

Review

Biogenesis and Mechanism of Action of Small Non-Coding RNAs: Insights from the Point of View of Structural Biology

Marina C. Costa¹, **Ana Lúcia Leitão**² and **Francisco J. Enguita**^{1,*}

¹ Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, Lisboa 1649-028, Portugal; E-Mail: marinacosta@fm.ul.pt

² Departamento de Ciências e Tecnologia da Biomassa, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Campus da Caparica, Caparica 2829-516, Portugal; E-Mail: aldl@fct.unl.pt

* Author to whom correspondence should be addressed; E-Mail: fenguita@fm.ul.pt; Tel.: +351-21-7999503; Fax: +351-21-7999412.

Received: 30 May 2012; in revised form: 17 July 2012 / Accepted: 2 August 2012 /

Published: 17 August 2012

Abstract: Non-coding RNAs are dominant in the genomic output of the higher organisms being not simply occasional transcripts with idiosyncratic functions, but constituting an extensive regulatory network. Among all the species of non-coding RNAs, small non-coding RNAs (miRNAs, siRNAs and piRNAs) have been shown to be in the core of the regulatory machinery of all the genomic output in eukaryotic cells. Small non-coding RNAs are produced by several pathways containing specialized enzymes that process RNA transcripts. The mechanism of action of these molecules is also ensured by a group of effector proteins that are commonly engaged within high molecular weight protein-RNA complexes. In the last decade, the contribution of structural biology has been essential to the dissection of the molecular mechanisms involved in the biosynthesis and function of small non-coding RNAs.

Keywords: non-coding RNA; structural biology; X-ray crystallography; microprocessor; miRNAs; Dicer; DGCR8; argonaute

1. Introduction

The central dogma of biology holds that genetic information normally flows from DNA to RNA and to proteins. As a consequence it has been generally assumed that genes code for proteins, and that proteins fulfill not only most structural and catalytic but also most regulatory functions in cells [1]. This is essentially true in prokaryotic organisms whose genomes are almost entirely composed of closely packed protein coding sequences. However, this is not the case in higher organisms in which proteomes and their coding sequences occupy only a tiny fraction of the genome. Around 97–98% of the transcriptional output of the human genome is non-protein coding RNA (ncRNA). This estimate is based upon the fact that intronic RNA constitutes 95% of primary protein coding transcripts (pre-mRNAs) and on a range of observations that suggest that there are a large number of ncRNA transcripts that do not contain substantial open reading frames and which may represent at least half of all the transcripts [2,3]. However, it is hard to escape the general conclusion that either the human genome is replete with useless transcription units, or that these RNAs are fulfilling some unexpected functions [4].

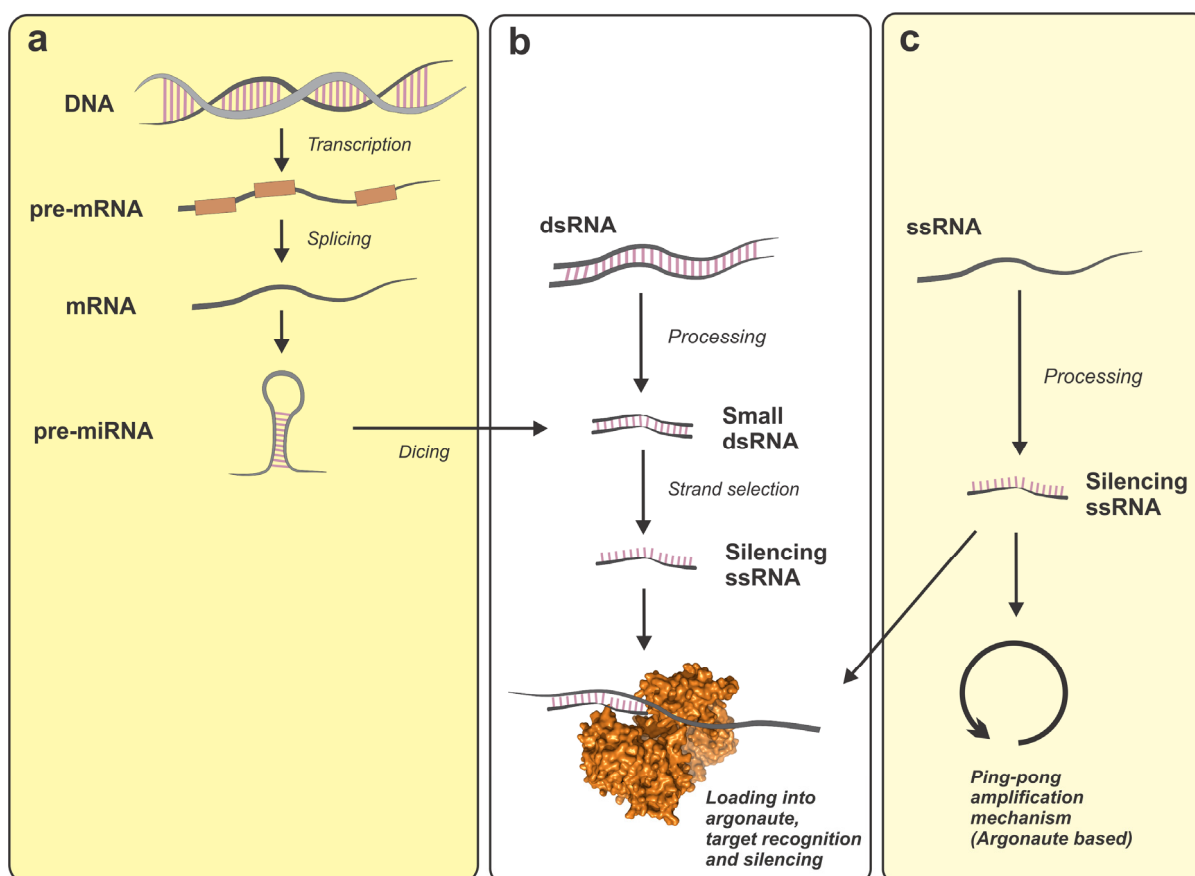
Many different ncRNAs with different functions in eukaryotic cell and developmental biology have already been described [5–7]. The fact that noncoding RNAs carry out the majority of the transcription of the genomes of humans and other complex organisms suggests that a second tier of genetic output and a network of parallel RNA-mediated interactions has evolved in these organisms, which may enable the integration and coordination of sophisticated suites of gene expression required for differentiation and development [8–10]. The expansion of the complement of ncRNAs in the higher organisms also suggests that the evolution of complexity may not have been simply dependent on an expanded repertoire of proteins and protein isoforms, but on a (much) larger set of genomic design instructions embedded in trans-acting RNAs (and cis-acting receiver sequences), which form the basis of a cascade of programmed response networks capable of implementing stored sequences of dynamical activities in response to internal and external stimuli. It is also likely that alteration in this control architecture is responsible for much of the phenotypic variation that is observed between individuals and species (such as vertebrates) that use a relatively common set of functional components, with the remainder of the variation (and the majority of catastrophic problems) due to variation in the components (the proteins) themselves [11,12]. Many ncRNAs are simply unprocessed primary transcripts but in other cases there are formed from the exons of spliced transcripts, which may also be alternatively spliced or polyadenylated. Some ncRNAs are derived from the further processing of exons such the miRNAs produced by some transcripts, and of introns of both protein coding and ncRNA genes as exemplified by snoRNAs.

There are three major families of small non-coding RNAs (ncRNAs) in eukaryotic cells: micro-RNAs (miRNAs), piRNAs and small-interfering RNAs (siRNAs). These families are different in their origin, but they share specific steps in their biosynthetic pathways and regulatory mechanisms (Figure 1). In general, small ncRNAs are regulators of genomic output either at the transcriptional or at the post-transcriptional levels [13,14]. Regulatory control by small ncRNAs requires two main groups of proteins: processors and effectors. The processors are enzymes with nuclease activity able to excise small RNAs from specific RNA transcripts. On the other hand, the group of the effectors is constituted

by a diverse cohort of RNA-binding proteins responsible for the stabilization, transport and regulatory activity of the small ncRNA over its cognate target [15].

Over the last decade, Structural Biology methods such as X-ray crystallography, NMR spectroscopy and more recently electron microscopy, have deeply contributed to the overall knowledge of the structure-function relationships among the components of the functional pathways of small ncRNAs [16–18]. In this review we will summarize these contributions, with special emphasis to the human key proteins involved in the biosynthesis and function of small ncRNAs.

Figure 1. Pathways of small non-coding RNA generation in eukaryotic cells showing their main steps and relationships between groups. micro-RNA (miRNA) and small-interfering RNAs (siRNAs) are generated by nucleases (Drosha and Dicer) from dsRNA precursors, whereas PIWI-interacting RNAs (piRNAs) are produced from long single-stranded RNA precursors by an already not completely understood mechanism. (a) miRNA canonical pathway; (b) small-interfering RNA pathway; (c) piRNA pathway. dsRNA: double-stranded RNA; ssRNA: single-stranded RNA.



2. Micro-RNAs (miRNAs)

Among non-coding RNAs, the best known family is constituted by miRNAs. These RNA molecules firstly discovered in the worm *Caenorhabditis elegans*, are short ncRNAs (21–23 nt) generated by a complex cellular pathway which starts with the transcription of specific genomic loci and flows from the nucleus to the cytoplasm with the help of specialized nucleases that cleave the RNA transcripts to produce a mature miRNA [19,20]. In the nucleus a tandem of proteins, Drosha and DGCR8, will

recognize and excise small RNA hairpin loops formed during transcription in specific transcriptional units [21,22]. Drosha is a nuclease whereas DGCR8 recognizes the stem-ssRNA junction in pri-miRNAs. The excised loops called pre-miRNAs are exported to the cytoplasm through the nuclear pores with the help of Exportin-5 [23]. In the cytoplasm, the pre-miRNAs are recognized by a type III ribonuclease called Dicer that will cut the loop to generate a small dsRNA fragment of 19–23 pbs that will remain bound to the enzyme [24,25]. In this situation, the complex between Dicer and the small dsRNA will recruit several proteins of the argonaute family that will form the RNA induced silencing complex (RISC) that will select one of the chains of the dsRNA to generate a mature miRNA molecule [26,27]. Mature RISC complexes will be targeted to complementary sequences in mRNA transcripts, interfering with the translation process and consequently reducing the protein expression from mRNA transcripts [28–30]. In animal cells the complementarity of the miRNA and its target is not perfect, whereas in plants miRNAs are often 90–100% homologous to the target mRNAs. In this last case the RISC complex can induce mRNA targeted degradation, catalyzed by one of the argonautes, AGO2 protein [31,32]. An alternative pathway for the formation of miRNAs from intronic RNA transcripts has been also described. In this pathway, nuclear processing of RNA hairpins by Drosha/DGCR8 is skipped and the pre-miRNA are fully constituted by small introns that are excised exclusively by the splicing machinery. These miRNAs generated by skipping of the nuclear part of the native pathway are called Mirtrons [33,34].

Genes encoding miRNAs are frequently clustered together, and they are very abundant in the human genome having a preferential location within introns [35,36]. However it is already known that some of these miRNAs are localized in exons. Alternative splicing of these exons in a cellular or tissue specific manner could control the production of some miRNAs that will regulate different genes also in a specific way.

Regulatory activity exerted by miRNAs is a complex process in which a single mRNA transcript can be targeted by several miRNAs at the same time, and also a single miRNA can regulate hundreds of different mRNA targets [37–39].

