nature COMMUNICATIONS

REVIEW ARTICLE

DOI: 10.1038/s41467-018-07449-7 OPEN

DNA interference and beyond: structure and functions of prokaryotic Argonaute proteins

Lidiya Lisitskaya[1], Alexei A. Aravin[1,2] & Andrey Kulbachinskiy[1]

Recognition and repression of RNA targets by Argonaute proteins guided by small RNAs is the essence of RNA interference in eukaryotes. Argonaute proteins with diverse structures are also found in many bacterial and archaeal genomes. Recent studies revealed that, similarly to their eukaryotic counterparts, prokaryotic Argonautes (pAgos) may function in cell defense against foreign genetic elements but, in contrast, preferably act on DNA targets. Many crucial details of the pAgo action, and the roles of a plethora of pAgos with non-conventional architecture remain unknown. Here, we review available structural and biochemical data on pAgos and discuss their possible functions in host defense and other genetic processes in prokaryotic cells.

Small noncoding RNAs are essential players in the control of gene expression and maintenance of genome stability in both prokaryotes and eukaryotes. In eukaryotes, several classes of small noncoding RNAs regulate gene expression and protect cells against exogenous and endogenous harmful genetic elements through specific recognition of complementary RNA targets, in a group of processes collectively called RNA interference (RNAi). RNAi pathways are diverse between species, and several distinct pathways can also operate within the same organism and even single cell[1]. Despite such diversity, all RNAi processes rely on a common core complex, composed of small guide RNA tightly bound to a protein from the Argonaute (Ago) family[2–7] (Box 1). This complex (sometimes called RNA-induced silencing complex, RISC) recognizes complementary RNA targets and either directly cleaves them through endonuclease activity of Ago (slicer activity) or performs other functions—such as cleavage-independent RNA destabilization, repression of translation and transcription—by interacting with other proteins[7,8].

Analysis of prokaryotic genomes revealed broad distribution of Ago proteins in both archaea (~30% of all sequenced genomes) and bacteria (~10% of genomes)[9–11]. Remarkably, pAgos are much more diverse than eAgos, and the latter form only a small branch on the pAgo tree suggesting their origin from pAgos[9,10]. Structural and biochemical studies of pAgos, in particular, from thermophilic prokaryotes, revealed a detailed pathway of guide binding, target recognition and slicer activity that provided crucial insight into the molecular mechanisms of RNAi in eukaryotes. However, until recently no information about the functions of these proteins in their prokaryotic hosts was available and their natural nucleic acid partners in the cell were unknown. Here, we review available data on the complexes of various pAgos with nucleic acids, and describe known biochemical activities of pAgos. We further discuss their emerging role in the genome defense against foreign genetic elements and hypothesize that they may

[1] Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia. [2] Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA. Correspondence and requests for materials should be addressed to A.A.A. (email: aaa@caltech.edu) or to A.K. (email: akulb@img.ras.ru)

perform additional functions in the regulation of genetic processes (e.g., DNA transcription, replication and repair) in the prokaryotic cell.

Structural organization of Ago proteins

All eAgo proteins contain six structural segments, including N-terminal, L1 (Linker 1), PAZ (PIWI–Argonaute–Zwille), L2 (linker 2), MID (Middle) and PIWI (P-element Induced Wimpy Testis) domains (Fig. 1). pAgo proteins have diverse structures and can be divided into two large phylogenetic groups[9,10,12,13]. One group, denoted long pAgos, predominantly includes pAgos that contain all domains present in eukaryotic proteins, although some members of this group (e.g., AfAgo) have lost the N-PAZ domains (Fig. 1)[9,10,12]. The second group of so-called short pAgos harbors proteins that have only MID and PIWI domains. All studied eAgos and long pAgos have a bilobal structure, consisting of the N-PAZ and MID-PIWI lobes, with nucleic acids—the guide and the target—accommodated between the lobes (Fig. 1). The catalytic site is formed by the RNaseH fold of the PIWI domain; it is located in the middle of the nucleic acid binding cleft and binds two divalent metal ions for catalysis. Many pAgos, including all short pAgos, contain substitutions of essential catalytic residues suggesting that they lack endonucleolytic activity. The genes of inactive pAgos often adjoin to genes encoding putative nucleases that were proposed to play a role in biogenesis of nucleic acid guides and/or repression of their genetic targets. The genes that are next to short pAgos also always contain the APAZ ("analog of PAZ") domain of unknown functions[9,10].

Biochemical and structural studies of pAgo proteins from several species revealed that they can bind either DNA or RNA guides but, in contrast to eAgos, preferably recognize DNA targets (Table 1 and references therein). Several pAgos were also shown to cleave RNA targets but the functional significance of this activity remains unknown (Table 1). To date, structural models of pAgo proteins and their complexes with guide and target nucleic acids were obtained for seven proteins, including DNA-guided (AfAgo[14–17], AaAgo[18–20], MjAgo[21–23], PfAgo[24–26], and TtAgo[27–33]) and RNA-guided (RsAgo[34,35] and MpAgo[36,37]) pAgos (see Table 1 for pAgo abbreviations). The most complete structural information was obtained for TtAgo that was crystallized with guide (DNA) and target (DNA or RNA) molecules at different steps of its functional cycle. The compendium of all published structures of pAgos and the summary of their functional properties are presented in Supplementary Fig. 1, 2, 4, Table 1 and Supplementary Table 1. For comparison, we also include eukaryotic Argonautes KpAgo[38] (yeast *Kluyveromyces polysporus*), hAgo1[39], hAgo2[38,40–42], and hAgo3[43] (human) from the AGO-clade and SIWI[44] (silkworm *Bombyx mori*) from the PIWI-clade, all the eAgos for which three-dimensional structures have been determined to date. Below, we outline common features and structural variations observed for these proteins.

