RNA Processing Activities of the *Arabidopsis* Argonaute Protein Family

Ki Wook Kim, Andrew L Eamens and Peter M Waterhouse

*University of Sydney, School of Molecular Biosciences*

Australia

1. Introduction

Stringent regulation of gene expression is essential for all organisms, and eukaryotes rely on diverse RNA silencing mechanisms for this regulation at both the transcriptional and post-transcriptional level. For example, transcriptional gene silencing (TGS) maintains genome integrity by controlling the replication of transposons and other repetitive DNA elements, as well as preserving chromatin states and epigenetic imprinting. Post-transcriptional gene silencing (PTGS) mechanisms on the other hand control the expression of messenger RNA (mRNA) transcripts of protein-coding genes in order to regulate developmental transitions and responses to environmental stresses. In plants, both transcriptional and post-transcriptional RNA silencing mechanisms are also involved in the defence against invading pathogens, especially viruses.

RNA silencing pathways are directed by a specific class of small RNA (sRNA) which are predominantly 20 to 25 nucleotides (nt) in length. These sRNAs are processed from longer precursor molecules of either perfectly or imperfectly double-stranded RNA (dsRNA) by a member of the DICER RNase III-like endonuclease family (Bernstein et al., 2001; Gregory et al., 2005). Once processed from its dsRNA substrate, the sRNA is subsequently modified and loaded into an RNaseH-like Argonaute (AGO) protein to form the catalytic core of an RNA-induced silencing complex (RISC). RISC uses the loaded sRNA as a sequence-specificity guide to direct RNA silencing of the targeted sequence at either the transcriptional or post-transcriptional level depending on; i) the class of sRNA loaded by AGO, and; ii) the AGO protein family member loaded with the sRNA.

Although the effector function mediated by the sRNA-loaded AGO protein is highly conserved amongst eukaryotes, the number of AGO proteins encoded by different species varies widely (Tolia & Joshua-Tor, 2007; Hutvagner & Simard, 2008; Vaucheret, 2008). For example, *Caenorhabditis elegans* (*C. elegans*) has twenty seven AGOs, *Drosophila melanogaster* (*Drosophila*) five, humans four and the yeast *Schizosaccharomyces pombe* one. In *Arabidopsis thaliana* (*Arabidopsis*), the model dicotyledonous plant species, the AGO protein family consists of ten members, that mediate the parallel RNA silencing pathways of Arabidopsis, and which are directed by numerous classes of endogenous sRNA, including the microRNA (miRNA; Lee & Ambros, 2001), small-interfering RNA (siRNA; Hamilton & Baulcombe, 1999), repeat-associated small-interfering RNA (rasiRNA; Meister & Tuschl, 2004), trans-acting small-interfering RNA (tasiRNA; Adenot et al., 2006; Xie et al., 2005) and natural antisense transcript small-interfering RNA (natsiRNA; Borsani et al., 2005) classes of sRNA.
In this chapter we discuss the RNA silencing-related activities of the ten Arabidopsis AGO protein family members, and by comparison with the known functions of characterised animal AGOs, as well as in other plant species, we suggest where future insights may be made.

2. AGO protein domain structure

The crystal structure of plant AGO proteins remains to be determined, therefore most of our current knowledge is based on studies of AGO purified from the bacteria *Thermus thermophilus* (Fig. 1). These analyses have revealed that AGOs are large proteins (ca 90-100 kDa) comprised of a single variable N-terminal domain and three conserved C-terminal domains, including the PAZ, MID and PIWI domains (Vaucheret, 2008). The N-terminal domain is thought to facilitate the separation of the sRNA/target transcript duplex post cleavage. The conserved PAZ and MID domains of the C-terminus recognize and anchor the 3’ and 5’ ends of the bound sRNA to its target mRNA respectively (Wang et al., 2008, Wang et al., 2009; Parker, 2010). The third C-terminal domain, the PIWI domain, specifies the endonuclease or “Slicer” activity of cleavage-competent AGOs. This domain adopts a folded structure that closely resembles the catalytic domain of the *Bacillus holodurans* RNaseH enzyme and which usually carries an Asp-Asp-His (DDH) motif in its active site (Rivas et al., 2005). Mutagenesis studies have demonstrated that altering the amino acid composition of this motif abolishes the Slicer activity of several target transcript-cleaving AGOs (Liu et al., 2004; Rivas et al., 2005). However, presence of the DDH motif does not guarantee cleavage activity. For example, the DDH Slicer motif is present in human AGO2 and AGO3, however only AGO2 is capable of catalysing sRNA-directed target transcript cleavage (Liu et al., 2004; Meister et al., 2004). Conversely, absence of the DDH Slicer motif does not preclude the AGO from cleavage-based RNA silencing. This is demonstrated by the

![Domains and crystal structure of AGO proteins.](image-url)
Drosophila AGO, DmPIWI which instead encodes an Asp-Asp-Lys (DDK) Slicer motif and is cleavage-competent (Saito et al., 2006). In addition to mediating the Slicer activity of cleavage-competent AGOs, the PIWI domain has been demonstrated to serve as an interface for protein-protein interactions, namely interaction with glycine-tryptophan (GW) repeat proteins. In animals for example, the interaction of AGOs with GW182 has been shown to be essential for translational repression (Liu et al., 2005; Eulalio et al., 2008). To date, no GW182 orthologues have been identified in plants. However, in the RNA-directed DNA methylation (RdDM) RNA silencing pathway of Arabidopsis, AGO4 has been shown to interact with the GW protein NRPD1 to recruit the DNA methylation machinery required for the effector step of this pathway (El-shami et al., 2007). Whether other Arabidopsis AGO family members also interact with GW proteins remains to be determined.

