.1.                                     | Création d'ADNc .....   | 57 |
| 1.2.                                     | Clonage .....   | 57 |
| 1.2.1.                                   | Plasmides et amorces utilisés .....                           | 57 |
| 1.2.2.                                   | Kits utilisés .....   | 58 |
| 1.2.3.                                   | Transformation bactérienne .....                              | 58 |
| 1.3.                                     | Purification de protéines .....                               | 59 |
| 1.4.                                     | Tests de clivage et «Filter Binding Assay» .....              | 61 |
| 1.4.1.                                   | Radiomarquage des ARN de 21 nt ou 32 nt en 5' .....           | 61 |
| 1.4.2.                                   | Réaction de clivage in vitro et «Filter Binding Assay» .....  | 61 |
| 2.                                       | Manipulation de <i>C. elegans</i> .....                       | 63 |
| 2.1.                                     | Création de lignées transgéniques .....                       | 63 |
| 2.2.                                     | Extraction d'ARN total .....                                  | 63 |
| 2.3.                                     | Analyse de type Northern .....                                | 63 |
| 2.4.                                     | Clonage de courts ARN par 3' et 5' RACE : .....               | 65 |
| 2.5.                                     | RT-PCR .....  | 65 |
| 2.6.                                     | Extraction de protéines totales .....                         | 66 |

|  |  |     |
|--|--|-----|
| 2.7.   | Identification de miRISC par oligonucléotides 2'-O-méthylés .....  | 67  |
| 2.8.   | Immunoprécipitation de protéines Argonaute .....   | 67  |
| Chapitre 3 : Caractérisation de l'activité des protéines Argonaute <i>in vitro</i> .....                     |  | 69  |
| 1.   | Détermination de l'activité endonucléasique des Argonaute possédant le motif catalytique .....   | 70  |
| 1.1.   | Détermination de l'activité endonucléasique potentielle de rRDE-1 et rSAGO-1 .....   | 70  |
| 1.2.   | Détermination de l'activité endonucléasique potentielle de rPRG-1 .....  | 72  |
| 1.3.   | Détermination de l'activité endonucléasique potentielle de rALG-1 et rALG-2 .....  | 74  |
| 1.3.1.   | Mise en évidence de l'activité endonucléasique de rALG-1 et rALG-2 .....   | 75  |
| 1.3.2.   | Comparaison de l'activité de rALG-1 et rALG-2 .....  | 76  |
| 1.3.3.   | Dépendance du motif DDH pour l'activité endonucléasique .....  | 77  |
| 1.3.4.   | Autres domaines influençant l'activité de rALG-1 et rALG-2 .....   | 80  |
| 1.3.5.   | Importance du nucléotide 5'terminal pour le clivage de rALG-1/2 .....  | 81  |
| 2.   | Activité de rALG-1 et rALG-2 sur les duplexes de courts ARN et les pré-miARN .....   | 83  |
| 2.1.   | Test de l'activité endonucléasique sur les duplexes de courts ARN .....  | 84  |
| 2.2.   | Courbes de liaison aux duplexes de courts ARN .....  | 85  |
| 2.3.   | Détermination d'une activité de relarguage du brin passager d'un duplex de courts ARN .....  | 86  |
| 2.4.   | Mise en évidence de l'implication du domaine PAZ dans la prise en charge des duplexes .....  | 88  |
| 2.5.   | Activité de rALG-1 et rALG-2 sur le pré-miARN let-7 .....  | 89  |
| 3.   | Conclusions sur l'étude <i>in vitro</i> des Argonaute .....  | 90  |
| Chapitre 4 : Caractérisation de l'activité des protéines ALG-1/2 dans la voie des miARN <i>in vivo</i> ..... |  | 91  |
| 1.   | Analyse des mutants nuls <i>alg-1</i> et <i>alg-2</i> .....  | 91  |
| 1.1.   | Phénotypes observés .....  | 91  |
| 1.2.   | Dépendance des miARN let-7 et lin-4 vis-à-vis de ALG-1 .....   | 93  |
| 2.   | Analyse phénotypique des lignées transgéniques <i>alg-1</i> et <i>alg-2</i> sauvages et mutantes catalytiques .....  | 95  |
| 2.1.   | Analyse de la restitution de fonction des lignées transgéniques <i>alg-1</i> et <i>alg-2</i> ....  | 95  |
| 2.2.   | Analyse de la perte de <i>alg-1</i> et <i>alg-2</i> endogène par ARNi sur les lignées transgéniques .....  | 96  |
| 2.3.   | Analyse biochimique des lignées transgéniques <i>alg-1</i> et <i>alg-2</i> sauvages et mutantes catalytiques .....   | 98  |
| 2.3.1.   | Effets de la perte d'activité endonucléasique de <i>alg-1</i> et <i>alg-2</i> sur la maturation des miARN .....  | 99  |
| 2.3.1.1.   | Les miARN let-7 et lin-4 .....   | 99  |
| 2.3.1.2.   | Les miARN miR-90 et miR-40 .....   | 101 |
| 2.3.1.3.   | Les miARN introniques miR-50 et miR-62 .....   | 103 |
| 2.4.   | Détermination du site de troncation du précurseur de miARN .....   | 104 |
| 2.5.   | Effets de la perte d'activité endonucléasique de <i>alg-1/2</i> sur l'association aux miARN .....  | 106 |
| 2.5.1.   | miARN associés aux protéines Argonaute chez <i>alg-1 3a</i> ; <i>alg-2(ARNi)</i> et <i>alg-2 3a</i> ; <i>alg-1(ARNi)</i> : le cas particulier de lin-4 ..... | 107 |

|   |  |     |
|---|--|-----|
| 2.5.2.  | miARN associées aux protéines Argonaute chez <i>alg-1 3a</i> ; <i>alg-2(ARNi)</i> et <i>alg-2 3a</i> ; <i>alg-1(ARNi)</i> : analyse de l'association de <i>let-7</i> ..... | 108 |
| 2.5.3.  | Investigation de l'accumulation de miARN chez <i>alg-1 3a</i> ; <i>alg-2(ARNi)</i> et <i>alg-2 3a</i> ; <i>alg-1(ARNi)</i> .....   | 109 |
| 2.5.4.  | Conséquence de la perte d'activité endonucléasique sur la formation de complexe miRISC détecté par 2'-O-méthyl .....   | 110 |
| 2.6.  | Effet de la perte d'activité endonucléasique de ALG-1 et ALG-2 sur la stabilité d'ARNm régulée par des miARN.....  | 111 |
| 2.6.1.  | Stabilité des ARNm régulé par des miARN dans les mutants nuls <i>alg-1</i> et <i>alg-2</i> .....   | 112 |
| 2.6.2.  | Stabilité des ARNm régulé par des miARN dans les lignées transgéniques .....   | 113 |
| 2.7.  | Exploration des différents complexes formés par les protéines Argonaute chez <i>alg-1 3a</i> ; <i>alg-2(ARNi)</i> et <i>alg-2 3a</i> ; <i>alg-1(ARNi)</i> .....            | 114 |
| Chapitre 5 : Discussion et perspectives ..... |  | 117 |
| 1.  | Détermination de l'activité endonucléasique des protéines Argonaute de <i>C. elegans</i> .....   | 117 |
| 1.1.  | Activité endonucléasique des protéines Argonaute recombinantes rRDE-1, rSAGO-1 et rPRG-1.....  | 117 |
| 1.1.1.  | Activité de rRDE-1 et rSAGO-1.....   | 117 |
| 1.1.2.  | Activité de rPRG-1 .....   | 118 |
| 1.2.  | Activité endonucléasique des protéines Argonaute recombinantes rALG-1 et rALG-2.....   | 118 |
| 1.2.1.  | Implication de différents domaines pour l'activité endonucléasique .....   | 120 |
| 1.2.2.  | Déterminants nécessaires au niveau de la molécule d'ARNsb guide permettant le clivage d'un ARN cible .....   | 121 |
| 2.  | Signification <i>in vivo</i> de l'utilisation de différents substrats d'ARN <i>in vitro</i> par ALG-1 et ALG-2.....  | 122 |
| 2.1.  | Action au niveau du précurseur .....   | 122 |
| 2.2.  | Action au niveau du duplex de miARN.....   | 123 |
| 2.3.  | Origine du précurseur tronqué .....  | 124 |
| 2.4.  | Modèle de l'intervention catalytique de ALG-1/2.....   | 126 |
| 3.  | Différences fonctionnelles entre ALG-1/2 .....   | 127 |
| 3.1.  | miARN spécifiques à ALG-1 ou ALG-2.....  | 127 |
| 3.2.  | Rôle différentiel de l'activité de ALG-1 et ALG-2 .....  | 128 |
| 3.3.  | Rôle de ALG-1 à différents niveaux.....  | 129 |
| 4.  | Conclusions et perspectives.....   | 130 |
| Conclusion .....                              |  | 131 |
| Bibliographie .....                           |  | 132 |
| Annexe 1 .....                                |  | 152 |

## Liste des tableaux

|  |    |
|--|----|
| Tableau 1 : Description de différentes classes de courts ARN.....  | 4  |
| Tableau 2 : Amorces utilisées pour les différentes constructions.....                                    | 59 |
| Tableau 3 : Présentation des différents mutants pour les protéines rALG-1 et rALG-2 .....                | 80 |
| Tableau 4 : Phénotypes et interaction génétique entre les gènes <i>alg-1</i> et <i>alg-2</i> . .....     | 92 |
| Tableau 5 : Liaison génétique des gènes <i>alg-1</i> , <i>alg-2</i> , <i>let-7</i> et <i>lin-4</i> ..... | 95 |
| Tableau 6 : Expérience d'ARNi sur les différentes lignées transgéniques.....                             | 98 |

## Liste des figures

|   |     |
|---|-----|
| Figure 1 : Description schématique des différents organes composant un nématode hermaphrodite au stade adulte .....                                       | 5   |
| Figure 2 : Représentation des différentes phases de développement du ver.. .....  | 6   |
| Figure 3 : Cascade de régulation hétérochronique contrôlant la division des cellules HSC chez <i>C. elegans</i> .....                                     | 10  |
| Figure 4 : Biogenèse des miARN chez les mammifères.. .....  | 20  |
| Figure 5 : Organisation de transcrits de miARN. ....  | 21  |
| Figure 6 : Domaines représentés chez les 3 classes d'enzymes RNases III. ....   | 23  |
| Figure 7 : Complexe Drosha-DGCR8-pri-miARN.....   | 24  |
| Figure 8 : Complexe Dicer/pré-miARN.. .....   | 26  |
| Figure 9 : Représentation de l'appariement 3'UTR-miARN.....   | 29  |
| Figure 10 : Complexe d'initiation de la traduction.....   | 30  |
| Figure 11 : Voie des siARN chez <i>C. elegans</i> .....   | 37  |
| Figure 12 : Modèle ping-pong de production des piARN.. .....  | 40  |
| Figure 13 : Arbre phylogénétique des protéines Argonaute.....   | 43  |
| Figure 14 : Domaines caractéristiques d'une protéine AGO .....  | 44  |
| Figure 15 : Comparaison entre la structure du domaine PAZ de la protéine AGO2 de drosophile et de la protéine Rho chez <i>E. coli</i> .....               | 45  |
| Figure 16 : Repliement du domaine Mid de la protéine Ago2 humaine.....  | 47  |
| Figure 17 : Comparaison des repliements adoptés par le domaine PIWI des protéines Argonaute et les enzymes de types RNase H .....                         | 49  |
| Figure 18 : The Argonaute silencing complex at work.. .....   | 53  |
| Figure 19 : Activité endonucléasique des protéines rRDE-1 et rSAGO-1. ....  | 72  |
| Figure 20 : Activité endonucléasique de la protéine rPRG-1.. .....  | 74  |
| Figure 21 : Activité endonucléasique des protéines rALG-1 et rALG-2.....  | 76  |
| Figure 22 : Activité endonucléasique et courbe de liaison à l'ARNsb P pour les protéines rALG-1 et rALG-2.....  | 77  |
| Figure 23 : Activité endonucléasique et courbe de liaison à l'ARNsb P pour les protéines rALG-1, rALG-1 3A, rALG-2 et rALG-2 3A .....                     | 79  |
| Figure 24 : Activité endonucléasique pour la protéine rALG-1 avec différents ARNsb.. ...  | 82  |
| Figure 25 : Duplexes utilisés dans les différentes expériences.....   | 84  |
| Figure 26 : Essai de clivage des duplexes d'ARN.....  | 85  |
| Figure 27 : Courbe de liaison à différents substrats d'ARN des protéines rALG-1 3A et rALG-2 3A.....  | 86  |
| Figure 28 : Essai de relargage d'ARNsb à partir de différents duplexes.....   | 87  |
| Figure 29 : Activité de clivage des duplexes par les mutants rALG-1/2 ΔPAZ.....   | 88  |
| Figure 30 : Courbes de liaisons à pré-let-7 et l'ARNsb P pour rALG-1 et rALG-2. pré-let-7 est représenté en trait foncé, et l'ARNsb P en trait clair..... | 89  |
| Figure 31 : Phénotypes observés les mutants <i>alg-1/2</i> .....  | 93  |
| Figure 32 : Analyse des miARN let-7 et lin-4 par Northern Blot.....   | 94  |
| Figure 33 : Analyse des miARN lin-4 et let-7 par Northern Blot.....   | 100 |
| Figure 34 : Analyse des miARN miR-90 et miR-40 par Northern Blot. ....  | 102 |
| Figure 35 : Analyse des miARN miR-50 et miR-62 par Northern Blot.....   | 104 |

|  |     |
|--|-----|
| Figure 36 : Détermination de la troncation de pré-miR-50 .....   | 106 |
| Figure 37 : Expérience de détection de lin-4 associé aux protéines ALG-1 et GFP-ALG-2<br>en condition d'ARNi.....  | 108 |
| Figure 38 : Expérience de détection de let-7 associé aux protéines ALG-1 et GFP-ALG-2<br>en condition d'ARNi.....  | 109 |
| Figure 39 : Expérience de détection de miARN associé aux protéines ALG-1 et GFP-ALG-<br>2 en condition d'ARNi..... | 110 |
| Figure 40: Expérience de détection de complexe miRISC ALG-1/let-7 et ALG-1/lin-4...                                | 111 |
| Figure 41: RT-PCR sur les souches N2, <i>alg-1</i> et <i>alg-2</i> , au stade adulte. ....                         | 113 |
| Figure 42 : RT-PCR sur les lignées transgéniques au stade adulte .....   | 114 |
| Figure 43: Protéines associées à ALG-1 et GFP-ALG-2 en condition d'ARNi.....                                       | 116 |

## Listes des Abréviations

A : Adénosine  
AUB : Aubergine  
ADN : Acide désoxyribonucléique  
AGO : Argonaute  
AIN-1/2 : *ALG-1/2 INteracting protein*  
ALG-1 : *Argonaute-Like Gene*  
AML-1 : Acute Myeloid Leukemia  
ARN : Acide ribonucléique  
ARNad : ARN adaptateur  
ARNadP : ARN adaptateur phosphorylé  
ARNdb : ARN double brin  
ARNi : ARN interférence  
ARNm : ARN messenger  
ARNnc : ARN non-codant  
ARNr : ARN ribosomal  
ARNn : *Small nuclear RNA*  
ARNno : *Small nucleolar RNA*  
ARNt : ARN de transfert  
BMP : *Bone Morphogenetic Protein*  
C : Cytosine  
CasiARN : *cis-acting siRNA*  
Cdk6 : *Cyclin-Dependant Kinase 6*  
CMH : Complexe majeur d'histocompatibilité  
CS : Chalcone synthase  
DCL : *Dicer-like protein*  
DGCR8 : *DiGeorge syndrome critical region gene 8*  
dsRBD : *Double-stranded RNA binding domain*  
DTC : *Distal tip cells*  
EGO-1 : *Enhancer of Glp-1*  
FGF : *Fibroblast growth factor*  
G : Guanine  
GCY : Guanylyl cyclase  
Glp-1 : *Glucagon-like peptide-1*  
h : heure  
HBL : *HunchBack Like-1*  
HDC : *Hypodermal blast cell*  
HIF : *Hypoxia-inducible factor*  
HMGA-2 : *High mobility group AT-hook 2*  
hnRNPK : *Heterogenous nuclear ribonucleoprotein K*  
HSC : *Hypodermal seam cell*  
IRES : *Internal ribosome entry site*  
Let : *Lethal*  
Lin : *Lineage*

LINE : *Long interspersed nuclear elements*  
 LLA : *Leucémie lymphocytaire aigüe*  
 LLC : *Leucémie lymphocytaire chronique*  
 LMC : *Leucémie myéloïde chronique*  
 LOQS : *Loquacious*  
 m7G : *Methylguanosine*  
 M-CSFR : *Macrophage colony-stimulating factor receptor*  
 miARN : *microARN*  
 MID : *Middle*  
 MMP-1G : *Matrix metalloproteinase 1G*  
 NatsiARN : *Natural antisense transcript-derived siRNA*  
 nt : *nucléotides*  
 OB : *oligonucleotides/oligosaccharides binding*  
 PABP1 : *Poly(A) binding protein 1*  
 PACT : *Protein activator et the interferon-induced protein kinase*  
 PAZ : *PIWI Argonaute Zwillie*  
 PDCD4 : *Program cell death protein 4*  
 piARN : *Piwi-interacting RNA*  
 PIWI : *P-element induced winpy testis*  
 poly (A) : *Polyadénylation*  
 pré-miARN : *précurseur de miRNA*  
 PRG : *PIWI related gene*  
 pri-miARN : *transcrit primaire de miRNA*  
 PTEN : *Phosphatase and tensin homolog*  
 PTGS : *Prostaglandin-endoperoxide synthase*  
 RAS : *Rat sarcoma*  
 RBCC : *Ring finger, B-Box and coiled-coil protein*  
 RDE-1 : *RNAi defective-1*  
 RDE-4 : *RNAi defective-4*  
 RNase : *Ribonucléase*  
 RRF-1 : *RNA dependent RNA polymerase-1*  
 shARN : *Short hairpin RNA*  
 siARN : *Small interfering RNA*  
 SID-1/2 : *Systemic RNAi defective*  
 TAp63 : *transformation related protein 63*  
 TasiARN : *trans-acting siRNA*  
 TCR : *récepteur des cellules T*  
 TGF : *Transforming growth factor*  
 TNRC6A : *Trinucleotide repeat-containing gene 6A protein*  
 TRBP : *TAR RNA binding protein*  
 TUTase : *Terminal uridylyl transferase*  
 U : *Uracyl*  
 UTR : *Untranslated region*  
 VEGF : *Vascular endothelial growth factor*

# Introduction

Les molécules d'acides ribonucléiques non-codants (ARNnc) sont de natures multiples et présentent une large variété de fonctions. La fonction des ARN non-codant dans la biologie est essentielle : lors de la traduction, pour les ARN de transferts (ARNt), lors de la synthèse protéique pour les ARN ribosomiques (ARNr), lors de l'épissage pour les petits ARN nucléaires (snARN) et lors de la modification de certains ARNr et ARNt pour les petits ARN nucléolaires (snoARN). Plus récemment, un autre domaine de grande envergure s'est ouvert avec la découverte de l'ARN interférence (ARNi). L'ARNi englobe les mécanismes de régulation de gènes par de courtes molécules d'ARN d'une vingtaine de nucléotides (nt). Ces courts ARN agissent en trans sur des ARN messagers (ARNm) par hybridation plus ou moins complémentaire et permettent l'inhibition de la traduction par blocage ou par dégradation de l'ARNm.

L'histoire de l'ARNi débute par pur hasard en 1990, chez les plantes. Le groupe de Jorgensen, chercheur dans le domaine de la biologie végétale, tente alors de manipuler l'expression du gène de la chalcone synthase (CS), une enzyme intervenant dans la synthèse du pigment conférant la couleur pourpre chez le pétunia (Napoli et al., 1990). Le but est alors d'introduire de nombreuses copies de l'enzyme afin d'obtenir des fleurs d'une couleur pourpre plus intense. À leur grande surprise, aucune des plantes transgéniques créées ne montre une intensité de coloration accrue. La majorité des plantes obtenues ont des pétales blancs tandis que d'autres sont blancs avec quelques taches pourpres. Une petite minorité est néanmoins pourpre, mais avec une coloration très pâle. L'analyse moléculaire révélera que chez les plantes avec des pétales blancs, il y a une baisse radicale du niveau d'ARNm de la CS. Un peu plus tard, on découvre que la surexpression de la CS provoquait la transcription d'un ARNm aberrant capable de s'apparier pour donner une forme double brin et que l'inhibition de la CS se faisait de façon posttranscriptionnelle (PTGS pour «post-transcriptionnal gene silencing»), le taux de transcription étant inchangé (Kooter et al., 1999).

Chez *Caenorhabditis elegans*, un nématode utilisé en génétique, Andrew Fire et Craig Mello découvrent un phénomène similaire de régulation de type PTGS. L'injection d'un

ARN double brin (ARNdb) exogène dans le corps du ver provoque la dégradation spécifique de l'ARNm du gène correspondant (Fire et al., 1998). Ce phénomène est nommé ARNi. Il est également découvert que l'injection n'est pas la seule façon de distribuer cet ARN db : incuber des vers dans une solution d'ARN db (Tabara et al., 1998) ou nourrir les vers avec des bactéries exprimant cet ARN db arrive au même résultat (Timmons and Fire, 1998). Par la suite, ce mécanisme conservé chez les modèles de cellules en cultures de mammifères et de drosophiles est utilisé pour inhiber le produit de certains gènes. Pour la première fois, on sort du cadre de la génétique classique où la recherche de mutants correspondait à l'étape initiale. Désormais, par cette génétique nommée génétique inverse, on est en mesure d'inhiber un gène d'intérêt par ARNi, ce qui offre un outil encore inédit. Avec la découverte des miARN et des autres courts ARN, un autre monde s'ouvre et on se rend alors compte que de nombreux gènes de mammifères et d'organismes supérieurs sont contrôlés par de courts ARN produits de façon endogène, changeant totalement la vision de la communauté scientifique quant à la façon de voir la régulation de l'expression de gène.

# Chapitre 1 : Revue de littérature

Les courts ARNnc sont des séquences d'ARN de 20 à 30 nt généralement monophosphorylés en 5' (Tolia and Joshua-Tor, 2007). Il y a trois grandes classes de courts ARN : les microARN (miARN), les «small interfering» ARN (siARN) et les «Piwi-interacting» ARN (piARN) (Kim et al., 2009). Comme nous le verrons plus loin, ces différents courts ARN (siARN, miARN et piARN) sont pris en charge par des protéines Argonaute particulières (Hutvagner and Simard, 2008). Bien que d'autres catégories de courts ARN soient connues dans la littérature (Tableau 1) et qu'ils exercent des fonctions importantes, ils ne seront pas décrits dans cette thèse. Les «cis-acting small interfering» ARN (casiARN), les «trans-acting small interfering» (tasiARN) et les «natural antisense transcript-derived small interfering» (natsiARN) sont spécifiques aux plantes et ne sont donc pas retrouvés chez d'autres organismes. Il apparait très nettement dans ce tableau que les plantes ont le plus large éventail de courts ARN. Leur spécificité est essentiellement reliée aux différentes protéines «Dicer Like Proteins» (DCL) qui les synthétisent, ainsi qu'à la source de l'ARNdb initiateur et à leurs fonctions. Les miARN sont l'espèce la plus conservée à travers l'évolution et ils sont retrouvés chez les plantes, les algues, les animaux ainsi que chez certains protistes (Ghildiyal and Zamore, 2009). Chez notre modèle d'étude, *Caenorhabditis elegans*, les trois grands groupes de courts ARN sont présents et les piARN n'ont été découvert que tardivement (Batista et al., 2008; Das et al., 2008).

**Tableau 1 : Description de différentes classes de courts ARN.** Ils sont retrouvés chez différents organismes. *Arabidopsis thaliana* (At), *Drosophila melanogaster* (Dm), *Schizosaccharomyces pombe* (Sp), *Caenorhabditis elegans* (Ce), *Thermus thermophilus* (Tt). Bien que non représenté dans le tableau, les piARN sont présents chez *C. elegans*. Tiré de Tolia et Joshua Tor, 2007.

| Class   | Description                             | Size (nt) | Observed in             |
|---------|---|-----------|-------------------------|
| siRNA   | Small interfering RNA                   | ~20-24    | Mammals, At, Dm, Sp, Ce |
| miRNA   | MicroRNA                                | ~20-24    | Mammals, At, Dm, Sp, Ce |
| tasiRNA | trans-acting siRNA                      | ~24       | At, Dm, Sp              |
| tncRNA  | Tiny noncoding RNA                      | ~20-21    | Ce                      |
| scnRNA  | Small scan RNA                          | ~28       | Tt                      |
| rasiRNA | Repeat-associated small interfering RNA | ~24-29    | At, Dm, Sp, Ce          |
| piRNA   | Piwi-interacting RNA                    | ~26-31    | Mammals                 |

## 1. Les miARN

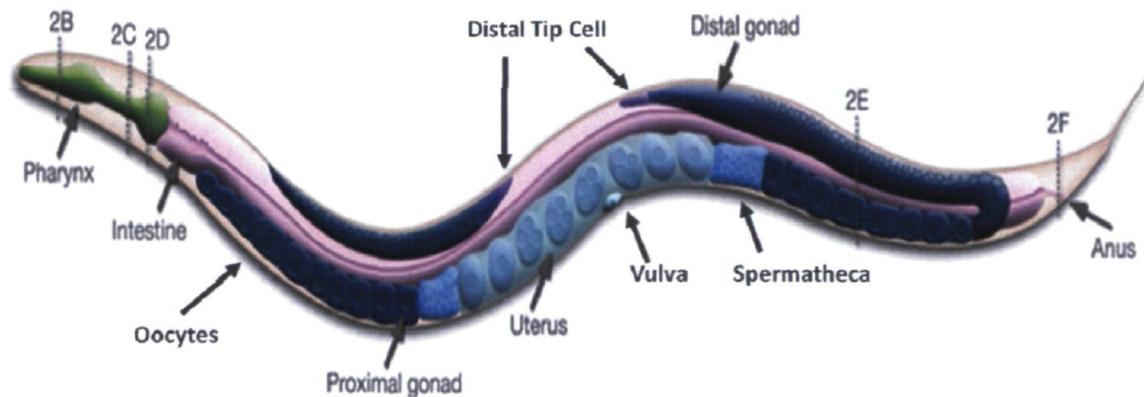
### 1.1. Le nématode *C. elegans*

#### 1.1.1. Généralités

Ce petit ver de 1 mm de long, transparent et très mobile a pour milieu de vie naturel le sol et est retrouvé dans toutes les régions du monde. Il se nourrit de bactéries et appartient au règne animal, embranchement des nématodes. Il a été sélectionné comme modèle « parfait » pour l'étude d'un organisme pluricellulaire par Sydney Brenner en 1965. Il possède deux genres, mâles (XO) et hermaphrodites (XX), mais les mâles ne représentent que 0,05 à 1 % des individus et sont hémizygote pour le chromosome sexuel. Le génome est composé de 5 paires de chromosomes et d'un chromosome sexuel. La reproduction de l'animal peut se faire par autofécondation ou par fécondation d'un hermaphrodite par un mâle. Le nématode possède un nombre invariable de cellules et, en 1983, le lignage

cellulaire du ver est alors entièrement élucidé : depuis le stade embryonnaire à une cellule jusqu'aux 959 noyaux somatiques qui composent un ver adulte hermaphrodite (Sulston et al., 1983). En 1999, le génome a été entièrement séquencé. Il devient un modèle de référence pour l'étude de l'apoptose, de la programmation cellulaire, du vieillissement et de la neurobiologie étant donné que 302 cellules sont des neurones. De plus, c'est un puissant outil génétique. L'hermaphrodite adulte a une morphologie très sommaire (Fig.1). La tête est composée d'une cavité buccale et d'un pharynx contenant une pompe lui servant à se nourrir. Un long tube lui servant d'intestin traverse le ver pour finir à l'anus. L'appareil reproducteur du vers est bidirectionnel : on voit de part et d'autre les deux gonades, distales et proximales, débutant chacune par la cellule «Distal Tip Cell» (DTC). Les oocytes sont produits dans les deux gonades et fécondés par le sperme contenu dans la spermatheca. Les œufs obtenus se retrouvent dans l'utérus et sortent par la vulve. Le nématode passe par plusieurs stades de développement (Fig.2), un stade embryonnaire, 4 stades larvaires (L1, L2, L3 et L4) et un stade adulte. Je me concentrerai uniquement sur la description du développement des hermaphrodites.

**Figure 1 : Description schématique des différents organes composant un nématode hermaphrodite au stade adulte.** Modifié de WORMATLAS.

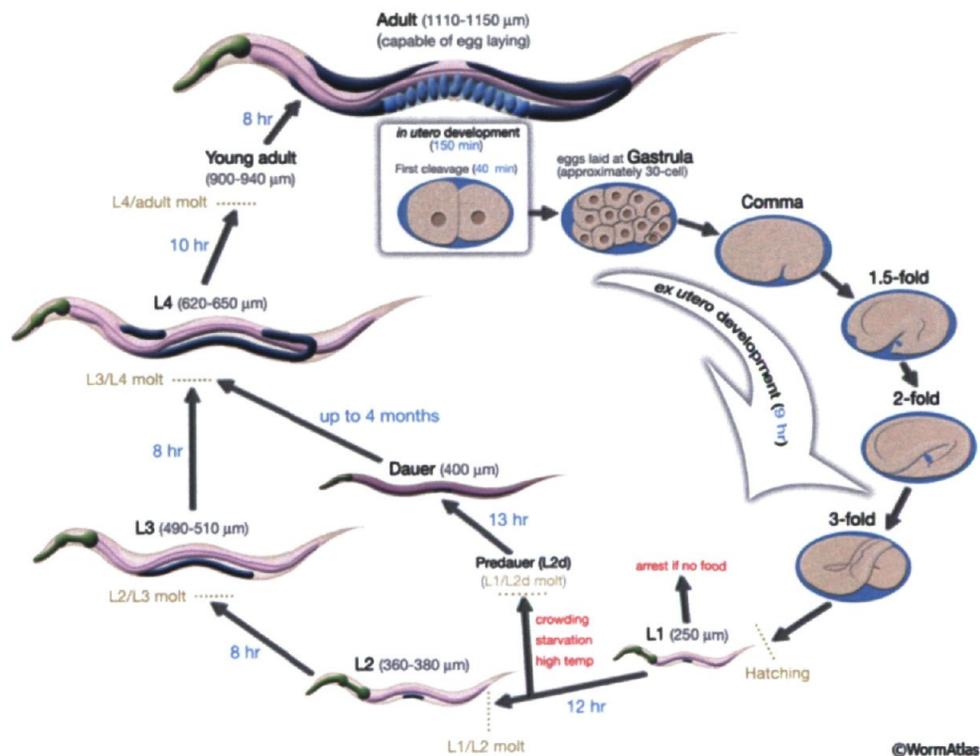


### 1.1.2. Développement embryonnaire

À 22 °C, après la fécondation, le développement embryonnaire se distingue par deux phases : la première est une phase de prolifération, tandis que la seconde est une phase appelée morphogénétique ou organogénétique (Von Ehrenstein, 1980). La phase de prolifération commence *in utero* et est suivie de nombreuses divisions depuis le stade une

cellule pour atteindre un total de 550 cellules non différenciées. Après seulement 150 min, au stade 30 cellules, la ponte a lieu, peu après le stade Gastrula (Fig.2; 26 cellules) (Altun, 2009). Vers la fin de la phase proliférative, l'embryon adopte une forme sphérique et 3 couches de tissus sont présentes : l'ectoderme, qui donnera l'hypoderme et les neurones, le mésoderme, précurseur du pharynx et des muscles; et l'endoderme, qui initiera l'intestin et les lignées germinales. La phase morphogénétique débute après environ 6 h, soit la fin de la gastrulation, au cours de laquelle on assistera à la formation de la structure larvaire et l'apparition d'une structure en haricot appelée «Comma». On peut déjà distinguer mâle et femelle après 8,5 h par le biais d'un développement neuronal différent (Altun, 2009). Par la suite, la différenciation se poursuit à travers les stades 1,5 « fold », 2 « fold » et 3 « fold » (Fig.2), au cours desquels la forme finale se dessine et le ver commence à se mouvoir (Altun, 2009). Après 14 h, il y a éclosion d'une larve composée d'exactly 558 cellules.

**Figure 2 : Représentation des différentes phases de développement du ver.** Tiré de WormAtlas.



### 1.1.3. Les différents stades larvaires

Le démarrage du développement larvaire se fait 3 h après éclosion, s'il y a présence de nourriture. La phase L1 débute 11,5 h après la ponte de l'œuf, où il y a génération de 5 des 8 neurones moteurs ainsi que le développement de certains muscles. On observe également le début de la synthèse des gonades somatiques durant la fin de ce stade. Au stade L2, 23,5 h après la ponte, on aura formation des 3 autres neurones. Les cellules germinales continuent à se diviser puis il y a réarrangement de celles-ci avec les cellules somatiques pour former la structure de la future gonade. Ce processus est contrôlé par des signaux envoyés par la cellule DTC, présente à chaque extrémité, et qui provoque l'élongation de la gonade (Kimble and Hirsh, 1979). À la fin du stade L2, en absence de nourriture ou sous l'influence de certains facteurs environnementaux tels que des phéromones, le ver peut entrer dans un stade de résistance que l'on appelle «Dauer» (Wood, 1988). Le métabolisme est alors fortement ralenti, il n'y a plus de croissance et la locomotion est réduite. Sous cette forme, le ver peut alors survivre plusieurs semaines et peut reprendre son développement quand l'environnement redevient favorable. Au stade L3, 31,5 h après la ponte, on voit apparaître la spermathèque (contenant les spermatozoïdes) et l'utérus, tandis que la migration des gonades continue le long du corps dans des directions opposées et les cellules précurseurs de la vulve sont présentes (Kimble and Hirsh, 1979). À 39,5 h, en L4, la gonadogénèse est terminée et la migration des gonades s'arrête proche de l'endroit où sera la vulve. L'esquisse de la future vulve est déjà visible, car sa morphogénèse est en cours, et ne se terminera qu'au stade adulte. La formation de la vulve se termine quand le ver est jeune adulte. La méiose des cellules germinales est alors très active et ne finira qu'à la transition L4/adulte pour la spermatogénèse. Le sperme sera stocké et utilisé pour l'autofécondation. En revanche, l'oocytogénèse se poursuivra jusqu'à la mort du ver. Près de 3 jours après la fécondation, on observe un ver fertile capable de pondre des œufs complétant ainsi son cycle de vie.

### 1.1.4. Les miARN et le développement du nématode *C. elegans*

#### 1.1.4.1. Le miARN *lin-4*

Au début des années 1990, l'équipe du Dr Victor Ambros travaille depuis plusieurs années sur des gènes hétérochroniques qui régulent la destinée et le lignage des cellules chez le nématode *C. elegans*. Ces gènes permettent le passage du nématode à travers les différents stades larvaires. Ils sont définis comme une cascade de gènes régulateurs qui vont déterminer le bon déroulement du développement dans le temps et un défaut dans l'expression de ces gènes altère les divisions cellulaires qui seront omises ou répétées (Ambros, 1989; Ambros and Horvitz, 1984). Le gène *lin-14* est alors connu pour coder une protéine nucléaire, qui s'avère être un facteur de transcription. Hautement exprimé dans l'embryon et au cours du premier stade larvaire, ce facteur permet le passage au stade suivant par sa diminution (Ruvkun and Giusto, 1989). Il a été également démontré que la régulation négative de ce gène se faisait via la région 3'UTR («UnTranslated Region») de l'ARNm. Le groupe s'intéresse alors au gène *lin-4*, nécessaire à la régulation négative de l'expression du gène *lin-14* (Arasu et al., 1991; Wightman et al., 1993) et, notamment, au mécanisme de l'interaction exacte entre ces deux gènes. Il décide tout d'abord de cloner le locus du gène *lin-4* et découvre que le produit de ce gène n'est pas une protéine, mais deux ARN de 22 nt et 61 nt dont une partie de la région 3'UTR de *lin-14* est complémentaire (Lee et al., 1993). Il arrive donc à l'élaboration d'un modèle où la régulation négative via les courts ARN *lin-4* se ferait par complémentarité de base sur la région 3'UTR de *lin-14* (Lee et al., 1993). La forme ARN active serait l'ARN de 22 nt qui découlerait d'un processus de maturation posttranscriptionnel de l'ARN de 61 nt. La découverte du premier miARN *lin-4* et de son précurseur est, à l'époque, vue comme une curiosité spécifique du nématode.

#### 1.1.4.2. Le miARN *let-7*

Quelques années plus tard, un autre miARN, *let-7*, est découvert. Il a tout d'abord été caractérisé phénotypiquement comme étant un gène hétérochronique. Dans un animal sauvage, les cellules «Hypodermal Blast Cell» (HBC) se divisent à chaque stade larvaire jusqu'au stade adulte, où ces cellules vont fusionner avec les cellules «Hypodermal Seam

Cell» (HSC) voisines pour former des alaes. Cette structure particulière située au niveau de la cuticule, est un marqueur du passage au stade adulte (Sulston and Horvitz, 1977). Dans un mutant de perte de fonction *let-7*, il y a une incapacité à former des alaes car les cellules HSC continuent de se diviser et donc réitèrent le patron observé au stade L4 (Reinhart et al., 2000). En revanche, la surexpression de ce gène va provoquer la formation des alaes dans un stade plus précoce, au stade L4. L'expression adéquate de ce gène va donc permettre le démarrage des programmes de divisions menant au stade adulte. De façon similaire à *lin-4* régulant *lin-14* de par sa région 3'UTR, *let-7*, dont l'expression commence en L3, possède une complémentarité partielle pour la région 3'UTR d'ARNm de nombreux gènes hétérochroniques, tels que *lin-41*, *lin-28* et *daf-12* (Arasu et al., 1991; Reinhart et al., 2000; Slack et al., 2000). Le produit de ces gènes intervient à différentes étapes. Par exemple, la protéine LIN-41, protéine à domaine RBCC («Ring finger-B box-Coiled coil»), est une protéine inhibitrice du facteur de transcription LIN-29, essentielle pour le déclenchement de la transcription de plusieurs gènes nécessaires pour la transition du stade L4 à adulte (Slack et al., 2000). La traduction de la protéine LIN-41 est négativement régulée par *let-7* durant cette transition (Reinhart et al., 2000; Vella et al., 2004). DAF-12 fait partie d'une famille de protéines réceptrices d'hormones stéroïdiennes importantes pour le développement dont la traduction est régulée par *let-7* (Hammell et al., 2009a). LIN-28, une protéine pouvant lier l'ARN, est régulée à la fois par *lin-4* et *let-7* (Reinhart et al., 2000; Seggerson et al., 2002).

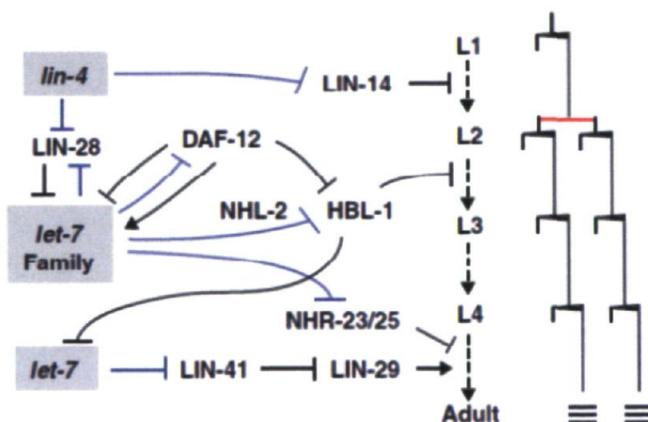
À la différence du premier miARN, il se trouve que *let-7* est hautement conservé chez plusieurs espèces, dont l'humain et la drosophile, mais également que l'expression de ce miARN est temporelle, tout comme chez le nématode. De plus, la principale cible de *let-7*, *lin-41*, possède également des homologues chez ces différentes espèces (Pasquinelli et al., 2000).

#### 1.1.4.3. Division des cellules HSC

Aujourd'hui, on connaît beaucoup plus précisément la finesse de la régulation de la division des cellules HSC chez *C. elegans*. Comme le montre la Figure 3, cette division est orchestrée et temporellement régulée par *lin-4*, *let-7* et d'autres membres de la famille *let-7*. Une cascade d'évènements a pour point de départ *lin-4*, qui va permettre le passage d'un

mode de division asymétrique des cellules HSC (L1) à un mode de division symétrique (L2) (Ambros, 2011). Cette régulation est assurée par la protéine HBL-1 entre les stades L2 et L3 (Abbott et al., 2005). HBL-1 est un facteur de transcription à doigts de zinc impliqué dans une boucle de rétrocontrôle avec *let-7* : il régule positivement l'expression de *let-7*, tandis que *let-7* réprime la traduction en protéine sur l'ARNm *hbl-1* (Abbott et al., 2005; Grosshans et al., 2005; Lin et al., 2003). HBL-1 est sous l'influence de la famille *let-7* via la protéine NHL-2. NHL-2 est un modulateur protéique qui augmente la répression posttranscriptionnelle de certaines cibles de miARN (Hammell et al., 2009b). L'ensemble de ce processus de régulation est, de manière continue, sous l'influence de l'expression de *lin-4*, qui inhibe LIN-28, lui-même régulateur négatif des membres de la famille *let-7*. Pour le passage L4-adulte, l'arrêt de division est activé par LIN-29 et, pour ce faire, une inhibition de *let-7* par HBL-1 est nécessaire. Comme décrit précédemment, l'arrêt de la division des cellules va provoquer la fusion des cellules HBC et HSC pour former les alae cuticulaires.

**Figure 3 : Cascade de régulation hétérochronique contrôlant la division des cellules HSC chez *C. elegans*.** Les voies de régulation (à gauche) et les lignages (à droite) sont schématisés avec la différenciation terminale (triple barre) au stade adulte. Les différents stades larvaires L1, L2, L3, L4 et adultes sont précisés. Tiré de Ambros, 2011.



#### 1.1.4.4. Développement de la vulve

Le développement de la vulve est initié par la division de 6 cellules précurseurs pluripotentes et est contrôlé par de nombreuses voies de signalisation dont la voie LIN-12/Notch qui assure spécifiquement l'entrée en division spécifique de certaines de ces cellules (Sternberg, 2005). Il a été démontré que le miARN miR-61 avait pour cible

potentielle l'ARNm de Vav, un inhibiteur de la voie LIN-12/Notch (Yoo and Greenwald, 2005). Par inhibition de la traduction de VAV, miR-61 va permettre l'activation du récepteur LIN-12/Notch et provoquer la translocation du domaine intracellulaire ; celui-ci va recruter des cofacteurs qui vont positivement activer la transcription au locus de miR-61, maintenant ainsi le niveau de la protéine inhibitrice, et permettant le maintien de la signalisation de LIN-12/Notch. Il en découlera l'entrée en division de certaines cellules de façon spécifique et miR-61 ferait donc partie d'une boucle de rétrocontrôle de la voie de signalisation LIN-12/Notch.

#### **1.1.4.5. Développement neuronal**

Au niveau des neurones de la tête, il a été découvert que, chez l'animal adulte, il y avait une asymétrie des neurones sensitifs : à gauche, on a le type ASEL («ASE left»), et à droite ASER («ASE right») (Hobert et al., 2002). Ils se différencient par l'expression de gènes codant pour des chimiorécepteurs à activité guanylyl cyclase (gcy). Les gènes gcy-6 et gcy-7 sont exprimés seulement dans le neurone ASEL, tandis que gcy-5 est seulement exprimé chez ASER. Plusieurs articles ont démontré que deux miARN, lsy-6 et miR-273, étaient impliqués dans des boucles de régulation des gènes contrôlant l'expression de ces chimiorécepteurs (Hobert, 2006; Hobert et al., 2002). Ainsi, le patron d'expression distinct gcy-6/7 et gcy-5 était principalement relié à une expression différentielle dans ASEL et ASER de ces deux miARN; lsy-6 réprimant les gènes ASER-spécifique dans le neurone ASEL et miR-273 réprimant les gènes ASEL-spécifique dans ASER (Chang et al., 2004).

Plus récemment, on a rapporté que l'expression du miARN miR-124 était majoritairement située au niveau du système nerveux sensor du ver (Clark et al., 2010). Une analyse bio-informatique a révélé que de nombreux gènes étaient différemment exprimés dans le système nerveux comparativement au reste de l'animal, pouvant expliquer un rôle de ce miARN dans la spécialisation tissulaire.

#### **1.1.4.6. Développement embryonnaire**

Une famille de miARN, dont l'expression est spécifique à l'embryon, a été récemment caractérisée. La famille miR-35, composée des miARN miR-35 à miR-42, ne montre

individuellement aucun phénotype de létalité (Miska et al., 2007). Ces miARN partagent une grande homologie de séquence et lorsque de multiples mutants sont créés, il s'avère qu'au stage de développement embryonnaire 2-fold, dans de nombreux mutants, l'élongation du pharynx n'a pas lieu. Le pharynx ne s'attache pas à l'hypoderme pour former le lien avec le tractus intestinal, bien que la formation des organes soit complète (Alvarez-Saavedra and Horvitz, 2010). Il s'ensuit une mortalité embryonnaire importante.

## **1.2. Fonction des miARN dans le développement des mammifères**

Comme décrit précédemment, les miARN interviennent dans pratiquement chaque étape de la vie du ver. En ce qui concerne les mammifères et notamment l'humain, de nombreuses études ont démontré l'implication directe des miARN dans le développement de certains tissus et organes. Je ferai donc ici une description de quelques mécanismes connus de contrôle de certains processus développementaux des miARN.

### **1.2.1. Régulation du système nerveux**

Des études récentes ont montré que les miARN sont nécessaires et jouent un rôle important dans la maturation et la survie neuronale. Ainsi, on a démontré l'implication (directe ou indirecte) des miARN dans la biologie neuronale. Par exemple, le miARN miR-134 pourrait contribuer à la croissance et au maintien de la plasticité synaptique (Sempere et al., 2004). miR-133b, un miARN spécifiquement exprimé dans les neurones dopaminergiques du mésencéphale, régule la maturation et la fonction de ce type de neurones dans une boucle de rétroaction négative impliquant le facteur de transcription Pitx3. Ce miARN est déficient chez les patients atteints de la maladie de Parkinson (Smirnova et al., 2005).

Un autre miARN spécifique des neurones, miR-124, vise directement l'ARNm du gène «Polypyrimidine tract-binding protein 1» (PTBP1), qui code pour un répresseur global de l'épissage alternatif, très actif chez les cellules non-neuronales. Pendant la différenciation neuronale, l'expression de miR-124 va réduire considérablement les niveaux de PTBP1 et cela va activer l'épissage alternatif neuronal. Cet épissage neurone-spécifique va engendrer la synthèse des facteurs importants dans le développement et la différenciation neuronale (Makeyev et al., 2007).

Chez l'humain et la drosophile, la protéine ATROPHINE-1, un répresseur transcriptionnel, est impliquée dans le contrôle de l'apoptose neuronale (Ellerby et al., 1999). Ce gène est une cible de miR-8, et les mutants pour ce miARN ont des phénotypes particuliers qui sont attribuables à un niveau d'ATROPHINE-1 élevé, entraînant chez la drosophile l'apoptose de certaines cellules nerveuses (Karres et al., 2007). Des mutations au niveau de ce gène sont connues chez l'homme : l'insertion de triplets CAG dans une région riche en glutamine a pour conséquence une augmentation de la durée de vie de la protéine, causant une pathologie neurodégénérative connue sous le nom de Syndrome de Naito-Oyanagi (Nagafuchi et al., 1994).

### **1.2.2. Développement du système cardio-vasculaire**

L'implication du miARN miR-126 a été mise en évidence dans les processus d'angiogénèse. Chez la souris, lorsque ce miARN est délété, plusieurs défauts apparaissent comme des vaisseaux défectueux sujets à l'hémorragie, et une perte générale de l'intégrité vasculaire suite à une prolifération des cellules endothéliales et une migration cellulaire anormales (Wang et al., 2008a). Les animaux survivants affichent une néovascularisation cardiaque défectueuse et sont fréquemment atteints d'infarctus du myocarde. L'expression de miR-126 est restreinte aux cellules endothéliales et il semblerait que sa mutation soit similaire à une diminution des deux principaux facteurs favorisant l'angiogénèse, le «Vascular Endothelial Growth Factor» (VEGF) et le «Fibroblast Growth Factor» (FGF). Ce miARN soutiendrait l'action pro-angiogénique du VEGF et du FGF et ce, en favorisant la formation de vaisseaux sanguins par répression de l'expression de SPRED-1, un inhibiteur intracellulaire de signalisation pro-angiogénique (Fish et al., 2008).

Toujours chez la souris, deux miARN, miR-143 et miR-145, sont cotranscrits au niveau des cellules progénitrices pluripotentes cardiaques avant de se localiser au niveau des cellules musculaires lisses du cœur (Cordes et al., 2009). Ces deux miARN sont induits directement par des facteurs de réponses spécifiques sériques telles que la myocardine et Nkx2-5. En combinaisons avec divers facteurs, ces deux miARN soutiennent la différenciation en cellules musculaires lisses comme réponse à des facteurs sérique. miR-145 s'est avéré suffisant pour la reprogrammation de fibroblastes en cellules musculaires lisses sous la

dépendance de la myocardine, et pour induire la différenciation de cellules pluripotentes en cellules musculaires lisses vasculaires (Cordes et al., 2009).

En ce qui concerne le contrôle entre la prolifération et la différenciation des cellules du coeur, plusieurs miARN jouent un rôle important. Une surexpression de miR-1 dans des cellules myoblastiques favoriserait la différenciation (Zhao et al., 2005). Un autre exemple, miR-21, cible l'ARNm de «Programmed Cell Death 4» (PDCD4) dans les myocytes cardiaques, une protéine qui est impliquée dans l'apoptose (Cheng et al., 2009). Ce même miARN est retrouvé augmenté dans les fibroblastes d'un cœur défaillant et sa fonction serait de réguler la voie des kinases MAP/ERK via l'inhibition de SPROUTY 1, un antagoniste de la signalisation VEGF/FGF (Thum et al., 2008). Dans le même registre, miR-133, présent sous 2 isoformes, est contrôlé durant le développement cardiaque par un facteur de transcription spécifique et les souris déficientes dans les deux isoformes affichent un phénotype létal dans la moitié des cas, correspondant à une malformation du septum ventriculaire (Liu et al., 2008). De plus, par utilisation d'antagomir (une séquence complémentaire au miARN qui va bloquer la fonction de celui-ci), il est possible de provoquer une hypertrophie cardiaque très marquée, preuve encore de l'influence de ce miARN dans le développement cardiaque (Care et al., 2007).

### **1.2.3. Développement des organes**

Le développement de certains organes est finement contrôlé par les miARN mais est très souvent spécifique à quelques types cellulaires. Ainsi, dans le pancréas, miR-375 est le miARN le plus abondant chez l'homme et il est retrouvé spécifiquement dans certains îlots cellulaires nommés  $\alpha$  et  $\beta$  (Poy et al., 2004). Il aurait un rôle primordial dans le contrôle de la différenciation cellulaire entre  $\alpha$  et  $\beta$ . Chez la souris, la délétion de ce miARN causerait une nette diminution de cellules  $\beta$  et une augmentation de cellules  $\alpha$ , suggérant ainsi que l'expression de ce miARN au niveau des cellules précurseurs pourrait influencer le devenir de celle-ci (Poy et al., 2009). Au niveau des reins, les néphrons sont régulés par une expression fine de miR-30 et une réduction de ce miARN chez *Xenopus* entraîne une malformation rénale avec des structures néphrotiques plus petite (Agrawal et al., 2009).

Plusieurs miARN régulent le développement pulmonaire et chez le rat, lors d'une étude de profilage des miARN dans des fœtus à différents stades, il a été observé que le niveau de miR-127 progresse au fur et à mesure en quantité, et que son expression se spécialise pour certains types cellulaires (Bhaskaran et al., 2009). Par hybridation *in situ*, il a été démontré qu'il est tout d'abord exprimé dans le tissu mésenchymal puis plus tard dans l'épithélium. Ce changement commence et s'opère dans les derniers jours de l'embryon et le premier jour après la naissance. Une surexpression de ce miARN plus précoce entraîne la formation de structure défectueuse au niveau des poumons. Le groupe miR-17~92, quant à lui, affiche un profil d'expression opposé et est très abondant dans les premiers stades du développement pulmonaire (Lu et al., 2007). Par surexpression de ces miARN au niveau de l'épithélium pulmonaire, il a été possible de provoquer une hyperplasie épithéliale au niveau des cavités alvéolaires. De plus, sa délétion complète est létale et parmi les nombreux phénotypes, on observe une hypoplasie du poumon sans défaut de structure particulière (Ventura et al., 2008).

#### **1.2.4. Système immunitaire et miARN**

Les premiers miARN découverts comme pouvant être impliqués dans la régulation du système immunitaire sont miR-15 et miR-16. Des études ont déterminé que dans 70 % des cas de leucémies lymphocytaires chroniques (LLC), ces miARN étaient absents ou sous-exprimés (Calin et al., 2002), suggérant donc un lien causal entre cette dérégulation et la maladie. Plusieurs exemples sont maintenant connus dans la biologie du système immunitaire où les miARN sont des acteurs importants. Chez la souris, le miARN miR-223 contrôle l'activation des granulocytes (Johnnidis et al., 2008). La perte de ce miARN entraîne une accumulation de cellules progénitrices de granulocytes et le processus de maturation en est affecté. Les granulocytes formés sont hypermatures, très sensibles aux stimuli du système immunitaire et ont une activité fongicide accrue. Chez les lymphocytes, le miARN miR-181 est également un élément de maturation essentiel. Il contrôle le passage de lymphocytes immatures en lymphocytes matures activés, ayant reconnu les molécules du Complexe Majeur d'Histocompatibilité (CMH). miR-181 est un élément activateur de l'expression des récepteurs à cellules T (TCR) permettant de reconnaître les molécules du CMH et la surexpression ce miARN engendre une hypersensibilité de lymphocytes

immatures pour la reconnaissance du CMH (Li et al., 2007). D'autres cellules immunitaires comme les monocytes sont aussi sous l'influence de miARN pendant leur maturation. Ainsi, miR-17~92 et ses paralogues miR-105~25 inhibent le facteur de transcription «Acute Myeloid Leukemia 1» (AML1) et sa principale cible transcriptionnelle, le récepteur «Monocyte-Colony Stimulating Factor Receptor» (M-CSFR) (Fontana et al., 2007).

### **1.3. miARN dans les cancers**

Nous avons vu la participation des miARN dans le développement animal, et pour les différents exemples précédemment décrits, il est aisé de comprendre qu'une dérégulation va souvent être associée à différentes pathologies, notamment des cancers. Généralement, un cancer est dû à un déséquilibre entre facteurs oncogéniques et facteurs suppresseurs de tumeurs. Or, l'implication des miARN dans le développement animal démontre qu'ils peuvent être l'un ou l'autre.

#### **1.3.1. miARN suppresseur de tumeurs**

Les miARN suppresseurs de tumeur agissent principalement sur des gènes pouvant prendre un caractère oncogénique lorsqu'ils sont dérégulés. Au niveau du cancer du poumon, de nombreuses études ont démontré l'implication de let-7 dans la pathologie et, de par sa sous-expression, il est signe de mauvais pronostic et d'agressivité tumorale (Takamizawa et al., 2004). De plus, l'oncogène RAS est la cible de let-7 et dans certaines cellules cancéreuses, la diminution du miARN coïncide avec une augmentation de l'expression de la protéine RAS (Johnson et al., 2005). Ce miARN joue le rôle d'un suppresseur de tumeur puisque sa surexpression dans des cellules d'adénocarcinomes pulmonaires influe négativement sur la prolifération, la croissance et la survie de ces cellules (Sakurai et al., 2011). Chez l'homme, il existe une dizaine d'homologues de let-7 et tous sont fréquemment sujets à des mutations ou des délétions dans différents types de cancers (Bussing et al., 2008). Ce miARN possède de nombreuses autres cibles, dont des facteurs intervenants dans le cycle cellulaire. En inhibant les proto-oncogènes «Cycline Dependent Kinase 6» (Cdk6) et «Cell Division Cycle 25a» (Cdc25a), let-7 participe au contrôle du cycle cellulaire en ayant un effet négatif sur la prolifération cellulaire (Johnson et al., 2007). Toujours dans le poumon, l'oncogène «High-Mobility Group AT-hook 2» (HMGA2), une protéine contrôlant la

transcription lors de la prolifération et de la différenciation cellulaire, est fortement surexprimé dans les cancers et est une cible de let-7 (Lee and Dutta, 2007).

Dans le glioblastome, on peut prendre l'exemple de miR-34, miR-124, miR-137, miR-15 et miR-128 tous très fortement sous exprimés (Sayed and Abdellatif, 2011). Cette sous-expression entraîne l'augmentation de la cycline Cdk6 (cible des 3 premiers), de la cycline E1 et du facteur de transcription E2F3a. Chez les lymphomes à cellules B, l'expression de l'oncogène c-MYC provoque une aggrésivité du lymphome chez la souris ainsi qu'une réduction de l'apoptose (Zhao et al., 2008). c-MYC réprime de nombreux miARN dont let-7, celui-ci étant notamment impliqué dans une boucle de rétroaction puisqu'il a pour cible c-Myc (O'Donnell et al., 2005). De plus, c-MYC active la transcription du cluster miR-17~92, considéré comme un oncomir puisque ce cluster aurait des effets antiapoptotiques en favorisant la prolifération cellulaire (Dews et al., 2006).