Box 1

Based on sequence and structural features, eukaryotic Ago (eAgo) proteins can be divided into two main clades, AGO and PIWI, with additional clades formed by worm-specific and trypanosomal Ago proteins[10,113–115]. Three main classes of small RNAs interacting with eAgos include microRNAs (miRNAs), small interfering RNAs (siRNAs), and PIWI-interacting RNAs (piRNAs). miRNAs and siRNAs share a partly common biogenesis pathway that relies on endonuclease Dicer that processes longer precursors to make mature small RNAs[116–118]. While siRNAs are processed by Dicer from exogenous or endogenous double-stranded RNA in the cytoplasm, miRNAs are processed from hairpin (pseudo-dsRNA) precursors that first have to be cleaved by another endonuclease, Drosha, in the nucleus[117,119,120]. piRNAs, found in Metazoa, are processed from longer precursors that—in contrast to miRNA and siRNA—can be single-stranded, and their biogenesis is independent of Dicer and Drosha[76–78,121]. miRNAs and siRNAs associate with AGO-clade Ago proteins. In some species like *Drosophila* miRNAs and siRNAs are predominantly sorted into different Ago proteins[122,123], while in others including mammals they can be loaded into the same protein. piRNAs associate with the PIWI clade Ago proteins[76,78,124–126]. The resulting effector complexes then regulate the expression of host genes, suppress transposons or combat viral infection[87,127–129].

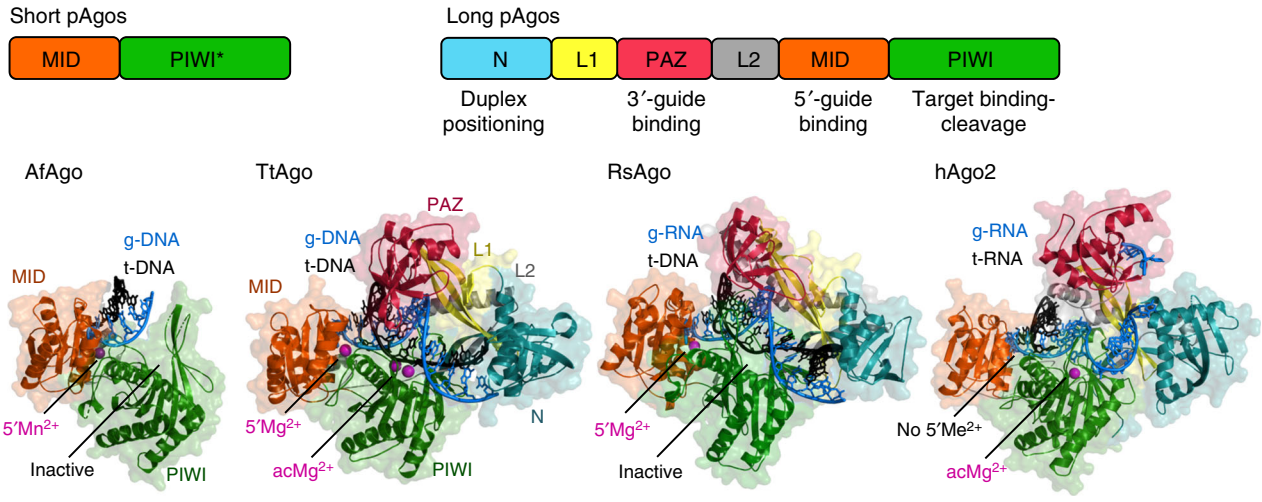

Fig. 1 Structural organization of Ago proteins. The domain architecture of short and long pAgos is schematically illustrated at the top. Short pAgos always contain inactive PIWI domain (PIWI*). The structures of four representative Ago proteins are shown in ternary complexes with guide ("g-") and target ("t-") nucleic acids: short inactive AfAgo (PDB: 2W42[17]) and long active TtAgo with g-DNA and t-DNA (PDB: 4NCB[30]), long inactive RsAgo with g-RNA and t-DNA (PDB: 5AWH[35]) and active human Ago2 with g-RNA and t-RNA (PDB: 4W5O[54]). The N-domain is turquoise, L1 is yellow, PAZ is magenta, L2 is gray, MID is orange, PIWI is green. The guide strand is blue, the target strand is black. Metal ions bound in the MID-pocket (5′Me^{2+}) or in the active center (acMe^{2+}) are indicated

Table 1 Functional properties of analyzed pAgo proteins in comparison with eAgos

Host	Argonaute	Guide	5′-end; nucleotide preference	Target	Catalytic activity	Functional activity	References
Aquifex aeolicus	AaAgo	DNA	5′-P; Unknown	(RNA; DNA not tested)	Guide-dependent	–	18–20
Archaeoglobus fulgidus	AfAgo (short pAgo)	DNA (RNA)	5′-P; Unknown	DNA (RNA)	Inactive	–	14–17
Marinitoga piezophila *Thermotoga profunda*	MpAgo TpAgo	RNA	5′-OH; None	DNA (RNA)	Guide-dependent	–	36,37
Methanocaldococcus jannaschii	MjAgo	DNA	5′-P; Purines	DNA	Guide-dependent; Chopping	Reduced plasmid content and transformation efficiency	21–23
Pyrococcus furiosus	PfAgo	DNA	5′-P; None	DNA	Guide-dependent; Guide-independent	Reduced transformation efficiency	24–26
Rhodobacter sphaeroides	RsAgo	RNA[a]	5′-P; g1U	DNA[a]	Inactive	Reduced transcription of reporter genes and plasmid content	34,35
Thermus thermophilus	TtAgo	DNA[a]	5′-P; g1C/t1′G	DNA[a] (RNA)	Guide-dependent; Chopping	Reduced plasmid content and transformation efficiency; changes in gene expression	27–33
Homo sapiens	hAgo1	RNA	5′-P; g1U or g1A	RNA	Inactive	miRNA pathway	39
	hAgo2	RNA	5′-P; g1U/t1′A	RNA	Guide-dependent	miRNA pathway	38,40–42
Kluyveromyces polysporus	KpAgo	RNA	5′-P; g1U	RNA	Guide-dependent	miRNA pathway	38
Bombyx mori	SIWI	RNA	5′-P; g1U	RNA	Guide-dependent	piRNA pathway	44

[a]Both in vitro and in vivo

The catalytic cycle of Ago proteins

The main steps in the catalytic cycle of Ago proteins established in vitro include guide binding, target recognition and annealing, target cleavage and target release (Fig. 2). These steps are likely similar for catalytically active eAgos and pAgos, however, the activity cycles of various pAgo proteins may include additional functional steps, as discussed below. Catalytically inactive Agos do not cleave their targets but are similar to the active Agos in guide binding and target recognition. Molecular mechanisms of the Ago action have been covered by several recent reviews[10,12,45–47]. We therefore briefly overview the main steps of the catalytic cycle of Ago proteins with particular emphasis on pAgos.