3. The Arabidopsis Argonaute protein family

AGO1 is the founding member of the Arabidopsis AGO protein family. The ago1 mutant plant was originally identified in a forward genetics screen, exhibiting pleiotropic developmental defects, characterized by tubular shaped leaves that were thought to closely resemble the tentacles of a small squid of the Argonauta genus (Bohmert et al., 1998). Subsequent phenotypic and/or molecular studies of plant lines defective for the activity of the miRNA biogenesis machinery proteins SERRATE (se), DICER-LIKE1 (dcl1) and dsRNA BINDING PROTEIN (drb1), or of transformed plant lines expressing miRNA resistant targets implicated the involvement of AGO1 in miRNA biogenesis (Lobbes et al., Park et al., 2002; Vaucheret et al., 2004). Alleles of ago1 were also isolated in genetic screens identifying plant lines where the expression of a post-transcriptionally silenced sense transgene (termed S-PTGS in plants) was reactivated (Fagard et al., 2000). Taken together, these studies identified AGO1 as playing an integral, central role in the parallel sRNA-directed RNA silencing pathways of Arabidopsis.

The identification of the Arabidopsis AGO1 protein prompted extensive searches for orthologues in other model organisms. Indeed, AGO proteins were found to be highly conserved across plant and animal kingdoms (Catalanotto et al., 2000; Fagard et al., 2000). In Arabidopsis, the complete annotation of its genome identified nine other AGO family members (termed AGO2 to AGO10 respectively), and phylogenetic analysis of this protein family at the amino acid level identified three distinct clades, namely the AGO1/AGO5/AGO10, AGO2/AGO3/AGO7, and AGO4/AGO6/AGO8/AGO9 clades (Vaucheret, 2008; Fig.2). It is important to note that although the distribution of the 10 Arabidopsis AGO proteins into three distinct clades is purely based on amino acid sequence homology, and does not directly infer similarities in activity or redundancies in function, several examples of functional redundancy have been identified between AGO clade members, namely between AGO1 and AGO10 of the AGO1/5/10 clade (Mallory et al., 2009), and AGO4, AGO6 and AGO9 of the AGO4/6/8/9 clade (Havecker et al., 2010).

In addition to the identification of functional redundancy amongst family members, some Arabidopsis AGO proteins have also been shown to exhibit strong preferences to load sRNA species of a particular size and/or 5' terminal nucleotide (Havecker et al., 2010; Mi et al., 2008; Takeda et al., 2008). For example, AGO1 binds sRNAs that are predominantly of the 21-nt size class with have a 5' terminal uracil, whereas AGO2 preferentially binds 21-nt sRNAs, with a 5' terminal adenine. Unlike AGOs 1 and 2, AGO5 predominantly binds sRNAs of the 24-nt size class and with a cytosine at their 5' terminal residue. AGOs 4, 6 and
Fig. 2. Phylogenetic tree constructed from the full-length amino acid sequences of the ten Arabidopsis AGO family members. The phylogeny clade divisions are marked in different colours. Drawing not to scale.

9 also preferentially associate with 24-nt sRNAs but prefer to bind sRNAs of this size class with 5’ adenine residues. The 5’ terminal nucleotide preference for AGOs 3, 7, 8 and 10 remain to be determined. All known 5’ terminal nucleotide and size class preferences are summarized in Table 1.

<table>
<thead>
<tr>
<th>Family member</th>
<th>Clade</th>
<th>Protein size (aa)</th>
<th>Molecular Weight</th>
<th>Protein Accession No.</th>
<th>5’ terminal nucleotide preference</th>
<th>sRNA length preference (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGO1</td>
<td>1/5/10</td>
<td>1048</td>
<td>116190.3</td>
<td>NP_175274</td>
<td>U</td>
<td>21</td>
</tr>
<tr>
<td>AGO2</td>
<td>2/3/7</td>
<td>1014</td>
<td>113422.7</td>
<td>NP_174413</td>
<td>A</td>
<td>21</td>
</tr>
<tr>
<td>AGO3</td>
<td>2/3/7</td>
<td>1194</td>
<td>129184.1</td>
<td>NP_174414</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>AGO4</td>
<td>4/6/8/9</td>
<td>924</td>
<td>102839.7</td>
<td>NP_565633</td>
<td>A</td>
<td>24</td>
</tr>
<tr>
<td>AGO5</td>
<td>1/5/10</td>
<td>997</td>
<td>111087.8</td>
<td>NP_850110</td>
<td>C</td>
<td>24</td>
</tr>
<tr>
<td>AGO6</td>
<td>4/6/8/9</td>
<td>878</td>
<td>98680.7</td>
<td>NP_180853</td>
<td>A</td>
<td>24</td>
</tr>
<tr>
<td>AGO7</td>
<td>2/3/7</td>
<td>990</td>
<td>113395.8</td>
<td>NP_177103</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>AGO8</td>
<td>4/6/8/9</td>
<td>850</td>
<td>95506.0</td>
<td>NP_197602</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>AGO9</td>
<td>4/6/8/9</td>
<td>896</td>
<td>100523.6</td>
<td>NP_197613</td>
<td>A</td>
<td>24</td>
</tr>
<tr>
<td>AGO10</td>
<td>1/5/10</td>
<td>988</td>
<td>110866.7</td>
<td>NP_199194</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

