ARTICLE
Journal of Cellular Biochemistry 113:2576–2585 (2012)

Journal of Cellular Biochemistry

Structural Evolution and Functional Diversification Analyses of Argonaute Protein

Kai-Fa Wei,[1*] Ling-Juan Wu,[1] Juan Chen,[1] Yan-feng Chen,[1] and Dao-xin Xie[2]

[1] *Department of Biological Sciences and Biotechnology, Zhangzhou Normal University, Zhangzhou 363000, Fujian, China*

[2] *School of Life Sciences, Tsinghua University, Beijing 10084, China*

ABSTRACT

Argonaute (AGO) proteins are highly specialized small-RNA-binding modules and small RNAs are anchored to their specific binding pockets guiding AGO proteins to target mRNA molecules for silencing or destruction. The 135 full-length AGO protein sequences derived from 36 species covering prokaryote, archaea, and eukaryote are chosen for structural and functional analyses. The results show that bacteria and archaeal AGO proteins are clustered in the same clade and there exist multiple AGO proteins in most eukaryotic species, demonstrating that the increase of AGO gene copy number and horizontal gene transfer (HGT) have been the main evolutionary driving forces for adaptability and biodiversity. And the emergence of PAZ domain in AGO proteins is the unique evolutionary event. The analysis of middle domain (MID)-nucleotide contaction shows that either the position of sulfate I bond in Nc_QDE2 or the site of phosphate I bond in Hs_AGO2 represents the 5′-nucleotide binding site of miRNA. Also, H334, T335, and Y336 of Hs_AGO1 can form hydrogen bonds with 3′-overhanging ends of miRNAs and the same situation exists in Hs_AGO2, Hs_AGO3, Hs_AGO4, Dm_AGO1, and Ce_Alg1. Some PIWI domains containing conserved DDH motif have no slicer activity, and post-translational modifications may be associated with the endonucleolytic activities of AGOs. With the numbers of AGO genes increasing and fewer crystal structures available, the evolutionary and functional analyses of AGO proteins can help clarify the molecular mechanism of function diversification in response to environmental changes, and solve major issues including host defense mechanism against virus infection and molecular basis of disease. J. Cell. Biochem. 113: 2576–2585, 2012. © 2012 Wiley Periodicals, Inc.

KEY WORDS: sRNA; AGO PROTEIN; STRUCTURAL EVOLUTION; FUNCTIONAL DIVERSIFICATION; PHYSIOLOGICAL PROCESS

RNA interference (RNAi) and other gene silencing pathways are triggered by small RNAs (sRNAs), which function as sequence-specific guide to regulate mRNA stability, protein synthesis, transposon silencing, and chromatin organization. Nowadays, five major groups of sRNAs have been detected: tiRNAs, short interfering RNAs (siRNAs), microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs), and heterochromatin associated small RNAs (hsRNAs). sRNAs and their potential protein interaction partners assemble into a protein–RNA complex called the RNA-induced silencing complex (RISC) [Hammond et al., 2000; Meister and Tuschl, 2004; Carthew and Sontheimer, 2009; Saito and Siomi, 2010]. In the RISCs, the sRNA (the guide strand) is loaded into an Argonaute (AGO) protein which is the core of RISCs, and recruits AGO protein to complementary target RNA by base-pairing interactions. Perfect complementarity around the cleavage site in the guide–target duplex is a prerequisite for target RNA slicing. After slicing, the cleaved target RNA is released and the RISC is recycled for another round of slicing. Nonperfect complementarily prevents the target from being sliced by AGO protein, but its translation may be repressed and its poly (A) tails may be removed, which leads to mRNA degradation. miRNA regulation has emerged as a critical regulatory principle in the mammalian and plant physiological processes. miR-17∼92 in mice regulates components of phosphatidylinositol-3-kinase pathway (PI3K) and inhibits apoptosis; miR-155 promotes autoimmune inflammation by enhancing inflammatory T cell development in mice; miR417 of Arabidopsis has negatively regulating roles on the rate of the seed germination and seedling survival; miR393 is strongly upregulated by ABA, cold, dehydration, and salt stress, and the putative target gene encoding E3 ubiquitin ligase SCF complex F-box protein,

Abbreviations used: AGO, argonaute; Ago, Ago clade protein; HGT, horizontal gene transfer; WAGO, worm-specific clade protein; PIWI, PIWI clade protein; pAGO, prokaryote AGO; aAGO, archaeal AGO; QDE-2, quelling deficient 2; eIF4E, translation initiation factor 4; Aub, aubergine.

Additional supporting information may be found in the online version of this article.

*Correspondence to: Prof. Kai-Fa Wei, Ph.D., Department of Biological Sciences and Biotechnology, Zhangzhou Normal University, 36 Xian Qian Zhi Street, Zhangzhou 363000, Fujian, China. E-mail: kaifa-wei@163.com

Manuscript Received: 29 December 2011; Manuscript Accepted: 7 March 2012

Accepted manuscript online in Wiley Online Library (wileyonlinelibrary.com): 13 March 2012

DOI 10.1002/jcb.24133 • © 2012 Wiley Periodicals, Inc.

and so on (See the supplementary Table 1 for more details). Additionally, the anti-Su antibodies from both human patients with rheumatic diseases and a mouse model of autoimmunity recognize the endonucleolytic AGO and Dicer proteins, and a later study declared that autoantibodies to the Su antigen are found in a variety of systemic rheumatic diseases, suggesting that AGO proteins may be involved in autoimmune diseases. Each of these evidences highlights the vital roles of AGO proteins associated with sRNAs in regulating physiological and pathophysiological processes.

As an effector in sRNA-mediated gene regulation, although AGO proteins are conserved between species, they have undergone remarkable structural evolution and functional diversification. With more and more AGO gene sequences and protein crystal structures available, a need has arisen for molecular evolutionary analysis of AGO proteins in a large number and wide range of species. To investigate the relationship between structure and function, evolution, and diversification may help us understand the adaptability and biodiversity of organisms, and contribute to further clarify the molecular mechanisms underlying the physiological processes. However, the previous research did not cover a wide range of species, and did not study in a systematic manner. In this article, we perform a comprehensive analysis in order to tackle these and many more issues from the visual angle of evolutionism.

RESULTS

DIVERSITY OF EUKARYOTIC AGO PROTEINS

The AGO proteins are widespread among prokaryotes, archaea, and eucaryotes, and there are multiple AGO proteins in most eukaryotic species. The increase of AGO gene copies implies their roles in adaptability to environment, and resistance to kinds of diseases. AGO protein is a multi-domain protein. Eukaryotic AGO is usually composed of four sections: N-terminal, PAZ domain, MID domain, and PIWI domain, while most archeal and prokaryotic AGO proteins do not contain PAZ domain, and the domain architecture is variable. For instance, *Burkholderia xenovorans LB400* AGO protein consists of five domains, Sir2, APAZ, L2, MID, and PIWI, while *Archaeoglobus fulgidus DSM 4304* AGO protein contains the L2, MID, and PIWI domains [Makarova et al., 2009]. Homologs of prokaryotic and archeal AGO proteins are conveniently referred to as pAGOs and aAGOs [Makarova et al., 2009]. In eukaryotes, AGO proteins can be classified into three paralogous groups based on the phylogenetic analysis: the Ago subfamily, the PIWI subfamily, and the worm-specific WAGO subfamily [Tolia and Joshua-Tor, 2007; Czech and Hannon, 2011]. Ago proteins are associated with siRNAs and miRNAs which direct RISCs towards appropriate target [Czech and Hannon, 2011], whereas the PIWI subfamily proteins bind piRNAs [Jinek and Doudna, 2009], and the WAGO subfamily members bind secondary siRNAs [Yigit et al., 2006]. In order to study the potential function and molecular basis underlying the physiological processes of AGO proteins in different subfamilies, 135 AGO protein sequences derived from 36 species covering prokaryotes, archaea, and eukaryotes were chosen from relational databases for further analysis.