Guide binding and target recognition. All studied Ago proteins bind guide nucleic acid molecules (18–21 nt in analyzed structures) in a similar way, with the 5′- and 3′-ends of the guide fixed in protein pockets formed by the MID and PAZ domains, respectively (Figs. 1, 2, Supplementary Fig. 1 and Fig. 2). Analysis of eAgos suggested that the guide is subdivided into several functional segments, including the 5′ (anchor) nucleotide, the seed region (nucleotides 2–8), the site of cleavage (positions 10–11), the 3′ supplementary site (positions 12–16) and the 3′ tail, and a similar subdivision likely occurs in pAgos (Fig. 2)[12,47–50].

Several studied pAgos have preferences for specific 5′-nucleotides in the guide molecule (Table 1), including RsAgo (prefers 5′-uridine guides, similarly to hAgo2, KpAgo and SIWI[34,35,38,40,44]), TtAgo (5′-cytosine)[31] and MjAgo (5′-purines)[22]; other pAgos (MpAgo, PfAgo) show no 5′-end specificity[26,36]. In ternary complexes, the 5′-guide residue remains unpaired with the target and the corresponding target nucleotide (t1) can be bound and specifically recognized in a separate pocket in the PIWI domain (t1′G for TtAgo[33], t1′A for hAgo2[41], and RsAgo[51]).

Most pAgos bind 5′-phosphorylated guides and use a Mg^{2+} ion bound in the MID pocket for interaction with the first guide phosphate (Fig. 1, Supplementary Fig. 1; Table 1 and Supplementary Table 1). In contrast, eAgos of the AGO clade rely on a conserved lysine residue for the 5′-phosphate binding[38–40,52,53]. Unexpectedly, recent structural analysis of the silkworm SIWI protein from the PIWI clade of eAgos revealed that its MID pocket is similar to pAgos, with the Mg^{2+} ion involved in guide interactions (Supplementary Fig. 1)[44]. In contrast to other Agos, MpAgo binds unphosphorylated 5′-OH-guides and has a more hydrophobic pocket without metal ions or positively charged residues[36,37].

In all Ago-guide complexes, several nucleotide bases from the seed region are preoriented in a helical conformation and exposed to the solution (positions 2–4 to 2–6 in various Agos[22,27,36,40]). Initial target pairing with this region induces conformational changes that expose downstream nucleotides for further target recognition (Fig. 2)[29,30,54]. The downstream part of the seed region (positions 6–8) is kinked in available structures, depending on the geometry of the nucleic acid binding cleft (Supplementary Fig. 1)[22,28,36,38–40,52]. In eukaryotic Ago2, the resulting subdivision of the seed is important for the stepwise target recognition[55], and a similar role for guide kinking was proposed for pAgos[12].

The 3′-proximal part of the guide, except few last nucleotides that are bound in the PAZ pocket, is disordered in all binary Ago-guide complexes suggesting that it is structurally flexible (Figs. 1, 2). This includes the 3′ supplementary site that plays an important role in the recognition of mRNA targets by eAgos[48–50] and of DNA targets by analyzed pAgos (e.g., MpAgo[37] and RsAgo[51]). Such flexibility may likely facilitate helix formation during target annealing.

The 3′-end of the guide is bound in the PAZ pocket in binary complexes but is extruded upon target annealing (Fig. 2, Supplementary Fig. 1 and 2)[21,29,30,37,56]. For TtAgo, the 3′-guide release was observed after formation of a 12 bp g-DNA/t-RNA duplex or a 16 bp g-DNA/t-DNA duplex (Supplementary Fig. 2), suggesting that these processes are tightly coordinated and depend on the structure of the target strand[29,30]. Indeed, the guide-PAZ interactions are important for specific target recognition[21,56,57] and may also prevent guide degradation by cellular nucleases[57]. The PAZ pocket may exhibit certain preferences toward 3′-guide nucleotides in some pAgos (e.g., pyrimidine bases in MjAgo)[22] but the functional importance of this remains to be investigated. The conformational mobility of the PAZ domain (indicated with arrows in Fig. 2, Supplementary Fig. 1) may also contribute to the ability of various Agos to interact with populations of short RNAs or DNAs with different length distributions. At the same time, some pAgos have an incomplete PAZ domain (RsAgo[35,51], MpAgo[36]) or completely