3. piRNAs

Many organisms including mammals produce short-RNAs to regulate gene expression. piRNAs derive from long single-stranded RNAs transcribed from specific regions within the genome. In fact, piRNAs are generated from regions harboring transposons, and they were firstly described as a mechanism to protect the cells against the internal attacks of transposons [40]. They are preferentially expressed in germinal lines, however recent evidences showed that they could have expanded regulatory roles also in somatic cells [41–44]. piRNAs are able to interact with a specialized family of argonaute proteins called PIWI that will guide them to their targets and silence the transposon transcripts by their slicing activity [45,46].

piRNAs are slightly longer than miRNAs (24–31 nt in length), typically 3'-methylated and their biogenesis is still not well understood. The fact is that piRNAs are generated from longer RNAs transcribed from direct and reverse DNA strands from transposon regions, and sliced into small mature piRNAs by a mechanism that in *Drosophila* involves different proteins in somatic and germ lines. In somatic cells, the piRNA precursor is cleaved into mature piRNAs by a coordinated reaction

involving the putative helicase Armitage and the putative nuclease Zucchini [44,47]. In germ cells, the biogenesis of piRNAs is dependent of two proteins, AUB and AGO3. Mature piRNAs captured by AGO3 will recognize their targets that will be cleaved by the AGO3 slicer activity and subsequently the generated fragments will constitute secondary piRNAs that amplify the signaling loop in a “ping-pong” mechanism [48–50]. This simple mechanism allows a rapid and efficient transposon silencing in germ cells even with a small amount of generated primary piRNA [51–53]. Some authors considered that the transposon rich-clusters in the genome together with the ping-pong amplification cycle constitutes a RNA-based immune system against RNA threads [54].

piRNA are considered as the precursor of an ancient mechanism of defense against genetic threads suffered by cells. In fact these small non-coding RNAs have been also found in primitive organisms such as cnidarians and sponges [55]. Despite their relatively well understood role in germ lines, the functions of piRNAs in somatic cells are still far from being understood, and more investigation is needed.

4. siRNAs

The small-interfering RNAs (siRNAs) are small dsRNAs generated by Dicer from longer precursors [56]. They were first described and further characterized in worms and plants as products of the catalytic action of RNA-dependent RNA polymerases (RdRPs) [57–59]. RdRPs are able to generate long dsRNAs that will be subsequently exported to the cytoplasm, processed by Dicer and recruited to the specific target by the cytoplasmic silencing complexes [60–62]. In plants, siRNAs are sometimes produced as a response to an external stress or as defense mechanism against genetic mobile elements or viruses [63,64].

However, in animals the apparent absence of RdRPs in their genomes prevented the search for endogenous siRNAs until the accidental discovery of LINE-1, a retro-transposon detected in human cell cultures able to produce a bidirectional RNA transcript using a double promoter system in the sense and antisense orientations [65,66]. In flies, next generation sequencing data of RNA pools obtained from AGO2 immunoprecipitation allowed also to identify a siRNA population clearly distinguishable from the miRNAs and piRNAs. Small interfering-RNAs from *Drosophila* are 21 nucleotides long, have modifications at the 3' ends and are double-stranded [45,67]. In many cases, siRNAs are related with pseudogenes containing regions with tandem inverted repeats that allow the formation of long intramolecular dsRNA structures susceptible to be processed by Dicer [45].

Moreover, siRNAs have been also identified in mouse oocytes. As in flies, mouse siRNAs are 21 nucleotide long and Dicer-dependent products and in some occasions their targets are within protein coding genes [40,60]. However, the regulatory functions of siRNAs are still not clear in the eukaryotic context. They probably have evolved from a more primitive defense system, and some authors consider that the transcriptional units producing siRNAs are still under evolutionary pressure [55,68]. Indeed, the key challenge will be to understand the real role of endogenous siRNA, more specifically those that will target protein coding genes and how they regulate mRNA expression.

5. Small Non-Coding RNA Processors

Typically, small ncRNAs are generated from bigger RNAs by the help of specific endonucleases. These endonucleases belong to the RNase III family, and constitute hot-spots in the production of small ncRNAs. Small ncRNA processing enzymes are modular proteins, harboring domains for the binding and recognition of the RNA precursor and nuclease domains for its processing [69,70]. In plants and mammals, the small ncRNA production is compartmentalized by the eukaryotic cell structure in two different locations: nucleus and cytoplasm.

5.1. Droscha and the Nuclear Microprocessor

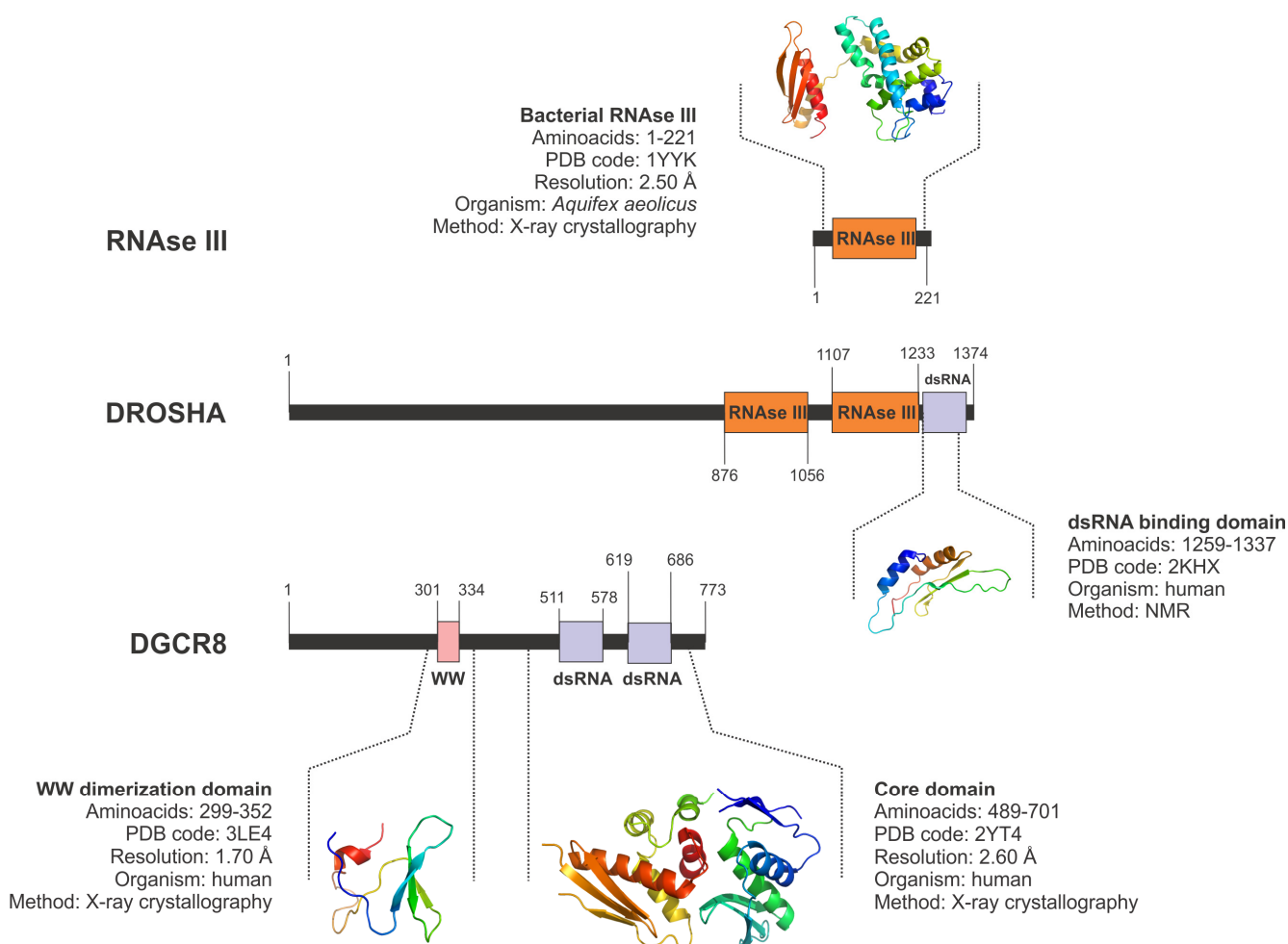
The first step in the biosynthesis of miRNAs in animal and insect cells is catalyzed by a tandem of proteins that form the microprocessor complex: Drosha, a type III RNase and DGCR8, a dsRNA-binding protein responsible for the recognition of specific hairpin loops [20,71]. These two proteins represent the essential requirement for the initial processing of pri-miRNA transcripts [21]. However, in human cell extracts, Drosha has been described as the core of two types of microprocessor complexes [72,73]: a binary complex composed only of Drosha and DGCR8 with a consistent pri-miRNA processing activity, and a second larger complex containing Drosha, DGCR8 and several accessory proteins [72,73]. It is not clear whether this bigger complex is assembled, because some of the accessory factors include hnRNP proteins, the dead box helicases DDX5 and DDX17, and the p68 and p72 proteins [71,72,74]. In plants, this binary microprocessor complex is substituted by a nuclear Dicer-like protein, DCL1, a type III RNase that generates precursor dsRNAs from longer transcripts [75].

The involvement of Drosha in the nuclear processing of pri-miRNAs to generate pre-miRNAs was discovered by Lee and coworkers in 2003 [76]. A year later, the partner of Drosha (Pasha) was discovered in *Drosophila* defining the so-called “microprocessor” [77]. RNA interference experiments of both elements of the microprocessor produced an accumulation of pri-miRNAs in the cell nucleus, suggesting their pivotal role in miRNA generation [77]. In humans, DGCR8 protein was determined to be the partner of Drosha and core component of the microprocessor [72]. Interestingly, deletions in the chromosome 22 that affect the gene encoding DGCR8 protein are the cause of the DiGeorge syndrome and have been well known since the early 80s [78,79]. The discovery of the relationships between the locus encoding DGCR8 protein and the miRNA processing suggested that one of the main causes of DiGeorge syndrome is an impairment in miRNA biogenesis [72].

In humans, Drosha has 1374 aminoacid residues that can be clearly divided in two regions: a N-terminal segment from aminoacids 1 to 550, highly unstructured and flexible and probably involved in protein-protein interactions with other partners of the microprocessor complex; and the C-terminal segment from aminoacids 550 to 1374 that comprises two catalytic RNase III domains and a terminal dsRNA binding domain. RNase III domains in Drosha have been characterized on the base of their sequence homology with bacterial enzymes of the same family. The presence of a highly disordered N-terminal region has prevented the enzyme to be studied by X-ray crystallography (Figure S1). Until now, the only structural data from Drosha came from NMR experiments that determined the tridimensional arrangement of the C-terminal dsRNA binding domain (Figure 2) [80]. Because of the

nature of Drosha, further structural studies to characterize the enzyme should probably employ a different approach combining methods such as NMR and electron microscopy to determine the structure of the enzyme in complex with other proteins.

Figure 2. Domain map and available structural information of the human microprocessor components, Drosha and DiGeorge critical region 8 (DGCR8), and the bacterial RNase III. Colored boxes along the sequence represent the functional domains present in each protein. The regions with already known three-dimensional structure are depicted, including information about the flanking aminoacids, PDB code for the atomic coordinates, resolution of the data used for refinement of the models, protein source and the experimental technique employed to determine the structure.



On the other hand, DGCR8 is an RNA binding protein that in humans has 773 aminoacids. In this protein, the structure of two regions has been determined by X-ray crystallography: a WW-dimerization domain located close to the N-terminal region of the protein [81], and a dsRNA binding region comprising two different dsRNA binding domains and located in the C-terminal segment (Figure 2) [80,82]. The dsRNA binding domains showed a tridimensional structure composed of an alpha-helical core fused to a small beta-sheet segment that is arranged in tandem to ensure a wider coverage of the target RNA molecule [81,82].

Unexpectedly, DGCR8 has been characterized as a heme-binding protein, being this cofactor directly involved in the efficiency of pri-miRNA processing by the nuclear complex [83–85]. In fact, spectroscopic analysis of recombinant DGCR8 showed that it is the first known example of a heme-containing protein that harbors two axial cysteine ligands to complex a ferric iron [85]. The heme binding-motif is embedded as an independent region within the dimerization domain, suggesting also a contribution to the DGCR8 monomer interactions [81]. However, the exact role of the heme cofactor in the miRNA biosynthetic pathway remains elusive and further investigation is needed.