Ago SUBFAMILY

Many proteins are clustered into metazoan Ago clade (Supplementary Fig. 1), showing that there are multiple Ago paralogs in one species, and the RISC variants can be distinguished by their Ago proteins. Among four human Ago genes, *Hs_AGO1*, *Hs_AGO3*, and *Hs_AGO4* are closely clustered together on chromosome 1, while the *Hs_AGO2* is located on chromosome 8 [Höck and Meister, 2008]. Hs_AGO1 has been proved to play a role in post-transcriptional gene silencing (PTGS), but it does not cleave mRNA efficiently [Janowski et al., 2006]. Instead, Hs_AGO2 is recruited to the noncoding transcript that overlaps the promoter during both gene silencing and activation [Chu et al., 2010]. Hs_AGO3 and Hs_AGO4 have been reported to interact with a similar spectrum of RNA transcripts, even though they do not function as slicer [Azuma-Mukai et al., 2008; Landthaler et al., 2008; Chu et al., 2010]. In *Drosophila*, there are five AGO proteins: Dm_AGO1, Dm_AGO2, Dm_AGO3, Piwi, and Aub [Hutvagner and Simard, 2008]. Dm_AGO1 and Dm_AGO2 belong to the Ago subfamily, while the rest of the Dm_AGOs are contained in germline-specific PIWI subfamily (Supplementary Fig. 1). Dm_AGO1 functions in miRNA processing, and is involved in miRNA mediated-translational inhibition without altering levels of the mRNA [Okamura et al., 2004]. Fly mutant embryos lacking Dm_AGO2 are defective in siRNA-directed RNAi but are still capable of miRNA-directed target mRNA cleavage [Okamura et al., 2004], indicating that Dm_AGO1 is primarily involved in miRNA mediated gene silencing while Dm_AGO2 is the catalytic core of RISC associated with RNAi [Okamura et al., 2004; Miyoshi et al., 2005]. Moreover, Dm_AGO1 and Dm_AGO2 are placed in the different evolutionary branches. The result implicates that the heritable variations have occurred during evolution and thereby aid in preserving functional differences. There are about 1–4 AGO isoforms in other mammals and insects as well. It is notable that Ago-containing animals (e.g., *Caenorhabditis elegans*, *Oikopleura dioica*, *Bombyx mori*, and *Chlamydomonas reinhardtii*) are presented in the same clade except human and fly. In fungus, *Schizosaccaromyces pombe* has only one AGO protein that possesses the slicer activity in siRNA pathway and participates in heterochromatic silencing. Especially, QDE2 of *Neurospora crassa* binding with QDE2-interacting RNAs (qiRNAs), with the aid of QDE2-interacting protein (QIP) exonuclease, can facilitate removal of the passenger strand of siRNA duplex from RISC [Maiti et al., 2007; Lee et al., 2009].

In plants, there is a wide diversity of sRNA-directed Ago protein functions, and specific member has distinct biochemical activities. *Arabidopsis* contains ten AGO proteins, which belong to three different phylogenetic branches shown in Supplementary Figure 1. At_AGO1, At_AGO5, and At_AGO10 are clustered in the first clade, At_AGO2, At_AGO3, and At_AGO7 are present in the second clade, and At_AGO4, At_AGO6, At_AGO8, and At_AGO9 belong to the third clade [Vaucheret, 2008]. By contrast, Ago proteins of *Oryza sativa* (*Os*), *Populus trichocarpa* (*Pt*), and *Physcomitrella patens* subsp. *Patens* (*Pps*) are distributed among three major phylogenetic clades. The Ago proteins of different species scattering in one clade suggest that multiple horizontal gene transfers (HGT) have occurred between different species. In *Arabidopsis* clade 1, At_AGO1 and At_AGO10 (ZWILLE) function in the miRNA and siRNA pathways and are essential for multiple processes in development [Mallory

et al., 2009]. Likewise, At_AGO5 shows preference for sRNAs but it tends to load those with a cytosine at the 5′ terminus [Takeda et al., 2008]. In clade 2, only At_AGO7 has slicer activity involved in silencing-based antiviral defense [Qu et al., 2008], together with miR390, functions at two distinct trans-acting siRNA (tasiRNA) biogenesis steps [Montgomery et al., 2008]. At_AGO2 and At_AGO3 are very similar to each other, suggesting that they derived from a recent gene duplication event, while both At_AGO1 and At_AGO2 have a crucial role in the host defence against viral infections [Harvey et al., 2011]. Thus, it can be concluded that At_ AGO3 may perform substantially the same function. In *Arabidopsis* clade 3, At_AGO4 depends on 24-nt siRNAs, acting like fission yeast siRNAs that are associated with Scp_Ago1, and is capable of guiding chromatin remodeling proteins to act in *cis* and cleave heterochromatin [Irvine et al., 2006; Vaucheret, 2008]. At_AGO6 specifically functions in the heterochromatin siRNA and TGS pathways, and the activity is partially redundant with that of At_AGO4 [Zheng et al., 2007; Vaucheret, 2008]. At_AGO9 acts to repress germ cell fate in somatic tissues [Singh et al., 2011]. Notably, it has been reported that At_AGO4, At_AGO6, and At_AGO9 are implicated in RNA-directed DNA methylation, all three AGO proteins can bind to 5′-adenosine of sRNAs, but each exhibited different preferences for sRNAs [Havecker et al., 2010]. At_AGO8 and At_AGO9 are similar in sequence, suggesting that they derived from a recent gene duplication. *At_AGO8* was determined to be a pseudogene, and expressed at levels lower than *At_AGO9* [Takeda et al., 2008].

PIWI SUBFAMILY

The second subfamily of AGO proteins is the PIWI subfamily. The Supplementary Figure 1 shows that the PIWI family genes are highly conserved during evolution, which are found in all animals examined so far but not in plants. Previous studies showed that PIWI subfamily proteins are involved in biogenesis and function of piRNA in animal gonads, and also participate in germline development and gametogenesis of many metazoan species [Thomson and Lin, 2009]. In this study, four out of eight human AGO proteins belong to PIWI family, including Hs_HILI, Hs_HIWI1, Hs_HIWI2, and Hs_HIWI3, which are encoded by genes distributed on chromosomes 12, 11, 22, and 8, respectively. As shown in Supplementary Figure 1, there are three PIWI proteins in mouse, such as Ms_MIWI, Ms_MIWI2, and Ms_MILI. And in flies, the homologs of PIWI proteins are Dm_AGO3, Dm_PIWI, and Dm_Aub. In three species described above, Hs_HILI, Hs_HIWI1, Ms_MILI, Ms_MIWI2, and Dm_PIWI are capable of promoting the formation of heterochromatin, while Dm_AGO3 and Dm_Aub can lead to RNA cleavage and initiate an amplification loop of piRNA biogenesis (called ping-pong circle) [Siomi and Siomi, 2008; Kim et al., 2009], but no studies have been performed on the rest. Recent studies demonstrated that Zebrafish has two PIWI proteins: Ziwi (zebrafish PIWI) and Zili (zebrafish PIWI like) [Houwing et al., 2008]. Ziwi is a cytoplasmic protein present in gonads and very early stage embryos, suggesting that Ziwi is maternally deposited. Zili is relocalized to the nucleus during germline development, like Dm_PIWI, is not present at the sites of the most intense DAPI-staining, and Zili-null animals are agametic [Houwing et al., 2008], suggesting that Zili might be required for germline/stem cell maintenance [Kim et al., 2009]. Most importantly, the over and ectopic expression of Hs_HIWI genes was shown to be highly correlated to several types of cancers [Chen et al., 2007]. The most direct evidence of the roles of Hs_HIWI in cancer is that inhibition of Hs_HIWI gene expression can repress cancer cell growth in cell culture [Liu et al., 2006; Thomson and Lin, 2009].