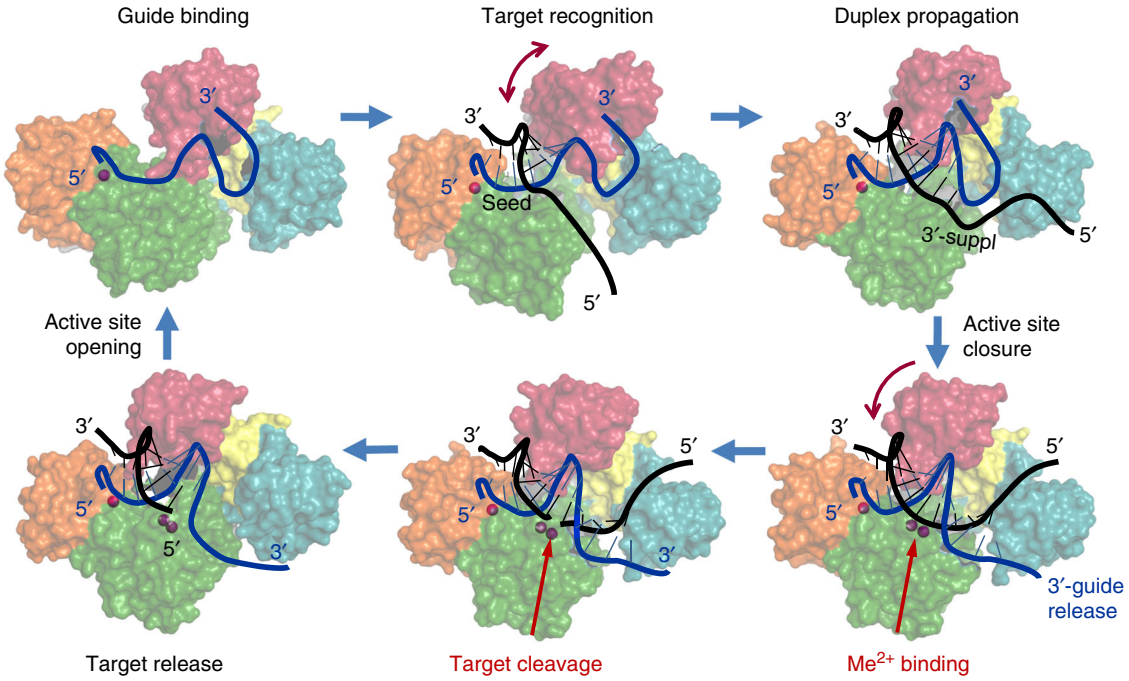

Fig. 2 The catalytic cycle of Ago proteins. Guide-loaded Ago performs search for a complementary target through base-pairing with the seed region of the guide strand, followed by duplex propagation through the central part and the 3′-supplementary site of the guide, thus checking for possible mismatches. Conformational mobility of the PAZ domain (shown by arc-shaped arrows) likely facilitates correct base-pairing, through controlled release of the guide 3′-end and active site closure. Conformational changes in the active site allow binding of catalytic metal ions, followed by cleavage of the target strand and its stepwise release from the complex. The drawings are based on the structures of TtAgo at different steps of its functional cycle (PDBs, from the upper left corner, clockwise: 3DLH, 3F73, 4N41, 4NCB, 4NCA, 4N76, see Supplementary Fig. 2). The guide strand is blue, the target strand is black; only the target strand of DNA substrate is shown (the structure of complexes with double-stranded DNA remains unknown for any pAgo)

lack it (short pAgos, such as AfAgo, Fig. 1). It remains to be established whether additional proteins may be involved in 3′-guide interactions in such pAgos.

Catalysis and target release. The binding of complementary nucleic acid target is accompanied by structural changes of the Ago molecule that include rotations of the PAZ domain and changes in the conformations of several loops in the PIWI domain, resulting in closure of the nucleic acid duplex within the catalytic cleft of pAgo and activation of catalysis, as described below (Figs. 2, 3)[27,29,30,36,37].

The catalytic site of all active Ago proteins contains a conserved tetrad of negatively charged amino acid residues, DEDX (where X is D, H, or K) that chelate catalytic divalent metal ions, Mg^{2+} or Mn^{2+} (Fig. 3)[9,10]. Mn^{2+} usually increases pAgo activity, and some pAgos (PfAgo) were shown to be active only in the presence of manganese ions[26,29,31,36]. The catalytic glutamate residue is located in the so-called glutamic finger that can adopt different conformations. In the absence of a target, it is located away from the catalytic site ("unplugged"), the complete tetrad is not formed, and no metal ions are bound in the active site (or only a single ion is observed) (Fig. 3, Supplementary Fig. 2; Supplementary Table 1)[22,28,36]. Formation of the extended guide-target duplex is accompanied by its closure within the nucleic acid binding cleft of pAgo, due to conformational changes in the PIWI and PAZ domains (indicated with red arrows in Fig. 3), insertion of the glutamic residue into the active site ("plugged in" conformation), binding of catalytic metal ions and activation of target cleavage (Fig. 3 and Supplementary Fig. 2)[29,30].

Catalytically inactive pAgos, such as RsAgo, contain substitutions of one or more negatively charged residues in the active site[9,10]. In addition, RsAgo remains in the unplugged conformation even after ternary complex formation, which also prevents catalytic metal binding (Supplementary Fig. 3)[35,51]. In contrast to pAgos, the catalytic site of the AGO-clade eAgos (hAgo2, hAgo3, KpAgo) was always found in the "plugged in" conformation, independently of the guide and target binding (Supplementary Fig. 3)[38,40,43,53,54]. At the same time, the PIWI-clade SIWI protein adopted the unplugged conformation in the absence of a target, suggesting that it may be more closely related to pAgos[44,58] (see below).

For most studied catalytically active Agos, the target is cleaved precisely between positions complementary to the 10th and 11th nucleotides of the guide strand (Figs. 2, 3)[18,22,23,27,29,36]. Intriguingly, more than one cleavage site was observed for MjAgo[21–23] but the structural basis for this remains unknown. Analysis of catalytically active eukaryotic and prokaryotic Ago proteins demonstrated that they are multiple turnover enzymes. Target release was shown to be the rate-limiting step in the action of eAgo proteins, due to persisting complementary guide-target interactions after target cleavage[25,48,49,59,60]. Mismatches in both the seed region and the 3′-supplementary guide site increase the enzyme turnover, although at the cost of decreased target binding[48,49,60]. At the same time, target release is not rate limiting for catalysis by the thermophilic TtAgo protein[29], for which the high temperature used in the assays likely promotes target dissociation. It remains to be established whether other protein factors may assist target release for pAgos from mesophilic prokaryotes.

A structural insight into the process of target release was obtained from the analysis of a ternary complex of TtAgo that was incubated at high temperature after target cleavage before crystallization (Figs. 2, 3, Supplementary Fig. 2)[30]. As revealed in the structure, the cleaved 5′-part of the target strand has dissociated from TtAgo and the corresponding 3′-portion of