### 1.3.2. OncomiRs

Les miARN potentiellement oncongénique sont appelés oncomiRs. Le cluster miR-17~92 est un locus polycistronique codant pour plusieurs miARN dont miR-17-5, miR-18a, miR-19a, miR-20a, miR-19b, et miR-92. Ce cluster est essentiel car sa délétion chez la souris entraîne une mort prématurée avec notamment une hypoplasie pulmonaire et des anomalies cardiaques (Ventura et al., 2008). Il est situé sur le chromosome 13 dans une région très sensible et fortement amplifiée dans différents types de cancer du poumon, de la prostate et du côlon (Sayed and Abdellatif, 2011). L'action de ce cluster se situe au niveau de la transition G1-S où il régule le facteur de transcription E2F1, favorisant la transition de la phase G1 à la phase S (O'Donnell et al., 2005). De plus, comme précédemment énoncé, l'oncogène c-MYC, autre acteur de la transition G1-S, induit l'expression de ce cluster (Dews et al., 2006).

Le miARN miR-21 est surexprimé dans plusieurs cancers de type très variés aussi bien des cas de lymphomes que des cancers ayant pour origine le sein, l'ovaire, le cerveau, le foie, les poumons, le pancréas ou encore le cou (Sayed and Abdellatif, 2011). Dans le glioblastome, la tumeur cérébrale la plus fréquemment rencontrée chez l'adulte, ce miARN (et d'autres comme miR-26a et miR-221/222) est très hautement exprimé (Silber et al.,

2009). On les considère comme faisant parties des oncomiRs, car parmi leurs cibles fréquentes, on retrouve de nombreux gènes qui régule négativement la prolifération, la survie et l'invasion tumorale. Dans le cancer du pancréas, miR-21 a été montré comme étant augmenté dans 80 % des cas et il ciblerait la protéine «Phosphatase and TENSin homologue» (PTEN), un inhibiteur de la division cellulaire (Sayed and Abdellatif, 2011). Son augmentation provoquerait la levée de l'inhibition cellulaire maintenue par PTEN et aurait ainsi un rôle de promotion de la division cellulaire.

Dans les lymphomes à cellules B, miR-155 est un court ARN situé proche du site d'intégration du virus de la leucose aviaire, responsable de la transformation cellulaire (Eis et al., 2005). Cette intégration engendre une augmentation de la transcription de ce miARN. Par expression ectopique de miR-155 dans des cellules B chez une souris transgénique, il a été démontré qu'il y avait prolifération clonale de précurseurs à cellules B (Costinean et al., 2006). Les symptômes entraînés sont une splénomégalie et une lymphopénie, signes avant-coureurs d'un lymphome.

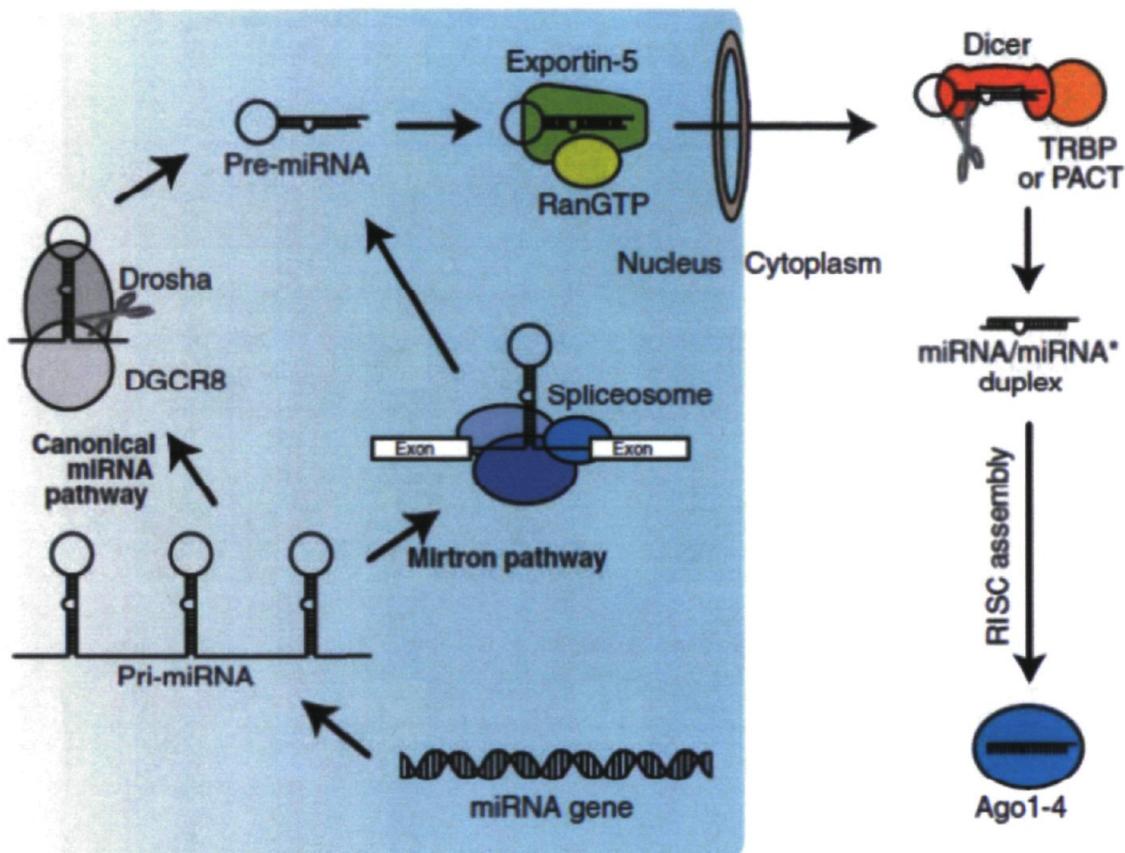
### **1.3.3. Les miARN et la spécificité cellulaire dans les cancers**

L'action des miARN dans le cancer se trouve être très spécifique et dépendante du type cellulaire en question. Ainsi, certains miARN, en fonction du cancer dans lequel ils sont dérégulés, peuvent agir comme oncogènes ou suppresseur de tumeur. De façon étonnante, le cluster miR-17~92 agit comme suppresseur de tumeur dans les cancers du sein, de l'ovaire et dans certains mélanomes où son expression est très réduite (Zhang et al., 2006). D'autre part, cette spécificité d'expression permet de discriminer précisément entre différents types cancéreux. Ainsi, dans toute une série de leucémies lymphocytaires, de nombreux miARN ont été découverts comme étant surexprimés ou réprimés. Outre la LLC, la leucémie lymphocytaire aiguë (LLA) ou encore les leucémies myéloïdes chroniques (LMC) ou aiguës (LMA), on observe des profils de miARN distincts (Sayed and Abdellatif, 2011). Ainsi, on est capable de discriminer la LLA de la LMA à 97% en comparant l'expression des miARN miR-128, let-7, et miR-223. miR-128 est hautement exprimé dans la LLA alors que dans la LMA, ce sont let-7 et miR-233 qui prédominent.

## 1.4. La biogénèse des miARN

Généralement, les miARN sont encodés dans le génome, transcrits par l'ARN polymérase II et maturés par deux enzymes de type RNases III (RiboNucléases III), Drosha, au niveau nucléaire, et Dicer, au niveau cytoplasmique. Drosha va prendre en charge le transcrit primaire de miARN (pri-miARN) pour le transformer en précurseur de miARN (pré-miARN). Dicer, par le recrutement de la protéine Argonaute appropriée et de cofacteurs, formera le complexe RLC (R pour RISC «RNA-Induced Silencing Complex» et LC pour «Loading Complex») qui va prendre en charge le pré-miARN pour former le complexe miRISC (Fig.4). Les miRISC cibleront l'extrémité 3'UTR des ARNm par complémentarité imparfaite et mèneront à une inhibition de la traduction par blocage de l'initiation et/ou dégradation de la chaîne peptidique naissante et/ou déstabilisation de la queue polyadénylée.

**Figure 4 : Biogenèse des miARN chez les mammifères.** Après transcription, les transcrits primaires de miARN (pri-miARN) sont pris en charge par le complexe Drosha/DGCR8 qui les converti en précurseurs de miARN (pré-miARN). Ces précurseurs peuvent aussi provenir directement du spliceosome dans le cas de mirtrons. Le complexe Exportin-5/RanGTP exporte le pré-miARN du noyau vers le cytoplasme, où Dicer interagit avec le complexe TRBP/PACT/Ago. Le modèle établi préconise que le pré-miARN soit clivé en duplex miARN/miARN\*. Ces duplexes sont chargés sur les protéines Ago pour former le complexe RISC mature. La paire de ciseaux schématise les coupures endonucléolytique. Tiré de Kwak *et al.*, 2010.

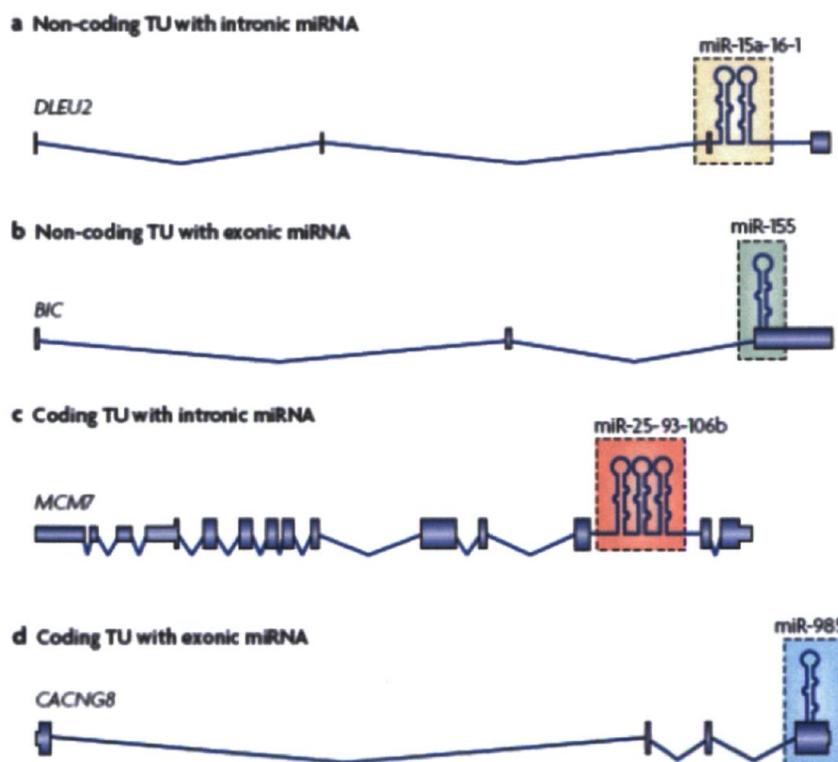


#### 1.4.1. Organisation génomique et transcription

En ce qui concerne la localisation génomique des miARN, elle est diverse et variée mais, généralement, les miARN sont très redondants au niveau génomique. Il existe de nombreux isoformes de miARN dans un génome et, de plus, pour un même miARN l'expression peut provenir de différents loci (Lee and Dutta, 2009). Chez les animaux, les miARN sont hautement conservés, et environ la moitié sont partagés entre *C. elegans* et l'humain (Ibanez-Ventoso et al., 2008). Chez ce dernier, la moitié des gènes codant pour des miARN sont organisés sous forme polycistronique (Lee et al., 2002). De ces miARN regroupés,

environ 50% contiennent plusieurs miARN de la même famille ayant pour cible les mêmes ARNm (Yu et al., 2006). Cette organisation permet une régulation fine et efficace, les miARN ciblant des ARNm identiques seront transcrits simultanément, au besoin. D'autre part, pas moins de 50% des miARN sont présents dans des transcrits non-codants, la plupart dans des introns (Rodriguez et al., 2004) (Fig.5). Une petite fraction seulement est présente dans des séquences génomiques répétées, telles que les séquences «Long Interspersed Nuclear Elements LINE» (LINE). En ce qui concerne les miARN provenant des transcrits codants, environ 40 % sont issues des introns (Kim et al., 2009).

**Figure 5 : Organisation de transcrits de miARN.** Représentation des différentes unités de transcription primaire (TU) contenant les précurseurs de miR-15a/16-1, miR-155, miR-25/93/106b et miR-985. Les précurseurs peuvent être isolés (b, d) ou regroupés (a,c). Ils peuvent être contenus dans un transcrit non-codant (a), dans un exon (b, d), dans un intron d'un transcrit codant (c). Tiré de Kim et al., 2009.



La transcription des miARN est généralement soumise à l'action de l'ARN polymérase II (Lee et al., 2004). Comme il s'agit de la même enzyme que pour les ARNm, les transcrits de miARN arborent les mêmes caractéristiques c'est-à-dire qu'ils sont coiffés et polyadénylés (Cai et al., 2004). Plusieurs facteurs de transcriptions, connus pour s'associer

à la polymérase II, sont impliqués dans le contrôle transcriptionnel des miARN. Ces facteurs ont souvent été découverts lors de dérégulation associée aux cancers, et pour ne citer que quelques exemples, on a vu précédemment les protéines MYC et E2F, mais on peut ajouter d'autres éléments régulateurs importants comme p53, «Signal transducer and activator of transcription 3» (STAT3) et «Hypoxia Inducible Factor» (HIF) (Lee and Dutta, 2009). Un faible nombre de miARN est transcrit par l'ARN polymérase III, par le biais de séquences promotrices riches en séquences appelées Alu (Borchert et al., 2006). Ces répétitions Alu induisent spécifiquement le recrutement de la polymérase III et environ 50 miARN, représentés par 19 groupes de transcrits, sont sujets à ce type de transcription.

Le transcrit obtenu est appelé transcrit primaire de miARN (pri-miARN). Il est composé d'une structure en tige boucle et de séquences simple brin de part et d'autre de la base de la tige. La plupart des pri-miARN contiennent plus d'une tige boucle pouvant donner naissance à un miARN (Fig.5).

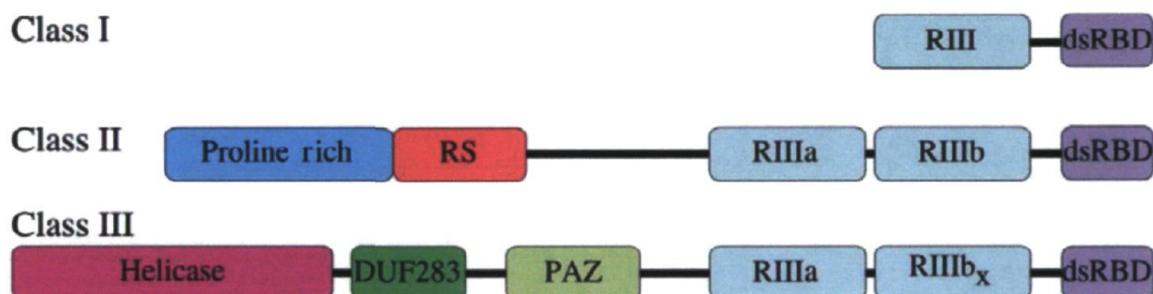
## **1.4.2. Maturation des miARN**

### **1.4.2.1. Description des RNases de type III**

Les formes précurseurs de miARN (pri-miARN et pré-miARN) vont être pris en charge par deux enzymes de la famille des RNases III, qui sont des endoribonucléases capables de couper l'ARN double brin. Elles sont composées au minimum d'un domaine RNase III et d'un domaine «Double-stranded RNA Binding domain dsRBD» (dsRBD) (Carmell and Hannon, 2004). Le dsRBD est un domaine de liaison à l'ARN double brin et le domaine RNase III contient le site catalytique. Elles sont classées en 3 groupes, selon les domaines fonctionnels (Fig.6). Drosha dans le noyau, puis Dicer dans le cytoplasme vont successivement couper la molécule d'ARN. Drosha et Dicer possèdent deux domaines RNases III qui vont permettre d'effectuer un clivage sur chacun des brins de la tige boucle. L'orientation tridimensionnelle des deux sites actifs, combinée à certains cofacteurs protéiques, va contribuer à créer un site de coupure arborant une extrémité 3'OH sortante lors de la coupure (Gan et al., 2006). Drosha possède un domaine riche en proline impliquée dans les interactions protéines-protéines et une région riche en sérine-arginine (SR) qui est fréquemment retrouvée chez les protéines impliquées dans le métabolisme de

l'ARN dont font parties les protéines SR impliquées dans l'épissage (Carmell and Hannon, 2004). Dicer se distingue de Drosha par un domaine PAZ (PIWI-Argonaute-Zwille) et un domaine hélicase en N-terminal. Le domaine PAZ est un domaine de liaison à l'ARN possédant une affinité pour les extrémités 3' hydroxylée (3'OH) (Lingel et al., 2004; Yan et al., 2003) et va reconnaître le site de coupure créé par Drosha sur la tige boucle. Le domaine PAZ est aussi retrouvé chez les protéines Argonaute et fera l'objet d'une description détaillée.

**Figure 6 : Domaines représentés chez les 3 classes d'enzymes RNases III.** RIII, domaine endonucléasique; dsRBD, domaine de liaison à l'ARN double brin; PAZ, domaine de liaison à l'extrémité 3'OH de l'ARN; RS, région riche Ser/Arg. La classe I comprend la RNase III d'*E. coli*, la classe II Drosha et la classe III Dicer. Tiré de Carmell and Hannon, 2004.



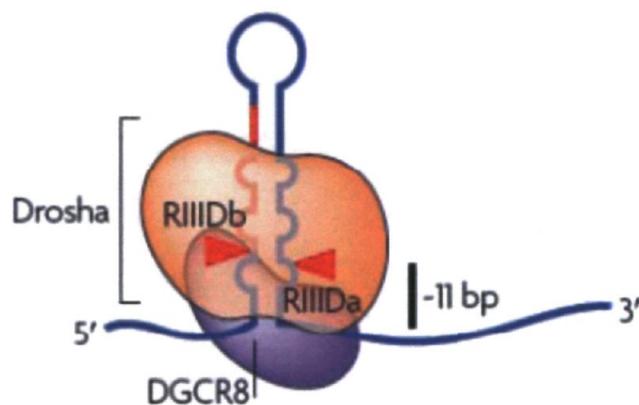
#### 1.4.2.2. Maturation nucléaire par Drosha

Le transcrit pri-miARN va être pris en charge et mûré par le complexe appelé microprocesseur. Il est composé principalement de l'enzyme Drosha et de son cofacteur «DiGeorge syndrome critical region gene 8» (DGCR8), mieux connu sous le nom de Pasha chez le nématode *C. elegans* et la drosophile (Denli et al., 2004; Han et al., 2004). Le complexe Drosha/DGCR8 reconnaît le pri-miARN en deux temps (Fig.7). Premièrement, DGCR8 contient deux domaines de liaison à l'ARN double brin qui sont essentiels à la reconnaissance de la portion appariée au niveau de la tige-boucle, qui est généralement d'une longueur de 33 nt. Deuxièmement, DGCR8 reconnaît les deux extrémités simples brins du pri-miARN. Suite à cette reconnaissance spécifique, DGCR8 stabilise Drosha pour permettre le clivage au bon endroit à 11 paires de bases (pb) de la jonction simple brin-double brin du pri-miARN, pour donner le pré-miARN. Le clivage par le complexe Drosha-DGCR8/Pasha se fait simultanément à la transcription et à l'épissage de l'ARNm

dans lequel se trouve le précurseur de miARN (Kim et al., 2009). Son excision de séquences introniques ne perturbe en aucun cas le recrutement du spliceosome et le processus d'épissage de l'intron. Au niveau des introns, certains miARN, les mirtrons, représentant environ 5 à 10% de tous les miARN chez les vertébrés et invertébrés, ne sont pas générés par Drosha-DGCR8 mais directement par la machinerie d'épissage (Kim et al., 2009).

Le produit de clivage du microprocesseur (ou du spliceosome dans le cas des mirtrons) se nomme pré-miARN. Sa structure se résume en un ARN de 60 à 100 nt ayant un aspect d'épingle à cheveux avec, à sa base, 2 nt de plus à l'extrémité 3'OH qu'à l'extrémité 5' phosphate (5'P). Cette marque est caractéristique du clivage par les enzymes de type RNase III (MacRae and Doudna, 2007) et sert au transport du pré-miARN hors du noyau par le complexe Exportin 5/Ran GTPase, qui reconnaît spécifiquement l'extrémité sortante 3'OH de molécule d'ARN en tige boucle (Bohnsack et al., 2004; Yi et al., 2003).

**Figure 7 : Complexe Drosha-DGCR8-pri-miARN.** L'orientation des deux domaines RNases III (RIII) va permettre la création d'une extrémité 3'OH sortante. La protéine DGCR8 se fixe à Drosha et permet l'ancrage correct de celle-ci sur le pri-miARN par l'intermédiaire de sa liaison avec la base simple brin. Tiré de Kim et al., 2009.



Chez le nématode, la délétion de Drosha est létale et est associée à une incapacité à produire des miARN matures. Les expériences d'ARNi sur ce gène ont montré des phénotypes sévères de stérilité montrant le caractère essentiel des miARN dans les processus développementaux du ver (Denli et al., 2004). Au niveau embryonnaire, la descendance qui porte deux allèles mutés aura des arrêts développementaux à un stade très précoce lors de la

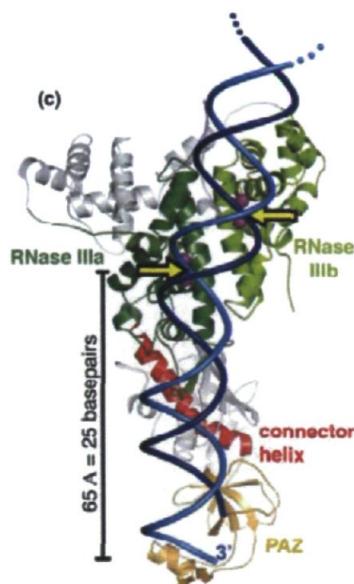
phase de prolifération. Les expériences d'ARNi sur le partenaire, Pasha, ont montré des défauts au niveau de la production de miARN avec une accumulation de pri-miARN. De façon similaire au mutant *let-7*, le développement de la vulve et des ailes est affecté (Denli et al., 2004).

### 1.4.2.3. Maturation cytoplasmique par Dicer

La formation du complexe miRISC mature, qui va cibler un ARNm afin d'inhiber la traduction, consiste en une Argonaute chargée du brin miARN mature associé à d'autres protéines. Pour ce faire, les pré-miARN vont être pris en charge par un premier complexe nommé RLC dont l'acteur majoritaire, Dicer, élimine la boucle du précurseur et conduit à un duplex dont les deux extrémités arborent des extrémités 3'OH (Bernstein et al., 2001). L'unique structure cristalline de Dicer provient de *Giardia intestinalis*, un protozoaire flagellé parasite de l'intestin, dont la protéine ne possède que les deux domaines RNase III et le domaine PAZ (Fig.8). Les informations de cette structure ont permis de bâtir un modèle où la protéine Dicer fonctionne comme une règle moléculaire : la reconnaissance de l'extrémité 3'OH par le domaine PAZ crée une coupure s'effectuant exactement à deux tours d'hélices des extrémités 5' et 3' de la tige boucle et correspond à un duplex d'environ 21 à 25 nt (Macrae et al., 2006). Il a été déterminé très récemment que la protéine recombinante Dicer humaine avait également besoin de l'extrémité 5'P pour la précision et l'efficacité de clivage des molécules d'ARN (Park et al., 2011).

Chez le nématode, la mutation du gène de Dicer montre des phénotypes plus forts que ceux observés pour les mutants *Droscha*. Par ARNi, on observe une accumulation de la forme pré-miARN, substrats de Dicer, et une baisse drastique des miARN produits (Grishok et al., 2001). Phénotypiquement, on observe des défauts de développement de la vulve qui vont la rendre très fragile et une certaine population de vers vont mourir par rupture de cet organe. La synthèse des ailes est très affectée et les vers traités par ARNi affichent des défauts ou une perte totale de ces structures.

**Figure 8 : Complexe Dicer/pré-miARN.** Le domaine PAZ lie l'extrémité 3'OH de son substrat. Les deux domaines RNase III sont orientés, tout comme Drosha, pour un clivage décalé de l'ARN db. Le premier site de clivage se trouve à 25 pb de l'extrémité 3'OH. Modifié de MacRae et Doudna, 2007.



#### 1.4.2.4. Les cofacteurs de Dicer : le complexe RLC

L'isolement du RLC a démontré la présence de plusieurs cofacteurs protéiques liés à Dicer qui vont contribuer à la suite des réactions menant au chargement du miARN mature sur la protéine Argonaute. Chez l'homme, «TAR RNA Binding Protein» (TRBP) est retrouvé dans un complexe où sa présence est nécessaire à Dicer pour le recrutement de la protéine Ago2 (Chendrimada et al., 2005). TRBP possède 3 domaines dsRBDs dont le dernier domaine en C-terminal est impliqué dans la liaison avec la protéine Dicer (Haase et al., 2005). Cependant, l'absence de TRBP n'est pas aussi essentielle pour Dicer que DGCR8 l'est pour Drosha. En effet, en l'absence de son cofacteur, la production de miARN par Dicer est réduite, mais demeure fonctionnel (Lee et al., 2006). Une seconde protéine, «PKR-activating protein» (PACT) a été retrouvé interagissant avec Dicer (Lee et al., 2006). Une inhibition de PACT provoque aussi une baisse de la formation de miARN mature, mais il semblerait que ce dernier cofacteur intervienne après l'action de Dicer en favorisant la formation du complexe miRISC (Lee et al., 2006).

Chez la drosophile, il existe deux paralogues, DCR-1 et DCR-2. Alors que DCR-2 intervient dans la voie des siARN, DCR-1 est requise pour la voie des miARN. Cette

dernière possède comme cofacteur la protéine «Loquacious» (LOQS), qui contient 3 domaines dsRBDs (Forstemann et al., 2005; Saito et al., 2005). Dans un extrait cellulaire de drosophile, une inhibition de la protéine LOQS entraîne une baisse significative du niveau de miARN matures. Comme chez l'humain, la protéine à dsRBD n'est pas absolument requise pour la maturation d'un pré-miARN. En ce qui concerne *C. elegans*, aucun homologue à cette famille de protéines n'est connu comme participant à l'action de Dicer lors de la maturation du pré-miARN, ce qui laisse envisager que le complexe RLC serait composé uniquement de protéines Argonaute et de Dicer.

#### 1.4.2.5. Dicer et les protéines Argonautes

Le mécanisme exact par lequel la maturation par Dicer engendre le chargement du miARN sur la protéine Argonaute reste à élucider. Néanmoins, il a été montré que Dicer se lie aux protéines Argonaute par une interaction entre les domaines RNase III et le domaine «P-element Induced Wimpy Testis» (PIWI) des protéines Argonautes (Tahbaz et al., 2004). Les protéines Argonaute, douées d'activité catalytique, effectueraient un clivage du brin miARN\* (miARN passager) pour permettre sa dégradation. Néanmoins, comme nous le verrons par la suite, certaines protéines Argonaute ne possèdent pas d'activité endonucléasique et, pour certains miARN, les mésappariements du duplex ne permettent pas toujours le clivage (Kim et al., 2009). Selon certaines études, les protéines Argonaute auraient la capacité de séparer le duplex de miARN dans un processus indépendant de l'activité endonucléasique (Wang et al., 2009a). Ainsi, deux modèles sont communément admis pour le passage à la forme simple brin : clivage/dégradation du passager et/ou intervention d'une activité hélicase présumée (Khvorova et al., 2003). Concernant la sélection du brin miARN dans le duplex, elle est déterminée par les extrémités de la molécule d'ARN. Un duplex de miARN possède toujours une asymétrie au niveau de la stabilité thermodynamique de ses extrémités et le brin dont le 5' est le plus instable sera favorablement sélectionné pour former le miRISC (Schwarz et al., 2003).

Il existe deux protéines Argonaute impliquées dans la voie des miARN chez *C. elegans* : ALG-1 et ALG-2. Le double mutant *alg-1; alg-2(ARNi)* affiche des phénotypes étroitement liés aux défauts observés pour les mutants des deux RNases III impliquées dans la maturation des miARN (Grishok et al., 2001). Le double mutant *alg-1; alg-2* est létal et des

expériences menées dans le laboratoire indiquent que, au niveau embryonnaire, le développement s'arrête entre les stades 2-fold et 3-fold (Vasquez-Rifo et al., 2012). Les expériences d'ARNi sur ALG-1/2 ont pour conséquence une absence d'alaie et des animaux mourant par rupture au niveau de la vulve (Grishok et al., 2001). Au niveau moléculaire, on voit une baisse importante de miARN matures avec, parfois, une accumulation des formes précurseurs, notamment dans le cas de *lin-4*. Cette accumulation de précurseurs peut s'expliquer par le fait que les protéines Argonaute sont requises pour l'activité de Dicer. La chute du niveau des miARN peut être due à une baisse de la stabilité des miARN en l'absence des protéines Argonaute, ou par le fait que ces protéines sont impliquées dans la maturation des miARN.

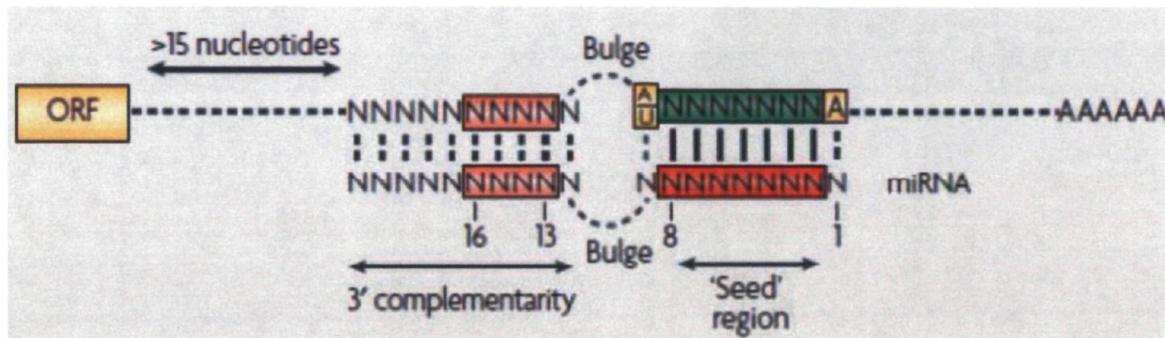
### **1.4.3. Mécanisme d'inhibition de la traduction par les miARN**

#### **1.4.3.1. Interaction entre le miARN et sa cible**

Afin d'inhiber la traduction, le miRISC formé devra cibler de façon spécifique un ARNm. L'interaction entre le miARN et la région ciblée de l'ARNm s'effectue généralement dans la région 3'UTR, comme il a été décrit initialement lors de la découverte de *lin-4* et *let-7* (Lee et al., 1993; Reinhart et al., 2000). Ces études ont confirmé que les régions complémentaires entre le miARN et le 3'UTR n'étaient que de quelques nt, de la position 2 à 8 en partant du 5'P du miARN (Fig.9). De plus, parmi les miARN conservées entre les espèces, cette région appelée «seed» est celle qui est la moins variable et il en est de même pour les régions de 3'UTR de cible également conservée (Bartel, 2004). En revanche, chez les plantes, la plupart des miARN sont parfaitement complémentaires dans toute leur séquence au 3'UTR ciblé et le miRISC des plantes effectue un clivage de l'ARNm (Bartel, 2004; Filipowicz et al., 2008).

Plus tardivement, des miARN agissant sur d'autres régions ont été décrits. On peut citer l'exemple du miARN miR-10a, dont l'action sur les régions 5'UTR des ARNm de protéines ribosomales, conduit à une activation de la traduction lorsqu'il est lié à sa cible (Orom et al., 2008). Au niveau des régions codantes, le miARN miR-148 cible un variant d'épissage codant pour l'ADN méthyltransférase 3 en induisant l'inhibition de ce variant de manière spécifique (Duursma et al., 2008).

**Figure 9 : Représentation de l'appariement 3'UTR-miARN.** Généralement, le premier nucléotide du miARN n'est pas apparié. Il y a deux régions appariées, surlignées en rouge, soit la région «seed» formée des nucléotides 2 à 8 (foncé) et une région 3' complémentaire de 13 à 16 (clair). La région centrale ne présente pas de complémentarité. Tiré de Filipowicz et al., 2008.



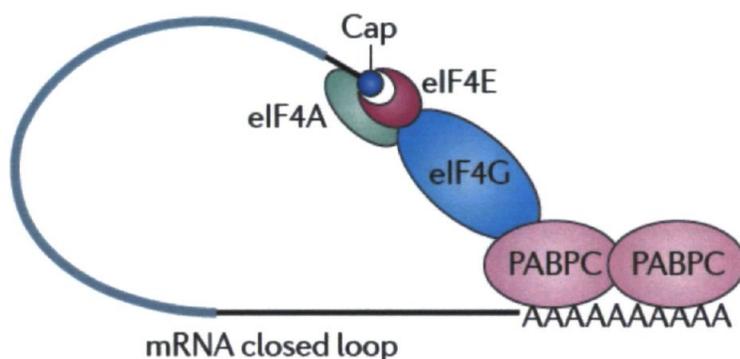
#### 1.4.3.2. La machinerie de traduction

Avant d'entrer dans le détail de l'inhibition de la traduction, il est important d'introduire brièvement la traduction des ARNm et notamment l'étape de formation du complexe d'initiation (Fig.10). La traduction des ARNm se déroule en 3 étapes : l'initiation, l'élongation et la terminaison. L'étape d'initiation commence par la reconnaissance de la coiffe, un motif  $m^7G$  (7-méthylguanosine) en 5' de l'ARNm, par le facteur d'initiation eIF4E, lui-même sous-unité du complexe eIF4F, comprenant également eIF4A et eIF4G (Filipowicz et al., 2008). Le facteur eIF4A possède une activité hélicase nécessaire pour éliminer les structures secondaires de la molécule d'ARNm. Toutes les protéines impliquées au niveau de l'initiation sont liées au facteur eIF4G, fonctionnant comme une protéine adaptatrice pour la formation du complexe et la circularisation de la structure du complexe d'initiation. Toutes ces étapes sont nécessaires pour initier la traduction. eIF4G permet la liaison de façon spécifique à la queue polyadénylée (polyA), par l'intermédiaire de la protéine «PolyA Binding Protein 1» (PABP1).

Le complexe d'initiation formé va permettre l'arrimage de la sous-unité 40S du ribosome à eIF4F. Ce recrutement se fait par l'intermédiaire de la liaison d'eIF4G à un autre complexe eIF3. Le 40S se déplace alors dans la région 5'UTR, le début de la lecture commence jusqu'à la détection du codon initiateur AUG. La lecture de ce codon va provoquer le début de la traduction avec la formation du ribosome 80S, suite à l'arrimage du 60S au 40S. Dans

la formation du 80S, un facteur joue un rôle important, il s'agit d'eIF6. Ce facteur anti-traduction séquestre le 60S afin d'éviter la formation précoce du ribosome 80S (Russell and Spremulli, 1978).

**Figure 10 : Complexe d'initiation de la traduction.** eIF4F est composé de eIF4A, eIF4G et eIF4E. La reconnaissance du motif  $m^7G$  se fait via eIF4E, le facteur eIF4A élimine les structures secondaires de la molécule d'ARNm. eIF4G est une protéine adaptatrice qui permet la circularisation du complexe en liant eIF4E et PABP. Tiré de Filipowicz et al., 2008.



#### 1.4.3.3. Inhibition de la traduction médié par les miARN

Il y a plusieurs modèles d'inhibition de la traduction par les miARN. Je décrirai ici quelques exemples qui ont pu conduire à favoriser un modèle de blocage de l'initiation ou post-initiation. Concernant le modèle d'inhibition à l'étape d'initiation, les exemples les plus convaincants sont les études biochimiques qui ont démontré que la répression dépend de la coiffe et la queue polyA (Humphreys et al., 2005; Pillai et al., 2005), deux éléments qui permettent la circularisation de l'ARNm. L'absence d'un des deux éléments abolit la répression médiée par les miARN. Ce modèle a été indirectement appuyé par les travaux menés sur certains ARNm cellulaires et viraux possédant des «Internal Ribosome Entry Site» (IRES). Ces ARNm ont la particularité d'avoir un mode de recrutement de la sous-unité 40S du ribosome indépendant de la coiffe  $m^7G$  et d'eIF4E et il a été démontré que la répression traductionnelle par les miARN était peu efficace dans ces conditions (Mathonnet et al., 2007; Thermann and Hentze, 2007). Nous verrons dans la section traitant des protéines Argonaute, l'implication des protéines dans l'interaction avec le complexe d'initiation.

Plusieurs études ont aussi conduit à élaborer un modèle post-initiation. Tout d'abord, le premier miARN découvert, *lin-4*, inhibe la traduction sur l'ARNm de *lin-14* sans pour autant affecter le niveau de l'ARNm ou le profil des polysomes (Olsen and Ambros, 1999). Ces observations suggèrent que l'étape d'élongation peut être affectée lors d'une inhibition par les miARN. Des études basées sur des expériences de co-immunoprécipitation font état de l'existence d'un complexe formé par TRBP/Ago2/Dicer et d'eIF6 (Chendrimada et al., 2007). Selon ces données, une réduction d'eIF6 s'associe, chez *C. elegans* et chez les cellules humaines en culture, à une réduction de la répression traductionnelle. L'interaction du miRISC avec eIF6 favoriserait un modèle de répression post-initiation. Dans ce sens, des études ont montré que le miRISC se retrouvait avec les polyribosomes dans des expériences de fractionnements par gradient de sucrose et qu'il inhibait la traduction à ce niveau et non au niveau de l'initiation (Thermann and Hentze, 2007). Dans le laboratoire, nous avons mis en évidence l'interaction directe entre le miRISC et le ribosome en montrant qu'un composant de la sous-unité 40S, RACK1 («Receptor for Activated protein kinase C 1») est important la régulation miARN-dépendante des gènes chez les animaux (Jannot et al., 2011a). Ces données indiquent que RACK1 peut contribuer au recrutement du miRISC sur le site de la traduction via une interaction directe avec les protéines Argonaute.

#### 1.4.3.4. Devenir de l'ARNm ciblé

Initialement, la communauté scientifique s'accordait à dire que la répression traductionnelle par les miARN mène uniquement à un blocage traductionnel, n'agissant pas sur la stabilité de l'ARNm (Olsen and Ambros, 1999; Seggerson et al., 2002). Cependant, des études récentes ont démontré certains cas où l'intégrité de l'ARNm est altérée. Chez le poisson-zèbre, le miARN *miR-430* joue un rôle majeur dans le développement embryonnaire, puisqu'il cible plusieurs centaines d'ARNm hérités de la mère (Shyu et al., 2008). Pour que l'embryon accède à d'autres stades de développement, les ARNm maternels doivent être réprimés de manière efficace. Pour ce faire, tous les ARNm ciblés par *miR-430* sont entraînés vers un processus de dégradation, initié en premier lieu par une déadénylation (Giraldez et al., 2006). Le même mécanisme de déadénylation a été observé chez des cellules humaines en cultures (Wu et al., 2006). Chez le nématode *C. elegans*, les ARNm

lin-41, lin-14 et lin-28, ciblés par les miARN let-7 et lin-4, sont dégradés (Bagga et al., 2005). Toujours chez le nématode, une autre étude a permis de démontrer que certains miARN embryonnaires entraînaient la dégradation de l'ARNm ciblé, notamment par une déadénylation (Wu et al., 2010). Les résultats démontrent également que la déadénylation des ARNm ciblés n'est pas un phénomène constant et que pour un même miARN, le ciblage des ARNm inhibe systématiquement la traduction alors que, dépendamment du contexte de la région 3'UTR, la dégradation par déadénylation est plus variable (Wu et al., 2010).

Les études portant sur l'inhibition traductionnelle par les miARN ont conduit à l'isolement du complexe miRISC dans des foyers cellulaires particuliers, les «P bodies», qui interviennent dans le métabolisme de l'ARN. Ces foyers sont également appelés «GW-bodies» car ils contiennent des protéines riches en motifs GW, dont les protéines «Trinucleotide repeat-containing gene 6» (TNRC6) chez l'homme (Jakymiw et al., 2005; Liu et al., 2005). Les protéines de cette famille ont la capacité de lier l'ARN et les protéines Argonaute, et sont nécessaires pour la répression traductionnelle médiée par le miRISC (Eulalio et al., 2008; Liu et al., 2005). La fonction de cette famille de protéines est conservée chez les organismes possédant une voie des miARN. Chez *C. elegans*, deux homologues, «ALG-1 INteracting» 1 et 2 (AIN-1 et AIN-2), jouent ce rôle (Ding et al., 2005; Zhang et al., 2007). De par un motif conservé d'une vingtaine de résidus, les protéines GW lient spécifiquement les protéines Argonaute par leur domaine PIWI pour les conduire au «P-bodies» (Till et al., 2007). Au niveau de ces foyers, il y a présence d'enzymes de dégradation de l'ARN, comme la déadénylase Ccr4 (Andrei et al., 2005; Sheth and Parker, 2003), le complexe de décoiffage DCP1-DCP2 et des exoribonucléases 5'-3', tel que l'enzyme XRN-1 (Ingelfinger et al., 2002). L'importance des P-bodies a été appuyé par une étude qui indique que le complexe RCK/p54, intervenant dans leurs formations, est nécessaire pour la répression traductionnelle par les miARN (Chu and Rana, 2006), mais n'a que peu d'effet sur la répression par les siARN. Néanmoins, plusieurs zones d'ombres demeurent car il est également démontré que l'abolition de la formation des «P-bodies» par une inhibition de Lsm-1, un autre facteur intervenant dans la formation de ces foyers, n'a aucun effet sur la répression traductionnelle par les miARN (Chu and Rana, 2006). Ces foyers sont également dépourvus de composants ribosomiaux, de la

plupart des facteurs d'initiation de la traduction et des protéines PABP, ce qui soutient l'idée que les «P-bodies» sont des sites où la traduction est réprimée. Toutes ces données indiquent que les «P-bodies» sont plus certainement une conséquence qu'une cause de la répression traductionnelle : de par leur nature réversible et dynamique, les «P-bodies» sont un lieu de stockage des ARNm ciblées par les miARN, qui peuvent mener à la répression puis à la dégradation, ou bien encore à un retour à la traduction (Eulalio et al., 2007).

#### 1.4.4. Stabilité des miARN

Outre le contrôle de la transcription qui influence directement les niveaux de miARN, plusieurs facteurs affectent la stabilité des miARN de façon post-transcriptionnelle. Une étude a décrit un mécanisme d'augmentation des niveaux de miR-21, sans pour autant activer sa transcription (Davis et al., 2008). Sous l'influence des voies de signalisation de «Transforming growth factor beta» (TGF- $\beta$ ) et «Bone Morphogenetic Protein» (BMP), la protéine SMAD et l'hélicase p68 rendraient Drosha plus efficace dans la maturation du pri-miR-21 en pré-miR-21. À l'opposé, l'exemple de let-7 et de la protéine LIN-28, dont la traduction est régulée par let-7 est assez intéressant. LIN-28 est une protéine de liaison à l'ARN et va, par reconnaissance de la boucle de la structure de pré-let-7, recruter l'enzyme «Terminal Uridyl Transferase» (TUTase) pour ajouter une série de résidus U à l'extrémité 3'OH de la molécule d'ARN (Heo et al., 2008; Heo et al., 2009). Cette modification cytoplasmique a pour effet de bloquer la maturation du pré-miARN let-7 par l'enzyme Dicer (Newman et al., 2008; Viswanathan et al., 2008).

Chez le ver, un groupe a récemment identifié les protéines XRN-1/2 comme fonctionnellement similaires à SDN-1, connues chez la plante pour être directement impliquées dans la dégradation des miARN (Chatterjee and Grosshans, 2009). Ces enzymes, par leur activité exonucléasique, sont capables de dégrader de courtes molécules d'ARN. Les expériences de ce groupe ont démontré que cette dégradation était dépendante de la présence d'une cible pour le miARN (Chatterjee et al., 2011). L'augmentation de la quantité d'ARN cibles a pour effet de diminuer la dégradation des miARN. Ils ont ainsi pu prévenir la dégradation des brins passagers de miARN, par l'addition de séquences complémentaires et les rendre assez abondants pour être détectables par analyse de type

Northern. D'après les auteurs, ce processus est intimement lié avec le processus de sélection des brins après le clivage par Dicer.

## **2. Les siARN**

### **2.1. Généralités**

Les siARN sont une classe de courts ARN fortement sujets à confusion. Ce qui distingue les siARN des miARN, c'est leur complémentarité parfaite avec leur cible qui induit le clivage de celle-ci, et ce, chez tous les organismes, à l'exception des plantes. Chez les mammifères, les miARN ayant ces caractéristiques existent, mais sont peu nombreux dans la littérature. Chez le nématode, l'inhibition de gènes par des molécules d'ARN parfaitement appariées est devenue une méthode courante (Fire et al., 1998; Montgomery et al., 1998; Timmons and Fire, 1998). Le nématode possède deux voies, l'une prenant en charge les courts ARN exogènes (exo-ARNi), dont on discutera dans cette thèse, et l'autre intervient pour les courts ARN endogènes (endo-ARNi), un sujet qui ne sera pas abordé.

Chez les organismes possédant une voie de siARN, plusieurs caractéristiques sont partagées. La première est la fonction de protection contre les virus et les éléments transposables. Chez *C. elegans*, la fonction de la voie de l'exo-ARNi n'a été déterminée que très récemment avec la découverte d'une réponse antivirale ARNi-dépendante (Felix et al., 2011; Rechavi et al., 2011). La seconde caractéristique, décrite précédemment, est la complémentarité parfaite avec l'ARN cible, qui conduit à la dégradation de celui-ci et qui est initié par un clivage d'une protéine Argonaute. Pour ce qui est des plantes et du nématode, ils partagent deux processus très importants : la nécessité d'un ARN déclencheur qui va initier la réponse ARNi, et l'étape d'amplification du signal, nécessitant l'activité d'une ARN polymérase ARN dépendante.

### **2.2. La voie de l'exo-ARNi**

Chez le nématode *C. elegans*, l'introduction de l'ARN exogène peut se faire de quatre manières : la micro-injection des molécules d'ARNdb directement à l'intérieur de l'animal (Fire et al., 1998), l'incubation des vers dans une solution concentrée d'ARNdb, l'ingestion

de bactérie exprimant l'ARN db (Tabara et al., 1998; Timmons and Fire, 1998) ou la création de lignées de vers exprimant un ARN db (Tabara et al., 1999). Un récapitulatif est présenté Figure 11.

### **2.2.1. La transmission systémique de l'ARNdb**

L'ARN déclencheur de l'ARNi ingéré, micro-injecté ou absorbé doit passer de cellule en cellule pour pouvoir affecter le ver entier. Pour ce faire, deux protéines sont nécessaires pour le transport des molécules d'ARN : «Systemic RNAi Defective» 1 et 2 (SID-1 et SID-2). Ces protéines transmembranaires ont été découvertes par un criblage génétique visant des gènes de résistance à la diffusion de l'ARNi au niveau des tissus (Winston et al., 2002). La protéine SID-1 est exprimée dans toutes les cellules. SID-2 est exprimée au niveau des cellules de l'intestin, et cette dernière est nécessaire pour l'ARNi par ingestion de bactérie ou d'ARN db (Winston et al., 2007), mais n'est pas nécessaire à la transmission de l'ARNi de cellule à cellule. En ce qui concerne la nature de l'ARNdb transmis, celui-ci reste une énigme. Une étude génétique a récemment conclu que la transmission de l'ARNi dans le ver nécessitait les protéines impliquées très en amont dans la voie, tels que «RNAi defective-4» (RDE-4), DCR-1 et MUT-2, ce qui suggère que l'ARN transmis est l'ARN déclencheur ou une forme transformée de cet ARN (Jose et al., 2011).

### **2.2.2. Synthèse des siARN primaires**

Les siARN primaires sont les premiers à être synthétisés et dérivent de l'ARN déclencheur. Les longs ARNdb vont être maturés par DCR-1 en courts duplexes de 21 nt (Bernstein et al., 2001). Cette maturation commence toujours à une extrémité du long ARNdb (Ketting et al., 2001; Warf et al., 2011) mais, à la différence de la voie des miARN, une protéine de liaison à l'ARNdb, RDE-4, est requise pour que DCR-1 génère ces duplexes (Ketting et al., 2001; Tabara et al., 2002). RDE-4 interagit directement avec la protéine Argonaute impliqué dans cette initiation, « RNAi defective-1» (RDE-1) (Tabara et al., 1999). RDE-4 lie également un autre facteur ayant potentiellement une activité hélicase DRH-1 (DEXH-box helicase) (Tabara et al., 2002). Toutes les protéines citées se retrouvent dans le complexe immunoprécipité de DCR-1, enzyme partagée pour les voies des miARN et de l'exo-ARNi (Duchaine et al., 2006).

Dans ce processus, la présence de RDE-1 est nécessaire, car le mutant est incapable d'une réponse ARNi en contact d'ARNdb (Tabara et al., 1999). En utilisant des colonnes d'ARN 2'-O-méthylés (voir Matériels et méthodes), il a été démontré que RDE-1, lors de l'ARNi, pouvait être isolé avec des ARN 2'-O-méthylés dirigés contre des siARN primaires (Yigit et al., 2006). L'activité endonucléasique de RDE-1 est nécessaire pour le passage de la forme duplex à ARNsb par le clivage du brin passager du duplex (Steiner et al., 2009).

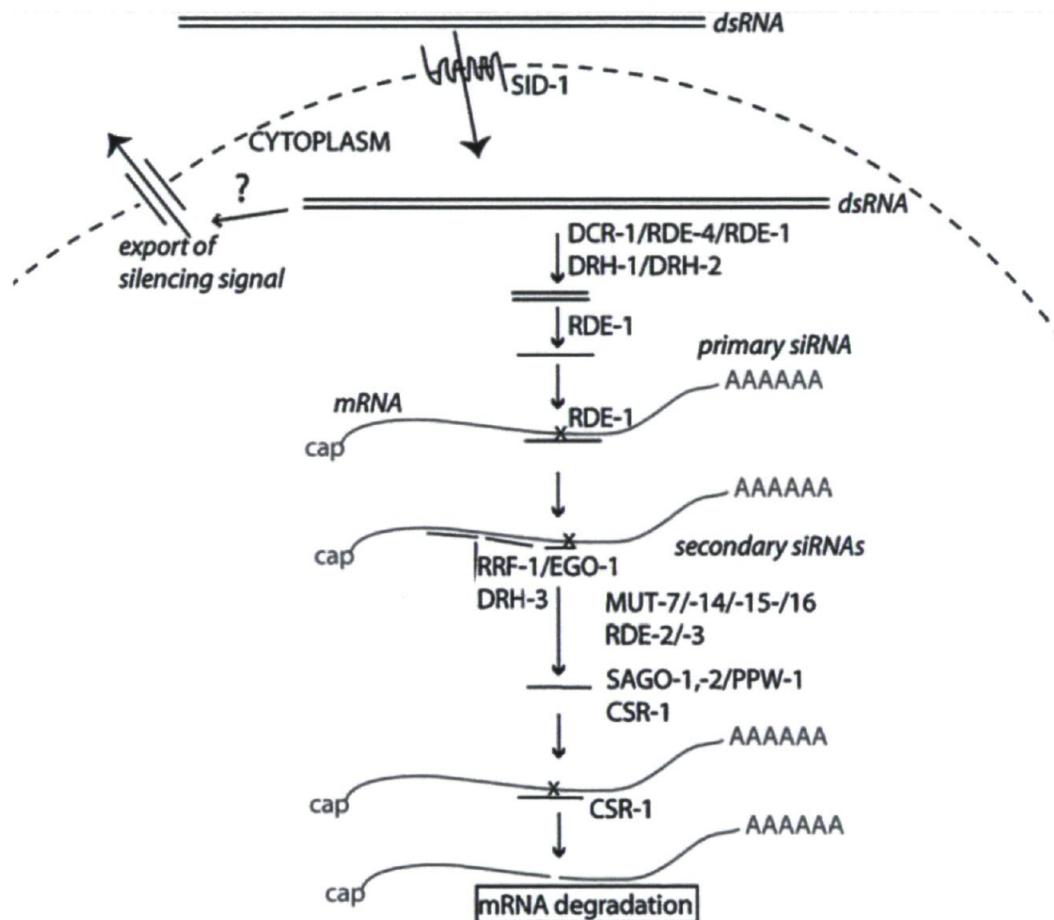
### 2.2.3. Synthèse des siARN secondaires

Après le clivage des duplexes de 21 nt par RDE-1, le mécanisme est peu connu. On ne sait pas si le complexe RDE-1-siARN primaire va cibler un ARNm et le dégrader. Mais il est clair que le complexe RDE-1-siARN-ARNm va permettre le recrutement de différents facteurs pour procéder à une étape essentielle : l'amplification du signal. Elle consiste en l'utilisation de l'ARNm ciblé pour produire une multitude de siARN secondaires complémentaire à la séquence de l'ARNm en amont (en 5') des sites ciblés par les siARN primaires (Alder et al., 2003). La production de siARN secondaires est indépendante de DCR-1 (Aoki et al., 2007) et pour la voie de l'exo-ARNi, les polymérase nécessaires sont RRF-1 (Sijen et al., 2001) et, spécifiquement dans les lignées germinales, EGO-1 (Smardon et al., 2000). Ils sont souvent di- ou tri-phosphorylés, signe d'une incorporation de nucléotides *de novo* impliquant l'activité d'une ARN polymérase ARN dépendante (Pak and Fire, 2007; Sijen et al., 2007).

La prise en charge des siARN secondaires est également très peu connue. Des expériences d'immunoprécipitation d'une catégorie de protéines Argonaute spécifique à *C. elegans*, la famille MAGO, ont montré un enrichissement pour les siARN secondaires (Yigit et al., 2006). L'analyse bioinformatique des 6 membres de la famille MAGO a révélé l'absence des sites requis pour supporter une activité endonucléasique, et un sextuple mutant de ces 6 gènes Argonaute est déficient dans les voies de l'exo-ARNi et de l'endo-ARNi (l'exo-ARNi et l'endoARNi partagent ces protéines Argonaute dites secondaires). La réintroduction de SAGO-1 (membre de la famille MAGO) dans le sextuple mutant est suffisante pour le rétablissement de la fonction. Le fait que les protéines Argonaute secondaires ne conservent pas de résidus essentiels pour une activité endonucléasique a orienté vers un modèle de répression traductionnelle et/ou transcriptionnelle par ces

protéines (Yigit et al., 2006). Enfin, une dernière étude biochimique a mis en doute ce modèle en démontrant que la protéine Argonaute CSR-1 avait une affinité particulière pour les siARN secondaires et une activité endonucléasique importante en utilisant ce type de molécules, suggérant ainsi que ces protéines pourraient dégrader les ARNm ciblés par ce moyen (Aoki et al., 2007). Cependant, étant donné que CSR-1 est associé aux chromosomes, son rôle reste à éclaircir.

**Figure 11 : Voie des siARN chez *C. elegans*.** Récapitulatif des différents acteurs de la voie des siARN primaires et secondaires. Modifié de Fisher, 2011.



## 3. Les piARN

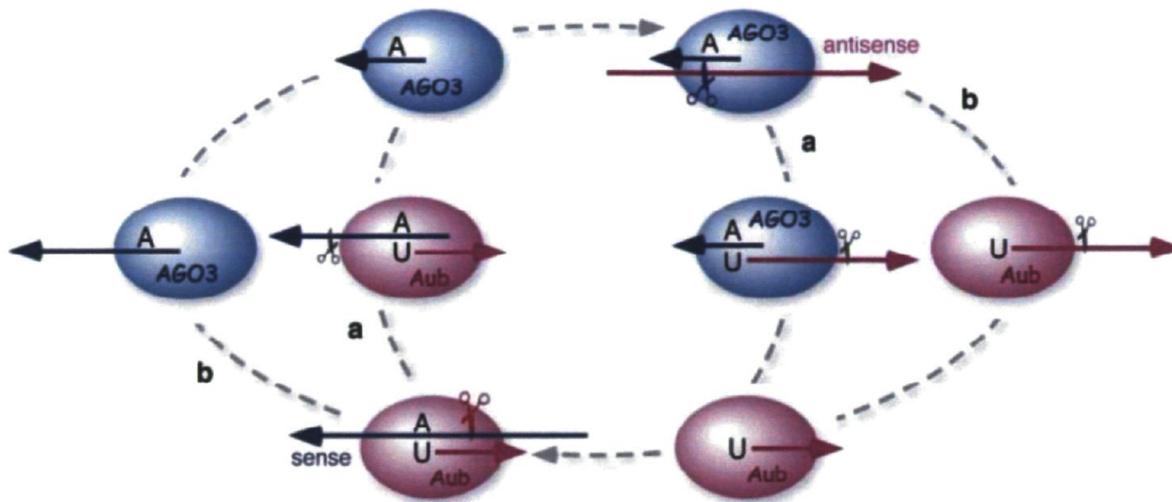
### 3.1. Découverte et fonction des piARN

Les piARN ont été découverts chez la drosophile en étudiant la relation du gène *Stellate* et son suppresseur *Su(Ste)*, composés de séquences répétées intergéniques qui interviennent dans le maintien de la fertilité chez les mâles (Aravin et al., 2001). Il est découvert que le transcrite de *Su(Ste)* forme un ARNdb spécifique aux lignées germinales et que de petits ARN, se distinguant des miARN de part leur taille de 24 à 29 nt, complémentaires aux brins sens et antisens du transcrite *Su(Ste)* en sont probablement dérivés. Il est connu que certains mutants de gènes codant pour des protéines Argonaute ont un rôle dans les lignées germinales : la protéine «P-element induced wimpy testis» (PIWI) est impliquée dans le maintien des cellules souches germinales (Cox et al., 1998) et une autre protéine, Aubergine (AUB) est importante dans le développement embryonnaire (Harris and Macdonald, 2001). De façon expérimentale, la protéine PIWI est montrée comme étant essentielle au contrôle du rétrovirus *gypsy* par le locus hétérochromatique *Flamenco*, et on découvre l'existence de courts ARN complémentaires à *gypsy* (Sarot et al., 2004). La protéine Aubergine intervient dans un mécanisme quasi identique au niveau des ovaires. Elle contrôle la rétrotransposition en coopération avec le locus *vasa* et une protéine à activité hélicase potentielle Spindle-E (Vagin et al., 2004). Par immunoprécipitation dans des tissus ovariens, l'association entre les protéines PIWI et cette classe de courts ARN est alors confirmée (Saito et al., 2006). Ce groupe démontre l'existence d'une activité endonucléasique de la protéine PIWI recombinante, suggérant le clivage potentiel des transcrits ciblés. D'autre part, il a été démontré que la protéine AUB lie ces petits ARN, et que le complexe immunoprécipité est capable de dégrader les transcrits de *Stellate* (Nishida et al., 2007). Ce mode de répression spécifique aux lignées germinales est conservé chez les mammifères, et deux groupes publient des travaux où les courts ARN associés à différentes protéines Argonaute sont identifiés, par immunoprécipitation, dans les testicules de souris (Aravin et al., 2006; Girard et al., 2006). Ces courts ARN sont appelés piARN, car ils interagissent avec les membres de la famille PIWI.