WORM-SPECIFIC WAGO SUBFAMILY

As discussed previously, AGO proteins have undergone several rounds of duplication and gene retention resulting in diversification of protein functions, especially in plants and metazoans. The largest number of expansion is in *C. elegans*, we performed a cluster analysis of 25 AGO amino acid sequences, except for C06A1.4 that has been certificated as pseudogene. The result is that 25 AGO proteins scatter in different subfamilies, such as Ce_Alg1, Ce_Rde1 in Ago subfamily, Ce_PRG1, and Ce_R09a1.1 in PIWI subfamily, CSR-1, Ppw-1, and other 14 AGO proteins in WAGO subfamily.

Of 25 AGO proteins, Ce_Rde1 functions as a scavenger protein by taking up sRNAs from many different sources, including the miRNA pathway. In the absence of Ce_Rde-1, the maturation of siRNA duplexes can be affected, but miRNAs are not influenced, indicating its significant roles in the biogenesis of siRNAs [Tabara et al., 1999; Steiner et al., 2009; Correa et al., 2010]. Ce_Alg1 and Ce_Alg2 are essential for the miRNA pathway, and the loss of their functions causes miRNA defects in the timing of animal development [Hutvagner and Simard, 2008]. Recently, it has been identified that Ce_Alg1 can interact with specific 3′-untranslated region (UTR) and coding exon sequence with cross-linking immunoprecipitation coupled with high throughput sequencing (CLIP-seq) [Zisoulis et al., 2010]. In PIWI subfamily, Ce_PRG1 localizes to P granules in germ cells and is required for successful spermatogenesis. Further, the mutation of Ce_PRG1 gene causes a marked reduction in expression of a subset of genes during spermatogenesis, and generates extensive defects in activation and fertilization. Moreover, Ce_PRG1 activity is most important to the piRNA–PIWI complex assembly required for fertility [Wang and Reinke, 2008]. Both Ce_PRG2 and Ce_PRG1 have similar sequences, and placed in the same evolutionary branch, suggesting the overlapping functions may be homologous. And in WAGO subfamily, most AGO proteins have diverged their catalytic motif to a degree, and endonuclease activity are probably impaired based on a multiple sequence alignment analysis (Supplementary Fig. 2). As Ce_Csr1/F20D12.1 is required for chromosome segregation, most Ce_Csr1 deletion homozygotes are sterile, and some hermaphrodites may produce few embryos with chromosome segregation defects identical to those observed in Ce_Csr1 RNAi embryos [Yigit et al., 2006]. The functional patterns of other worm-specific WAGO subfamily members are not characterized clearly, to further assess the biological roles of WAGO subfamily proteins need more experimental evidence.

EVOLUTION OF PROKARYOTIC AND ARCHAEAL AGO PROTEINS

Some previous works have reported that eukaryotic genomes may contain one or more AGO genes, few genomes of either bacterial or archaeal origin encode AGO or PIWI proteins [Kitamura et al., 2010]. Recently, a couple of AGO protein homologs have been reported in some archaeal species, indicating the possibility of existence of eukaryotic RNA silencing like pathways in archaea. As archaea

represents a novel life form placed between bacteria and eukarya [Londei, 2005; Li et al., 2010], archaeal AGO protein homologs may provide new insight into the origin and evolution of AGO proteins. Research in this article shows that archaeal and bacterial AGO proteins are in the same clade. Five of them belong to *Pyrococcus furiosus DSM*, *Methanopyrus_kandleri_AV19*, *Archaeoglobus fulgidus DSM*, *Halorubrum lacusprofundi ATCC 49239*, and *Methanosarcina acetivorans C2A*, respectively, and the rest is nine bacterial AGO proteins. It is worth noting that Tt_Ago is similar to the RNase H, and identified to be DNA strand-mediated site-specific RNA endonuclease responsible for defense responses against exogenous nucleic acids such as viruses and plasmids [Wang et al., 2008; Makarova et al., 2009]. AGO protein of archaeal strain *Archaeoglobus fulgidus* binds to DNA more tightly than it does to RNA, by contrast, eukaryotic AGO proteins use RNA as guide strand [Yuan et al., 2005]. Considering the fact that bacteria and archaeal AGO proteins scatter in the same clade, we can infer that multiple HGT events have occurred between bacteria and archaea genomes. Comparative analysis based on the protein secondary structures of the identified aAGOs and pAGOs shows that Pf_AGO is the only PAZ domain-containing protein among 14 proteins, which is also the first crystal structure of full length AGO protein [Jinek and Doudna, 2009]. As PAZ domain is a signature of eukaryotic AGO proteins, Pf_AGO provides us new clues about origin and evolution of PAZ domain-containing proteins.

STRUCTURAL EVOLUTION AND FUNCTIONAL DIVERSIFICATION OF AGO DOMAINS

As protein structures are more highly conserved than protein sequences, vestigial structures can be viewed as evidence for evolution, and the evolution of AGO structures is highly associated with the functional diversification. Even within a subfamily, each member seems functionally distinct, so this need to make certain why these AGO proteins do what they end up doing. Compared to phylogenetic analysis as discussed previously, it is more important to perform comparative studies of structural evolution and functional diversification. Hence, 29 AGO protein sequences from the phylogenetic tree (Supplementary Fig. 1) were chosen for further study, which represent the different subfamilies. As AGO proteins are defined by four sections: an N-terminal, PAZ, MID, and a C-terminal PIWI domain (Supplementary Fig. 3), in which PAZ, MID, and PIWI domain belong to main functional domains [Parker, 2010], we systematically examine the functional diversification of the three domains based on the structural evolution.

PAZ DOMAIN

PAZ domains (~140 residues) are found in both AGO proteins and Dicer enzymes, and structures of PAZ domains either in their isolated form or in complex with RNA have been determined using X-ray crystallography and NMR spectroscopy. The three-dimensional structure of isolated PAZ domain from Dm_AGO protein is like a deviant oligonucleotide/oligosaccharide-binding (OB) fold containing a central cleft lined with conserved aromatic residues which can bind specifically to single-stranded 3′ ends [Lingel et al., 2003; Hall, 2005]. That is to say, PAZ domain predominantly contacts the 3′-overhang-containing strand. Moreover, PAZ domain shows a preference for RNA over DNA and apparently lack of sequence dependence during recognition [Yan et al., 2003]. Figure 1A shows that the PAZ domain of PfAgo anchors the 2-nucleotide 3′-overhang of the siRNA-like duplex within a binding