Table 1. Summary of the 5’ terminal nucleotide and size (nt) preferences of Arabidopsis AGOs determined through immunoprecipitation experiments. The protein size, clade member, molecular weight, 5’ terminal nucleotide and Genbank protein accession number also listed.
3.1 The AGO1/AGO5/AGO10 clade

3.1.1 AGO1

AGO1 has been shown to direct sRNA-mediated gene expression regulation for all currently characterized Arabidopsis miRNAs (Baumberger & Baulcombe, 2005; Vaucheret et al., 2004). Most ago1 mutants exhibit pleiotropic developmental defects characteristic of perturbed miRNA function. In these plants, the miRNA levels are reduced, and their target transcript expression levels increased. In addition, the majority of plant miRNAs have a 5’ terminal uracil residue and are preferentially loaded by AGO1 (Mi et al, 2008). The high level of miRNA/target transcript sequence complementarity in plants results in AGO1 repressing target gene expression via miRNA-mediated target transcript cleavage. A recent study has suggested that AGO1 can also repress target gene expression via translational repression (Brodersen et al., 2008), however, it remains unclear whether this is a widespread RNA silencing mechanism in plants. To maintain steady-state expression of AGO1, the expression of the Ago1 transcript is itself regulated by a miRNA. Regulation of Ago1 by miR168 ensures that AGO1 levels remain constant, in turn ensuring normal plant development. AGO1 homeostasis is indeed crucial for normal plant development as the expression of a miRNA-resistant AGO1 transgene caused severe developmental defects that led to the eventual death of the plant (Vaucheret et al., 2004, 2006).

In addition to its role in the miRNA biogenesis pathway, AGO1 also performs a dual function in the biogenesis of the closely related endogenous sRNA class, the tasiRNAs. AGO1 uses loaded miRNAs, namely miR173 and miR828 (Allen et al., 2005; Rajagopalan, et al., 2006; Yoshikawa et al., 2005) to target the non-protein-coding transcripts Tas1, Tas2 and Tas4 for miRNA-mediated cleavage respectively. This initial cleavage event identifies these cleavage products for dsRNA synthesis by the RNA-directed RNA polymerase RDR6 (Peragine et al., 2004). Following dsRNA synthesis and processing of these molecules, the resulting tasiRNAs are loaded by AGO1-catalyzed RISC for sRNA-mediated target transcript cleavage (Yoshikawa, et al. 2005). More recently, AGO1 has also been shown to be involved in the generation of ‘secondary’ or ‘transitive’ siRNAs from sRNA cleaved transcripts (Chen et al., 2010; Cuperus et al., 2010). Furthermore, AGO1 also mediates the effector step for siRNA-directed RNA silencing. These siRNAs may be derived from either an infecting virus, or from introduced transgenes (including sense, antisense or hairpin RNA transgenes). AGO1 is the primary AGO family member involved in the antiviral response, and ago1 plants are hyper-susceptible to several viruses, including Cucumber mosaic virus (CMV) (Morel et al., 2002). CMV encodes the silencing suppressor protein (SSP) 2b, which directly impairs the function of AGO1 (Zhang et al., 2006). Members of the Polerovirus family have also been shown to encode SSPs which target the action of AGO1 and Arabidopsis ago1 plants have also been demonstrated to be hyper-susceptible to these viruses (Baumberger et al., 2007; Bortolamiol et al., 2007).

The central role played by AGO1 in sRNA-directed RNA silencing is mirrored by its expression domain. Array data reveals that Ago1 is ubiquitously expressed at high levels throughout development (Schmid et al., 2005). Experiments in Arabidopsis using the AGO1 promoter fused to a GUS reporter gene revealed that the promoter is active in all aerial tissues, but its activity appears to be highest in meristematic and vascular tissue (Vaucheret et al., 2006). In addition to being expressed ubiquitously, AGO1 seems to function in both the cytoplasm and nucleus of the plant cell. It appears to be strictly cytoplasmic when processing viral RNAs but Fang & Spector (2007) and Song et al., (2007) have reported that AGO1 is in the nucleus and is most concentrated around small nuclear bodies termed...
nuclear dicing bodies or D-bodies. The miRNA biogenesis machinery proteins DCL1, HYL1 and SE are also found in D-bodies, where miRNA precursor transcripts are processed prior to loading of the mature miRNA guide strand into AGO1.

3.1.2 AGO5
To date, no ago5 mutant alleles have been identified in any forward genetic screens. Furthermore, T-DNA knockout lines are wild-type in appearance and the role of this AGO family member in sRNA-directed RNA silencing remains to be determined. No changes in endogenous or exogenous sRNA classes were detected in an ago5 mutant (Takeda et al., 2008). In contrast to AGO1, the expression profile for AGO5 is highly specific to reproductive tissues (Schmid et al., 2005), accumulating in the sperm cell cytoplasm in mature pollen and growing pollen tubes (Borges et al., 2011).