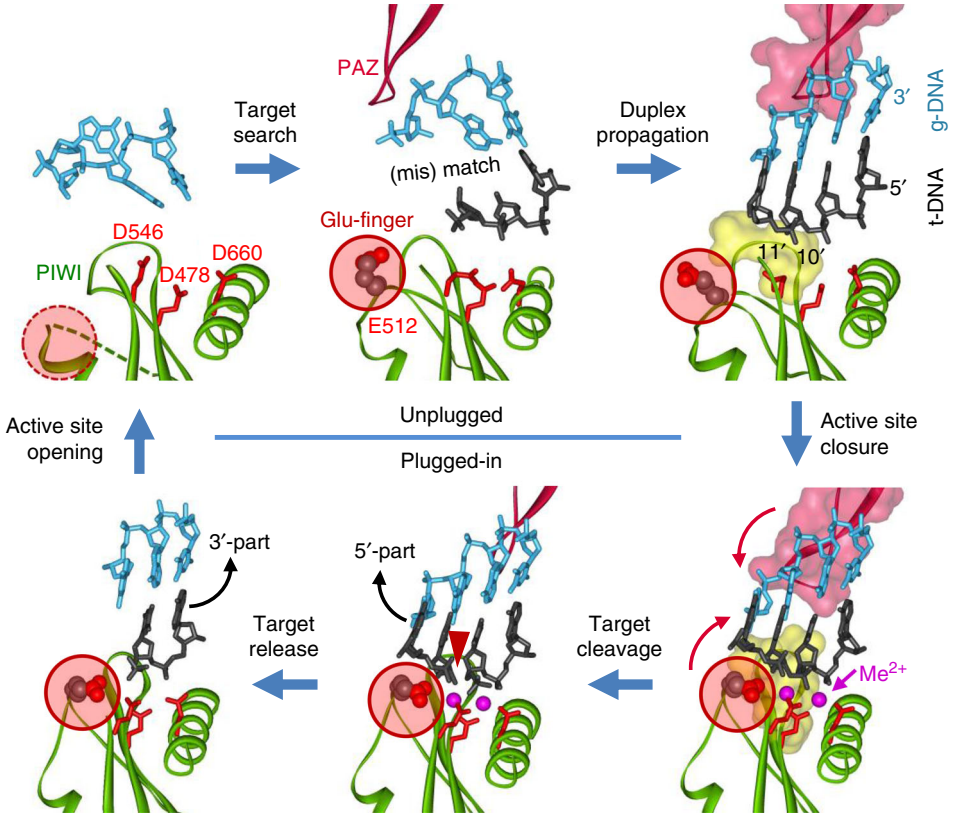

Fig. 3 Conformational changes in the active site of TtAgo during target recognition and catalysis[29,30]. The active site residues are shown in red; the glutamate finger is indicated with a red circle. During first steps of guide binding and target recognition, the active site is unplugged (upper raw); duplex propagation is accompanied by changes in the conformations of the PIWI and PAZ domains (indicated with red arrows), plugging-in of the glutamate finger, catalytic metal binding, and activation of catalysis (lower raw). Finally, stepwise target release leads to unplugging of the active site, thus making possible recognition of the next target molecule. The PDB accession numbers (from the upper left corner, clockwise): 3DLH, 3F73, 4N41, 4NCB, 4NCA, 4N76 (see Fig. 2 and Supplementary Fig. 2)

the guide is disordered (Fig. 3, bottom left). FRET measurements demonstrated that dynamic 3′-guide re-association with the PAZ pocket likely promotes target release[56]. This is likely followed by dissociation of the 3′-part of the target strand and unplugging of the active site, thus regenerating the binary guide-pAgo complex for the next round of catalysis. Analysis of eukaryotic Agos revealed the same sequential pathway of target dissociation, which can change depending on the presence of mismatches in the seed and 3′-supplementary guide sites[48,49].

Recognition of mismatched vs. matched targets. In eukaryotes, the efficiency of target repression by Ago-containing effector complexes greatly depends on the extent of complementarity between the guide and target RNAs[61–67]. Although a possible functional importance of the mismatched target recognition by pAgos remains unknown (in the context of their cellular functions discussed below), their further analysis may shed light onto the mechanisms of target recognition and various silencing pathways in both prokaryotes and eukaryotes.

Mismatches in the seed region between miRNAs and siRNAs and their targets have the most deleterious effects on the efficiency of silencing in eukaryotes[48–50,68–70]. Similarly, mismatches and bulges within the seed region significantly impair target binding and cleavage by studied pAgos[17,27,37,51,71]. No information on the structure of mismatched complexes is available for eAgo proteins. However, recent studies unexpectedly revealed that TtAgo and RsAgo can accommodate helical imperfections within the seed region in ternary complexes with only moderate structural perturbations (Fig. 4 and Supplementary Fig. 4)[51,71]. It was shown that purine-purine mismatches in the seed region can be bound without significant distortions of the duplex (e.g., mm A3-A3′, mm G8-A8′, mm A8-G8′ for RsAgo, Fig. 4). Nucleotide bulges in the guide strand in the ternary complexes of TtAgo stack-in between adjacent bases resulting in local distortions of the double helix (e.g., bulges g-4-A-5 and g-7-T-8, Fig. 4). In contrast, bulges in the target strand, which is more solvent-exposed, were shown to be looped-out of the duplex (e.g., bulges t-6′-A-7′ and t-9′-U-10′ for TtAgo, t-3′-AA-4′ for RsAgo, Fig. 4), resulting in stronger helix distortion and, in some cases, shifting of the cleavage site[51,71].

Intriguingly, the presence of bulges or mismatches in the seed region was shown to stimulate release of the imperfect guide-target hybrid from RsAgo, thus providing a mechanism for rapid guide exchange and Ago recycling[51]. Similarly, it was recently shown that mismatches in the seed region promote unloading of miRNAs from human Ago2, suggesting that such mechanism of guide exchange may be conserved in evolution[72].