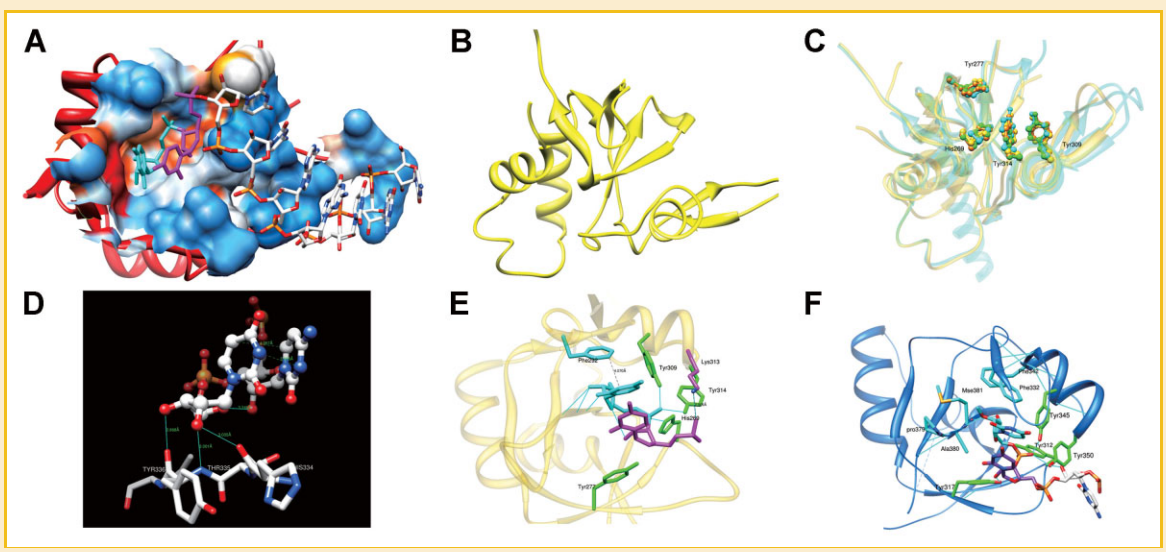

Fig. 1. Comparative analysis of structure and function of PAZ domains. A: Binding pocket of PAZ domain of Hs_AGO1 anchors the 2-nucleotide 3′overhang of the siRNA (PDB: 1SI3.pdb). B: Ribbon diagram representation of a homology model of Hs_AGO2 based on the structure of Hs_AGO1. C: The homology model of Hs_AGO2 PAZ domain, Hs_HILI (PDB: 3O7X.pdb), and Hs_ HIWI1 (PDB: 3O6E.pdb) are superposed onto the Hs_AGO1 PAZ domain (PDB: 1SI3.pdb). D: H334, T335, and Y336 in β sheet of Hs_AGO1 interact with ligand, and the green line represents hydrogen bond. E: Detailed view of the binding pocket in Hs_AGO1 (PDB: 1SI3.pdb). Conserved residues that are involved in 3′ end binding are shown in stick format, labeled by black, colored by green. The other two residues in PAZ domain interacting with the ligand are color-coded as follows: cyan, the residue interacting with the terminal nucleotide; aubergine, the residue interacting with the penultimate nucleotide, respectively. F: Detailed view of the binding pocket in Hs_HIWI1 (PDB: 3O6E.pdb). The color holds same significance as E, two residues interact with the terminal nucleotide. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

pocket composed of four highly conserved amino acids, such as Y212, Y216, H217, and Y190 [Song et al., 2004]. In human, being lack of Hs_AGO2 crystal structure, the reliable homology model of Hs_AGO2 protein shown in Figure 1B is constructed by threading the sequence of Hs_AGO2 onto the Hs_AGO1 structure using SWISS-MODEL (http://www.expasy.org/spdbv/; [Guex and Peitsch, 1997]). When the PAZ domain structures of Hs_HILI, Hs_HIWI1, and Hs_AGO2 are superimposed onto Hs_AGO1, the result shows that H269, Y277, Y309, and Y314 in Hs_AGO1 are structurally equivalent to H277, Y279, Y311, and H316 in Hs_AGO2. So it is with Y422, Y427, Y455, and Y460 in Hs_HILI, and Y312, Y317, Y345, and Y350 in Hs_HIWI1 (Fig. 1C). When combined with the primary structure analyses (Supplementary Fig. 4), we come to a conclusion that the positions of the conserved amino acid residues may vary in different species. At the bottom of binding pocket of Hs_AGO1, aromatic residue F292 stacks against the terminal nucleotide (Fig. 1E), while the functional site is usually occupied by the aromatic Phe in most organisms. Simultaneously, K313 contacts the penultimate nucleotide, but it is not so conserved as F292. These indicate that the terminal nucleotide has a rather fixed position, but the penultimate nucleotide shows a large flexibility change during the binding process [Lingel et al., 2004; Ma et al., 2004; Hall, 2005]. Further comparative analysis shows that only one aromatic residue F292 interacts with the terminal nucleotide in Hs_AGO1, but two aromatic residue F332 and F342 stack against the terminal nucleotide in Hs_Hiwi1 (Fig. 1E, F). For another, Hs_AGO1 PAZ domain favor 2′-OH over 2′-OCH_3 3′ ends during binding process, while the Hs_Hiwi1 PAZ-binding pocket is capable of accommodating 3′ ends that are either 2′-OH or 2′-OCH3 [Tian et al., 2011]. Obviously, methyl group has different roles in the binding process of 3′ ends to PAZ domains of Hs_AGO1 and Hs_HIWI1. The exact molecular mechanism remained elusive so far, we merely know that 2′-*O*-methylation of the piRNA 3′ terminus contributes to maintain stability of piRNA and makes it less susceptible to degradation [Zeng et al., 2011]. Our investigation into the structural basis of PAZ-nucleotide contaction showed that H334, T335, and Y336 in the β sheet of Hs_AGO1 also interact with the ligand (Fig. 1D, E). The same situation exists in the Hs_AGO2, Hs_AGO3, Hs_AGO4, Dm_AGO1, and Ce_Alg1. So do P379, A380, and M381 with the ligand in Hs_HIWI1 (Fig. 1F). All of them are able to form hydrogen bonds with 3′-overhanging ends of siRNA or miRNA, which may enhance the binding force.

An intriguing question is that why the PAZ domain is mainly found in eukaryotes, whereas most prokaryotes are short of PAZ domain? The PAZ-lacking pAGO does not have slicer activity, conversely, almost all PAZ-domain containing pAgos are predicted to be directly involved in antiviral responses by nuclease activity, such as destroy viruses or plasmid nucleic acids [Makarova et al., 2009]. The foregoing analysis may induce us to propose that the emergence of PAZ domain, which is previously described as an anchoring site for 3′-ends of guide RNA, may be the shared derived characters representing unique evolutionary events.

