

Review

A Prokaryotic Twist on Argonaute Function

Sarah Willkomm ¹, Adrian Zander ², Alexander Gust ² and Dina Grohmann ^{2,*}

¹ Institute of Molecular Medicine, Universitätsklinikum Schleswig-Holstein, University of Lübeck, 23538 Lübeck, Germany; E-Mail: willkomm@imm.uni-luebeck.de

² Physikalische und Theoretische Chemie-NanoBioSciences, Technische Universität Braunschweig, Hans-Sommer-Strasse 10, 38106 Braunschweig, Germany; E-Mails: a.zander@tu-bs.de (A.Z.); a.gust@tu-bs.de (A.G.)

* Author to whom correspondence should be addressed; E-Mail: d.grohmann@tu-braunschweig.de; Tel.: +49-531-391-5395.

Academic Editors: Hans-Peter Klenk, Michael W. W. Adams and Roger A. Garrett

Received: 21 December 2014 / Accepted: 5 February 2015 / Published: 16 February 2015

Abstract: Argonaute proteins can be found in all three domains of life. In eukaryotic organisms, Argonaute is, as the functional core of the RNA-silencing machinery, critically involved in the regulation of gene expression. Despite the mechanistic and structural similarities between archaeal, bacterial and eukaryotic Argonaute proteins, the biological function of bacterial and archaeal Argonautes has remained elusive. This review discusses new findings in the field that shed light on the structure and function of Argonaute. We especially focus on archaeal Argonautes when discussing the details of the structural and dynamic features in Argonaute that promote substrate recognition and cleavage, thereby revealing differences and similarities in Argonaute biology.

Keywords: Argonaute; archaeal; interference, guide strand

1. Introduction

The Argonaute (Ago) protein family was initially discovered in eukaryotes [1,2], but orthologs were found in many archaeal and bacterial organisms [3–5]. In eukaryotic organisms, Argonaute represents the principal component of the RNA silencing machinery. Despite the advancements in the understanding of Argonaute function in the eukaryotic field, the biological role of prokaryotic Argonaute

proteins (pAgo) remained unknown for a long time. Argonaute proteins are encoded in ~32% and 9% of the sequenced archaeal and bacterial genomes, respectively [6]. PAgos were found to cluster in two groups distinguished by the presence or absence of the PAZ domain [5]. A lack of the PAZ domain often coincides with an apparent inactivation of the nuclease activity [5,6]. Interestingly, pAgos are often found in operons with a diverse range of endonucleolytic DNases (nucleases of the restriction endonuclease fold, a distinctive Sirtuin family domain or TIR domain proteins) and/or helicases (e.g., of the DinG-class) [3,5,7], leading to the hypothesis that the co-action of pAgo and an endo-DNase might act as a plasmid/phage restriction system. However, the subset of pAgos that exhibit a high sequence similarity to their eukaryotic counterpart does not seem to show conserved operonic associations with any other genes. Ago is composed of the N-terminal, PAZ (Piwi-Argonaute-Zwille), middle (MID) and PIWI (P-element-induced wimpy testis) domains interconnected by two structured linker regions (Figures 1 and 2). This review mainly discusses the “long” pAgo variants, which share a comparable domain organization as determined for the eukaryotic Agos. In contrast, short pAgos variants contain the MID and PIWI domain only [5]. Recently, two studies have shed light on the biological role of bacterial Agos [8,9]. Together with our findings on the substrate specificity of an archaeal Ago variant [10], these data point to a paradigm shift in the field, as the spectrum of Argonaute silencing activities now also includes DNA- or RNA-guided DNA interference in prokaryotic organisms.

2. Structural Organization of Argonaute

The structures of archaeal [11,12] and bacterial [13–15] Agos provided valuable insights into the structure-function relationship of the protein before structures of eukaryotic Ago became available [16–18]. The structure of Argonaute from the archaeal organism *Pyrococcus furiosus* (PfAgo) was the first reported full-length structure [12], quickly followed by the structures of bacterial and, finally, eukaryotic Argonaute variants. Interestingly, the Argonaute structures from all three domains of life show a high degree of similarity (Figure 1). The domains of the protein are arranged in a bilobal fashion with the PAZ and N-terminal domain forming one and PIWI and MID domains the other lobe. The 5'-end of the guide strand is buried in a deep pocket at the interface between the MID and PIWI domains in a highly conserved region where side chains of four invariant residues contact the obligatory 5'-end phosphate group of the first nucleotide. These four residues are highly conserved among archaeal and eukaryotic Agos. In the *A. fulgidus* PIWI (AfPIWI) protein, these residues are Y123, K127, Q137 and K163 [19], whereas in human Ago2 (hAgo2), the terminal phosphate is stabilized by Y529, K533, Q545 and K570 [17,20,21]. The *T. thermophilus* Ago (TtAgo), in contrast, uses different residues, which are R418, K422, S432, Q433 and K457 [15]. The insertion of the 5' nucleotide into the binding pocket of TtAgo leads to a distortion of the RNA/DNA duplex, which results in an unwinding of the first base pair of the loaded nucleic acid duplex [15,19,22]. Interestingly, there are no extensive base-specific contacts between amino acid side-chain residues and the bases. However, structures of TtAgo with a 21 mer guide DNA with complementary 12-mer target RNA or target DNA show that the terminal 5'-end nucleotide of the let-7 guide (a thymine) undergoes hydrogen-bonding contacts with the backbone amide carbonyl of residue M413 and the side chain of N436 [14,23]. For the eukaryotic Argonaute, a nucleotide specificity loop contributes to the recognition of the first nucleotide (Figure 2a). The specificity loop can be found in archaeal and bacterial Agos, but not all residues in the loop are conserved [21]. The loop is pulled

away from the nucleotide of the base in the AfPIWI structure [24], while AMP and UMP interact with the backbone atoms of the hAgo2 specificity loop (G524 and T526). In contrast, GMP and CMP are repulsed by the loop [21]. The structures of AfPIWI and TtAgo show that a divalent metal ion is coordinated with the C-terminus of Argonaute. This metal ion is involved in 5'-end binding [15,22], while eukaryotic Argonautes do not make use of a metal ion and neutralize the charge by a lysine side chain [16]. The electron density of the guide strand can be detected easily for nucleotides 2 to 7 or 8 (the “seed” region [13,18,20,25,26]). Here, amino acids located in the MID, PIWI domain (R792, K709, Y804, S798, R761) and L1 linker (A221) of hAgo2 contact the guide strand via salt linkages to the phosphate backbone and hydrogen bonding [20]. The nucleotides in the seed region are continuously base-stacked and solvent-exposed, while the nucleotides beyond nucleotide 7 are threaded into Ago [17,18]. None of the contacts establish any sequence specificity congruent with Argonaute’s ability to bind a multitude of different guide sequences.