Mismatches and bulges around the active site greatly decrease the efficiency of target cleavage by most studied eAgos[48,49,68,70] and pAgos alike[27,37,71]. From the structural perspective, mismatches at the cleavage site disrupt protein-nucleic acid interactions in the ternary complexes of TtAgo (in some mismatched complexes, the downstream part of the duplex is completely disordered) and the active site remains in the open unplugged conformation (Fig. 3, Supplementary Figs. 2 and 4)[27,29,71]. Thus, formation of the perfect guide-target duplex in the active site is a critical checkpoint in the specific target

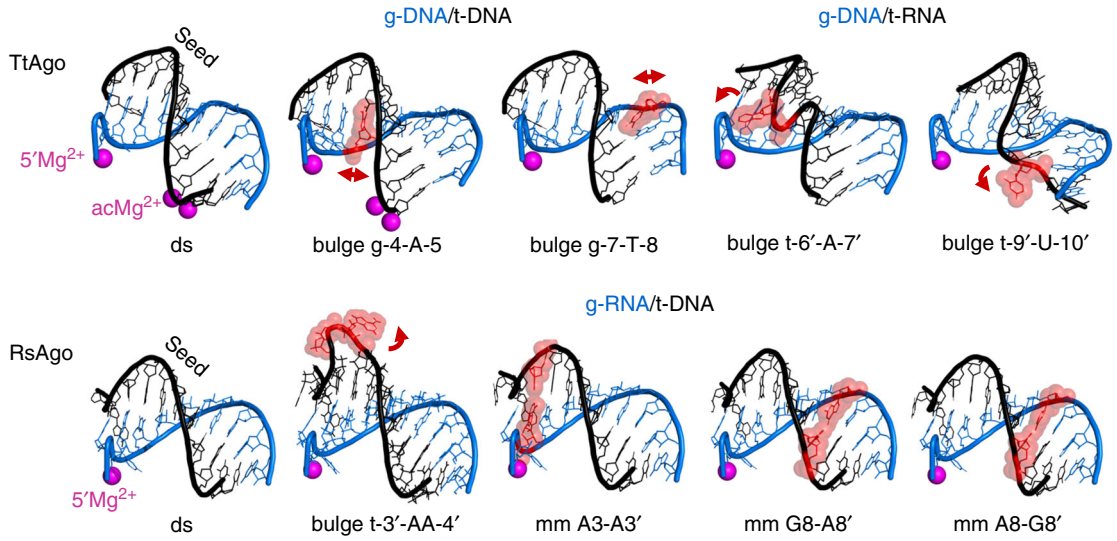

Fig. 4 Accommodation of helical imperfections in the ternary complexes of pAgo proteins. Structural features of the duplexes formed in the seed region in ternary complexes of TtAgo[71] (upper raw) and RsAgo[51] (bottom) containing bulges or mismatches (shown in red) in the guide or target strand, in comparison with fully double-stranded duplex ("ds"). Only the part of the duplex between the guide 5′-end and the active site in the PIWI domain is shown (guide positions 1 through 10–12 for various complexes); Mg^{2+} ions bound in the MID-pocket ($5'Mg^{2+}$) and in the active site ($acMg^{2+}$) are indicated; some complexes of TtAgo were obtained with a catalytically inactive mutant and thus lack catalytic metal ions. The distortions of the double-helix are shown with red arrowheads; the nucleotide bulges can be either stacked-in (bulges in the guide strand; g-4-A-5 and g-7-T-8 in TtAgo) or flipped-out of the duplex (bulges in the target strand; t-6′-A-7′, t-9′-U-10′ for TtAgo, t-3′-AA-4′ for RsAgo). The ternary complexes were obtained with g-DNA/t-DNA or g-DNA/t-RNA for TtAgo, or g-RNA/t-DNA for RsAgo, as indicated. The PDB accession numbers are (from left to right): TtAgo, 4NCB, 5XP8, 5XOU, 5XOW, 5XPA; RsAgo, 6D8P, 6D8A, 6D92, 6D9L, 6D9K. See Supplementary Table 1 and Supplementary Fig. 3 for full description of each complex

cleavage by Ago proteins, and the presence of helical imperfections hampers structural transitions required for activation of catalysis.

Functional activities of pAgos

It was initially proposed that pAgos might provide defense against foreign genetic elements such as transposons, phages and plasmids[9]. This hypothesis has found experimental support in recent studies of two long pAgos, catalytically active TtAgo and inactive RsAgo. The properties of these two proteins were most extensively studied in vitro and in vivo thus making them favorable models to understand functional activities of pAgos.

DNA-guided interference by TtAgo. TtAgo is an active endonuclease that binds DNA guides to cleave complementary DNA or RNA targets in vitro[27,29–31]. When purified from bacterial cells, TtAgo is associated exclusively with short DNA molecules[31]. The preferable substrate for TtAgo in vitro is ssDNA but it can also cleave plasmid substrates, when provided with guide molecules complementary to the two DNA strands[31]. The plasmid cleavage depends on DNA supercoiling or the presence of A/T-rich regions and occurs only at elevated temperatures, suggesting that it requires local DNA melting[31,33]. Deletion of TtAgo from the genome of *T. thermophilus* increases the efficiency of natural transformation and plasmid yield suggesting that TtAgo can also target plasmid DNA in vivo[31].

One of the most intriguing questions is how target-specific DNA guides associated with TtAgo and other DNA-loaded pAgos are generated. Cloning and sequencing of small DNAs (13–25 nucleotides in size) associated with TtAgo during expression in a heterologous *E. coli* system revealed that they predominantly originate from plasmids and are uniformly distributed over replicons, independently of the G/C-richness, gene content and orientation[31]. Importantly, these small DNAs were absent upon expression of a mutant TtAgo with substitutions of catalytic residues in the active site indicating that guide DNA formation depends on its catalytic activity.

Small DNA molecules associated with TtAgo in vivo have a strong preference for cytosine at their 5′-end (g1C)[31] but in vitro analysis demonstrated that TtAgo rather recognizes complementary guanosine residue in the target DNA strand (t1G′). This suggests that initial substrate for TtAgo is dsDNA and that selection of 5′-C-containing guides occurs during guide loading prior to removal of the complementary strand[33]. Indeed, prolonged incubation of guide-free TtAgo with double-stranded substrates, but not ssDNA, resulted in their cleavage[33]. This activity, termed DNA 'chopping', required the presence of A/T-rich or mismatched DNA regions, preferably located in the 5′-direction relative to the site of cleavage.

Other studied DNA-guided pAgo proteins revealed similar activities in vitro (Table 1). Thermophilic AaAgo, MjAgo and PfAgo exhibited efficient guide-dependent cleavage of single-stranded or supercoiled plasmid substrates[18,22,23,26]. At elevated temperatures (≥75 °C), MjAgo and PfAgo also cleaved linear or plasmid double-stranded DNA substrates without the addition of guide molecules[23,26]. Although hyperthermophiles (such as *P. furiosis* and *M. piezophila*) usually contain reverse gyrase to positively supercoil their DNA, the extreme temperatures of their habitats likely promote local DNA melting. Thus, catalytically active pAgos can autonomously initiate DNA cleavage and produce specific guide molecules for the same target, and may not require additional factors for initiation of DNA interference in vivo.