5.2. Cytoplasmic Processors: Dicer

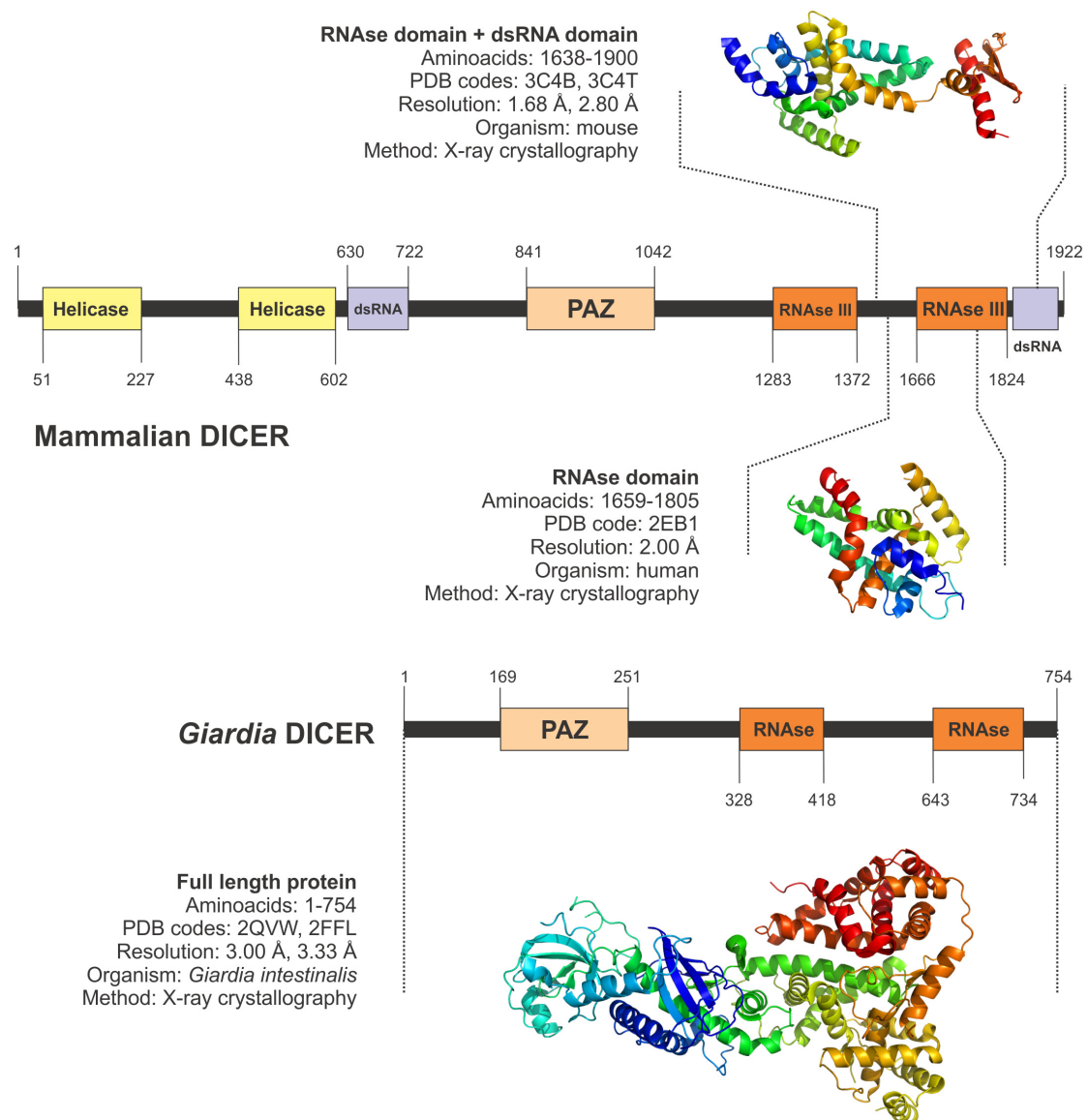
The molecular mechanism of miRNAs and siRNAs is exerted mainly in the cytoplasm over mRNA transcripts. Dicer is an enzyme with a complex duty; first it has to capture pre-miRNA loops exported from the nuclear processing machinery by recognizing the end of dsRNA, cleave them to generate 21–23 nt dsRNAs, and second, it has to act as a meeting point for the recruitment of the proteins involved in the gene silencing (RNA silencing complexes) [86–88].

Dicer-like proteins are well conserved in the entire eukaryotic world, suggesting a decisive role of these enzymes in the cell physiology. However, in complex eukaryotic cells such those belonging to mammals, Dicer proteins have evolved in a composite way comprising several duplicated domains that will perform all the protein functions (Figures S2 and S3). The most complete structural information available for a Dicer protein comes from the *Giardia intestinalis* protein. That is in fact a primitive Dicer, that has about half of the size of the human protein, and only contains two catalytic domains (RNase domains) and a RNA binding domain (PAZ domain) (See Figure 3 for details) [89,90]. Data obtained from X-ray crystallography experiments in *Giardia's* Dicer, allowed to determine that the protein acts as a molecular ruler, measuring the distance from the end of a precursor dsRNA and covering approximately 23–25 nucleotides, which is the distance between the PAZ domain and the catalytic RNase region [89–91]. Dicer can process dsRNAs in a sequence-independent manner, allowing base pairing mismatches and different internal RNA structures. In fact this is a physiologically relevant feature of the enzyme, which allows virtually any dsRNA to enter the gene silencing pathways. The only exception to the general rule is the dsRNA lacking an open terminal region, which Dicer cannot process [91]. Interestingly, the analysis of the protein from *Giardia* revealed an important conformational flexibility driven by the presence of a flexible hinge in the interface regions between PAZ and RNase domains. Authors have proposed a model of “induced folding” for the enzyme in response to the presence of a dsRNA substrate, in which the molecular hinge will adapt the PAZ and RNase domains to embrace the RNA chain [89,91,92].

Human Dicer is a molecular machine with a much more complex structure than the *Giardia's* protein (Figure 3). It conserves the catalytic core of PAZ and RNase III domains, but also includes two dsRNA binding domains and a tandem of helicase domains in its N-terminal region (DExD domain). The helicase domains are also present in fly, worm, plants and yeast Dicers, but its function in the whole catalytic process remains unclear, being suggested to be involved in the discrimination of dsRNAs termini to promote an altered reaction mode [93]. Recent structural studies using a combination of electron microscopy with X-ray crystallography data allowed us to determine that the helicase domains in human Dicer are arranged in a clamp-like shape close to the RNase III active site [70].

This architecture is also conserved in the protein from *Drosophila*, and is expected to appear also in other complex Dicers. The possible function of the helicase clamp is related with the recognition of pre-miRNA hairpin loops and long non-coding RNAs [70,94]. The absence of the helicase domain apparently does not affect the catalytic efficiency of human Dicer.

Figure 3. Domain map and available structural information of the mammalian and *Giardia* Dicers. Colored boxes represent the functional or structural domains present in each protein. The regions with already known tridimensional structure are depicted, including information about the limiting aminoacids, protein source, PDB code of the atomic coordinates, resolution and the experimental technique used to determine the structure.



Recently, the X-ray structure of a catalytic fragment of mouse Dicer has been reported. Results showed the presence of a highly conserved lysine residue in the boundary of RNase III domains that is involved in the dicing mechanism. In fact this catalytic lysine has been also described in other related RNase III enzymes such as Drosha [95].

6. Effectors of the Small Non-Coding RNA Regulatory Networks: Proteins from the Argonaute Family

Argonaute proteins are widespread family of RNA-binding proteins described in organisms ranging from bacteria to humans, and always present in the core of RNA silencing complexes. Members of the argonaute family are able to bind small guide RNAs and to direct them to specific RNA transcripts for silencing or degradation of the messenger signal. In eukaryotic organisms, argonaute proteins are main effectors of the RNA-dependent regulatory machinery, being important players in the RNA interference, miRNA regulation, and piRNA function. Some eukaryotic argonautes harbor nuclease activity, being able to directly slice or degrade specific RNA molecules. Moreover their roles in prokaryotic organisms are still far from being clarified; however, it is assumed that they could play an important role in the defense systems against external genetic threads, namely bacteriophages and more recently mobile genetic elements [96].

Structurally, all the proteins belonging to the argonaute family consist of a highly variable N-terminal domain and three conserved domains: the PAZ domain, the MID or intermediate domain, and the PIWI domain. PAZ and MID domains are involved in the proper recognition and interactions with the small guiding RNA. PAZ domain interacts with the 3' end of the small RNA and MID domain is responsible for the recognition and interaction with the phosphate group at the 5' end of the small non-coding RNAs. Meanwhile the PIWI domain will guide the RNA-AGO complex to the target RNA.

Argonaute family of proteins is divided in three different phylogenetic subfamilies: the AGO subfamily, named after the discovery of Argonaute 1 protein in *Arabidopsis thaliana*; the PIWI subfamily, firstly described in *D. melanogaster* (P-element induced wimpy testis); and the WAGO subfamily (Worm-specific AGO proteins), only present in *C. elegans*.

6.1. AGO Sub-Family

AGO proteins are present in different extend from budding yeasts to mammals. They constitute the core of the RNA induced silencing complexes in the cytoplasm (RISC complex) and in the nucleus (RITS complex) [97,98]. Moreover, AGO proteins could be consider as a molecular bridge since their function requires a simultaneous interaction with RNA and proteins. In humans and other vertebrates there are four AGO paralogs designated as AGO1-4, harboring the characteristic domain structure of the argonaute family members (Figure S4). Among them, AGO2 is unique since it is the only member of the family able to catalyze selective RNA slicing *in vivo* as the functional core of the cytoplasmic RISC complex [99]. This slicing activity is only exerted when the complementarity between the small guiding RNA to its cognate target is almost complete. On the other hand, AGO1 is an essential component of the nuclear RITS complex in yeasts, regulating the chromatin structure in response to the presence of small non-coding RNAs complementary to nascent mRNA transcripts [100,101]. Functional analysis of human AGOs using epitope-tagging techniques has shown that the population of small non-coding RNAs that bind to AGO1 and AGO2 are different, suggesting a diversity in their target specificities [102]. The other members of the family, AGO3 and AGO4 are not well characterized in terms of function and specificity of action. Recently, AGO3 has been pointed out as a

backup system for channeling miRNA action in the absence of AGO1 and AGO2 [103]. Taking into consideration all these facts and aside from these specialized functions, all mammalian Argonautes appear to cooperate in the small non-coding RNA pathways in a largely redundant and overlapping way.