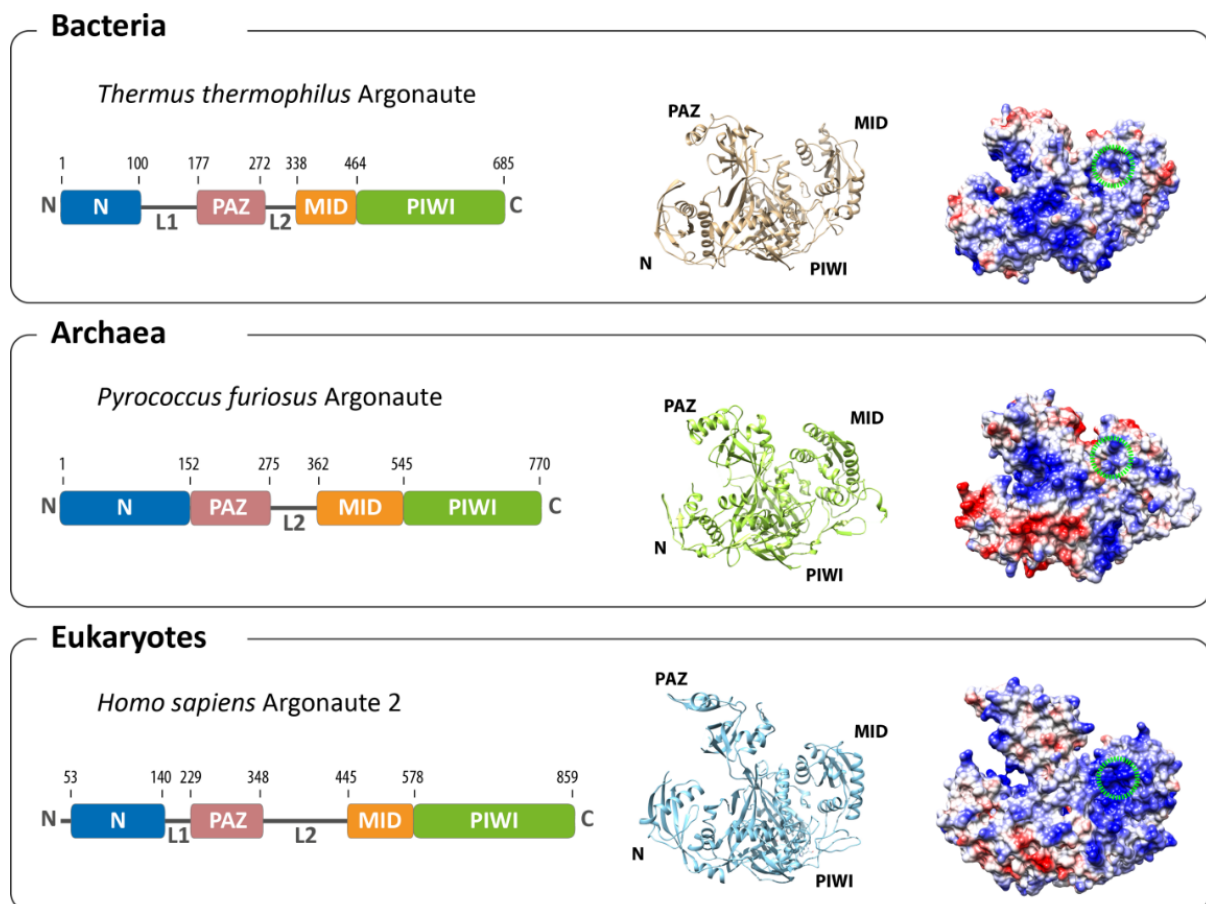


Figure 1. Overall architecture of Argonaute from the three domains of life. The domain composition (left) and structures (middle) of the bacterial (based on *Thermus thermophilus*, PDB: 3DLH), the archaeal (based on *Pyrococcus furiosus*, PDB: 1U04) and the eukaryotic (based on human Argonaute 2, PDB: 4E13) Argonaute reveal an evolutionarily conserved architecture. Differences can be found in the surface charge distribution of Argonaute proteins (negatively-charged surfaces in red; positively-charged surfaces in blue). The binding pocket for the 5'-end of the guide in the MID domain is highlighted with a green circle.