MID DOMAIN

MID domain is the second functional domain of AGO proteins, and is similar to Lac repressor [Friedman et al., 1995; Song et al., 2004]. Although it was reported that the MID domain contains the binding sites for the 5′ end of sRNAs, direct structural information on eukaryotic AGO domains, including crystal structure, are seldom available. The structural analyses of Hs_AGO2 and Nc_QDE2 show clearly that the 5′-terminal nucleotide of the guide molecule needs to adjust its conformation to accommodate a preformed pocket of the MID domain (Fig. 2). Analyses of receptor–ligand interactions demonstrate that residues surrounding the ligand are Y529, K533, Q545, and K570 in Hs_AGO2 MID domain shown in Figure 2A, while the residues participating in the interactions are Y595, K599, K634, and K638 in the MID domain of Nc_QDE2 (Fig. 2B). The study on the multiple sequence alignment (Supplementary Fig. 5) shows that the 5′ phosphate binding pockets are highly conserved. Another direct evidence comes from the conservation surface-mapping of Hs_AGO2 (Fig. 2C). Accordingly, the slicer activity of AGO proteins may be impaired by the mutation of one out of four residues that is involved in metal-ion coordination and 5′ phosphate binding [Ma et al., 2005; Jinek and Doudna, 2009]. The experimental study of Hs_AGO2 functional domain in complex with nucleoside monophosphates (AMP, CMP, GMP, and UMP) mimicking the 5′ ends of miRNAs showed that the specific interaction exists between 5′ nucleotide and a nucleotide specificity loop of the MID domain, and confirmed that eukaryotic MID domain has a distinct preference for U or A rather than C and G at the 5′ ends of miRNAs [Frank et al., 2010]. Obviously, the predilection of base may influence the integrity of the seed sequence and miRNA strand selection. Further analysis shows that five residues, which are P523, G524, K525, and T526, consist of a rigid loop of Hs_AGO2 MID domain. Among them, P523 and P527 are located at either end of the loop, contributing to maintain the specific loop conformation, and G524 functions as a sharp kink in its trajectory. Although the loop is absolutely conserved in all four human AGO proteins, as well as in the miRNA-associated Dm_ AGO1, Ce_ Alg1, and Ce_Alg2, it is not easy to extrapolate the results obtained from structural analysis of the Dm_AGO2 that has little overall sequence similarity with Hs_AGO2 [Frank et al., 2010]. Consequently, the conclusion can be drawn that the residues shown as black letters on cyan background may be the amino acids forming the special loop in other species (Supplementary Fig. 5).

Furthermore, as MC sequence region described in Hs_AGO2 is similar to eIF4E-like mRNA cap binding motif [Kiriakidou et al., 2007], this implies that MC sequence may compete with it on mRNA, resulting in translation repression and mRNA degradation. However, Kinch and Grishin [2009] mapped the MC sequence to the MID domain structure, found that the Hs_AGO2 MC region does not contain the motif, and proposed that the phosphates of the m7Gppp cap may adopt a similar conformation like the pattern of miRNA complementary to the binding sites in AGO2. It was supported by another evidence that the Nc_QDE2 MID domain has no structural similarity to eukaryotic eIF4E [Boland et al., 2010]. This binding mode suggests that AGO protein would compete with guide RNA, so there is no binding site for mRNA. In this study, Hs_AGO2 MID domain structure was superimposed onto Nc_QDE2 and the results showed that both are highly homologous with a root-mean-square deviation (RMSD) difference of 2.1 Å (Fig. 2D). Also, we find that the first sulfate (sulfate I) of Nc_QDE2 is coordinated by the highly

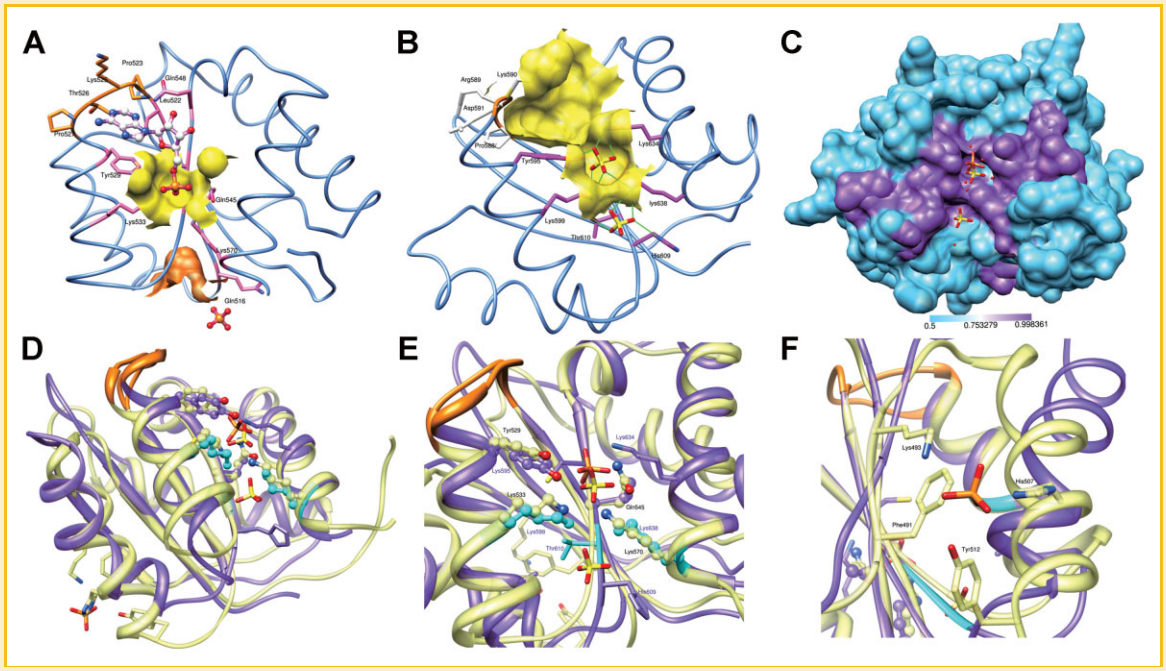

Fig. 2. Comparative analysis of structure and function of MID domains. A: The 3-D structure of Hs_AGO2 (PDB: 3LUD.pdb) including ligand AMP and phosphate ions (II) both shown as ball sticks (red: oxygen; orange: phosphorus). The two-ligand binding pockets are show in yellow and orange, respectively. Residues consisting of specificity loop are shown as stick, in orange. B: Ribbon diagram of the Nc_QDE2 MID domain (PDB: 2XDY.pdb) exhibits the positions of two binding sulfate ions (I and II) shown as stick (red: oxygen; yellow: sulphur). And the nucleotide specificity loop is highlighted in orange. C: The conservation surface mapping of Hs_AGO2. D: Superposition of Hs_AGO2 MID domain structure (PDB: 3LUC.pdb) onto Nc_QDE2 (PDB: 2XDY.pdb). E: The close-up of the position of the first phosphate of Hs_AGO2 ((PDB: 3LUC.pdb) and two sulfate of Nc_QDE2 (PDB: 2XDY.pdb), as well as their surrounding residues. The residues Lys599, Lys638 and Thr610 shared by two sulfate of Nc_QDE2 are highlighted in cyan. F: The close-up of the second phosphate position in Hs_AGO2 (PDB: 3LUC.pdb) and surrounding residues. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

conserved amino acids Lys595, Lys599, and Lys638, and it is located in the similar position as the 5′ phosphate of AMP observed in the Hs_AGO2 MID domain (Fig. 2E). Therefore, either the sulfate I bond in Nc_QDE2 or phosphate I bond in Hs_AGO2 is likely to represent the 5′-nucleotide binding site of miRNA. However, the positions of the second sulfate (sulfate II) and phosphate II are significantly different in both MID domains (Fig. 2D–F). In Nc_QDE2, sulfate II is 6.3 Å from sulfate I, shares coordination with Lys599 and Lys638, and is further coordinated by Thr610 [Faehnle and Joshua-Tor, 2010]. And sulfate II is bound in the side of the MID domain opposite from where the guide RNA extends from the 5′-nucleotide binding site, so sulfate II cannot compete with sulfate I [Faehnle and Joshua-Tor, 2010]. Except for the 5′-phosphate binding pocket for siRNA or miRNA, Djuranovic et al. [2010] proposed that Dm_AGO1 has a second miRNA-dependent site that can bind nucleotides such as the 5′-cap. The second ligand-binding site of Dm_AGO1 might be under allosteric control, whereas the other AGO proteins might be regulated by distinct ligand or might no longer be regulated by a second ligand [Boland et al., 2010]. According to this, it induces us to speculate that the second sulfate ion in Nc_QDE2 MID domain may occupy the second ligand binding site. And is there any similarity between the second ligand binding site and eIF4E-like mRNA cap binding motif? If so, the second site may be occupied by the phosphates of the m7Gppp cap. The results of crystallographic and NMR analysis of the Hs_AGO2 MID domain and full-length Hs_AGO2 come to the conclusion that the cap analogues do not bind significantly to the isolated MID domain, but the full-length AGO can bind the cap-like nucleotides or noncap-like nucleotides [Frank et al., 2011]. Consequently, the controversial view was proposed that the second site may be involved in binding with the 5′-cap, which may be another way to understand the mechanism of translational suppression. As discussed previously, the two ligand binding sites of Hs_AGO2 are different from that of Nc_QDE2, then what about other species?