### 3.2. Synthèse des piARN chez les mammifères et la drosophile

La synthèse de ces courts ARN est différente de ce qui est connu pour les miARN et les siARN. Tout d'abord, comparée aux autres voies, la caractéristique la plus frappante est que la synthèse des piARN est totalement indépendante de Dicer (Vagin et al., 2006). Ensuite, ils possèdent une modification à leur extrémité 3' : une méthylation sur le 2'OH du ribose (Yang et al., 2006). Cette modification et la protéine qui en est responsable, HEN-1, sont conservées chez les mammifères et la drosophile (Horwich et al., 2007; Kirino and Mourelatos, 2007; Ohara et al., 2007). Concernant leur synthèse, ils sont produits par un mécanisme nommé ping-pong, caractérisé chez la drosophile. Les piARN que lie AGO3 sont, en très grande majorité, orientés du même sens que les transcrits de rétrotransposons (Gunawardane et al., 2007), alors que les piARN associés à PIWI et AUB ont des séquences antisens par rapport au transcrit de rétrotransposons réprimés (Nishida et al., 2007; Saito et al., 2006). L'analyse de la fréquence d'apparitions des nucléotides dans les piARN sens et antisens a mis en évidence un biais pour un A en position centrale des piARN sens et un U en position 5' des piARN antisens, ce qui a conduit les auteurs à établir le modèle de synthèse ping-pong des piARN (Fig.12). Ce modèle est basé sur la caractéristique des protéines Argonaute, que nous verrons par la suite, de cliver l'ARN cible face à la position 10-11 de l'ARN guide (Hammond et al., 2000). Les transcrits de piARN seront reconnus par une première protéine avec un piARN sens. De là, un clivage du transcrit va s'effectuer après le A en position centrale formant un produit antisens commençant par un U. Ce long ARN sera alors dégradé en 3' par une enzyme encore inconnue jusqu'à atteindre la taille des piARN. Ce piARN antisens sera pris en charge par un deuxième groupe de protéines PIWI qui va permettre de cliver les transcrits de rétrotransposons. Le transcrit de rétrotransposons sera mûré de façon similaire en 3' pour permettre la production de piARN sens, fermant ainsi la boucle.

**Figure 12 : Modèle ping-pong de production des piARN.** AGO3 lie les piARN sens (bleu), tandis que AUB lie les piARN antisens. La voie a et b représentent les deux possibilités concernant la maturation en 3'. Le long transcrit clivé peut rester sur la protéine responsable de son clivage (a) ou être transmis à la protéine Argonaute qui va l'utiliser comme ARN guide (b). Tiré de Gunawardane et al., 2007.



### 3.3. Les piARN chez *C. elegans*

Chez le nématode, deux paralogues à la protéine PIWI, «PIWI Related Gene» 1 et 2 (PRG-1 et PRG-2), sont connus (Cox et al., 1998). Parmi les nombreuses classes de courts ARN chez le nématode, de courts ARN de 21 nt, produits de deux larges régions sur le chromosome IV, se distinguent par leur U en 5'OH et une probable modification à l'extrémité 3' (Ruby et al., 2006). Plusieurs groupes arrivent à la conclusion que cette classe de courts ARN sont les piARN de *C. elegans*, bien que leur taille diffère de celle des piARN d'autres espèces (Batista et al., 2008; Das et al., 2008; Wang and Reinke, 2008). Il est montré que les piARN sont exprimés uniquement dans les cellules germinales : on les retrouve à partir du stade L4, où la méiose des gamètes est très active jusqu'au stade adulte. Ils sont présents dans les embryons reflétant une provenance paternelle et/ou maternelle. Ces piARN s'associent spécifiquement aux protéines PIWI et cela est démontré par leur co-immunoprécipitation de PRG-1 (Batista et al., 2008). L'analyse de mutant nul révèle que PRG-1 à un rôle prépondérant par rapport à PRG-2, la mutation du gène codant pour PRG-1 provoque la chute drastique de la quantité de piARN. Comme phénotype majeur, on observe de nombreux défauts dans les lignées germinales. Les mutants *prg-1* et *prg-2* affichent une baisse significative du nombre d'embryons formés par individus et, pour *prg-*

*1*, on observe une stérilité lorsque la température de croissance est augmenté à 25 °C. PRG-1 est localisée dans une structure périnucléaire appelée P-granule, qui joue un rôle de régulation pendant la spermatogénèse. La stérilité s'explique par les défauts occasionnés durant la spermatogénèse, la formation des ovocytes n'étant pas affectée (Wang and Reinke, 2008).

La fonction des piARN chez le nématode consiste également à l'extinction des éléments transposables dans les lignées germinales. Dans les mutants *prg-1*, on observe une augmentation de la transposition de Tc3 (Das et al., 2008). Il n'y a aucune preuve d'un mécanisme de ping-pong dans la production des piARN chez le nématode, et un groupe suggère la collaboration des protéines Argonaute secondaires de la voie des endo-siARN pour l'amplification de ce mécanisme, qui demeure obscur (Das et al., 2008).

## 4. Les protéines Argonaute

Comme on a pu le constater jusqu'ici, les protéines Argonaute sont les effecteurs majeurs des courts ARN. Chez le nématode, depuis l'identification de la première protéine Argonaute RDE-1 (Tabara et al., 1999), il est apparu que cette protéine appartient à une très grande famille de gènes conservés. La plupart des mutants de ces protéines conduisent à des défauts dans le développement : chez les plantes tout d'abord, avec les protéines ARGONAUTE-1 (Bohmert et al., 1998) et ZWILLE (Moussian et al., 1998) mais aussi chez la drosophile avec les protéines PIWI et STING (Lin and Spradling, 1997).

### 4.1. Généralités

#### 4.1.1. Classification des protéines Argonaute

Les protéines Argonaute sont généralement classées en trois groupes paralogues selon leur homologie : le groupe «Argonaute-like» similaire à AGO-1 d'*A. thaliana*, le groupe «PIWI-like» similaire à la protéine PIWI de *D. melanogaster*, le groupe 3 spécifique à *C. elegans* (Hutvagner and Simard, 2008) (Fig.13) .

Ces trois groupes vont lier différentes classes de courts ARN. Le groupe «Argonaute-like» comprend les protéines intervenant dans la voie des miARN. Chez *C.elegans*, les protéines

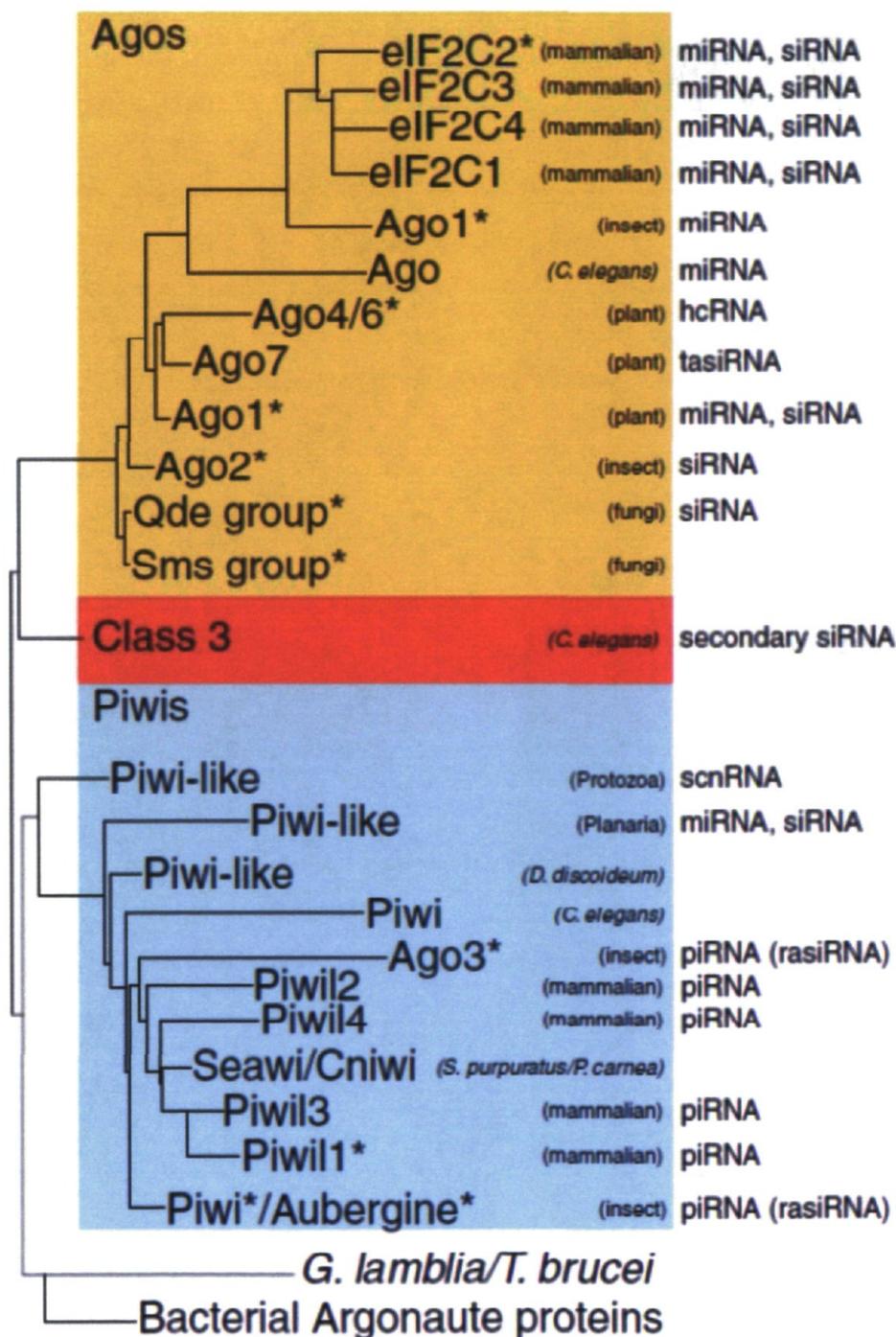
ALG-1 et ALG-2 sont les deux acteurs de cette voie. Les protéines du groupe «PIWI-like» lient les piARN et les deux représentants du nématode sont PRG-1 et PRG-2. Quant au dernier groupe spécifique de *C. elegans*, ils regroupent les protéines pouvant lier les siARN qui sont de deux types : les endo-siARN et les exo-siARN. Ce dernier groupe comprend de nombreuses protéines, dont la protéine RDE-1, effectrice de l'exo-ARNi chez le ver. Il est intéressant de noter que, parmi toutes les espèces chez lesquelles les mécanismes d'ARNi sont présents, *C. elegans* est l'organisme possédant le plus grand nombre de protéines Argonaute avec un total de 26 protéines, dont 5 appartiennent au groupe «AGO-like», 3 au groupe PIWI et 18 au groupe spécifique à *C. elegans* (Hutvagner and Simard, 2008; Yigit et al., 2006). À titre comparatif, l'humain ne possède que 8 protéines Argonaute dont 4 sont apparentées au groupe «AGO-like» et les 4 autres au groupe PIWI. D'après leur degré de similarité, les plus proches des protéines «AGO-like» humaines sont ALG-1 et ALG-2. Pour ce qui est des autres grands modèles d'études où l'ARNi est présent, la drosophile utilise 5 protéines AGO, dont 2 «AGO-like» et 3 PIWI, *A. thaliana* exprime 10 protéines «AGO-like», alors que *S. Pombe* ne comporte qu'une seule protéine «AGO-like».

#### 4.1.2. Les protéines Argonaute procaryotiques

Avant de continuer, il faudrait ouvrir une petite parenthèse concernant les protéines Argonaute procaryotiques. En effet, certaines bactéries possèdent des protéines ayant une forte homologie avec les protéines Argonaute. D'ailleurs, de nombreuses études cristallographiques sont basées sur ces protéines, étant donné qu'elles sont plus faciles à produire chez la bactérie. Cependant, aucun mécanisme d'ARNi médié par des protéines Argonaute n'a encore été mis en évidence chez ces organismes. Très récemment, une nouvelle voie de régulation par de courts ARN a été découverte chez certaines bactéries : il s'agit de la voie CRISPR/Cas. Connue comme étant un phénomène de défense contre les virus, les plasmides envahissants ou les éléments transposables, il a été démontré pour la première fois que les principaux effecteurs de ce mécanisme sont de courts ARN dérivés d'un précurseur, le transcrit bactérien CRISPR (Brouns et al., 2008). Malgré certaines similitudes fonctionnelles, la comparaison des protéines de l'ARNi et CRISPR/Cas indique que les deux mécanismes ont évolué indépendamment. Cependant, une récente étude bio-

informatique suggère que les protéines Argonaute procaryotique, dont la fonction demeure inconnue, pourraient avoir un rôle à jouer (Makarova et al., 2009).

**Figure 13 : Arbre phylogénétique des protéines Argonaute.** Les groupe «Argonaute-like» similaire à AGO-1 d'*A. thaliana* (orange), le groupe similaire à la protéine PIWI de *D. melanogaster* (bleu), et le groupe 3 spécifique à *C. elegans* (rouge) sont indiqués. En bas sont représentés certaines protéines d'organismes particuliers. Tiré de Farazi et al., 2008.



## 4.2. Structure et fonction des protéines Argonaute

Les protéines Argonaute sont composées de 3 domaines fonctionnels : un domaine PAZ, un domaine intermédiaire nommé Mid et un domaine PIWI (Fig.14). Un quatrième domaine très variable en N-terminal est également présent mais ne sera pas décrit pour la simple raison que son rôle est encore très peu documenté.

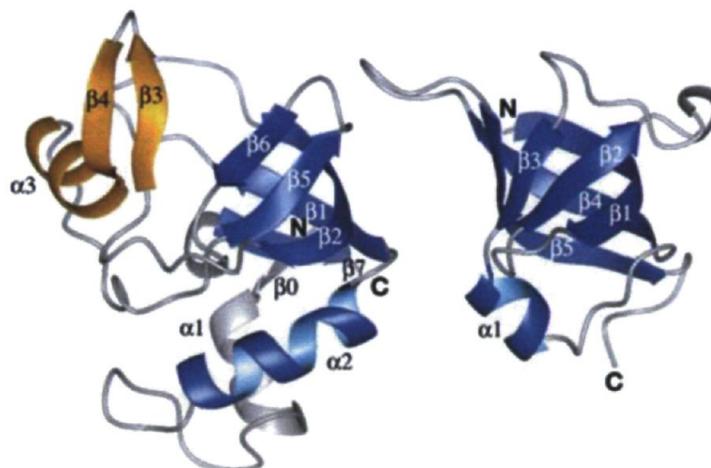
Figure 14 : Domaines caractéristiques d'une protéine AGO. Les domaines N-ter, PAZ, Mid et PIWI sont représentés. Modifié de Hutvagner and Simard, 2008.



### 4.2.1. Le domaine PAZ

Ce domaine est un domaine fonctionnel conservé et présent également chez la protéine Dicer (Cerutti et al., 2000). Il a été le premier domaine des protéines Argonaute dont la structure a été résolue (protéine AGO2 de drosophile) et, qui plus est, c'est un domaine conservé uniquement chez les protéines Argonaute eucaryotes (Lingel et al., 2003). La structure cristalline de ce domaine nous apprend qu'il est composé de 5 feuillets  $\beta$  centraux qui vont former une structure en tonneau ouvert, de deux hélices  $\alpha$  en N-terminal et d'un module spécifique contenant les feuillets  $\beta_3$ ,  $\beta_4$  et  $\alpha_3$  (Fig.15). Ce genre de structure est très similaire à l'architecture des repliements de types «oligosaccharides-oligonucleotides binding» (OB), retrouvés principalement chez des protéines liant des acides nucléiques (Theobald et al., 2003). Cependant, quelques différences sont présentes entre ces deux structures, notamment la présence d'un module spécifique et l'orientation des extrémités N et C-terminale.

**Figure 15 : Comparaison entre la structure du domaine PAZ de la protéine AGO2 de drosophile et de la protéine Rho chez *E. coli*.** AGO2 est à gauche, Rho est à droite. Les structures conservées sont représentées en bleu, alors que le module spécifique à AGO2 est en jaune. Tiré de Lingel et al., 2003.



Il a été démontré que le domaine PAZ lie les acides nucléiques, que ce soit l'ARN ou l'ADN, simple ou double brin, avec une préférence pour l'ARN (Lingel et al., 2003, 2004). Cette interaction est due à une affinité du domaine pour l'extrémité 3'OH des molécules d'acides nucléiques (Song et al., 2003). La reconnaissance du 3'OH se fait par une poche basique située entre le module spécifique des protéines Argonaute et le tonneau  $\beta$ , et n'est pas vraiment dépendant de la présence d'un groupement OH en position 2' du ribose, d'où la capacité de lier également des courtes molécules d'ADN (Ma et al., 2004). Plusieurs acides aminés basiques de cette poche coordonnent et forment des interactions hydrogènes et électrostatiques avec les 5 derniers groupes phosphodiester qui composent l'acide nucléique dont l'extrémité 3' est fixée.

Enfin, deux structures cristallines d'une protéine Argonaute de la bactérie thermophile *Aquifex Aeolicus* nous apprend que le domaine PAZ (et potentiellement le domaine N-terminal en amont) est très flexible (Rashid et al., 2007; Yuan et al., 2005). En effet, les différentes structures obtenues montrent que ce domaine peut prendre différentes orientations et, par conséquent, serait l'élément important et régulateur d'un possible changement de conformation. Une autre étude portant sur la résolution d'un complexe ARN-protéine Argonaute entière de *Archaeoglobus fulgidus* confirme les conclusions des

études précédentes, quant à la poche du domaine PAZ qui fixe l'extrémité 3'OH du court ARN utilisé comme guide (Ma et al., 2005).

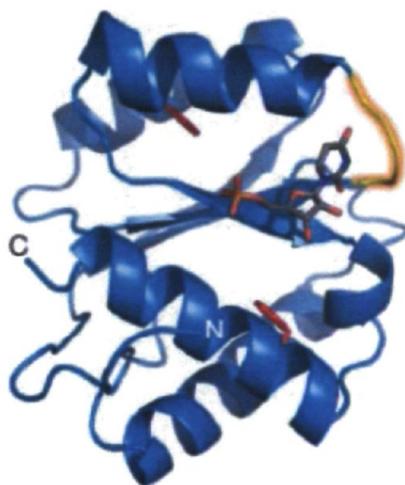
#### 4.2.2. Le domaine Mid

Initialement, ce domaine a été décrit comme un intermédiaire (Mid pour Middle) entre les domaines PAZ et PIWI sans vraiment avoir de fonction assignée (Song et al., 2004). La première évidence de l'existence d'une fonction pour ce domaine provient de cristaux de protéines Argonaute d'*Archeoglobus fulgidus*. Deux études du complexe ARN-Argonaute révèlent les bases d'une spécificité de liaison à l'extrémité 5'P du court ARN guide à une poche hyperbasique formée dans la région Mid (Ma et al., 2005; Parker et al., 2005). Ces études ont identifié les résidus-clés pour la fonctionnalité de la poche et par mutation, la liaison à de courts duplexes d'ARN est affectée. Cette poche est facilement accessible pour le groupement phosphate en 5', quand le premier nucléotide d'un duplex d'ARN n'est pas apparié à sa cible, comme c'est souvent le cas pour les interactions miARN-ARNm (Filipowicz et al., 2008).

Nous avons vu précédemment comment le modèle de la répression traductionnelle à l'étape de l'initiation avait été établi au vu de la dépendance du miRISC pour la coiffe, la queue poly-A et eIF4E. Ce modèle va être appuyé ou réfuté par une série d'études contradictoires sur les caractéristiques du domaine Mid. Une première étude va appuyer ce modèle sur la base d'un motif de reconnaissance dans le domaine Mid pour la coiffe (Kiriakidou et al., 2007) qui s'avère être finalement éronné (Kinch and Grishin, 2009). Une autre étude biochimique plus récente revendique l'existence de deux sites de reconnaissance spécifique dans le domaine Mid des protéines Argonaute apparentées à la classe «AGO-like» : l'un pour la coiffe et l'autre pour les extrémités 5'P des miARN, et suggère un possible changement allostérique (Djuranovic et al., 2010). La structure cristalline du domaine Mid de QDE-2 de *Neurospora crassa*, publié l'année suivante, vient confirmer ces conclusions en identifiant deux poches voisines l'une de l'autre (Boland et al., 2010). La première poche identifiée correspond à la poche de liaison des extrémités 5'P des ARN, mis en évidence par d'autres groupes chez des protéines Argonaute procaryotiques (Ma et al., 2005; Parker et al., 2004, 2005), reflétant ainsi une haute conservation. Enfin, une analyse de différents complexes nucléotides-Mid par résonance magnétique, démontre que la

protéine Ago2 humaine ne possède qu'un seul site de liaison de nucléotides dans le domaine Mid (Frank et al., 2011). Ce même groupe a démontré les bases moléculaires de cette reconnaissance via deux motifs (Fig.16): le premier coordonnant un phosphate grâce à des résidus chargés positivement, le second créant des liaisons hydrogènes préférentiellement avec les U et A (Frank et al., 2010). En conclusion, la présence de un ou plusieurs sites fait toujours débat. Il est très probable qu'en fonction des protéines Argonautes étudiées et/ou des méthodes utilisées, l'interprétation soit différente.

**Figure 16 : Repliement du domaine Mid de la protéine Ago2 humaine.** On voit, représentée en jaune, la boucle importante pour la sélection du nucléotide en 5' ainsi que le placement d'une molécule d'UMP. En rouge, deux phénylalanines, importantes pour la liaison de la coiffe, sont représentées. Tiré de Frank et al., 2009.



#### 4.2.3. Le domaine PIWI

Le domaine PIWI est le domaine le plus conservé parmi les protéines Argonaute eucaryotiques même parmi les représentants procaryotiques qui ne possèdent pas le domaine PAZ. Il est le domaine connu pour être le lieu de l'activité endonucléasique des protéines Argonaute. Depuis la mise en évidence que certains mécanismes d'ARNi provoque la dégradation de l'ARNm (Montgomery et al., 1998; Ngo et al., 1998), une enzyme capable d'initier cette réaction était recherchée. Ainsi, de nombreuses études *in vitro* ont été menées et ont permis de mettre en évidence cette activité endonucléasique. Pour ce faire, le développement d'essais biochimiques sur des lysats d'embryons de drosophile a permis de comprendre les mécanismes exacts de l'ARNi par récapitulation *in*

*vitro*. La caractérisation de l'activité du complexe RISC sur une cible d'ARN a tracé la voie des études portant sur les propriétés du domaine PIWI.

#### 4.2.3.1. Caractérisation de l'activité RISC

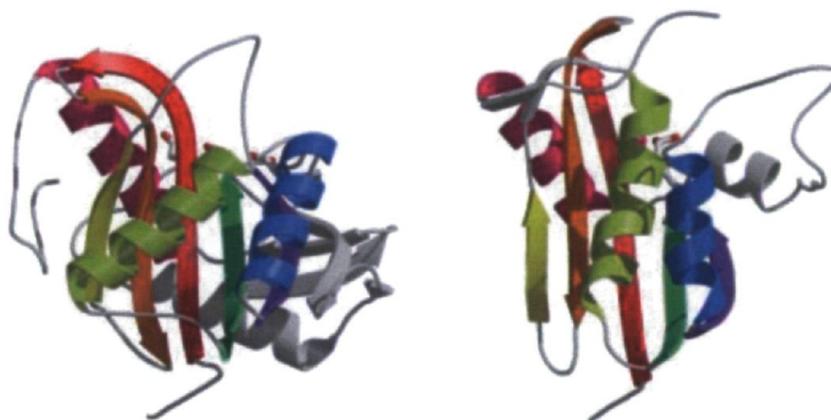
La caractérisation du complexe RISC est intimement liée à l'étude de son activité enzymatique. En premier lieu, il a été montré que l'enzyme du RISC responsable du clivage provoquait une coupure laissant des extrémités 5'P et un 3'OH, et qu'elle nécessitait un cofacteur métallique, le  $Mg^{2+}$  (Hammond et al., 2000). La réaction implique une rupture de la liaison phosphodiester de façon spécifique, et tout changement de type de liaison inhibe la réaction, comme l'utilisation d'une liaison phosphorotioate. De plus, cette étude a également mis en évidence que la coupure s'effectuait sur l'ARN cible en position 10-11 de l'ARN guide utilisé. Le lien entre cette activité et les protéines Argonaute ne s'est pas fait tout de suite. La découverte des protéines Argonaute comme étant impliquées dans l'ARN interférence s'est faite essentiellement par des études génétiques, qui montre que certains mutants perdent la capacité à médier l'ARNi. Chez *C. elegans*, un mutant de la protéine Argonaute RDE-1 montre comme phénotype principal la perte de l'exo-ARNi (Tabara et al., 1999). Les protéines ALG-1 et ALG-2 du vers, deux paralogues de RDE-1, sont ensuite montrées comme étant essentielles au développement du vers et comme intervenant dans la voie des miARN (Grishok et al., 2001). Par des méthodes de purification et l'utilisation d'anticorps spécifiques, il est montré que la protéine AGO2, initialement nommée eIF2C, est toujours présente dans les fractions contenant l'activité endonucléasique séquence-spécifique (Hammond et al., 2001).

#### 4.2.3.2. Caractérisation du domaine PIWI

La première étude d'importance a été menée par des cristallographes sur une protéine proche des Argonaute chez *Pyrococcus furiosus*. La structure du domaine PIWI a révélé un repliement similaire à une famille de ribonucléase, les RNases H (Song et al., 2004). Présent chez la bactérie *E. coli* et chez certains virus, cette famille d'enzyme utilise un brin d'ADN guide afin de rompre la liaison phosphodiester d'une molécule d'ARN (Haren et al., 1999; Song et al., 2004). Ce domaine se replie généralement en 5 feuillets  $\beta$  entrecoupés d'hélices  $\alpha$ . Le site actif de ces enzymes est composé de 3 acides aminés Asp-

Asp-Glu/Asp qui sont capables, de par leur charge négative, de coordonner deux cations métalliques bivalents. Pour les protéines Argonaute eucaryotiques, la troisième position varie quelque peu et peut exposer d'autres acides aminés (Asp/Glu/His/Lys) qui coordonnent un seul cation bivalent pour leur site actif (Tolia and Joshua-Tor, 2007).

**Figure 17 : Comparaison des repliements adoptés par le domaine PIWI des protéines Argonautes et RNases H.** Le repliement des protéines Argonaute est à gauche et celui des enzymes de types RNase H est à droite. Tiré de Song et al., 2004.



Par la suite, des études d'immunoprécipitation d'Argonaute chez des cellules humaines en culture ont permis de démontrer que Ago2 possède une activité endonucléasique dépendante d'un court ARN, et que la triade DDH forme le site actif. Si ce site actif est muté, la protéine perd la capacité de couper l'ARN cible (Liu et al., 2004). Ago3, qui possède cette triade, mais n'exerce pas cette activité, laisse suggérer que la présence de ce domaine n'est pas synonyme d'activité endonucléasique. Deux études basées sur la production de protéine Argonaute recombinante chez la drosophile et l'humain ont permis de confirmer ces résultats (Miyoshi et al., 2005; Rivas et al., 2005).

#### 4.2.4. Structures cristallines de complexes Argonaute-ARN

Les structures cristallines de complexes Argonaute-ARN ont éclairé grandement les mécanismes par lesquels la protéine Argonaute était capable de se lier avec un court ARN, puis de lier son ARN cible. Le groupe de recherche du Dr Patel, à lui seul, a publié une série de structures cristallines de protéines Argonaute complexés avec différentes molécules

d'ARN et d'ADN qui ont permis d'élaborer un modèle quant au fonctionnement de ces protéines.

#### **4.2.4.1. Avant propos**

Nous avons été invités par le journal Nature, en 2009, à introduire un article primordial dans le domaine de l'étude des protéines Argonaute, à savoir la première structure cristalline d'un complexe Argonaute-ADN guide-ARN cible. J'ai participé à la rédaction de cet article de synthèse avec le Dr Simard dans lequel nous avons décrit les principaux travaux cristallographiques ayant permis de comprendre les changements structuraux des protéines Argonaute lors de la liaison à différents substrats d'ARN.

#### **4.2.4.2. Tracing Argonaute binding**

Samir Bouasker<sup>1</sup> and Martin J. Simard<sup>1</sup>

<sup>1</sup>Laval University Cancer Research Centre, Hôtel-Dieu de Québec, 9 McMahon Street, Quebec City, Québec G1R 2J6, Canada

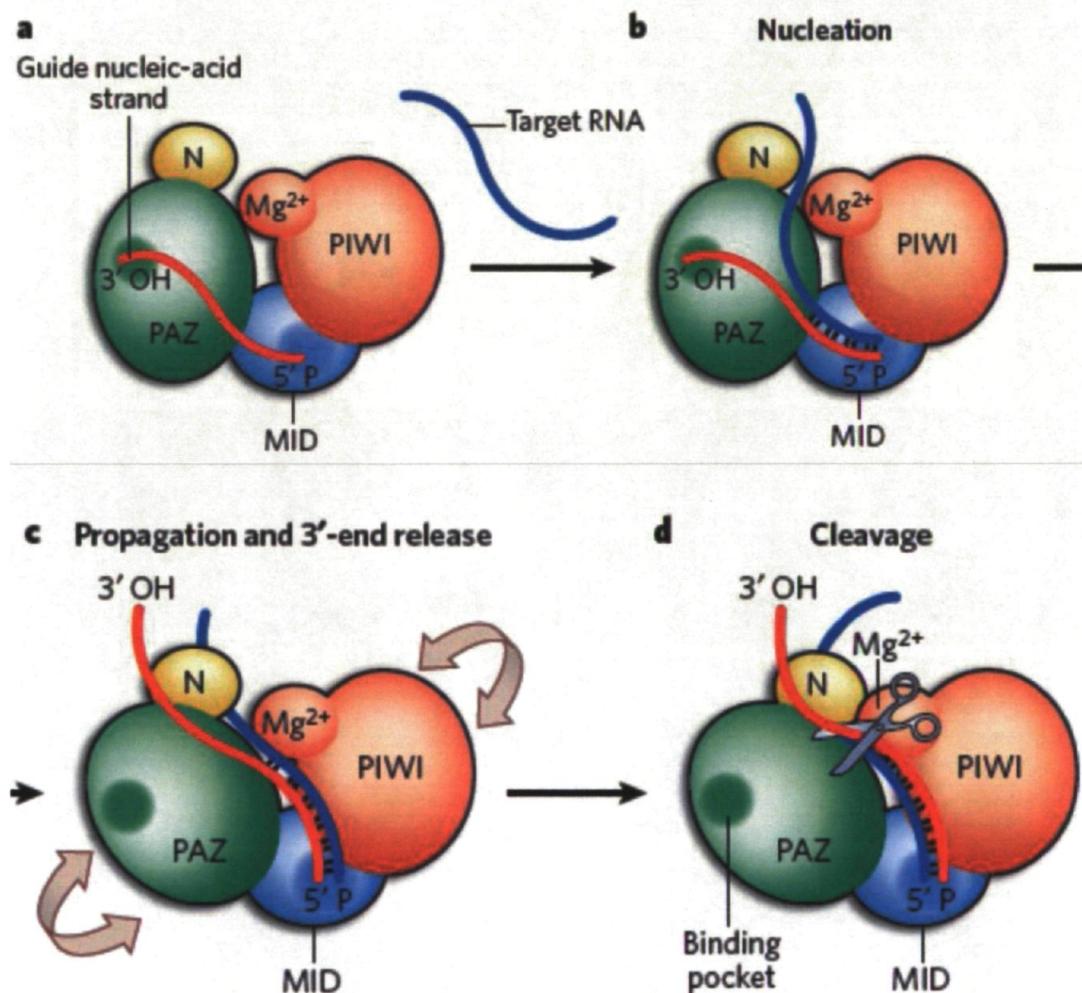
Argonaute proteins inhibit gene expression by binding to messenger RNA via a small nucleic-acid guide. Structures of the Argonaute complex bound to target RNA reveal snapshots of a silencing machine at work.

Argonaute proteins are essential regulators of gene expression, being key components of gene-silencing pathways mediated by RNA. In plants and animals, Argonaute proteins interact with small RNA molecules to form an enzymatically active complex called RISC. The RISC complex silences target genes by binding to messenger RNA that has sequence complementarity to the Argonaute-bound RNA. Some Argonaute proteins cleave mRNA ; others have lost their catalytic activity and regulate gene expression by some other means, most likely by inhibiting mRNA translation<sup>1</sup>. Although previous biochemical studies have revealed much about the molecular mechanisms of RNA silencing, exactly how Argonaute proteins interact with small RNA molecules as well as with their target mRNAs has been unclear. The paper by Wang and colleagues<sup>2</sup> from the Patel and Tuschl labs (page 754 of this issue) is therefore especially welcome — the authors report the results of a series of impressive structural studies that reveal the molecular dynamics of the Argonaute protein

as it binds to, and slices up, its RNA target. Argonaute proteins have four domains: the amino-terminal, PAZ, MID and PIWI domains (Fig. 1). The PAZ domain is implicated in the binding of single-stranded RNA, whereas the PIWI domain has endonucleolytic activity — the ability to cleave nucleic acid at internal bonds. The first crystal structures of Argonaute proteins from prokaryotes<sup>3-5</sup> (bacteria or archaea) revealed that the PIWI domain is structurally similar to members of the RNase H family of endonucleases, which use DNA as a guiding template to target RNA molecules. Although prokaryotic Argonaute proteins use DNA as the nucleic-acid template strand, plant and animal Argonaute proteins have evolved to use single-stranded RNA, rather than DNA, as a template to target RNA. The Patel and Tuschl labs recently reported<sup>6</sup> the structure of a binary complex of bacterial Argonaute bound to a 5'-phosphorylated DNA guide strand. In this work, they showed that both ends of the 21-nucleotide DNA guide strand are anchored within the Argonaute protein — the 5'-phosphate end is anchored in a binding pocket in the MID domain, and two nucleotides at the extreme 3'-hydroxyl end of the guide DNA strand are anchored in a pocket of the PAZ domain. The same two groups described<sup>7</sup> the first ternary complex of bacterial Argonaute, consisting of the Argonaute protein bound to guide DNA plus a 20-nucleotide target RNA. To solve this ternary structure, Patel, Tuschl and colleagues had to prevent Argonaute from cleaving the target RNA; they did this by introducing base mismatches between the guide DNA and the target RNA. However, this manoeuvre made it impossible to assess guide DNA–target RNA base-pairing at and beyond the mispaired cleavage site. So although this study<sup>7</sup> showed that the Argonaute complex undergoes conformational changes when it binds to target RNA, the researchers could not clearly observe the behaviour of Argonaute during the nucleation and propagation of the guide DNA–target RNA complex or during cleavage of target RNA. In their latest study, the authors<sup>2</sup> circumvent this problem by using a catalytic mutant of bacterial Argonaute, which is unable to slice up target RNA. The authors analysed several structures of the mutant Argonaute protein bound to a guide DNA that is fully base-paired to target RNA molecules of different lengths. As previously observed, both ends of the guide strand are anchored in the Argonaute binding pockets (Fig. 1a). The nucleation step begins with the binding of the guide DNA to target RNA at the DNA 5'-phosphate end (Fig. 1b). The guide DNA and target RNA form base pairs and 'zipper up', forming a DNA–RNA double helix that

extends from the 5' end of the guide DNA to DNA nucleotide 16. Pivoting of the Argonaute protein allows double-helix formation while both ends of the guide DNA are anchored in the Argonaute binding sites. Beyond nucleotide 16, Argonaute's N-terminal domain blocks additional base-pairing towards the 3'-hydroxyl end of the guide strand. The propagation of base-pairing between the guide DNA and the target RNA before this obstruction results in the release of the 3' end of the DNA from its anchor site in the PAZ domain (Fig. 1c). This release allows rotation of the PAZ domain, leading to a conformational change that favours the positioning of the cleavage site of target RNA close to the catalytic residues in the PIWI domain (Fig. 1d). The structures solved by Wang *et al.*<sup>2</sup> also clearly show that two magnesium ions that are essential for cleavage activity are located one on either side of the cleavage site. This requirement for cations to facilitate site-specific cleavage is also a feature of RNase H endonucleases, confirming that Argonaute cleavage activity is highly similar to that of RNase H enzymes. Although studies of Argonaute proteins in prokaryotes are informative, prokaryotic Argonaute has not been implicated in small RNA-mediated silencing pathways. The next big challenge will be to solve the structures of Argonaute proteins in plants and animals. It will be interesting to know whether plant and animal Argonaute proteins promote nucleation of all of the guide strand's nucleotides with the RNA target to increase binding and silencing specificity, or whether they nucleate only up to position 16, like bacterial Argonaute. Another question is whether interaction between animal microRNA (a type of small RNA encoded in the genome that is used as a guide strand) and target mRNA can be accommodated by the Argonaute protein, because microRNA typically binds imprecisely to target mRNA and forms an imperfectly paired RNA double helix. Comparing the structural features of Argonaute proteins from different organisms will help us to further understand their functions within the RNA silencing pathways and might even uncover new roles for this versatile protein family.

**Figure 18 : The Argonaute silencing complex at work.** a, Argonaute proteins have four domains: the amino-terminal domain (N), PAZ, MID and PIWI. Each Argonaute protein binds to a small nucleic-acid molecule (red; RNA in plants and animals, and DNA in bacteria), which functions as a template for binding to complementary target RNA. The 5'-phosphate (5' P) end of the guide nucleic acid is anchored in the MID domain, and the 3'-hydroxyl end (3' OH) is anchored in the PAZ domain. b, Structural studies by Wang and colleagues<sup>2</sup> reveal that when the Argonaute complex binds to target RNA, the nucleation step begins with formation of a double helix by base pairing between the guide nucleic acid and the target RNA, commencing at the 5'-phosphate end of the guide strand. c, Pivotal movement of the Argonaute protein allows extension of the double helix while the guide DNA is anchored at both ends. The 3'-hydroxyl end of the guide strand is then released from the PAZ domain, allowing its rotation. d, This conformational change favours the exact positioning of the target RNA cleavage site close to the Argonaute PIWI domain. Magnesium ions in the PIWI domain facilitate precise cleavage of the target.



1. Hutvagner, G. & Simard, M. J. *Nature Rev. Mol. Cell Biol.* 9, 22–32 (2008).
2. Wang, Y. *et al. Nature* 461, 754–761 (2009).
3. Song, J.-J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. *Science* 305, 1434–1437 (2004).
4. Yuan, Y.-R. *et al. Mol. Cell* 19, 405–419 (2005).
5. Ma, J.-B. *et al. Nature* 434, 666–670 (2005).
6. Wang, Y. *et al. Nature* 456, 209–213 (2008).
7. Wang, Y. *et al. Nature* 456, 921–926 (2008).

## 5. Contexte et hypothèse de recherche

La description des différentes voies des courts ARN montre que, pour chaque voie, des protéines Argonaute spécifiques sont impliquées. Nous avons vu que les protéines Argonaute interviennent notamment dans la formation d'un complexe ciblant les ARNm qui vont permettre l'inhibition de gène de façon post-transcriptionnelle. Leur activité endonucléasique est souvent apparue comme nécessaire pour dégrader leur cible mais, dans la voie des miARN, rares sont les études décortiquant l'implication *in vivo* de cette activité endonucléasique. Chez l'homme, quatre protéines Argonaute sont impliquées dans la voie des miARN et seule la protéine Ago2, tel que confirmée par plusieurs études, possède une activité endonucléasique. Chez la souris, la mutation du gène Ago2 est létale tout comme la mutation du motif catalytique. Un groupe a récemment découvert un miARN, miR-451, dont le processus de biogénèse Dicer-indépendant, requiert l'activité de la protéine Ago2 (Cheloufi et al., 2010; Cifuentes et al., 2010b). Chez le nématode, deux protéines orthologues des protéines Argonaute du groupe «Ago-like» humain, ALG-1 et ALG-2, assurent la fonction des miARN. Ces deux protéines conservent un motif DDH, pouvant exercer une activité endonucléasique. L'ARNi contre *alg-1* entraîne une baisse des niveaux des miARN matures *let-7* et *lin-4*, indiquant que la protéine ALG-1 participe au processus de maturation des miARN (Grishok et al., 2001).

Notre hypothèse est que toutes ces données convergent vers une participation active des protéines Argonaute dans la maturation des miARN et que cette participation requiert, outre la spécificité de la protéine, une activité endonucléasique. L'exemple de Ago2 démontre que les processus évolutifs poussants à conserver une activité enzymatique, telle qu'une activité endonucléasique, sont toujours en lien avec une fonction essentielle (Cheloufi et al., 2010; Cifuentes et al., 2010b).

Les avantages d'utiliser *C. elegans* pour étudier la participation de l'activité endonucléasique dans la maturation des miARN sont nombreux. Ce modèle nous permet d'étudier ce mécanisme à l'échelle d'un animal entier en alliant la génétique, l'ARNi, et les analyses biochimiques. De plus, ALG-1 et ALG-2 sont phylogénétiquement très proches des protéines Argonaute humaines.

## 6. Objectifs de recherche

Les objectifs pendant mon doctorat étaient les suivants :

**1-** Explorer l'activité endonucléasique de protéines Argonaute représentant des 3 sous-groupes différents (ALG-1/2, PRG-1, RDE-1 et SAGO-1). Cet objectif sera réalisé par une approche biochimique en utilisant des protéines Argonaute recombinantes produites chez *E. coli*. Nous utiliserons, pour purifier les protéines recombinantes, des étiquettes Glutathion-S-transférase (GST), déjà utilisées avec succès pour produire des protéines Argonaute de drosophile (Miyoshi et al., 2005). Les protéines recombinantes seront testées pour déterminer si elles sont douées d'une activité endonucléasique. L'étude sera portée principalement sur les protéines ALG-1/2 et la comparaison avec les autres protéines Argonaute nous permettra de les caractériser. Nous déterminerons également, par des tests de liaisons à l'ARN, les molécules ayant le plus d'affinité pour les protéines ALG-1 et ALG-2.

**2-** Explorer l'activité endonucléasique des protéines Argonautes *in vivo*. Par une approche génétique, nous allons créer des lignées transgéniques mutées au niveau de ce motif. Cela va nous permettre de déterminer le caractère essentiel ou non de l'activité dans la voie des miARN. L'analyse biochimique de ces mutants, que j'appelle mutants catalytiques, va nous permettre de déterminer la ou les étapes nécessitant cette activité.

## Chapitre 2 : Matériels et méthodes

### 1. Biologie moléculaire

#### 1.1. Création d'ADNc

Les différents ADNc ont été synthétisés avec le kit «SuperScript III First-Strand Synthesis System» (Roche). Cinq ug d'extraits d'ARN totaux de vers N2 ont été utilisés et l'amorce oligo(dT)<sub>20</sub> a été utilisé pour synthétiser tous les ADNc de ALG-1, ALG-2, RDE-1, SAGO-1 et PRG-1. Le milieu de réaction est composé de 5 µg d'ARN total, 2,5 µM d'amorce, 0,5 mM de dNTPs, dans un volume initial de 13,3 µl. Ce milieu réactionnel a été dénaturé à 65°C pendant 5 min et refroidi sur la glace pendant 5 min. Ensuite, le tampon RT 1X, ainsi que 10 mM de DTT, 1 U/µl de RNaseOut et, finalement, 1 µl de Superscript III ont été ajoutés, puis mélangés. Par la suite, la réaction a été centrifugée brièvement et incubée à 50°C pendant une heure. L'enzyme Superscript III a ensuite été inactivée à 85°C pendant 5 min. Finalement, 1 µl de RNase H a été ajouté par échantillon et la réaction a été incubée à 37°C pendant 20 min. Les ADNc ainsi obtenus sont conservés à -20°C.

#### 1.2. Clonage

##### 1.2.1. Plasmides et amorces utilisés

Les ADNc de alg-1, alg-2, rde-1, sago-1 et prg-1 ont été clonées dans les plasmides pGEX 6P-1 ou pGEX 5X3.

Les réactions de PCR ont été effectuées avec 2 µL d'ADNc et l'ADN polymérase utilisée est la KOD Hot Start (Novagen). Les conditions d'amplification sont les suivantes : 5µl de tampon 10X PCR, 3 µL de MgCl<sub>2</sub> 25 mM, 5 µl dNTP 8 mM, 0,15 µl oligo sens 100 µM, 0,15 µl oligo antisens 100 µM, 1 µl Taq, 2 µl d'ADNc et de l'eau jusqu'à un volume final de 50 µl.

Les réactions de mutagénèses dirigées : la même enzyme et les mêmes conditions ont été utilisées. Après la réaction de PCR, le temps d'élongation est de 9 min. La digestion est

effectuée avec 2  $\mu$ l d'enzyme Dpn I (New England Biolabs) pendant 1 h 30 à 37°C, suivi de la transformation bactérienne dans une souche DH5 $\alpha$ .

Les réactions de mutagénèses par délétions : la même enzyme et les mêmes conditions ont été utilisées. Le temps d'élongation de la PCR est de 9 min. On effectue par la suite une réaction de ligation avec la T4 DNA ligase (New England Biolabs) pendant 16 h à 16°C, suivi de la transformation bactérienne dans une souche DH5 $\alpha$ .

### **1.2.2. Kits utilisés**

Purification des fragments de PCR : le kit utilisé est le EZ-10 Spin Column PCR products purification Kit (Biobasics).

Purification des produits de PCR digérés : après migration du produit sur gel d'agarose, la bande est découpée sous UV et l'ADN est élué du gel à l'aide du kit EZ-10 Spin Column DNA gel Extraction Kit (Biobasics).

Digestion enzymatique des produits de PCR et des plasmides : les enzymes de restrictions utilisés proviennent toutes de la compagnie New England Biolab.

### **1.2.3. Transformation bactérienne**

Pour les clonages, 100  $\mu$ l de bactéries DH5 $\alpha$  compétentes sont ajoutées à 50-100 ng de plasmide et incubé 15 min sur glace. Après un choc thermique de 1,5 min à 42°C, 500  $\mu$ l de LB (1% (p/v) bacto-peptone, 1% (p/v) chlorure de sodium et 0,5% (p/v) extrait de levures) sont ajoutés et les bactéries sont régénérées une heure à 37°C sous agitation. Environ 100  $\mu$ l sont étalés sur un milieu LB-agar (LB additionné de 1,7% (p/v) d'agar) contenant de l'ampicilline (50  $\mu$ g/ml).

BL21 codon + : la même stratégie est utilisé mais les antibiotiques sont différents. LB-Agar avec ampicilline (50  $\mu$ g/ml) et chloramphenicol (100 $\mu$ g/ml).

**Tableau 2 : Amorces utilisées pour les différentes constructions.** a) clonage. b) mutation DDH. c) délétion des domaines PAZ, C-terminal et insertion de G.

**a**

| ADNc(plasmide) | Numéro amorce | Séquence   |
|----------------|---------------|--|
| rde-1 (5x3)    | mso833        | 5'-ATC TAT GGA TCC TA ATGCTCTCGAATTTCCCGAATTG    |
|                | mso829        | 3'-AATGACGATATCTTATGCGAACGACATTCAGGGTAC          |
| sago-1 (6p1)   | mso574        | 5'-TTTGGATCCATGTCCAATATCACCCAAGTCAC              |
|                | mso575        | 3'-AGATGCGGCCGCTCATTATGCATTGAATCTCGTACCGAA       |
| prg-1 (6p1)    | mso572        | 5'-TTTGGATCCATGGCATCTGGAAGTGGTCG                 |
|                | mso573        | 3'-AGATGCGGCCGCTCATTACAAGAAGAAGACAGCTTGTCCAG     |
| alg-1 (6p1)    | mso568        | 5'-TTTGGATCCATGTCCGGCCGGCCGAATAT                 |
|                | mso569        | 3'-AGATGCGGCCGCTCATTAAAGCAAAGTACATGACGTTGGCATCCG |
| alg-2 (6p1)    | mso565        | 5'-TGTATTGTCGACCCATGTCCCTCTGCTGTACACAAT          |
|                | mso566        | 3'-AGATGCGGCCGCTCATTAGGCAAATACATGACGTTGTTCCAGTC  |

**b**

| ADNc      | Numéro amorce | Séquence                             |
|-----------|---------------|--------------------------------------|
| alg-1 adh | mso412        | 5'-ATTTCTTTGGATGTGCTATCACTCATCCACCA  |
|           | mso413        | 3'-TGGTGGATGAGTGATAGCACATCCAAAGAAAAT |
| alg-1 dah | mso414        | 5'-CGTTGTCTACCGTGCCGGAGTTTCTGAGG     |
|           | mso415        | 3'-CCTCAGAACTCCGGCACGGTAGACAACG      |
| alg-1 dda | mso416        | 5'-AGCATATTATGCTGCTTGGTAGCGTTCC      |
|           | mso417        | 3'-GGAACGCTACCAAAGCAGCATAATATGCT     |
| alg-2 adh | mso418        | 5'-TCCTGGTTGCGCTATCACCCATCC          |
|           | mso419        | 3'-GGATGGGTGATAGCGCAACCCAGGA         |
| alg-2 dah | mso420        | 5'-TGTGGTCTACAGAGCCGGAGTGTCTGAGG     |
|           | mso421        | 3'-CCTCAGACTCCGGCTCTGTAGACCACA       |
| alg-2 dda | mso422        | 5'-GCGTACTATGCACTCTGGTGGCGTTC        |
|           | mso423        | 3'-GAACGCCACAGAGCTGCATAGTACGC        |
| rde-1 adh | mso238        | 5'-TATGTTGGAATTGCTGTAACCTACCAA       |
|           | mso239        | 3'-TTGGATGAGTTACAGCAATCCAACATA       |
| rde-1 dah | mso240        | 5'-GTAGTCTATCGAGCCGGAGTTAGCGAT       |
|           | mso241        | 3'-ATCGCTAACCCGCTCGATAGACTAC         |
| rde-1 dda | mso242        | 5'-CCGGTTCATTATGCTGCTTTATCATGTGAA    |
|           | mso243        | 3'-TTCACATGATAAAGCAGCATAATGAACCGG    |
| prg-1 dah | mso1047       | 5'-TTGATCCTCTACAGAGCCGGTGTGCGGACGGA  |
|           | mso1048       | 3'-TCCGTCGCCAGCACCGGCTCTGTAGAGGATCAA |

**c**

| ADNc         | Numéro amorce | Séquence                                |
|--------------|---------------|---|
| alg-1 g      | mso1347       | 5'-GTTCTCCCTGAAAAGTACTCCAGTTATGCTGAA    |
|              | mso1348       | 3'-TTCAGCATAAAGTGGAGTACCTTTCCAGGGAGAAC  |
| alg-2 g      | mso1349       | 5'-GTGCTGCCAGGAAAGGGTACTCCGATCTATGCCGAA |
|              | mso1350       | 3'-TTCGGCATAGATCGGAGTACCCTTTCTGGCAGCAC  |
| alg-1 Δ PAZ  | mso708        | 5'-GCTCCAGAACGTGAGCGTGAG                |
|              | mso707        | 3'-GGAACGATAGAAAAGCAGTCGAGAG            |
| alg-2 Δ PAZ  | mso712        | 5'-GCTCCAGAGCGTGAGCGTGAG                |
|              | mso711        | 3'-CGCGCGTAAAAGGCCGT                    |
| alg-1 Δ Cter | mso954        | 5'-GCTCGTTATCATCTTTAAGATCGCGAGCATGAC    |
|              | mso955        | 3'-GTCATGCTCGCGATCTTAAAGATGATAACGAGC    |
| alg-2 Δ Cter | mso1234       | 5'-GGATCATGGATCCGGTAAGGAAGGATCACAACC    |
|              | mso1235       | 3'-GGTTGTGATCTTCTTACCAGGATCCATGATCC     |

### 1.3. Purification de protéines

Culture : Ensemencement avec une colonie de bactérie BL21 codon plus transformée dans 4 ml de LB additionné d'ampicilline (50 µg/mL) et de chloramphenicol (100 µg/ml). Culture 16 h à 37°C puis 2 ml sont utilisés pour ensemercer 1 L de milieu avec la même

concentration d'antibiotiques. Culture à 37°C puis à DO=0,4, 1 ml d'IPTG à 1 M est additionné au milieu. Après deux heures de cultures, les bactéries sont récupérées par centrifugation en utilisant des bouteilles à centrifuger de 500 ml. Les bactéries sont récoltés avec 25 ml de LB puis sont centrifugés dans un falcon 50 ml et le culot est gardé à - 80°C jusqu' à utilisation.

Purification : le culot est récupéré dans 30 ml de PBS 1X froid additionné de cocktail d'inhibiteurs de protéases (complete, EDTA-free Protease Inhibitor Cocktail Tablets, Roche) et de 100 mM de PMSF. La solution est lysée dans un appareil de broyage Constant Cell Disruption System de Constant Systems à 27,5 KPSI. Le lysat est récupéré est additionné de détergeant Empigen BB (Sigma-Aldrich) pour une concentration finale de 0,25% et de 10 ml de PBS 1X froid. Après 1 h 30 sous agitation à 4°C, le lysat est centrifugé à 15000 g durant 15 min. Le lysat est ensuite directement utilisé pour l'étape de purification.

Les billes utilisées pour la purification sont les billes Glutathione Sepharose 4B (GE Amersham). Après équilibrage des billes, le lysat additionné de bille (équivalent de 665 µL de bille d'origine pour 1 L de culture) et incubé au minimum 8 h à 4°C. Après 3 lavages de 10 ml de PBS froid, les billes sont incubées dans 4 ml de tampon d'élution pendant au minimum 4 h. L'éluat est récupéré puis passer sur colonne Amicon de 100 kDa de Millipore qui est lavé 2 fois avec le PBS froid. Le volume retenu est récupéré et les parois des membranes sont rincées avec 200 µL de PBS froid. Ces deux extraits sont analysés sur gel de polyacrylamide, coloré au bleu de Coomassie et la concentration en protéine est estimée par dosage en utilisant le Dc Protein Assay (Biorad).

Tampon d'élution : 10 mM de glutathion, 50 mM Tris-HCl pH 8, 10% glycérol, 5 mM MgAc additionné de cocktail d'inhibiteur de protéases (Complete, EDTA-free Protease Inhibitor Cocktail Tablets, Roche)

## 1.4. Tests de clivage et «Filter Binding Assay»

### 1.4.1. Radiomarquage des ARN de 21 nt ou 32 nt en 5'

Un  $\mu\text{l}$  de T4 polynucléotide kinase (NEB), 1  $\mu\text{l}$  de tampon 10 X, 5  $\mu\text{l}$  d'ARN à 100  $\mu\text{M}$  et 3  $\mu\text{l}$  de gamma-32P-ATP. Incubation à 37°C durant 1 h 30. Après ajout de 40  $\mu\text{l}$  d'eau, l'ARN radiomarqué est purifié sur colonne Quick Spin Columns for radiolabeled RNA purification Sephadex G-25 (Roche), puis l'éluat est migré sur gel d'ARN dénaturant Séquagel (National Diagnostics). Après exposition sur un film autoradiographique (Fujifilm), l'ARN radiomarqué est découpé et élué dans du tampon Proteinase K (PK) 2X pendant au moins 2 h. L'éluat est traité au phénol chloroforme pour éliminer les protéines, puis les molécules d'ARN sont précipitées à l'éthanol 100% à -20°C. Le culot d'ARN est récupéré dans un volume de 50  $\mu\text{L}$  d'eau et dosé par compteur à scintillation. 2500 à 5000 cpm sont suffisants pour les réactions de clivage *in vitro*.

Pour la formation des duplexes : une quantité équimolaire de brin sens radiomarqué et de brin antisens phosphorylé en 5' est chauffé à 90°C pendant 15 min, puis refroidit lentement (par extinction du bain marie et attente de 30 min). Une fraction de la réaction est chargée sur un gel non dénaturant de polyacrylamide 15% Sequagel en même temps qu'un contrôle simple brin pour vérifier la formation de duplexes.

Tampon PK 2X: 200 mM Tris-Cl, pH 7.5, 25 mM EDTA, pH 8.0, 300 mM de NaCl et 2% (p/v) sodium dodecyl sulfate.

### 1.4.2. Réaction de clivage *in vitro* et «Filter Binding Assay»

Réaction de clivage *in vitro* : 30 mM HEPES pH 7,4, 40 mM acétate de potassium, 5 mM acétate de magnésium, 5 mM DTT, 0,5  $\mu\text{g}$  d'ARN de levure, 40 U de RNAsine, 1  $\mu\text{g}$  de BSA acétylée par réaction, 100 nM d'ARNsb ou de duplex. Différentes concentrations de protéines ont été utilisées mais la concentration des autres éléments ne varient pas. 2000 à 5000 cpm d'ARN radiomarqué sont nécessaire.

Clivage ARN cible-ARNsb et relargage : le mélange réactionnel est incubé 1 h 30 à 20°C puis on ajoute l'ARN radiomarqué, suivi d'une seconde incubation d'une durée de 1 h 30 à

20°C. Le milieu réactionnel est digéré ensuite dans 100 µl de PK 2X, 20 µg de glycogène et 50 µg de PK. Puis, on isole l'ARN par traitement au phénol/chloroforme et on le précipite à l'éthanol. On visualise le résultat par migration sur un gel de polyacrylamide dénaturant (Séquagel).