6.1.1. AGO2

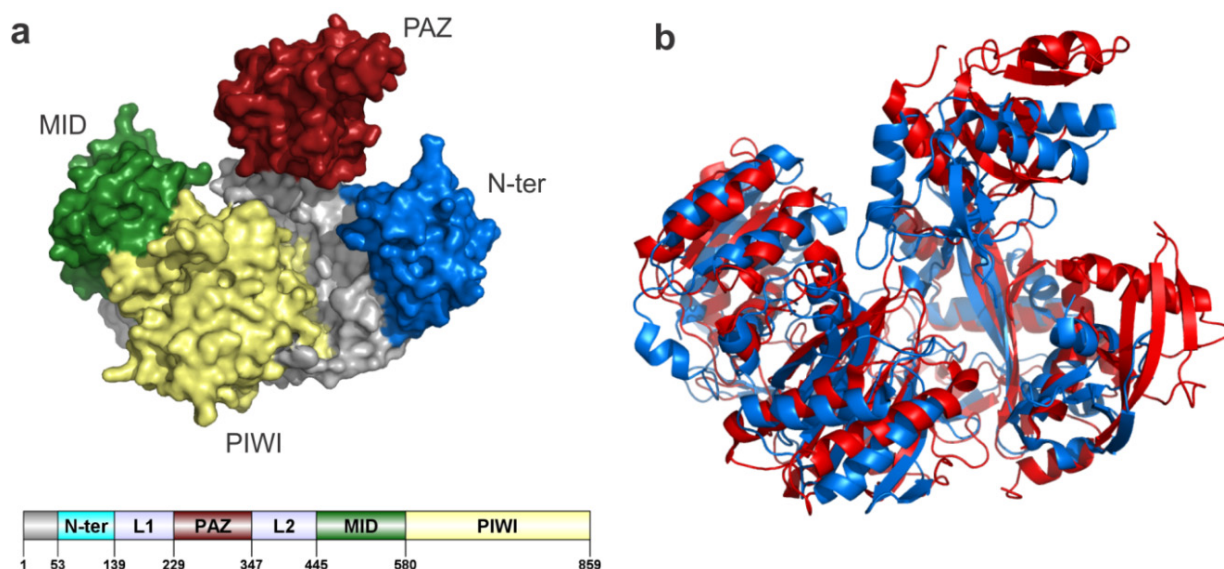
Classical structural studies of full-length AGO proteins were initially performed on the prokaryotic family members because of their favorable behavior for overexpression and purification in bacterial platforms [104–107]. However, a few preliminary studies using isolated protein domains also showed the high structural homology among the building blocks of the family [108]. Recently, the full-length crystal structure of human AGO2 has been determined at 2.3 Å resolution in complex with RNA [109].

Human AGO2 structure has the typical four domain arrangement also observed in other argonautes (Figure 4). N-terminal, MID and PIWI domains are organized to form a cavity that is partially covered by the PAZ domain lid. In comparison with its prokaryotic relatives, human AGO2 showed major architectural differences. Additional secondary structure elements present in the N-terminal, PIWI and PAZ domains will likely play a role in the recognition and binding of AGO-associated protein factors (Figure 4). However, structural alignment of human AGO2 (PDB code: 4EI1) and the argonaute protein from *Pyrococcus furiosus* (PDB code: 1U04) showed a common structural core that is evolutionarily connected also with the PIWI protein family [110].

Data obtained from the structural studies of bacterial argonautes mainly by Patel and coworkers, showed that argonaute proteins are carriers of small RNAs irrespective to the sequence, which is reflected in the absence of specific contacts between RNA bases and the protein chain [105–107,111]. Crystal structure of argonaute and the respective protein RNA complexes were determined at different resolution levels, taking advantage of the previous characterization of point mutants lacking RNA slicer activity [112,113]. Despite the natural preference of bacterial argonautes for DNA as guide strand instead of RNA, the determination of the structure of these complexes has increased the overall knowledge of its catalytic cycle. Indeed data from x-ray crystallography studies showed that bacterial argonautes embraces a two-state based mechanism for guide strand selection, target recognition and slicing.

The junction between the MID and PIWI domains is responsible for the formation of a binding pocket for the guide strand, that it is highly stabilized by interactions in the 5' end of the nucleic acid chain. Moreover, the orientation of the guide strand allows its tethering of the 3' end within the PAZ domain. Structural studies performed in mammalian PAZ domains isolated from full-length protein have confirmed the presence of a similar mechanism for attaching of the small RNA strand to the AGO2 pocket [108,114]. Interestingly, the recently solved structure of the human AGO2 showed the presence of electronic density across the MID-PIWI interface that the authors modeled as small cellular RNAs that remained bound to the protein along the purification and crystallization processes [109]. The modeled RNA is bound in a similar conformation that the DNA guiding strand found in the bacterial argonautes [109].

Figure 4. Structural elements of human AGO2 and comparison with prokaryotic members of the family. All the protein structure figures and diagrams were prepared with Pymol. (a) Molecular surface of human AGO2 with the four structural domains depicted in different colors, and relative position of the domains within the primary structure. (b) Three dimensional superposition of human AGO2 (red) and *P. furiosus* argonaute (blue) performed by the SuperPose server [115] and showing the relative position of secondary structure elements in both proteins. Sequence identity of both Argonautes is 17.0%. Global Root Mean Square Deviation (RMSD) of the three-dimensional superposition for a total of 657 alpha-carbons was 10.78.



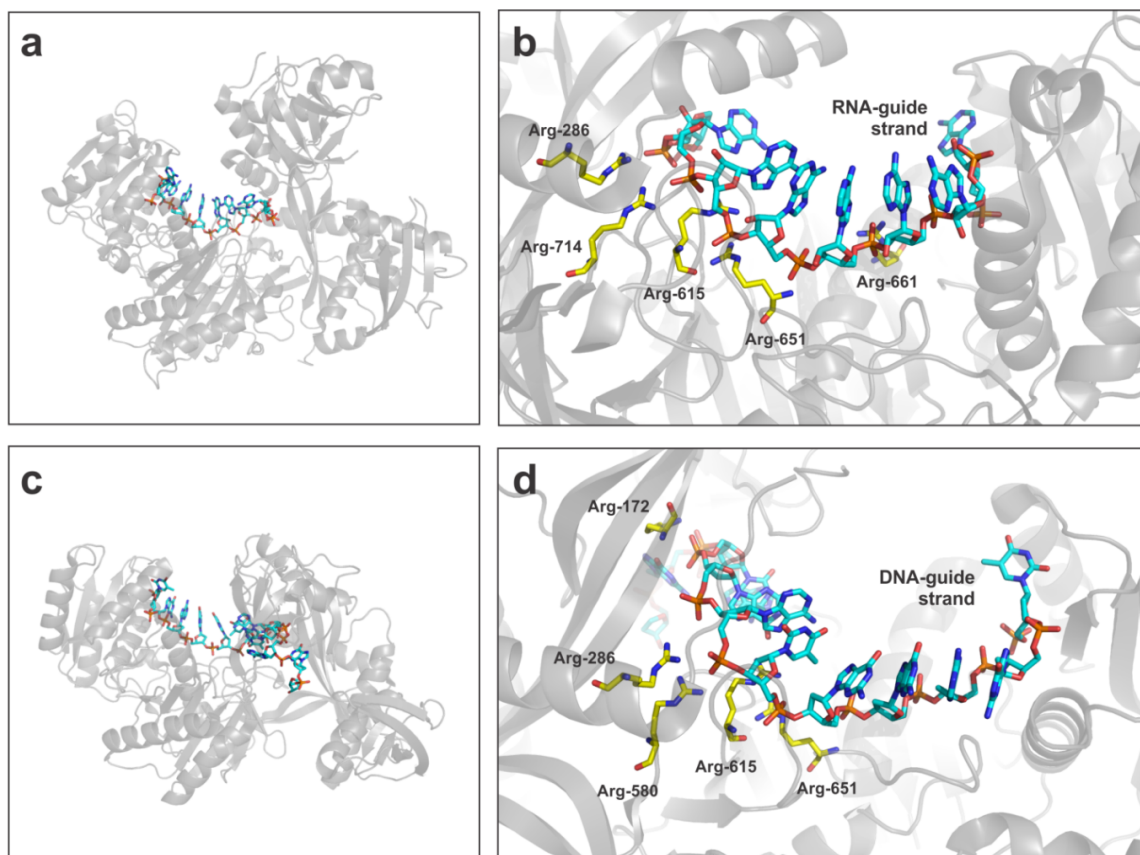
The spatial orientation of the guide strand within the argonaute pocket will favor the molecular interactions with its cognate target. The nucleic acid guiding strand is oriented exposing the edges of the nucleotides from the seed sequence to the outer part of the protein, ready to capture a target. This orientation is accomplished by direct polar interactions between the phosphodiester skeleton and positive residues in the argonaute pocket. The seed sequence of the guide strand is captured in the core of the protein, meanwhile the 5' end of the guide nucleic acid is in a flexible conformation, showing the absolute importance of the seed sequence in the global target recognition process [112,116,117] (Figure 5).

Comparison of key RNA interacting residues in human AGO2 versus bacterial argonaute showed a clear conserved pattern comprising Arg-286, -615 and -651 suggesting a similar mechanism of binding to RNA in both the members of the family (Figure 5). In human AGO2 the slicer activity is ensured by the presence of three catalytic residues, Asp-597, Asp-699 and His-807 [99,104]. Interestingly, the majority of argonaute proteins are catalytically inactive under physiological conditions, in despite of the relative conservation of the three active residues [99,102,118]. The underlying mechanistic explanations for this fact are still in the matter of the speculation, but probably there are additional external factors that would contribute to the slicer activity of AGO2 like post-translational modifications or helper proteins [103].

6.1.2. Other Members of the AGO Sub-Family

In humans and other mammals, the argonaute sub-family comprises also three additional members named AGO1, AGO3 and AGO4. Their probable redundant functions in the cell are in contradiction with their RNA-binding specificity. In fact, RIP-seq studies with tagged human argonautes showed clear differences in the cohort of small non-coding RNAs that are bound to them [102]. Human AGO2 and AGO3 appeared to be specific for miRNA binding and targeting, whereas AGO1 is more specific for siRNAs [102,119,120]. On the other hand, the function of AGO4 has been mainly characterized in plants. In fact, in *Arabidopsis thaliana* AGO4 is responsible for the guiding of several DNA methyltransferases to chromatin [62,121,122]. Moreover, AGO4 in plants is also related with the silencing of repetitive genome elements by a mechanism involving also a long intergenic non-coding RNA (lincRNA) [123,124]. In mammals, a proposed cooperative model of argonautes action based on high-throughput sequencing data is currently arising [125].

Figure 5. Mechanism of guide-strand capture by human AGO2 and bacterial argonaute from *P. furiosus* revealed by X-ray crystallography studies. **(a)** Overall folding of human AGO2 structure in complex with a RNA guide strand (PDB code: 4E11). **(b)** Structure of the binding pocket of the RNA-guide strand in human AGO2 showing the arginine residues involved in the interaction with the phosphodiester backbone. **(c)** Overall folding of argonaute from *P. furiosus* in complex with a DNA guide strand (PDB code: 3DLH). **(d)** Structure of the binding pocket of the DNA-guide strand in *P. furiosus* argonaute with the arginine residues interacting with the negatively charged groups in the DNA skeleton. Arg-286, -615 and -651 are conserved between both structures.

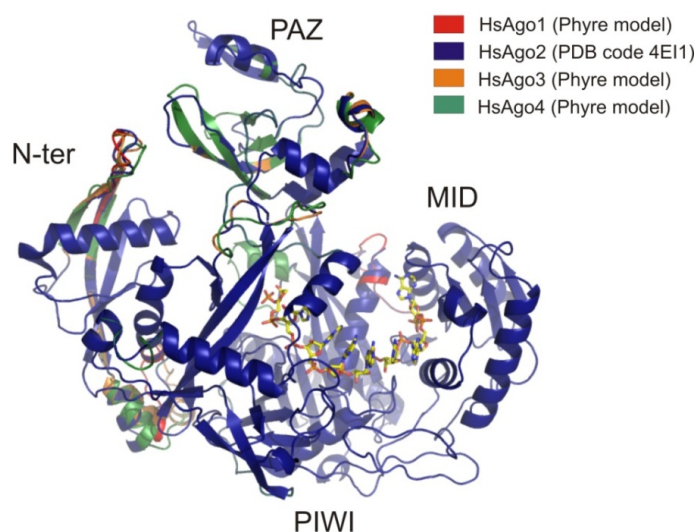


As we previously pointed out, AGO1, AGO3 and AGO4 lack slicer activity over mRNA targets, however two of the catalytic residues Asp-597 and Asp-699 found in the nuclease active AGO2 are conserved in all the family members. The last residue of the catalytic center, His-807, is only conserved in AGO2 and AGO3, being substituted by an arginine in AGO1 and AGO4.