These studies have led to the model of specific DNA targeting by TtAgo and other DNA-guided pAgos schematically shown in Fig. 5 [31,33]. Guide-free TtAgo initially attacks double-stranded DNA substrates (step a) and makes distributed nicks on each DNA strand, thus resulting in generation of double-stranded fragments of varying length (step b). This is a low-efficiency process that may be stimulated by the presence of partially single-stranded regions or noncanonical DNA structures. Next, guide molecules are selected from the pool of these fragments based on

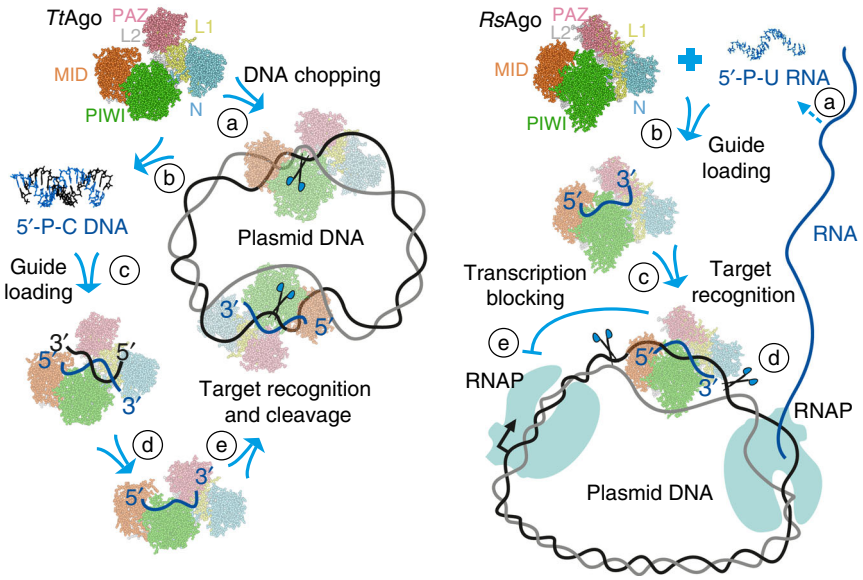

Fig. 5 Proposed mechanisms of DNA interference by DNA-guided (TtAgo, left) and RNA-guided (RsAgo, right) pAgos. TtAgo was proposed firstly to process invader DNA in a guide-independent manner ("DNA chopping", a), resulting in slow DNA fragmentation (b) and binding of short DNA duplexes (c), followed by dissociation of the passenger strand (d). Guide-loaded TtAgo can then attack the target DNA with high efficiency (e)[31,33]. RsAgo was proposed to bind short RNAs processed from mRNAs by Ago-associated or cellular nucleases (a, b), followed by target DNA recognition (c), which can result in DNA degradation by accessory nucleases (d) and/or inhibition of transcription (e)[34]

the presence of guanine in the passenger strand opposite first guide cytosine (step c), whose binding in separate protein pockets may facilitate strand separation. This is followed by dissociation of the passenger strand, either with or without its cleavage, stimulated by the presence of an A/T-rich segment in its 5′-part (step d). Guide-loaded TtAgo then attacks the same DNA target with high efficiency and specificity, resulting in the decrease in plasmid transcription and its further degradation (step e).

RNA-guided interference by RsAgo. RsAgo uses RNA guides to recognize complementary DNA targets in vitro but lacks the slicer activity due to substitutions of key catalytic resides in the active site (Supplementary Fig. 3)[34,35]. However, when purified from the host cells, RsAgo is associated with small 15–19 nt RNA and complementary 20–25 nt DNA molecules of diverse sequences[34]. The RsAgo-bound guide RNAs contain a 5′-uridine residue (gU1) and complementary DNAs have an adenine at corresponding position (tA1′) close to their 3′-end[34]; these residues are specifically recognized by RsAgo in vitro[35,51].

Small RsAgo-associated RNAs correspond to the sense strand of the genes suggesting that they are processed from cellular RNA transcripts. Little gene specificity was observed for these RNAs, though moderate enrichment for plasmid-derived and transposon transcripts, and depletion of noncoding RNAs was reported[34]. In the *R. sphaeroides* genome, RsAgo is located in the same operon with a downstream gene encoding putative nuclease. However, RsAgo still associates with small RNAs and DNAs when expressed without nuclease either in *R. sphaeroides* or in *E. coli* cells suggesting that the nuclease is not essential for nucleic acid processing and RsAgo may "collect" short RNAs from the pool of cellular RNAs processed by various RNases.

In *R. sphaeroides*, RsAgo decreases the expression of plasmid genes without obvious plasmid degradation[34]. When expressed at high levels in *E. coli*, it also decreases plasmid content and causes plasmid degradation, suggesting that it can affect not only transcription but also DNA integrity[34,35]. The mechanism of DNA processing remains unknown; however, since RsAgo lacks catalytic activity and small DNAs are processed outside of the region of complementarity to guide RNAs, the involvement of other cellular DNases was proposed[34]. An even bigger mystery is the observed specificity of target DNA recognition, since despite promiscuous association of RsAgo with RNA guides, the complex seems to target foreign DNA, particularly transposons, plasmids and prophages[34].

Overall, these studies suggested the model of RNA-guided interference by RsAgo shown in Fig. 5[34]. Initial processing of RNA transcripts by cellular nucleases results in generation of a pool of RNA fragments corresponding to both host and foreign genes (step a). Guide molecules are selected by RsAgo from this pool by their size and the presence of 5′-uridine, probably followed by the 3′-end trimming (step b). At this stage, certain properties of foreign RNA transcripts, such as low efficiency of translation, may distinguish them from host protein-coding genes (which have optimal expression patterns) or structured noncoding RNAs (protected from degradation), thus allowing preferable guide loading. At the next step, the RsAgo-RNA complex binds target DNA of corresponding genetic loci (step c). This process may be facilitated by gene transcription, which promotes local negative DNA supercoiling and melting behind RNA polymerase[73]. The presence of bound pAgo may directly affect gene transcription, by imposing a roadblock to RNA polymerase (step d). Finally, DNA-bound RsAgo complexes can be removed from the genome by the action of unknown nucleases, resulting in the appearance of single-stranded gaps and double-stranded breaks in the DNA target (step e). Similarly to DNA-guided pAgos, this may lead to degradation of the target replicons.