Sequencing of the sRNAs bound to AGO5 revealed its preference for species 24-nts in length and with a 5’ terminal cytosine residue. However, AGO5 is also able to bind 21-nt sRNAs. MiR169 is one of a small number Arabidopsis miRNAs that do not have uracil as the 5’ terminal nucleotide, and this 21-nt miRNA preferentially associates with AGO5 rather than AGO1 (Mi et al., 2008; Takeda et al., 2008). The biological function of miR169 remains to be determined in Arabidopsis, but this highly conserved miRNA is important for floral development in petunia and anthirinum (Cartolano et al., 2007; Combier et al., 2006), to suggest that AGO5/miR169 may be involved in regulating gene expression in Arabidopsis.

Furthermore, AGO5 has been shown to bind both 21 and 24-nt viral siRNA size classes (Takeda et al., 2008). However, ago5 plants do not appear to be hyper-susceptible to plant viruses (Harvey et al., 2011; Wang et al., 2011), suggesting that this AGO family member may only play a minor or subservient role in viral defence under normal conditions (e.g. in the presence of AGO1 activity).

3.1.3 AGO10
The AGO10 mutant alleles, pinhead and zwille were identified through forward genetic screens (Lynn et al., 1999; Moussian et al., 1998). Both alleles are characterized by abnormal shoot apical meristem (SAM) development, but these mutants do not display any other readily observable developmental defects. Despite the high level of amino acid sequence similarity between AGO10 and AGO1, ago10 mutants are not impaired in S-PTGS and show no reduction in the accumulation miRNAs, tasiRNAs or any other siRNA class assessed in this mutant background (Morel et al., 2002; Takeda et al., 2008).

As the closest parologue of AGO1, it had been postulated that AGO10 may have similar activities and function redundantly with AGO1. Indeed, previous studies have shown that the ago1:ago10 double mutant is embryo lethal, strongly suggesting functional redundancy between AGO10 and AGO1 during post-embryonic development (Lynn et al., 1999). Fusion of the AGO1 promoter and coding sequence to a reporter gene revealed that AGO1 is expressed in whole embryos, with its expression highest in provascular cells from the globular to early torpedo stages. The expression of AGO10 partially overlaps the expressional domain of AGO1. Fusion of the AGO10 promoter and coding sequence to a reporter gene showed that its expression is more restricted in whole embryos than observed for AGO1, becoming limited to provascular strands and the adaxial side of cotyledons at the globular stage (Mallory et al., 2009). Moreover, the expression of the AGO10 coding sequence, fused to the AGO1 promoter, revealed that this family member can partially
compensate for AGO1 activity, to again suggest that AGO10 may be involved in sRNA-mediated gene expression regulation in specific cells and/or tissues. Consistent with the SAM defects observed in pinhead and zwille mutants, recent studies have demonstrated that AGO10 acts as a critical regulator of SAM maintenance by specifically interacting with miR165 and miR166 (Liu et al., 2009; Zhu et al., 2011). Members of both miRNA families regulate the expression of class III homeodomain-Leucine Zipper (HD-Zip III) transcription factors, which in turn determine the fate of the SAM (Jung & Park, 2007; Zhou et al., 2007). AGO10 exhibits a higher binding affinity for miR165/166 than AGO1, and when miR165/166 loading to AGO10 is perturbed, plants exhibit a defective SAM. Although the exact mechanism of how AGO10 regulates SAM development via miR165/166 regulation remains elusive, Zhou et al., 2011 demonstrated that the miRNA-binding activity of AGO10, and not its miRNA-directed Slicer activity is the important determinant of this interaction. The authors went on to suggest that AGO10 may in fact be specifically sequestering miR165/166 duplexes to prevent their incorporation into AGO1 and subsequent repression of the HD-Zip III transcription factors.

3.2 The AGO2/AGO3/AGO7 clade

3.2.1 AGO2/AGO3

AGO2 and AGO3 are thought to have arisen from a recent duplication event, as these two family members share a very high level of amino acid sequence similarity and are adjacent to one another in the Arabidopsis genome. Array data reveals that all three family members of the AGO2/3/7 clade have overlapping expression domains (Schmid et al., 2005). AGO2 and 3 are most highly expressed in developing seeds and siliques, and at lower levels in senescing leaves and flowers. Both family members also have dynamic cellular localization, being expressed in both the nucleus and cytoplasm (Takeda et al., 2008). Although, to date, no functional similarity or redundancy has been reported for AGO2 and 3, their high level of sequence similarity, proximal genomic positioning and shared expression patterns strongly suggests that they have the same or similar RNA silencing roles in Arabidopsis. However, no forward genetic mutants have been identified for either family member and T-DNA knockout mutants of AGO2 and AGO3 are wild-type in appearance (Lobbes et al., 2006). Furthermore, northern blotting has shown wild-type accumulation for all sRNA species assessed in ago2 and ago3 plants (Katiyar-Agarwal et al., 2007; Takeda et al., 2008). AGO2 is preferentially loaded with sRNA species, including viral sRNAs, possessing a 5’ terminal adenine residue (Mi et al., 2008; Takeda et al., 2008). More recent studies have implicated the involvement of AGO2 in antiviral defence, showing that ago2 mutants are hyper-susceptible to Turnip crinkle virus (TCV) and CMV infection, and that AGO2 expression is induced upon TCV and CMV infection of wild-type plants (Harvey et al., 2011; Wang et al., 2011). Furthermore, AGO2 has been shown to act downstream of the viral secondary siRNA biogenesis together with AGO1 in a non-redundant manner, essential for defence against CMV infection (Wang et al., 2011). Despite AGO2 playing such an important antiviral role in the defence against TCV and CMV, ago2 mutants are not hyper-susceptible to all plant viruses. For example, ago2 mutants show wild-type-like symptoms upon Tobacco mosaic virus (TMV) infection. Unlike TCV and CMV, TMV does not encode a SSP that impairs the function of AGO1. The induction of AGO2 upon TCV and CMV infection is therefore thought to result from a decreased accumulation of the AGO1-dependent, AGO2-regulating miRNA, miR403 (Harvey et al., 2011). This may also be true for Ago3 as its transcript is also regulated by miR403. This raises the question: is this a system that has
evolved to provide back-up protection against viruses that target AGO1 with their SSPs (e.g., 2b of CMV and P38 of TCV) or an unselected for, accidental consequence of reduced miR403 accumulation? Whereas the former seems likely, it is interesting that rice orthologs of Arabidopsis AGOs 2 and 3 do not possess the 3' UTR miR403 target site and would not provide an elevatable back-up system.