PIWI DOMAIN

PIWI domain is the third and most important structure of AGO proteins, which is the catalytic center for rendering their target cleavage activity similar to RNase H. PIWI domain sequence alignment analysis shows that most of PIWI domains contain DDH motif composed of two aspartates and a histidine, which constitute a conserved catalytic core similar to the DDE catalytic triad in RNase H. The two aspartate residues are invariant whereas the third residue can be substituted by other residue (Supplementary Fig. 6). The overall tertiary structures of the PIWI domains are rather rigid requiring divalent metal ions for structural stability and catalytic activity, and B-factor analysis indicates that the PIWI domain of *Pyrococcus furiosus* is highly stable (Fig. 3). Then we construct an electron-density map of the DDH motif, which shows a direct static picture of charge distribution and occupied volume of Asp558,

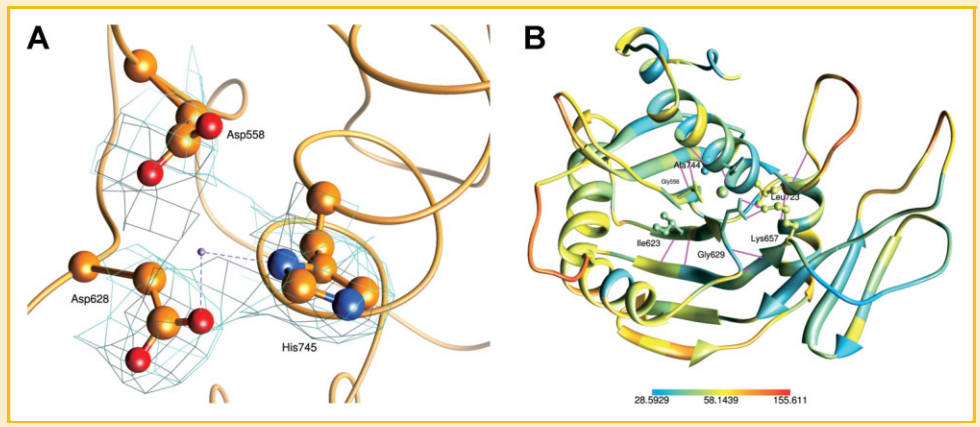

Fig. 3. Structual analysis of catalytic center in PIWI domain of *Pyrococcus furiosus*. A: The electron-density map of three highly conserved amino acids composed of DDH motif in catalytic center (PDB: 1z25.pdb) are shown in ball stick format and labeled. B: The B-factor analysis of PIWI domain (PDB: 1z25.pdb) and the six residues making up scaffold are labeled. [Color figure can be seen in the online version of this article, available at http://wileyonlinelibrary.com/journal/jcb]

Asp628, and His745, and the result demonstrates that the interaction interface is determined between the catalytic site and the target. The structural analysis showed that six residues, which are G556, I623, G629, K657, L723, and A744, may serve as a scaffold that stabilizes correctly folded conformation. So do other species (Supplementary Fig. 6). Surprisingly, some PIWI domains contain DDH motif, but do not have slicer activity. For example, only Ago2 in human four Ago proteins has slicer activity. And, furthermore, At_AGO1 and At_AGO7 have been identified as slicers that use either miRNAs or siRNAs as guide, and At_AGO4 has both catalytic and noncatalytic functions, whereas, the other seven At_AGOs have no catalytic activity [Qi et al., 2006; Höck and Meister, 2008]. All of the evidences actually suggest that some AGO proteins with the conserved catalytic residues may lose their slicer activity during evolution, but the specific mechanism remains unclear. We speculate that post-translational modifications, such as acetylation, methylation, phosphorylation, ubiquitination, and etc., may influence endonucleolytic activities.

In order to illustrate the mechanism of target recognition and cleavage mediated by AGO proteins, two models have been proposed [Filipowicz, 2005]. The two-state model assumed that both 5′ and 3′ ends of the guide are tethered to PAZ domain and PIWI domain, respectively, then, the 3′ end of the guide strand docking at the PAZ domain needs to release in order to make the guide 3′ region available for base pairing. The fixed-end model does not need recurrent disruption of guide-PAZ contact and both ends of the guide strand remain docked in their binding pockets during slicing. For example, being anchored would limit the seed sequence interaction with the target mRNA, and being released could eliminate the topological constraint in the binary complex, so the 5′ end is fixed in Tt_Ago, whereas the 3′ end of the guide switches on and off PAZ during the catalytic cycle [Parker, 2010]. Further crystal structure analysis of the full-length Tt_AGO protein in complex with guide DNA strand showed the molecular details of protein–RNA recognition [Wang et al., 2008]: a 5′-phosphorylated 21-base DNA guide strand is bound to Tt_AGO, and both ends are anchored in the pockets, meanwhile, R172 and R548 are able to lock guide bases 10 and 11 at the cleavage site [Wang et al., 2008]. Then, the seed region of the guide adopts a stacked, partially helical conformation, and pre-organize to initiate base pairing with the target. In the absence of target, the guide is kinked at the 10–11 base step. This indicates that further structural rearrangement is required during target-RNA recognition [Jinek and Doudna, 2009], which may orientate the target in the nucleic-acid-binding channel between the PAZ- and PIWI-containing lobes. Simultaneously, AGO protein undergoes a significant conformational change for better accommodation and catalytically activity [Ming et al., 2007].

DISCUSSION

As AGO proteins are linked to command center of major physiological processes, evolutionary and functional analysis of AGO proteins will help us understand how evolutionary stresses modified functional role of AGO proteins. Prokaryotic AGO proteins function as key components in novel defense system against exogenous genetic elements, and archea has apparent functional counterparts to the RNAi system. In contrast, eukaryotic AGO proteins play crucial roles in gene expression regulation that broadly participate in numerous biological processes, including developmental timing, cell differentiation, cell proliferation, cell death, metabolic control, immunity, and transposon silencing etc. And following the trail of protein signals mediated by AGO proteins, we often lead to the major command center of many physiological processes, especially link to molecular basis of many diseases formation (Supplementary Table 1), exactly as Hs_HIWI give us a new sight into PIWI subfamily proteins that not only are required in the maintenance of germline stem cell (GSC) but also have been shown to be involved in cancer formation.