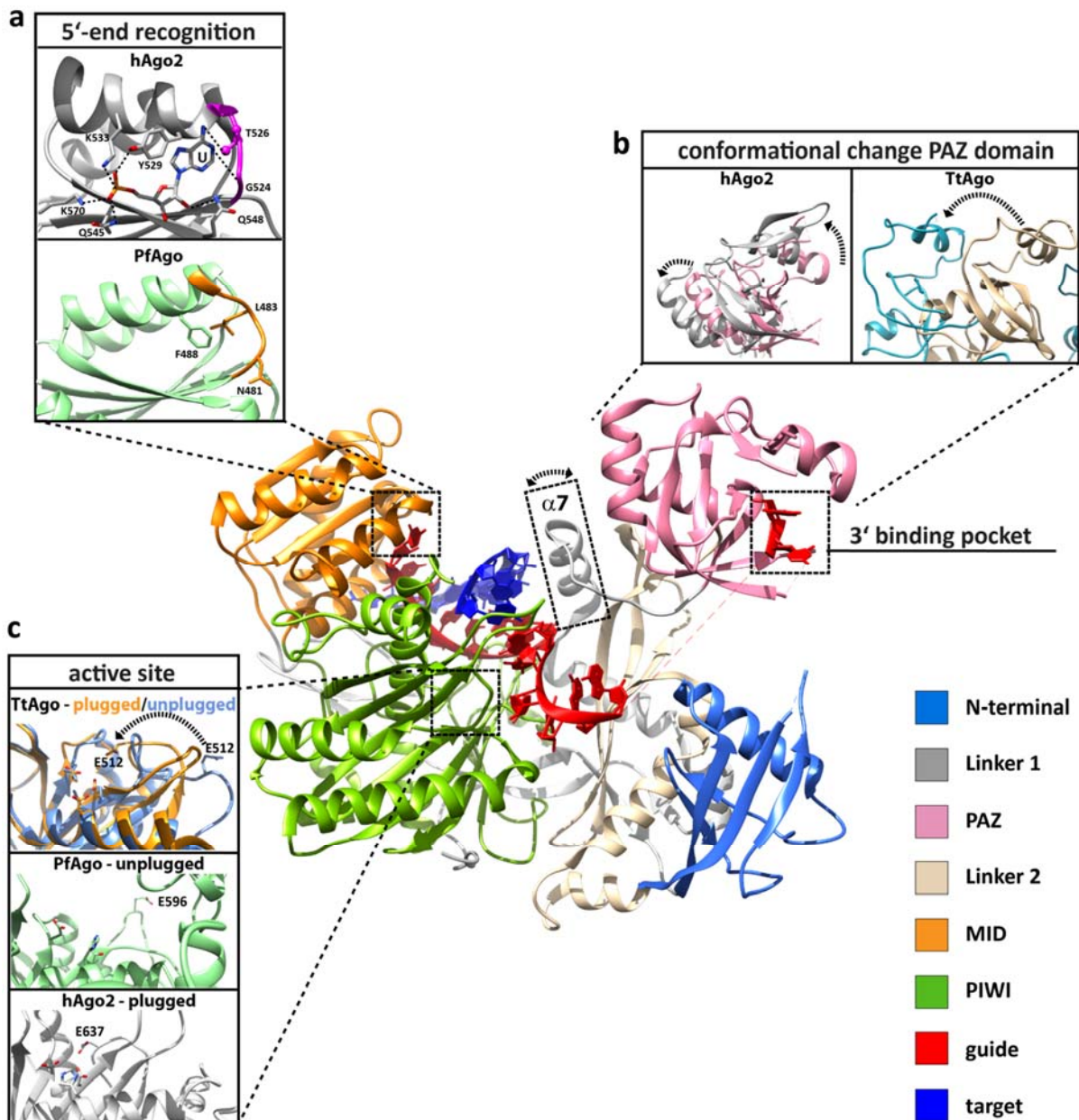


Figure 2. Important structural, functional and dynamic features of Argonaute. Structural elements that are important for Argonaute function are highlighted based on the human Argonaute 2 (hAgo2) structure in complex with a guide (red) and target (blue) strand (PDB: 4W5T). **(a)** The 5'-end is buried in a binding pocket in the MID domain (orange), where specific interactions with the terminal phosphate of the guide strand and interactions between the protein backbone of the specificity loop (highlighted in purple or orange) contribute to the specific recognition of the first nucleotide (PDB: 3LUD). This interaction network leads to the stable positioning of UTP in hAgo2. In contrast, the Argonaute structure from *Pyrococcus furiosus* (PfAgo) shows that the specificity loop (orange) is pulled away from the first nucleotide (PDB: 1U04). **(b)** The PAZ domain (pink) of all Argonaute variants is a mobile element, as revealed by structural, kinetic and single-molecule studies. Shown are the conformational changes (highlighted by a broken arrow) of the PAZ domain between the RNA guide-associated hAgo2 (pink, PDB: 4EI3) and hAgo2 in complex with an RNA guide and an 11-nucleotide RNA target (grey, PDB: 4W5T). The movement of the PAZ domain is

more pronounced when comparing the structure of DNA guide-associated *Thermus thermophilus* Ago (TtAgo, PDB: 3DLH) and the ternary TtAgo complex, which also includes a 19-nucleotide RNA target (PDB: 3HVR). Progression to the ternary complex leads to the release of the 3'-end of the guide from its binding pocket in the PAZ domain. Another flexible element that undergoes a structural change upon ternary complex formation is helix $\alpha 7$ (boxed), which is only found in archaeal-eukaryotic Argonautes. (c) The PIWI domain (green) harbors the active site where the glutamate finger can be found in an “unplugged” or “plugged” conformation (PfAgo in its free state (mint green) with the “unplugged” glutamate finger, PDB: 1U04; cleavage-incompatible ternary TtAgo complex with “unplugged” glutamate finger (PDB: 3F73, corn blue); cleavage-compatible ternary TtAgo complex with “plugged” glutamate finger (PDB: 3DLH, orange); ternary hAgo2 complex with “plugged” glutamate finger (PDB: 4W5T, grey). In the “plugged” conformation, an invariant glutamate sidechain is inserted to complete the tetrad in the catalytic pocket (the broken arrow indicates the relocation of E512).

Beyond the seed region, the structure of the guide strand appears to be largely disordered, with the exception of the guide's 3'-end. The structures of the TtAgo and hAgo2 binary complex showed that both proteins introduce a pronounced kink in the guide strand. Kinking after nucleotide 6 at the end of the seed region in hAgo2 is caused by I365 [14,16,17,27]. In TtAgo, a similar kink is positioned at nucleotide 10 of the guide [15]. The last two nucleotides of the guide strand are contacted by conserved aromatic and basic residues anchoring the 3'-end in the PAZ domain [15–17,19,20,28,29] (Figure 2b). There is no structural information of the archaeal Argonaute protein from *Methanocaldococcus jannaschii* (MjAgo) provide evidence for a 3'-end anchoring of the guide strand in the archaeal domain [10]. Kinetic studies revealed that the binding of hAgo2 to a guide strand follows a consecutive pathway. The association occurs in a three-phase process. The first phase, which is only limited by diffusion, represents a collision between hAgo2 and a guide strand. The following phase corresponds to the subsequent anchoring of the guide's 5'-end in the MID domain, which is a pre-requisite for the third phase, representing the docking of the guide's 3'-end in the PAZ-domain [30]. Therefore, Argonaute provides a scaffold for the integration of a guide strand, making it plausible that guide recognition follows a highly conserved mechanism.

While the guide strand is firmly anchored in Ago, integration of the target strand is mainly mediated via Watson–Crick base pairing with the guide strand. The bases of the guide in the seed region face outwards, ready to base pair with a target. Thus, the target strand associates with the pre-arranged guide strand, forming an A-form helix in the seed-region. This nucleation process does not require an extensive interaction network between the protein and the target strand [13,14]. Instead, specificity is mediated by the shape complementarity between Argonaute and the minor groove of the duplex, which allows for hydrophobic and van der Waals interactions of the linker 2 domain with the guide-target duplex in the seed region [18]. The guide-target duplex extends from the seed towards the 3'-end binding site. The duplex remains in a helical form before further base pairing is disrupted in the N-PAZ channel, where the helical form is blocked by the N-terminal domain. Bacterial structures showed that the strand separation occurs at nucleotide 16 of the guide strand [14]. The unstacked bases of the guide face into

the interior of the complex, preventing an association of the target strand in the 3' half of the duplex. The catalytic center resides in the PIWI domain, which adopts a typical RNase H fold. The catalytic activity is mediated by a catalytic tetrad composed of a highly conserved DEDX (X being an aspartate or histidine). In addition, parts of the N-terminal domain influence the cleavage activity of human Argonautes [31–34]. Here, unstructured loops seem to arrange the target strand in a cleavage-compatible geometry.

While the overall structural organization of Ago across the domains of life is highly similar, there are also significant differences noticeable. The individual Ago domains align well, but the relative position of the lobes differs significantly. PfAgo adopts a more compacted form than hAgo2 and TtAgo, with the PAZ and MID module moved towards each other (Figure 1). However, the unstructured loops found in the eukaryotic N-terminal domain necessary for cleavage activity are significantly shortened or non-existent in the archaeal counterparts. Eukaryotic Ago proteins show a number of additional loops and unstructured elements [16,17,35]. Most of these elements are surface-exposed, representing contact sites for eukaryotic-specific interaction partners [36]. A significant difference between eukaryotic and archaeal-bacterial Argonautes can be found in the surface charge distribution. The surface of the duplex binding channel, as well as the binding pocket in the MID domain of eukaryotic Ago is strongly positively charged (Figure 1). In contrast, the binding channel of TtAgo is less positively charged (Figure 1), and the binding pocket is more hydrophobic [17]. This is even more pronounced in PfAgo. An extended negatively-charged surface can be found in the N-PAZ tunnel of PfAgo, which might hint at an alternative pathway for the guide or target strand after the release of the guide strand from the PAZ domain.