3.2.2 AGO7
Alleles of the ago7 mutant were originally identified in a reverse genetics screen for plant lines exhibiting accelerated juvenile-to-adult phase change (Hunter et al., 2003; Peragine et al., 2004; Yoshikawa et al., 2005). This screen also identified mutant alleles of DCL4, RDR6 and SGS3 (dcl4, rdr6 and sgs3 mutant plant lines respectively). Subsequent studies revealed that DCL4, RDR6 and SGS3 are essential players in the biogenesis of tasiRNA sRNAs from the non-protein-coding TAS transcripts, Tas1, Tas2, Tas3 and Tas4. In addition to expressing an accelerated juvenile-to-adult transition, ago7 plants display floral morphogenesis defects, a phenotypic characteristic subsequently associated with plant lines where TAS3 biogenesis is disrupted (Adenot et al., 2006; Garcia et al., 2006). AGO7 has since been demonstrated to exclusively function in the TAS3 biogenesis pathway (Montgomery et al., 2008). In TAS3 tasiRNA biogenesis, miR390 is specifically loaded to AGO7 to direct AGO7 binding to the two miR390 target sites within the Tas3 mRNA. AGO7 cleaves the targeted transcript at only the 3' target site and this event identifies the cleaved mRNA for RDR6-directed dsRNA synthesis (Yoshikawa et al., 2005; Montgomery et al., 2008). A subset of the TAS3-specific tasiRNAs are subsequently loaded to AGO1 to target the auxin response factor family members Arf3 and Arf4 for cleavage-based repression. ARF3 and ARF4 are required for specification of the adaxial fate of Arabidopsis rosette leaves (Fahlgren et al., 2006; Garcia et al., 2006), therefore, AGO7-mediated, miR390-directed regulation of gene expression is essential for normal plant development in Arabidopsis.

As mentioned above, array data suggests that AGO7 shares an overlapping expressional domain with the two other AGO2/AGO3/AGO7 clade members (Schmid et al., 2005). Fusion of the AGO7 promoter to the GUS reporter gene revealed that AGO7 is predominantly expressed in the vasculature of seedlings and in the cells and tissues immediately surrounding the SAM (Montgomery et al., 2008). GUS expression was also observed in the adaxial-most cells of developing leaf primordial to again demonstrate the importance of AGO7 expression for normal leaf development (Fahlgren et al., 2006; Garcia et al., 2006). As described for ago2 plants, the ago7 mutant is hyper-susceptible to TCV infection (Qu et al., 2008). This suggests a possible additional antiviral role for AGO7. However, ago7 mutants are not hyper-susceptible to any other plant virus, and furthermore, direct association of AGO7 with the accumulation of viral-specific siRNAs remains to be demonstrated.

Besides its preferential association with miR390, the 5' terminal nucleotide preference of AGO7 is unknown. Unlike AGO1, AGO2, AGO4 and AGO5, the AGO7/miR390 association is not based on the 5' terminal nucleotide of the sRNA. Replacing the 5' terminal adenine of miR390 with a cytosine residue does not influence the preferential association of this sRNA with AGO7 (Montgomery et al., 2008), suggesting a specialized association mechanism.

3.3 The AGO4/AGO6/AGO8/AGO9 clade
3.3.1 AGO4
The ago4 mutant was originally identified in a forward genetics screen for mutants impaired in TGS of the SUPERMAN locus, along with the RdDM machinery proteins
CHROMOMETHYLASE3 (CMT3) and KYPTONITE (KYP; Zilberman et al., 2003). Subsequent research has shown that AGO4 functions in the effector step of RdDM to maintain sRNA-directed DNA methylation of repetitive genomic sequences (e.g. maintains transposons in their epigenetically silent state; Xie et al., 2004; Zilberman et al., 2004). Array data shows that AGO4 is expressed ubiquitously (Schmid et al., 2005). This is consistent with the AGO4 promoter-GUS reporter gene expression pattern observed in transgenic Arabidopsis plants, where GUS was found to be expressed throughout developing embryos, mature leaves and flowers (Havecker et al., 2010).