Clivage de duplexe : incubation pendant 2 h à 20°C. Le milieu réactionnel est digéré ensuite dans un 100 µl de PK 2X, 5 µL de PK additionné de glycogène. Puis, on isole l'ARN par traitement au phénol/chloroforme et on le précipite à l'éthanol. On visualise le résultat par migration sur un gel de polyacrylamide dénaturant (Séquagel).

Courbe de liaison par «Filter Binding Assay» : le même milieu réactionnel que pour la réaction de clivage est utilisée. Un appareil BioDot (Bio-Rad) est requis pour cette expérience ainsi que deux types de membranes. Une membrane de nitrocellulose (Trans-Blot de 0,45 µM de Bio-Rad) va retenir les protéines tandis que la seconde membrane (GE Amersham Hybond-N+) va retenir les acides nucléiques libres. 20000 cpm d'ARN radiomarqué sont utilisés. Après une exposition sur un écran phosphoimager, l'intensité de la radioactivité retenue est déterminée par le logiciel Multigauge (Fujifilm).

## **2. Manipulation de *C. elegans***

### **2.1. Création de lignées transgéniques**

Les plasmides d'ADN génomiques pBS-*alg-1* et pBS-GFP::*alg-2* étaient déjà disponibles au laboratoire. La mutagenèse dirigée a été faite comme précédemment décrit en utilisant les amorces que pour les ADNs de *alg-1* et *alg-2*. Pour les lignées transgéniques, les marqueurs *rol-6* et *m-cherry* ont été utilisés. L'intégration a été faite par la méthode standard aux UV selon Mello et Fire (1995).

### **2.2. Extraction d'ARN total**

D'un culot de vers, l'ARN est extrait en ajoutant quatre volumes de TRI-Reagent (Sigma) par volume de nématodes et mélangé vigoureusement pendant 30 s. Le mélange est plongé dans l'azote liquide pendant 30 s, puis décongelé dans un bain à 37°C. Cette étape est répétée 3 fois. On rajoute ensuite 1 ml de TRI-Reagent et on mélange vigoureusement, puis 1 ml de chloroforme est ajouté et on mélange. On incube à température pièce pendant 2 min. On centrifuge à 12 000 g pendant 15 min à 4°C. La phase aqueuse est isolée et un volume égal d'isopropanol est ajouté, puis mélangé et incubé à température pièce pendant 10 min. On centrifuge à 12 000 g pendant 10 min à 4°C. Le culot est lavé avec de l'éthanol 75%, centrifugé à 7 500 g pendant 5 min à 4°C. Le culot est séché à l'air libre et resuspendu dans de l'eau «RNase-free». La concentration et la qualité de l'ARN ont été déterminées par spectrométrie à 260 nm, par le ratio 260/280 nm et par migration sur gel d'agarose. Les échantillons sont conservés -80°C.

### **2.3. Analyse de type Northern**

Le protocole provient du laboratoire du Dr Bartel (Voir protocole détaillé sur le site web du laboratoire du Dr David Bartel). Dix µg d'ARN sont chargés sur gel 15% de polyacrylamide dénaturant (Séquagel). Après transfert sur une membrane (Genescreen hybridization membrane) pendant 45 min à 20 V, l'ARN est fixé sur la membrane par chauffage à 60°C dans une solution de méthylimidazole et d'EDC (1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride) pendant 2 h. Après rinçage de la

membrane, l'incubation se fait dans un tampon de pré-hybridation classique. La sonde est synthétisée par le système Starfire (Integrated DnaTechnologies) et la polymérase Klenow (New England Biolabs). Après une incubation d'au moins 2 h dans le tampon de pré-hybridation, la moitié de la sonde synthétisée est ajoutée après chauffage pendant 5 min à 95°C. Après une incubation qui ne doit jamais excéder 16 h, trois rinçages de 10 min avec une solution de lavage non-stringent et un dernier rinçage stringent sont effectués (on utilise un volume de pré-hybridation par rinçage). On expose ensuite la membrane à une plaque phosphorimager (Fujifilm), de 1 à 3 jours, dépendamment de l'abondance du miARN.

Liste des amorces Starfire utilisés :

Let-7 : 5'-AACTATAACAACCTACTACCTCA/Starfire/-3'  
 Let-7 \* : 5'-GTAAGGTAGAAAATTGCATAGTTCA/Starfire/-3'  
 Lin-4 : 5'-TCACACTTGAGGTCTCAGGGA/Starfire/-3'  
 Lin-4 \* : 5'-GTACCCGGAGAGCCCAGGTGT/Starfire/-3'  
 miR-50 : 5'-AACCCAAGAATACCAGACATATCA/Starfire/-3'  
 miR-50 \* : 5'-GTCGATACGTCTAATATGCGGG/Starfire/-3'  
 miR-40 : 5'-TTAGCTGATGTACACCCGGTGA/Starfire/-3'  
 miR-40 \* : 5'-TATCATCATGGCATAACATCCACT/Starfire/-3'  
 miR-90 : 5'-AGGGGCATTCAAACAACATATCA/Starfire/-3'  
 miR-90\* : 5'-GTTGATATCGTCGTTGAAAGCCG/Starfire/-3'  
 miR-62 : TGA TAT GTA ATC TAG CTT ACA G/Starfire/-3'

Solution Pré-hybridation

| <u>Final:</u>                                 | <u>Start:</u> | <u>50 ml</u> |
|---|---------------|--------------|
| 5X SSC  | 20X           | 12.5 ml      |
| 20 mM Na <sub>2</sub> HPO <sub>4</sub> pH 7.2 | 1M            | 1 ml         |
| 7% SDS  | 10%           | 35 ml        |
| 2X Denhardt's Solution                        | 100X          | 1.5 ml       |

Solution Non-Stringent

| <u>Final:</u>                                 | <u>Start:</u> | <u>200 ml</u> |
|---|---------------|---------------|
| 3X SSC  | 20X           | 30 ml         |
| 25 mM NaH <sub>2</sub> PO <sub>4</sub> pH 7.5 | 1M            | 5 ml          |
| 5% SDS  | 10%           | 100 ml        |
| H <sub>2</sub> O distillé                     |               | 65 ml         |

### Solution Stringent

|                   |               |               |
|-------------------|---------------|---------------|
| <u>Final:</u>     | <u>Start:</u> | <u>200 ml</u> |
| 1X SSC            | 20X           | 10 ml         |
| 1% SDS            | 10%           | 20 ml         |
| dH <sub>2</sub> O |               | 170 ml        |

## 2.4. Clonage de courts ARN par 3' et 5' RACE :

Les petits ARN ont été sélectionnés par le Kit Mirvana selon le protocole du fabricant (Life Technologies). Après purification, les petits ARN (taille inférieure à 200 pb) ont été soumis, soit à une ligation en 3' (3' RACE) ou en 5' (5' RACE) en utilisant un ARN phosphorylé ou non phosphorylé (ARN adaptateur), respectivement. Les deux produits ligués, ont été rétrotranscrit en utilisant soit un oligonucléotide complémentaire à l'adaptateur ARN (3' RACE) ou un oligonucléotide complémentaire à l'extrémité 3' du précurseur de miR-50 ou miR-90 (5' RACE). Les ADNc générés ont été soumis à deux réactions de PCR et les produits amplifiés ont été sélectionnés par leur taille, isolés et clonés dans le vecteur TOPO (Life Technologies) pour le séquençage.

### 3' RACE :

ARNad : 5'-PCGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUA

RT : Primer RNA A: 5'-TCTCCTTCGTCCTGTCGTGTC-3'

PCR-1 : primer RNA-A and primer 50: TGATATGTCTGGTATTCTTGGGTT

PCR-2 : primer RNA-B GGACACTGACATGGACTGAAGGAGTA and primer 50

### 5' RACE:

RNA adaptor : 5'-CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUA

RT : Primer 50 \*: 5'-GTCGATACGTCTAATATGCGGG-3'

PCR-1 : primer RNA-A and primer 50 \*

PCR-2 : primer RNA-B and primer 50 \*

## 2.5. RT-PCR

Les réactions de PCR ont été faites à partir de 1 µl d'ADNc (comme précédemment décrit sur des extraits d'ARN total). La réaction de PCR est une réaction classique additionnée de réactif SYBR green I (Roche). L'appareil utilisé est un appareil thermocycler (Applied Biosystems).

Amorce RT-PCR :

lin-14

msol1626 : 5' TGCAAATCTTCCAATCAAAGG

msol1627 : 3' TTCTGCCTGAGCCTCTTCTC

lin-28

msol1624 : 5' GTGTTATACCTGTGGATCCGAG

msol1625 : 3' CGTCAGATGTAGTCGGTGATG

lin-41

msol1620 : 5' GGATTGTTTCGACACCAACG

msol1621 : 3' ACCATGATGTCAAACCTGCTGTC

daf-12

msol1618 : 5' GATCCTCCGATGAACGAAAA

msol1619 : 3' CTCTTCGGCTTCACCAGAAC

hbl-1

msol1622 : 5' ACTGCACATATGCCACCAAA

msol1623 : 3' TGATGTAACCGGCTCAACTG

tba-1

msol1164 : 5' GTACACTCCACTGATCTCTGCTGACAAG

msol1165 : 3' CTCTGTACAAGAGGCAAACAGCCATG

## 2.6. Extraction de protéines totales

Solution de lyse : 100 mM Acétate de K, 30 mM Hepes K-OH pH 7,4, 8 mM Acétate deMg, 1 M DTT et 50 µl Triton X-100, additionné d'inhibiteur de protéases («Complete, EDTA-free Protease Inhibitor Cocktail Tablets», Roche) et Supersine (Ambion) pour les immunoprécipitations.

Le culot de vers est lavé trois fois avec trois volumes de solution de lyse, puis récupéré dans un volume de culot. La solution est passée dans un appareil de broyage «Constant Cell Disruption System» (Constant Systems) à 7 KPSI. Le lysat est ensuite centrifugé à 17000 g pendant 15 min à 4°C, et on obtient l'extrait protéique qui correspond au surnageant. La concentration en protéines est déterminée par dosage en utilisant le «Dc Protein Assay» (Bio-Rad).

## 2.7. Identification de miRISC par oligonucléotides 2'-O-méthylés

Pour les expériences d'isolement du miRISC, des billes Streptavidine M-280 (Invitrogen), avec un support magnétique, sont utilisés. L'expérience se déroule comme ceci. La première étape consiste à la liaison des différents oligonucléotides aux billes de Streptavidine : pour un extrait 60 µl de billes sont incubés avec 120 µl d'oligonucléotides-contrôles ou d'oligonucléotides-miARN à 1µM, à température pièce et sous agitation, pendant 45 min. Après rinçage des billes, 4 mg d'extraits de protéines sont ajoutés à 30 µl de billes/oligonucléotides-contrôles et sont incubés pendant 45 min. On récupère le mélange dans deux tubes différents et après 3 rincages en utilisant le tampon de lyse (300 µl), on ajoute 30 µl de billes/oligonucléotides-contrôles dans une moitié et 30 µl de billes/oligonucléotides-miARN dans l'autre. On ré-incube 45 min à température pièce sous agitation. Après 3 rincages, on récupère les billes dans 30 µl de Tampon 2X Laemmli, on chauffe à 95°C pendant 10 min. Après centrifugation brève, on charge sur un gel SDS-PAGE le surnageant, avec comme contrôle 10% de l'extrait total utilisé. Une analyse de type Western est ensuite effectuée en utilisant l'anticorps anti-ALG-1 ou GFP (anti-AFP de QBiogene).

Oligonucléotides 2'-O-méthylés utilisés :

oligo ctl : 5'-bio-CAUCACGUACGCGGAAUACUUCGAAAUGUC-3'

let-7 : 5'-bio-GTAAGGTAGAAAATTGCATAGTTCA-3'

lin-4 : 5'-bio-GTACCCGGAGAGCCCAGGTGT-3'

## 2.8. Immunoprécipitation de protéines Argonaute

Quatre mg d'extraits de protéines totales sont incubées avec 10 µg d'anticorps anti-ALG-1 couplé à des billes protéines G agarose (Roche) ou des billes protéines G agarose couplé à un anticorps anti-GFP (Chromotek), pendant 1 h à 4°C. Les billes sont lavées 3 fois avec le tampon de lyse et récupéré dans 50 µl de tampon. Une fraction de 10% du culot sont analysés par immunobuvardage de type Western pour apprécier la qualité de l'IP. Pour l'analyse Northern, le 90% restant subi une extraction d'ARN par la PK (20 µg de

glycogène et 50  $\mu\text{g}$  de PK) pendant 45 minutes à 37°C. On ajoute ensuite 200  $\mu\text{l}$ , on récupère la phase aqueuse par phénol-chloroforme et on précipite l'ARN à l'éthanol 100%. Le culot est récupéré dans 20  $\mu\text{l}$  d'eau.

## Chapitre 3 : Caractérisation de l'activité des protéines Argonaute *in vitro*

Pour caractériser l'activité endonucléasique des protéines Argonaute, nous avons choisi une approche de protéines recombinantes produites chez la bactérie. Nous avons choisi de produire ces protéines chez la bactérie pour plusieurs raisons : 1- ce système permettra de déterminer avec précision le rôle *in vitro* de la protéine Argonaute seule, car les voies de l'ARNi sont absentes chez la bactérie (miARN, siARN, piARN) et nous serons en mesure d'éliminer les interactions avec de courts ARN ou des partenaires potentiels; 2- ce système a déjà été utilisé avec succès pour investiguer l'activité endonucléasique des protéines AGO1 et AGO2 chez la drosophile (Miyoshi et al., 2005). Nous nous sommes basés sur cette étude pour mettre au point un moyen de production de ces protéines étiquetées GST. La première étape fut de concevoir une méthode de production de protéines Argonaute. Nous avons vu dans l'introduction qu'il existe 26 protéines Argonaute chez le nématode *C. elegans* et pour démontrer l'activité de certaines protéines Argonaute, j'ai décidé d'utiliser PRG-1, RDE-1, SAGO-1 et ALG-1/2, qui sont représentatives des différents sous-groupes de protéines Argonaute chez le nématode. J'ai donc cloné les différents ADNc dans le vecteur d'expression pGEX. J'ai également muté les acides aminés responsables de l'activité endonucléasique afin de pouvoir confirmer que, *in vitro*, ces acides aminés étaient requis pour l'activité endonucléasique. Après de nombreux efforts et différents protocoles de purifications testés, j'ai finalement réussi à obtenir ces protéines (voir matériels et méthodes). Plusieurs de ces résultats ont été utilisés pour la rédaction de manuscrits. Concernant l'étude de l'activité de PRG-1, j'ai participé à un article publié dans la revue Science et intitulé «Function, targets and evolution of *Caenorhabditis elegans* piRNAs». Mon projet de recherche principal porte sur l'étude des protéines ALG-1 et ALG-2 dans la voie des miARN et fait l'objet d'un manuscrit en seconde révision dans le journal «Nucleic Acids Research» et intitulé «The slicing activity of microRNA-specific Argonautes is essential for the miRNA pathway in *C. elegans*» (en annexe 1).

# 1. Détermination de l'activité endonucléasique des Argonaute possédant le motif catalytique

D'après la littérature, le clivage par les protéines Argonaute se caractérise par plusieurs caractéristiques (Hammond et al., 2000). En premier lieu, l'activité endonucléasique nécessite la participation d'une molécule d'ARN simple brin (ARNsb) guide, qui va se lier par complémentarité parfaite à une cible d'ARN. Deuxièmement, le clivage aura lieu entre la position 10-11 sur la cible en comptant à partir du 5' de l'ARN guide; la protéine Argonaute utilisant l'ARN guide comme une règle lui permet de couper sa cible en fonction de celui-ci. Troisièmement, cette activité requiert du  $Mg^{2+}$  comme cofacteur, qui formera, avec la triade DDH, le site actif (Liu et al., 2004). J'ai donc commencé par vérifier ces points. Pour vérifier cette activité, nous avons utilisé comme cible une molécule d'ARNsb ayant les caractéristiques des molécules de courts ARN, c'est-à-dire une extrémité 5'P et 3'OH ainsi qu'une longueur de 21 nt (Fig.19b). Nous avons utilisé comme cible une molécule d'ARN de 32 nt radiomarquée en 5'. Nous avons effectué un essai de clivage en utilisant les conditions standards utilisées pour les protéines Argonaute de drosophile en présence de  $Mg^{2+}$  (Miyoshi et al., 2005) en modifiant la température pour 20°C, qui permet la croissance optimale de *C. elegans*. La réaction se déroule en deux temps : 1- incubation de l'ARNsb guide avec la protéine Argonaute dans le milieu réactionnel additionné de  $Mg^{2+}$  pendant 1 h 30. Cette incubation va permettre le chargement de l'ARN guide sur la protéine Argonaute; 2- ajout de l'ARN cible radiomarqué pendant 1 h 30. Le clivage donnera naissance à une molécule d'ARN cible plus courte d'une longueur de 15 nt qui sera détectable sur un gel de polyacrylamide dénaturant.

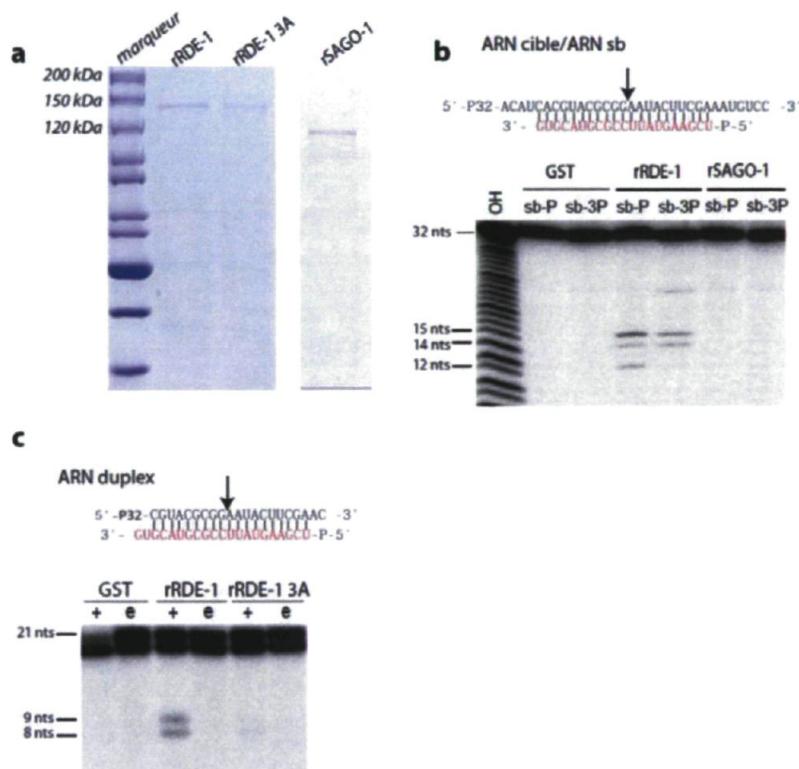
## 1.1. Détermination de l'activité endonucléasique potentielle de rRDE-1 et rSAGO-1

Pour explorer l'activité endonucléasique des protéines Argonaute, ils étaient importants de tester les protéines RDE-1 et SAGO-1 du groupe 3, spécifique à *C. elegans* pour plusieurs raisons. *In vivo*, il a été démontré que RDE-1 est responsable du clivage du duplex de siARN primaire dans la voie de l'exo-ARNi (Steiner et al., 2009). Quant à SAGO-1, elle ne

conserve pas le motif catalytique DDH (Yigit et al., 2006) et représentera un excellent contrôle négatif si elle ne démontre pas d'activité. Cette dernière protéine a la particularité de prendre en charge des siARN qui sont tri-phosphorylés.

Nous avons produit les protéines étiquetées GST-RDE-1 et GST-SAGO-1 (Fig.19a). Par souci de simplification, les protéines étiquetées GST seront nommées r (par exemple la protéine GST-RDE-1 est nommée rRDE-1). Étant donné que la substitution des acides aminés DDH par 3 alanines (A) a été démontrée comme suffisante pour abolir l'activité catalytique des protéines Argonaute (Liu et al., 2004), nous avons également produit la protéine mutante rRDE-1 3A (Fig.19a). Nous avons testé leur activité en utilisant deux ARNs<sub>b</sub> guide : l'un monophosphorylé (ARNs<sub>b</sub>-P) et l'autre triphosphorylé (ARNs<sub>b</sub>-3P). Ces deux molécules sont identiques en termes de séquence, mais diffèrent par le nombre de phosphates à leurs extrémités 5'. Le clivage spécifique de l'ARN cible conduirait à une molécule de 15 nt (Fig.19b). Pour les deux types d'ARNs<sub>b</sub> guide testés, rRDE-1 est doué d'une activité endonucléasique (Fig.19b). Pour l'ARNs<sub>b</sub>-P, nous voyons 3 produits, dont l'un majoritaire, représentant le produit de clivage en position 10-11 de l'ARNs<sub>b</sub>-P. Ce produit est également présent avec l'ARNs<sub>b</sub>-3P, ce qui indique que rRDE-1 est capable de prendre en charge ce type de molécule et que le tri-P n'a pas d'effet sur l'efficacité de clivage du substrat. Les produits de 14 et 12 nt correspondent à des produits inattendus qui nécessitent l'activité des protéines Argonautes. Pour rSAGO-1, que ce soit avec l'ARNs<sub>b</sub>-P ou l'ARNs<sub>b</sub>-3P, nous ne détectons pas d'activité dans les conditions utilisées. Nous avons ensuite examiné si la protéine rRDE-1 peut récapituler ce qui a préalablement été démontré *in vivo*, à savoir le clivage d'un duplex de 21 nt parfaitement apparié (Steiner et al., 2009). *In vitro*, rRDE-1 clive le brin passager du duplex et deux produits de 9 et 8 nt sont visibles (Fig.19c), le produit de 9 nt correspond au produit attendu. Lorsque l'on séquestre les ions Mg<sup>2+</sup> du milieu réactionnel par ajout d'EDTA, on perd l'activité de la protéine de même que la mutation du motif catalytique de rRDE-1 affecte le clivage (Fig.19c). Cependant, un produit demeure visible. La présence du motif DDH et du Mg<sup>2+</sup> sont nécessaires pour l'activité des protéines Argonaute et, la mutation de ce site, la séquestration des ions Mg<sup>2+</sup> ou le test d'une protéine ne conservant pas ce motif (rSAGO-1) le démontre. Le produit majoritaire demeure de la taille attendue que ce soit pour le clivage d'un ARN cible (Fig.19b) ou le clivage d'une molécule d'ARN duplex (Fig.19c).

**Figure 19 : Activité endonucléasique des protéines rRDE-1 et rSAGO-1.** a) Gel SDS-PAGE 8% avec 0,5 à 1 µg de protéines recombinantes révélées par coloration au bleu de Coomassie. b) Séquences des molécules utilisées et essai de clivage avec 100 nM de GST, rRDE-1 et rSAGO-1. sb-P, avec un ARNsb P, sb-3-P, avec un ARNsb 3-P. Le schéma du haut représente les séquences des molécules, le brin guide est représenté en rouge tandis que le brin d'ARN cible est représenté en noir. Le marqueur de taille utilisé est obtenu par hydrolyse alcaline de l'ARN cible radiomarqué. c) Essai de clivage avec 100 nM de GST et rRDE-1 sur un duplex de 21 nt parfaitement apparié. + avec duplex, e avec duplex et EDTA. Migration sur gel de polyacrylamide dénaturant 15%.

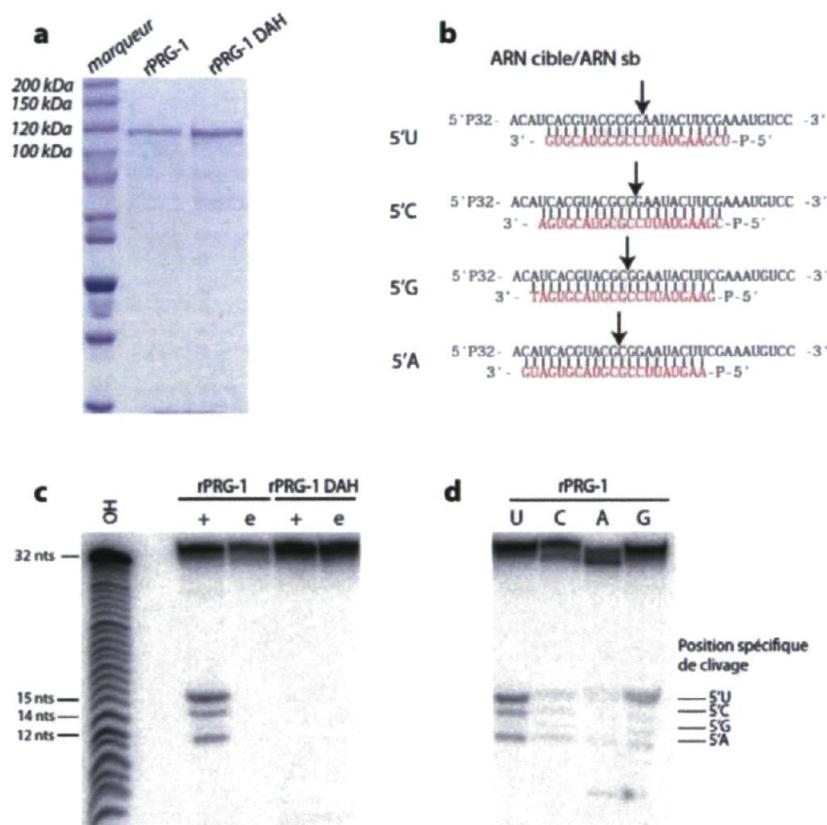


## 1.2. Détermination de l'activité endonucléasique potentielle de rPRG-1

L'autre membre de la famille des protéines Argonaute auquel je me suis intéressé est PRG-1. Cette protéine possède le motif DDH et est impliquée dans la voie des piARN. *In vivo*, il a été démontré qu'elle liait les piARN chez *C. elegans*, molécules de 21 nt avec la caractéristique de porter un U en 5' (Batista et al., 2008; Wang and Reinke, 2008). En utilisant le même système, nous avons produit PRG-1 fusionnée à la GST (Fig.20a). Tout comme rRDE-1, rPRG-1 possède une activité catalytique (Fig.20c). Nous avons muté le second résidu D du motif DDH (Fig.20b) pour déterminer si cette mutation affecte l'activité de rPRG-1. La mutation au niveau du motif DDH engendre la perte de l'activité spécifique, démontrant ainsi que la substitution de ce seul résidu est suffisante pour affecter

l'activité de cette protéine (Fig.20c). Étant donné que PRG-1 lie de courts ARN portant un U en 5', nous avons décidé de tester si son activité en était dépendante en utilisant différents ARNsb 100% complémentaire à la même molécule d'ARN cible radiomarké (Fig.20b). D'après les résultats de clivage, cette protéine possède une spécificité d'action, puisque les produits spécifiques correspondants à la coupure en position 10-11 des ARNsb sont quasi absents lorsque le premier nucléotide en 5' n'est pas un résidu U. Ceci démontre que l'utilisation de ces molécules est peu efficace comparativement à l'ARNsb portant un U en 5'. Ces résultats indiquent que rPRG-1 possèdent une activité catalytique due à son motif DDH et qu'elle affiche une préférence pour le premier nucléotide en 5' en ce qui concerne son activité de clivage. De plus, cela démontre assez bien que la méthode de production de protéines et l'étiquette GST n'affectent pas les protéines Argonaute puisque le test d'activité avec différents ARNsb reflète la préférence observée *in vivo* pour le U en 5' terminal.

**Figure 20 : Activité endonucléasique de la protéine rPRG-1.** a) Gel SDS-PAGE 8% avec 0,5 à 1 µg de protéines recombinantes révélé par une coloration au bleu de Coomassie. b) Séquences des molécules utilisées pour les essais de clivages, le brin guide est représenté en rouge tandis que le brin d'ARN cible est représenté en noir. La nature du premier nucléotide de l'ARNsb P guide y est indiquée. c) essai de clivage avec 100 nM de rPRG-1 ou rPRG-1 DAH. + avec ARNsb P, e ARNsb P et EDTA. Le marqueur de taille utilisé est l'ARN cible radiomarqué et lysé par hydrolyse alcaline. d) essai de clivage avec différents ARNsb P; U, U en 5'; C, C en 5'; A, A en 5'; G, G en 5'. Migration sur gel de polyacrylamide dénaturant 15%.



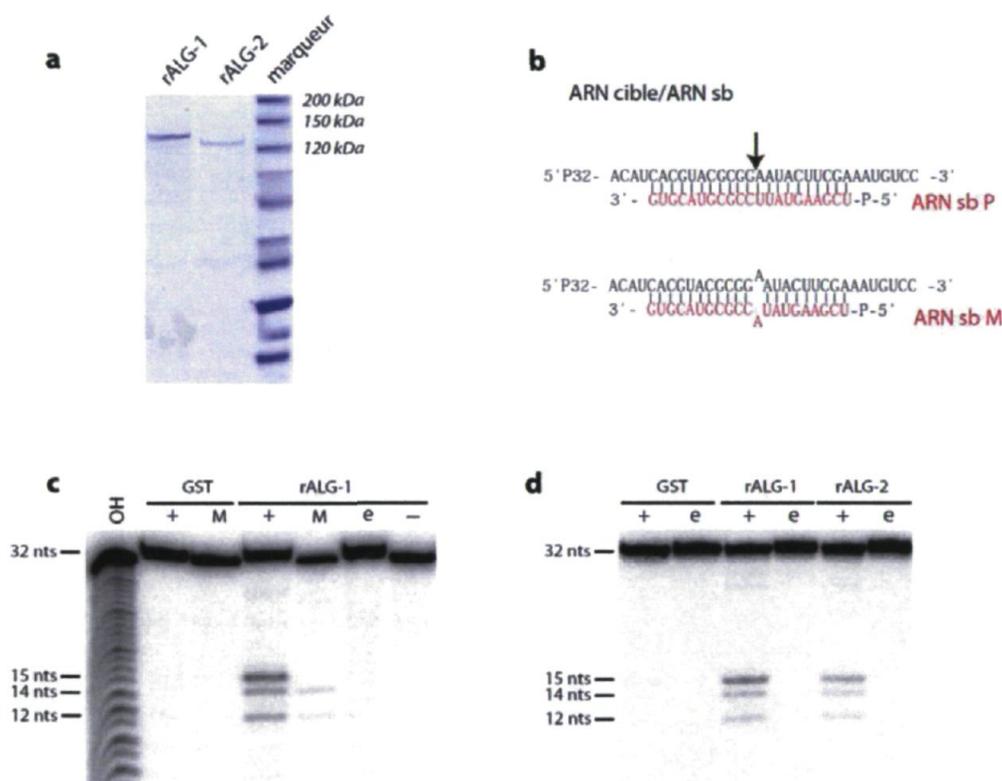
### 1.3. Détermination de l'activité endonucléasique potentielle de rALG-1 et rALG-2

Concernant ALG-1 et ALG-2, nous avons suivi la même méthode. Nous avons produit les protéines recombinantes rALG-1 et rALG-2, sous forme de fusion GST, pour déterminer leurs activités endonucléasiques (Fig.21a). Étant donné que les protéines rALG-1/2 sont hautement homologues avec une similarité de 81% (Vasquez-Rifo et al., 2012), beaucoup d'expériences effectuées ont donné des résultats similaires.

### 1.3.1. Mise en évidence de l'activité endonucléasique de rALG-1 et rALG-2

Tout comme rRDE-1 et rPRG-1, rALG-1 et rALG-2 sont capables de cliver une molécule d'ARN cible en utilisant un ARNsb guide P (Fig.21c et 21d). Comme précédemment, la réaction dépend des ions  $Mg^{2+}$  et les produits de clivage obtenus sont similaires et correspondent à des tailles de 15, 14 et 12 nt. Le produit majoritaire est celui de 15 nt et qui correspond à la taille attendue si le clivage s'effectue face à la position 10-11 de l'ARNsb guide. Pour déterminer si les produits de 14 et 12 nt sont issus de la dégradation ou de site alternatif de clivage, nous avons effectué une mutation d'un nucléotide au niveau de l'ARNsb en position 10-11 (Fig.21b et 21c). Cette mutation a pour effet d'abolir la production de la molécule de 15 nt indiquant bien que cette bande est due à l'activité des protéines (Fig.21c) et qu'elle est générée uniquement par appariement parfait entre l'ARN cible et l'ARN guide. Elle génère aussi moins de produits de 12 et 14 nt, suggérant que leur production est affectée par un mésappariement en position 10-11 et qu'ils proviennent certainement de clivage face aux positions 11-12 et 13-14 de l'ARN cible. On peut donc conclure que rALG-1 et rALG-2 répondent aux critères et ont une activité endonucléasique site-spécifique.

**Figure 21 : Activité endonucléasique des protéines rALG-1 et rALG-2.** a) Gel SDS-PAGE 8% avec 0,5 à 1 µg de protéines recombinantes révélé par une coloration au bleu de Coomassie. b) Séquence des molécules d'ARN utilisées pour les essais de clivage, le brin guide est représenté en rouge tandis que le brin d'ARN cible est représenté en noir. ARNs b P représente l'ARNsb parfaitement apparié et ARNs b M représente l'ARNsb avec un mésappariement en position 10-11. c) Essai de clivage avec 100 nM de GST, rALG-1 dans différentes conditions. P avec l'ARNsb P, M avec l'ARNsb muté en position 10-11, e avec EDTA, - sans ARNs b. OH marqueur de taille par lyse alcaline de l'ARN cible. d) Essai de clivage avec 100 nM de GST, rALG-1 ou rALG-2 dans différentes conditions. + avec ARNs b P, e ARNs b P et EDTA. Migration sur gel de polyacrylamide dénaturant 15%.

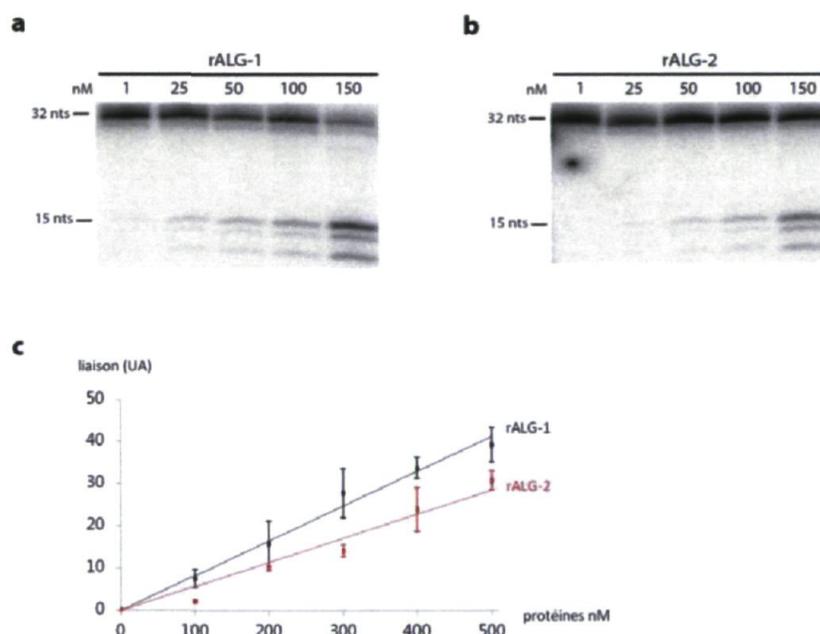


### 1.3.2. Comparaison de l'activité de rALG-1 et rALG-2

Nous avons ensuite comparé l'apparition des produits de clivage de rALG-1 et rALG-2, en utilisant différentes concentrations de protéines. Le produit de clivage est plus abondant pour rALG-1 que pour rALG-2. Lorsque l'on utilise une gamme de concentrations de rALG-1 et rALG-2, il s'avère que le produit détecté est toujours plus abondant pour rALG-1 (Fig.22a et 22b). Les produits sont détectables à partir de 25 nM pour rALG-1 alors que pour rALG-2, pour la même quantité de protéines, les produits ne sont pas détectables. Utilisant le même ARNs b que pour le clivage, nous avons déterminé par «Filter Binding Assay» (FBA) les courbes de liaison ARN-protéines afin de juger si l'affinité des protéines

pour l'ARNsb P peut expliquer la différence d'activité. Pour une même quantité de protéines, rALG-1 lie toujours d'avantages d'ARNsb que rALG-2 (Fig.22c). À 100 nM, la différence de liaison est négligeable et cette quantité correspond à la concentration la plus élevée utilisée pour les expériences de clivages. Plus on augmente la quantité de protéines, plus la différence est marquée et on arrive à une différence d'environ 25% de moins pour rALG-2 (10 unités Arbitraire (UA) sur 40) pour 500 nM de protéines. L'affinité pour l'ARNsb P pourrait donc expliquer l'activité plus forte de rALG-1.

**Figure 22 : Activité endonucléasique et courbe de liaison à l'ARNsb P pour les protéines rALG-1 et rALG-2.** a) Essai de clivage ARN cible/ARNsb P pour rALG-1, les concentrations de protéines varient de 1 à 150 nM de protéines et l'ARN guide utilisé est l'ARNsb P. b) Essai de clivage ARN cible/ARNsb P pour rALG-2, les concentrations de protéines varient de 1 à 150 nM de protéines et l'ARN guide utilisé est l'ARNsb P. Migration sur gel dénaturant 15%. c) Test de FBA de rALG-1 et rALG-2 sur l'ARNsb P. Comparaison des courbes de liaison de rALG-1 et rALG-2 avec le substrat ARNsb P.

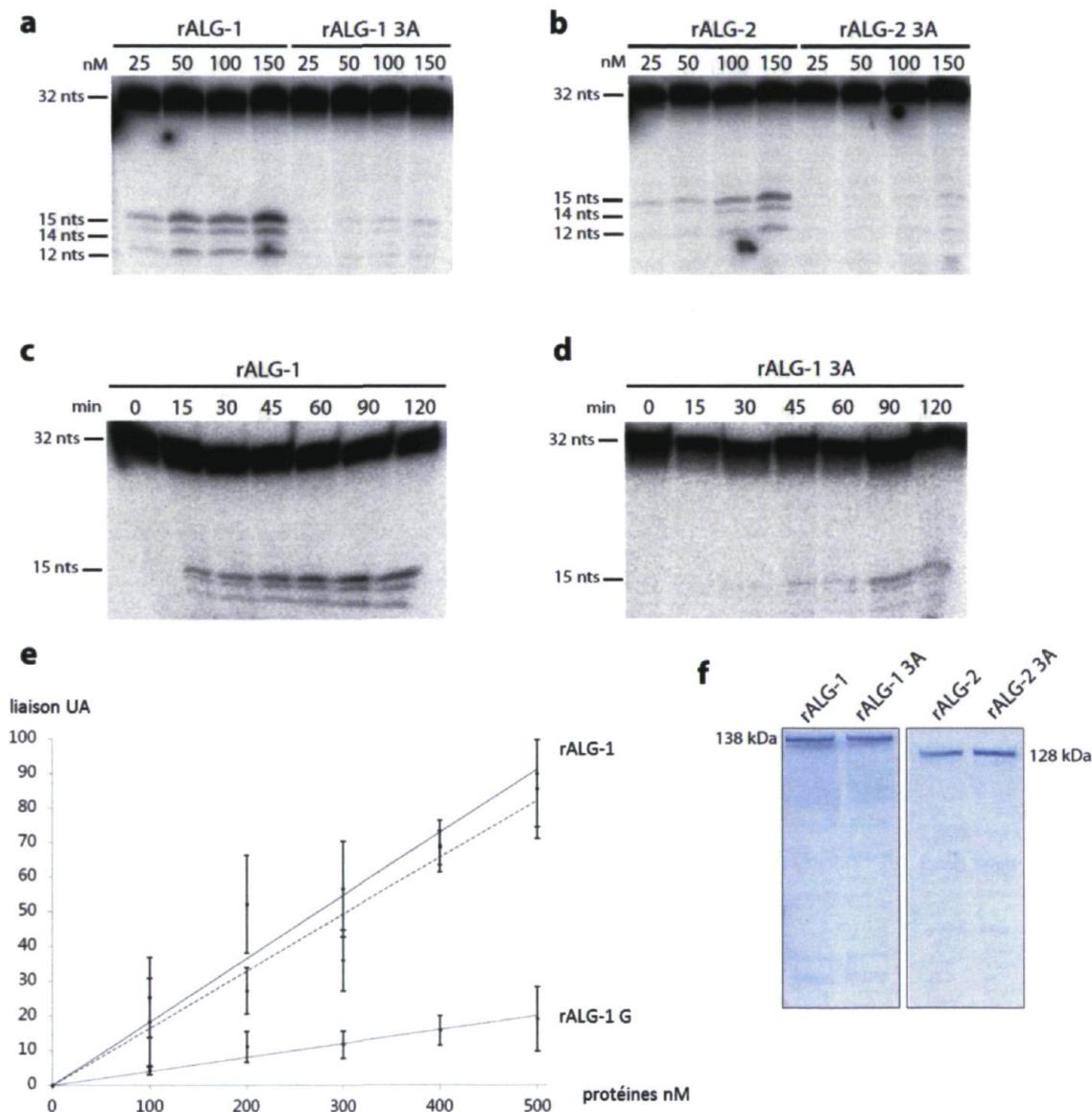


### 1.3.3. Dépendance du motif DDH pour l'activité endonucléasique

Pour vérifier que cette activité est spécifique, nous avons investigué si la perte du site actif entraîne une perte de fonction. Nous avons donc muté le motif DDH pour 3A, et produit rALG-1 3A et rALG-2 3A (Fig.23f). Étant donné qu'une activité résiduelle était détectable pour rRDE-1 3A à 100 nM, nous avons utilisé des concentrations de protéines de 25 à 150 nM. À des concentrations inférieures à 100nM de protéines rALG-1 3A et rALG-2 3A,

nous ne détectons pas d'activité résiduelle (Fig.23a et 23b). Cependant, à forte concentration (100 nM et plus), on voit apparaître les mêmes produits que pour les protéines sauvages mais à l'état de trace (Fig.23a et 23.b). Nous avons donc suivi au cours du temps la formation de ces produits à une concentration de 200 nM de rALG-1 3A (Fig.23c et 23d). Comparativement à rALG-1 où le produit apparaît en début de réaction à 15 min, le produit de rALG-1 3A apparaît à 60 min en fin de réaction. De plus, le produit est moins abondant à la fin de la réaction pour rALG-1 3A qu'en début de réaction pour rALG-1. On peut conclure que le motif DDH est essentiel pour l'activité endonucléasique des protéines Argonaute et que la perte de ce motif abolie presque totalement cette activité. Par la suite, nous avons utilisé ces mutants pour déterminer leur capacité à lier l'ARNsb afin de déterminer si la perte d'activité était due à une perte de liaison. Utilisant le même ARNsb que pour le clivage, nous avons déterminé par FBA les courbes de liaison ARN-protéines afin de juger si l'affinité des mutants 3A pour l'ARNsb était affectée. La liaison de rALG-1 3A n'est pas significativement affectée (Fig.23e) (rALG-2 3A donne des résultats similaires). À titre comparatif, nous avons inclus un contrôle qui consiste en une protéine mutée dans la région Mid au niveau d'une boucle importante pour la sélection du premier nucléotide en 5' (Frank et al., 2010). L'insertion d'un G (rALG-1 G) perturbe la reconnaissance du premier nucléotide et conduit à une baisse de la liaison à l'ARNsbP. La qualité de la protéine recombinante est comparable à celle obtenu pour rALG-1 et RALG-2 (données non montrées). Par rapport à rALG-1/rALG-1 3A, rALG-1 G a une capacité de liaison de l'ARNsbP réduite (Fig.23e). De ces expériences, on peut conclure que la mutation du motif DDH affecte seulement l'activité de clivage sur une cible et non la liaison des ARNsb guides.

**Figure 23 : Activité endonucléasique et courbe de liaison à l'ARNsb P pour les protéines rALG-1, rALG-1 3A, rALG-2 et rALG-2 3A.** Les concentrations de protéines varient de 25 à 150 nM de protéines et l'ARN guide utilisé est l'ARNsb P. a) avec rALG-1 et rALG-1 3A. b) avec rALG-2 et rALG-2 3A. c) et d) comparaison au cours du temps de l'apparition du produit entre 0 et 120 minutes avec 200 nM de protéines rALG-1 et rALG-1 3A. Migration sur gel dénaturant 15%. e) Test de FBA de rALG-1, rALG-1 3A et rALG-1 G sur l'ARNsb P. Comparaison des courbes de liaison de ces protéines avec le substrat ARNsb P. f) Gel SDS-PAGE 8% avec 0,5 à 1 µg de protéines recombinantes révélé par une coloration au bleu de Coomassie.



### 1.3.4. Autres domaines influençant l'activité de rALG-1 et rALG-2

Dans la mesure de mieux comprendre comment cette activité est soutenue, nous avons créé des délétions de ces protéines et fait une analyse systématique de l'activité que j'ai présenté sous forme de tableau (Tableau 3). Certains mutants de délétions de protéines Argonaute sont très sujets à la dégradation pour une raison inconnue et n'ont pas été inclus (des tests sur ces protéines ont montré la dégradation de l'ARN cible). D'autres sont de bonne qualité et ont pu être testé. La première observation intéressante est l'incapacité du domaine PIWI seul à pouvoir effectuer un clivage, soulignant la participation nécessaire des autres domaines. Le deuxième point est l'incapacité de la protéine déléetée du domaine Mid de pouvoir effectuer un clivage. Ceci n'est pas le cas des mutants déléetés du domaine PAZ (ce point sera présenté dans une partie ultérieure) qui affichent, comme la protéine déléetée du C-ter, une baisse significative du clivage.

**Tableau 3 : Présentation des différents mutants créés pour les protéines rALG-1 et rALG-2.** La nature des mutations y est présentée, ainsi que la qualité de la protéine et les résultats correspondants du test de clivage ARNsb/ARN cible. Les cases grises correspondent aux mutants de mauvaise qualité et dont l'activité n'a pu être validée. Des résultats similaires ont été obtenus pour rALG-1 et rALG-2.

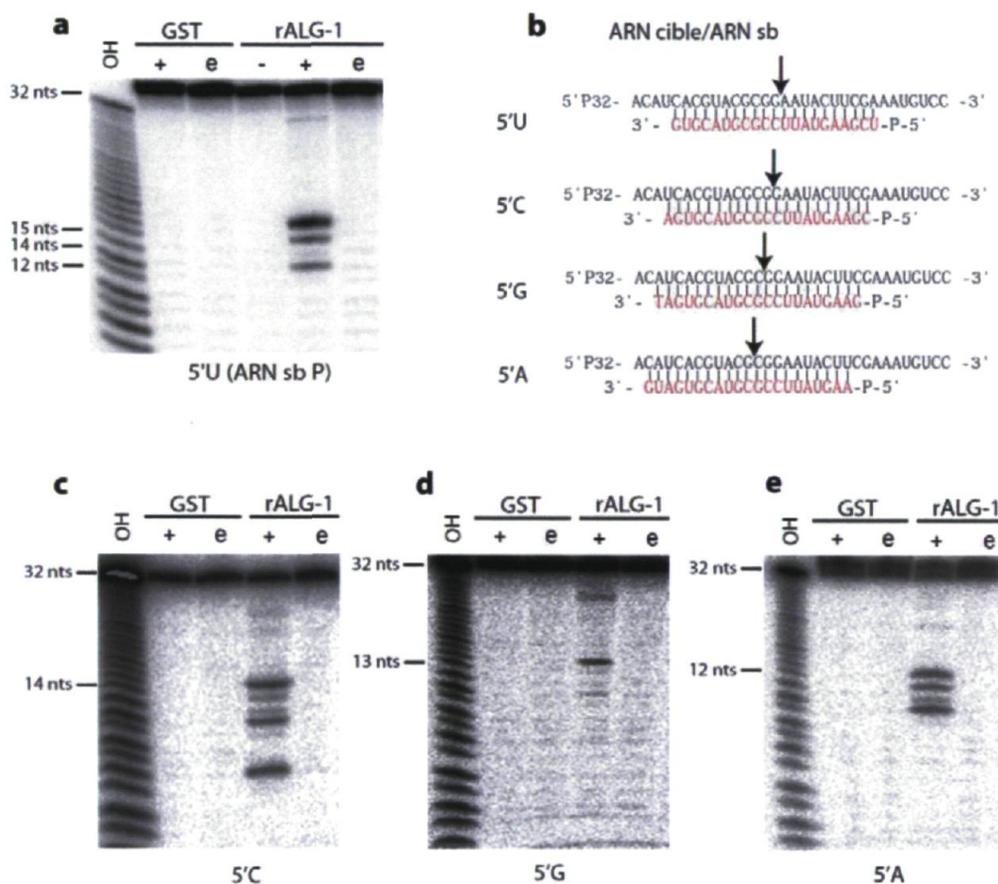
| MUTATION        | QUALITÉ | ARN sb/ARN cible |
|-----------------|---------|------------------|
| PIWI            | bonne   | pas de clivage   |
| PIWI-Cter       | bonne   | pas de clivage   |
| Mid-PIWI-Cter   | dégradé |                  |
| Δ PAZ           | bonne   | clivage affecté  |
| Δ Mid           | bonne   | pas de clivage   |
| Δ PIWI          | dégradé |                  |
| Δ Cter-         | bonne   | clivage affecté  |
| Δ PIWI-Cter     | dégradé |                  |
| Δ Mid-PIWI-Cter | dégradé |                  |

### 1.3.5. Importance du nucléotide 5'terminal pour le clivage de rALG-1/2

Nous avons testé si la nature du premier nucléotide en 5' influence l'efficacité du clivage comme c'est le cas pour la protéine rPRG-1. Pour cela, les molécules d'ARNsb complémentaire de 21 nt, portant en 5' un C, G ou A, ont été utilisées (Fig.24b) avec toujours le même ARN cible que précédemment. Il en résulte une différence au niveau du site théorique de clivage de l'ARN cible. Aucune différence n'a pu être détectée entre rALG-1 et rALG-2 (données non-montrées). La Figure 24a montre le patron observé avec l'ARNsb P, qui porte un U en 5', et les trois produits de 15, 14 et 12 nt. Comme le montre les Figure 24c, 24d et 24e, des produits de clivage pour tous les ARNsb guides utilisés ont été observés contrairement aux résultats obtenus avec rPRG-1 (Fig.20d). Cela démontre que rALG-1/2 prend en charge ces différents types de molécules.

On retrouve plusieurs bandes de clivage et/ou dégradation pour chacun des ARNsb testés. Pour l'ARNsb G, une bande plus faible correspond à la taille attendue (Fig.24d). Dans le cas des ARNsb A et C, on observe un profil similaire à l'ARNsb U avec plusieurs bandes dont la bande la plus haute correspond au produit spécifique de clivage de 12 nt et 14 nt respectivement (Fig.24e et 24c). Cependant, le nombre de bandes contaminantes varie avec 5 bandes pour l'ARNsb 5'C et 4 bandes pour l'ARNsb 5'A. Pour les ARNsb où des produits multiples sont observés, les bandes correspondent toutes à des troncations d'un ou deux nt, sauf pour la bande la plus basse de l'ARNsb C où l'on observe un «saut» de 3 nt. En conclusion, cette expérience indique que les protéines rALG-1/2 sont capables d'utiliser tous types de molécules d'ARNsb de 21 nt, indépendamment de leur 5'terminal, mais qu'un biais existe concernant le 5'G dont l'activité endonucléasique est réduite comparé aux 5'U/A/C.

**Figure 24 : Activité endonucléasique pour la protéine rALG-1 avec différents ARNsb.**  
 a) Essai de clivage pour 100 nM de GST ou rALG-1 avec l'ARNsb P (5'U). +, avec ARNsb P; e, avec ARNsb P et EDTA. b) Représentation des différentes molécules d'ARNsb utilisées portant un U, A, G ou C en 5'. c) Essai de clivage pour 100 nM de GST ou rALG-1 avec l'ARN 5'C. +, avec ARN 5'C; e, avec ARN 5'C et EDTA. d) Essai de clivage pour 100 nM de GST ou rALG-1 avec l'ARN 5'G. +, avec ARN 5'G; e, avec ARN 5'G et EDTA. e) Essai de clivage pour 100 nM de GST ou rALG-1 avec l'ARN 5'A. +, avec ARN 5'A; e, avec ARNsb A et EDTA. OH, marqueur de taille par hydrolyse alcaline de l'ARN cible. Migration sur gel dénaturant 15%.

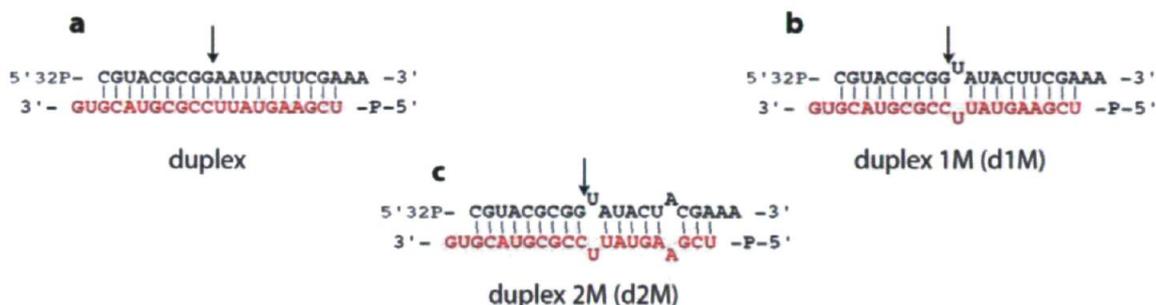


## 2. Activité de rALG-1 et rALG-2 sur les duplexes de courts ARN et les pré-miARN

Dans la première section des résultats, nous avons démontré que les différentes protéines Argonaute testées, possédant un motif DDH, étaient douées d'activités catalytiques. Dans un contexte *in vivo*, le ciblage et la dégradation d'un ARNm cible d'un miARN nécessite l'intervention de l'activité endonucléasique des protéines Argonaute dans des cas isolés (Yekta et al., 2004). Il est donc très probable que les substrats potentiels des protéines Argonaute de cette voie soient des duplexes de miARN, après la maturation des pré-miARN par Dicer. Nous avons donc étudié l'action de rALG-1 et rALG-2 sur ces différentes molécules.

Pour l'étude des duplexes de miARN, nous avons utilisé des molécules qui sont similaires aux produits de Dicer (Fig.25). Ainsi, les duplexes sont de 21 nt avec une extrémité 3' OH sortante de 2 nt, marque du clivage par Dicer. Les extrémités 5' sont monophosphorylées et l'identité des séquences d'ARN aux extrémités est de nature à établir une asymétrie de stabilité favorable à la sélection d'un des deux brins. Trois molécules ont été testées, un duplex 100% complémentaire (duplex), un duplex mésapparié en position 10 (d1M) et un duplexe mésapparié aux positions 4 et 10 (d2M). Ces duplexes avec un ou deux mésappariements miment la grande majorité des miARN qui sont généralement mésappariés en position centrale (d1M) et/ou à l'extrémité 5' du brin miARN (d2M). Je tiens à préciser ici que les molécules duplexes sont appariées par chauffage (95°C) puis refroidissement lent d'un mélange équimolaire de chaque ARNs. L'intégrité de chaque duplex est ensuite vérifiée sur gel de polyacrylamide non-dénaturant (voir la section matériels et méthodes).

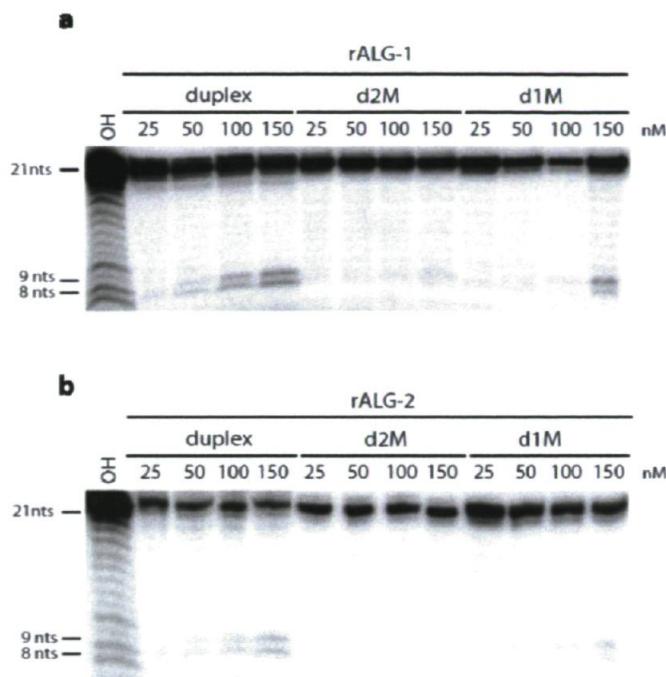
**Figure 25 : Duplexes utilisés dans les différentes expériences.** a) duplex parfaitement apparié (duplex). b) duplex mésapparié en position 10 (d1M). c) duplex mésapparié en position 4 et 10 (d2M). Le brin passager radiomarqué est indiqué en noir et le brin guide en rouge. La flèche indique la position théorique de clivage face à la position 10-11.



## 2.1. Test de l'activité endonucléasique sur les duplexes de courts ARN

Nous avons radiomarqué en 5' les brins passagers (en noir Fig.25) et apparié ces brins avec l'ARNsb P, utilisé précédemment lors des essais de clivage ARNsb/ARN cible (en rouge, Fig.25). Nous avons ensuite incubé ces différents duplexes durant 2 h en présence des protéines rALG-1/2 à différentes concentrations. Les résultats de ces expériences ont montré que rALG-1 est capable de cliver le duplex parfaitement apparié (Fig.26.a, duplex). Deux produits sont présents, une bande de 9 nt et une autre de 8 nt. Le produit issu du clivage en position 10-11 correspond au produit de 9 nt, le site spécifique de clivage. L'autre produit de 8 nt correspond à un second site de coupure face à la position 11-12 de l'ARN guide. L'introduction d'un mésappariement en position 10 (d1M) abolit le produit de clivage de 9 nt (Fig.26.a, d1M) et l'ajout d'un autre mésappariement en position 4 (d2M) abolit presque totalement le clivage du brin passager (Fig.26.a, d2M). Le clivage est donc efficace sur un duplex sans mésappariement, mais sur les duplexes s'apparentant aux duplexes de miARN, d1M et d2M, l'activité est réduite en fonction du degré de mésappariement. Pour rALG-2, on obtient pratiquement les mêmes résultats bien que les produits à concentration équivalente de protéines soient toujours plus faibles que pour rALG-1 (Fig.26.b). Il en résulte une activité sur les duplexes d1M et d2M très faible comparée à rALG-1.