3. Conformational Flexibility Facilitates Argonaute Function

Argonaute does not merely represent a static scaffold structure that assists nucleic acid strand association and dissociation, but the protein itself undergoes dynamic changes throughout its activity cycle. The crystal structures of the binary and ternary complex, pre-steady state kinetics and single-molecule experiments revealed the presence of flexible domains, loops and helices that promote the positioning of the target strand in a cleavage-compatible configuration. The most pronounced conformational change occurs on the progression from the binary to the ternary complex, as captured for TtAgo (Figure 2b). TtAgo structures in complex with a 21-mer DNA guide and DNA or RNA target strands showed that upon loading of a sufficiently long target strand, the 3'-end of the guide is released from the PAZ domain [14,23]. Single-molecule studies using TtAgo, a DNA guide and an RNA target strand even suggested that the release of the 3'-end might be a dynamic event [37]. The rearrangement of the 3' half of the guide is accompanied by a significant movement of the PAZ domain. Meanwhile, the 5'-end of the guide strand remains firmly anchored in the MID binding pocket [14]. The release from the PAZ domain is a direct consequence of the helical structure of the DNA that prevents the 3'-end from reaching the PAZ binding pocket. Hence, PAZ release is correlated with the length of the duplex. TtAgo structures reveal the influence of the target substrate on the duplex length. A 15-mer target RNA leads to a DNA guide-RNA target duplex with a length of 14 bp, which is accompanied by the release of the guide 3'-end from the PAZ domain [14]. In contrast with a 15-mer target DNA, the DNA guide-DNA target duplex spans only 13 bp, with the guide 3'-end still being anchored in the PAZ domain.

A 16-mer DNA target leads to the formation of 15 bp between guide and target DNA and induces the release of the guide 3'-end from the PAZ domain. Therefore, with TtAgo, the rearrangement of a cleavage-incompatible to a cleavage-compatible conformation differs for RNA and DNA target substrates [23]. However, incorporation of a short target strand already induces a conformational change in the protein, leading to the opening of the PAZ domain [14]. A more pronounced rotation of the PAZ domain in a situation where the nucleic acid duplex is further extended results in a widening of the nucleic acid binding channel to accommodate the target strand [14,23]. A comparable ternary structure of eukaryotic Argonaute variants, including a full-length target RNA, could not be solved yet. However, analogous to the bacterial structures, hAgo2 loaded with an RNA guide only or a short RNA duplex (11-nt target) shows the 3'-end still anchored in the PAZ domain [18]. Nevertheless, kinetic experiments suggest that extended pairing of the guide and target most likely also leads to 3'-end release [30]. Even though there are no structures of the binary or ternary complex available for the archaeal domain, recent single-molecule fluorescence resonance energy transfer (FRET) studies on the archaeal Argonaute protein from *Methanocaldococcus jannaschii* in solution support the common theme of PAZ release upon the formation of the ternary complex using a DNA duplex (21-mer guide, 20-mer target) [10]. This study also showed that the release of the guide 3'-end in the archaeal enzyme does not require cleavage of the target strand. Taken together, these data support the two-state model of Argonaute action [38]; a mechanism conserved in Argonaute variants from all domains of life.

Integration of the target strand in TtAgo leads furthermore to the straightening of the guide DNA to fully adopt an A-form with consecutive base-stacking. As a result, bases 10 and 11 of the guide strand bound by TtAgo stack on top, allowing the correct orientation of the scissile phosphate of the target strand relative to the catalytic residues [14]. The hAgo2 structures also show that the kink between nucleotides 6 and 7 of the guide is relieved due to a movement of helix $\alpha 7$, which is necessary to avoid steric clashes with the target strand (Figure 2). Helix $\alpha 7$ and the PAZ domain move as discrete rigid bodies relative to the MID, PIWI and N-terminal domain upon target loading [18]. Interestingly, helix $\alpha 7$ is conserved in archaeal, but not bacterial Ago [17]. A significant conformational transition occurs close to the active center of the enzyme with direct consequences on the active site configuration. In TtAgo, PIWI loop 2 undergoes a conformational switch, thereby inserting a “glutamate finger” into the active site, completing the catalytic tetrad (Figure 2c). Strikingly, all TtAgo structures with the 3'-end located in the PAZ domain show the “unplugged” conformation, while all PAZ-released structures show the “plugged” conformation of the glutamate finger. Hence, the PAZ release seems to be coupled to the activation of the slicing activity of the enzyme. The archaeal structure of unliganded Argonaute exhibits an unplugged conformation. The unplugged to plugged transition is not found in eukaryotic Agos: all structures available show the glutamate finger in the plugged conformation, irrespective of the PAZ release [16–18,35].

4. Molecular Mechanism of the Silencing Process

4.1. RNA Interference Mediated by Human Argonaute 2

In eukaryotes, Argonaute constitutes the principal component of the eukaryotic RNA interference (RNAi) pathway, a mechanism fundamental to posttranscriptional regulation of gene expression [39]. Generally, in the RNAi system, Argonaute 2 is loaded with short double-stranded RNAs within the RLC (RISC-loading complex) [40,41]. One strand of the RNA duplex (the guide strand) is retained in the Argonaute protein, which is part of a multiprotein complex, called RISC (RNA-induced silencing complex). The non-guide strand (passenger strand) is cleaved by Argonaute and eventually ejected, allowing the guide strand to find its cognate mRNA target [25,31,42]. The fully complementary target RNA of a siRNA-guided RISC is cleaved by hAgo2 (slicing) and released, and RISC can engage in another round of slicing [43,44]. In contrast, miRNAs typically guide hAgo2 to partially complementary targets in the 3' untranslated regions (UTR) of mRNAs, which leads either to mRNA degradation by hAgo2 or to translational inhibition [45–47]. Extensive research over the last two decades has shown that the function of Argonaute exceeds its role in RNA silencing (for recent reviews, see [36,48,49]) and revealed its role in chromatin dynamics [50], transcriptional regulation [51,52], alternative splicing [53–55] and double-strand break repair [56]. Eukaryotic organisms frequently encode more than one Argonaute gene. The nematode *C. elegans* encodes an impressive Argonaute family of 27 members [57,58]. However, not all Argonautes proteins are catalytically active variants. Among the four human Agos, only hAgo2 acts as a nuclease [59,60].