In correlation with its role in sRNA-directed DNA methylation, AGO4 appears to be exclusively located in the nucleus (Li et al., 2006). In the nucleus, AGO4 co-localizes with RdDM machinery proteins, including the plant specific DNA-dependent RNA polymerases, PolIV and PolV, as well as RDR2, DCL3 and DOMAINS REARRANGED METHYLASE2 (DRM2) in two types of specialized nuclear compartments, namely Cajal-bodies and AB-bodies (Li et al., 2006; Pontes et al., 2006). Co-localization of AGO4 to two specialized nuclear bodies suggests that AGO4 is not only required for sRNA-directed DNA methylation, but also for the maintenance of heterochromatin (Irvine et al., 2006). Accordingly, AGO4 preferentially binds repeat-associated (rasiRNAs) and heterochromatin-specific (hcsiRNAs) siRNAs of the 24-nt size class. Although there is an even distribution of 24-nt rasiRNAs and hcsiRNAs with 5' terminal adenine, cytosine, guanine and uracil residues in Arabidopsis, AGO4 preferentially binds sRNAs of this size class with 5' terminal adenine residues (Mi et al., 2008; Havecker et al., 2010). Curiously, AGO4 does not appear to be involved in either the biogenesis or effector step of another class of endogenous 24-nt sRNA, the natsiRNA class. This class of 24-nt siRNA was demonstrated to accumulate to wild-type levels in the absence of AGO4 activity (Xie et al., 2004). Similar observations were made for plant lines deficient in the activity of a number of other RdDM machinery proteins to suggest that the rasiRNA/hcsiRNA and natsiRNA silencing pathways operate through different AGO-catalysed effector complexes.

In virus-infected wild-type Arabidopsis plants virus-specific 24-nt siRNAs accumulate to readily detectable levels, although at much lower levels than those of virus-specific 21-nt siRNAs. However, ago4 mutants do not appear to be hyper-susceptible to any plant virus, to suggest that another 24-nt binding AGO family member may be responsible for sRNA-directed methylation of viral transcripts. Intriguingly, ago4 plants are hyper-susceptible to the bacterial pathogen *Pseudomonas syringae* (Agorio & Vera, 2007), suggesting that AGO4 may be involved in directing a defence response against only specific pathogens. Alternatively, the epigenetic de-repression of other genes in this mutant background could be causing this hyper-susceptibility effect.

### 3.3.2 AGO6

AGO6 specifically acts at the transcriptional level in the hcsiRNA-directed RNA silencing pathway (Zheng et al., 2007; Havecker et al., 2010). The ago6 mutant was originally identified in a forward genetics screen for plant lines where the expression of a transcriptionally-silent transgene was reactivated in the *ros1* mutant background (Zheng et al., 2007). The authors showed that the level of transcriptional reactivation was higher in the *ago4/ros1* double mutant background than in the *ago6/ros1* mutant. This suggests that AGO6 does not play as wide a role in sRNA-directed heterochromatin RNA silencing as that directed by AGO4 in Arabidopsis. However, AGO6 does appear to be partially redundant...
with AGO4 function as the level of transgene reactivation was demonstrated to be even higher in the ago4/ago6/ros1 triple mutant, compared to either of the analysed double mutants. Furthermore, array and reporter gene expression data reveal that the expression domain of AGO6 overlaps that of AGO4 (Schmid et al., 2005; Havecker et al., 2008). Taken together, these analyses suggest that these two AGO family members act on a shared subset of repeat elements, and that their overlapping function occurs in similar tissues and at the same developmental time point.

### 3.3.3 AGO8/AGO9

As with AGOs 2 and 3, AGO8 and AGO9 are predicted to have arisen from a recent gene duplication event (Vaucheret, 2008). The amino-acid sequences of these two AGOs are very similar (although AGO8 is not annotated in TAIR), and the AGO8 and AGO9 genes are almost adjacent to one another on chromosome 5 of the Arabidopsis genome. According to array data, the tissue-specific expression patterns of Ago8 and Ago9 mRNAs are also highly similar (Schmid et al., 2005). However, the Ago8 transcript is expressed at a much lower level than Ago9 and contains a splicing-induced frame-shift, which is predicted to render the AGO8 protein non-functional (Takeda et al., 2008).

To date, no forward genetic ago8 or ago9 alleles have been identified in any mutant screening population. T-DNA insertion-mutant lines of AGO8 or AGO9 are wild-type in appearance and have unchanged miRNA, tasiRNA and siRNA accumulation levels. However, a recent study has implicated the involvement of AGO9 in the siRNA-directed maintenance of the silencing state of several classes of repetitive DNA element (Havecker et al., 2010), and closer examination of the ago9 mutant has revealed a previously overlooked apomixes-like fertilization-independent seed production phenotype (Olmedo-Monfil, 2010).