In this article, we systematically chose 135 full-length AGO protein sequences, covering prokaryote, archaea, and eukaryote, for the molecular phylogenetic analysis. Phylogenetic analysis shows that many organisms have evolved multiple members of AGO proteins, and the number may vary across species, suggesting that

AGO genes have evolved by gene duplication followed by sequence divergence. As a result of environmental response, the specific member of the AGO proteins has distinct biochemical activities, and shows us the diversity of structure and function as well. And as for guider strand selection and recognition, prokaryotic AGO proteins select DNA as a guide strand, similarly, archeal AGO protein binds DNA more tightly than it does to RNA, while eukaryotic AGO proteins use RNA as guide. As described previously, PAZ domain bind to the 3′ overhanging ends of siRNA or miRNA, but most prokaryotic AGO proteins do not contain PAZ domain, and have no miRNAs, additionally, pAGO genes are adjacent to defence genes on chromosomes; although the proposed defense mechanism is not clear, we may infer that PAZ domain may be the product of evolution. MID domain contains two binding sites, the first site is proved to be 5′ nucleotide position of the guide strand, and the function of the second site has still been studying. Other issues, including whether the special ring exists and what of the binding preference of DNA guide to prokaryotic PIWI domains, are also discussed. It is entirely clear that AGO proteins associated with sRNAs play an irreplaceable and crucical role in physiological and pathophysiological processes. As described previously, overexpression of miR-17∼92 may activate the PI3K/AKT pathway and inhibit chemotherapy-induced apoptosis in MCL cell lines, and miR-17∼92 has been considered as a therapeutic target. For another, the stress down-regulated miRNAs may result in accumulation of their target mRNAs, which may positively regulate stress tolerance in plants. Although the exact mechanisms underlying the AGO protein–target complex are difficult to dissect at a molecular level, any progress being achieved in this field will lead us to the core physiological processes, and help us take effective measures to influence pathophysiological processes.

METHODS

The AGO protein sequences were downloaded from the NCBI database (http://www.ncbi.nlm.nih.gov/), the Uniport database (http://www.uniprot.org) and Worm base (http://www.wormbase.org/).The BLASTP search was done with default settings using eight human AGO proteins as the query. Homologous sequences to the query were identified using E-value < 0.01 as the cutoff threshold for the BLASTP output. One hundred and thirty five AGO protein sequences were chosen for the phylogenetic analysis. The species and its corresponding number of AGO protein sequence in parentheses are as follows: *Homo sapiens* (8), *Sus scrofa* (6), *Mus musculus* (7), *Gallus gallus* (3), *Drosophila melanogaster* (5), *Bombyx mori* (3), *Danio rerio* (3), *Xenopus tropicalis* (2), *Oikopleura dioica* (3), *Strongylocentrotus purpuratus* (1), *Echinococcus multilocularis* (1), *Caenorhabditis elegans* (25), *Nematostella vectensis* (2), *Giardia intestinalis* (1), *Populus trichocarpa* (10), *Oryza sativa* (19), *Arabidopsis thaliana* (10), *Physcomitrella patens subsp. Patens* (6), *Ectocarpus siliculosus* (1), *Chlamydomonas reinhardtii* (3), *Schizosaccaromyces pombe* (1), *Neurospora crass* (1), *Aquifex aeolicus VF5* (1), *Anoxybacillus flavithermus WK* (1), *Thermus thermophilus HB8* (1), *Archaeoglobus fulgidus DSM 4304* (1), *Methanosarcina acetivorans C2A* (1), *Pyrococcus furiosus DSM 3638* (1), *Syntrophobacter fumaroxidans MPOB* (1), *Pelotomaculum thermopropionicum SI* (1), *Methanopyrus_kandleri_AV19* (1), *Halorubrum lacusprofundi ATCC 49239* (1), *Aromatoleum aromaticum EbN1* (1), *Xanthomonas campestris str_B100* (1), *Mesorhizobium loti MAFF303099* (1) *and Rhizobium etli CIAT 652* (1).

Multiple sequence alignment was performed using ClustalX1.83 with default settings, and modified by the GeneDoc (http://www.nrbsc.org/gfx/genedoc/). The neighbor-joining tree of AGO family was constructed with 1,000 bootstrap replicates using the MEGA 5.0 program. Figtree 1.3.1 was used for graphical view of phylogenetic tree. The homology model building for Hs_AGO2 PAZ domain was constructed using the SwissModel Alignment Mode (http://swissmodel.expasy.org/). The crystal structure data of Hs_AGO1 including PDB (PDB: 1SI3.pdb), mmCIF and XML format are available from the Protein Data Bank on the website (http://www.pdb.org/pdb/home/home.do). All three-dimensional figures were generated by UCSF Chimera (http://www.cgl.ucsf.edu/chimera/).

REFERENCES

Azuma-Mukai A, Oguri H, Mituyama T, Qian ZR, Asai K, Siomi H, Siomi MC. 2008. Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. Proc Natl Acad Sci USA 105:7964–7969.

Boland A, Tritschler F, Heimstadt S, Izaurralde E, Weichenrieder O. 2010. Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. EMBO Rep 11:522–527.

Carthew RW, Sontheimer EJ. 2009. Origins and mechanisms of miRNAs and siRNAs. Cell 136:642–655.

Chen Y, Pane A, Schupbach T. 2007. Cutoff and aubergine mutations result in retrotransposon upregulation and checkpoint activation in *Drosophila*. Curr Biol 17:637–642.

Chu Y, Yue X, Younger ST, Janowski BA, Corey DR. 2010. 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 38:7736–7748.

Correa RL, Steiner FA, Berezikov E, Ketting RF. 2010. MicroRNA-directed siRNA biogenesis in *Caenorhabditis elegans*. PLoS Genet 6:e1000903.

Czech B, Hannon GJ. 2011. Small RNA sorting: Matchmaking for Argonautes. Nat Rev Genet 12:19–31.

Djuranovic S, Zinchenko MK, Hur JK, Nahvi A, Brunelle JL, Rogers EJ, Green R. 2010. Allosteric regulation of Argonaute proteins by miRNAs. Nat Struct Mol Biol 17:144–150.

Faehnle CR, Joshua-Tor L. 2010. Argonaute MID domain takes centre stage. EMBO Rep 11:564–565.

Filipowicz W. 2005. RNAi: The nuts and bolts of the RISC machine. Cell 122:17–20.

Frank F, Sonenberg N, Nagar B. 2010. Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. Nature 465:818–822.

Frank F, Fabian MR, Stepinski J, Jemielity J, Darzynkiewicz E, Sonenberg N, Nagar B. 2011. Structural analysis of 5′-mRNA-cap interactions with the human AGO2 MID domain. EMBO Rep 12:415–420.

Friedman AM, Fischmann TO, Steitz TA. 1995. Crystal structure of lac repressor core tetramer and its implications for DNA looping. Science 268:1721–1727.

Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18:2714–2723.

Hall TM. 2005. Structure and function of argonaute proteins. Structure 13:1403–1408.

Hammond SM, Bernstein E, Beach D, Hannon GJ. 2000. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature 404:293–296.

Harvey JJ, Lewsey MG, Patel K, Westwood J, Heimstadt S, Carr JP, Baulcombe DC. 2011. An antiviral defense role of AGO2 in plants. PLoS ONE 6:e14639.

Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, Schwach F, Doonan JH, Baulcombe DC. 2010. The *Arabidopsis* RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell 22:321–334.

Höck J, Meister G. 2008. The Argonaute protein family. Genome Biol 9:210.

Houwing S, Berezikov E, Ketting RF. 2008. Zili is required for germ cell differentiation and meiosis in zebrafish. EMBO J 27:2702–2711.

Hutvagner G, Simard MJ. 2008. Argonaute proteins: Key players in RNA silencing. Nat Rev Mol Cell Biol 9:22–32.

Irvine DV, Zaratiegui M, Tolia NH, Goto DB, Chitwood DH, Vaughn MW, Joshua-Tor L, Martienssen RA. 2006. Argonaute slicing is required for heterochromatic silencing and spreading. Science 313:1134–1137.