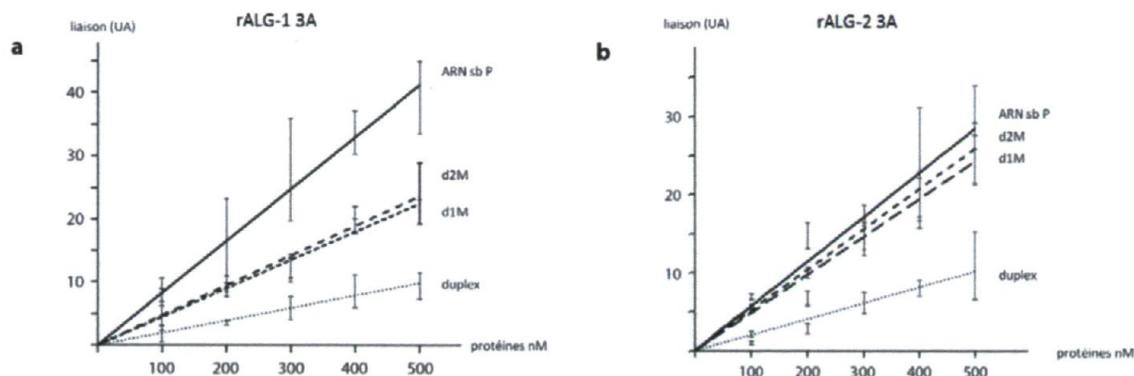
**Figure 26 : Essai de clivage des duplexes d'ARN. 25 à 150 nM de recombinantes ont été utilisés.** Le marqueur de taille (OH) a été obtenu par hydrolyse alcaline du substrat d'ARN de 21 nt. Les molécules utilisées sont ceux de la figure 25. a) rALG-1. b) rALG-2. Migration sur gel de polyacrylamide dénaturant 15%.



## 2.2. Courbes de liaison aux duplexes de courts ARN

Afin de déterminer si l'efficacité de clivage observée pour ces types de duplexes sont dues à des différences au niveau de l'affinité de rALG-1/2, nous avons analysé par FBA la liaison à ces trois molécules. Pour cette étude, des protéines mutantes 3A ont été utilisées pour éviter le clivage des différents duplexes. Ces protéines mutantes ne devraient pas interférer dans la liaison, puisque nous avons vu que pour l'ARNsb, il n'y avait pas de différence de liaison entre WT et 3A (Fig.23e). En ce qui concerne rALG-1 3A, il s'avère que la molécule qu'elle lie le moins est le duplex 100% complémentaire (Fig.27a). La liaison est plus forte pour les duplexes possédant des mésappariements, indiquant une affinité plus forte pour ce type de molécule. Finalement, la préférence de rALG-1 se porte davantage sur l'ARNsb. rALG-2 affiche les mêmes caractéristiques (Fig.27b). Cependant, pour rALG-2, la différence de liaison entre l'ARNsb et d1M/d2M n'est pas significative.

**Figure 27 : Courbe de liaison à différents substrats d'ARN des protéines rALG-1 3A et rALG-2 3A.** Les ARN utilisés sont les ARNs b P, duplex, d1M et d2M (Fig.25). a) rALG-1 3A. b) rALG-2 3A.



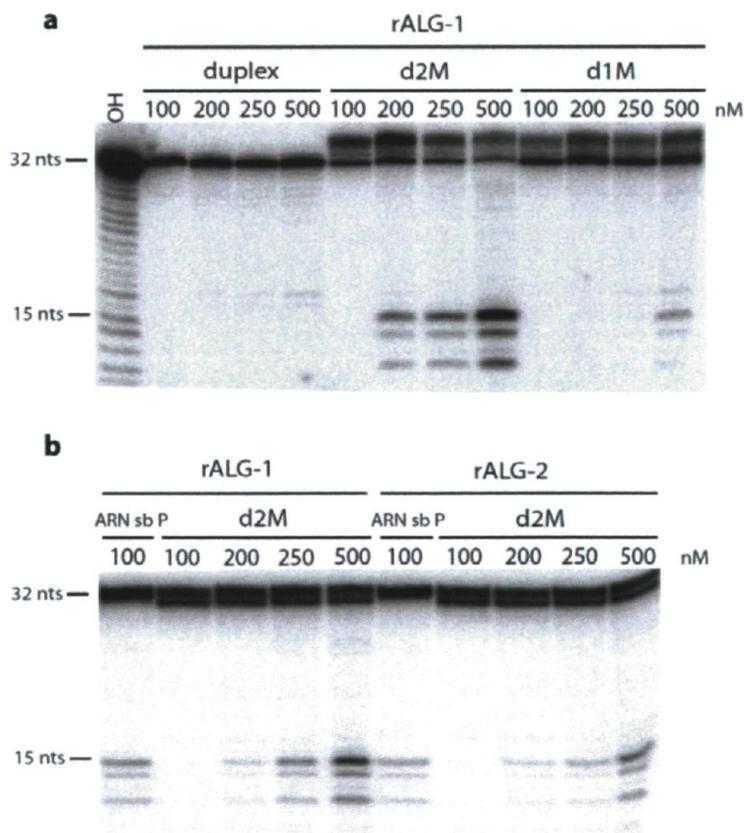
### 2.3. Détermination d'une activité de relargage du brin passager d'un duplex de courts ARN

La comparaison entre la liaison et les données du clivage des duplexes apporte un point intrigant. Les molécules ayant le plus d'affinité pour rALG-1/2 possèdent des mésappariements et ne sont que faiblement clivées. On s'est donc demandé quelle pouvait être la signification biologique de cette observation. Une étude a démontré que les protéines Argonaute humaines étaient en mesure de lier des duplexes d'ARN et de se défaire de l'un des brins pour former un complexe Argonaute-ARNsb de façon clivage indépendant (Wang et al., 2009a). Nous avons testé si rALG-1/2 étaient capable de cette fonction. Pour mettre en évidence cette forme ARNb, nous avons utilisé la capacité de clivage de rALG-1/2 dans un test de clivage modifié en incubant tout d'abord rALG-1/2 avec les différents duplexes, puis en ajoutant au milieu réactionnel une cible radiomarquée complémentaire au brin guide. Ainsi, si un produit de clivage apparaît, cela est dû à la présence de complexe Argonaute-ARNsb issue du relargage d'un des brins du duplex.

Les 3 duplexes décrits Figure 25 ont été testées. Il s'est avéré que le duplex 100% complémentaire nous indique que ce relargage est totalement indépendant de l'activité de clivage (Fig.28a) car lorsque l'on utilise ce duplex, lequel est clivé par rALG-1/2 (Fig.26a et 26b), il n'y a pas de formation de complexe Argonaute-ARNsb capable d'aller couper l'ARN cible (Fig.28a). Autrement dit, il n'y a pas de relargage du brin passager clivé du duplex. En revanche, l'utilisation des duplexes d1M et d2M (Fig.28a) permet la formation

de complexe Argonaute-ARNsb qui clivera l'ARN cible. Cet effet est dépendant de la concentration en protéines et est totalement indépendant du clivage étant donné que le duplex clivé ne peut être utilisé pour former un complexe Argonaute-ARNsb. Il requiert également des mésappariements. Pour le d2M, on détecte un produit à partir de 200 nM de protéines tandis que pour d1M on ne détecte les produits qu'à 500 nM. Pour d2M, on observe le même résultat pour rALG-2 (Fig.28b). Ainsi, on peut dire que rALG-1/2 possède une fonction de relargage du brin passager indépendante de l'activité endonucléasique.

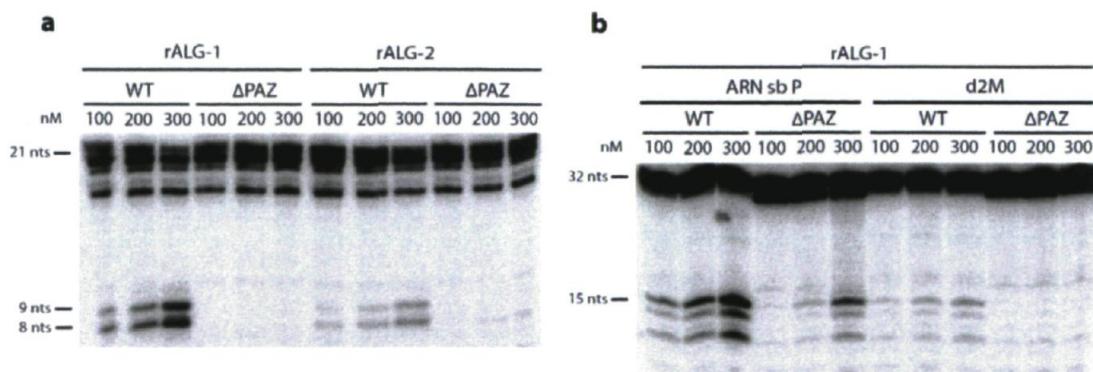
**Figure 28 : Essai de relargage d'ARNsb à partir de différents duplexes.** a) Les duplexes (duplex, d1M et d2M) ont été incubés avec rALG-1 (100 à 500 nM) pendant 1 h 30, puis une cible radiomarquée complémentaire au brin guide a été ajoutée au milieu réactionnel pendant 1 h 30. Le marqueur de taille (OH) a été obtenu par hydrolyse alcaline de l'ARN cible de 32 nt. b) l'ARNsb P et le duplex d2M ont été incubés avec respectivement 100 nM ou de 100 à 500 nM de protéines rALG-1 ou rALG-2. La réaction est similaire à a. Migration sur gel de polyacrylamide dénaturant 15%.



## 2.4. Mise en évidence de l'implication du domaine PAZ dans la prise en charge des duplexes

Nous avons noté que, chez les mutants de délétions de rALG-1 et rALG-2 testés pour le clivage ARNsb/ARN cible (Tableau 3), certains mutants conservaient une activité endonucléasique, notamment le mutant  $\Delta$ PAZ. Ce domaine est partagé avec la protéine Dicer et est nécessaire pour la reconnaissance de l'extrémité 3'OH des pré-miARN (Macrae et al., 2006). Il était donc intéressant de tester si ce mutant de délétion était nécessaire pour la reconnaissance des différents duplexes de 21 nt. Les tests de clivage ARNsb/ARN cible ont démontré la capacité de ce mutant de délétion à cliver l'ARN cible (Fig.29b, ARNsb P). Le produit est cependant nettement plus abondant chez la protéine sauvage que chez ce mutant, suggérant que l'activité endonucléasique et/ou la liaison à l'ARNsb sont affectées. Pour rALG-1  $\Delta$ PAZ et rALG-2  $\Delta$ PAZ, nous avons noté que ces mutants perdaient la capacité de cliver le duplex 100% complémentaire (Fig.29a) aux mêmes concentrations pour lesquelles nous avons encore une activité de clivage ARNsb/ARN cible. Dès lors, nous avons testé si ces mutants étaient fonctionnels pour l'essai de relarguage du brin passager et le résultat est négatif (Fig.29b, d2M). Cependant, ces résultats doivent tenir compte du fait que l'activité est réduite pour le clivage de l'ARNsb mais ils vont dans le sens d'une fonction de ce domaine dans la prise en charge des molécules d'ARN duplex.

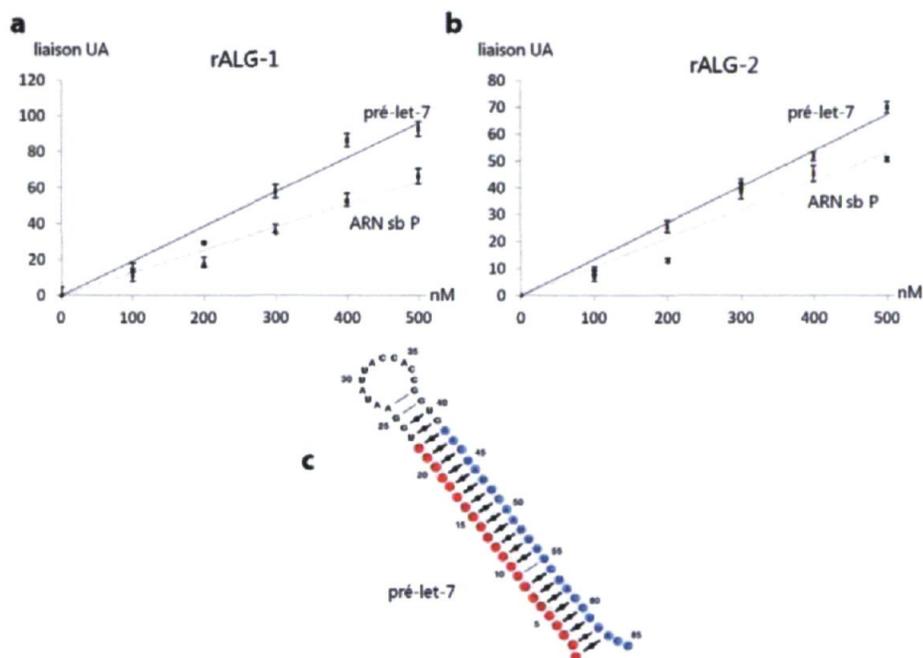
**Figure 29 : Activité de clivage des duplexes par les mutants rALG-1/2  $\Delta$ PAZ.** a) Comparaison des activités sur un duplexe 100% complémentaire des protéines rALG-1/2 sauvages (WT) et mutantes ( $\Delta$ PAZ). De 100 à 300 nM de protéines ont été utilisées. b) Comparaison des activités de relarguage de rALG-1 WT et rALG-1  $\Delta$ PAZ sur le duplex d2M. L'ARNsb est utilisé comme contrôle d'activité endonucléasique de rALG-1  $\Delta$ PAZ. De 100 à 300 nM de protéines ont été utilisées.



## 2.5. Activité de rALG-1 et rALG-2 sur le pré-miARN let-7

Selon la littérature, la protéine Ago2 humaine peut prendre en charge pré-let-7 et créer une coupure dans la molécule de pré-let-7 au niveau du brin correspondant au futur brin passager du duplex de miARN (Diederichs and Haber, 2007). Ce précurseur porte le miARN let-7 sur son bras 5' et, au niveau de la tige contenant le duplex de miARN, cette molécule ne possède qu'un seul mésappariement en position 9 (en partant du 5'), rendant théoriquement possible une coupure aux positions 10-11 (Fig.30c). À la fois pour rALG-1/2, nous n'avons observé aucune activité de clivage sur cette molécule, et ce, dans toutes les conditions utilisées (données non-montrées). L'ajout d'ATP, la variation de la concentration de protéines, de sels, du temps de réaction, le mélange des deux protéines n'apportent aucun changement (données non-montrées). Malgré tout, nous avons constaté, parmi tous les substrats d'ARN utilisés, que le pré-let-7 était la molécule que rALG-1/2 liait avec le plus d'efficacité (Fig.30a et b). Comparés à la liaison de l'ARNsb P, nous avons une augmentation plus prononcée pour rALG-1 que rALG-2. Ainsi, l'affinité est plus grande pour les précurseurs que pour toutes autres espèces d'ARN.

**Figure 30 : Courbes de liaisons à pré-let-7 et l'ARNsb P pour rALG-1 et rALG-2.** pré-let-7 est représenté en trait foncé, et l'ARNsb P en trait clair. a) rALG-1. b) rALG-2. c) Structure prédite de pré-let-7 selon le logiciel MC-fold. La structure secondaire du pré-let-7 est représentée en rouge sur le bras 5' le miARN let-7, en bleu le brin passager. On voit que le degré de complémentarité est presque parfait.



### **3. Conclusions sur l'étude *in vitro* des Argonaute**

Cette étude biochimique a permis de mettre en évidence plusieurs caractéristiques des protéines Argonaute et notamment leur activité endonucléasique médiée par le motif DDH. Concernant rALG-1/2, ces protéines prennent en charge différentes molécules d'ARN et ont une grande affinité pour les structures précurseurs de miARN et les courts ARNs<sub>b</sub>. Pour les duplexes, elles lient davantage ceux possédant des mésappariements et sont capables, pour ce type de molécule, de relarguer le brin passager indépendamment d'une activité de clivage. Les pré-miARN, *in vitro* du moins, ne sont pas sujets à l'action des protéines Argonaute.

# Chapitre 4 : Caractérisation de l'activité des protéines ALG-1/2 dans la voie des miARN *in vivo*

Nos données *in vitro*, qui démontrent que les protéines ALG-1 et ALG-2 possèdent une activité catalytique, nous portent à penser que cette activité est conservée et nécessaire pour la voie des miARN chez *C. elegans*. Nous posons l'hypothèse que ce motif a été conservé, chez ALG-1/2, car l'activité qui en découle intervient dans une étape clé de la synthèse des miARN. Par comparaison aux protéines RDE-1 et SAGO-1 notamment, la contribution d'une activité catalytique dans la voie de l'exo-ARNi est très informative. Bien que RDE-1 et SAGO-1 (et les Argonaute secondaires) soit essentiels pour la voie de l'exo-ARNi, l'une des protéines a conservé ce motif DDH tandis que l'autre non. Le fait que l'activité catalytique de RDE-1 est nécessaire pour l'exo-ARNi (Steiner et al., 2009), tout comme la présence de la protéine non-catalytique SAGO-1 (et les Argonaute secondaires) (Yigit et al., 2006), suggère que les protéines Argonaute peuvent intervenir dans deux registres distincts. Pour étudier si l'activité d'ALG-1/2 est nécessaire dans la voie des miARN, nous avons créé plusieurs lignées transgéniques sauvages et mutées au niveau du motif DDH, à la fois pour les gènes *alg-1* et *alg-2*. Deux mutants nuls sont disponibles pour ces gènes, la souche *alg-1(gk214)* et la souche *alg-2(ok304)* (ces deux souches seront appelées *alg-1* et *alg-2*). La souche sauvage de référence utilisée est la souche N2.

## 1. Analyse des mutants nuls *alg-1* et *alg-2*

### 1.1. Phénotypes observés

La souche *alg-1* est malade et plusieurs processus biologiques sont défectueux. Ainsi, on observe une certaine létalité avec notamment explosion de la vulve chez 19% des individus (Tableau 4). Il est possible d'obtenir partiellement ce phénotype par ARNi contre *alg-1* (selon la nomenclature, ils seront appelés *alg-1(ARNi)*) sur des vers sauvages. Le défaut occasionné au niveau de la production d'alaes, ces structures cuticulaires du stade adulte, est présent chez 45% des vers analysés. Il s'agit bien souvent de rupture d'alaes, mais aucune absence n'a été reportée. D'autres problèmes comme une croissance ralentie et une rétention d'œufs, allant jusqu'à l'éclosion *in utero* (Fig.31a), ont été observés. Récemment,

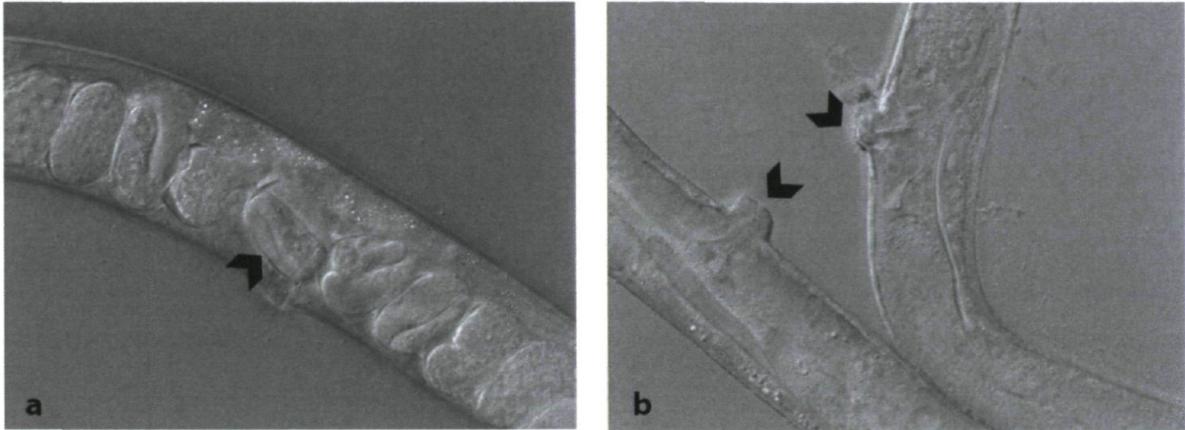
notre groupe de recherche a démontré qu'une baisse du nombre de descendants était associée à la souche *alg-1* (Bukhari et al., 2012). Les vers *alg-2* sont eux exempts de tout phénotype apparent. Ils ne présentent aucun des défauts mentionnés ci-dessus.

**Tableau 4 : Phénotypes et interaction génétique entre les gènes *alg-1* et *alg-2*.** Les souches N2, *alg-1*, *alg-2* ont été utilisées seules ou avec ARNi. La létalité, la survie et l'intégrité des alae a été vérifiée. Les défauts d'alae correspondent à des ruptures ou des absences. La létalité est associée à une rupture de la vulve où un arrêt de développement est indiqué. N correspond au nombre d'individus analysés.

|                                    | <u>% Létalité</u> | <u>% Survie</u> | <u>N</u> | <u>% Défaut alae</u> | <u>N</u> |
|------------------------------------|-------------------|-----------------|----------|----------------------|----------|
| N2                                 | -                 | 100             | 47       | -                    | 24       |
| <i>alg-1</i>                       | 19                | 80              | 54       | 45                   | 66       |
| N2; <i>alg-1</i> (ARNi)            | 9                 | 91              | 42       | 37                   | 28       |
| <i>alg-2</i>                       | -                 | 100             | 35       | -                    | 23       |
| N2; <i>alg-2</i> (ARNi)            | -                 | 100             | 49       | -                    | 19       |
| <i>alg-1</i> ; <i>alg-2</i> (ARNi) | 100               | -               | 24       | pas d'alae           | 17       |
| <i>alg-2</i> ; <i>alg-1</i> (ARNi) | 100               | -               | 25       | pas d'alae           | 15       |

Cependant, le double mutant est embryoniquement létal (Vasquez-Rifo et al., 2012). Cette létalité est récapitulée et est visible au stade postembryonnaire sur les souches *alg-1*; *alg-2*(ARNi) et *alg-2*; *alg-1*(ARNi). Une létalité de 100%, par arrêt de développement au stade larvaire, est alors observée (Tableau 4). Ces vers ne se rendent jamais au stade adulte et par conséquent, aucun des vers analysés ne développent d'alae. Ils meurent prématurément par éclatement de la vulve (Fig.31b) ou arrêt de développement. Au niveau moléculaire, une baisse drastique des miARN, *let-7* et *lin-4* notamment, a été reportée dans le double mutant *alg-1*; *alg-2*(ARNi) (Grishok et al., 2001).

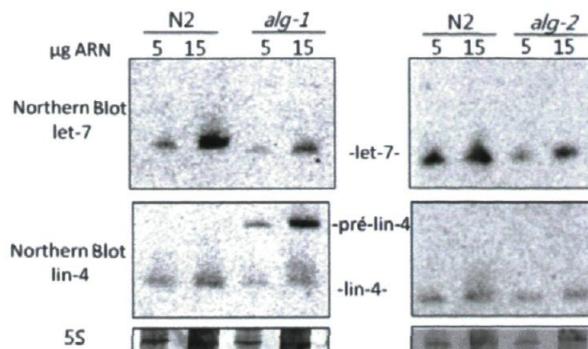
**Figure 31: Phénotypes observés chez les mutants *alg-1* et *alg-1; alg-2(ARNi)*.** a) Phénotype de rétention d'œufs du simple mutant *alg-1*. La flèche indique un œuf ayant éclos dans l'utérus de la mère. b) *alg-1; alg-2(ARNi)* stérile. Les flèches indiquent une protubérance anormale de la vulve pour le mutant *alg-1; alg-2(ARNi)* et une vulve rompue entraînant la mort du ver.



## 1.2. Dépendance des miARN *let-7* et *lin-4* vis-à-vis de *ALG-1*

Les phénotypes observés pour le mutant *alg-1* sont étroitement liés aux phénotypes observés chez différents mutants de miARN, notamment les souches mutantes pour *let-7*. Ainsi, on observe une mortalité par rupture de la vulve et un défaut d'œuf pour la souche thermosensible *let-7(n2853)*, à température non-permissive de 20°C (Reinhart et al., 2000). Cette liaison phénotypique est visible au niveau moléculaire lors de la biogenèse des miARN où, pour une souche *alg-1*, on détecte une baisse du miARN *let-7* mature comparé à N2, au stade adulte (Grishok et al., 2001). Concernant *lin-4*, *alg-1* affiche une baisse du niveau de miARN mature et une accumulation de la forme pré-*lin-4*. Ceci suggère fortement que l'absence de la protéine ALG-1 provoque des défauts dans la maturation de ces deux miARN. Ces résultats ont été confirmés par analyse Northern Blot (Fig.32).

**Figure 32 : Analyse des miARN let-7 et lin-4 par Northern Blot.** Des extraits d'ARN totaux de vers N2, *alg-1* et *alg-2* d'une population mixte de vers, ont été utilisés à raison de 5 et 15 µg. La quantité d'ARN est contrôlée par comparaison avec l'ARN 5S.



Génétiquement, nous avons ensuite testé l'interaction entre les gènes *alg-1* et *alg-2*, et les gènes *let-7* et *lin-4*, par des expériences d'ARNi en utilisant les souches mutantes *let-7(n2853)* et *lin-4(e912)*. Les résultats ont démontré une synergie évidente pour *alg-1* alors que la perte de la protéine ALG-2 ne semble pas affecter ces deux mutants (Tableau 5). *let-7(n2853)*, à la température non-permissive de 20°C, présente environ 83% de mortalité par explosion de la vulve. Ce phénotype est moins pénétrant à 15°C où seulement 24% des vers sont affectés. Pour les vers *let-7(n2853); alg-1(ARNi)*, à température permissive de 15°C, on voit une synergie reflétée par un phénotype plus fort, jusqu'à 87% de mortalité. À 20°C, tous les vers meurent. Chez *let-7(n2853); alg-2(ARNi)*, il n'y a pas d'augmentation de la force du phénotype à 15°C et 20°C, où respectivement 19% et 87% des vers sont affectés. Les expériences sur *lin-4(e912)* sont plus délicates étant donné que la souche est très malade. Elle croit uniquement à 15°C et elle est sans vulve. Il n'y a donc pas de ponte et les œufs éclosent dans les adultes et sortent par lyse de ceux-ci. Néanmoins, tout comme pour *let-7*, le phénotype de létalité est amplifié par la perte de ALG-1 où tous les vers sont affectés par arrêt précoce de développement. La perte de ALG-2 n'a pas d'effet apparent. Ceci indique un lien génétique entre ALG-1 et les miARN *let-7* et *lin-4*.

**Tableau 5 : Liaison génétique des gènes *alg-1*, *alg-2*, *let-7* et *lin-4*.** Les expériences ont été faites à 15°C et/ou 20°C et des larves en L1 ont été en contact avec l'ARNi pendant 48 à 72 h. Les vers affectés correspondent à une létalité larvaire par explosion de la larve pour *let-7* et arrêt de développement pour *lin-4*. N correspond au nombre d'individus analysés.

|                                  | 15°C       |                |    | 20°C       |                |    |
|----------------------------------|------------|----------------|----|------------|----------------|----|
|                                  | % affectés | % non affectés | N  | % affectés | % non affectés | N  |
| <i>let-7(n2853)</i>              | 24         | 76             | 39 | 83         | 17             | 45 |
| <i>let-7(n2853); alg-1(ARNi)</i> | 87         | 13             | 47 | 100        | -              | 56 |
| <i>let-7(n2853); alg-2(ARNi)</i> | 19         | 81             | 51 | 87         | 13             | 43 |
| <i>lin-4(e912)</i>               | -          | 100            | 32 |            |                |    |
| <i>lin-4(e912); alg-1(ARNi)</i>  | 100        | -              | 35 |            |                |    |
| <i>lin-4(e912); alg-2(ARNi)</i>  | -          | 100            | 28 |            |                |    |

## 2. Analyse phénotypique des lignées transgéniques *alg-1* et *alg-2* sauvages et mutantes catalytiques

Pour pouvoir étudier l'effet d'une inactivation de l'activité endonucléasique des protéines Argonaute *in vivo*, les deux souches mutantes nulles décrites précédemment ont été utilisées afin de pouvoir créer quatre lignées transgéniques. Dans un mutant *alg-1*, nous avons intégré un ADN plasmidique exprimant une copie génomique d'*alg-1* sauvage (*alg-1* WT) ou *alg-1* muté (*alg-1* 3a) sous la dépendance du promoteur et de la région 3'UTR d'endogène. Pour *alg-2*, nous avons suivi la même stratégie, avec l'intégration d'ADN plasmidique portant une copie génomique d'*alg-2* sauvage (*alg-2* WT) ou *alg-2* muté (*alg-2* 3a), sous la dépendance des séquences régulatrices d'*alg-2*, intégré dans une souche *alg-2*. Étant donné que nous ne disposons pas d'anticorps pour ALG-2, une étiquette GFP en position N-terminal a été introduite pour les analyses biochimiques.

### 2.1. Analyse de la restitution de fonction des lignées transgéniques *alg-1* et *alg-2*

Pour *alg-1* WT, il y a une très forte restitution de la fonction. Le ver est phénotypiquement sauvage, au niveau de la croissance, de la mobilité et de la résistance à plus haute température (25°C). Cependant, la synthèse des alae demeure déficiente (37% d'alaes

défectueux pour N=51). Pour *alg-2* WT, cette souche présente quelques défauts visibles de croissance, notamment à 25°C. Cependant, aucun défaut d'alae n'a été détecté. Pour les mutants catalytiques, *alg-1* 3a est très affecté. Les vers sont de très petite taille, ont une croissance lente et sont très affectés à 25°C. De plus, ce mutant présente des défauts au niveau des mues à chaque stade larvaire et entre dans une forme léthargique. Ce défaut est souvent associé à la perte de fonction de certains miARN comme ceux de la famille *let-7* (Abbott et al., 2005). La lignée *alg-2* 3a est phénotypiquement proche du mutant *alg-2* et aucun défaut dans la croissance ou le développement n'a été détecté. Par comparaison entre les transgènes 3a, on peut conclure que les défauts sévères affichés de la protéine ALG-1 muté par rapport à ALG-2 muté montrent un rôle prédominant de l'activité catalytique de la protéine ALG-1. Le fait que l'introduction du transgène *alg-2* WT apporte des défauts développementaux alors qu'ils sont inexistantes dans les souches mutantes *alg-2* et *alg-2* 3a apporte un nouvel élément.

## **2.2. Analyse de la perte de *alg-1* et *alg-2* endogène par ARNi sur les lignées transgéniques**

Pour établir si les protéines ALG-1 3a et ALG-2 3a sont capables de soutenir la voie des miARN, nous avons testé si la perte de fonction de la protéine Argonaute endogène restante est létale pour le ver. Dans ce contexte, seule l'expression du transgène supportera la voie des miARN, sauvage ou mutant catalytique. Pour inhiber l'Argonaute endogène, nous avons utilisé l'ARNi, car les essais de croisement ont été infructueux pour une raison principale : l'obtention de mâle fonctionnel pour le mutant *alg-1* et pour *alg-1* 3a n'est pas possible.

Les résultats sont résumés dans le Tableau 6. Pour les souches transgéniques WT, les expériences d'ARNi ont montré que la souche *alg-1* WT; *alg-2*(ARNi) est viable. Les défauts affichés sont minimes et seuls 14% des vers sont affectés. La moitié des vers affectés meurent par arrêt de développement au stade larvaire (létalité larvaire), alors que l'autre moitié des vers meurent par rupture de la vulve. Cela prouve que le transgène restitue la fonction de *alg-1* endogène, mais la restitution n'est pas totale. Pour *alg-2* WT; *alg-1*(ARNi), la restitution de fonction est beaucoup plus faible puisque, dans notre

expérience, 72% des vers sont affectés. Néanmoins, 28% des vers n'affichent aucun phénotype particulier. Cette restitution de fonction faible peut être expliquée par le rôle prédominant de *alg-1* par rapport à *alg-2*. De plus, il est peu probable que l'étiquette GFP affecte la fonction de ALG-2 étant donné qu'au laboratoire une souche GFP-ALG-1 restitue pleinement la fonction de cette protéine.

En ce qui concerne les souches mutées *alg-1 3a; alg-2(ARNi)* et *alg-2 3a; alg-1(ARNi)*, l'effet est très clair. Lorsque seule une protéine Argonaute mutée est exprimée, la totalité des vers est affectée. Respectivement, 27 et 28% des vers ont d'importants défauts de développement, mais sont viables. Cependant, ces vers n'atteignent jamais le stade adulte et sont par conséquent stériles. Pour *alg-1 3a; alg-2(ARNi)*, la mortalité observée correspond à un arrêt développemental au stade larvaire pour certains individus (18%) et la grande majorité d'entre eux meurt par rupture de la vulve. Pour la souche *alg-2 3a; alg-1(ARNi)*, la majorité des vers meurent par arrêt de développement (72%). Ces phénotypes sont similaires à ceux observés pour les doubles mutants de protéines Argonaute (Fig.31b). Ceci nous conduit à la conclusion qu'au moins une protéine Argonaute avec une fonction catalytique est nécessaire pour le développement et la viabilité des nématodes. Parallèlement à ces expériences, l'ARNi contrôle, exprimant un ARN db sans cible, a été effectué (Tableau 6c). Ces résultats appuient l'analyse phénotypique des lignées transgéniques et la restitution de la fonction des Argonaute. On observe une certaine létalité pour les lignées *alg-1 3a*, *alg-2 WT* et *alg-2 3a*, qui n'est cependant sans commune mesure avec celle observée chez *alg-1 3a; alg-2(ARNi)* et *alg-2 3a; alg-1(ARNi)*.

**Tableau 6 : Expérience d'ARNi sur les différentes lignées transgéniques.** a) ARNi contre *alg-2* sur N2, *alg-1*, *alg-1* WT et *alg-1* 3a. b) ARNi contre *alg-1* sur N2, *alg-2*, *alg-2* WT et *alg-2* 3a. c) ARNi contrôle avec le vecteur contrôle sur les différentes souches citées précédemment. Le pourcentage de vers affectés est indiqué en gris. Les différents phénotypes comptabilisés sont la stérilité, la létalité larvaire, la létalité par rupture de la vulve et la rétention d'œufs. La section autre phénotype pour c) correspond à des vers *alg-2* WT malades pour lequel les phénotypes classiques d'affection de la voie des miARN n'ont pas été détectés.

**a** *alg-2(ARNi)*

|                 | vers affectés % | stérilité % | létalité larvaire % | rupture vulve % | rétention d'œufs % | N   |
|-----------------|-----------------|-------------|---------------------|-----------------|--------------------|-----|
| N2              | 0               | -           | -                   | -               | -                  | 126 |
| <i>alg-1</i>    | 100             | -           | 100                 | -               | -                  | 151 |
| <i>alg-1</i> WT | 14              | 0           | 7                   | 7               | -                  | 235 |
| <i>alg-1</i> 3a | 100             | 27          | 18                  | 55              | -                  | 207 |

**b** *alg-1(ARNi)*

|                 | vers affectés % | stérilité % | létalité larvaire % | rupture vulve % | rétention d'œufs % | N   |
|-----------------|-----------------|-------------|---------------------|-----------------|--------------------|-----|
| N2              | 20              | 0           | 0                   | 10              | 10                 | 226 |
| <i>alg-2</i>    | 100             | 0           | 100                 | -               | -                  | 153 |
| <i>alg-2</i> WT | 72              | 22          | 50                  | -               | -                  | 176 |
| <i>alg-2</i> 3a | 100             | 28          | 72                  | -               | -                  | 200 |

**c** *contrôle(ARNi)*

|                 | vers affectés % | autres phénotypes % | létalité larvaire % | rupture vulve % | rétention d'œufs % | N   |
|-----------------|-----------------|---------------------|---------------------|-----------------|--------------------|-----|
| N2              | 0               | -                   | -                   | -               | -                  | 130 |
| <i>alg-1</i>    | 50              | -                   | -                   | 14              | 36                 | 125 |
| <i>alg-1</i> WT | 0               | -                   | -                   | -               | -                  | 122 |
| <i>alg-1</i> 3a | 10              | -                   | 10                  | -               | -                  | 157 |
| <i>alg-2</i>    | 4               | 4                   | -                   | -               | -                  | 130 |
| <i>alg-2</i> WT | 7               | -                   | 7                   | -               | -                  | 137 |
| <i>alg-2</i> 3a | 14              | -                   | 14                  | -               | -                  | 111 |

### 2.3. Analyse biochimique des lignées transgéniques *alg-1* et *alg-2* sauvages et mutantes catalytiques

Dans le but de comprendre le phénotype de létalité observé chez lignées mutantes *alg-1* 3a; *alg-2(ARNi)* et *alg-2* 3a; *alg-1(ARNi)*, nous avons ciblé plusieurs étapes dans lequel les protéines ALG-1 et ALG-2 peuvent intervenir. Ainsi, nous avons étudié le processus de maturation des miARN, la formation de complexe Argonaute-miARN et la régulation de certains ARNm régulés par des miARN.

### 2.3.1. Effets de la perte d'activité endonucléasique de *alg-1* et *alg-2* sur la maturation des miARN

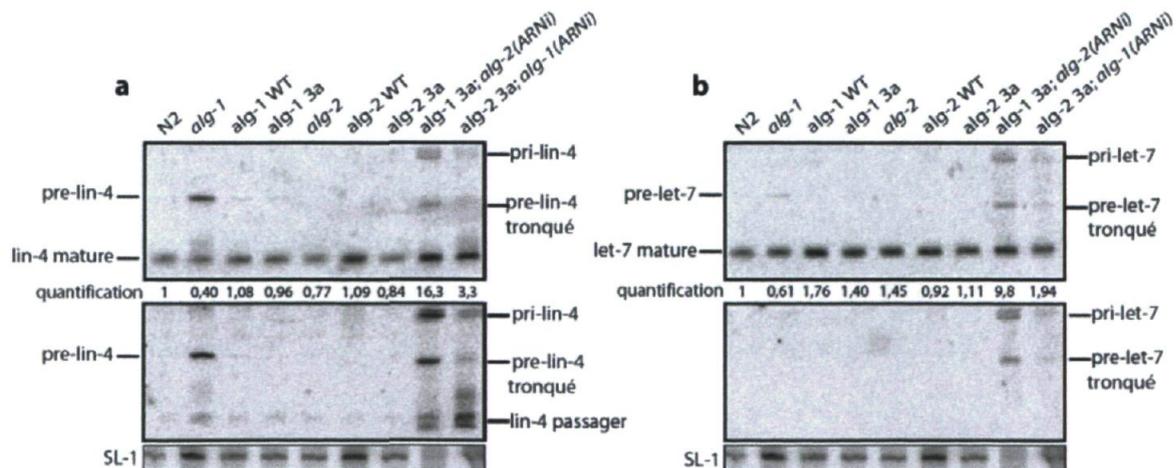
#### 2.3.1.1. Les miARN *let-7* et *lin-4*

Nous avons analysé, par des expériences de type Northern Blot, si la synthèse des miARN était affectée dans ces différentes souches. Nous avons tout d'abord vérifié l'intégrité du processus de maturation des miARN *lin-4* et *let-7* (Fig.33), lesquels affichaient des défauts de maturation chez la souche *alg-1* (Fig.32). Ces deux miARN ont la particularité d'être positionné sur le bras 5' de leurs précurseurs respectifs. Pour les différentes expériences de Northern Blot, nous avons utilisé «Spliced Leader 1» (SL-1) comme référence. Il s'agit d'une séquence d'ARN intervenant dans l'épissage en trans et ajouté en 5' d'environ 70% des ARNm chez le nématode (Conrad et al., 1991). Comme le montre la Figure 33a, toutes les lignées transgéniques restituent bien le patron observé pour les vers sauvages N2. Les quantifications donnent lieu à des niveaux de miARN mature semblables pour *alg-1* WT et *alg-1* 3a (Fig.33a, panel du haut). Par conséquent, on peut conclure que l'expression de ces transgènes, qu'ils soient sauvages ou catalytiques mutants, rétablit la maturation du miARN *lin-4* à un niveau comparable à la souche sauvage. De plus, le précurseur n'est plus détectable dans *alg-1* WT et *alg-1* 3a. Bien que ce soit plus faible, on observe une baisse du mature pour le mutant *alg-2*, et ce niveau est rétabli pour les lignées transgéniques *alg-2* WT et *alg-2* 3a (Fig.33a, panel du haut). Ces résultats nous indiquent qu'en présence des deux protéines Argonaute, si seulement une seule est mutante 3a, on observe une restitution complète de la maturation du miARN *lin-4*. Dans la condition létale, pour les vers *alg-1* 3a; *alg-2*(ARNi) et *alg-2* 3a; *alg-1*(ARNi) (Fig.33a, panel du haut), nous voyons que les niveaux de miARN mature sont plus élevés que pour le sauvage N2 (16,3 et 3,3). On détecte également une forme précurseur plus petite que la forme détectée dans la souche *alg-1* (*pré-lin-4* tronqué) et une forme de taille similaire à la forme primaire (*pri-lin-4*), formes qui ne sont pas détectables dans les autres conditions. Cela démontre que dans la condition où seule une protéine Argonaute catalytiquement inactive est exprimée, la maturation du miARN est aberrante, avec des niveaux hauts de miARN mature et la détection de forme précurseur inattendue. Dans ces animaux, on observe aussi une stabilisation du brin passager (Fig.33a, panel du bas), alors qu'il est presque totalement

indécelable dans le sauvage et dans toutes les autres souches, excepté pour le mutant *alg-1*. Il est également intéressant de noter que la forme pré-lin-4 tronquée est également détectable dans cette condition.

Concernant *let-7*, le processus de maturation est affecté de manière significative dans le mutant *alg-1* avec une baisse du niveau de miARN mature (Fig.32). Chez toutes les lignées, par l'expression de transgènes catalytiques mutants ou sauvages, nous voyons une restitution du niveau de *let-7* mature (Fig.33b, panel du haut) et, dans la plupart des cas, il y a même une augmentation comparée au niveau des animaux sauvages N2. Cette augmentation appuie encore le fait qu'une fine régulation des protéines Argonaute pour la biogenèse des miARN est nécessaire, *let-7* étant peut-être plus sensible que *lin-4* à ce niveau. Les mutants *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)* affichent le même patron que *lin-4* avec notamment une accumulation de mature (9,8 et 1,94) et la détection des formes primaires (*pri-let-7*) et d'un précurseur de taille plus courte que la normale (*pré-lin-4 tronqué*). À la différence de *lin-4*, le brin passager de *let-7* n'est pas détectable (Fig.33b, panel du bas), suggérant une stabilité faible de ce passager, mais la forme *pré-let-7 tronquée* est visible.

**Figure 33 : Analyse des miARN *lin-4* (a) et *let-7* (b) par Northern Blot.** Dix µg d'extrait d'ARN total de vers synchronisés aux stades adultes ont été utilisés. L'ARN SL-1 a été utilisé comme contrôle endogène pour établir la quantification qui est relative à N2. Les panels du haut détectent le brin miARN mature, les panels du bas les brins passagers. Les différentes formes mature, précurseur (pré), précurseur primaire (pri) et précurseur tronqué (pré tronqué) sont indiqués. Les génotypes sont indiqués ainsi que le type d'ARNi utilisé, le cas échéant.



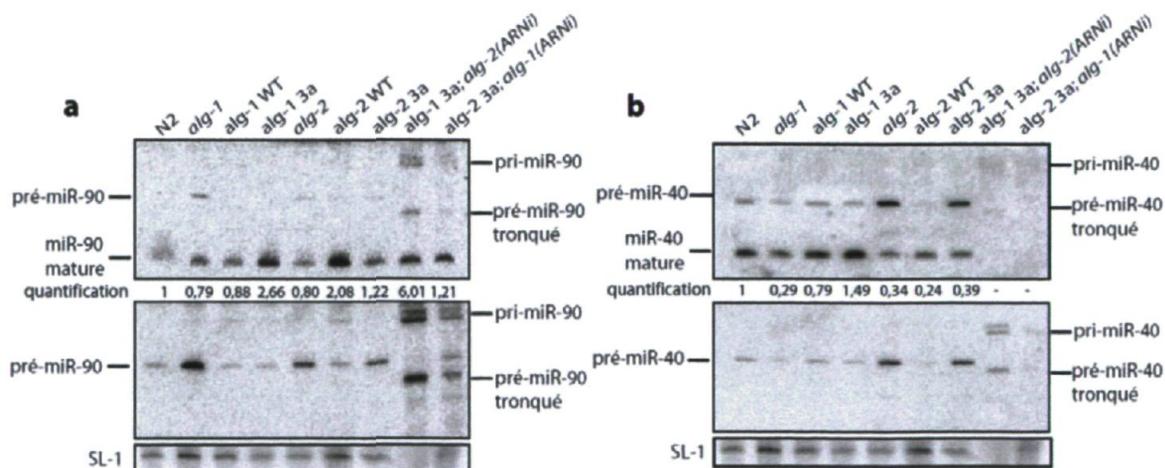
### 2.3.1.2. Les miARN miR-90 et miR-40

Nous nous sommes ensuite intéressés à la maturation des miARN dont la forme mature se trouvait sur le brin 3' du précurseur. Nous avons choisi les miARN miR-90 et miR-40 pour cette caractéristique, mais également pour le fait que la tige boucle de leur forme précurseur possède un haut degré de complémentarité. miR-40 fait partie de la famille miR-35~41, impliquée dans le contrôle du développement embryonnaire, et il est essentiellement exprimé à ce stade de développement (Alvarez-Saavedra and Horvitz, 2010). miR-90 fait partie de la famille miR-50 et est exprimé dans tous les stades embryonnaires (Martinez et al., 2008). Comme le montre la figure 34, nous observons un patron différent de *let-7* et *lin-4*. Pour miR-90 (Fig.34a, panel du haut), dans les mutants nuls *alg-1* et *alg-2*, nous détectons une légère baisse de miARN mature. Cette diminution est rétablie par l'addition de tous les types de transgènes. Cependant, pour les transgènes *alg-1* 3a (2,66) et *alg-2* WT (2,08), une forte accumulation est observée tandis que les niveaux sont proches du niveau de N2 pour *alg-1* WT (0,88) et *alg-2* 3a (1,22). Il est intéressant de noter que l'accumulation du précurseur affiché dans le mutant *alg-1* est rétablie pour les lignées *alg-1* WT et *alg-1* 3a. Cette forme est détectable pour les deux mutants nuls bien que le niveau est très élevé pour *alg-1* par rapport à *alg-2*, et que la détection soit plus forte avec la sonde contre le passager (Fig.34a, panel du bas). Cependant, alors que cette accumulation disparaît à la fois pour les deux transgènes sauvages et mutants d'*alg-1*, elle ne disparaît que pour le transgène d'*alg-2* WT et ceci, indépendamment du niveau de mature. On peut en conclure, dans ce cas précis, que l'activité endonucléasique de ALG-2 est requise pour rétablir le niveau de précurseur puisque le transgène *alg-2* 3a affiche une accumulation similaire aux mutants nuls *alg-1* et *alg-2*. À l'opposé, la présence de la protéine ALG-1 est nécessaire, et le fait qu'elle soit possède ou non le motif DDH rétablit le niveau de précurseurs. Dans le cas des souches *alg-1* 3a; *alg-2*(ARNi) et *alg-2* 3a; *alg-1*(ARNi), on observe néanmoins le même patron que pour les autres miARN testés avec une accumulation de miARN mature et la détection des formes pri-miR-90 et pré-miR-90 tronquée.

En ce qui concerne miR-40, il présente la particularité que la forme précurseur est bien détectable pour toutes les souches testées (Fig.34b, panel du haut). De plus, les mutants

nuls *alg-1* et *alg-2* présentent des défauts sévères dans la synthèse de miARN mature. Pour les transgènes *alg-1* WT et *alg-1* 3a, le niveau est rétabli bien qu'une baisse légère soit observable pour le premier (0,79) et qu'une augmentation soit détectable pour le second (1,49). En revanche, les transgènes *alg-2* WT et 3a ne rétablissent pas les niveaux de mature (0,24 et 0,39). Cela montre encore une fois que le transgène *alg-2* WT ne restitue pas pleinement la fonction de l'endogène comparativement au transgène *alg-1* WT. Cependant, en ce qui concerne l'accumulation de précurseur, on observe le même phénomène que pour miR-90, où l'on observe une accumulation dans *alg-2* et dans *alg-2* 3a. L'analyse des vers *alg-1* 3a; *alg-2*(ARNi) et *alg-2* 3a; *alg-1*(ARNi) montre cette fois-ci un patron différent des autres miARN. Alors que les formes pri-miR-40 et pré-miR-40 sont détectables pour *alg-1* 3a; *alg-2*(ARNi), la forme mature n'est pas présente pour ce miARN, ceci s'expliquant certainement par l'expression de ce miARN, qui est embryonnaire. En effet, étant donné qu'il est présent uniquement dans les embryons et dans les adultes gravides, les doubles mutants n'atteignent jamais ce dernier stade de développement car ils sont stériles.

**Figure 34 : Analyse des miARN miR-90 (a) et miR-40 (b) par Northern Blot.** Dix  $\mu$ g d'extrait d'ARN total de vers synchronisés aux stades adultes ont été utilisés. L'ARN SL-1 a été utilisé comme contrôle endogène pour établir la quantification qui est relative à N2. Les panels du haut détectent le brin miARN mature, les panels du bas les brins passagers. Les différentes formes mature, précurseur (pré), précurseur primaire (pri) et précurseur tronqué (pré tronqué) sont indiqués. Les génotypes sont indiqués ainsi que le type d'ARNi utilisé, le cas échéant.

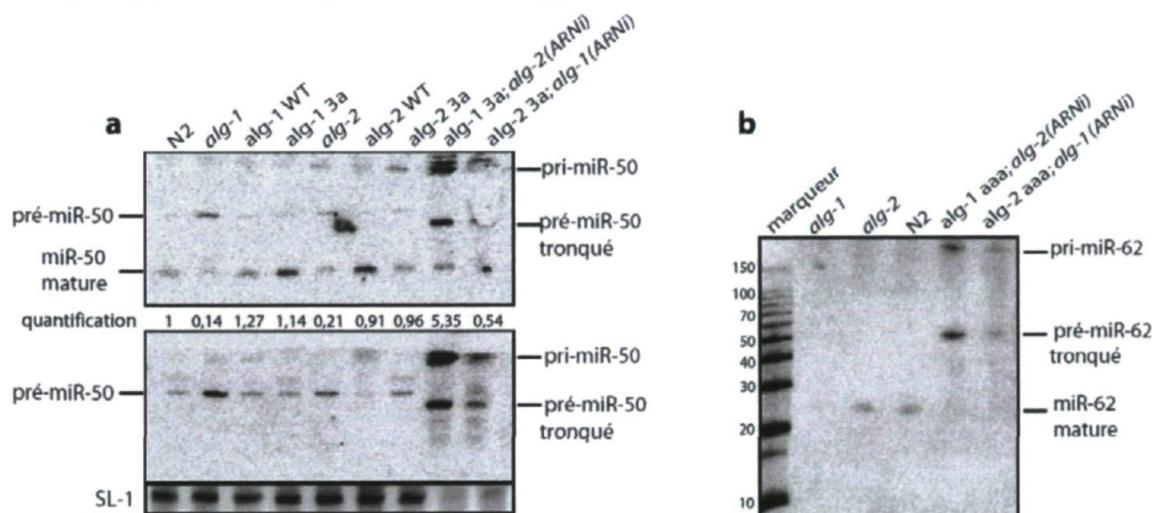


### 2.3.1.3. Les miARN introniques miR-50 et miR-62

Afin de déterminer si les formes primaires et précurseurs tronqués dans les lignées *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)* sont détectables pour une large population de miARN, nous avons analysé deux miARN présents dans les introns. Le premier est miR-50, le représentant type de la famille miR-50, est présent dans un intron du gène Y71G12B.11a (Miska et al., 2007). Le début de la séquence du précurseur étant très éloigné du site d'épissage, ce miARN n'est pas considéré comme un mirtron car il requiert certainement l'intervention de Drosha/Pasha pour sa maturation (Isik et al., 2010). D'autre part, nous avons choisi un mirtron confirmé, miR-62, inclus dans un intron du gène T07C5.1, gène codant pour une UDP-Glucuronosyl Transférase. L'intron dans lequel il se trouve, une fois épissé, peut directement être utilisé par l'enzyme DCR-1 et ne requiert pas le couple Drosha/Pasha (Isik et al., 2010). Le niveau de miR-50 mature se trouve sévèrement affecté pour les deux mutants nuls *alg-1* (0,14) et *alg-2* (0,21) (Fig.35a, panel du haut) mais est rétabli à des niveaux similaires aux animaux sauvages N2 dans toutes les lignées transgéniques utilisées. L'accumulation du précurseur est conséquente pour *alg-1*, plus faible pour *alg-2* et est détectable pour les différentes lignées, notamment avec la sonde dirigée contre miR-50 passager (Fig.35a, panel du bas), excepté pour la lignée *alg-2* WT. Comme observé pour miR-40 et miR-90, ce résultat suggère un besoin de l'activité de la protéine rALG-2 pour éviter cette accumulation. Pour les lignées *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*, on observe les mêmes défauts dans la synthèse que les autres miARN précédemment étudiés. Le fait que le miARN se trouve dans un intron n'influence donc pas la dépendance d'une activité endonucléasique pour sa maturation.

Le miARN miR-62 étant très faiblement abondant (Warf et al., 2011), nous avons enrichi les échantillons en courts ARN et 50 µg ont été utilisés afin d'avoir un signal. La synthèse de ce miARN, indépendant de Drosha, est totalement abolie chez les souches *alg-1*, *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)* (Fig.35.b). Il n'est en revanche pas affecté dans le mutant *alg-2*. Chez *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*, la détection des formes pri-miR-62 et pré-miR-62 tronqué prouve que les produits observés ne sont pas dus à une interférence avec l'activité de Drosha. De plus, le marqueur de taille nous confirme que le précurseur tronqué est de 50 nt.

**Figure 35 : Analyse des miARN miR-50 (a) et miR-62 (b) par Northern Blot.** a) Dix  $\mu$ g d'extrait d'ARN total de vers synchronisés aux stades adultes ont été utilisés. L'ARN SL-1 a été utilisé comme contrôle endogène pour établir la quantification qui est relative à N2. Les différentes formes mature, précurseur (pré), précurseur primaire (pri) et précurseur tronqué (pré tronqué) sont indiquées. Les génotypes sont indiqués ainsi que le type d'ARNi utilisé, le cas échéant. Les panels du haut détectent le brin miARN mature, le panel du bas les brins passagers. b) 50  $\mu$ g de courts ARN purifiés ont été utilisés pour détecter miR-62. Les différentes formes mature, précurseur primaire (pri) et précurseur tronqué (pré tronqué) et les génotypes sont indiqués ainsi que le type d'ARNi.



## 2.4. Détermination du site de troncation du précurseur de miARN

Nous avons vu que pour tous les miARN testés, nous étions en mesure de détecter une forme tronquée de précurseur de miARN. Cette forme tronquée est détectable à la fois avec les sondes dirigées contre les miARN matures et passagers, ce qui indique qu'en fait il y a deux formes tronquées : celles détectables avec les sondes appariant les miARN matures contiennent le brin mature et celles détectables avec les sondes appariant les miARN passagers contiennent le brin passager. Chez les souches *alg-1* WT; *alg-2*(ARNi) ou *alg-2* WT; *alg-1*(ARNi), la troncation du pré-miARN n'est jamais détectable pour tous les miARN testés comme le montre l'exemple de miR-50 (Fig.36a). Comparé à un marqueur de taille, le pré-miR-50 tronqué est de 50 nt (Fig.36a) et la taille de pré-miR-50 d'environ 65-70 nt (Fig.36b). La troncation est donc de 15 à 20 nt.

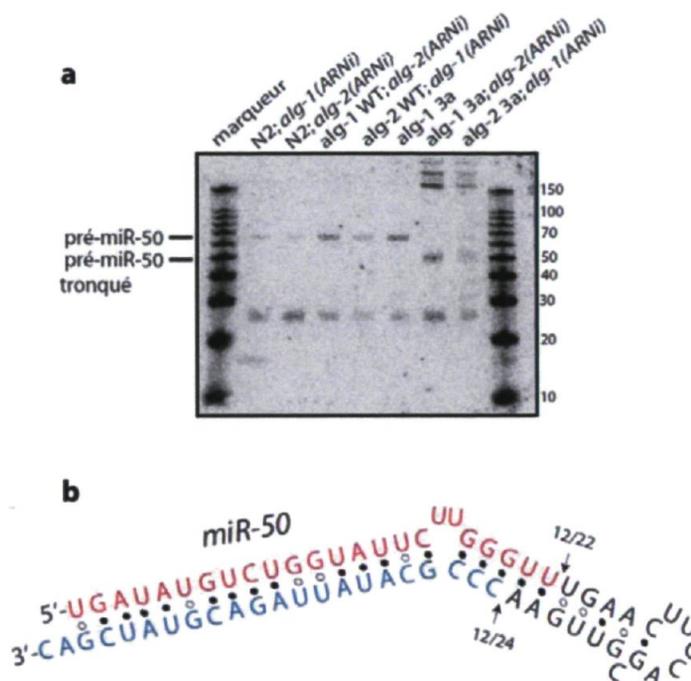
Par des techniques de 5' et 3' RACE, nous avons cloné le pré-miR-50 tronqué, car au niveau des analyses Northern, il paraissait le plus abondant. La stratégie utilisée est la suivante. Nous avons tout d'abord purifié sur colonne miRVana les courts ARN inférieurs à

200 pb, afin d'éliminer les ARNm. L'extrait d'ARN utilisé est alg-1 3a; *alg-2(ARNi)*. Nous avons ensuite ligué en 3' ou en 5', un ARN adaptateur (ARNad) et deux conditions ont été utilisées pour l'ARNad : phosphorylé (ARNadP) en 5' ou non phosphorylé (ARNad). Il faut savoir qu'une réaction de ligation par la T4 ARN ligase s'effectue uniquement entre une extrémité 5'P et une extrémité 3'OH. Nous avons donc utilisé deux réactions de ligation, l'une avec l'ARNad qui va liquer toutes les molécules d'ARN ayant une extrémité 5'P et l'autre avec l'ARNadP qui va joindre les extrémités 3'OH. Les ARNad et ARNadP vont servir de séquence de référence pour déterminer les extrémités des molécules d'ARN isolés. Après ligation, une réaction de transcription inverse a été effectuée avec différents oligonucléotides (Matériels et méthodes), suivi de deux réactions de PCR pour amplifier le produit de ligation. Les différents fragments correspondants aux tailles attendues de pré-miR-50, pré-miR-50 et pré-miR-50 tronqué ont été purifiés sur gel, clonés et séquencés.