4.2. Prokaryotic Argonaute Acts in DNA-Silencing Pathways

Despite the mechanistic and structural similarities between the archaeal, prokaryotic and eukaryotic Argonaute proteins, a major difference in the fundamental mechanism of silencing was revealed recently. Bacterial Argonautes use either DNA or RNA guide strands to silence complementary DNA strands (Figure 3). In native cells Argonaute from the alphaproteobacterium *Rhodobacter sphaeroides* (RsAgo) is associated with small RNAs and DNAs [8]. Interestingly, RsAgo belongs to the Ago class with an inactivated catalytic tetrad. However, RsAgo is encoded in an operon with a predicted DNA nuclease. The small RNAs are derived from mRNA precursors that can be mapped to the majority of cellular transcripts and most likely are generated from mRNA degradation products. DNAs associated with RsAgo are largely complementary to the bound RNAs. A current model for the generation of RNA-interacting DNAs (riDNA) proposes that the small RNA directs RsAgo to the complementary DNA target followed by the nucleolytic cleavage of the DNA by a yet unidentified nuclease. Alternatively, RsAgo-RNA complexes loaded onto a complementary stretch of DNA inhibit RNA polymerase loading or block RNA polymerase elongation, leading to transcriptional repression of the target DNA. The enrichment of riDNA for foreign sequences, like plasmid DNA and transposons, suggests that the RsAgo-mediated DNA silencing mechanism is in place to destroy foreign genetic elements. However, the molecular mechanisms that allow the discrimination between self and foreign DNA are not known. Another example of bacterial Ago-mediated DNA silencing was described for TtAgo. Expression of TtAgo in *E. coli* and subsequent characterization of TtAgo-bound nucleic acids revealed that TtAgo associates primarily with DNA sequences (small interfering DNAs) preferentially

derived from its own expression plasmid [9]. The underlying mechanism for DNA guide processing from foreign DNA (e.g., plasmids) has not been deciphered yet. However, loading of TtAgo with guide DNA and subsequent cleavage of target DNA is only observed if the catalytic center of the enzyme is intact, indicating that the nuclease activity of TtAgo is required for guide processing. Even though the experimental characterization of TtAgo *in vivo* was mainly carried out using plasmid DNA, it is feasible that the DNA-guided DNA silencing mechanism targets replication intermediates from invading genetic elements and DNA taken up by the natural competence system present in *Thermus thermophilus*. The functional role of archaeal Argonaute still remains elusive. However, *in vitro* studies showed that MjAgo exclusively cleaves DNA targets when using a DNA guide [10], suggesting that this archaeal Argonaute variant, like the bacterial counterparts, is involved in DNA silencing processes.

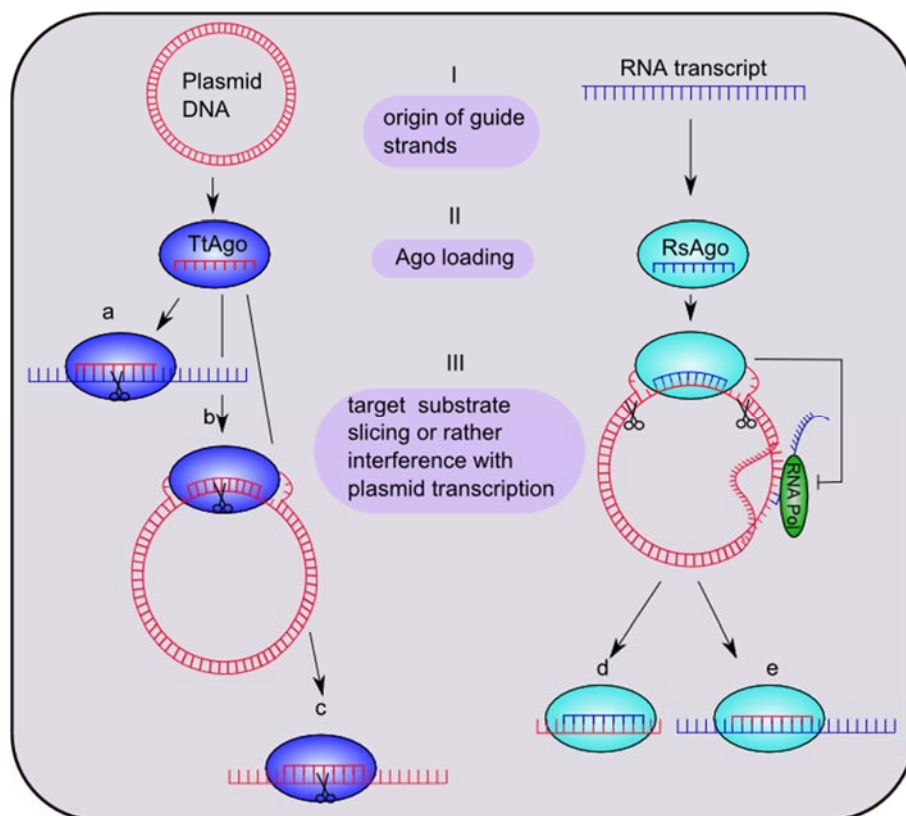


Figure 3. Putative mechanisms of bacterial Ago-mediated silencing pathways. (I) Guide sequences of *Thermus thermophilus* Ago (TtAgo) and *Rhodobacter sphaeroides* Ago (RsAgo) are derived from plasmid DNA or RNA transcripts, respectively. (II) TtAgo is loaded with a 13–25-nt guide DNA and RsAgo with a 15–19-nt guide RNA. (III) The target substrates are (a) ssRNA, (b) negatively-supercoiled plasmid DNA and (c) ssDNAs in the case of TtAgo. RsAgo binds plasmid DNA, which will be cleaved, yielding 22–24-nt DNA fragments. Binding of the guide-Ago complex to plasmid DNA furthermore possibly leads to an inhibition of plasmid transcription. The short fragments either stay bound to Argonaute (d) or interact with other RsAgo molecules to constitute DNA-RsAgo complexes and regulate plasmid transcription (e). Figure in part modified from Olovnikov *et al.* [8].

4.2. Diversity in Recognition and Selection of Guide and Target Strands

Most eukaryotic and prokaryotic Agos are able to associate with RNA, as well as with DNA substrates [10,14,61], but act in a very selective way *in vivo*. Thus, what are the determinants that guide the selection of small nucleic acid interaction partners for Ago proteins? Signatures like length, phosphorylation status and the identity of the first nucleotide, play an important role in selection and sorting (see Table 1 for an overview). Hago2 preferentially associates with small RNAs 20–25 nucleotides in length. Even though hAgo2 tolerates a variety of guide lengths, lengths of 21 nt (siRNA) [43,62] or 22 nt (miRNA) [44] are most abundantly associated with hAgo2. Furthermore, dsRNAs associated with hAgo2 are distinguished by 2-nt overhangs, a 5'-end phosphate group and a 3'-end hydroxyl group [63]. These signatures are a result of the biogenesis pathway [64]. Central to si/miRNA biogenesis are Drosha and Dicer, two nucleases that process dsRNA substrates into short dsRNA fragments, typically of 21–25 nt in length [65]. Drosha is involved in the maturation of the 5'-end, creating the characteristic phosphate group at the 5'-end. Structural information on Dicer came from the Doudna lab, which was able to solve the structure of the enzyme derived from the unicellular eukaryote *Giardia intestinalis* [66]. The structure revealed why Dicer cleavage products (i) have a certain length and (ii) possess a 2-nt overhang at the 3'-end. The RNase III domains form an intramolecular dimer, and the active centers are located opposite, but slightly shifted to each other, separated by 17.5 Å, which matches the width of a major groove in dsRNA, making it plausible that Dicer generates dsRNAs with a 2-nt overhang. Recent structural and functional studies revealed that Dicer not only anchors the 3'-end, but simultaneously the 5'-end of the RNA substrate in a pocket in the platform domain of Dicer [67,68]. 5'-end docking is especially important to generate uniform 22 nt-long products (5' counting rule).