Janowski BA, Huffman KE, Schwartz JC, Ram R, Nordsell R, Shames DS, Minna JD, Corey DR. 2006. Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. Nat Struct Mol Biol 13:787–792.

Jinek M, Doudna JA. 2009. A three-dimensional view of the molecular machinery of RNA interference. Nature 457:405–412.

Kim VN, Han J, Siomi MC. 2009. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 10:126–139.

Kinch LN, Grishin NV. 2009. The human Ago2 MC region does not contain an eIF4E-like mRNA cap binding motif. Biol Direct 4:2.

Kiriakidou M, Tan GS, Lamprinaki S, De Planell-Saguer M, Nelson PT, Mourelatos Z. 2007. An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell 129:1141–1151.

Kitamura S, Fujishima K, Sato A, Tsuchiya D, Tomita M, Kanai A. 2010. Characterization of RNase HII substrate recognition using RNase HII-argonaute chimaeric enzymes from *Pyrococcus furiosus*. Biochem J 426: 337–344.

Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, Dinic L, Ojo T, Hafner M, Zavolan M, Tuschl T. 2008. Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA 14:2580–2596.

Lee HC, Chang SS, Choudhary S, Aalto AP, Maiti M, Bamford DH, Liu Y. 2009. qiRNA is a new type of small interfering RNA induced by DNA damage. Nature 459:274–277.

Li Y, Liu X, Huang L, Guo H, Wang XJ. 2010. Potential coexistence of both bacterial and eukaryotic small RNA biogenesis and functional related protein homologs in Archaea. J Genet Genomics 37:493–503.

Lingel A, Simon B, Izaurralde E, Sattler M. 2003. Structure and nucleic-acid binding of the *Drosophila* Argonaute 2 PAZ domain. Nature 426:465–469.

Lingel A, Simon B, Izaurralde E, Sattler M. 2004. Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. Nat Struct Mol Biol 11:576–577.

Liu X, Sun Y, Guo J, Ma H, Li J, Dong B, Jin G, Zhang J, Wu J, Meng L, Shou C. 2006. Expression of hiwi gene in human gastric cancer was associated with proliferation of cancer cells. Int J Cancer 118:1922–1929.

Londei P. 2005. Evolution of translational initiation: new insights from the archaea. FEMS Microbiol Rev 29:185–200.

Ma JB, Ye K, Patel DJ. 2004. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature 429:318–322.

Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ. 2005. Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. Nature 434:666–670.

Maiti M, Lee HC, Liu Y. 2007. QIP, a putative exonuclease, interacts with the Neurospora Argonaute protein and facilitates conversion of duplex siRNA into single strands. Genes Dev 21:590–600.

Makarova KS, Wolf YI, van der Oost J, Koonin EV. 2009. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. Biol Direct 4: 1–15.

Mallory AC, Hinze A, Tucker MR, Bouche N, Gasciolli V, Elmayan T, Lauressergues D, Jauvion V, Vaucheret H, Laux T. 2009. Redundant and specific roles of the ARGONAUTE proteins AGO1 and ZLL in development and small RNA-directed gene silencing. PLoS Genet 5:e1000646.

Meister G, Tuschl T. 2004. Mechanisms of gene silencing by double-stranded RNA. Nature 431:343–349.

Ming D, Wall ME, Sanbonmatsu KY. 2007. Domain motions of Argonaute, the catalytic engine of RNA interference. BMC Bioinform 8:470.

Miyoshi K, Tsukumo H, Nagami T, Siomi H, Siomi MC. 2005. Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. Genes Dev 19:2837–2848.

Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, Chapman EJ, Fahlgren N, Allen E, Carrington JC. 2008. Specificity of ARGONAUTE7-miR390 interaction and dual functionality in TAS3 trans-acting siRNA formation. Cell 133:128–141.

Okamura K, Ishizuka A, Siomi H, Siomi MC. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev 18:1655–1666.

Parker JS. 2010. How to slice: Snapshots of Argonaute in action. Silence 1:3.

Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ. 2006. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature 443:1008–1012.

Qu F, Ye X, Morris TJ. 2008. *Arabidopsis* DRB4, AGO1, AGO7, and RDR6 participate in a DCL4-initiated antiviral RNA silencing pathway negatively regulated by DCL1. Proc Natl Acad Sci USA 105:14732–14737.

Saito K, Siomi MC. 2010. Small RNA-mediated quiescence of transposable elements in animals. Dev Cell 19:687–697.

Singh M, Goel S, Meeley RB, Dantec C, Parrinello H, Michaud C, Leblanc O, Grimanelli D. 2011. Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. Plant Cell 23:443–458.

Siomi H, Siomi MC. 2008. Interactions between transposable elements and Argonautes have (probably) been shaping the *Drosophila* genome throughout evolution. Curr Opin Genet Dev 18:181–187.

Song JJ, Smith SK, Hannon GJ, Joshua-Tor L. 2004. Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305:1434–1437.

Steiner FA, Okihara KL, Hoogstrate SW, Sijen T, Ketting RF. 2009. RDE-1 slicer activity is required only for passenger-strand cleavage during RNAi in *Caenorhabditis elegans*. Nat Struct Mol Biol 16:207–211.

Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC. 1999. The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell 99:123–132.

Takeda A, Iwasaki S, Watanabe T, Utsumi M, Watanabe Y. 2008. The mechanism selecting the guide strand from small RNA duplexes is different among argonaute proteins. Plant Cell Physiol 49:493–500.

Thomson T, Lin H. 2009. The biogenesis and function of PIWI proteins and piRNAs: Progress and prospect. Annu Rev Cell Dev Biol 25:355–376.

Tian Y, Simanshu DK, Ma JB, Patel DJ. 2011. Structural basis for piRNA 2′-*O*-methylated 3′-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains. Proc Natl Acad Sci USA 108:903–910.

Tolia NH, Joshua-Tor L. 2007. Slicer and the Argonautes. Nat Chem Biol 3:36–43.

Vaucheret H. 2008. Plant ARGONAUTES. Trends Plant Sci 13:350–358.

Wang G, Reinke V. 2008. A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. Curr Biol 18:861–867.

Wang Y, Juranek S, Li H, Sheng G, Tuschl T, Patel DJ. 2008. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature 456:921–926.

Yan KS, Yan S, Farooq A, Han A, Zeng L, Zhou MM. 2003. Structure and conserved RNA binding of the PAZ domain. Nature 426:468–474.

Yigit E, Batista PJ, Bei Y, Pang KM, Chen CC, Tolia NH, Joshua-Tor L, Mitani S, Simard MJ, Mello CC. 2006. Analysis of the *C. elegans* Argonaute family reveals that distinct Argonautes act sequentially during RNAi. Cell 127:747–757.

Yuan YR, Pei Y, Ma JB, Kuryavyi V, Zhadina M, Meister G, Chen HY, Dauter Z, Tuschl T, Patel DJ. 2005. Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. Mol Cell 19:405–419.

Zeng L, Zhang Q, Yan K, Zhou MM. 2011. Structural insights into piRNA recognition by the human PIWI-like 1 PAZ domain. Proteins 79:2004–2009.

Zheng X, Zhu J, Kapoor A, Zhu JK. 2007. Role of *Arabidopsis* AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. EMBO J 26:1691–1701.

Zisoulis DG, Lovci MT, Wilbert ML, Hutt KR, Liang TY, Pasquinelli AE, Yeo GW. 2010. Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. Nat Struct Mol Biol 17:173–179.