Pour les tailles correspondantes aux pré-miR-50, nous avons pu cloner la totalité de ce précurseur. Sur 20 clones séquencés de la réaction de 5' RACE, 9 correspondaient à la séquence exacte du précurseur, 6 était amputé de un ou deux nucléotides et les 5 autres correspondaient à des séquences de pré-miR-50 amputé de 3 à 10 nt. La réaction de 3'RACE nous a donné à peu de chose près les mêmes résultats sur les différents clones envoyés à séquencer (données non montrées).

Pour les tailles du pré-miR-50 tronqué, nous avons pu obtenir approximativement le même niveau de précision que pour le précurseur. Nous avons été en mesure d'identifier pour la moitié des clones testés, à savoir 12/22 pour la 5'RACE et 12/24 pour la 3'RACE (Fig.36b), les séquences de pré-miR-50 tronqué qui correspondent au précurseur amputé du miARN miR-50 mature ou passager. Ces sites de coupures sont ceux qui sont initiés par l'enzyme DICER à la fois pour le bras 5' et le bras 3' du précurseur. Sur ce séquençage, 4/22 (5'RACE) et 7/24 (3'RACE) correspondaient à des séquences identiques à un ou deux nucléotides près, le reste des séquences était généralement amputé de plus larges portions supérieures à 5' nt et incluant parfois jusqu'à la boucle du précurseur. Nous avons donc identifié par ces expériences, avec une précision relative, le site de troncation. Étant donné qu'on détecte des formes pré-tronquées chez les miARN testés, nous pouvons penser que le site de troncation est conservé.

**Figure 36 : Détermination de la troncation de pré-miR-50.** a) Dix  $\mu\text{g}$  d'extrait d'ARN total de vers synchronisés aux stades adultes ont été utilisés. Les différentes formes mature, précurseur (pré), précurseur primaire (pri) et précurseur tronqué (pré tronqué) sont indiqués. Les génotypes sont indiqués ainsi que le type d'ARNi utilisé. b) Clonage de pré-miR-50 tronqué. Schéma de pré-mR-50 avec en rouge la séquence du miARN et en bleu la séquence du brin passager correspondant aux clivages théoriques de Dicer. En noir, la boucle potentiellement éliminée par Dicer est représentée. Le point entre les pb indique s'il y a appariement (point noir plein) ou non (cercle). Le résultat des expériences de clonage identifiant le pré-miR-50 tronqué est indiqué (12/22 5'RACE et 12/24 3'RACE).



## 2.5. Effets de la perte d'activité endonucléasique de *alg-1/2* sur l'association aux miARN

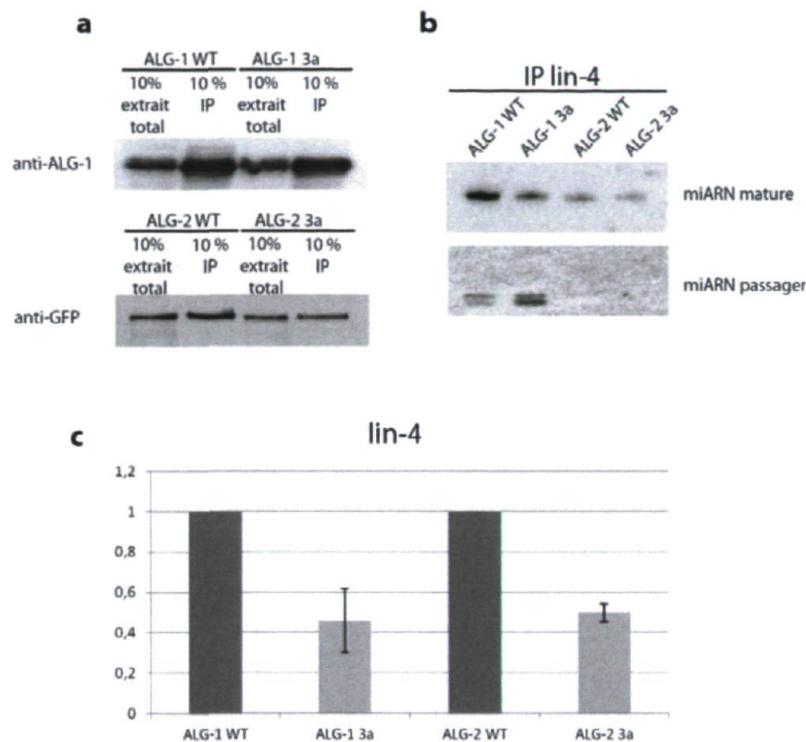
Chez *alg-1 3a; alg-2(ARNi)* et *alg-2 3a; alg-1(ARNi)*, il y a accumulation de pri-miARN, de pré-miARN tronqué, une accumulation forte de miARN mature et une accumulation de lin-4 passager. Bien que les phénotypes de mutants nuls pour *alg-1*, *alg-2* ou *Dicer* et ceux obtenus pour *alg-1 3a; alg-2(ARNi)* et *alg-2 3a; alg-1(ARNi)* soient similaires, dans le premier cas on observe une abolition ou une baisse de la synthèse de miARN mature (Grishok et al., 2001) tandis que pour *alg-1 3a; alg-2(ARNi)* et *alg-2 3a; alg-1(ARNi)*, la synthèse semble maintenue.

Pour mieux comprendre ce phénomène, nous avons examiné les miARN associées à ALG-1 ou ALG-2 dans les lignées *alg-1* WT; *alg-2(ARNi)*, *alg-1 3a*; *alg-2(ARNi)*, *alg-2* WT; *alg-1(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*. Dans ces conditions, nous pourrions comparer les WT versus 3a dans la condition où seule une protéine Argonaute est exprimée et vérifier l'association des miARN matures, passagers et des formes précurseurs, le cas échéant. Deux anticorps ont été utilisés pour cette immunoprécipitation (IP), un anticorps anti-ALG-1 purifié dans notre laboratoire (Vasquez-Rifo et al., 2012), et un anticorps commercial anti-GFP, pour immunoprécipiter GFP-ALG-2.

### **2.5.1. miARN associés aux protéines Argonaute chez *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)* : le cas particulier de *lin-4***

Le premier miARN que nous avons étudié par IPs est le miARN *lin-4*. Ce miARN, son passager tout comme les formes *pri-lin-4* et *pré-lin-4* tronqué s'accumulent chez les lignées *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)* (Fig.33a). Pour ALG-1 et ALG-2, les IP effectuées dans ces conditions montrent que l'association de miARN *lin-4* mature était supérieure dans le WT comparé au mutant 3a (Fig.37b, panel du haut), bien que les niveaux de protéines dans les extraits ou dans l'IP soient comparables (Fig.37a). La compilation de trois IP indépendantes pour le brin mature de *lin-4* a confirmé la tendance d'une association réduite pour ALG-1 3a et GFP-ALG-2 3a comparativement à ALG-1 WT et GFP-ALG-2 WT (Fig.37c). Concernant la détection du brin passager, l'association semble spécifique à ALG-1 avec notamment une accumulation pour ALG-1 3a (Fig.37b, panel du bas). Cependant, pour GFP-ALG-2 WT et GFP-ALG-2 3a, nous n'avons pas détecté le brin passager de *lin-4*, bien que l'accumulation soit visible chez *alg-2 3a*; *alg-1(ARNi)*. Enfin, à la fois pour l'IP ALG-1 3a et GFP-ALG-2 3a, nous n'avons pas été en mesure de détecter les formes *pré-lin-4*, *pré-lin-4* tronqué ou *pri-lin-4*. Cette donnée indique que ces formes précurseurs ne sont pas associées aux protéines ALG-1/2.

**Figure 37 : Expérience de détection de lin-4 associé aux protéines ALG-1 et GFP-ALG-2 en condition d'ARNi.** a) Analyse de type Western Blot contre les protéines ALG-1 et GFP-ALG-2 sur 10% de l'extrait total de protéines utilisées (4 mg) et 10% de l'IP. b) Northern blot sur lin-4 mature et lin-4 passager sur 90% des IP d'ALG-1 WT, 3a, GFP-ALG-2 WT et 3a c) Association de lin-4 mature à ALG-1 3a et ALG-2 3a relativement aux protéines WT par analyse Northern. L'expérience a été faite en triplicata et la barre d'erreur correspond à un intervalle de confiance de 95%.

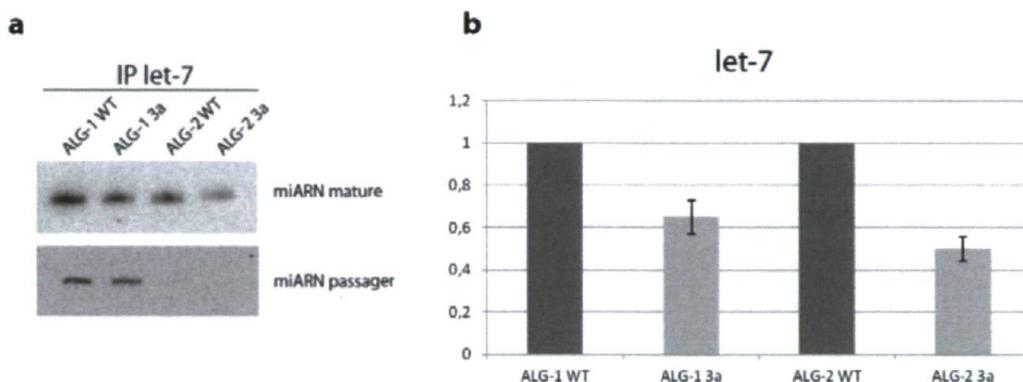


### 2.5.2. miARN associées aux protéines Argonaute chez *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)* : analyse de l'association de *let-7*

Concernant le miARN *let-7*, nous avons utilisé la même stratégie pour évaluer l'association de ce miARN aux protéines Argonaute chez *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*. Tout comme *lin-4*, il nous a été impossible d'immunoprécipiter les formes pri-*let-7*, pré-*let-7* et pré-*let-7* tronqué (Fig.38b). En revanche la tendance d'une association plus faible entre le mutant 3a et la protéine WT est conservée (Fig.38a, panel du haut) et la compilation de trois IP indépendantes confirme ce résultat (Fig.38b). Après une exposition prolongée, nous avons été en mesure de détecter le brin passager de *let-7* (Fig.38a, panel du bas). Tout comme *lin-4*, il n'est détectable que pour l'IP ALG-1 mais en revanche, il ne semble pas sujet à une accumulation dans la lignée *alg-1 3a*. Toutefois, il faut prendre en

considération que l'on observe une baisse du niveau de mature mais que le niveau détecté de passager semble similaire, ce qui peut suggérer qu'il y a une stabilisation du brin passager pour *alg-1 3a*. Des résultats similaires ont été obtenus pour miR-90 et miR-50 (données non-montrées).

**Figure 38 : Expérience de détection de *let-7* associé aux protéines ALG-1 et GFP-ALG-2 en condition d'ARNi.** a) Northern blot sur *let-7* mature et *let-7* passager sur 90% des IP d'ALG-1 WT, 3a, GFP-ALG-2 WT et 3a. b) Association de *let-7* mature à ALG-1 3a et ALG-2 3a relativement aux protéines WT par analyse Northern. L'expérience a été faite en triplicata et la barre d'erreur correspond à un intervalle de confiance à 95%.

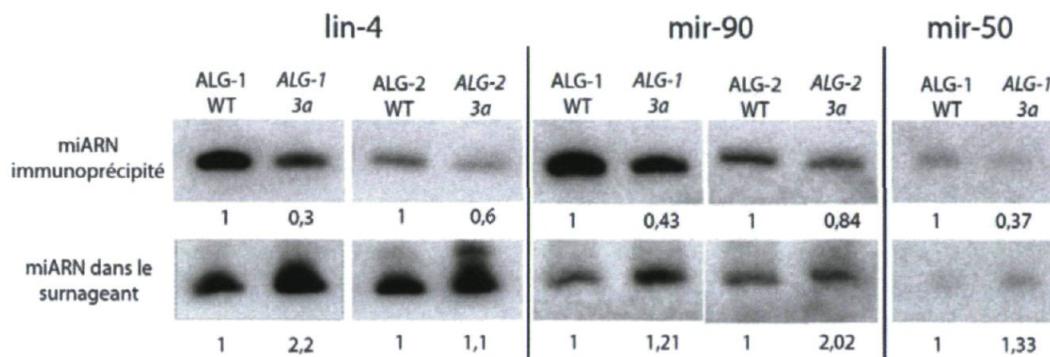


### 2.5.3. Investigation de l'accumulation de miARN chez *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*

Les résultats d'IP indiquent que chez *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*, on observe une baisse de miARN mature pour les différents miARN testés et parallèlement, l'analyse Northern sur l'ARN total démontre une accumulation de ces miARN matures. Dans le but de mieux comprendre ce point, nous avons analysé pour différents miARN le surnageant des IP faites sur ALG-1 et GFP-ALG-2. Comme le montre la Figure 39 pour *lin-4*, miR-90 et miR-50, ces différents miARN sont toujours plus abondants dans l'IP des protéines WT que dans celles des protéines mutantes 3a (panel du haut). Mais si l'on regarde les miARN restants dans les différents surnageants, on observe l'effet inverse, autrement dit une présence plus forte de miARN non immunoprécipité dans le surnageant des extraits protéiques des lignées mutantes 3a (panel du bas). Bien qu'il y ait accumulation de miARN mature pour les lignées *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*, l'association des miARN matures aux protéines ALG-1 3a et GFP-ALG-2 3a est plus

faible, bien que la concentration des protéines Argonaute dans les différents extraits et l'efficacité des IP entre WT et 3a soient comparables (Fig.37a). En ce qui concerne miR-50, seul l'IP ALG-1 est montré, car aucune détection n'a été possible pour l'IP GFP-ALG-2 ou GFP-ALG-2 3a et ce, même après une longue exposition.

**Figure 39 : Expérience de détection de miARN associé aux protéines ALG-1 et GFP-ALG-2 en condition d'ARNi.** 90% des IP ont été utilisés. Une extraction d'ARN suivi d'une précipitation a été faite sur la moitié du surnageant restant (2 mg d'extraits de protéines totales).

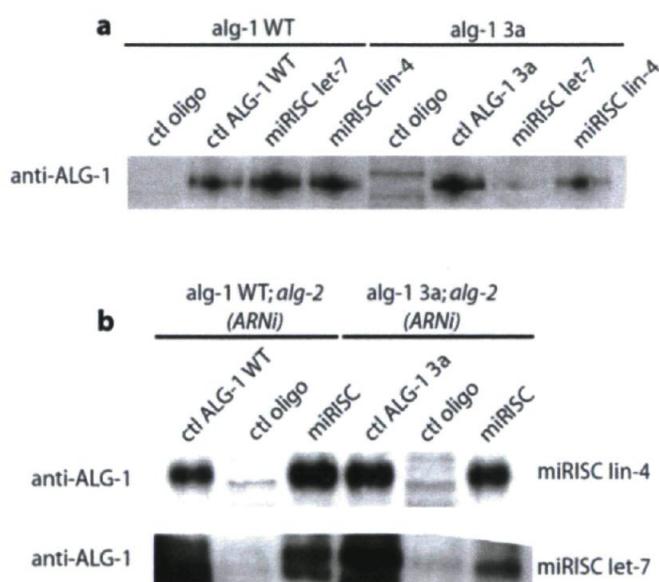


#### 2.5.4. Conséquence de la perte d'activité endonucléasique sur la formation de complexe miRISC détecté par 2'-O-méthyl

Pour confronter les résultats obtenus par IP, nous avons réalisé des expériences de 2'-O-méthyl. Cette méthode consiste à purifier directement les protéines qui lient les miARN en passant des extraits de protéines sur des colonnes sur lesquelles sont fixés des oligonucléotides, complémentaires à la séquence d'intérêt, et méthylés en position 2' du ribose, ce qui leur confère une résistance à la dégradation par les nucléases (Jannot et al., 2011b). Les colonnes sont couplées à la Streptavidine et les oligonucléotides 2'-O-méthylés sont biotinylés. Nous avons utilisé pour cette expérience deux types d'oligonucléotides 2'-O-méthylés, l'un complémentaire à lin-4 et l'autre à let-7. La première observation est le fait qu'aucun complexe n'a pu être isolé avec GFP-ALG-2 (données non-montrées). À la fois pour lin-4 et pour let-7, et dans différentes conditions, avec ou sans ARNi, GFP-ALG-2 ne semble pas être en mesure d'être isolé par cette méthode. La deuxième observation est le fait que l'effet le plus fort sur la formation du complexe miRISC semble être fait sur le simple mutant alg-1 3a (Fig.40a). On voit ici que le niveau de détection de la protéine

ALG-1 est plus fort dans la souche *alg-1* WT que dans la souche *alg-1 3a* et la baisse est plus forte pour le miARN *let-7* que pour le miARN *lin-4*. Cette expérience nous démontre donc que la formation du complexe miRISC est très affectée dans cette souche et que l'activité d'ALG-1 est nécessaire pour la formation du miRISC. Pour *alg-1 3a*; *alg-2(ARNi)*, il ne semble pas y avoir de différence significative et on observe une différence très légère, à la fois pour *let-7* et *lin-4* (Fig.40b). Comparativement aux données issues d'IP où l'on observe une baisse des miARN associées à ALG-1 3a et GFP-ALG-2 3a dans les doubles mutants ARNi, la détection quasi similaire observée pour l'expérience de 2'-O-méthyl indique que les complexes purifiés à l'aide d'anticorps ou les complexes purifiés par l'intermédiaire d'oligonucléotides 2'-O-méthylés sont distincts.

**Figure 40 : Expérience de détection de complexe miRISC ALG-1/*let-7* et ALG-1/*lin-4*.** Analyse de type western sur ALG-1 sur des extraits de protéines totales. Quatre mg de protéines ont été utilisés. a) vers récoltés au stade adulte. b) vers sur *alg-2* ARNi après 4 jours de traitement. ctl oligo, 2'-O-méthyl non spécifique; ctl protéine, 100 µg d'extraits de protéines utilisés; miRISC *let-7*, miRISC *lin-4* et miRISC, protéines liées spécifiquement par les oligonucléotides complémentaires à *let-7* et *lin-4*.



## 2.6. Effet de la perte d'activité endonucléasique de ALG-1 et ALG-2 sur la stabilité d'ARNm régulée par des miARN

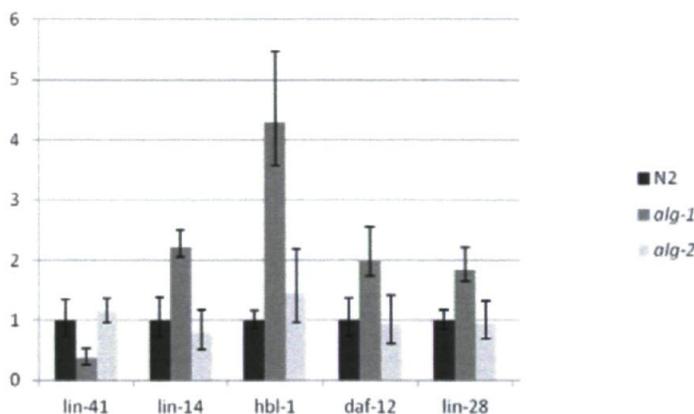
Pour étudier l'effet des mutations de protéines Argonaute sur certains ARNm cibles de miARN, nous avons porté notre attention sur plusieurs gènes dont la régulation par des

miARN était confirmée. Ainsi, nous avons contrôlé les niveaux de *lin-41*, *lin-14*, *hbl-1*, *daf-12* et *lin-28* par PCR en temps réel (RT-PCR) quantitative. L'ARNm du gène hétérochronique *lin-41* est négativement régulé par *let-7* durant les stades L4 à adultes (Reinhart et al., 2000; Vella et al., 2004). Celui de *hbl-1* est également réprimé par *let-7* (Abbott et al., 2005; Grosshans et al., 2005; Lin et al., 2003). La traduction de *lin-14* est réprimée par le miARN *lin-4* (Lee et al., 1993; Olsen and Ambros, 1999). La traduction de *lin-28* est régulée par *lin-4* et *let-7* (Reinhart et al., 2000; Seggerson et al., 2002) et *daf-12* code pour une protéine qui fait partie d'une famille de protéines réceptrices d'hormones stéroïdiennes et sa traduction est régulé par le miARN *let-7* (Hammell et al., 2009a; Reinhart et al., 2000).

### **2.6.1. Stabilité des ARNm régulé par des miARN dans les mutants nuls *alg-1* et *alg-2***

Nous avons vu précédemment les principaux phénotypes chez les mutants nuls *alg-1* et *alg-2*. La souche *alg-1* affiche le phénotype le plus sévère alors que *alg-2* ne semblent pas présenter de phénotype particulier. Au niveau moléculaire, les niveaux des miARN *let-7* et *lin-4* sont considérablement affectés chez le mutant *alg-1*, alors que chez *alg-2*, ils ne semblent pas être affectés. Les résultats de RT-PCR montrent une augmentation des cibles des miARN *let-7* et *lin-4* dans le mutant *alg-1* comparativement à la souche sauvage (Fig.41). Le niveau des ARNm de *lin-14*, *daf-12*, *hbl-1* et *lin-28* est plus abondant chez le mutant *alg-1* que chez la souche de référence N2. Seul l'ARNm de *lin-41* baisse de façon significative chez le mutant *alg-1*. Au vu de ses résultats, on peut en conclure que les défauts phénotypiques observés dans le mutant *alg-1* et la baisse de la formation de miARN mature vont de pairs avec un problème dans la stabilité d'ARNm ciblés par *let-7* et *lin-4*. En revanche, pour le mutant *alg-2*, aucun changement significatif n'est observable.

**Figure 41: RT-PCR sur les souches N2, *alg-1* et *alg-2*, au stade adulte.** Les résultats compilent trois cultures indépendantes et les RT-PCR ont été réalisées en duplicata. Les ARNm de *lin-41*, *lin-14*, *hbl-1*, *daf-12* et *lin-28* ont été testés. L'erreur barre représente l'intervalle de confiance à 95%.



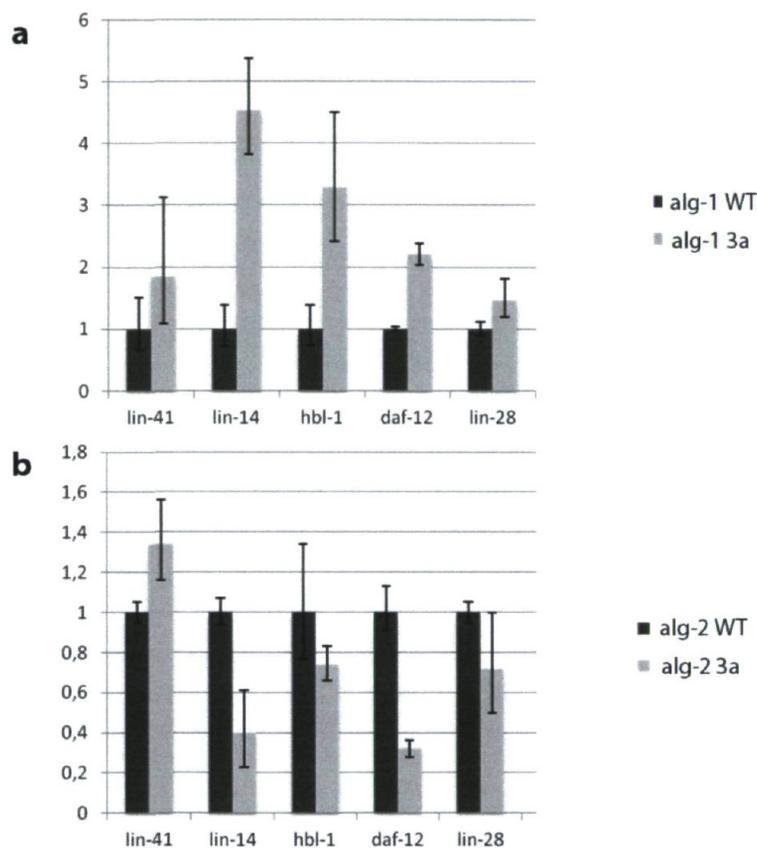
### 2.6.2. Stabilité des ARNm régulé par des miARN dans les lignées transgéniques

Pour rationaliser l'utilisation des transgènes, nous avons pris, pour *alg-1* et *alg-2*, le transgène WT comme référence afin de déterminer précisément l'effet de la perte d'activité des protéines Argonaute sur la stabilité des différents ARNm ciblés par *let-7* et *lin-4*. Il faut rappeler que ces résultats doivent être analysés dans leur contexte : 1- pour *let-7* et *lin-4*, nous avons des niveaux de miARN proches de celles du sauvage pour nos lignées transgéniques, excepté pour *let-7* pour lequel les lignées *alg-1* WT et *alg-1* 3a montraient une surproduction de miARN mature (Fig.33); 2- les expériences de 2'-O-méthyl ont démontré que la souche *alg-1* 3a était déficiente dans la formation de complexe miRISC *let-7* et *lin-4* (Fig.40a).

Les résultats de RT-PCR affichent deux profils très différents pour ces lignées (Fig.42). Chez la souche *alg-1* 3a, seuls les ARNm *lin-41* et *lin-28* varient très peu, mais *lin-14*, *hbl-1* et *daf-12* subissent une augmentation de leurs niveaux relativement à la souche transgénique WT (Fig.42a). Il y donc une nette tendance à l'augmentation des niveaux d'ARNm. Pour ce qui est de la souche *alg-2* 3a, nous voyons une nette différence (Fig.42.b). Alors que certaines cibles comme *lin-41*, *hbl-1* et *lin-28* sont très proches de la lignée *alg-2* WT, d'autres cibles comme *lin-14* et *daf-12* baissent de façon significative.

Les analyses par RT-PCR sur les lignées *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)* donnent des résultats très aberrants avec notamment une variation importante du gène contrôle de la tubuline (données non-montrées). Pour cette raison, l'interprétation des résultats est impossible et nous avons abandonné cette analyse.

**Figure 42 : RT-PCR sur les lignées transgéniques au stade adulte.** a) Analyse RT-PCR sur les lignées *alg-1* WT et *alg-1 3a*. b) Analyse RT-PCR sur les lignées *alg-2* WT et *alg-2 3a*. Les résultats compilent trois cultures indépendantes et les RT-PCR ont été réalisées en duplicata. Les ARNm de *lin-41*, *lin-14*, *hbl-1*, *daf-12* et *lin-28* ont été testés. L'erreur barre représente l'intervalle de confiance à 95%.



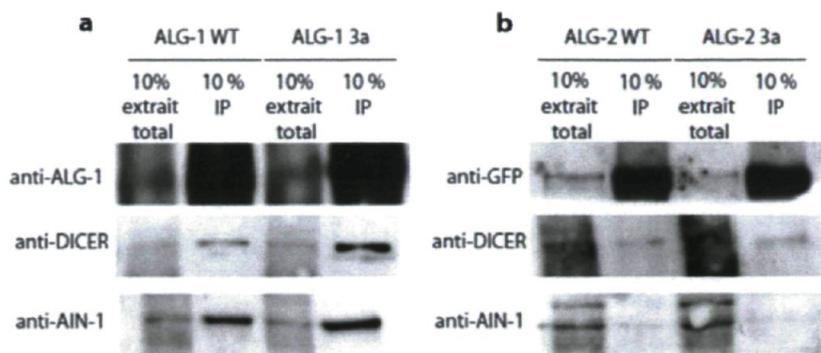
## 2.7. Exploration des différents complexes formés par les protéines Argonaute chez *alg-1 3a*; *alg-2(ARNi)* et *alg-2 3a*; *alg-1(ARNi)*

Plusieurs données du laboratoire et dans la littérature indiquent que ALG-1 et ALG-2 font partie d'un même complexe. Ainsi, par une étude protéomique, il a été identifié que, par l'intermédiaire de DCR-1, plusieurs protéines dont ALG-1, ALG-2 et AIN-1 font partie

d'un même complexe (Duchaine et al., 2006). Dans le laboratoire, la présence de ALG-1 et GFP-ALG-2 dans un même complexe immunoprécipité a également été reportée (G. Jannot). Cette interaction ne semble toutefois pas être directe puisque par expériences de double hybride, l'interaction entre ALG-1 et ALG-2 n'a pas été observée (Nellie Giguère). L'ensemble des résultats que j'ai obtenus jusqu'ici a permis de tirer des conclusions quant aux rôles de l'activité endonucléasique des protéines Argonaute aussi bien dans la maturation, dans la formation de complexes miRISC et dans la stabilité d'ARNm ciblés par des miARN, et nous en discuterons ultérieurement. Cependant, certaines données suggèrent que les protéines ALG-1 et ALG-2 ont un rôle distinct. Ainsi, l'absence de détection, par 2-O-méthyl, de la protéine GFP-ALG-2 associée aux miARN *let-7* et *lin-4*, bien que ces mêmes miARN soient détectables par IP de GFP-ALG-2, reflètent la présence de plusieurs types de complexes. De plus, la faible restitution de fonction de *alg-2* WT; *alg-1(ARNi)* peut être due à certaines fonctions spécifiques à *alg-1*. Nous avons donc effectué plusieurs IP afin de vérifier si deux des composants majeurs des complexes associés aux protéines Argonaute à savoir DCR-1 et AIN-1 respectivement, font partie des protéines associées à ALG-1 et/ou GFP-ALG-2. De plus, cette analyse nous permettra de déterminer si la mutation 3a influence la liaison avec ces deux protéines.

Nous avons donc immunoprécipité les protéines Argonaute chez *alg-1* WT; *alg-2(ARNi)*, *alg-1* 3a; *alg-2(ARNi)*, *alg-2* WT; *alg-1(ARNi)* et *alg-2* 3a; *alg-1(ARNi)*. Comme le montre la Figure 43, nous sommes capables de détecter une interaction de la protéine DCR-1 à la fois pour ALG-1 et ALG-2. Que ce soit pour les protéines Argonaute WT ou mutantes 3a, il semble que la mutation n'affecte pas l'interaction avec DCR-1. En revanche, concernant AIN-1, l'interaction apparaît exclusive à la protéine ALG-1. Comme pour DCR-1, la mutation n'affecte aucunement l'interaction. De façon surprenante, GFP-ALG-2 ne semble pas interagir avec l'homologue de GW182, ou de façon mineure, puisque l'on détecte une bande résiduelle. Cependant, il est possible que l'étiquette GFP affecte cette liaison. Ces IP supportent que les protéines ALG-1 et ALG-2 agissent à différents niveaux dans la voie des miARN.

**Figure 43: Protéines associées à ALG-1 et GFP-ALG-2 en condition d'ARNi.** a) Analyse de type Western Blot d'IP anti-ALG-1 contre les protéines ALG-1, DCR-1 et AIN-1. b) Analyse de type Western Blot d'IP anti-GFP contre les protéines GFP-ALG-2, DCR-1 et AIN-1. 10% de l'extrait total de protéines utilisés (4 mg) a été utilisé comme contrôle de la présence des protéines et 10% de l'IP a été utilisé pour vérifier la présence de protéines associées.



# Chapitre 5 : Discussion et perspectives

## 1. Détermination de l'activité endonucléasique des protéines Argonaute de *C. elegans*

### 1.1. Activité endonucléasique des protéines Argonaute recombinantes rRDE-1, rSAGO-1 et rPRG-1

#### 1.1.1. Activité de rRDE-1 et rSAGO-1

Certaines études génétiques ou biochimiques ont étayés le rôle des protéines RDE-1 et SAGO-1 dans la voie de l'exo-ARNi. *In vivo*, des études démontrent que l'activité de RDE-1 est nécessaire pour le passage du duplex au simple brin de siARN (Steiner et al., 2009). SAGO-1, membre représentatif des Argonaute secondaires et ne possédant pas de motif DDH, est impliquée dans la prise en charge des siARN secondaires (Yigit et al., 2006). D'autres données indiquent que CSR-1, une autre protéine Argonaute spécifique au nématode, est responsable de l'activité catalytique médiée par les siARN secondaires (Aoki et al., 2007). Mes résultats indiquent que rRDE-1 peut cliver des molécules d'ARN et que rSAGO-1 n'a pas d'activité catalytique. rRDE-1 est en mesure de cliver un ARN cible en présence d'un ARNsb P ou 3-P et clive des substrats d'ARN duplex parfaitement apparié.

De ces éléments, on peut en conclure que RDE-1 est la protéine Argonaute nécessaire pour cliver le duplex de siARN, après le clivage de l'ARN déclencheur par DCR-1 puisque les données *in vivo* et *in vitro* concordent. Néanmoins, il n'y a aucune certitude quant au fait que l'ARNm est clivé ou non par RDE-1. Pour ce faire, il serait important de démontrer, par des expériences de type Northern, si l'ARNm ciblé est coupé par RDE-1. Durant mon doctorat, j'ai créé les lignées transgéniques *rde-1* WT et *rde-1* 3a. Par l'intermédiaire d'un ARNm reporteur comme la GFP, nous serions en mesure de définir, en induisant une réponse ARNi contre la GFP, s'il y a une coupure endonucléolytique de l'ARNm de la GFP dépendante de RDE-1. Concernant la prise en charge des siARN secondaire tri-P, mes résultats montrent que, *in vitro*, rRDE-1 peut cliver un ARN cible par l'intermédiaire d'un siARN tri-P. Ces résultats apportent un nouvel élément dans certaines hypothèses posées

dans la littérature, notamment celle indiquant que CSR-1 peut cliver les ARNm par les siARN secondaires tri-P. Enfin, concernant SAGO-1, et le rôle des autres protéines Argonaute secondaires ne conservant pas d'activité endonucléasique, il semble clair que l'absence du motif DDH conduit à une absence d'activité. En complexe avec les siARN secondaires, ils pourraient agir comme inhibiteur transcriptionnel pour maintenir la réponse ARNi.

### **1.1.2. Activité de rPRG-1**

Concernant rPRG-1, cette protéine a été initialement identifiée comme interagissant avec de courts ARN de 21 nt portant un U en 5' et s'avérant être les piARN spécifiques à *C. elegans* (Batista et al., 2008; Wang and Reinke, 2008). Contrairement au mécanisme chez la drosophile et chez l'humain, il ne semble pas y avoir de production par un processus de ping-pong chez *C. elegans*. En effet, il n'a pas été démontré de séquence de courts ARN sens et antisens à un transcrit ciblé comme pour la drosophile. Mes travaux sur la protéine recombinante ont permis de démontrer l'activité catalytique de rPRG-1 mais aussi que l'utilisation d'ARN guide pour les expériences de clivage requiert un U en 5' terminal. Le fait que les piARN soient exclusivement uridylés en 5' peut expliquer pourquoi le clivage d'une cible n'est possible qu'en utilisant ce type de molécule.

Ces travaux ont été utilisés par le groupe du Dr Miska, dans le cadre d'une collaboration, où on a tenté de déterminer l'implication de l'activité endonucléasique de PRG-1 dans la voie des piARN. Contrairement à la voie des miARN et de l'exo-ARNi, l'activité de PRG-1 ne semble pas intervenir dans la production des piARN ou dans la régulation de certaines cibles. De plus, cette étude révèle également que la régulation des ARN ciblés par les piARN s'effectue par une complémentarité imparfaite et la stabilité des cibles n'est pas influencée par la perte d'activité de PRG-1.

## **1.2. Activité endonucléasique des protéines Argonaute recombinantes rALG-1 et rALG-2**

Mes travaux ont démontré l'activité endonucléasique de deux protéines Argonaute du groupe «AGO-like», ALG-1 et ALG-2, dont le motif DDH est essentiel au maintien de

cette activité. Il a été démontré préalablement, chez d'autres organismes, que ce motif était nécessaire pour que les protéines Argonaute soient douées d'activités catalytiques, et ce, dans deux études de références (Liu et al., 2004; Song et al., 2004). Il était important, pour l'étude de la participation de l'activité de ALG-1 et ALG-2 à la voie des miARN, de démontrer l'activité de ALG-1 et ALG-2 *in vitro* pour plusieurs raisons. Les deux études de références, démontrant que le DDH est nécessaire à l'activité endonucléasique, sont basés sur la similarité observée entre le domaine PIWI et la famille des RNases H, et sur le fait que la mutation de ce motif abolit le clivage. Néanmoins, plusieurs points demeurent encore mystérieux et d'autres études indiquent que ce domaine n'est pas le seul déterminant conférant une activité catalytique. Ainsi, certains membres de la famille des Argonaute humaines, comme Ago3, conservent ces trois résidus, mais ne sont pas catalytiquement active. Cette absence d'activité a été déterminée à la fois par des expériences d'IP mais également par essai sur des protéines recombinantes (Liu et al., 2004; Wang et al., 2009a). Ces deux études obtiennent des résultats divergents concernant un autre membre, Ago1, qui possède une triade DDR. Alors que l'IP de Ago1 ne permet pas de visualiser une quelconque activité, l'essai sur des protéines recombinantes produites chez *E. coli* indique le contraire. On peut donc conclure que l'activité peut être certainement régulée par certaines modifications post-traductionnelles n'ayant pas lieu chez la bactérie. Il était donc important de démontrer l'effet de la perte du motif DDH de ALG-1 et ALG-2, *in vitro* et *in vivo*. De plus, il était également important de produire d'autres membres de la famille des protéines Argonaute comme rRDE-1, rPRG-1 et rSAGO-1. Premièrement, pour confirmer que le motif DDH est nécessaire pour l'activité endonucléasique, et rSAGO-1, n'en conservant pas, a servi d'un excellent contrôle. Et deuxièmement, par comparaison des différentes études faites pour RDE-1 et PRG-1, je voulais déterminer si certaines caractéristiques importantes comme le clivage de molécules duplex ou la préférence pour un 5'U étaient conservées chez les protéines recombinantes. Par ce biais, les caractéristiques issues des expériences pour rALG-1/2 peuvent ainsi avoir plus de poids quant à leur signification biologique.

### 1.2.1. Implication de différents domaines pour l'activité endonucléasique

Certains résultats, que j'ai partiellement décrits dans cette thèse, sont très intrigants et les conclusions vont quelque peu à l'encontre d'une étude faite chez la drosophile où notamment des portions de protéines Argonaute démontrent une activité notamment pour le domaine PIWI (Miyoshi et al., 2005). Le domaine PIWI avec une portion du domaine MID et la partie C-terminale conserve une activité. Lors de la tentative de production de différents mutants de délétion, résumés dans le Tableau 3, il s'est avéré que certaines protéines étaient très instables, sujettes à la dégradation et à la précipitation. La portion PIWI et la portion Mid-PIWI-Cter ont montré, dans les expériences de clivages, une capacité de clivage abolie pour le premier, et une instabilité complète de la protéine, pour le second. Au final, je suis arrivé à la conclusion que ces différents mutants ne permettaient, chez *C. elegans* et pour rALG-1 et rALG-2, de conclure à une quelconque activité de ces portions.

Il s'est avéré que certaines délétions/mutations abolissent ou réduisent considérablement l'activité, d'autant que la pureté de ces mutants et la stabilité étaient bonnes. Ainsi, les délétions du PAZ et du Mid ont une activité réduite, pour le premier, ou abolie, pour le second. Le rôle respectif de ces deux domaines explique aisément les résultats observés. Le domaine Mid est important pour lier spécifiquement le 5' terminale du court ARN guide (Frank et al., 2011; Frank et al., 2010) et cet ancrage apparaît comme essentiel à l'activité de clivage. Le domaine PAZ est connu pour lier l'extrémité 3'OH du court ARN guide (Lingel et al., 2004) et, d'après de récentes études structurales, est très mobile durant le processus de nucléation entre un ARN guide et un ARN cible (Wang et al., 2009b). La délétion de ce domaine affecte le clivage et ceci peut être explicable par un ancrage de l'extrémité 3'OH absent, et à une nucléation entre ARN cible-ARN guide moins efficace, probablement due à un problème conformationnel. De plus, mes résultats indiquent que ce domaine est possiblement très important dans la reconnaissance des duplexes de 21 nt.

Dans le mutant  $\Delta$ PAZ, l'activité de clivage et de relargage du duplex de 21 nt est abolie. L'explication la plus logique serait de penser que la liaison est affectée mais d'après des résultats de FBA (données non-montrées), la liaison aux différentes molécules d'ARN testées n'est pas affectée. Il est toutefois possible que d'autres domaines permettant cette liaison, comme le domaine Mid et PIWI, puissent masquer la baisse d'affinité de  $\Delta$ PAZ, le

test de FBA n'étant pas très sensible. Mais, si la liaison n'est pas affectée, étant donné la mobilité de ce domaine lors de la nucléation ARN guide-ARN cible, on peut logiquement penser que lors du relargage, le processus inverse, ce domaine est également mobile et participe au déroulement des deux brins. Cependant, concernant le clivage du duplex, l'implication de la mobilité du domaine PAZ est plus difficilement envisageable. On peut toutefois imaginer que l'absence du domaine PAZ ne permet pas le clivage par une absence d'ancrage d'une extrémité 3'OH du duplex. Cette absence d'ancrage empêcherait le clivage du duplex de manière plus forte que pour le clivage d'un ARN cible par l'intermédiaire d'un ARNsb.

Pour ce qui est du domaine C-terminal, le fait que l'activité soit affectée n'est appuyé par aucune donnée dans la littérature. Néanmoins, plusieurs structures font état d'une position centrale de ce domaine dans la structure tertiaire des Argonaute (Song et al., 2004; Wang et al., 2008b; Wang et al., 2009b). Récemment, un groupe, travaillant sur une portion de la protéine QDE-2 de *Neurospora Crassa*, a déterminé la présence d'un élément hydrophobique important en C-terminal ainsi que deux hélices présentes uniquement dans les protéines Argonaute eucaryotiques du groupe «AGO-like». La mutation de ces hélices a pour effet, *in vivo*, la perte de liaison aux miARN (Boland et al., 2011). Dans un contexte de protéine pleine longueur, ces deux hélices seraient en mesure de créer des interactions avec le domaine PAZ et le domaine Mid. Ces différents éléments pourraient expliquer pourquoi la perte du domaine C-ter mènerait à une réduction de l'activité.

### **1.2.2. Déterminants nécessaires au niveau de la molécule d'ARNsb guide permettant le clivage d'un ARN cible**

Les conditions de bases pour visualiser le clivage des protéines Argonaute sont de fournir une molécule d'ARNsb de 21 nt, suivi d'une période d'incubation qui va permettre la liaison de l'ARNsb à la protéine Argonaute. Ensuite, on ajoute au milieu réactionnel un ARN radiomarqué. Bien que n'ayant pas pu par FBA déterminer une différence de liaison entre les ARNsb utilisés (ayant en 5' un U, C, A ou G), certainement par un manque de sensibilité de la méthode, les expériences de clivage nous ont montré une préférence concernant l'utilisation de ces molécules pour les protéines rALG-1, rALG-2 et rPRG-1 notamment. Concernant rALG-1 et rALG-2, il semble que nous soyons en mesure de

détecter un produit spécifique de clivage pour toutes les molécules d'ARNsb utilisé. Néanmoins, le produit de clivage de taille spécifique est obtenu en quantité équivalente pour les ARNsb portant en 5' un U, C et A, alors qu'il est minoritaire pour ceux portant un G. Il est connu que chez d'autres espèces comme la drosophile, la majorité des miARN possède un U en 5' terminal (70%), ou encore un A (15-20%), et les miARN possédant un C ou G sont minoritaires (Ghildiyal et al., 2010). Chez la plante, il a été déterminé que la sélection par les différentes et nombreuses protéines Argonaute se faisait via le 5' terminal des différents court ARN (Mi et al., 2008). Enfin, plus spécifiquement chez *C. elegans*, le séquençage des miARN confirme la tendance d'un U en position 5'terminal (Warf et al., 2011; Zisoulis et al., 2010). Le résultat des expériences de clivage va donc dans la même direction que la littérature.

## **2. Signification in vivo de l'utilisation de différents substrats d'ARN in vitro par ALG-1 et ALG-2**

Les travaux présentés dans cette thèse ont démontré que les protéines Argonaute rALG-1 et rALG-2 sont capables de prendre en charge différents substrats d'ARN. Ainsi, les duplexes de 21 nt parfaitement appariés ou avec des mésappariements, tout comme les molécules sb et les structures précurseurs de miARN, sont des substrats pour rALG-1 et rALG-2. La confrontation entre nos données in vitro, in vivo et la littérature nous permet de construire plusieurs modèles quant à l'étape où ce clivage intervient.

### **2.1. Action au niveau du précurseur**

Dans la littérature, une étude sur la surexpression à la fois de Ago2 humaine et de pré-let-7, a démontré le clivage du pré-let-7 par Ago2, menant à la synthèse d'un nouvel intermédiaire, utilisable par Dicer (Diederichs and Haber, 2007). Cet intermédiaire est clivé au niveau du brin passager de let-7 directement dans le précurseur, face à la position 10-11 en partant du 5' du brin miARN. Ce clivage est aboli par la mutation du motif DDH. Détecté par surexpression de pré-let-7 et de Ago2, cette équipe a néanmoins été capable de cloner des précurseurs clivés endogènes. Mes travaux ont démontré que les protéines recombinantes rALG-1 et rALG-2, bien que capable de lier efficacement ce type de

molécule, étaient incapable de les cliver. Dans la littérature, aucune évidence n'indique que les protéines Argonaute recombinantes seraient capables d'effectuer un clivage du précurseur. De plus, je n'ai pas pu confirmer *in vivo* par des expériences d'IP que les protéines Argonaute lient ce type d'ARN. Plusieurs interprétations sont possibles : 1-le clivage du précurseur n'est pas possible avec les protéines Argonaute recombinantes; 2-le clivage n'est possible qu'avec l'aide de cofacteurs comme Dicer et, en l'occurrence pour les cellules humaines, TRBP (pas d'homologue chez *C. elegans*); 3-le clivage par les protéines Argonaute, dans les conditions endogènes, n'a pas lieu chez *C. elegans*, et ce phénomène n'est pas conservé. Cependant, l'occurrence de ce phénomène *in vivo* ne semble pas être essentielle, car l'article, cité précédemment, n'a pas démontré que Dicer utilise ce produit de façon exclusive ou préférentielle par rapport au pré-let-7.

## 2.2. Action au niveau du duplex de miARN

L'ensemble de mes résultats indique que l'activité de clivage des protéines ALG-1 et ALG-2 se situe au niveau du duplex de miARN. *In vitro*, nous avons déterminé que le clivage de duplexes plus ou moins apparié était possible. *In vivo*, pour différentes raisons possibles, le seul miARN dont le brin passager s'accumule est le miARN lin-4. Cette accumulation peut être due à une stabilité élevée de ce miARN passager comparée à tous les autres miARN testés. Il a été démontré dans la littérature que la dégradation du brin passager pouvait être freinée par la présence d'ARN complémentaire, rendant ainsi des passagers très peu abondants, détectables par analyse de type Northern (Chatterjee et al., 2011). Il en est de même pour des miARN mature peu abondants. On pourrait donc penser que lin-4 passager pourrait exercer une fonction de régulation inconnue expliquant ainsi sa détection. Pour ce miARN, nous sommes capables de voir une différence très claire entre les souches alg-1 WT et alg-1 3a. Dans un extrait total d'ARN, l'accumulation du brin passager dans le mutant 3a indique clairement que la stabilité et la quantité de ce miARN se trouve positivement affecté par la mutation. Dans les IP, alors que l'on voit une baisse significative de lin-4 mature dans ALG-1 3a, on observe, en même temps et comparativement au WT, une augmentation de lin-4 passager associé au mutant. Ceci est une preuve directe qu'il y a un défaut de clivage ou de dégradation dans le mutant catalytique.

Pour les autres miARN comme let-7, miR-90 et miR-50, la détection du brin passager n'est visible que par IP. La tendance au niveau du brin mature est la même que pour lin-4, à savoir une baisse significative de miARN mature entre le WT et le mutant, et ce pour ALG-1 et rALG-2. Si l'on compare la baisse d'association du brin mature dans le mutant ALG-1 3a et la détection similaire du brin passager, on peut considérer qu'il s'agit d'une stabilité accrue de passager. Cela irait dans le même sens que pour lin-4 où une perte de l'activité endonucléasique serait visible par une augmentation de l'association du passager relativement au miARN mature. Cependant, l'absence de détection de ces différents brins passagers, dans les extraits d'ARN total, peut s'expliquer par une stabilité ou une quantité de passagers ne permettant pas de les visualiser, même dans les conditions où le clivage *in vivo* est déficient.

Enfin, il n'est pas non plus exclu que certains duplexes de miARN ne soient pas sous l'influence du clivage des protéines Argonaute. De nombreux miARN possèdent de multiples mésappariements et impliquent que les protéines Argonaute interviennent dans leur maturation par une fonction non catalytique au niveau des duplexes. Mes résultats démontrent que le relarguage que j'observe est indépendant du clivage, puisque les protéines Argonaute ne clivent pas ou clivent faiblement les duplexes à haut degré de mésappariement. Dans la littérature, il a été démontré que les protéines Argonaute étaient capables de séparer les deux brins d'un duplex de façon indépendante du clivage du passager (Wang et al., 2009a). Ainsi, on pourrait penser que les protéines Argonaute mutées, dans certaines conditions, soient capables de relarguer le brin passager, qui sera entraîné vers la dégradation.

### **2.3. Origine du précurseur tronqué**

*In vivo*, la perte d'activité des protéines ALG-1 ou ALG-2 conduit à la production de précurseur tronqué pour tous les miARN testés, comme lin-4, let-7 mais aussi le mitron miR-62, qui est indépendant de Drosha. L'origine de cette maturation aberrante peut être inhérente aux protéines Argonaute mutées ou à DCR-1. L'activité aberrante de la protéine Drosha est exclue car miR-62 affiche également une bande de taille correspondante à un précurseur tronqué. Deux hypothèses sont donc possible concernant la présence du

précurseur tronqué: 1-le précurseur tronqué provient d'une activité aberrante des protéines Argonaute mutées; 2-ce précurseur provient d'une activité aberrante de DCR-1. La première hypothèse est difficilement soutenable étant donné que, *in vitro*, les protéines Argonaute n'ont pas d'activité sur le précurseur, que ce soit pour les protéines WT ou mutantes. La deuxième hypothèse, une activité aberrante de DCR-1, semble la plus probable. Premièrement, les précurseurs tronqués sont observables à la fois avec des sondes dirigées contre le miARN mature et contre le passager, ce qui indique que ces séquences contiennent soit le brin mature, soit le brin passager. Ensuite, la taille de ces précurseurs est d'environ 50 nts, soit une vingtaine de nt de moins que la majorité des précurseurs (70 nts), taille qui correspond aux fragments générés par DCR-1. Enfin, le clonage et le séquençage des précurseurs tronqués indiquent qu'ils correspondent au précurseur amputé de leur miARN ou de leur brin passager. Or, la seule protéine capable d'une telle activité est DCR-1. Ces données nous conduisent à penser que l'activité catalytique abolie des Argonaute peut mener à un dysfonctionnement de DCR-1. En conséquence, la production de ce précurseur aberrant ne va pas permettre une fine régulation de la synthèse de miARN, essentielle au développement, ce qui expliquerait le phénotype observé lors de la perte d'activité des protéines Argonaute dans le double mutant. La première évidence est le fait que la protéine DCR-1 agit de façon séquentielle pour cliver le précurseur de miARN. Une étude biochimique chez la drosophile a établi que chaque domaine RNases III clive chaque brin du précurseur de façon indépendant (Ye et al., 2007). La mutation du motif catalytique d'un domaine RNase III a pour conséquence la production d'un précurseur tronqué du miARN ou du passager, dépendamment que la mutation soit faite sur l'un ou l'autre des domaines. Une autre étude souligne que DCR-1 agit de façon séquentielle, clivant d'abord le précurseur sur le bras 5' ou 3', rendant ainsi possible la détection d'un intermédiaire tronqué, puis clive le second bras, afin de produire un miARN mature (Flores-Jasso et al., 2009). Enfin, il a été déterminé que DCR-1 et les protéines Argonaute interagissent ensemble via les domaines RNases III et PIWI (Tahbaz et al., 2004). Dans la même étude, les auteurs mentionnent que l'ajout de domaine PIWI peut inhiber la maturation de précurseur de miARN par DCR-1. Dès lors, on peut imaginer que l'interaction de domaines PIWI d'une protéine Argonaute non fonctionnelle avec les domaines RNases III bloque la réaction de maturation.

## 2.4. Modèle de l'intervention catalytique de ALG-1/2

Les IP de ALG-1/2 montrent que les protéines Argonaute font partie d'un complexe avec DCR-1 et ce, avec les protéines WT ou mutées. Chez *C. elegans*, des expériences de protéomiques ont préalablement démontré que ALG-1 et ALG-2 font partie d'un complexe immunoprécipité de DCR-1 (Duchaine et al., 2006). Chez d'autres organismes, comme la drosophile ou les mammifères, la présence de protéines Argonaute dans un complexe avec la protéine Dicer est également clairement établie (Kim et al., 2009). Un autre cofacteur important, nécessaire pour la maturation du précurseur, est également toujours associé à ce complexe et a une fonction de liaison à l'ARN. Il s'agit de TRBP chez les mammifères (Chendrimada et al., 2005; Haase et al., 2005) et de LOQS chez la drosophile (Forstemann et al., 2005; Saito et al., 2005). Contrairement à ces organismes, *C. elegans* n'a pas de protéine similaire identifiée qui participe à la voie des miARN, laissant penser que cette fonction pourrait être assurée par ALG-1 et ALG-2, ou qu'elle n'est pas nécessaire. Cependant, l'absence de ALG-1 et ALG-2, chez le nématode, conduit à l'abolition de la maturation des miARN par Dicer et à une accumulation de précurseur (Grishok et al., 2001) et tous ces éléments nous permettent d'envisager que les protéines Argonaute soient nécessaires pour la maturation des précurseurs par Dicer.

Mes travaux de thèse vont plus loin quant au mécanisme par lequel les protéines Argonaute sont impliquées dans la maturation des précurseurs de miARN et permettent de bâtir un modèle du mécanisme de maturation des miARN. L'ensemble de mes résultats suggèrent que l'activité catalytique de DCR-1 et celles des protéines Argonaute est couplée. Dans la condition WT, DCR-1, interagissant avec ALG-1 et ALG-2, clive le précurseur de façon séquentielle, sur le brin 5' puis sur le brin 3'. Le fait que mes expériences d'IP n'ont pas permis d'identifier de précurseur tronqué associé aux protéines Argonaute indique que le chargement sur les protéines Argonaute se fait au niveau de l'intermédiaire duplex de miARN. Le rôle des protéines Argonaute sera de cliver ou relarguer le brin passager, selon le degré de mésappariement des deux brins du duplex. Le complexe miRISC, ainsi formé, peut cibler un ARNm et n'est alors plus lié à DCR-1. Dans la condition mutante, le duplex est également chargé sur les protéines Argonaute, mais les protéines sont déficientes dans la dégradation du brin passager. L'équilibre entre le clivage du précurseur et le

clivage/relargage du duplex est rompu et la maturation est alors limitée dans l'étape de formation du miRISC (protéines Argonaute chargés avec un miARN). Ce frein dans le processus de maturation a pour conséquence un ralentissement de l'activité de DCR-1, rendant possible la détection de la forme tronquée, l'accumulation de brin passager associé aux protéines ALG-1 et ALG-2 et une réduction de l'association entre les protéines Argonaute et le brin de miARN mature.

### 3. Différences fonctionnelles entre ALG-1/2

Mes travaux de thèse ont également pointé quelques différences fonctionnelles entre ALG-1 et ALG-2. Bien qu'*in vitro* nous n'ayons pu déterminer de différences entre les protéines recombinantes rALG-1 et rALG-2, l'étude *in vivo* nous apporte quelques éléments de réponses. Le fait que les simples mutants *alg-1* et *alg-2* soient viables et que le double mutant *alg-1; alg-2* soit létal indique bien évidemment une redondance de fonction. Mais la pénétrance du phénotype du mutant *alg-1* comparé à *alg-2* indique que cette redondance est partielle, avec un rôle prépondérant d'ALG-1 dans la voie des miARN et pour comprendre ce rôle, plusieurs hypothèses seront étayées.