In order to determine some structural features of all AGO family members, we have performed a homology modeling to determine the predicted three dimensional structure of the human AGO family members using AGO2 atomic coordinates as a reference. Homology modeling is a well-established and reliable protocol for structure determination when the sequence homology is higher than 40%. Some recently developed methods such as Phyre algorithm take also into account secondary structure information in order to increase the probability to get a reliable tridimensional model [126,127]. Phyre was able to build tridimensional models of human AGO1, AGO3 and AGO4 with a 99.9% confidence over the 92% of the aminoacid sequence using AGO2 coordinates as a template (PDB code: 4EI3). Atomic coordinates of Phyre models for AGO1, AGO3 and AGO4 are available as supplementary files.

Aligned models generated by Phyre are represented in Figure 6. Observed structural differences among the human AGOs are mainly present in the N-terminal and PAZ domains. MID and PIWI domains are structurally very similar in all the members of the family. Additional secondary structure elements are clearly observable in N-terminal and PAZ domains in AGO4 and AGO3 proteins in comparison with the reference model (Figure 6). Since the RNA binding pocket is homogeneously conserved in all AGOs, we can hypothesize that the structural differences observed in PAZ and N-terminal domains could be related with their different partner specificity [71,128]. However, more detailed studies are required to understand the physiological role of all the AGO proteins and their target specificity.

Figure 6. Structural superposition of the human AGO family members performed by the SuperPose server [115]. Homology models for AGO1, AGO3 and AGO4 were built with Phyre software [127] using coordinates of the human AGO2 as a template (PDB code: 4EI1). Structural domains are depicted in the figure. Calculated RMSD of the global structural alignment was 0.818 for all the alpha-carbons and 0.820 for the backbone aminoacid chains.



7. Helper Proteins and Additional Members of the Non-Coding RNA Effector Complexes

7.1. TRBP

TAR RNA binding protein 2 (TRBP2) was discovered and characterized as a trans-activation responsive protein against HIV infection in humans [129,130]. In 2005 the group of Shiekhattar discovered the interaction of TRBP with Dicer and its ability to recruit AGO2 to form a RISC loading complex (RLC) [131]. The RLC is responsible for the selection of one of the strands of the small ncRNA to generate a mature RISC complex [26]. TRBP is a 366 aminoacid protein comprising three dsRNA binding domains in tandem. Its function appeared to be related with the recruitment of argonaute proteins to the mature RISC in a RNA-dependent manner [116,132,133]. In other organisms RLC contains R2D2, an ortholog of TRBP2 [134,135]. R2D2 protein senses the relative stability of dsRNA strands, selecting one strand depending on the hybridization energy and 3'-end of the strand [133]. Specifically, the strand with the less stable 5' end at the siRNA duplex will be selected for loading the RISC complex and the other strand discarded being considered as a passenger strand.

Structural information of the isolated TRBP2 protein is limited to one of its dsRNA binding domains, determined by X-ray crystallography (PDB code: 3ADL) [136]. The structural features of the second dsRNA binding domain of the human TRBP2 are also observed in other similar domains and include a small strand segment flanked by two alpha helices [116,136]. In despite of the well-structured dsRNA binding domains, the connecting segments between domains are extremely flexible and have prevented the crystallization of the full length protein. However, more recent studies have characterized the RLC by cryo-electron microscopy; Lau and coworkers determined the structure of a TRBP2-Dicer complex at 20 Å resolution [137]. Docking of the X-ray structures of *Giardia's* Dicer onto the obtained density, allowed the authors to exactly locate the position of the enzyme within the L-shaped complex [137]. More recently the group of Nogales, has characterized the whole RLC by electron microscopy showing that the complex is established on the basis of core Dicer interactions. In the RLC, Dicer interacts with TRBP2 by its N-terminal region and with AGO2 protein with the C-terminal catalytic domain [116].

7.2. GW182

In vertebrates, miRNA regulatory functions are exerted by the loaded RISC complex that it is recruited to its cognate mRNA targets. The partial complementarity of the miRNAs with their targets induces a translational repression of the mRNA and consequently reduced levels of translated protein. In *Drosophila*, the GW182 protein is recruited to the miRNA regulatory complex by its affinity to argonaute proteins. In *Drosophila*, GW182 protein has been characterized as a partner of the argonautes engaged in the RISC complex, able to interact with AGO proteins by a domain containing GW/WG repeats [138]. These molecular interactions are needed for a productive translational repression via degradation of the poly-A tail or simultaneous recruitment of the involved mRNA transcripts to P-bodies [136,137].

In humans and other vertebrates, there are at least three GW182 paralogs named TNRC6A, TNRC6B and TNRC6C with apparently redundant roles [139]. Proteins from the TNRC6 family showed a complex primary structure with several putative argonaute-interaction domains present along

the aminoacid sequence. They also harbor a RNA-binding domain close to the C-terminal end of the protein and two glutamine-rich and glycine-rich regions in the N-terminal segment [118,139,140].

TNRC6B contains three binding sites for AGO2, within the amino-terminal glycine tryptophan (GW/WG)-repeated region that is characteristic of the GW182 family proteins. Multiple argonaute proteins are expected to be recruited to the RISC complex via interaction with GW182 protein [118,138]. Interestingly, X-ray crystallography experiments have demonstrated the direct interaction between TNRC6C and the cytoplasmic poly-adenine binding protein (PABPC1). This interaction is postulated to be directly involved in the deadenylation process and subsequent translational repression of mRNA transcripts induced by miRNAs [141,142]. However, the TNRC6C-PABPC1 does not have an observable deadenylation activity, it could be involved in the further recruitment of other specific enzymes that will catalyze the poly-A tail shortening and the transport of the selected mRNA to the P-bodies for degradation [143–145]. These interactions which are described as being critical for mRNA silencing by miRNAs are also conserved in *Drosophila* [144,145].

Structural information about GW182 family proteins will be limited by the intrinsic disorder propensity of all the members of the group. In fact, more than 60% of the aminoacids in TNRC6 group are predicted to be in disordered and present flexible regions. Disordered segments are concentrated in the first 500 residues of the protein (Figure S5), which is in agreement with the enrichment of argonaute-binding motifs in this region that will facilitate transient interactions with these proteins and different ways to build silence complexes [139,140].

7.3. PIWI Proteins

PIWI proteins were described because of their intrinsic binding affinity for piRNAs. The PIWI family of proteins is a sub-family of the argonaute group. In flies, the PIWI family is composed of Piwi, Aubergine (AUB) and AGO3 proteins; in mice MILI, MIWI and MIWI2; and in humans of HILI, HIWI1, HIWI2 and HIWI3 [146,147]. PIWI family is required for efficient transposon silencing in germinal cell lines, however they are also produced in somatic cells [67,148]. All the family members are similar in sequence to the AGO group, harboring the four characteristic argonaute domains N-terminal, PAZ, MID and PIWI.

The role of individual PIWI proteins has been extensively studied in *Drosophila*. In this model organism it is well documented that every PIWI protein has a different binding specificity for sense or antisense RNAs derived from transposon regions in the genome; PIWI and AUB have affinity for antisense transcripts, whereas AGO3 is mainly bound to sense RNAs [149]. Because PIWI proteins have slicer activity, any of them can initiate the “ping-pong” amplification cycle for transposon silencing already described in previous sections.

Structural information available about PIWI proteins is still limited in humans to the X-ray crystallographic structure of the PAZ domain of MIWI and HIWI proteins in complex with a single stranded nucleic acid (PDB code: 2XFM). PAZ domains from PIWI proteins are extremely similar to those observed in AGO proteins and also in Dicer [150].

Interestingly, recent observations described the presence of post-translational modifications in PIWI proteins. Arginine methylation has been characterized as a modulation mechanism to produce specific

signatures of biological processes. In fact, several arginines present in the C-terminal domain of PIWI proteins are symmetrically dimethylated (sDMA) [148,149]. This arginine dimethylation could regulate the function of several proteins, including transcription factors and proteins belonging to the splicing machinery. In PIWI proteins, methylation of terminal arginines is catalyzed by PRMT5 methyltransferase. Methylated arginines are recognized by the Tudor family of proteins, classically linked to the gametogenesis process even before the discovery of PIWI proteins and piRNAs [151,152]. Tudor domains are protein modules that usually are mediating protein-protein interactions, potentially by binding to methylated aminoacids.

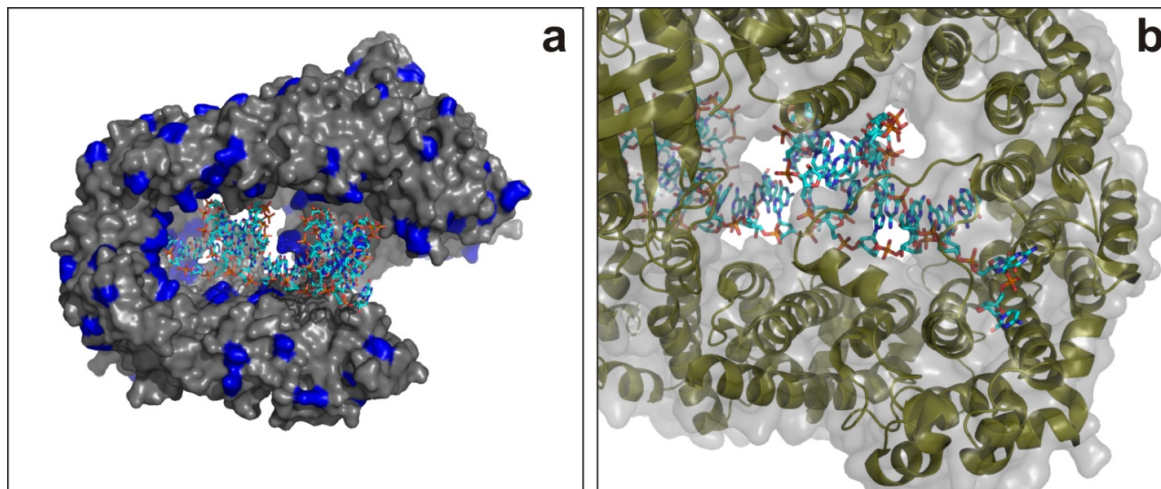
Recent published X-ray crystallography data, determined the molecular basis for the specificity of Tudor domain binding to methylated arginines in the C-terminal region of human MIWI protein [153]. Liu and coworkers determine that the specific recognition of dimethylated arginines by the Tudor domain is ensured by ionic interactions of a negatively charged groove in Tudor domain that recognize the positive charged arginine patch [152,153]. Additional data from X-ray crystallography experiments allowed also characterizing complexes between several additional proteins from the Tudor family. Chen and coworkers described the structure of the Tdrkh Tudor domain, and inferred its interaction mechanism with methylated arginines by mutagenesis analysis [154].

8. Other Helper Proteins: Exportin 5

Pre-miRNAs are actively exported from the nucleus to the cytoplasm to originate mature miRNAs. This transport is ensured by the Exp5-RanGTP system, firstly characterized by using *Xenopus* oocytes [155]. Exportin 5 is a ds-RNA binding protein that can translocate these molecules from the nucleus to the cytoplasm. This system is also capable of transporting other dsRNAs such as short-hairpin RNAs (shRNAs) [13,155,156]. RanGTPase-dependent export mediators (exportins) constitute the largest class of these carriers and are functionally highly versatile. As other exportins, Exp5 load its cognate pre-miRNA substrates in response to RanGTP binding in the nucleus and traverse the nuclear pores as ternary RanGTP-exportin-cargo complexes to the cytoplasm, where GTP hydrolysis leads to export complex disassembly [23,157,158].