Table 1. Substrate preferences of prokaryotic Argonautes proteins *in vitro* and *in vivo*. n.d. = not determined; inactive = no catalytic tetrad present; () = significantly reduced affinity; # = determined *in vitro*; * = determined *in vivo*; { } unpublished.

Argonaute Variant	Guide Strand Bound	Preference/Enrichment For 1st Guide Nucleotide	Target Strands Bound	Target Strands Cleaved	Ref.
Archaea					
<i>M. jannaschii</i> Ago	DNA #, (RNA) #	{G} #	DNA #, RNA #	DNA #	[10]
<i>A. fulgidus</i> Ago	DNA #, (RNA) #	n.d.	DNA #, RNA #	n.d.	[19]
Bacteria					
<i>A. aeolicus</i> Ago	DNA *, (RNA) *	n.d.	RNA #, DNA #	n.d.	[69]
<i>R. sphaeroides</i> Ago	RNA *, DNA *	U *	RNA *, plasmid DNA #	inactive	[8]
<i>T. thermophilus</i> Ago	DNA *,# , (RNA) #	C *	DNA *,# , RNA # , plasmid DNA *,#	DNA #, RNA # , plasmid DNA *,#	[9,14]

Stable anchoring of the 5'-end requires a 5' phosphate, which is the result of pre-miRNA processing by Drosha. The isolated MID domain of hAgo2 binds AMP and UMP with up to 30-fold higher affinity as compared to CMP and GMP [21]. A preference for a uridine as the first nucleotide in eukaryotic Ago-associated RNA was observed in several studies indicating that the first nucleotide serves as a

determinant for guide selection [16,57,64,70,71]. In plants, small RNA sorting in the different Agos is predominantly determined by the identity of the 5' nucleotide [59,72]. Structural studies have provided the structural basis for the 5' nucleotide specificity [21,73,74] that is conferred by the nucleotide specificity loop (Figure 2a). The 5' terminal nucleotide packs against a tyrosine residue (Y529), which contributes to the non-specific recognition of the base. In addition, specific interactions with a threonine (T526) and glycine residue in the specificity loop are only possible if the base is an UMP or AMP. The specificity loop is missing in the bacterial Ago variants, and only residue N436 undergoes a specific interaction with the first nucleotide (TtAgo in complex with the let-7 DNA guide and the DNA or RNA target) [14,23]. However, specific interactions between the protein backbone and the nucleotide cannot be established in AfPIWI and PfAgo, as the specificity loop is arranged too far from the 5'-nucleotide [22,24] (Figure 2a). It remains to be determined whether there are additional factors that determine the nature of the 5'-nucleotide where a 5'-nucleotide preference occurs, albeit the nucleotide bias appears not to follow a common pattern. For example, it would be conceivable that a pre-processing enzyme generates DNA guides with a defined 5'-nucleotide.

Information about the nucleic acid interactions partners associated with prokaryotic Ago *in vivo* are only available for bacterial Ago variants [8,9], as the growth and manipulation of archaeal organisms that encode potentially active Argonautes are notoriously difficult. However, *in vitro* studies on the recombinant archaeal MjAgo showed that the protein binds short nucleic acids with the characteristic siRNA/miRNA signatures, e.g., a 5'-end phosphate on the guide strand and two nucleotide overhangs, irrespective of the nucleic acid chemistry. However, MjAgo preferentially binds and exclusively cleaves DNA/DNA hybrids [10]. Furthermore, MjAgo exhibits a preference for a deoxyguanosine as the first nucleotide in the guide strand (unpublished data). A similar preference for DNA substrates was found for the isolated PIWI domain of *A. fulgidus* [19]. These data hint to the possibility that archaeal Agos, like their bacterial counterparts, act in a DNA-guided DNA silencing mechanism.

DNA-guided DNA silencing has been described for the bacterial Ago from *Thermus thermophilus* *in vitro* and *in vivo* experiments [9]. TtAgo loaded with DNA cleaves DNA and RNA substrates *in vitro* [14]. An overhang of the target at the 3'-end of the target is not required for efficient cleavage [14]. When expressed in *E. coli*, TtAgo associates with DNA guides 13–25 nt in length with a strong bias towards 15-nt guides. Furthermore, the guide DNA strands appear to carry the typical 5'-end phosphate. Interestingly, the majority of guides have a deoxycytidine and a deoxyadenosine at the first and second nucleotide, respectively. These guides enable TtAgo to cleave double-stranded DNA, supporting the idea that TtAgo follows a DNA-guided DNA silencing mechanism [9] (Figure 3). However, bacterial Argonautes do not follow a common selection pattern. *A. aeolicus* Ago shows the highest affinity for DNA as a guide and for a DNA/RNA as a hybrid [69]. In contrast, RsAgo favors 5'-end phosphorylated 18-nt RNA guide strands and 24-nt DNA target strands. The first and second nucleotide of the guide shows a strong enrichment for a uridine. The target strand is fully complementary to the guide strand, but shows an unusual 3-nt overhang on both sides. It is difficult to rationalize the selection patterns of prokaryotic Argonautes, as no homologues of Drosha or Dicer, which act further upstream of Argonaute and could distinguish between RNA or DNA precursors, are encoded. Heterologous expression of RsAgo in *E. coli* leads to specific loading of 5'U-RNAs with the correct length from a pool of available sequences and, hence, might indicate that no pre-processing machinery is needed for RNA-guided DNA silencing by RsAgo [8].

5. An Archaeal Perspective

While the biological function of Argonaute proteins is not conserved, many structural features and conformational changes that support the activity cycle of Argonaute are well-preserved across domains. Among them is the ability to bind short DNA or RNA guide strands with a preference for a terminal phosphate at the 5'-end of the guide that ensures stable incorporation of the guide into the MID domain. In addition to structural features, conformational dynamics are also conserved, suggesting that they play an important role for the function of the protein. Structural and biophysical studies demonstrated that the PAZ domain is a mobile element in all Argonaute variants [10,13,30,37]. However, some of the structural features are exclusively shared between archaeal and eukaryotic Argonaute variants. Helix $\alpha 7$ is not conserved in the bacterial domain, and the interaction network that allows the recognition of the first nucleotide at the 5'-end of the guide is well preserved between archaeal and eukaryotic Argonautes. The recently discovered DNA silencing function of bacterial Agos raises the question of whether comparable mechanisms are in place in archaeal organisms. However, while archaeal proteins served as an excellent model system to elucidate the structure of the Argonaute family, catalytically active forms were rarely characterized *in vivo* or *in vitro*. In addition, the structure of an archaeal Argonaute in complex with nucleic acid substrates is still lacking. Experiments employing the Argonaute protein from the hyperthermophilic organism *Methanocaldococcus jannaschii* suggest, however, that some archaeal Argonautes support a DNA-guided DNA silencing mechanism. From an archaeal perspective, there is still much to explore about the structure and function of Argonautes, and further studies are needed to discover the full picture of the Argonaute family.