#### 3.1. miARN spécifiques à ALG-1 ou ALG-2

L'une des explications possibles est de penser que ALG-1 lie ou est impliquée dans la synthèse de miARN ayant un rôle majeur dans le développement. Les deux miARN ayant le rôle le plus important sont *lin-4* et *let-7*. Ainsi, les analyses de type Northern nous indiquent une différence notable quant à l'accumulation de différents précurseurs et le niveau de miARN mature. Alors que tous les miARN testés semblent affectés dans le mutant *alg-1*, quelques-uns le sont sévèrement dans le mutant *alg-2* (miR-50 et miR-40). Les miARN *lin-4* et *let-7*, de première importance, semblent très dépendants de ALG-1, que ce soit au niveau du miARN mature, que dans la maturation par DCR-1 (augmentation du niveau de précurseur). Les expériences d'ARNi contre *alg-1* ou *alg-2* sur les mutants *lin-4* et *let-7 ts* vont dans le sens d'une participation majeure de ALG-1. L'absence de ALG-2 ne semble pas augmenter le phénotype de ces deux mutants alors que l'absence de ALG-1 a un effet drastique, à savoir une létalité de 100%. Cette spécificité particulière d'ALG-1

pour ces deux miARN pourrait donc à elle seule expliquer la sévérité du phénotype du mutant *alg-1*, comparativement à *alg-2*.

Ce qui rend plus compliqué ce modèle est le fait de ne pas trouver de miARN affectés spécifiquement par l'absence de ALG-2. Ainsi, même pour les miARN sévèrement touchés dans le mutant *alg-2* (miR-50 et miR-40), ces mêmes miARN sont atteints dans le mutant *alg-1*. Cela semble indiquer que ces miARN sont partagés par les deux protéines Argonaute et que la spécificité n'est que partielle. De plus, les expériences d'IP, dans des conditions d'ARNi, n'ont jamais démontré que certains miARN n'étaient immunoprécipités que par une seule protéine Argonaute, excepté miR-50. Pour ce miARN, il est également difficile de comprendre pourquoi, du moins au stade adulte, il n'est pas immunoprécipitable par GFP-ALG-2 alors qu'il semble clairement affecté par l'absence de ALG-2. Cela indique que l'absence de ALG-2 a un effet important sur la stabilité de miR-50 alors que la protéine ne semble pas pouvoir lier ce miARN. Ainsi, il semble évident que ALG-2 participe à la production de ce miARN sans pour autant s'y lier, ce qui indiquerait que la production de miARN par le biais du couple DCR-1-ALG-1-ALG-2 et l'association entre ALG-1, ALG-2 et le miARN est distincte.

### 3.2. Rôle différentiel de l'activité de ALG-1 et ALG-2

Bien que nous ayons démontré que l'activité catalytique d'au moins une protéine Argonaute est nécessaire, quelques données intrigantes peuvent nous faire penser à une autre hypothèse, à savoir une action différentielle de l'activité des protéines Argonaute. Ainsi, la principale donnée à prendre en compte pour cette hypothèse est l'accumulation de précurseur. Alors que l'accumulation de précurseur est restaurée pour les miARN affectés dans le mutant *alg-1* (ex *lin-4*, *let-7*, miR-90 et miR-50) par les transgènes *alg-1* WT et 3a, ceux dont l'accumulation est visible dans le mutant *alg-2*, ne le sont pas pour le transgène *alg-2* 3a (miR-90, miR-40 et miR-50). Autrement dit, cela correspond à dire que l'activité catalytique de ALG-2 est requise pour rétablir le niveau de précurseur des miARN affectés, mais que la présence de la protéine ALG-1, mutée ou active, est suffisante pour rétablir le niveau de précurseur.

### 3.3. Rôle de ALG-1 à différents niveaux

Certaines données peuvent indiquer un rôle de ALG-1 à plusieurs niveaux. Il a été démontré que ALG-1 et AIN-1 agissent de concert pour conduire certains ARNm ciblés par les miARN vers les «P-bodies» (Ding et al., 2005). Dans une étude plus récente, le même groupe a démontré que deux protéines homologues AIN-1 et AIN-2, participent à la régulation de l'expression des cibles de miARN (Zhang et al., 2007). Le double mutant affiche des défauts majeurs dans la régulation de gènes ciblés par les miARN, notamment lors du développement de la vulve et des cellules HSC et au niveau des simples mutants, le phénotype est plus sévère pour *ain-1*. Cette fonction est donc importante pour la régulation par les miARN. L'interaction physique de AIN-1 avec les protéines étiquetées GFP-ALG-1 et GFP-ALG-2 a été déterminée par IP. Mes résultats ont démontré que dans des conditions d'ARNi, GFP-ALG-2 ne semble pas être en mesure de former un complexe avec AIN-1 alors que le complexe avec DCR-1 est détectable. Le fait que ALG-1 et ALG-2 co-immunoprécipite nous porte à croire que la détection observée dans cette étude est indirecte et il se peut donc que ALG-2 interagissent à AIN-1 via ALG-1. Le fait que l'interaction ALG-2-AIN-1 ne soit pas détectable peut s'expliquer par le fait que ALG-1 est la protéine Argonaute principale vectrice du miRISC. Cependant, on ne peut pas exclure que ALG-2 lie la protéine AIN-2, dont la fonction précise reste à étudier. L'hypothèse du miRISC-ALG-1 pourrait également préciser pourquoi la restitution de fonction de *alg-2* WT transgénique est si faible par rapport à *alg-1* WT. De plus, au niveau de la régulation de cibles connus, le mutant *alg-1* montrent une augmentation de l'expression des ARNm alors que le mutant *alg-2* est proche du niveau du WT. L'incapacité de détecter, par 2'-O-méthyl, un complexe avec GFP-ALG-2, dans des conditions de cultures normales ou d'ARNi, vont dans ce sens. Cependant, étant donné que les miARN régulant ces ARNm sont *let-7* et *lin-4*, et que les expériences de 2'-O-méthyl sont également dirigés contre le miRISC *let-7* et *lin-4*, cette conclusion ne peut être générale et peut nous mener à penser que le rôle de ALG-1 est prépondérant, de par son importance pour les miARN *let-7* et *lin-4*.

## 4. Conclusions et perspectives

Mes travaux de doctorat ont démontré que les protéines ALG-1 et ALG-2 sont douées d'activité endonucléasique et que cette activité était essentielle *in vivo* pour la maturation du duplex de miARN en miARN simple brin. Nous avons également identifié une coopération *in vivo* entre les activités de DCR-1 et les protéines Argonaute, et que ces dernières sont un cofacteur essentiel pour la protéine DCR-1. Bien que partielles, nous avons pu identifier quelques différences fonctionnelles entre ces deux protéines, mais aucune de façon biochimique. Ainsi, la protéine ALG-1 est très importante dans le développement du nématode puisqu'elle pourrait agir en amont de la voie, avec DCR-1, et en aval avec AIN-1 alors qu'il semble que la protéine ALG-2 intervient en amont. L'importance de ALG-1 vient également du fait qu'elle est la protéine vectrice principale de *let-7* et *lin-4*, les deux miARN les plus importants pour le contrôle du développement du nématode.

Pour approfondir cette étude, il serait néanmoins très intéressant d'effectuer un séquençage à haut débit sur un extrait d'ARN total des différentes souches transgéniques créées. Cela nous permettrait de déterminer avec précision quelles miARN nécessitent l'activité de ALG-1 ou de ALG-2 et si une loi d'ordre moléculaire déterminerait le chargement préférentiel vers l'une ou l'autre des protéines. De plus, cela pourrait nous permettre de déterminer s'il y a des miARN dont la synthèse est entièrement dépendante de l'activité des protéines Argonaute, mais indépendante de DCR-1, comme récemment découvert chez les mammifères (Cheloufi et al., 2010; Cifuentes et al., 2010b).

## Conclusion

Mes travaux ont permis de déterminer *in vitro* l'activité endonucléasique de protéines représentatives de trois groupes phylogénétiques de protéine Argonaute. *In vivo*, nous avons pu mettre en évidence l'étape dans laquelle cette activité intervient dans la voie des miARN. Cette activité est essentielle et des défauts majeurs sont observés et l'activité de la protéine DCR-1 est perturbée. La dérégulation des miARN dans certains cancers et maladies, de même que l'implication des protéines PIWI dans la fertilité et la possibilité d'utiliser les mécanismes d'ARNi dans un but thérapeutique nécessite une connaissance profonde de ces voies de régulation. Ainsi, je pense que mes travaux ont contribué, modestement, à faire avancer les connaissances fondamentales quant à la participation de l'activité endonucléasique des protéines Argonaute chez l'animal.

## Bibliographie

- Abbott, A.L., Alvarez-Saavedra, E., Miska, E.A., Lau, N.C., Bartel, D.P., Horvitz, H.R., and Ambros, V. (2005). The let-7 MicroRNA family members mir-48, mir-84, and mir-241 function together to regulate developmental timing in *Caenorhabditis elegans*. *Dev Cell* *9*, 403-414.
- Agrawal, R., Tran, U., and Wessely, O. (2009). The miR-30 miRNA family regulates *Xenopus* pronephros development and targets the transcription factor Xlim1/Lhx1. *Development* *136*, 3927-3936.
- Alder, M.N., Dames, S., Gaudet, J., and Mango, S.E. (2003). Gene silencing in *Caenorhabditis elegans* by transitive RNA interference. *RNA* *9*, 25-32.
- Altun, Z.F. and Hall, D.H. 2008. Handbook of *C. elegans* Anatomy. In WormAtlas.
- Alvarez-Saavedra, E., and Horvitz, H.R. (2010). Many families of *C. elegans* microRNAs are not essential for development or viability. *Curr Biol* *20*, 367-373.
- Ambros, V. (1989). A hierarchy of regulatory genes controls a larva-to-adult developmental switch in *C. elegans*. *Cell* *57*, 49-57.
- Ambros, V. (2011). MicroRNAs and developmental timing. *Curr Opin Genet Dev* *21*, 511-517.
- Ambros, V., and Horvitz, H.R. (1984). Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* *226*, 409-416.
- Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R., and Luhrmann, R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA* *11*, 717-727.
- Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K., and Tabara, H. (2007). In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. *Embo J* *26*, 5007-5019.
- Arasu, P., Wightman, B., and Ruvkun, G. (1991). Temporal regulation of *lin-14* by the antagonistic action of two other heterochronic genes, *lin-4* and *lin-28*. *Genes Dev* *5*, 1825-1833.

- Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T., *et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* *442*, 203-207.
- Aravin, A.A., Naumova, N.M., Tulin, A.V., Vagin, V.V., Rozovsky, Y.M., and Gvozdev, V.A. (2001). Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol* *11*, 1017-1027.
- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* *122*, 553-563.
- Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* *116*, 281-297.
- Batista, P.J., Ruby, J.G., Claycomb, J.M., Chiang, R., Fahlgren, N., Kasschau, K.D., Chaves, D.A., Gu, W., Vasale, J.J., Duan, S., *et al.* (2008). PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in *C. elegans*. *Mol Cell* *31*, 67-78.
- Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* *409*, 363-366.
- Bhaskaran, M., Wang, Y., Zhang, H., Weng, T., Baviskar, P., Guo, Y., Gou, D., and Liu, L. (2009). MicroRNA-127 modulates fetal lung development. *Physiol Genomics* *37*, 268-278.
- Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. *Embo J* *17*, 170-180.
- Bohnsack, M.T., Czapinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* *10*, 185-191.
- Boland, A., Huntzinger, E., Schmidt, S., Izaurralde, E., and Weichenrieder, O. (2011). Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. *Proc Natl Acad Sci U S A* *108*, 10466-10471.
- Boland, A., Tritschler, F., Heimstadt, S., Izaurralde, E., and Weichenrieder, O. (2010). Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. *EMBO Rep* *11*, 522-527.

- Borchert, G.M., Lanier, W., and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* 13, 1097-1101.
- Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuys, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960-964.
- Bukhari SI, Vasquez-Rifo A, Gagné D, Paquet ER, Zetka M, Robert C, Masson JY, Simard MJ. (2012). The microRNA pathway controls germ cell proliferation and differentiation in *C. elegans*. *Cell Res* 6, 1034-45.
- Bussing, I., Yang, J.S., Lai, E.C., and Grosshans, H. The nuclear export receptor XPO-1 supports primary miRNA processing in *C. elegans* and *Drosophila*. *Embo J* 29, 1830-1839.
- Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10, 1957-1966.
- Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., *et al.* (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* 99, 15524-15529.
- Care, A., Catalucci, D., Felicetti, F., Bonci, D., Addario, A., Gallo, P., Bang, M.L., Segnalini, P., Gu, Y., Dalton, N.D., *et al.* (2007). MicroRNA-133 controls cardiac hypertrophy. *Nat Med* 13, 613-618.
- Carmell, M.A., and Hannon, G.J. (2004). RNase III enzymes and the initiation of gene silencing. *Nat Struct Mol Biol* 11, 214-218.
- Cerutti, L., Mian, N., and Bateman, A. (2000). Domains in gene silencing and cell differentiation proteins: the novel PAZ domain and redefinition of the Piwi domain. *Trends Biochem Sci* 25, 481-482.
- Chang, S., Johnston, R.J., Jr., Frokjaer-Jensen, C., Lockery, S., and Hobert, O. (2004). MicroRNAs act sequentially and asymmetrically to control chemosensory laterality in the nematode. *Nature* 430, 785-789.
- Chatterjee, S., Fasler, M., Bussing, I., and Grosshans, H. (2011). Target-mediated protection of endogenous microRNAs in *C. elegans*. *Dev Cell* 20, 388-396.
- Chatterjee, S., and Grosshans, H. (2009). Active turnover modulates mature microRNA activity in *Caenorhabditis elegans*. *Nature* 461, 546-549.

- Cheloufi, S., Dos Santos, C.O., Chong, M.M., and Hannon, G.J. (2010). A Dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465, 584-589.
- Chendrimada, T.P., Finn, K.J., Ji, X., Baillat, D., Gregory, R.I., Liebhaber, S.A., Pasquinelli, A.E., and Shiekhattar, R. (2007). MicroRNA silencing through RISC recruitment of eIF6. *Nature* 447, 823-828.
- Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740-744.
- Cheng, Y., Liu, X., Zhang, S., Lin, Y., Yang, J., and Zhang, C. (2009). MicroRNA-21 protects against the H<sub>2</sub>O<sub>2</sub>-induced injury on cardiac myocytes via its target gene PDCD4. *J Mol Cell Cardiol* 47, 5-14.
- Chu, C.Y., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol* 4, e210.
- Cifuentes, D., Xue, H., Taylor, D.W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G.J., Lawson, N.D., *et al.* (2010). A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328, 1694-1698.
- Clark, A.M., Goldstein, L.D., Tevlin, M., Tavare, S., Shaham, S., and Miska, E.A. (2010). The microRNA miR-124 controls gene expression in the sensory nervous system of *Caenorhabditis elegans*. *Nucleic Acids Res* 38, 3780-3793.
- Conrad, R., Thomas, J., Spieth, J., and Blumenthal, T. (1991). Insertion of part of an intron into the 5' untranslated region of a *Caenorhabditis elegans* gene converts it into a trans-spliced gene. *Mol Cell Biol* 11, 1921-1926.
- Cordes, K.R., Sheehy, N.T., White, M.P., Berry, E.C., Morton, S.U., Muth, A.N., Lee, T.H., Miano, J.M., Ivey, K.N., and Srivastava, D. (2009). miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* 460, 705-710.
- Costinean, S., Zanesi, N., Pekarsky, Y., Tili, E., Volinia, S., Heerema, N., and Croce, C.M. (2006). Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. *Proc Natl Acad Sci U S A* 103, 7024-7029.
- Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev* 12, 3715-3727.

- Das, P.P., Bagijn, M.P., Goldstein, L.D., Woolford, J.R., Lehrbach, N.J., Sapetschnig, A., Buhecha, H.R., Gilchrist, M.J., Howe, K.L., Stark, R., *et al.* (2008). Piwi and piRNAs act upstream of an endogenous siRNA pathway to suppress Tc3 transposon mobility in the *Caenorhabditis elegans* germline. *Mol Cell* 31, 79-90.
- Davis, B.N., Hilyard, A.C., Lagna, G., and Hata, A. (2008). SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* 454, 56-61.
- Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231-235.
- Dews, M., Homayouni, A., Yu, D., Murphy, D., Sevignani, C., Wentzel, E., Furth, E.E., Lee, W.M., Enders, G.H., Mendell, J.T., *et al.* (2006). Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. *Nat Genet* 38, 1060-1065.
- Diederichs, S., and Haber, D.A. (2007). Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* 131, 1097-1108.
- Dillhoff, M., Liu, J., Frankel, W., Croce, C., and Bloomston, M. (2008). MicroRNA-21 is overexpressed in pancreatic cancer and a potential predictor of survival. *J Gastrointest Surg* 12, 2171-2176.
- Ding, L., Spencer, A., Morita, K., and Han, M. (2005). The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in *C. elegans*. *Mol Cell* 19, 437-447.
- Djuranovic, S., Zinchenko, M.K., Hur, J.K., Nahvi, A., Brunelle, J.L., Rogers, E.J., and Green, R. (2010). Allosteric regulation of Argonaute proteins by miRNAs. *Nat Struct Mol Biol* 17, 144-150.
- 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.
- Eis, P.S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M.F., Lund, E., and Dahlberg, J.E. (2005). Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc Natl Acad Sci U S A* 102, 3627-3632.
- Ellerby, L.M., Andrusiak, R.L., Wellington, C.L., Hackam, A.S., Propp, S.S., Wood, J.D., Sharp, A.H., Margolis, R.L., Ross, C.A., Salvesen, G.S., *et al.* (1999). Cleavage of

atrophin-1 at caspase site aspartic acid 109 modulates cytotoxicity. *J Biol Chem* 274, 8730-8736.

Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* 8, 9-22.

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

Felix, M.A., Ashe, A., Piffaretti, J., Wu, G., Nuez, I., Belicard, T., Jiang, Y., Zhao, G., Franz, C.J., Goldstein, L.D., *et al.* (2011). Natural and experimental infection of *Caenorhabditis* nematodes by novel viruses related to nodaviruses. *PLoS Biol* 9, e1000586.

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

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

Fish, J.E., Santoro, M.M., Morton, S.U., Yu, S., Yeh, R.F., Wythe, J.D., Ivey, K.N., Bruneau, B.G., Stainier, D.Y., and Srivastava, D. (2008). miR-126 regulates angiogenic signaling and vascular integrity. *Dev Cell* 15, 272-284.

Flores-Jasso, C.F., Arenas-Huertero, C., Reyes, J.L., Contreras-Cubas, C., Covarrubias, A., and Vaca, L. (2009). First step in pre-miRNAs processing by human Dicer. *Acta Pharmacol Sin* 30, 1177-1185.

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

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

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

- Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* *124*, 355-366.
- Ghildiyal, M., Xu, J., Seitz, H., Weng, Z., and Zamore, P.D. (2010). Sorting of *Drosophila* small silencing RNAs partitions microRNA\* strands into the RNA interference pathway. *RNA* *16*, 43-56.
- Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: an expanding universe. *Nat Rev Genet* *10*, 94-108.
- Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* *312*, 75-79.
- Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* *442*, 199-202.
- 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.
- Grosshans, H., Johnson, T., Reinert, K.L., Gerstein, M., and Slack, F.J. (2005). The temporal patterning microRNA let-7 regulates several transcription factors at the larval to adult transition in *C. elegans*. *Dev Cell* *8*, 321-330.
- Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H., and Siomi, M.C. (2007). A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science* *315*, 1587-1590.
- Haase, A.D., Jaskiewicz, L., Zhang, H., Laine, S., Sack, R., Gatignol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep* *6*, 961-967.
- Hammell, C.M., Karp, X., and Ambros, V. (2009a). A feedback circuit involving let-7-family miRNAs and DAF-12 integrates environmental signals and developmental timing in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* *106*, 18668-18673.
- Hammell, C.M., Lubin, I., Boag, P.R., Blackwell, T.K., and Ambros, V. (2009b). *nhl-2* Modulates microRNA activity in *Caenorhabditis elegans*. *Cell* *136*, 926-938.

- Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 404, 293-296.
- Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 293, 1146-1150.
- Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 18, 3016-3027.
- Haren, L., Ton-Hoang, B., and Chandler, M. (1999). Integrating DNA: transposases and retroviral integrases. *Annu Rev Microbiol* 53, 245-281.
- Harris, A.N., and Macdonald, P.M. (2001). Aubergine encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. *Development* 128, 2823-2832.
- Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Mol Cell* 32, 276-284.
- Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* 138, 696-708.
- Hobert, O. (2006). Architecture of a microRNA-controlled gene regulatory network that diversifies neuronal cell fates. *Cold Spring Harb Symp Quant Biol* 71, 181-188.
- Hobert, O., Johnston, R.J., Jr., and Chang, S. (2002). Left-right asymmetry in the nervous system: the *Caenorhabditis elegans* model. *Nat Rev Neurosci* 3, 629-640.
- Horwich, M.D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., and Zamore, P.D. (2007). The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol* 17, 1265-1272.
- Humphreys, D.T., Westman, B.J., Martin, D.I., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc Natl Acad Sci U S A* 102, 16961-16966.
- Hutvagner, G., and Simard, M.J. (2008). Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol* 9, 22-32.

- Ibanez-Ventoso, C., Vora, M., and Driscoll, M. (2008). Sequence relationships among *C. elegans*, *D. melanogaster* and human microRNAs highlight the extensive conservation of microRNAs in biology. *PLoS One* 3, e2818.
- Ingelfinger, D., Arndt-Jovin, D.J., Luhrmann, R., and Achsel, T. (2002). The human LSM1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *RNA* 8, 1489-1501.
- Isik, M., Korswagen, H.C., and Berezikov, E. (2010). Expression patterns of intronic microRNAs in *Caenorhabditis elegans*. *Silence* 1, 5.
- Jakymiw, A., Lian, S., Eystathiou, T., Li, S., Satoh, M., Hamel, J.C., Fritzler, M.J., and Chan, E.K. (2005). Disruption of GW bodies impairs mammalian RNA interference. *Nat Cell Biol* 7, 1267-1274.
- Johnnidis, J.B., Harris, M.H., Wheeler, R.T., Stehling-Sun, S., Lam, M.H., Kirak, O., Brummelkamp, T.R., Fleming, M.D., and Camargo, F.D. (2008). Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. *Nature* 451, 1125-1129.
- Johnson, C.D., Esquela-Kerscher, A., Stefani, G., Byrom, M., Kelnar, K., Ovcharenko, D., Wilson, M., Wang, X., Shelton, J., Shingara, J., *et al.* (2007). The let-7 microRNA represses cell proliferation pathways in human cells. *Cancer Res* 67, 7713-7722.
- Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS is regulated by the let-7 microRNA family. *Cell* 120, 635-647.
- Jose, A.M., Garcia, G.A., and Hunter, C.P. (2011). Two classes of silencing RNAs move between *Caenorhabditis elegans* tissues. *Nat Struct Mol Biol*.
- Karres, J.S., Hilgers, V., Carrera, I., Treisman, J., and Cohen, S.M. (2007). The conserved microRNA miR-8 tunes atrophin levels to prevent neurodegeneration in *Drosophila*. *Cell* 131, 136-145.
- Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* 15, 2654-2659.
- Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209-216.

- Kim, V.N. (2004). MicroRNA precursors in motion: exportin-5 mediates their nuclear export. *Trends Cell Biol* *14*, 156-159.
- Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* *10*, 126-139.
- Kimble, J., and Hirsh, D. (1979). The postembryonic cell lineages of the hermaphrodite and male gonads in *Caenorhabditis elegans*. *Dev Biol* *70*, 396-417.
- Kinch, L.N., and Grishin, N.V. (2009). The human Ago2 MC region does not contain an eIF4E-like mRNA cap binding motif. *Biol Direct* *4*, 2.
- Kiriakidou, M., Tan, G.S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P.T., and Mourelatos, Z. (2007). An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* *129*, 1141-1151.
- Kirino, Y., and Mourelatos, Z. (2007). Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3' termini. *Nat Struct Mol Biol* *14*, 347-348.
- Kooter, J.M., Matzke, M.A., and Meyer, P. (1999). Listening to the silent genes: transgene silencing, gene regulation and pathogen control. *Trends Plant Sci* *4*, 340-347.
- Lagos-Quintana, M., Rauhut, R., Yalcin, A., Meyer, J., Lendeckel, W., and Tuschl, T. (2002). Identification of tissue-specific microRNAs from mouse. *Curr Biol* *12*, 735-739.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* *75*, 843-854.
- Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., *et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* *425*, 415-419.
- Lee, Y., Hur, I., Park, S.Y., Kim, Y.K., Suh, M.R., and Kim, V.N. (2006). The role of PACT in the RNA silencing pathway. *Embo J* *25*, 522-532.
- Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *Embo J* *21*, 4663-4670.
- Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *Embo J* *23*, 4051-4060.
- Lee, Y.S., and Dutta, A. (2007). The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev* *21*, 1025-1030.
- Lee, Y.S., and Dutta, A. (2009). MicroRNAs in cancer. *Annu Rev Pathol* *4*, 199-227.

- Li, Q.J., Chau, J., Ebert, P.J., Sylvester, G., Min, H., Liu, G., Braich, R., Manoharan, M., Soutschek, J., Skare, P., *et al.* (2007). miR-181a is an intrinsic modulator of T cell sensitivity and selection. *Cell* *129*, 147-161.
- Lin, H., and Spradling, A.C. (1997). A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. *Development* *124*, 2463-2476.
- Lin, S.Y., Johnson, S.M., Abraham, M., Vella, M.C., Pasquinelli, A., Gamberi, C., Gottlieb, E., and Slack, F.J. (2003). The *C. elegans* hunchback homolog, hbl-1, controls temporal patterning and is a probable microRNA target. *Dev Cell* *4*, 639-650.
- Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2003). Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. *Nature* *426*, 465-469.
- Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* *11*, 576-577.
- Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science* *305*, 1437-1441.
- Liu, J., Rivas, F.V., Wohlschlegel, J., Yates, J.R., 3rd, Parker, R., and Hannon, G.J. (2005). A role for the P-body component GW182 in microRNA function. *Nat Cell Biol* *7*, 1261-1266.
- Liu, N., Bezprozvannaya, S., Williams, A.H., Qi, X., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2008). microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev* *22*, 3242-3254.
- Lu, Y., Thomson, J.M., Wong, H.Y., Hammond, S.M., and Hogan, B.L. (2007). Transgenic over-expression of the microRNA miR-17-92 cluster promotes proliferation and inhibits differentiation of lung epithelial progenitor cells. *Dev Biol* *310*, 442-453.
- Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* *429*, 318-322.
- Ma, J.B., Yuan, Y.R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* *434*, 666-670.

- MacRae, I.J., and Doudna, J.A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr Opin Struct Biol* 17, 138-145.
- Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. *Science* 311, 195-198.
- Makarova, K.S., Wolf, Y.I., van der Oost, J., and Koonin, E.V. (2009). Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol Direct* 4, 29.
- Makeyev, E.V., Zhang, J., Carrasco, M.A., and Maniatis, T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Mol Cell* 27, 435-448.
- Martinez, N.J., Ow, M.C., Reece-Hoyes, J.S., Barrasa, M.I., Ambros, V.R., and Walhout, A.J. (2008). Genome-scale spatiotemporal analysis of *Caenorhabditis elegans* microRNA promoter activity. *Genome Res* 18, 2005-2015.
- Mathonnet, G., Fabian, M.R., Svitkin, Y.V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W.C., Darzynkiewicz, E., Pillai, R.S., *et al.* (2007). MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science* 317, 1764-1767.
- Mi, S., Cai, T., Hu, Y., Chen, Y., Hodges, E., Ni, F., Wu, L., Li, S., Zhou, H., Long, C., *et al.* (2008). Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide. *Cell* 133, 116-127.
- Miska, E.A., Alvarez-Saavedra, E., Abbott, A.L., Lau, N.C., Hellman, A.B., McGonagle, S.M., Bartel, D.P., Ambros, V.R., and Horvitz, H.R. (2007). Most *Caenorhabditis elegans* microRNAs are individually not essential for development or viability. *PLoS Genet* 3, e215.
- Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M.C. (2005). Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev* 19, 2837-2848.
- Montgomery, M.K., Xu, S., and Fire, A. (1998). RNA as a target of double-stranded RNA-mediated genetic interference in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A* 95, 15502-15507.

- Moriyama, T., Ohuchida, K., Mizumoto, K., Yu, J., Sato, N., Nabae, T., Takahata, S., Toma, H., Nagai, E., and Tanaka, M. (2009). MicroRNA-21 modulates biological functions of pancreatic cancer cells including their proliferation, invasion, and chemoresistance. *Mol Cancer Ther* 8, 1067-1074.
- Moussian, B., Schoof, H., Haecker, A., Jurgens, G., and Laux, T. (1998). Role of the ZWILLE gene in the regulation of central shoot meristem cell fate during Arabidopsis embryogenesis. *Embo J* 17, 1799-1809.
- Napoli, C., Lemieux, C., and Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* 2, 279-289.
- Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *Rna* 14, 1539-1549.
- Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 95, 14687-14692.
- Nishida, K.M., Saito, K., Mori, T., Kawamura, Y., Nagami-Okada, T., Inagaki, S., Siomi, H., and Siomi, M.C. (2007). Gene silencing mechanisms mediated by Aubergine piRNA complexes in *Drosophila* male gonad. *Rna* 13, 1911-1922.
- O'Donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 435, 839-843.
- Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., and Miyauchi, K. (2007). The 3' termini of mouse Piwi-interacting RNAs are 2'-O-methylated. *Nat Struct Mol Biol* 14, 349-350.
- Olsen, P.H., and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol* 216, 671-680.
- Pak, J., and Fire, A. (2007). Distinct populations of primary and secondary effectors during RNAi in *C. elegans*. *Science* 315, 241-244.
- Park, J.E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* 475, 201-205.
- Parker, J.S., Roe, S.M., and Barford, D. (2004). Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *Embo J* 23, 4727-4737.

- Parker, J.S., Roe, S.M., and Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* 434, 663-666.
- Parrish, S., and Fire, A. (2001). Distinct roles for RDE-1 and RDE-4 during RNA interference in *Caenorhabditis elegans*. *Rna* 7, 1397-1402.
- Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degan, B., Muller, P., *et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408, 86-89.
- Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* 309, 1573-1576.
- Poy, M.N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P.E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., *et al.* (2004). A pancreatic islet-specific microRNA regulates insulin secretion. *Nature* 432, 226-230.
- Poy, M.N., Hausser, J., Trajkovski, M., Braun, M., Collins, S., Rorsman, P., Zavolan, M., and Stoffel, M. (2009). miR-375 maintains normal pancreatic alpha- and beta-cell mass. *Proc Natl Acad Sci U S A* 106, 5813-5818.
- Rashid, U.J., Paterok, D., Koglin, A., Gohlke, H., Piehler, J., and Chen, J.C. (2007). Structure of *Aquifex aeolicus* argonaute highlights conformational flexibility of the PAZ domain as a potential regulator of RNA-induced silencing complex function. *J Biol Chem* 282, 13824-13832.
- Rechavi, O., Minevich, G., and Hobert, O. (2011). Transgenerational Inheritance of an Acquired Small RNA-Based Antiviral Response in *C. elegans*. *Cell* 147, 1248-1256.
- Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, A.E., Bettinger, J.C., Rougvie, A.E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901-906.
- Rivas, F.V., Tolia, N.H., Song, J.J., Aragon, J.P., Liu, J., Hannon, G.J., and Joshua-Tor, L. (2005). Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* 12, 340-349.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res* 14, 1902-1910.

- 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.
- Russell, D.W., and Spremulli, L.L. (1978). Identification of a wheat germ ribosome dissociation factor distinct from initiation factor eIF-3. *J Biol Chem* 253, 6647-6649.
- Ruvkun, G., and Giusto, J. (1989). The *Caenorhabditis elegans* heterochronic gene *lin-14* encodes a nuclear protein that forms a temporal developmental switch. *Nature* 338, 313-319.
- Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. *PLoS Biol* 3, e235.
- Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., Siomi, H., and Siomi, M.C. (2006). Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev* 20, 2214-2222.
- Sarot, E., Payen-Groschene, G., Bucheton, A., and Pelisson, A. (2004). Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the *Drosophila melanogaster* flamenco gene. *Genetics* 166, 1313-1321.
- Sayed, D., and Abdellatif, M. (2011). MicroRNAs in Development and Disease. *Physiol Rev* 91, 827-887.
- Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199-208.
- Seggerson, K., Tang, L., and Moss, E.G. (2002). Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene *lin-28* after translation initiation. *Dev Biol* 243, 215-225.
- Sempere, L.F., Freemantle, S., Pitha-Rowe, I., Moss, E., Dmitrovsky, E., and Ambros, V. (2004). Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation. *Genome Biol* 5, R13.
- Sheth, U., and Parker, R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 300, 805-808.

- Shyu, A.B., Wilkinson, M.F., and van Hoof, A. (2008). Messenger RNA regulation: to translate or to degrade. *Embo J* 27, 471-481.
- Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* 107, 465-476.
- Sijen, T., Steiner, F.A., Thijssen, K.L., and Plasterk, R.H. (2007). Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* 315, 244-247.
- Slack, F.J., Basson, M., Liu, Z., Ambros, V., Horvitz, H.R., and Ruvkun, G. (2000). The *lin-41* RBCC gene acts in the *C. elegans* heterochronic pathway between the *let-7* regulatory RNA and the LIN-29 transcription factor. *Mol Cell* 5, 659-669.
- Sardon, A., Spoerke, J.M., Stacey, S.C., Klein, M.E., Mackin, N., and Maine, E.M. (2000). EGO-1 is related to RNA-directed RNA polymerase and functions in germ-line development and RNA interference in *C. elegans*. *Curr Biol* 10, 169-178.
- Smirnova, L., Grafe, A., Seiler, A., Schumacher, S., Nitsch, R., and Wulczyn, F.G. (2005). Regulation of miRNA expression during neural cell specification. *Eur J Neurosci* 21, 1469-1477.
- Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* 10, 1026-1032.
- Song, J.J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* 305, 1434-1437.
- Steiner, F.A., Okihara, K.L., Hoogstrate, S.W., Sijen, T., and Ketting, R.F. (2009). RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in *Caenorhabditis elegans*. *Nat Struct Mol Biol* 16, 207-211.
- Sternberg, P.W. (2005). *Wormbook*
- Sulston, J.E., and Horvitz, H.R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. *Dev Biol* 56, 110-156.
- Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol* 100, 64-119.
- Tabara, H., Grishok, A., and Mello, C.C. (1998). RNAi in *C. elegans*: soaking in the genome sequence. *Science* 282, 430-431.

- Tabara, H., Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The *rde-1* gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* *99*, 123-132.
- Tabara, H., Yigit, E., Siomi, H., and Mello, C.C. (2002). The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-box helicase to direct RNAi in *C. elegans*. *Cell* *109*, 861-871.
- Tahbaz, N., Kolb, F.A., Zhang, H., Jaronczyk, K., Filipowicz, W., and Hobman, T.C. (2004). Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. *EMBO Rep* *5*, 189-194.
- Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., *et al.* (2004). Reduced expression of the *let-7* microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* *64*, 3753-3756.
- Theobald, D.L., Mitton-Fry, R.M., and Wuttke, D.S. (2003). Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* *32*, 115-133.
- Thermann, R., and Hentze, M.W. (2007). *Drosophila* miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature* *447*, 875-878.
- Till, S., Lejeune, E., Thermann, R., Bortfeld, M., Hothorn, M., Enderle, D., Heinrich, C., Hentze, M.W., and Ladurner, A.G. (2007). A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. *Nat Struct Mol Biol* *14*, 897-903.
- Timmons, L., and Fire, A. (1998). Specific interference by ingested dsRNA. *Nature* *395*, 854.
- Tolia, N.H., and Joshua-Tor, L. (2007). Slicer and the argonautes. *Nat Chem Biol* *3*, 36-43.
- Vagin, V.V., Klenov, M.S., Kalmykova, A.I., Stolyarenko, A.D., Kotelnikov, R.N., and Gvozdev, V.A. (2004). The RNA interference proteins and vasa locus are involved in the silencing of retrotransposons in the female germline of *Drosophila melanogaster*. *RNA Biol* *1*, 54-58.
- Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* *313*, 320-324.

- Vasquez-Rifo A, Jannot G, Armisen J, Labouesse M, Bukhari SI, Rondeau EL, Miska EA, Simard MJ. Developmental characterization of the microRNA-specific *C. elegans* Argonautes alg-1 and alg-2. *PLoS ONE*, **7**, e33750.
- Vella, M.C., Choi, E.Y., Lin, S.Y., Reinert, K., and Slack, F.J. (2004). The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes Dev* **18**, 132-137.
- Ventura, A., Young, A.G., Winslow, M.M., Lintault, L., Meissner, A., Erkeland, S.J., Newman, J., Bronson, R.T., Crowley, D., Stone, J.R., *et al.* (2008). Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. *Cell* **132**, 875-886.
- Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. *Science* **320**, 97-100.
- Von Ehrenstein, G.a.S.(1980). Cell lineages and development of *Caenorhabditis elegans* and other nematodes, Vol. I Chapter 1 (New York, Academic Press).
- Wang, B., Li, S., Qi, H.H., Chowdhury, D., Shi, Y., and Novina, C.D. (2009a). Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nat Struct Mol Biol* **16**, 1259-1266.
- Wang, G., and Reinke, V. (2008). A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. *Curr Biol* **18**, 861-867.
- Wang, S., Aurora, A.B., Johnson, B.A., Qi, X., McAnally, J., Hill, J.A., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2008a). The endothelial-specific microRNA miR-126 governs vascular integrity and angiogenesis. *Dev Cell* **15**, 261-271.
- Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T., and Patel, D.J. (2008b). Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921-926.
- Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G.S., Tuschl, T., and Patel, D.J. (2009b). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754-761.
- Warf, M.B., Johnson, W.E., and Bass, B.L. (2011). Improved annotation of *C. elegans* microRNAs by deep sequencing reveals structures associated with processing by Drosha and Dicer. *RNA* **17**, 563-577.

- Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862.
- Winston, W.M., Molodowitch, C., and Hunter, C.P. (2002). Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science* 295, 2456-2459.
- Winston, W.M., Sutherlin, M., Wright, A.J., Feinberg, E.H., and Hunter, C.P. (2007). *Caenorhabditis elegans* SID-2 is required for environmental RNA interference. *Proc Natl Acad Sci U S A* 104, 10565-10570.
- Wood, W.B. (1988). In *The nematode C. elegans*
- Wu, E., Thivierge, C., Flamand, M., Mathonnet, G., Vashisht, A.A., Wohlschlegel, J., Fabian, M.R., Sonenberg, N., and Duchaine, T.F. (2010). Pervasive and cooperative deadenylation of 3'UTRs by embryonic microRNA families. *Mol Cell* 40, 558-570.
- Wu, L., Fan, J., and Belasco, J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. *Proc Natl Acad Sci U S A* 103, 4034-4039.
- Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. *Nature* 426, 468-474.
- Yang, Z., Ebright, Y.W., Yu, B., and Chen, X. (2006). HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res* 34, 667-675.
- Ye, X., Paroo, Z., and Liu, Q. (2007). Functional anatomy of the *Drosophila* microRNA-generating enzyme. *J Biol Chem* 282, 28373-28378.
- Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 17, 3011-3016.
- Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J., and Mello, C.C. (2006). Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* 127, 747-757.
- Yoo, A.S., and Greenwald, I. (2005). LIN-12/Notch activation leads to microRNA-mediated down-regulation of Vav in *C. elegans*. *Science* 310, 1330-1333.
- Yu, J., Wang, F., Yang, G.H., Wang, F.L., Ma, Y.N., Du, Z.W., and Zhang, J.W. (2006). Human microRNA clusters: genomic organization and expression profile in leukemia cell lines. *Biochem Biophys Res Commun* 349, 59-68.

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

Zhang, L., Ding, L., Cheung, T.H., Dong, M.Q., Chen, J., Sewell, A.K., Liu, X., Yates, J.R., 3rd, and Han, M. (2007). Systematic identification of *C. elegans* miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol Cell* 28, 598-613.

Zhang, L., Huang, J., Yang, N., Greshock, J., Megraw, M.S., Giannakakis, A., Liang, S., Naylor, T.L., Barchetti, A., Ward, M.R., *et al.* (2006). microRNAs exhibit high frequency genomic alterations in human cancer. *Proc Natl Acad Sci U S A* 103, 9136-9141.

Zhao, Y., Samal, E., and Srivastava, D. (2005). Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* 436, 214-220.

Zisoulis, D.G., Lovci, M.T., Wilbert, M.L., Hutt, K.R., Liang, T.Y., Pasquinelli, A.E., and Yeo, G.W. (2010). Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat Struct Mol Biol* 17, 173-179.

# Annexe 1

**The slicing activity of miRNA-specific Argonautes is essential for the miRNA pathway  
in *C. elegans***

Samir Bouasker and Martin J. Simard\*

Laval University Cancer Research Centre, Hôtel-Dieu de Québec (CHUQ), Quebec  
City, Québec G1R 2J6, Canada

\*Corresponding Author:

Email: [Martin.Simard@crhdq.ulaval.ca](mailto:Martin.Simard@crhdq.ulaval.ca)

**Running title:** *The miRNA pathway requires slicer Argonautes ALGs*

**Keywords:** ALG-1; ALG-2; Argonaute; microRNA; *C. elegans*

## ABSTRACT

Among the set of Argonautes proteins encoded by metazoan genomes, some have conserved amino acids important for catalytic or slicing activity. The functional significance of these residues in miRNA-specific Argonautes in animals is still unclear since miRNAs do not induce site-specific cleavage of targeted mRNAs, unlike siRNAs. Here, we report that miRNA-specific ALG-1 and ALG-2 Argonautes from *C. elegans* possess the slicing activity normally implicated in the siRNA-silencing pathway. We also find that ALG-1/2 can bind and use a Dicer-processed miRNA duplex to target mRNAs, suggesting an ability to displace RNA strands. Importantly, the slicing activity of ALG-1 or ALG-2 is essential for the miRNA pathway *in vivo*, as shown by the accumulation of truncated miRNA precursors and altered miRISC formation. Taken together, our data demonstrate that the slicing activity of Argonautes contributes to a new and unexpected step in the canonical miRNA pathway that occurs prior to miRISC loading in animals.

## INTRODUCTION

Argonaute proteins act at different steps of the miRNA-mediated gene-silencing pathway. They form a complex with Dicer and interact with miRNAs to form the miRISC: the effector complex that targets complementary mRNA (reviewed in (1)). All members of this protein family share signature domains: PAZ, Mid and PIWI domains (2). While the PAZ and Mid provide the small RNA binding capacity to the protein (3-6), the PIWI domain can confer a catalytic or slicing activity to a specific subset of Argonautes that possesses three conserved amino acid residues (DDH) in the catalytic center (reviewed in (7)). In humans, Argonaute-encoding genes are phylogenetically grouped into two different clades called the Argonaute-like and the Piwi-like proteins that are respectively required for the miRNA and piRNA pathways (reviewed in (2)). In the nematode *Caenorhabditis elegans*, specific Argonaute genes found in the animal genome form a Worm specific protein clade called the WAGOs, which are important for endogenous (endo-) and exogenous (exo-) RNAi pathways (8). While a number of observations support the role of slicing Argonautes in the piRNA pathway as well as the exo- and endo-RNAi pathways (9-16), the importance of this enzymatic activity for miRNA-specific Argonautes remains unclear in animals.

Recently, two studies reported that the mouse and zebrafish slicer Ago2 is essential for the production of only one specific miRNA that is not subject to canonical Dicer-dependent biogenesis pathway (17,18). Aside from this unique case, current hypotheses suggest that the slicer activity of specific Argonaute proteins could be involved in the formation of miRISC, wherein they could cleave perfectly paired small RNA duplexes and promote the degradation of the passenger strand (11,14,16). It has also been proposed that the slicer activity of human Ago2 could cleave the precursor miRNA molecule to stimulate

production of the mature form by Dicer (19). Thus far, there is no *in vivo* evidence that supports the functional importance of Argonaute slicer activity in the canonical biogenesis of miRNAs.

To address the implications of the slicing activity of miRNA-specific Argonautes in animals, we set out to characterize *C. elegans* ALG-1 and ALG-2: the two Argonautes that carry a DDH motif and are essential for the miRNA pathway (20). Here, we demonstrate that ALG-1 and ALG-2 Argonautes possess slicer activity and, that this activity is essential for viability. We also determine that the slicing activity of ALG-1 and ALG-2 is important to coordinate Dicer cleavage and thus, is required for the formation of miRISC in *C. elegans*.

## MATERIALS & METHODS

### *Recombinant ALG-1 and ALG-2 protein preparation and purification*

BL-21 codon plus RIPL bacteria (Stratagene) transformed with GST-tagged ALG-1 and ALG-2 plasmids, were grown at 37°C in LB media containing ampicillin and chloramphenicol. When OD reached 0.4, 1mM IPTG was added to the media and the cultures were grown overnight at 16°C. The bacterial pellet was resuspended in cold PBS 1X with protease inhibitor cocktail from Roche and 100 µM of PMSF. The pellet was disrupted using a Cell Disruption machine (Constant System). The lysate was then incubated for 1h at 4°C in the presence of 0.25% Empigen (Sigma) and cleared by centrifugation. GST-tagged proteins were purified using Glutathione Sepharose 4B column (GE Healthcare) according to the manufacturer's instructions. It is important to take in consideration that a significant amount of RNA molecules from the expression systems

(either bacteria or insect cells) remain associated to purified recombinant Argonautes (data not shown; 21,22) and thus might impinge the efficacy of small RNA loading onto the proteins.

#### *Preparation of RNA molecules for slicing and binding assays*

Synthetic RNAs (Sigma) were phosphorylated and gel purified using a denaturing acrylamide gel. To produce duplexes, an equivalent amount of purified phosphorylated single-stranded RNAs were annealed as described in (23). To confirm the absence of single-stranded RNA in dsRNA duplexes, the nature of the samples were assessed by running a fraction of sample on a native acrylamide gel along with a single-stranded phosphorylated RNA as a size control.

#### *Slicing and binding assays*

The slicing assays were performed at 20°C following the protocol used in (16). The RNA binding experiments were performed as described in (24).

#### *C. elegans methods*

*C. elegans* strains were grown under standard conditions (25). Transgenic lines carrying *alg-1 wt*, *alg-1 AAA*, *gfp::alg-2 wt* or *GFP::alg-2 AAA* were produced using standard procedure (26). Transgenes carrying *alg-1 wt* and *alg-1 AAA* genes were introduced in *alg-1(gk214)* mutant and transgenes expressing tagged *alg-2 wt* and *AAA* genes introduced in *alg-2(ok304)* animals.

To generate RNAi knockdown population, synchronized L1 *C. elegans* population were exposed to bacteria expressing either control, *alg-1* or *alg-2* dsRNA molecules for 4 days at 15°C.

#### *RNA isolation and Northern Blot analysis*

The total RNA from a population of 50,000 animals was obtained by resuspending worm pellet in TRI Reagent (Sigma) and lysed by flash-freezing in liquid nitrogen. 10 µg purified total RNAs was then used for Northern Blot analysis following a protocol from the Bartel Lab (<http://web.wi.mit.edu/bartel/pub/protocols.html>) with EDC (1-ethyl-3-[3-dimethylaminopropyl] carbodiimide hydrochloride) chemical crosslinking (27). P<sup>32</sup> radiolabeled probes to detect miRNAs were synthesized using the IDT Starfire kit (probes sequences can be found in the supplemental online section).

#### *ALG-1 and ALG-2 complexes immunoprecipitation*

Whole worm extracts were prepared from either a population of transgenic *alg-1* animals carrying *alg-1 wt* or *alg-1 AAA* extrachromosomal transgene arrays exposed to *alg-2(RNAi)* or; a population of transgenic *alg-2* animals carrying *gfp::alg-2 wt* or *gfp::alg-2 AAA* extrachromosomal transgene arrays exposed to *alg-1(RNAi)*. Total protein extracts were made as described in (28).

Immunoprecipitations were performed by first pre-clearing whole worm lysate with 20 µl of Protein G Agarose beads (Roche) for 60 minutes at 4°C. The cleared extract was then incubated with beads coated with affinity purified anti-ALG-1 antibody or with beads

coated with monoclonal anti-GFP (Chromotek) for 1 hour at 4°C. Beads were then washed three times with ice-cold lysis buffer, a fraction kept to detect the presence of ALG-1 or GFP-ALG-2 proteins by Western blotting and RNA from the remaining beads was eluted and resolved to monitor *lin-4* mature and passenger strands by Northern blot analysis.

## RESULTS

### *ALG-1 and ALG-2 are slicer Argonautes*

To test whether ALG-1 and ALG-2 Argonaute proteins possess the slicing activity, we produced recombinant GST tagged ALG-1 and ALG-2 proteins (rALG-1 and rALG-2; Figure S1A). Once pre-loaded with single-stranded 21 nt guide RNA (ssRNA; Figure 1A-B), both purified rALG-1 and rALG-2 were able to induce cleavage of a fully complementary 32 nt RNA target only in presence of Magnesium (Figure 1C, Figure S1C, respectively). While we detected a major 15 nt long cleavage product corresponding to cleavage at the predicted site on the target RNA (16,29), two minor products were also produced (Figure 1C). Introducing a single mismatch at position 10 of the guide RNA abrogated the cleavage of the scissile bond between position 15 and 16 of the RNA target (Figure 1D) confirming that base-pairing at the cleavage site is essential for precise target cleavage. When the DDH residues found in the catalytic center of the PIWI domain were mutated to three Alanines (AAA), the slicing activity of both Argonautes was abrogated (Figure 1E and Figure S1D, respectively) without significantly affecting RNA binding ability (Figure S1B). Thus, we conclude that ALG-1 and ALG-2 possess the slicing activity able to induce target cleavage at a precise site. In addition, ALG-1 and ALG-2 display

some capacity to cleave at sites adjacent to the main target site, which may reflect a requirement for accessory factor(s) to “proofread” during ALG-1/2 slicing activity.

*ALG-1 and ALG-2 bind and cleave the passenger strand of Dicer-like products*

In flies and human cells, Argonautes have been involved in the loading, cleaving and unwinding of Dicer processed products (14,16,30). We therefore tested whether ALG-1 and ALG-2 share these functions. Taking in consideration the Dicer enzymatic signature (5' phosphate and 2 nt 3' overhang) along with the thermodynamic stability of the 5' end base-pairing of the guide strand as defined by the asymmetry rule (31), we designed three different small RNA duplexes (Figure 2A). These duplexes were: I) fully complementary (duplex); II) possessing a single central mismatch at position 10 of the guide strand (duplex 1M) and; III) possessing two mismatches at position 4 and 10 (duplex 2M) that mimic a miRNA duplex predicted to be produced by Dicer from a precursor miRNA. We first compared the capacity of ALG-1 and ALG-2 to bind these different small RNA duplexes, and observed that both Argonautes have a greater affinity for a ssRNA than small RNA duplexes. Furthermore, the binding affinity of ALGs to double-stranded small RNAs increases with the presence of mismatches in the duplex (Figure 2B and Figure S2A).

We next determined if ALG-1 and ALG-2 could cleave these small RNA duplexes (Figure 2C). As previously observed for target cleavage with ssRNA, increasing the concentration of Argonaute proteins led to the cleavage of the radiolabeled strand from perfectly paired duplex (Figure 2D and Figure S2B). The introduction of a mismatch altered the cleavage between position 10 and 11, and the miRNA-like duplex was not efficiently cleaved (Figure 2C and Figure S2B). Notably, it has been recently observed that small RNA

molecules remain associated with purified recombinant Argonautes and thus despite extensive purification methods (21,22). It is therefore important to consider that the RNA affinity measurement as well as the efficiency of duplex cleavage might be affected by the presence of bacterial RNA molecules associated to recombinant ALG-1 and ALG-2 proteins. Taken together, our data indicate that ALG-1 and ALG-2 proteins can bind Dicer-processed dsRNA products, and induce the cleavage of a perfectly paired duplex, but cannot cleave a miRNA-like duplex.

*ALG-1 and ALG-2 can use a miRNA-like duplex to induce target cleavage*

Recently, it has been reported that both recombinant and tagged purified human Argonautes can unwind small RNA duplexes containing mismatches (30,32). We therefore decided to test whether *C. elegans* ALG-1 and ALG-2 can use miRNA-like duplexes to cleave a target RNA. To achieve this, we first pre-incubated recombinant Argonautes with various small RNA duplexes (Figure 2A) followed by the addition of a radiolabeled target RNA complementary to the guide strand (Figure 3A). Hence, if ALG-1 and ALG-2 were capable of releasing the passenger strand of the duplex, cleavage of the target RNA would be detected. While the perfectly paired duplex and the duplex containing a central mismatch could barely induce target cleavage, the miRNA-like duplex efficiently triggered cleavage of the RNA target (Figure 3B). However, the miRNA-like duplex was still less effective than ssRNA at inducing ALG-mediated cleavage of the target RNA as at least twice the amount of proteins required to produce similar level of target cleavage when the miRNA-like duplex was used (Figures 3C and S3). These results indicate that *C. elegans* Argonautes can load miRNA-like duplexes and favor strand release to form a complex that is able to induce target cleavage.

*The slicer activities of ALG-1 and ALG-2 are essential for C. elegans viability*

To determine the contribution of ALG-1/2 slicing activities *in vivo*, we performed several rescue experiments in animals with plasmid arrays containing either wild-type or catalytically inactive *alg-1* and *alg-2* genes. While we observed that the *alg-1* wild-type transgene array rescued the *alg-1* mutant defects, the expression of catalytically inactive *alg-1* and *alg-2* genes in their respective mutant backgrounds led to noticeable phenotypes in the transgenic animal population (Table 1). Particularly, we also observed that expression of the *alg-2* wild-type transgene in the *alg-2* mutant background caused lethality in 7% of the animal population (Table 1), indicating that a tight regulation of the ALG-2 protein level is important to maintain miRNA homeostasis.

We then asked what is the contribution of Argonaute slicing activities when both endogenous *alg-1* and *alg-2* genes are absent. Because the genetic alteration of both *alg-1* and *alg-2* genes leads to embryonic lethality (33), we generated animals with reduced *alg-1* and *alg-2* post-embryonic expression by using RNAi to knockdown expression of either the *alg-1* or *alg-2* gene in *alg-2* and *alg-1* mutant animals, respectively. As previously described (20), we also observed that the simultaneous loss of both Argonautes resulted in lethality due to a variety of severe developmental defects (for instance, some animals die during larval stages, while others die from vulval bursting when they reach adulthood; Table 1). In addition, the phenotypes observed were similar to those observed when miRNAs themselves are lost or mutated (20,34-36). These results confirm the essential roles of ALG-1 and ALG-2 in the miRNA pathway. While the *alg-2* wild-type transgene partially rescued the developmental defects of the *alg-1; alg-2* double mutant, the

transgenic expression of the wild-type *alg-1* gene nearly abolished the lethality observed in this mutant population (Table 1). These data suggest a more important contribution of ALG-1 than ALG-2 to the miRNA pathway. Interestingly, when either an *alg-1* or *alg-2* catalytically inactive form was expressed in animals with both *alg-1*; *alg-2* reduced function, all animal populations displayed developmental defects (Table 1). Taken together, these data strongly support the notion that the slicing activity of at least one miRNA-specific Argonaute is essential for *C. elegans* viability.

*The slicer activities of ALG-1 and ALG-2 are critical for the production of miRNAs*

To gather insights about the molecular role(s) of ALG-1/2 slicing activity in an animal miRNA pathway, we monitored the processing of several different miRNAs. Using Northern blots, we examined both the guide and passenger strands of four different types of miRNAs: *lin-4* and *let-7* (miRNAs located on the 5' arm of their precursors); miR-90 (located on the 3' arm of the precursor) and; miR-50 (a miRNA encoded within an intron). When compared to wild-type animals, we first observed that animals carrying loss-of-function alleles of either *alg-1* or *alg-2* gene have generally less mature miRNAs (Figures 4 and S4; compared with SL1 RNA), as previously reported (20,28). We also observed that the loss of the *alg-1* Argonaute leads to the accumulation of precursor forms for all miRNA tested, while only miR-50 and miR-90 precursors accumulate in *alg-2* mutant animals (Figures 4A and S4A). The introduction of a transgene carrying either wild-type or catalytically inactive *alg-1* and *alg-2* genes (*alg-1 AAA* and *alg-2 AAA*) in their respective mutants, restored the level of mature miRNAs to that which was detected in wild-type animals (Figures 4A and S4A). Notably, the levels of *let-7* and miR-90 miRNAs were increased relative to wild-type levels upon the expression of the *alg-1/2* transgenes,

suggesting that some miRNAs may be sensitive to the expression level of miRNA-specific Argonautes (Figures 4A and S4A). Similarly, the accumulation of *lin-4* and *let-7* miRNA precursors that were detected in *alg-1* mutant animals was also abrogated in animals carrying both transgenes (Figure 4A). In contrast, only those animals expressing wild-type ALG-2 Argonaute had the levels of pre-miR-50 and pre-miR-90 restored to the levels observed in wild-type animals (Figure S4A). Thus, these data support that both ALG-1 and ALG-2 are important for the accumulation of miRNAs in *C. elegans* and that some specificity may exist between slicer ALG-1 and ALG-2 Argonautes for the processing of some miRNAs.

We then moved on to test to the impact on various miRNA species when neither ALG-1 nor ALG-2 are slicing competent. We observed that, in the absence of both *alg-1* and *alg-2*, mature miRNA levels decreased and the precursor forms accumulated (Figures 4B and S4B). The introduction of transgenes carrying a wild-type copy of *alg-1* or *alg-2* re-established mature miRNAs to a level equivalent to that of wild-type animals, while the pre-miRNAs species was almost completely depleted (Figure 4B and S4B). When we expressed catalytically inactive forms of ALG-1/ALG-2 (*alg-1 AAA* and *alg-2 AAA*) in animals lacking both endogenous *alg-1* and *alg-2* genes, we observed an accumulation of miRNAs (Figures 4B and S4B). However, an additional RNA molecule that corresponds to the size of a truncated form of precursor miRNA (pre-short) was also observed in these conditions, as supported by the presence of a hybridization signal when either sense or antisense probes for each miRNA were used (Figures 4B and S4B). Several lines of evidence suggested that these truncated forms of the precursor miRNAs correspond to pre-miRNAs lacking the Dicer-produced mature miRNA (or passenger strand): 1) for all

miRNAs tested, miRNAs accumulated to detectable levels (even the *lin-4* passenger strand is detected under these conditions; Figure 4B); 2) the size of the truncated precursor was ~21 nt shorter than full-length precursor miRNAs (Figure S5A) and; 3) sequencing of the cloned miR-50 truncated precursor indicated that more than 50% of the RNA species corresponded exactly to the precursor sequence minus the mature or passenger miRNA (Figure S5B). We therefore conclude that both Argonautes are important for the production and the maintenance of the level of miRNAs in animals. The absence of catalytically competent ALG-1 and ALG-2 leads to the accumulation of a new truncated precursor miRNA species without significantly affecting the production of mature miRNAs.