The crystal structure of Exportin 5 in complex with a shRNA was determined by Okada *et al.* in 2009 [159]. The ternary Exp5-RanGTP-pre-miRNA studied by X-ray crystallography indicates that the interaction between pre-miRNA and exportin 5 is mainly driven through ionic contacts (Figure 7). A narrow pocket within the protein recognizes the two nucleotide 3'-overhang in the pre-miRNA allowing a specific interaction with the protein (Figure 7). Indeed, pre-miRNA blocking by Exp-5 in both its terminal nucleotides ensures enhanced protection for degradation against nucleases during transport.

Figure 7. Crystal structure of Exportin 5 complex with a pre-miRNA. **(a)** Overview of the dsRNA-Exp5 complex with the positive charged residues (lysines and arginines) colored in blue (PDB code: 3A6P). **(b)** Detailed view of the protein-RNA interface with the 2-nucleotide 3'-overhang located in a narrow pocket of the protein.



9. Conclusions

Structural Biology has contributed to the global understanding of the regulatory mechanisms and biogenesis routes of small non-coding RNAs. Experimental data showed a group of proteins involved in the biogenesis and function of small ncRNAs, that are mainly modular. However the diversity of domains found in this family of proteins is limited, including only a small number of typologies: dsRNA binding, RNase and helicase domains. These domains are core components of two different groups of enzymes: processors and effectors. Processor enzymes, with the exception of Drosha, are mainly well structured polypeptides. The group of the effectors is more diverse, and includes globular proteins but also polypeptides with long disordered regions that are involved in protein-protein interactions and consequently in complex stability.

Because of the inherent nature of the protein complexes involved in the regulatory functions of small non-coding RNAs, future studies will need to combine methodologies for the determination of tridimensional structures. Recent approaches joined together electron microscopy with X-ray crystallography and have been successfully applied to the characterization of RISC complexes. Further studies are needed to understand nuclear processing complexes and also to determine the role of protein-protein and protein-RNA transient interactions in the global landscape of non-coding RNA regulatory mechanisms.

Acknowledgments

M.C.C. was supported by a post-doctoral fellowship from Fundação para a Ciência e Tecnologia, Portugal (Ref. SFRH/BPD/65131/2009). F.J.E. would like to thank Francisco Enguita Jr. for his support, friendship and excellent technical advice during the preparation of the manuscript. The authors would also like to acknowledge to M. Carmo-Fonseca for fruitful discussions and criticisms essential for the overall quality improvement of the manuscript.

References

1. Maniatis, T.; Tasic, B. Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* **2002**, *418*, 236–243.
2. Hui, J.; Bindereif, A. Alternative pre-mRNA splicing in the human system: Unexpected role of repetitive sequences as regulatory elements. *Biol. Chem.* **2005**, *386*, 1265–1271.
3. Smith, C.W.; Valcarcel, J. Alternative pre-mRNA splicing: The logic of combinatorial control. *Trends Biochem. Sci.* **2000**, *25*, 381–388.
4. Boue, S.; Letunic, I.; Bork, P. Alternative splicing and evolution. *Bioessays* **2003**, *25*, 1031–1034.
5. Mehler, M.F.; Mattick, J.S. Non-coding RNAs in the nervous system. *J. Physiol.* **2006**, *575*, 333–341.
6. Louro, R.; Nakaya, H.I.; Amaral, P.P.; Festa, F.; Sogayar, M.C.; da Silva, A.M.; Verjovski-Almeida, S.; Reis, E.M. Androgen responsive intronic non-coding RNAs. *BMC Biol.* **2007**, *5*, doi:10.1186/1741-7007-5-4.
7. Tomaru, Y.; Hayashizaki, Y. Cancer research with non-coding RNA. *Cancer Sci.* **2006**, *97*, 1285–1290.
8. Chen, Z.; Zhang, J.; Kong, J.; Li, S.; Fu, Y.; Li, S.; Zhang, H.; Li, Y.; Zhu, Y. Diversity of endogenous small non-coding RNAs in *Oryza sativa*. *Genetica* **2006**, *128*, 21–31.
9. Mattick, J.S.; Makunin, I.V. Non-coding RNA. *Hum. Mol. Genet.* **2006**, *15*, R17–R29.
10. Missal, K.; Rose, D.; Stadler, P.F. Non-coding RNAs in *Ciona intestinalis*. *Bioinformatics* **2005**, *21 Suppl 2*, ii77–ii78.
11. Royo, H.; Bortolin, M.L.; Seitz, H.; Cavaille, J. Small non-coding RNAs and genomic imprinting. *Cytogenet. Genome Res.* **2006**, *113*, 99–108.
12. Eckstein, F. Small non-coding RNAs as magic bullets. *Trends Biochem. Sci.* **2005**, *30*, 445–452.
13. Murchison, E.P.; Hannon, G.J. miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr. Opin. Cell Biol.* **2004**, *16*, 223–229.
14. Faehnle, C.R.; Joshua-Tor, L. Argonautes confront new small RNAs. *Curr. Opin. Chem. Biol.* **2007**, *11*, 569–577.
15. Perron, M.P.; Provost, P. Protein components of the microRNA pathway and human diseases. *Methods Mol. Biol.* **2009**, *487*, 369–385.
16. Patel, D.J.; Ma, J.B.; Yuan, Y.R.; Ye, K.; Pei, Y.; Kuryavyi, V.; Malinina, L.; Meister, G.; Tuschl, T. Structural biology of RNA silencing and its functional implications. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 81–93.
17. Ji, X. The mechanism of RNase III action: How dicer dices. *Curr. Top. Microbiol. Immunol.* **2008**, *320*, 99–116.
18. Wan, Y.; Kertesz, M.; Spitale, R.C.; Segal, E.; Chang, H.Y. Understanding the transcriptome through RNA structure. *Nat. Rev. Genet.* **2011**, *12*, 641–655.
19. Davis-Dusenbery, B.N.; Hata, A. Mechanisms of control of microRNA biogenesis. *J. Biochem.* **2010**, *148*, 381–392.
20. Newman, M.A.; Hammond, S.M. Emerging paradigms of regulated microRNA processing. *Genes Dev.* **2010**, *24*, 1086–1092.

21. Lee, Y.; Han, J.; Yeom, K.H.; Jin, H.; Kim, V.N. Drosha in primary microRNA processing. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 51–57.
22. Han, J.; Lee, Y.; Yeom, K.H.; Kim, Y.K.; Jin, H.; Kim, V.N. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* **2004**, *18*, 3016–3027.
23. Bohnsack, M.T.; Czaplinski, K.; Gorlich, D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **2004**, *10*, 185–191.
24. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297.
25. Di Leva, G.; Calin, G.A.; Croce, C.M. MicroRNAs: Fundamental facts and involvement in human diseases. *Birth Defects Res. C Embryo. Today* **2006**, *78*, 180–189.
26. De, N.; Macrae, I.J. Purification and assembly of human Argonaute, Dicer, and TRBP complexes. *Methods Mol. Biol.* **2011**, *725*, 107–119.
27. Ye, X.; Huang, N.; Liu, Y.; Paroo, Z.; Huerta, C.; Li, P.; Chen, S.; Liu, Q.; Zhang, H. Structure of C3PO and mechanism of human RISC activation. *Nat. Struct. Mol. Biol.* **2011**, *18*, 650–657.
28. Krutzfeldt, J.; Stoffel, M. MicroRNAs: A new class of regulatory genes affecting metabolism. *Cell Metab.* **2006**, *4*, 9–12.
29. Maroney, P.A.; Yu, Y.; Nilsen, T.W. MicroRNAs, mRNAs, and translation. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 531–535.
30. Osada, H.; Takahashi, T. MicroRNAs in biological processes and carcinogenesis. *Carcinogenesis* **2007**, *28*, 2–12.
31. Jones-Rhoades, M.W.; Bartel, D.P.; Bartel, B. MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.* **2006**, *57*, 19–53.
32. Zhang, B.; Wang, Q.; Pan, X. MicroRNAs and their regulatory roles in animals and plants. *J. Cell. Physiol.* **2007**, *210*, 279–289.
33. Ruby, J.G.; Jan, C.H.; Bartel, D.P. Intronic microRNA precursors that bypass Drosha processing. *Nature* **2007**, *448*, 83–86.
34. Okamura, K.; Hagen, J.W.; Duan, H.; Tyler, D.M.; Lai, E.C. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **2007**, *130*, 89–100.
35. Kim, Y.K.; Kim, V.N. Processing of intronic microRNAs. *EMBO J.* **2007**, *26*, 775–783.
36. Lin, S.L.; Miller, J.D.; Ying, S.Y. Intronic microRNA (miRNA). *J. Biomed. Biotechnol.* **2006**, *2006*, doi:10.1155/JBB/2006/26818.
37. De Yebenes, V.G.; Belver, L.; Pisano, D.G.; Gonzalez, S.; Villasante, A.; Croce, C.; He, L.; Ramiro, A.R. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J. Exp. Med.* **2008**, *205*, 2199–2206.
38. Di Leva, G.; Gasparini, P.; Piovan, C.; Nganheu, A.; Garofalo, M.; Taccioli, C.; Iorio, M.V.; Li, M.; Volinia, S.; Alder, H.; *et al.* MicroRNA cluster 221–222 and estrogen receptor alpha interactions in breast cancer. *J. Natl. Cancer Inst.* **2010**, *102*, 706–721.
39. Georges, S.A.; Biery, M.C.; Kim, S.Y.; Schelter, J.M.; Guo, J.; Chang, A.N.; Jackson, A.L.; Carleton, M.O.; Linsley, P.S.; Cleary, M.A.; *et al.* Coordinated regulation of cell cycle transcripts by p53-inducible microRNAs, miR-192 and miR-215. *Cancer Res.* **2008**, *68*, 10105–10112.

40. Watanabe, T.; Totoki, Y.; Toyoda, A.; Kaneda, M.; Kuramochi-Miyagawa, S.; Obata, Y.; Chiba, H.; Kohara, Y.; Kono, T.; Nakano, T.; *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **2008**, *453*, 539–543.
41. Grivna, S.T.; Pyhtila, B.; Lin, H. MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13415–13420.
42. Grivna, S.T.; Beyret, E.; Wang, Z.; Lin, H. A novel class of small RNAs in mouse spermatogenic cells. *Genes Dev.* **2006**, *20*, 1709–1714.
43. Kirino, Y.; Mourelatos, Z. The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. *RNA* **2007**, *13*, 1397–1401.
44. Malone, C.D.; Brennecke, J.; Dus, M.; Stark, A.; McCombie, W.R.; Sachidanandam, R.; Hannon, G.J. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell* **2009**, *137*, 522–535.
45. Lau, N.C.; Robine, N.; Martin, R.; Chung, W.J.; Niki, Y.; Berezikov, E.; Lai, E.C. Abundant primary piRNAs, endo-siRNAs, and microRNAs in a *Drosophila* ovary cell line. *Genome Res.* **2009**, *19*, 1776–1785.
46. He, Z.; Kokkinaki, M.; Pant, D.; Gallicano, G.I.; Dym, M. Small RNA molecules in the regulation of spermatogenesis. *Reproduction* **2009**, *137*, 901–911.
47. Couvillion, M.T.; Lee, S.R.; Hogstad, B.; Malone, C.D.; Tonkin, L.A.; Sachidanandam, R.; Hannon, G.J.; Collins, K. Sequence, biogenesis, and function of diverse small RNA classes bound to the Piwi family proteins of *Tetrahymena thermophila*. *Genes Dev.* **2009**, *23*, 2016–2032.
48. Castaneda, J.; Genzor, P.; Bortvin, A. piRNAs, transposon silencing, and germline genome integrity. *Mutat. Res.* **2011**, *714*, 95–104.
49. Kawaoka, S.; Izumi, N.; Katsuma, S.; Tomari, Y. 3' end formation of PIWI-interacting RNAs *in vitro*. *Mol. Cell* **2011**, *43*, 1015–1022.
50. Siomi, M.C.; Sato, K.; Pezic, D.; Aravin, A.A. PIWI-interacting small RNAs: The vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 246–258.
51. Zhang, C. Novel functions for small RNA molecules. *Curr. Opin. Mol. Ther.* **2009**, *11*, 641–651.
52. Nishida, K.M.; Okada, T.N.; Kawamura, T.; Mituyama, T.; Kawamura, Y.; Inagaki, S.; Huang, H.; Chen, D.; Kodama, T.; Siomi, H.; Siomi, M.C. Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in *Drosophila* germlines. *EMBO J.* **2009**, *28*, 3820–3831.
53. Kim, V.N.; Han, J.; Siomi, M.C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 126–139.
54. Malone, C.D.; Hannon, G.J. Small RNAs as guardians of the genome. *Cell* **2009**, *136*, 656–668.
55. Grimson, A.; Srivastava, M.; Fahey, B.; Woodcroft, B.J.; Chiang, H.R.; King, N.; Degnan, B.M.; Rokhsar, D.S.; Bartel, D.P. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* **2008**, *455*, 1193–1197.
56. Hamilton, A.J.; Baulcombe, D.C. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **1999**, *286*, 950–952.