Acknowledgments

We thank Tobias Restle for discussions. Research at the Grohmann laboratory is supported by the Deutsche Forschungsgemeinschaft (GR 3840/2-1), the Fond der Chemischen Industrie and the Boehringer Ingelheim Fonds. Sarah Willkomm is supported by the Graduate School for Computing in Medicine and Life Sciences at the University of Lübeck funded by Germany's Excellence Initiative [DFG GSC 235/1].

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Bohmert, K.; Camus, I.; Bellini, C.; Bouchez, D.; Caboche, M.; Benning, C. Ago1 defines a novel locus of arabidopsis controlling leaf development. *EMBO J.* **1998**, *17*, 170–180.
2. Lin, H.; Spradling, A.C. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the drosophila ovary. *Development* **1997**, *124*, 2463–2476.
3. Burroughs, A.M.; Ando, Y.; Aravind, L. New perspectives on the diversification of the RNA interference system: Insights from comparative genomics and small RNA sequencing. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 141–181.

4. Cerutti, L.; Mian, N.; Bateman, A. Domains in gene silencing and cell differentiation proteins: The novel PAZ domain and redefinition of the piwi domain. *Trends Biochem. Sci.* **2000**, *25*, 481–482.
5. 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.
6. Swarts, D.C.; Makarova, K.; Wang, Y.; Nakanishi, K.; Ketting, R.F.; Koonin, E.V.; Patel, D.J.; van der Oost, J. The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.* **2014**, *21*, 743–753.
7. Burroughs, A.M.; Iyer, L.M.; Aravind, L. Two novel piwi families: Roles in inter-genomic conflicts in bacteria and mediator-dependent modulation of transcription in eukaryotes. *Biol. Direct* **2013**, *8*, doi:10.1186/1745-6150-8-13.
8. Olovnikov, I.; Chan, K.; Sachidanandam, R.; Newman, D.K.; Aravin, A.A. Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell* **2013**, *51*, 594–605.
9. Swarts, D.C.; Jore, M.M.; Westra, E.R.; Zhu, Y.; Janssen, J.H.; Snijders, A.P.; Wang, Y.; Patel, D.J.; Berenguer, J.; Brouns, S.J.; *et al.* DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **2014**, *507*, 258–261.
10. Zander, A.; Holzmeister, P.; Klose, D.; Tinnefeld, P.; Grohmann, D. Single-molecule fret supports the two-state model of Argonaute action. *RNA Biol.* **2014**, *11*, 45–56.
11. Rashid, U.J.; Paterok, D.; Koglin, A.; Gohlke, H.; Piehler, J.; Chen, J.C. Structure of aquifex aeolicus Argonaute highlights conformational flexibility of the PAZ domain as a potential regulator of RNA-induced silencing complex function. *J. Biol. Chem.* **2007**, *282*, 13824–13832.
12. 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.
13. 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.
14. 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.
15. 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.
16. Nakanishi, K.; Weinberg, D.E.; Bartel, D.P.; Patel, D.J. Structure of yeast Argonaute with guide RNA. *Nature* **2012**, *486*, 368–374.
17. Schirle, N.T.; MacRae, I.J. The crystal structure of human Argonaute2. *Science* **2012**, *336*, 1037–1040.
18. Schirle, N.T.; Sheu-Gruttadauria, J.; MacRae, I.J. Gene regulation. Structural basis for microRNA targeting. *Science* **2014**, *346*, 608–613.
19. Ma, J.B.; Yuan, Y.R.; Meister, G.; Pei, Y.; Tuschl, T.; Patel, D.J. Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* piwi protein. *Nature* **2005**, *434*, 666–670.
20. Elkayam, E.; Kuhn, C.D.; Tocilj, A.; Haase, A.D.; Greene, E.M.; Hannon, G.J.; Joshua-Tor, L. The structure of human Argonaute-2 in complex with mir-20a. *Cell* **2012**, *150*, 100–110.
21. 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.

22. Parker, J.S.; Roe, S.M.; Barford, D. Crystal structure of a piwi protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J.* **2004**, *23*, 4727–4737.
23. Sheng, G.; Zhao, H.; Wang, J.; Rao, Y.; Tian, W.; Swarts, D.C.; van der Oost, J.; Patel, D.J.; Wang, Y. Structure-based cleavage mechanism of thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 652–657.
24. Parker, J.S.; Parizotto, E.A.; Wang, M.; Roe, S.M.; Barford, D. Enhancement of the seed-target recognition step in RNA silencing by a piwi/mid domain protein. *Mol. Cell* **2009**, *33*, 204–214.
25. Brennecke, J.; Stark, A.; Russell, R.B.; Cohen, S.M. Principles of microRNA-target recognition. *PLoS Biol.* **2005**, *3*, e85.
26. Lewis, B.P.; Burge, C.B.; Bartel, D.P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **2005**, *120*, 15–20.
27. Faehnle, C.R.; Elkayam, E.; Haase, A.D.; Hannon, G.J.; Joshua-Tor, L. The making of a slicer: Activation of human Argonaute-1. *Cell Rep.* **2013**, *3*, 1901–1909.
28. 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.
29. Lingel, A.; Simon, B.; Izaurralde, E.; Sattler, M. Structure and nucleic-acid binding of the drosophila Argonaute 2 PAZ domain. *Nature* **2003**, *426*, 465–469.
30. Deerberg, A.; Willkomm, S.; Restle, T. Minimal mechanistic model of siRNA-dependent target RNA slicing by recombinant human Argonaute 2 protein. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 17850–17855.
31. Leuschner, P.J.; Ameres, S.L.; Kueng, S.; Martinez, J. Cleavage of the siRNA passenger strand during risc assembly in human cells. *EMBO Rep.* **2006**, *7*, 314–320.
32. Schurmann, N.; Trabuco, L.G.; Bender, C.; Russell, R.B.; Grimm, D. Molecular dissection of human Argonaute proteins by DNA shuffling. *Nat. Struct. Mol. Biol.* **2013**, *20*, 818–826.
33. Hauptmann, J.; Dueck, A.; Harlander, S.; Pfaff, J.; Merkl, R.; Meister, G. Turning catalytically inactive human Argonaute proteins into active slicer enzymes. *Nat. Struct. Mol. Biol.* **2013**, *20*, 814–817.
34. Hauptmann, J.; Kater, L.; Loffler, P.; Merkl, R.; Meister, G. Generation of catalytic human Ago4 identifies structural elements important for RNA cleavage. *RNA* **2014**, *20*, 1532–1538.
35. Nakanishi, K.; Ascano, M.; Gogakos, T.; Ishibe-Murakami, S.; Serganov, A.A.; Briskin, D.; Morozov, P.; Tuschl, T.; Patel, D.J. Eukaryote-specific insertion elements control human Argonaute slicer activity. *Cell Rep.* **2013**, *3*, 1893–1900.
36. Meister, G. Argonaute proteins: Functional insights and emerging roles. *Nat. Rev. Genet.* **2013**, *14*, 447–459.
37. Jung, S.R.; Kim, E.; Hwang, W.; Shin, S.; Song, J.J.; Hohng, S. Dynamic anchoring of the 3'-end of the guide strand controls the target dissociation of Argonaute-guide complex. *J. Am. Chem. Soc.* **2013**, *135*, 16865–16871.
38. Tomari, Y.; Zamore, P.D. Perspective: Machines for RNAi. *Genes Dev.* **2005**, *19*, 517–529.
39. Hutvagner, G.; Simard, M.J. Argonaute proteins: Key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 22–32.
40. 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.