*ALG-1 and ALG-2 slicer activity is essential to generate miRISC*

Despite the fact that the phenotypes of animals lacking catalytically active ALG-1 and ALG-2 were similar to phenotypes resulting from the loss of miRNAs, we still observed an accumulation of mature miRNAs in these animals. Therefore, other molecular defects should be considered to explain the alteration of miRNA-mediated gene regulation in these animals. One possibility is that in the absence of catalytically active ALG-1/ALG-2, a significant fraction of the miRNA pool may remain associated with the truncated precursor and thus decrease the level of functional miRISC. To address this, we examined *lin-4* miRNA in ALG-1 immunoprecipitated complexes (miRISC) that were isolated from *alg-1*; *alg-2* animals expressing either the wild type or catalytically inactive ALG-1 protein. Both wild type and catalytically inactive ALG-1 proteins share similar affinities for ssRNAs (Figure S1B). While we could not detect truncated pre-miRNA in the catalytically inactive ALG-1 complexes (data not shown), we observed decreases in the level of mature of

miRNAs associated with this ALG-1 mutant (Figure 5). This decrease in miRISC-associated mature *lin-4* correlates with an increase of free *lin-4* miRNA when compared to the levels observed for the wild-type ALG-1 immunoprecipitation experiments (Figure 5). Similarly, less *lin-4* miRNA is associated with the catalytically inactive ALG-2 mutant complexes than with wild-type ALG-2 complexes isolated from *alg-1; alg-2* transgenic animals (Figure S5). Therefore, our data indicate that the production of miRISC is severely altered in animals lacking the miRNA-specific slicing Argonautes.

## DISCUSSION

The biochemical study of ALG-1 and ALG-2 confirms that these Argonautes have retained the molecular characteristics found in catalytically competent Argonautes of *S. pombe*, plants and other animals. The enzymatic activity of ALG-1/2 requires both magnesium (a divalent cation) (16), and the DDH motif within the PIWI domain of the proteins (37). We also observed that recombinant ALG-1/2 can bind small RNA duplexes and is capable of separating strands from miRNA-like molecules that contain mismatches. Similar data have been reported for human Argonautes (30,38). Recent studies performed with *Drosophila* embryo lysates demonstrated that the miRNA-specific Argonaute Ago1 protein also displays a preference for imperfectly paired duplexes and can only form functional miRISC complexes once loaded with these types of RNA molecules (39). In all these conditions, it is possible that the presence of a complementary target RNA to the guide strand contributes to the release of the passenger strand by Argonautes. Therefore, our results support the idea that miRNA-specific animal Argonautes have the capacity to release the miRNA passenger strand from a Dicer-processed miRNA duplex to form the functional miRISC.

The first study of *C. elegans alg-1* and *alg-2* Argonautes by Grishok and collaborators, demonstrated that the loss of both Argonautes led to a decrease of mature *let-7* and *lin-4* miRNAs and an accumulation of pre-*lin-4* molecules (20,40,41). While this initial observation suggested a role for ALG-1 and ALG-2 in the stability and production of miRNAs in *C. elegans*, it was still unclear how these proteins could contribute to miRNA biogenesis. The data presented herein reveal the role of ALG-1 and ALG-2 in miRNA production, and demonstrate for the first time the importance of ALG-1/2 slicing activity in processing multiple miRNAs. Furthermore, our studies unveil the functional differences of these miRNA-specific Argonautes in processing different miRNA species. While the replacement of either ALG-1 or ALG-2 with their respective catalytically inactive forms had no observable effect in animals, the loss of slicing activity for both miRNA-specific Argonautes led to severe developmental defects that are associated with the loss of miRNAs in *C. elegans* (20,34-36). At the molecular level, we observed that the loss of this enzymatic activity in miRNA-specific Argonautes led to the accumulation of mature miRNAs along with the appearance of a shorter form of miRNA precursor molecules. These observations suggest that in absence of slicing-competent miRNA-specific Argonautes, mature miRNAs remain associated to the complementary strand found in precursor miRNA molecules leading to a significant decrease of miRISC that is essential for animal viability. In addition, our data also support that Dicer can cleave independently RNA strands from pre-miRNA substrates.

Several lines of evidence in our data support the idea that Dicer can cleave only one strand of the precursor miRNA: 1) the size of the aberrant precursor miRNA form is about 21 nts

shorter than the pre-miRNA and; 2) the mature form for all miRNA tested was detectable by Northern blot. While many *in vitro* and *in vivo* studies demonstrate that the Dicer enzyme can cleave perfectly paired double-stranded RNA on both strands (40,42-44), some experimental evidence also supports that this enzyme can separately cleave both strands of a miRNA precursor. *In vitro* analysis of *Drosophila* Dcr-1, which is required for the miRNA pathway, demonstrated that a single mutation in the active site of one of RNaseIII domains produced a 21 nt miRNA and a truncated miRNA precursor (45). Recently, it has been shown that cells carrying specific point mutations within each RNaseIII domain of human Dicer display defects in the production of either 5p or 3p miRNAs (46). Biochemical studies on recombinant human Dicer also demonstrate its enzymatic capacity to produce a truncated miRNA precursor and a mature miRNA *in vitro* (47-49). In agreement with those observations, our data indicate that at least one Argonaute competent slicer is essential to enable complete Dicer cleavage of pre-miRNAs. Therefore, in absence of the slicing Argonautes, Dicer may cleave only once and thus affects the formation of miRISC. Alternatively, we can also consider that after Dicer produces a single 21 nt molecule from the pre-miRNA, slicing Argonautes may be important to cleave within the complementary sequence to trigger the degradation of the truncated precursor.

Previous observations indicated that Argonaute proteins and Dicer are found in the same complex and form the RISC loading complex (50-54). It has recently been observed that human Ago2 can bind some pre-miRNA molecules even in absence of Dicer (55). Therefore, the presence of co-factors such as slicing competent Argonautes might be essential to coordinate Dicer cleavage in animals. Future analysis with purified RISC loading complex will help in gaining insight about the contribution of slicing Argonaute

proteins to this process.

#### SUPPLEMENTARY DATA

Supplementary Data are available at NAR online: Supplementary Figures S1 to S5 and supplementary materials & methods.

#### FUNDING

This work was funded by The Cancer Research Society and in part by the Canadian Institutes of Health Research. M.J.S. is a Research Scholar Junior 2 of Fonds de Recherche du Québec-Santé.

#### ACKNOWLEDGMENTS

We would like to thank Dr. Julie Claycomb, Dr. Gyorgy Hutvagner, Dr. Jacques Côté and members of our laboratories for comments on the manuscript. Some nematode strains were provided by the *Caenorhabditis* Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR).

## REFERENCES

1. Kim, V.N., Han, J. and Siomi, M.C. (2009) Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol*, **10**, 126-139.
2. Hutvagner, G. and Simard, M.J. (2008) Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol*, **9**, 22-32.
3. Lingel, A., Simon, B., Izaurralde, E. and Sattler, M. (2004) Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol*, **11**, 576-577.
4. Frank, F., Sonenberg, N. and Nagar, B. (2010) Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*, **465**, 818-822.
5. Wang, Y., Juranek, S., Li, H., Sheng, G., Tuschl, T. and Patel, D.J. (2008) Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*, **456**, 921-926.
6. Ma, J.B., Ye, K. and Patel, D.J. (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*, **429**, 318-322.
7. Tolia, N.H. and Joshua-Tor, L. (2007) Slicer and the argonautes. *Nature chemical biology*, **3**, 36-43.
8. Yigit, E., Batista, P.J., Bei, Y., Pang, K.M., Chen, C.C., Tolia, N.H., Joshua-Tor, L., Mitani, S., Simard, M.J. and Mello, C.C. (2006) Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell*, **127**, 747-757.
9. Rand, T.A., Petersen, S., Du, F. and Wang, X. (2005) Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell*, **123**, 621-629.

10. Irvine, D.V., Zaratiegui, M., Tolia, N.H., Goto, D.B., Chitwood, D.H., Vaughn, M.W., Joshua-Tor, L. and Martienssen, R.A. (2006) Argonaute slicing is required for heterochromatic silencing and spreading. *Science*, **313**, 1134-1137.
11. Leuschner, P.J., Ameres, S.L., Kueng, S. and Martinez, J. (2006) Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep*, **7**, 314-320.
12. Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R. and Hannon, G.J. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*, **128**, 1089-1103.
13. Aoki, K., Moriguchi, H., Yoshioka, T., Okawa, K. and Tabara, H. (2007) In vitro analyses of the production and activity of secondary small interfering RNAs in *C. elegans*. *EMBO J*, **26**, 5007-5019.
14. Matranga, C., Tomari, Y., Shin, C., Bartel, D.P. and Zamore, P.D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell*, **123**, 607-620.
15. Gunawardane, L.S., Saito, K., Nishida, K.M., Miyoshi, K., Kawamura, Y., Nagami, T., Siomi, H. and Siomi, M.C. (2007) A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science*, **315**, 1587-1590.
16. Miyoshi, K., Tsukumo, H., Nagami, T., Siomi, H. and Siomi, M.C. (2005) Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev*, **19**, 2837-2848.
17. Cheloufi, S., Dos Santos, C.O., Chong, M.M. and Hannon, G.J. (2010) A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature*, **465**, 584-589.

18. Cifuentes, D., Xue, H., Taylor, D.W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G.J., Lawson, N. *et al.* (2010) A Novel miRNA Processing Pathway Independent of Dicer Requires Argonaute2 Catalytic Activity. *Science*, **328**, 1694-1698.
19. Diederichs, S. and Haber, D.A. (2007) Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell*, **131**, 1097-1108.
20. 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.
21. Schirle, N.T. and MacRae, I.J. (2012) The crystal structure of human Argonaute2. *Science*, **336**, 1037-1040.
22. Nakanishi, K., Weinberg, D.E., Bartel, D.P. and Patel, D.J. (2012) Structure of yeast Argonaute with guide RNA. *Nature*, **486**, 368-374.
23. Haley, B., Tang, G. and Zamore, P.D. (2003) In vitro analysis of RNA interference in *Drosophila melanogaster*. *Methods*, **30**, 330-336.
24. Djuranovic, S., Zinchenko, M.K., Hur, J.K., Nahvi, A., Brunelle, J.L., Rogers, E.J. and Green, R. (2010) Allosteric regulation of Argonaute proteins by miRNAs. *Nat Struct Mol Biol*, **17**, 144-150.
25. Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics*, **77**, 71-94.
26. Mello, C. and Fire, A. (1995) DNA transformation. *Methods Cell Biol*, **48**, 451-482.
27. Pall, G.S. and Hamilton, A.J. (2008) Improved northern blot method for enhanced detection of small RNA. *Nature protocols*, **3**, 1077-1084.

28. Jannot, G., Bajan, S., Giguere, N.J., Bouasker, S., Banville, I.H., Piquet, S., Hutvagner, G. and Simard, M.J. (2011) The ribosomal protein RACK1 is required for microRNA function in both *C. elegans* and humans. *EMBO Rep*, **12**, 581-586.
29. Rivas, F.V., Tolia, N.H., Song, J.J., Aragon, J.P., Liu, J., Hannon, G.J. and Joshua-Tor, L. (2005) Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol*, **12**, 340-349.
30. Wang, B., Li, S., Qi, H.H., Chowdhury, D., Shi, Y. and Novina, C.D. (2009) Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nat Struct Mol Biol*, **16**, 1259-1266.
31. Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N. and Zamore, P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199-208.
32. Yoda, M., Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q. and Tomari, Y. (2010) ATP-dependent human RISC assembly pathways. *Nat Struct Mol Biol*, **17**, 17-23.
33. Vasquez-Rifo, A., Jannot, G., Armisen, J., Labouesse, M., Bukhari, S.I.A., Rondeau, E.L., Miska, E.A. and Simard, M.J. (2012) Developmental characterization of the microRNA-specific *C. elegans* Argonautes alg-1 and alg-2. *PLoS ONE*, **7**, e33750.
34. Hammell, C.M., Lubin, I., Boag, P.R., Blackwell, T.K. and Ambros, V. (2009) nhl-2 Modulates microRNA activity in *Caenorhabditis elegans*. *Cell*, **136**, 926-938.
35. Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F. and Hannon, G.J. (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature*, **432**, 231-235.

36. Bussing, I., Yang, J.S., Lai, E.C. and Grosshans, H. (2010) The nuclear export receptor XPO-1 supports primary miRNA processing in *C. elegans* and *Drosophila*. *EMBO J*, **29**, 1830-1839.
37. Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L. and Hannon, G.J. (2004) Argonaute2 is the catalytic engine of mammalian RNAi. *Science*, **305**, 1437-1441.
38. Kwak, P.B. and Tomari, Y. (2012) The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat Struct Mol Biol*, **19**, 145-151.
39. Kawamata, T., Seitz, H. and Tomari, Y. (2009) Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat Struct Mol Biol*, **16**, 953-960.
40. Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J. and Plasterk, R.H. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev*, **15**, 2654-2659.
41. 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.
42. Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D. and Doudna, J.A. (2006) Structural basis for double-stranded RNA processing by Dicer. *Science*, **311**, 195-198.
43. Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E. and Filipowicz, W. (2004) Single processing center models for human Dicer and bacterial RNase III. *Cell*, **118**, 57-68.

44. Bernstein, E., Caudy, A.A., Hammond, S.M. and Hannon, G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363-366.
45. Ye, X., Paroo, Z. and Liu, Q. (2007) Functional anatomy of the *Drosophila* microRNA-generating enzyme. *J Biol Chem*, **282**, 28373-28378.
46. Gurtan, A.M., Lu, V., Bhutkar, A. and Sharp, P.A. (2012) In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs. *RNA*, **18**, 1116-1122.
47. Koscianska, E., Starega-Roslan, J. and Krzyzosiak, W.J. (2011) The Role of Dicer Protein Partners in the Processing of MicroRNA Precursors. *PLoS One*, **6**, e28548.
48. Starega-Roslan, J., Krol, J., Koscianska, E., Kozlowski, P., Szlachcic, W.J., Sobczak, K. and Krzyzosiak, W.J. (2011) Structural basis of microRNA length variety. *Nucleic Acids Res*, **39**, 257-268.
49. Flores-Jasso, C.F., Arenas-Huertero, C., Reyes, J.L., Contreras-Cubas, C., Covarrubias, A. and Vaca, L. (2009) First step in pre-miRNAs processing by human Dicer. *Acta Pharmacol Sin*, **30**, 1177-1185.
50. Tahbaz, N., Kolb, F.A., Zhang, H., Jaronczyk, K., Filipowicz, W. and Hobman, T.C. (2004) Characterization of the interactions between mammalian PAZ PIWI domain proteins and Dicer. *EMBO Rep*, **5**, 189-194.
51. 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.

52. Maniataki, E. and Mourelatos, Z. (2005) A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev*, **19**, 2979-2990.
53. Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K. and Shiekhattar, R. (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, **436**, 740-744.
54. Gregory, R.I., Chendrimada, T.P., Cooch, N. and Shiekhattar, R. (2005) Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell*, **123**, 631-640.
55. Liu, X., Jin, D.Y., McManus, M.T. and Mourelatos, Z. (2012) Precursor MicroRNA-Programmed Silencing Complex Assembly Pathways in Mammals. *Mol Cell*, **46**, 507-517.

## FIGURE LEGENDS

**Figure 1.** Purified ALG-1 and ALG-2 Argonautes have the slicing activity.

**(A)** Schematic of the assay. **(B)** A representation of the RNA molecules used for the enzymatic assay. A 32 nucleotide (nt) 5' end radiolabeled ( $^{32}\text{P}$ ) RNA target (in black) and fully complementary (P) or bulged next to the cleavage site (M) 21nt ssRNA are shown in red. **(C)** The radiolabeled RNA target is incubated with either GST, GST-ALG-1 or GST-ALG-2 proteins with (+) the fully complementary 21 nt ssRNAs. Reaction performed with (+) or without (-) 10mM EDTA to sequester Magnesium ions essential for the enzymatic reaction are indicated. **(D)** Alteration of basepairing at the cleavage site (ssRNA M) abrogates target specific cleavage. The radiolabeled RNA target is incubated with either GST or GST-ALG-1 proteins. Hydroxylation of the RNA target (OH) is used as size marker (the 3' end nucleotides of the radiolabeled RNA fragment are indicated). **(E)** Point mutations of the DDH motif in the PIWI domain of GST-ALG-1 abrogate cleavage activity. Amounts (nM) of recombinant proteins used in the assay are indicated.

**Figure 2.** Recombinant ALG-1 can bind and cleave small RNA duplexes.

**(A)** Representation of the small RNA duplexes used for the assay. Strands complementary to the target RNA are shown in red. **(B)** Determination of the affinity of ALG-1 for different types of small RNAs. Filter binding assays were performed with increasing amounts of GST-ALG-1 AAA (nM) incubated with different 5' radiolabeled ( $^{32}\text{P}$ ) 21nt RNAs illustrated in A (ssRNA is shown in Figure 1). The error bars represent a 95% confidence interval. **(C)** Schematic of the slicing assay. **(D)** ALG-1 is able to cleave a perfectly paired small RNA duplex. Increasing amounts of GST-ALG-1 (nM) are incubated with different 5' radiolabeled ( $^{32}\text{P}$ ) small RNA duplexes (as shown in A). Hydroxylation of

the 21 nt RNA (OH) is used as size marker (the 3' end nucleotides of the radiolabeled RNA fragment are indicated). 8G and 7C represent other cleavage products.

**Figure 3.** *ALG-1 can use a miRNA-like duplex to induce target cleavage.*

**(A)** Schematic of the assay. **(B)** The radiolabeled ( $^{32}\text{P}$ ) 32 nucleotides (nt) RNA target was incubated with different concentration of GST-ALG-1 protein (nM) along with RNA duplexes containing zero (duplex), one (d1M) or two mismatches (d2M). **(C)** Comparative analysis of the cleavage induced with either single stranded RNA (ssRNA) or miRNA small RNA duplex (d2M). Cleavage product sizes are indicated as Figure 1.

**Figure 4.** *The presence of the slicing-competent ALG-1 or ALG-2 is essential for the production of let-7 and lin-4 miRNAs.*

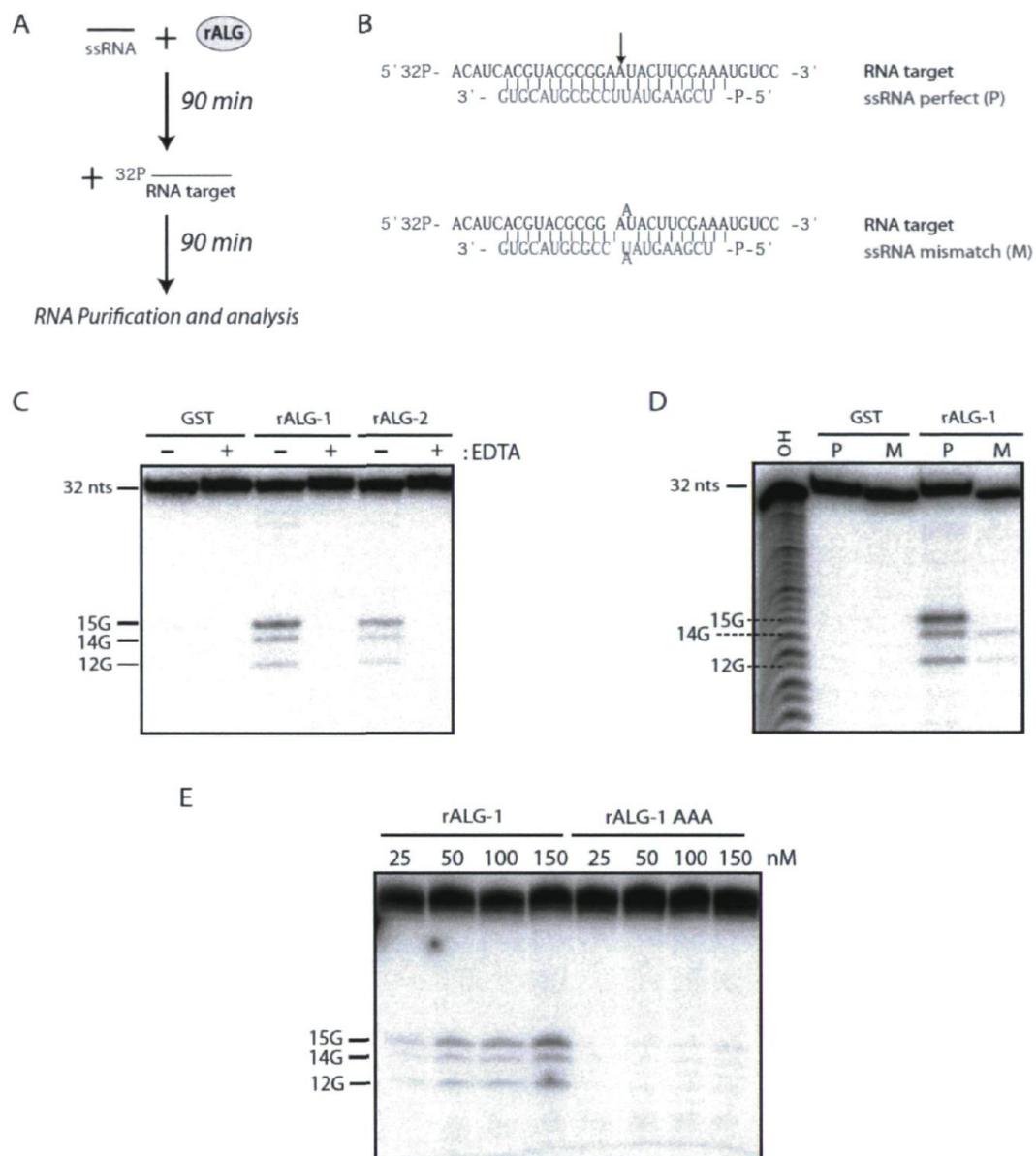
**(A-B)** Northern blot detection of miRNAs in animals expressing ALG-1 and ALG-2 transgenes. Upon total RNA extraction from different transgenic animals carrying either wild-type (wt) or catalytically defective (AAA) *alg-1* and *alg-2* genes, *lin-4* and *let-7* were detected using probes complementary to mature miRNA (antisense) and to the complementary strand (sense). SL1 RNA was probed and used as loading control (the detection of SL1 is only shown for one blot because the same membrane was stripped and re-used to detect RNA species from both miRNA). Precursor (pre), truncated precursor (pre short) and mature miRNA forms are indicated. The genotype and RNAi (below) as well as transgene (above) found in each animal strain are indicated. **(C)** Representation of the RNA molecules detected by Northern blotting.

**Figure 5.** *Slicing ALG-1 is required to generate functional miRISC.*

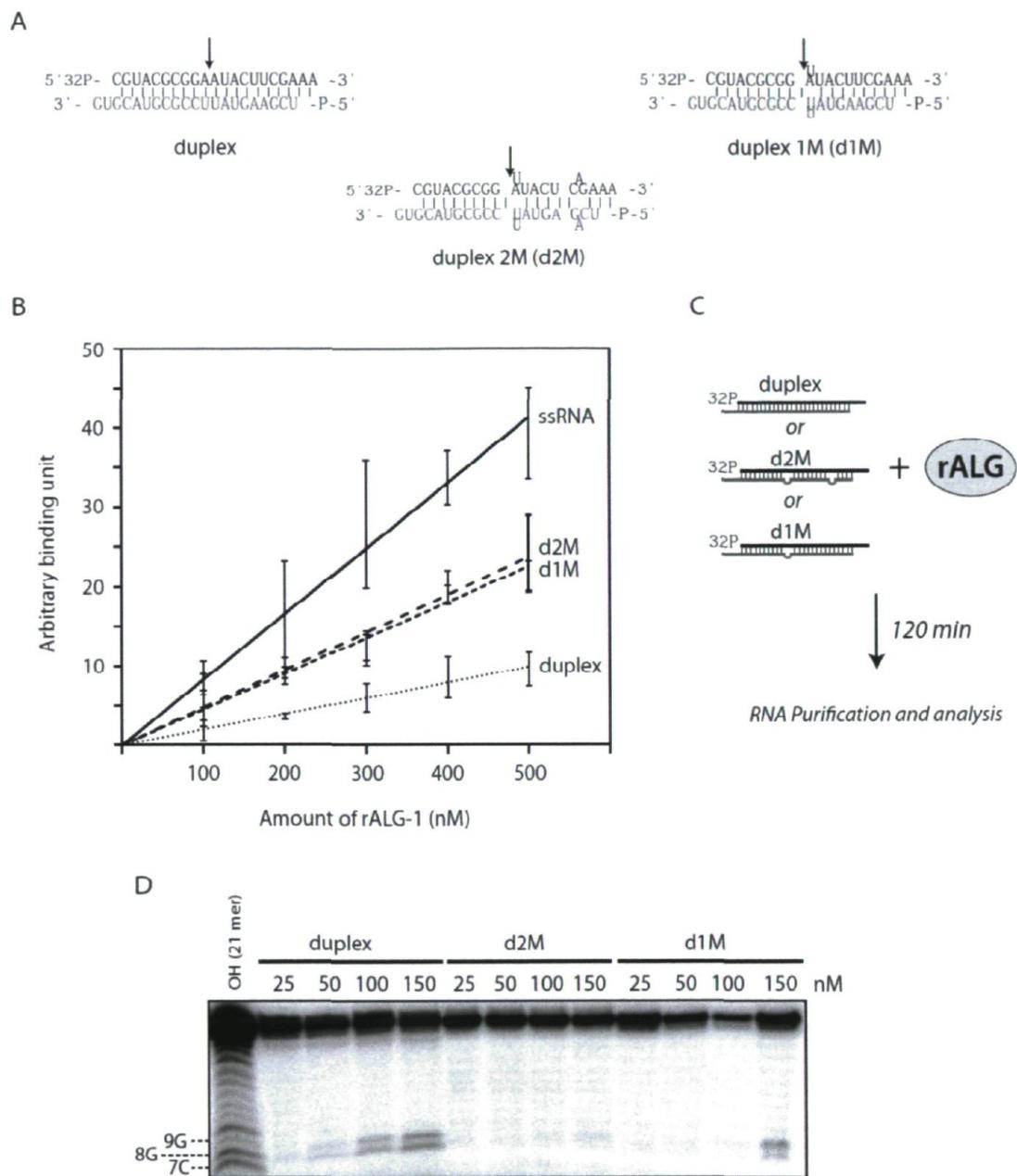
**(A)** Detection of *lin-4* in bound (IP) or unbound (FT) fractions after ALG-1 immunoprecipitations. Transgenically expressed ALG-1 wild-type (wt) or catalytically defective (AAA) protein were immunoprecipitated (IPed) from *alg-1(lf); alg-2(RNAi)* animals. The presence of *lin-4* strands was detected by Northern blotting (left panel) and 10% of immunopurified complexes were used for detection of ALG-1 (right panel). **(B)** The quantification of three independent experiments is shown in the lower panel. Error bars represent a 95% confidence interval and a Student's two-sided *t*-test was applied to obtain *P*-values.

**Table 1.** *Phenotypic analyses of ALG-1 and ALG-2 slicing defective animals.*

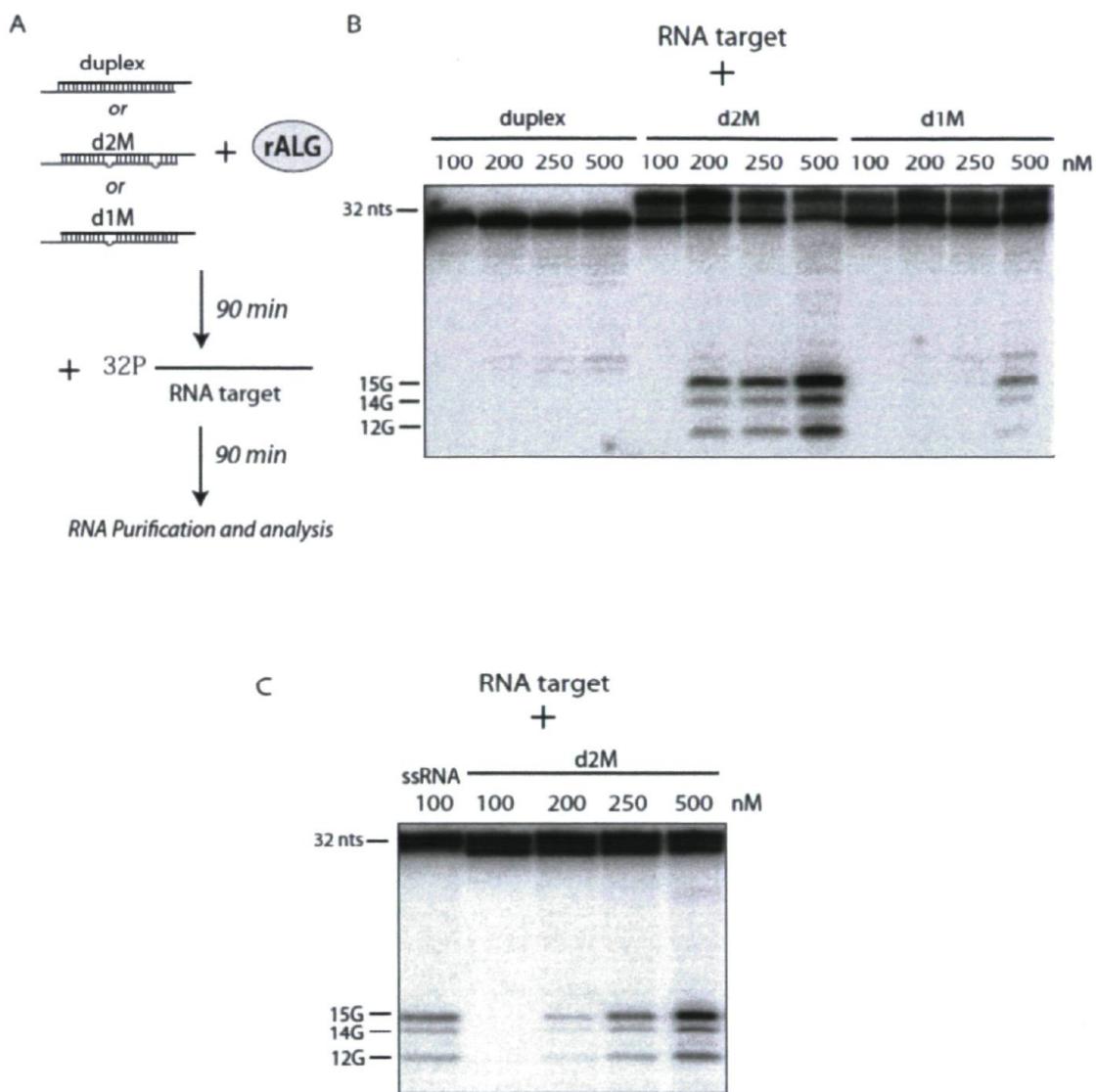
Transgenes expressing either wild-type or catalytically inactive ALG-1 and ALG-2 proteins were introduced in animals carrying loss-of-function (*lf*) alleles of *alg-1* or *alg-2* genes as indicated. Transgenic animals were then exposed to dsRNA targeting the other miRNA-specific Argonaute and phenotypes that are similar to the loss of the miRNA pathway were scored as percentage of *n* number of animals. Control RNAi corresponds to animals fed with bacteria containing the empty dsRNA-expressing plasmid. The lethality observed in animals is a consequence of a variety of severe developmental defects (some animals die during larval stages while others die from vulval bursting when they reach adulthood).



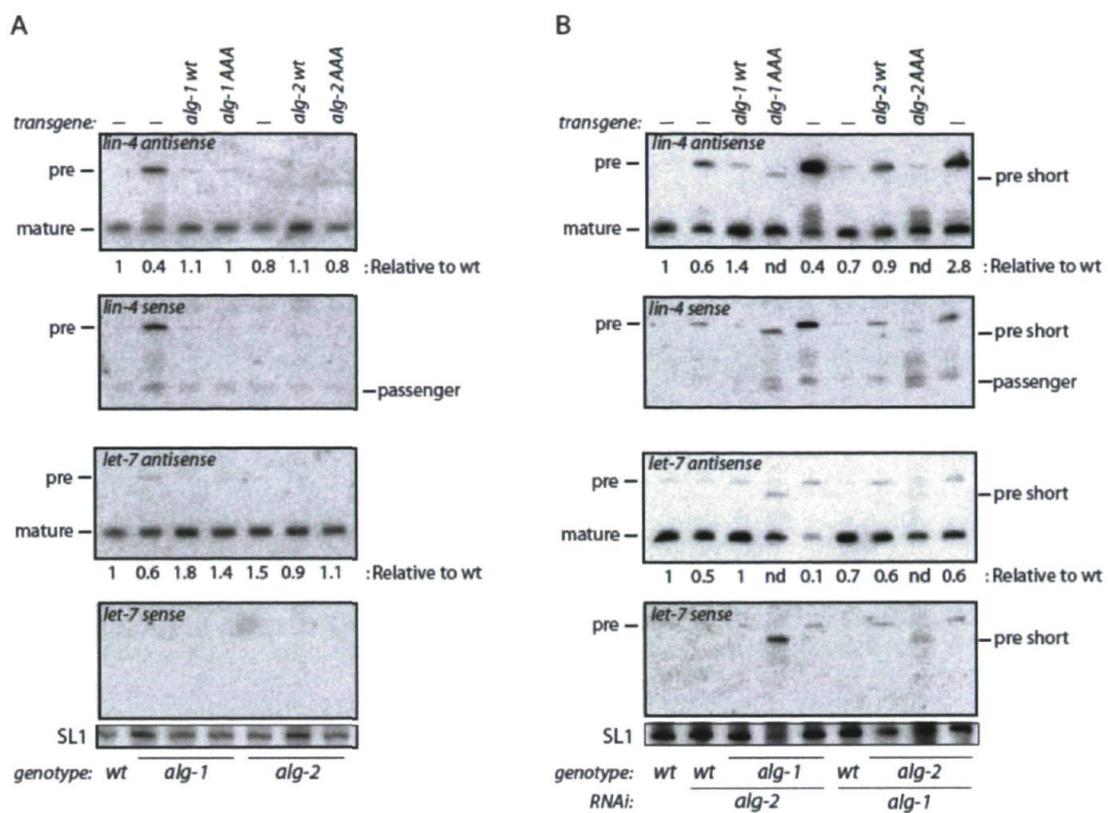
Bouasker, Fig. 1



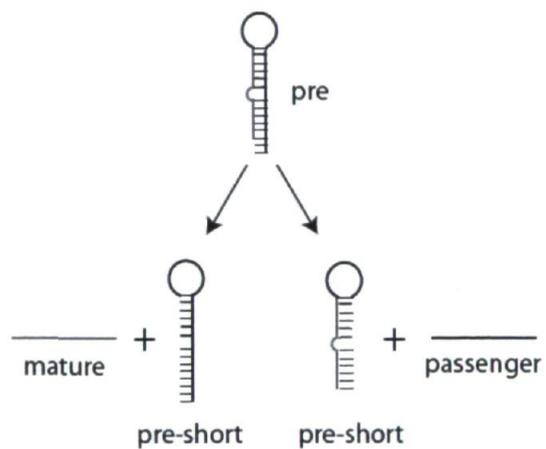
Bouasker, Fig. 2



Bouasker, Fig. 3



C



*types of miRNA molecules detected*

Bouasker, Fig. 4

Table 1: Effect of catalytically ineffective ALG-1 and ALG-2 on animals

| genotype             | Transgene        | RNAi         | Percentage of animal |                            |                            | (n) |
|----------------------|------------------|--------------|----------------------|----------------------------|----------------------------|-----|
|                      |                  |              | Fertile adults       | Sterility<br>(less severe) | Lethality<br>(most severe) |     |
| <i>alg-1 (gk214)</i> |                  | control      | 86                   | 0                          | 14                         | 125 |
| <i>alg-1 (gk214)</i> | <i>alg-1 wt</i>  | control      | 100                  | 0                          | 0                          | 122 |
| <i>alg-1 (gk214)</i> | <i>alg-1 AAA</i> | control      | 90                   | 0                          | 10                         | 157 |
| <i>alg-1 (gk214)</i> |                  | <i>alg-2</i> | 0                    | 0                          | 100                        | 151 |
| <i>alg-1 (gk214)</i> | <i>alg-1 wt</i>  | <i>alg-2</i> | 86                   | 0                          | 14                         | 235 |
| <i>alg-1 (gk214)</i> | <i>alg-1 AAA</i> | <i>alg-2</i> | 0                    | 27                         | 73                         | 207 |
| <i>alg-2 (ok304)</i> |                  | control      | 100                  | 0                          | 0                          | 130 |
| <i>alg-2 (ok304)</i> | <i>alg-2 wt</i>  | control      | 93                   | 0                          | 7                          | 137 |
| <i>alg-2 (ok304)</i> | <i>alg-2 AAA</i> | control      | 86                   | 0                          | 14                         | 111 |
| <i>alg-2 (ok304)</i> |                  | <i>alg-1</i> | 0                    | 0                          | 100                        | 153 |
| <i>alg-2 (ok304)</i> | <i>alg-2 wt</i>  | <i>alg-1</i> | 28                   | 22                         | 50                         | 176 |
| <i>alg-2 (ok304)</i> | <i>alg-2 AAA</i> | <i>alg-1</i> | 0                    | 28                         | 72                         | 200 |

**The slicing activity of microRNA-specific Argonautes is essential for the microRNA pathway in *C. elegans***

Samir Bouasker and Martin J. Simard

SUPPLEMENTARY DATA

SUPPLEMENTARY MATERIAL & METHODS

*Plasmids constructions*

cDNAs from *alg-1* and *alg-2* genes were cloned into plasmid pGEX-6P-1. The catalytic mutant (DDH to AAA) for both GST-expressing vectors and genomic *alg-1* and *alg-2* expressing transgenes were generated with Site directed mutagenesis kit (Stratagene) following manufacturer's instruction.

*C. elegans methods*

*C. elegans* strains were grown under standard conditions (44). Transgenic lines carrying *alg-1 wt*, *alg-1 AAA*, *gfp::alg-2 wt* or *GFP::alg-2 AAA* were produced using standard procedure (45). Transgenes carrying *alg-1 wt* and *alg-1 AAA* genes were introduced in *alg-1(gk214)* mutant and transgenes expressing tagged *alg-2 wt* and *AAA* genes introduced in *alg-2(ok304)* animals.

To generate RNAi knockdown population, synchronized L1 *C. elegans* population were exposed to bacteria expressing either control, *alg-1* or *alg-2* dsRNA molecules for 4 days at 15°C.

*Probes for Northern blot analysis*

Let-7 : 5'-AACTATACAACCTACTACCTCA/Starfire/-3'

Let-7 \* : 5'-GTAAGGTAGAAAATTGCATAGTTCA/Starfire/-3'

Lin-4 : 5'-TCACACTTGAGGTCTCAGGGA/Starfire/-3'

Lin-4 \* : 5'-GTACCCGGAGAGCCCAGGTGT/Starfire/-3'

miR-50 : 5'-AACCCAAGAATACCAGACATATCA/Starfire/-3'

miR-50 \* : 5'-GTCGATACGTCTAATATGCGGG/Starfire/-3'

miR-40 : 5'-TTAGCTGATGTACACCCGGTGA/Starfire/-3'

miR-40 \* : 5'-TATCATCATGGCATAACATCCACT/Starfire/-3'

miR-90 : 5'-AGGGGCATTCAAACAACATATCA/Starfire/-3'

miR-90\* : 5'-GTTGATATCGTCGTTGAAAGCCG/Starfire/-3'

*Cloning of the truncated pre-miR-50*

Small RNAs were size selected with miRVana Kit according to manufacturer protocol (Life Technologies). Upon purification, small RNAs were subjected either to 3' ligation (3'RACE) or 5' ligation (5'RACE) using a phosphorylated or unphosphorylated RNA adaptors, respectively.

Both ligated products, were reverse transcribed using either an oligonucleotide complementary to the RNA adaptor (3'RACE) or an oligonucleotide complementary to the 3' end of the miR-50 precursor (5'RACE). The generated cDNAs were subjected to two rounds of PCR and amplified products were size selected, isolated and cloned into TOPO vector (Life Technologies) for sequencing.

List of oligonucleotides used:

3'RACE :

RNA adaptor: 5'-CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUA-3'OH

RT: Primer RNA A: 5'-TCTCCTTCGTCCTGTCGTGTC-3'

PCR-1: primer RNA-A and primer 50: TGATATGTCTGGTATTCTTGGGTT

PCR-2: primer RNA-B GGACACTGACATGGACTGAAGGAGTA and primer 50

5'RACE:

RNA adaptor: 5'-CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUA-3'OH

RT: Primer 50 \*: 5'-GTCGATACGTCTAATATGCGGG-3'

PCR-1: primer RNA-A and primer 50 \*

PCR-2: primer RNA-B and primer 50 \*

#### SUPPLEMENTARY REFERENCES

44. Brenner, S. (1974) The genetics of *Caenorhabditis elegans*. *Genetics*, **77**, 71-94.
45. Mello, C. and Fire, A. (1995) DNA transformation. *Methods Cell Biol*, **48**, 451-482.

#### SUPPLEMENTARY FIGURE LEGENDS

**Figure S1.** (A) Production of recombinant ALG-1 and ALG-2 proteins. Purified GST-ALG-1 (138kDa) and GST-ALG-2 (128kDa) and variants (AAA) with substitutions of conserved aspartic acids (D) and the histidine (H) residues known to be essential for slicing activity to alanine residues were loaded onto 8% SDS-PAGE stained with PhastGel Blue R-350. (B) Filter binding assay to compare binding of wild-type and AAA variants of ALG-1

and ALG-2 proteins. **(C)** Comparative analysis of the cleavage activity of recombinant ALG-1 and ALG-2 proteins. **(D)** Point mutations of the DDH motif contains in the PIWI domain of GST-ALG-2 abrogate cleavage activity. Amounts (nM) of recombinant proteins used in the assay are indicated.

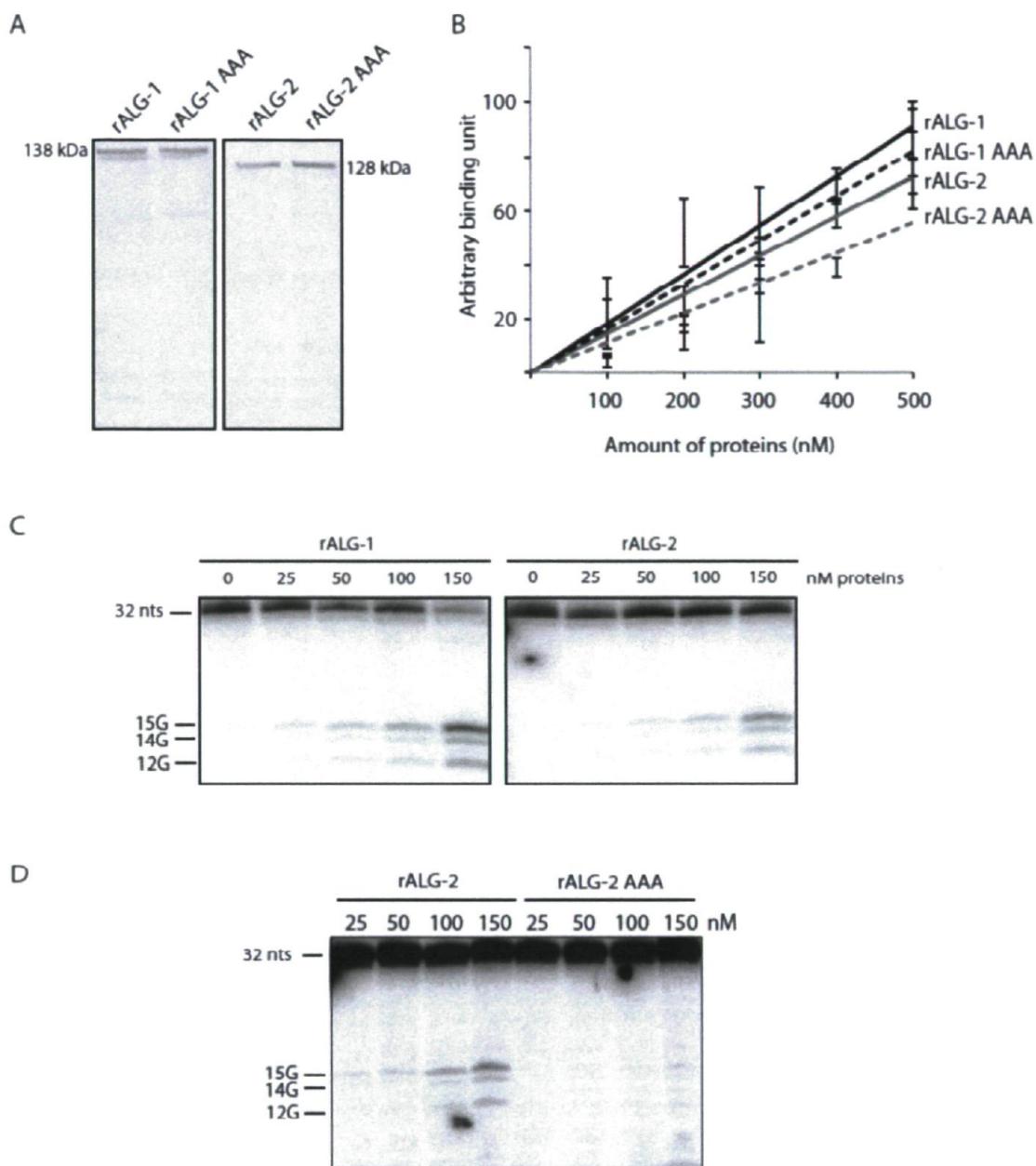
**Figure S2.** *Recombinant ALG-2 is able to bind and cleave small RNA duplexes.* **(A)** Determination of the affinity of ALG-1 for different types of small RNAs. Filter binding assays were performed with increasing amounts of GST-ALG-2 AAA (nM) incubated with different 5' radiolabeled (32P) 21nt long RNAs illustrated in Figure 2A (ssRNA is shown in Figure 1). The error bars represent the 95% confidence interval. **(B)** ALG-2 is able to cleave perfectly paired small RNA duplex. Increasing amounts of GST-ALG-2 (nM) are incubated with different 5' radiolabeled (32P) small RNA duplexes (as shown in Figure 2A). Hydroxylation of the 32 nts long RNA (OH) is used as size marker (the 3' end nucleotides of the radiolabeled RNA fragment are indicated). 8G and 7C represent other cleavage products.

**Figure S3.** *ALG-2 can use miRNA duplex to cleave RNA targets.* The radiolabeled (32P) 32 nucleotides (nt) long RNA target is incubated with different concentration of GST-ALG-2 protein (nM) along with either a single stranded RNA (ssRNA) or miRNA small RNA duplex (d2M). Cleavage product sizes are indicated as Figure 1.

**Figure S5.** **(A)** Determination of the size of miRNA molecules. Total RNA extraction from transgenic animals were loaded onto a 15% denaturing acrylamide gel and transferred for

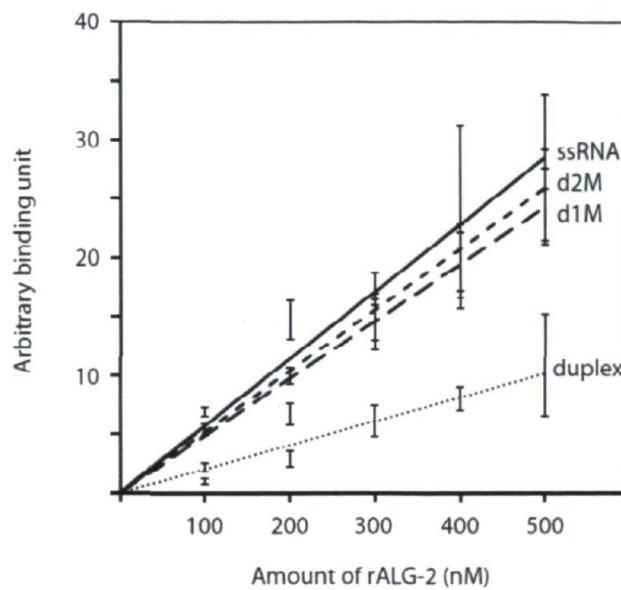
detection of *miR-50* miRNA molecules by Northern blotting using a probes complementary to mature miRNA. Radiolabeled 100 nt RNA ladder (Ambion) was used as size marker. Precursor (pre), truncated precursor (pre short) and mature miRNA forms are indicated. Genotype (below) and transgene (up) found in each animal strain are indicated. **(B)** Summary of the sequencing data for the truncated *miR-50* pre-miRNA shown in Figure 4B. 12 out of 22 sequenced clones generated by 5' RACE and 12 out 24 from the 3' RACE correspond to the exact predicted Dicer cleavage site that generate mature (red) and passenger (blue) strand, respectively. The remaining clones sequenced for each RACE experiment were products few nucleotides shorter, that most likely indicate degradation products.

**Figure S6.** Slicing ALG-2 is essential to generate functional miRISC. Detection of *lin-4* in bound (IP) or unbound (FT) fractions after ALG-2 immunoprecipitations. Transgenically expressed GFP::ALG-2 wild-type (wt) or catalytically defective (AAA) protein were immunoprecipitated from *alg-1(lf); alg-2(RNAi)* animals. The presence of *lin-4* strands was detected by Northern blotting (left panel) and 10% of immunopurified complexes were used for detection of GFP::ALG-2 (right panel). The quantification of three independent experiments is shown in the lower panel. Errors bars represent the 95% confidence interval and a Student's two-sided t-test was applied to obtain P.

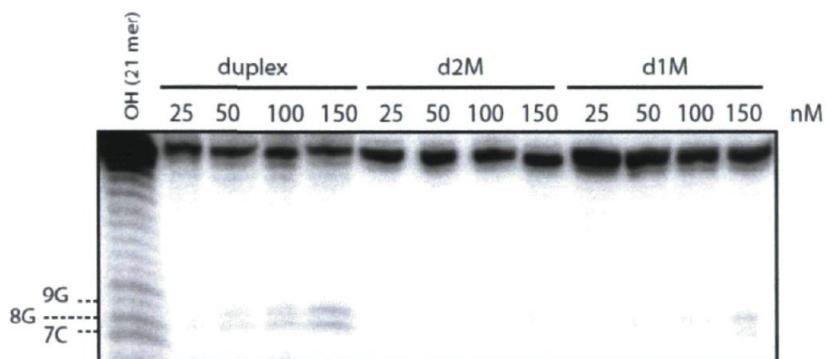


Bouasker, Figure S1

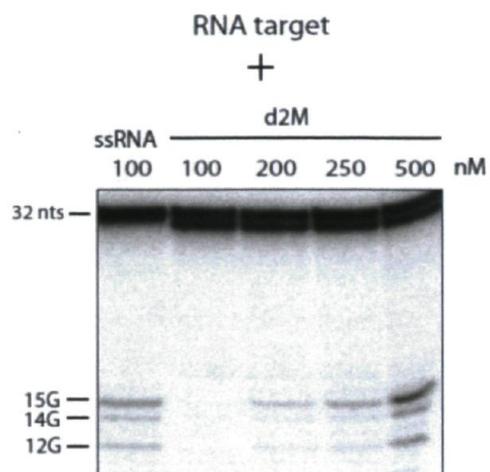
A



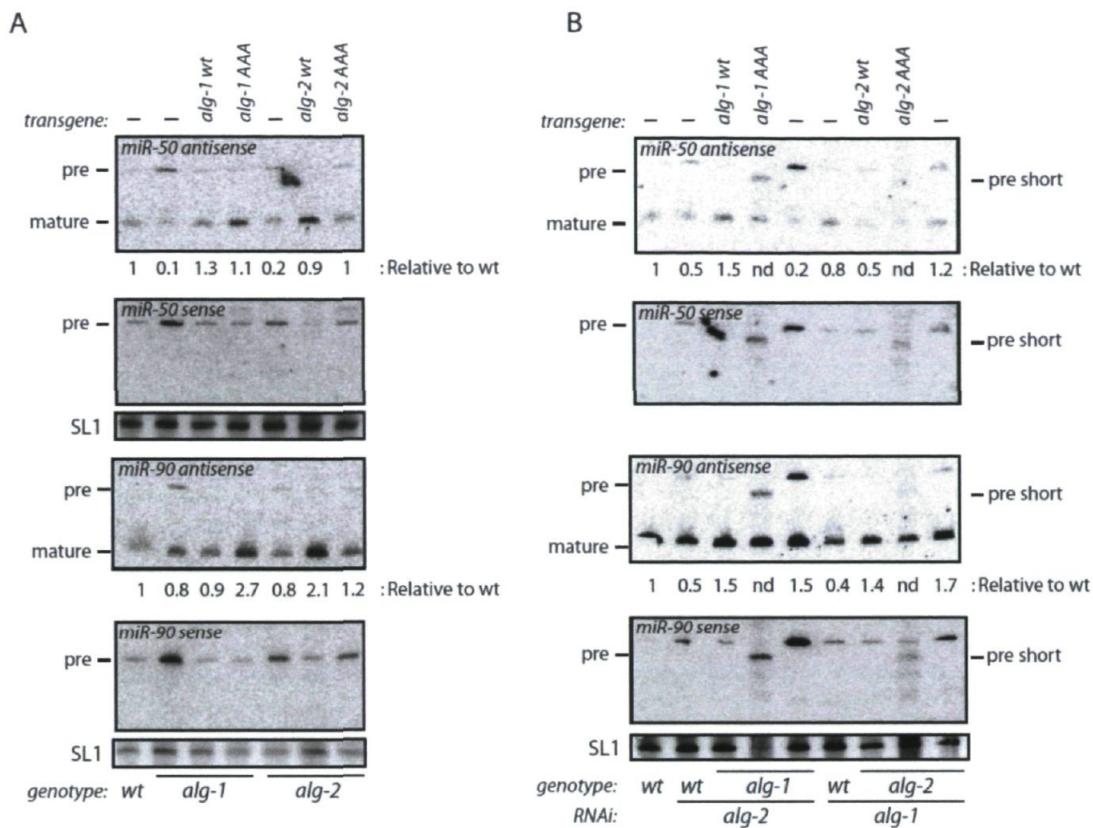
B



Bouasker, Figure S2

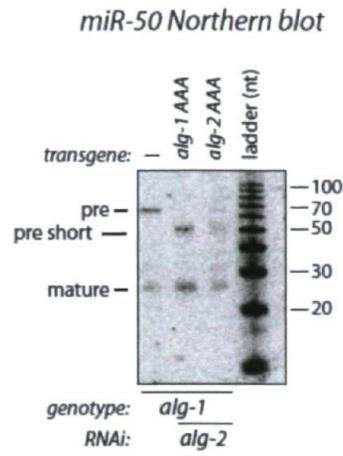


Bouasker, Figure S3

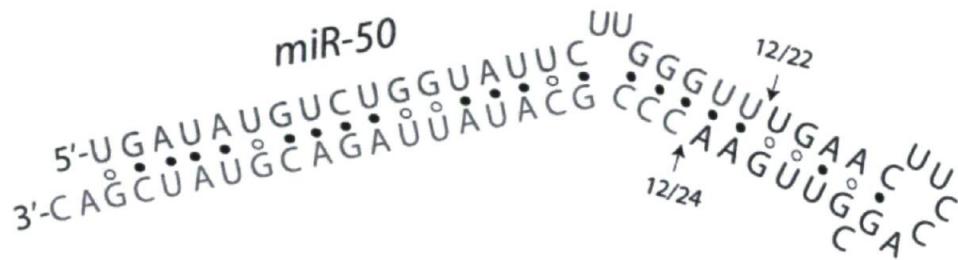


Bouasker, Figure S4

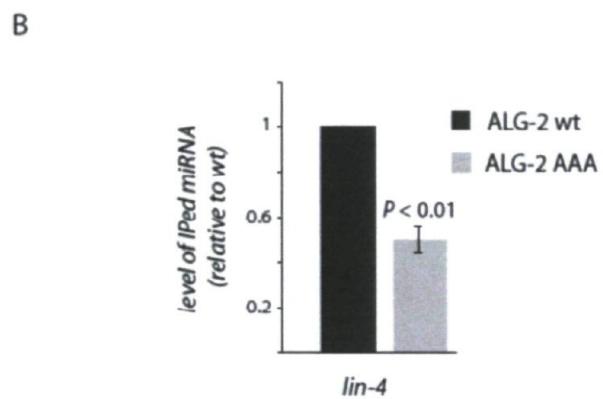
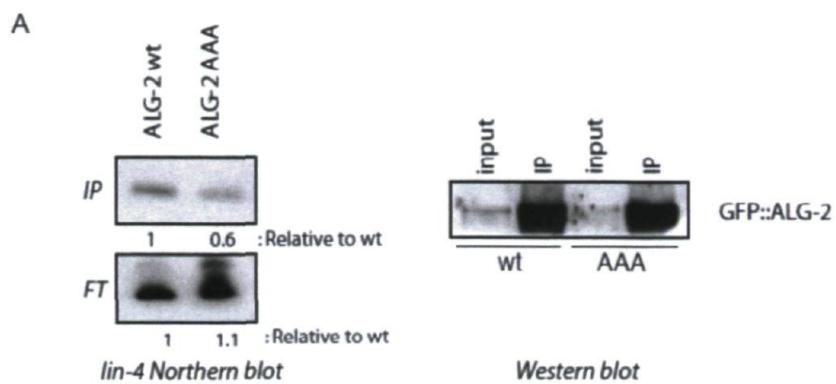
A



B



Bouasker, Figure S5



Bouasker, Figure S6