57. Sijen, T.; Steiner, F.A.; Thijssen, K.L.; Plasterk, R.H. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. *Science* **2007**, *315*, 244–247.
58. Samaha, H.; Delorme, V.; Pontvianne, F.; Cooke, R.; Delalande, F.; van Dorsselaer, A.; Echeverria, M.; Saez-Vasquez, J. Identification of protein factors and U3 snoRNAs from a *Brassica oleracea* RNP complex involved in the processing of pre-rRNA. *Plant J.* **2010**, *61*, 383–398.
59. Maniar, J.M.; Fire, A.Z. EGO-1, a *C. elegans* RdRP, modulates gene expression via production of mRNA-templated short antisense RNAs. *Curr. Biol.* **2011**, *21*, 449–459.
60. Tam, O.H.; Aravin, A.A.; Stein, P.; Girard, A.; Murchison, E.P.; Cheloufi, S.; Hodges, E.; Anger, M.; Sachidanandam, R.; Schultz, R.M.; Hannon, G.J. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **2008**, *453*, 534–538.
61. Correa, R.L.; Steiner, F.A.; Berezikov, E.; Ketting, R.F. MicroRNA-directed siRNA biogenesis in *Caenorhabditis elegans*. *PLoS Genet.* **2010**, *6*, doi:10.1371/journal.pgen.1000903.
62. Wang, H.; Zhang, X.; Liu, J.; Kiba, T.; Woo, J.; Ojo, T.; Hafner, M.; Tuschl, T.; Chua, N.H.; Wang, X.J. Deep sequencing of small RNAs specifically associated with *Arabidopsis* AGO1 and AGO4 uncovers new AGO functions. *Plant J.* **2011**, *67*, 292–304.
63. Chellappan, P.; Xia, J.; Zhou, X.; Gao, S.; Zhang, X.; Coutino, G.; Vazquez, F.; Zhang, W.; Jin, H. siRNAs from miRNA sites mediate DNA methylation of target genes. *Nucleic Acids Res.* **2010**, *38*, 6883–6894.
64. Chen, X. Small RNAs in development—Insights from plants. *Curr. Opin. Genet. Dev.* **2012**, in press.
65. Soifer, H.S.; Zaragoza, A.; Peyvan, M.; Behlke, M.A.; Rossi, J.J. A potential role for RNA interference in controlling the activity of the human LINE-1 retrotransposon. *Nucleic Acids Res.* **2005**, *33*, 846–856.
66. Apornthewan, C.; Phokaew, C.; Piriyaopongsa, J.; Ngamphiw, C.; Ittiwut, C.; Tongsimma, S.; Mutirangura, A. Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2. *PLoS One* **2011**, *6*, doi:10.1371/journal.pone.0017934.
67. Qi, H.; Watanabe, T.; Ku, H.Y.; Liu, N.; Zhong, M.; Lin, H. The Yb body, a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells. *J. Biol. Chem.* **2011**, *286*, 3789–3797.
68. Costa, F.F. Non-coding RNAs, epigenetics and complexity. *Gene* **2008**, *410*, 9–17.
69. Nowotny, M.; Yang, W. Structural and functional modules in RNA interference. *Curr. Opin. Struct. Biol.* **2009**, *19*, 286–293.
70. Lau, P.W.; Guiley, K.Z.; De, N.; Potter, C.S.; Carragher, B.; MacRae, I.J. The molecular architecture of human Dicer. *Nat. Struct. Mol. Biol.* **2012**, *19*, 436–440.
71. Perron, M.P.; Provost, P. Protein interactions and complexes in human microRNA biogenesis and function. *Front Biosci.* **2008**, *13*, 2537–2547.
72. Gregory, R.I.; Yan, K.P.; Amuthan, G.; Chendrimada, T.; Doratotaj, B.; Cooch, N.; Shiekhattar, R. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **2004**, *432*, 235–240.
73. Gregory, R.I.; Chendrimada, T.P.; Shiekhattar, R. MicroRNA biogenesis: Isolation and characterization of the microprocessor complex. *Methods Mol. Biol.* **2006**, *342*, 33–47.

74. Triboulet, R.; Gregory, R.I. Autoregulatory mechanisms controlling the microprocessor. *Adv. Exp. Med. Biol.* **2010**, *700*, 56–66.
75. Xie, Z.; Kasschau, K.D.; Carrington, J.C. Negative feedback regulation of Dicer-Like1 in *Arabidopsis* by microRNA-guided mRNA degradation. *Curr. Biol.* **2003**, *13*, 784–789.
76. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Radmark, O.; Kim, S.; Kim, V.N. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **2003**, *425*, 415–419.
77. Denli, A.M.; Tops, B.B.; Plasterk, R.H.; Ketting, R.F.; Hannon, G.J. Processing of primary microRNAs by the microprocessor complex. *Nature* **2004**, *432*, 231–235.
78. de la Chapelle, A.; Herva, R.; Koivisto, M.; Aula, P. A deletion in chromosome 22 can cause DiGeorge syndrome. *Hum. Genet.* **1981**, *57*, 253–256.
79. Kelley, R.I.; Zackai, E.H.; Emanuel, B.S.; Kistenmacher, M.; Greenberg, F.; Punnett, H.H. The association of the DiGeorge anomalad with partial monosomy of chromosome 22. *J. Pediatr.* **1982**, *101*, 197–200.
80. Mueller, G.A.; Miller, M.T.; Derose, E.F.; Ghosh, M.; London, R.E.; Hall, T.M. Solution structure of the Drosha double-stranded RNA-binding domain. *Silence* **2010**, *1*, doi:10.1186/1758-907X-1-2.
81. Senturia, R.; Faller, M.; Yin, S.; Loo, J.A.; Cascio, D.; Sawaya, M.R.; Hwang, D.; Clubb, R.T.; Guo, F. Structure of the dimerization domain of DiGeorge critical region 8. *Protein Sci.* **2010**, *19*, 1354–1365.
82. Sohn, S.Y.; Bae, W.J.; Kim, J.J.; Yeom, K.H.; Kim, V.N.; Cho, Y. Crystal structure of human DGCR8 core. *Nat. Struct. Mol. Biol.* **2007**, *14*, 847–853.
83. Faller, M.; Matsunaga, M.; Yin, S.; Loo, J.A.; Guo, F. Heme is involved in microRNA processing. *Nat. Struct. Mol. Biol.* **2007**, *14*, 23–29.
84. Barr, I.; Smith, A.T.; Senturia, R.; Chen, Y.; Scheidemantle, B.D.; Burstyn, J.N.; Guo, F. DiGeorge critical region 8 (DGCR8) is a double-cysteine-ligated heme protein. *J. Biol. Chem.* **2011**, *286*, 16716–16725.
85. Barr, I.; Smith, A.T.; Chen, Y.; Senturia, R.; Burstyn, J.N.; Guo, F. Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1919–1924.
86. Zhang, H.; Kolb, F.A.; Brondani, V.; Billy, E.; Filipowicz, W. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.* **2002**, *21*, 5875–5885.
87. Pellino, J.L.; Jaskiewicz, L.; Filipowicz, W.; Sontheimer, E.J. ATP modulates siRNA interactions with an endogenous human Dicer complex. *RNA* **2005**, *11*, 1719–1724.
88. Hammond, S.M. Dicing and slicing: The core machinery of the RNA interference pathway. *FEBS Lett.* **2005**, *579*, 5822–5829.
89. Macrae, I.J.; Li, F.; Zhou, K.; Cande, W.Z.; Doudna, J.A. Structure of Dicer and mechanistic implications for RNAi. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 73–80.
90. Macrae, I.J.; Zhou, K.; Li, F.; Repic, A.; Brooks, A.N.; Cande, W.Z.; Adams, P.D.; Doudna, J.A. Structural basis for double-stranded RNA processing by Dicer. *Science* **2006**, *311*, 195–198.
91. MacRae, I.J.; Zhou, K.; Doudna, J.A. Structural determinants of RNA recognition and cleavage by Dicer. *Nat. Struct. Mol. Biol.* **2007**, *14*, 934–940.

92. Cook, A.; Conti, E. Dicer measures up. *Nat. Struct. Mol. Biol.* **2006**, *13*, 190–192.
93. Welker, N.C.; Maity, T.S.; Ye, X.; Aruscavage, P.J.; Krauchuk, A.A.; Liu, Q.; Bass, B.L. Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. *Mol. Cell* **2011**, *41*, 589–599.
94. Park, J.E.; Heo, I.; Tian, Y.; Simanshu, D.K.; Chang, H.; Jee, D.; Patel, D.J.; Kim, V.N. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* **2011**, *475*, 201–205.
95. Du, Z.; Lee, J.K.; Tjhen, R.; Stroud, R.M.; James, T.L. Structural and biochemical insights into the dicing mechanism of mouse Dicer: A conserved lysine is critical for dsRNA cleavage. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 2391–2396.
96. Makarova, K.S.; Wolf, Y.I.; van der Oost, J.; Koonin, E.V. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct* **2009**, *4*, doi:10.1186/1745-6150-4-29.
97. Ekwall, K. The RITS complex-A direct link between small RNA and heterochromatin. *Mol. Cell* **2004**, *13*, 304–305.
98. Verdel, A.; Vavasseur, A.; Le Gorrec, M.; Touat-Todeschini, L. Common themes in siRNA-mediated epigenetic silencing pathways. *Int. J. Dev. Biol.* **2009**, *53*, 245–257.
99. Miyoshi, K.; Tsukumo, H.; Nagami, T.; Siomi, H.; Siomi, M.C. Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev.* **2005**, *19*, 2837–2848.
100. Verdel, A.; Jia, S.; Gerber, S.; Sugiyama, T.; Gygi, S.; Grewal, S.I.; Moazed, D. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **2004**, *303*, 672–676.
101. Mescalchin, A.; Detzer, A.; Weirauch, U.; Hahnel, M.J.; Engel, C.; Sczakiel, G. Antisense tools for functional studies of human Argonaute proteins. *RNA* **2010**, *16*, 2529–2536.
102. Azuma-Mukai, A.; Oguri, H.; Mituyama, T.; Qian, Z.R.; Asai, K.; Siomi, H.; Siomi, M.C. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7964–7969.
103. Wang, D.; Zhang, Z.; O'Loughlin, E.; Lee, T.; Houel, S.; O'Carroll, D.; Tarakhovskiy, A.; Ahn, N.G.; Yi, R. Quantitative functions of Argonaute proteins in mammalian development. *Genes Dev.* **2012**, *26*, 693–704.
104. Song, J.J.; Smith, S.K.; Hannon, G.J.; Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **2004**, *305*, 1434–1437.
105. Wang, Y.; Juranek, S.; Li, H.; Sheng, G.; Tuschl, T.; Patel, D.J. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **2008**, *456*, 921–926.
106. Wang, Y.; Sheng, G.; Juranek, S.; Tuschl, T.; Patel, D.J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **2008**, *456*, 209–213.
107. Yuan, Y.R.; Pei, Y.; Ma, J.B.; Kuryavyy, V.; Zhadina, M.; Meister, G.; Chen, H.Y.; Dauter, Z.; Tuschl, T.; Patel, D.J. Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell* **2005**, *19*, 405–419.
108. Song, J.J.; Liu, J.; Tolia, N.H.; Schneiderman, J.; Smith, S.K.; Martienssen, R.A.; Hannon, G.J.; Joshua-Tor, L. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat. Struct. Biol.* **2003**, *10*, 1026–1032.