41. Sontheimer, E.J. Assembly and function of RNA silencing complexes. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 127–138.
42. Matranga, C.; Tomari, Y.; Shin, C.; Bartel, D.P.; Zamore, P.D. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* **2005**, *123*, 607–620.
43. Haley, B.; Zamore, P.D. Kinetic analysis of the RNAi enzyme complex. *Nat. Struct. Mol. Biol.* **2004**, *11*, 599–606.
44. Hutvagner, G.; Zamore, P.D. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* **2002**, *297*, 2056–2060.
45. Filipowicz, W.; Bhattacharyya, S.N.; Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: Are the answers in sight? *Nat. Rev. Genet.* **2008**, *9*, 102–114.
46. Fukaya, T.; Iwakawa, H.O.; Tomari, Y. MicroRNAs block assembly of eIF4F translation initiation complex in drosophila. *Mol. Cell* **2014**, *56*, 67–78.
47. Fukao, A.; Mishima, Y.; Takizawa, N.; Oka, S.; Imataka, H.; Pelletier, J.; Sonenberg, N.; Thoma, C.; Fujiwara, T. MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. *Mol. Cell* **2014**, *56*, 79–89.
48. Mallory, A.; Vaucheret, H. Form, function, and regulation of Argonaute proteins. *Plant Cell* **2010**, *22*, 3879–3889.
49. Huang, V.; Li, L.C. Demystifying the nuclear function of Argonaute proteins. *RNA Biol.* **2014**, *11*, 18–24.
50. Li, L.C. Chromatin remodeling by the small RNA machinery in mammalian cells. *Epigenetics* **2014**, *9*, 45–52.
51. Cernilogar, F.M.; Onorati, M.C.; Kothe, G.O.; Burroughs, A.M.; Parsi, K.M.; Breiling, A.; Lo Sardo, F.; Saxena, A.; Miyoshi, K.; Siomi, H.; *et al.* Chromatin-associated RNA interference components contribute to transcriptional regulation in drosophila. *Nature* **2011**, *480*, 391–395.
52. Huang, V.; Zheng, J.; Qi, Z.; Wang, J.; Place, R.F.; Yu, J.; Li, H.; Li, L.C. Ago1 interacts with RNA polymerase ii and binds to the promoters of actively transcribed genes in human cancer cells. *PLoS Genet.* **2013**, *9*, e1003821.
53. Allo, M.; Buggiano, V.; Fededa, J.P.; Petrillo, E.; Schor, I.; de la Mata, M.; Agirre, E.; Plass, M.; Eyra, E.; Elela, S.A.; *et al.* Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* **2009**, *16*, 717–724.
54. Ameyar-Zazoua, M.; Rachez, C.; Souidi, M.; Robin, P.; Fritsch, L.; Young, R.; Morozova, N.; Fenouil, R.; Descostes, N.; Andrau, J.C.; *et al.* Argonaute proteins couple chromatin silencing to alternative splicing. *Nat. Struct. Mol. Biol.* **2012**, *19*, 998–1004.
55. Liu, J.; Hu, J.; Corey, D.R. Expanding the action of duplex RNAs into the nucleus: Redirecting alternative splicing. *Nucleic Acids Res.* **2012**, *40*, 1240–1250.
56. Gao, M.; Wei, W.; Li, M.M.; Wu, Y.S.; Ba, Z.; Jin, K.X.; Liao, Y.Q.; Adhikari, S.; Chong, Z.; Zhang, T.; *et al.* Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. *Cell Res.* **2014**, *24*, 532–541.
57. Peters, L.; Meister, G. Argonaute proteins: Mediators of RNA silencing. *Mol. Cell* **2007**, *26*, 611–623.
58. Yigit, E.; Batista, P.J.; Bei, Y.; Pang, K.M.; Chen, C.C.; Tolia, N.H.; Joshua-Tor, L.; Mitani, S.; Simard, M.J.; Mello, C.C. Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. *Cell* **2006**, *127*, 747–757.

59. Meister, G.; Landthaler, M.; Patkaniowska, A.; Dorsett, Y.; Teng, G.; Tuschl, T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **2004**, *15*, 185–197.
60. Liu, J.; Carmell, M.A.; Rivas, F.V.; Marsden, C.G.; Thomson, J.M.; Song, J.J.; Hammond, S.M.; Joshua-Tor, L.; Hannon, G.J. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **2004**, *305*, 1437–1441.
61. 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.
62. Elbashir, S.M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **2001**, *411*, 494–498.
63. Rivas, F.V.; Tolia, N.H.; Song, J.J.; Aragon, J.P.; Liu, J.; Hannon, G.J.; Joshua-Tor, L. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat. Struct. Mol. Biol.* **2005**, *12*, 340–349.
64. Ha, M.; Kim, V.N. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 509–524.
65. Bernstein, E.; Caudy, A.A.; Hammond, S.M.; Hannon, G.J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **2001**, *409*, 363–366.
66. 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.
67. 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.
68. Tian, Y.; Simanshu, D.K.; Ma, J.B.; Park, J.E.; Heo, I.; Kim, V.N.; Patel, D.J. A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Mol. Cell* **2014**, *53*, 606–616.
69. 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.
70. Hu, H.Y.; Yan, Z.; Xu, Y.; Hu, H.; Menzel, C.; Zhou, Y.H.; Chen, W.; Khaitovich, P. Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* **2009**, *10*, doi:10.1186/1471-2164-10-413.
71. Kawamata, T.; Seitz, H.; Tomari, Y. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat. Struct. Mol. Biol.* **2009**, *16*, 953–960.
72. Mi, S.; Cai, T.; Hu, Y.; Chen, Y.; Hodges, E.; Ni, F.; Wu, L.; Li, S.; Zhou, H.; Long, C.; *et al.* Sorting of small RNAs into Arabidopsis Argonaute complexes is directed by the 5' terminal nucleotide. *Cell* **2008**, *133*, 116–127.
73. Boland, A.; Tritschler, F.; Heimstadt, S.; Izaurralde, E.; Weichenrieder, O. Crystal structure and ligand binding of the mid domain of a eukaryotic Argonaute protein. *EMBO Rep.* **2010**, *11*, 522–527.
74. Frank, F.; Hauver, J.; Sonenberg, N.; Nagar, B. Arabidopsis Argonaute mid domains use their nucleotide specificity loop to sort small RNAs. *EMBO J.* **2012**, *31*, 3588–3595.