### 4. Function of mammalian AGO proteins

The four AGO proteins (AGOs 1 to 4) encoded by the mouse and human genomes perform key effector roles directed by the three endogenous sRNA classes of these respective species, namely the miRNA (Lee & Ambros, 2001), siRNA (Meister & Tuschl, 2004) and PIWI-interacting RNA (piRNA) classes (Vagin et al., 2006). Unlike members of the Arabidopsis AGO family, mammalian AGOs do not exhibit 5’ terminal nucleotide preferences for either the loading or sorting of sRNAs. The four mammalian AGOs appear to bind sRNAs with little discrimination, except for sense piRNAs which have been shown to specifically incorporate into AGO4 (Arvin et al., 2008). In Drosophila and C. elegans however, the degree of complementarity between sRNA duplex strands and/or the structure of the sRNA has been demonstrated to strongly influence the sorting of sRNAs into their respective AGO protein family members (Tomari et al., 2007; Forstemann et al., 2007; Steiner et al., 2007). Taken together, these studies suggest that in contrast to Arabidopsis, Drosophila and C. elegans, mammals lack specific rules for sRNA/AGO sorting.

Although all four mammalian AGOs exhibit similar siRNA binding affinities and each appear to specify a role in posttranscriptional regulation of miRNA expression, only AGO2 performs the effector function of siRNA-directed RNA silencing (Liu et al., 2004; Meister et al., 2004; Rivas et al., 2005). SiRNAs delivered into mammalian cells direct AGO2-mediated RNA silencing through target transcript cleavage and require a high degree of sRNA/target
RNA complementarity (Liu et al., 2004). In contrast, mammalian miRNAs have a low level of complementarity to their targets with the predominant mode of miRNA-directed gene expression regulation mediated via translational repression mechanisms. All four AGOs, including the Slicer efficient AGO2, direct translational repression through their shared partnership with GW182. With the assistance of GW182, miRNA-loaded mammalian AGOs typically target their regulated mRNAs in the 3' UTR. Translational repression is thought to occur through disruption of crucial interactions between the 5' Cap and the 3' poly-A tail of the regulated transcript, leading to a reduction in translational initiation and/or transcript destabilization (Liu et al., 2005; Eulalio et al., 2008).

In addition to performing the sRNA effector function, there is growing evidence to support important roles for mammalian AGOs in sRNA processing. For example, AGO2 has been implicated in miRNA biogenesis through direct binding of the precursor miRNA (pre-miRNA) molecule (Cheloufi et al., 2010; Cifuentes et al., 2010; Diderichs & Haber, 2007; Tan et al., 2009). In the canonical miRNA biogenesis pathway (Fig.3), the DNA-dependent RNA polymerase II (Pol II)-transcribed primary miRNA (pri-miRNA) transcript is recognized and bound by the dsRNA binding domain (dsRBD) protein Pasha. Pasha in turn recruits the RNase III-like endonuclease Drosha to form a multi-protein complex, the Microprocessor. Within the Microprocessor, and with the assistance of Pasha, Drosha cleaves the pri-miRNA to liberate a shorter precursor molecule of 60 to 70-nt in length, the pre-miRNA. The pre-miRNA is subsequently transported to the cytoplasm and further processed by Dicer and its dsRNA binding domain partner protein Laquacious to release a 22 to 23-nt mature miRNA duplex. Following the unwinding of duplex strands, the mature miRNA sRNA is then loaded by one of the four mammalian AGOs to form miRNA-loaded RISC (miRISC). However, and as mentioned above, recent studies have revealed that the direct binding of AGO2 to the pre-miRNA dsRNA and not the mature miRNA itself, can also form an active miRISC, capable of either; i) cleaving miRNA targets in vitro in the absence of Dicer activity (Tan et al., 2009), or; ii) processing of the pre-miRNA dsRNA into a shorter intermediate known as the AGO2-cleaved pre-miRNA (ac-pre-miRNA) (Diderichs & Haber, 2007). The significance of the AGO2-generated pre-miRNA remains to be determined, however, in mice and zebrafish the biogenesis of a conserved vertebrate miRNA, miR451, has been shown to also require the endonucleatic activity of AGO2, and that its biogenesis occurs independently of Dicer (Cheloufi et al., 2010; Cifuentes et al., 2010). In this non-canonical biogenesis pathway (Figure-3), AGO2 directly binds pre-miR451 and trims this molecule to produce the mature miR451 sRNA. The resulting mature sRNA can then be directly loaded by AGO2 to form an active miRISC.

The exact mechanism of how AGO2-mediated “trimming” occurs remains to be determined. Examination of the pre-miR451 sequence reveals two differences to most mammalian miRNA precursor transcripts, and Cheloufi et al. (2010) proposed that these differences may identify the miR451 precursor molecule for entry into the AGO2-mediated miRNA biogenesis pathway. Firstly, the pre-miR451 is 17-nt shorter than most pre-miRNAs. This marks pre-miR451 as an unlikely Dicer substrate since it has been shown in mice extracts that Dicer cannot process shorter pre-miRNA efficiently (Siolas et al., 2005). Secondly, the mature miR451 sequence includes some of the loop region and complementary arm of the stem-loop of the precursor molecule. This unique precursor structure, and mature miRNA position within the precursor, may interfere with Dicer’s ability to recognize this open-looped molecule for processing.
5. Conclusions

The availability of the Arabidopsis genome sequence and insertion mutants for the majority of its genes has been invaluable in advancing the understanding of plant cell and developmental biology. Knowing exactly how many Dicer-, Agonaute-, DRB-, and RDR-like genes are present in the genome, coupled with insertion and point mutants of these and associated genes (including those identified from forward and reverse genetic screens) have allowed plant biologists to generate an in depth picture of the parallel RNA silencing
pathways in Arabidopsis. It is tempting to conclude that we now have a close-to-full picture of AGO-mediated sRNA-directed regulation in plants. Some broad simplifications that can be made are:

1. AGO1 is the most important AGO, without which a plant cannot survive, regulate its development, or defend itself against viral infection.
2. The 10 members of the Arabidopsis AGO protein family can be divided into 3 functional groups:
   i. RNA slicers
   ii. RNA binders, and
   iii. chromatin modifiers
3. The above functional groupings largely follow the sequence-based phylogenetic clades described above. Members of the AGO1/5/10 clade are Slicers, AGOs 2, 3 and 7 bind sRNAs (however AGO7 has been demonstrated to direct Tas3 cleavage), and the four remaining family members of the AGO4/6/8/9 clade are modifiers
   i. AGO 1/5/10 - slicers
   ii. AGO 2/3/7 - binders (AGO7 also cuts)
   iii. AGO 4/6/8/9 - modifiers
4. Each clade has a main player with ubiquitous and high level expression (AGO1, AGO7 and AGO4) and a pair of important reproductive-specific (flower/embryo-specific) players (AGO5/10, AGO2/3 and AGO8/9). See Figure 4.
5. Arabidopsis AGO1 has similarities in Slicer activity with mammalian AGO2.

Fig. 4. Tissue expression domains of Arabidopsis AGO proteins. The expression of the ten Arabidopsis AGO family members was divided into tissue types, namely the flowers (FLOW), vasculature (VASC), roots (ROOT), rosette leaves (ROSS), cotyledons (COTY), embryos (EMBR), siliques (SILIQ) and shoot apical meristem (SAM). Segments shaded green represent expression, whereas grey shaded segments represent no expression.
However, there are many important questions, both in Arabidopsis and other plant species, yet to be answered before a full picture can be realised. No structural features have been directly determined for any eukaryotic AGO protein. Inferences about their structure/function relationships have been made from the crystal structure of a prokaryotic AGO-related protein and sequence-based structural predictions. Therefore:

i. are the structures of the Arabidopsis AGO proteins similar within a clade but divergent between clades?

ii. Are the structural differences between family members reflective of their different modes of action?

iii. Which Arabidopsis AGOs are responsible for translational repression of mRNAs and viral RNAs?

iv. Which family member(s) use the RDR6/DCL2-generated 22-nt secondary transitivity-inducing siRNAs?

v. What are the 5’ nucleotide preferences for AGOs 3,7,8 and 10?

vi. Is there an AGO that is preferentially loaded with sRNAs possessing a 5’ terminal guanine residue?

vii. How are different AGOs loaded with the appropriate sRNA and the “correct” dsRNA duplex strand chosen?

viii. Is there a non-canonical miRNA biogenesis pathway in Arabidopsis similar to the mammalian AGO2 system?

All of these and many other basic questions remain to be answered. There is also a broader question: are the Arabidopsis sRNA/AGO-mediated processes representative of those in other plant species? For example, rice has 18 AGOs (Fig. 5). While the Arabidopsis AGO2/3/7 and AGO4/6/8/9 clades appear to have almost exactly the same number of

![Phylogenetic tree](https://www.intechopen.com)

**Fig. 5.** Phylogenetic tree constructed from the full-length amino acid sequences of the ten Arabidopsis AGO family members and nineteen Rice AGO family members. The phylogeny clade divisions marked using different colour shades and labelled adjacently. Drawing not to scale.
counterpart rice genes, there are four AGO1 homologues in rice. Perhaps this expansion will be common in different plant species, reflective of the importance of the AGO1 function, as well as providing functional redundancy or tissue/cell-specific activities. As the sequences of completely assembled genomes of different plant species become known over the next few years, many of these questions will be answered. It is also possible that the different rice AGO1s have a spectrum of functions, and it is intriguing that the rice AGO5 family seems to be expanded and that AGO17 and especially AGO18 are orphans. It will be interesting to see whether these genes have functions absent in Arabidopsis, what roles they play, and whether these roles are monocot specific.

In conclusion, the study of Ago genes in Arabidopsis has revealed a family of proteins with elegant and complex activities that are essential for the normal development, genome stability and viral protection of the plant. However, it is clear that we still have much to learn about their mechanisms, actions and diversity across the plant kingdom.

6. References


www.intechopen.com


Mallory, AC., Hinze, A., Tucker, MR., et al. (2009) Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development: importance of pairing to the microRNA 5′ region. PLoS Genetics, Vol.5, No.9, (September 2009), e1000646, ISSN 1553-7404


www.intechopen.com


RNA functions broadly as informational molecule, genome, enzyme and machinery for RNA processing. While these functions reflect ancient activities, they also remain vital components of contemporary biochemical pathways. In eukaryotic cells RNA processing impacts the biogenesis of RNA molecules of essentially every shape and function. The collection of articles in this volume describes the current state of understanding of the broad array of RNA processing events in animal and plant cells, key unanswered questions, and cutting edge approaches available to address these questions. Some questions discussed in this volume include, how viruses subvert the RNA processing machinery of the host cell, how the coordination of co-transcriptional RNA processing is regulated at the level of chromatin, the status of RNA processing in plant organelles, and how micro RNA machinery is biosynthesized and regulated.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