109. Schirle, N.T.; Macrae, I.J. The crystal structure of human Argonaute2. *Science* **2012**, *336*, 1037–1040.
110. Boland, A.; Huntzinger, E.; Schmidt, S.; Izaurralde, E.; Weichenrieder, O. Crystal structure of the MID-PIWI lobe of a eukaryotic Argonaute protein. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10466–10471.
111. Wang, Y.; Juranek, S.; Li, H.; Sheng, G.; Wardle, G.S.; Tuschl, T.; Patel, D.J. Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **2009**, *461*, 754–761.
112. Lima, W.F.; Wu, H.; Nichols, J.G.; Sun, H.; Murray, H.M.; Crooke, S.T. Binding and cleavage specificities of human Argonaute2. *J. Biol. Chem.* **2009**, *284*, 26017–26028.
113. Zeng, Y.; Sankala, H.; Zhang, X.; Graves, P.R. Phosphorylation of Argonaute 2 at serine-387 facilitates its localization to processing bodies. *Biochem. J.* **2008**, *413*, 429–436.
114. Ma, J.B.; Ye, K.; Patel, D.J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **2004**, *429*, 318–322.
115. Maiti, R.; van Domselaar, G.H.; Zhang, H.; Wishart, D.S. SuperPose: A simple server for sophisticated structural superposition. *Nucleic Acids Res.* **2004**, *32*, W590–W594.
116. Wang, H.W.; Noland, C.; Siridechadilok, B.; Taylor, D.W.; Ma, E.; Felderer, K.; Doudna, J.A.; Nogales, E. Structural insights into RNA processing by the human RISC-loading complex. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1148–1153.
117. Frank, F.; Sonenberg, N.; Nagar, B. Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature* **2010**, *465*, 818–822.
118. Takimoto, K.; Wakiyama, M.; Yokoyama, S. Mammalian GW182 contains multiple Argonaute-binding sites and functions in microRNA-mediated translational repression. *RNA* **2009**, *15*, 1078–1089.
119. Wang, B.; Li, S.; Qi, H.H.; Chowdhury, D.; Shi, Y.; Novina, C.D. Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1259–1266.
120. Chu, Y.; Yue, X.; Younger, S.T.; Janowski, B.A.; Corey, D.R. Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the progesterone receptor promoter. *Nucleic Acids Res.* **2010**, *38*, 7736–7748.
121. Zilberman, D.; Cao, X.; Jacobsen, S.E. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **2003**, *299*, 716–719.
122. Zilberman, D.; Cao, X.; Johansen, L.K.; Xie, Z.; Carrington, J.C.; Jacobsen, S.E. Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr. Biol.* **2004**, *14*, 1214–1220.
123. Rowley, M.J.; Avrutsky, M.I.; Sifuentes, C.J.; Pereira, L.; Wierzbicki, A.T. Independent chromatin binding of ARGONAUTE4 and SPT5L/KTF1 mediates transcriptional gene silencing. *PLoS Genet.* **2011**, *7*, doi:10.1371/journal.pgen.1002120.
124. Tran, R.K.; Zilberman, D.; de Bustos, C.; Ditt, R.F.; Henikoff, J.G.; Lindroth, A.M.; Delrow, J.; Boyle, T.; Kwong, S.; Bryson, T.D.; Jacobsen, S.E.; Henikoff, S. Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in *Arabidopsis*. *Genome Biol.* **2005**, *6*, doi:10.1186/gb-2005-6-11-r90.

125. Broderick, J.A.; Salomon, W.E.; Ryder, S.P.; Aronin, N.; Zamore, P.D. Argonaute protein identity and pairing geometry determine cooperativity in mammalian RNA silencing. *RNA* **2011**, *17*, 1858–1869.
126. Bennett-Lovsey, R.M.; Herbert, A.D.; Sternberg, M.J.; Kelley, L.A. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins* **2008**, *70*, 611–625.
127. Kelley, L.A.; Sternberg, M.J. Protein structure prediction on the Web: A case study using the Phyre server. *Nat. Protoc.* **2009**, *4*, 363–371.
128. Parker, J.S.; Roe, S.M.; Barford, D. Molecular mechanism of target RNA transcript recognition by Argonaute-guide complexes. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 45–50.
129. Reddy, T.R.; Suhasini, M.; Rappaport, J.; Looney, D.J.; Kraus, G.; Wong-Staal, F. Molecular cloning and characterization of a TAR-binding nuclear factor from T cells. *AIDS Res. Hum. Retrovir.* **1995**, *11*, 663–669.
130. Gagnon, A.; Buckler-White, A.; Berkhout, B.; Jeang, K.T. Characterization of a human TAR RNA-binding protein that activates the HIV-1 LTR. *Science* **1991**, *251*, 1597–1600.
131. Chendrimada, T.P.; Gregory, R.I.; Kumaraswamy, E.; Norman, J.; Cooch, N.; Nishikura, K.; Shiekhattar, R. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **2005**, *436*, 740–744.
132. MacRae, I.J.; Ma, E.; Zhou, M.; Robinson, C.V.; Doudna, J.A. *In vitro* reconstitution of the human RISC-loading complex. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 512–517.
133. Gredell, J.A.; Dittmer, M.J.; Wu, M.; Chan, C.; Walton, S.P. Recognition of siRNA asymmetry by TAR RNA binding protein. *Biochemistry* **2010**, *49*, 3148–3155.
134. Kalidas, S.; Sanders, C.; Ye, X.; Strauss, T.; Kuhn, M.; Liu, Q.; Smith, D.P. *Drosophila* R2D2 mediates follicle formation in somatic tissues through interactions with Dicer-1. *Mech. Dev.* **2008**, *125*, 475–485.
135. Murphy, D.; Dancis, B.; Brown, J.R. The evolution of core proteins involved in microRNA biogenesis. *BMC Evol. Biol.* **2008**, *8*, doi:10.1186/1471-2148-8-92.
136. Yang, S.W.; Chen, H.Y.; Yang, J.; Machida, S.; Chua, N.H.; Yuan, Y.A. Structure of *Arabidopsis* HYPONASTIC LEAVES1 and its molecular implications for miRNA processing. *Structure* **2010**, *18*, 594–605.
137. Lau, P.W.; Potter, C.S.; Carragher, B.; MacRae, I.J. Structure of the human Dicer-TRBP complex by electron microscopy. *Structure* **2009**, *17*, 1326–1332.
138. Chekulaeva, M.; Parker, R.; Filipowicz, W. The GW/WG repeats of *Drosophila* GW182 function as effector motifs for miRNA-mediated repression. *Nucleic Acids Res.* **2010**, *38*, 6673–6683.
139. Lazzaretti, D.; Tournier, I.; Izaurralde, E. The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins. *RNA* **2009**, *15*, 1059–1066.
140. Zipprich, J.T.; Bhattacharyya, S.; Mathys, H.; Filipowicz, W. Importance of the C-terminal domain of the human GW182 protein TNRC6C for translational repression. *RNA* **2009**, *15*, 781–793.

141. Fabian, M.R.; Mathonnet, G.; Sundermeier, T.; Mathys, H.; Zipprich, J.T.; Svitkin, Y.V.; Rivas, F.; Jinek, M.; Wohlschlegel, J.; Doudna, J.A.; *et al.* Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Mol. Cell* **2009**, *35*, 868–880.
142. Jinek, M.; Fabian, M.R.; Coyle, S.M.; Sonenberg, N.; Doudna, J.A. Structural insights into the human GW182-PABC interaction in microRNA-mediated deadenylation. *Nat. Struct. Mol. Biol.* **2010**, *17*, 238–240.
143. Piao, X.; Zhang, X.; Wu, L.; Belasco, J.G. CCR4-NOT deadenylates mRNA associated with RNA-induced silencing complexes in human cells. *Mol. Cell Biol.* **2010**, *30*, 1486–1494.
144. Braun, J.E.; Huntzinger, E.; Fauser, M.; Izaurralde, E. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **2011**, *44*, 120–133.
145. Chekulaeva, M.; Mathys, H.; Zipprich, J.T.; Attig, J.; Colic, M.; Parker, R.; Filipowicz, W. miRNA repression involves GW182-mediated recruitment of CCR4-NOT through conserved W-containing motifs. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1218–1226.
146. Unhavaithaya, Y.; Hao, Y.; Beyret, E.; Yin, H.; Kuramochi-Miyagawa, S.; Nakano, T.; Lin, H. MILI, a PIWI-interacting RNA-binding protein, is required for germ line stem cell self-renewal and appears to positively regulate translation. *J. Biol. Chem.* **2009**, *284*, 6507–6519.
147. Siddiqi, S.; Matushansky, I. Piwis and piwi-interacting RNAs in the epigenetics of cancer. *J. Cell Biochem.* **2012**, *113*, 373–380.
148. Juliano, C.; Wang, J.; Lin, H. Uniting germline and stem cells: The function of PIWI proteins and the piRNA pathway in diverse organisms. *Annu. Rev. Genet.* **2011**, *45*, 447–469.
149. Brennecke, J.; Aravin, A.A.; Stark, A.; Dus, M.; Kellis, M.; Sachidanandam, R.; Hannon, G.J. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **2007**, *128*, 1089–1103.
150. Simon, B.; Kirkpatrick, J.P.; Eckhardt, S.; Reuter, M.; Rocha, E.A.; Andrade-Navarro, M.A.; Sehr, P.; Pillai, R.S.; Carlomagno, T. Recognition of 2'-O-methylated 3'-end of piRNA by the PAZ domain of a Piwi protein. *Structure* **2011**, *19*, 172–180.
151. Van der Heijden, G.W.; Bortvin, A. Defending the genome in tudor style. *Dev. Cell* **2009**, *17*, 745–746.
152. Liu, L.; Qi, H.; Wang, J.; Lin, H. PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition. *Development* **2011**, *138*, 1863–1873.
153. Liu, K.; Chen, C.; Guo, Y.; Lam, R.; Bian, C.; Xu, C.; Zhao, D.Y.; Jin, J.; MacKenzie, F.; Pawson, T.; Min, J. Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18398–18403.
154. Chen, C.; Jin, J.; James, D.A.; Adams-Cioaba, M.A.; Park, J.G.; Guo, Y.; Tenaglia, E.; Xu, C.; Gish, G.; Min, J.; Pawson, T. Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 20336–20341.
155. Yi, R.; Qin, Y.; Macara, I.G.; Cullen, B.R. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* **2003**, *17*, 3011–3016.
156. Zeng, Y.; Cullen, B.R. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res.* **2004**, *32*, 4776–4785.

157. Kim, V.N. MicroRNA precursors in motion: Exportin-5 mediates their nuclear export. *Trends Cell Biol.* **2004**, *14*, 156–159.
158. Lund, E.; Dahlberg, J.E. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb. Symp. Quant. Biol.* **2006**, *71*, 59–66.
159. Okada, C.; Yamashita, E.; Lee, S.J.; Shibata, S.; Katahira, J.; Nakagawa, A.; Yoneda, Y.; Tsukihara, T. A high-resolution structure of the pre-microRNA nuclear export machinery. *Science* **2009**, *326*, 1275–1279.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).