Commonalities and differences in the action of pAgo and eAgo proteins

At the molecular level, prokaryotic and eukaryotic Argonaute proteins are strikingly similar in the mechanisms of nucleic acid binding and slicer activity, suggesting that the basic function of Argonautes is conserved in evolution[10,11], but with certain variations discussed below. In eukaryotes, Ago proteins have evolved to use RNA guides (siRNA and miRNA) to regulate gene expression at post-transcriptional level through recognition of

RNA targets in the cytoplasm. In addition, nuclear Ago proteins in fission yeast and plants as well as nuclear PIWI-clade Argonautes in Metazoa induce transcriptional repression through binding to nascent RNAs in the nucleus[74–78]. In contrast, most studied pAgos, including archaeal proteins that likely served as predecessors of eAgos[10,11], use DNA guides to recognize DNA targets. Yet some pAgos like RsAgo and MpAgo utilize RNA guides, and it is not unlikely that RNA-targeting pAgos may also be discovered in the future, similarly to RNA-targeting CRISPR-Cas systems[79]. In fact, several pAgos including AaAgo, TtAgo and MpAgo, were shown to cleave RNA targets in vitro, although usually with lower activities in comparison with DNA targets[18,27,36,80]. The functional role of this activity in vivo remains to be established.

In contrast to eAgos, which require accessory proteins for guide generation and loading, small DNA or RNA guide loading into pAgos does not seem to depend on the action of additional proteins. Both TtAgo and RsAgo successfully associate with small nucleic acids in heterologous bacteria species[31,34], and initial DNA processing and guide loading by TtAgo and MjAgo in vitro does not require any accessory factors[23,33]. No chopping activity was reported for eAgos, but some specific miRNAs and synthetic siRNAs can be processed by the slicer activity of the Ago2 protein, without the need for Dicer, in a certain analogy with pAgos[61–63,81,82]. However, as shown for RsAgo, the mechanism of RNA-guided repression in prokaryotes is conceptually very different from RNAi in eukaryotes: while in eukaryotes guide RNAs are carefully selected to achieve the specificity of target recognition, in prokaryotes the selection is not driven simply by RNA guides and occurs —by as yet unknown mechanism—at the step of target (in this case, DNA) recognition by the guide-pAgo complex[34]. A specific group of RNA-guided CRISPR-associated pAgos, such as MpAgo, might use cellular memories of previous infections encoded in the CRISPR cassette for the recognition of foreign nucleic acids, but this has not been demonstrated experimentally yet[36,37].

The double-stranded nature of DNA implies that it should be premelted for guide-dependent recognition by pAgos, in contrast to eAgos that act on single-stranded RNA targets. DNA targeting seems to be a straightforward mechanism of gene silencing in prokaryotes, but may become inefficient in the case of eukaryotic cells, in which genomic DNA is tightly packed into chromatin, while gene activity is also highly regulated at post-transcriptional level—thus explaining the switch of eAgos to the RNA silencing activity. Indeed, DNA chromatinization was proposed to protect the genome (but not invader DNA) from the action of MjAgo in the archaeon *M. jannaschii*[23]. At the same time, some eAgos were proposed to recognize DNA in vivo (*A. thaliana* AGO4 and AGO1, mammalian Ago2)[83–85] and can use DNA guides for target recognition in vitro (hAgo2)[42], suggesting that their ability to interact with DNA might not be lost in evolution.

Suppression of foreign genetic elements by pAgos parallels the functions of the PIWI-clade eAgos and piRNAs in transposon silencing[86–89]. Furthermore, pAgos may possibly suppress gene expression at the transcriptional level[34], analogously to the piRNA pathway in eukaryotes[64–67,90,91] (see next section). Recent analysis of the SIWI protein from the PIWI clade revealed structural similarities with pAgos, including the unplugged conformation of the active site and the metal-mediated 5′-guide interactions in the MID pocket. PIWI proteins may therefore represent an ancient functional variant of eAgos[44,58].

Possible cellular functions of pAgos

While published studies proposed that elimination of foreign genetic elements through their nucleolytic cleavage may be the main mode of action for pAgos (Figs. 5, 6a; Table 1), we hypothesize that these proteins might also be implicated in the regulation of other genetic processes, not necessarily requiring DNA cleavage.

RNA-guided pAgos that lack endonuclease activity, such as RsAgo, may perform cleavage-independent repression of foreign genes (Fig. 6b). Indeed, repression of plasmid-encoded genes was observed in *R. sphaeroides* strains expressing wild-type RsAgo, without changes in the plasmid copy-number[34]. Small DNAs found in association with RsAgo in vivo[34] may possibly be a byproduct of RsAgo binding to specific DNA loci with occasional DNA cleavage, while its main function might be in transcriptional silencing. In particular, RsAgo could co-transcriptionally bind its genomic targets, thus preventing next rounds of transcription (Fig. 5)[34]. We hypothesize that such inhibition may be more efficient for foreign genes because of their inefficient translation, which is associated with lower speed of transcription and RNA polymerase backtracking[92]. Intriguingly, recent studies suggested that, similarly to RsAgo, the plant AGO4 protein may directly recognize DNA targets and prevent their further transcription through heterochromatinization[83].

Prokaryotic Ago proteins may also be involved in transcriptional regulation of host genes. In the case of eukaryotes, transcriptional repression is achieved through recognition of nascent RNA by a complex of nuclear eAgo and small RNA, followed by recruitment of chromatin modifiers that put repressive chromatin marks on the target locus[66,91,93–95]. Nuclear eAgos induce transcriptional silencing in fission yeast and plants, while the PIWI-clade Argonautes and associated piRNAs are responsible for transcriptional silencing of transposable elements in germ cells of Metazoa. In contrast to eAgos that bind nascent RNAs, loading of pAgos onto genomic loci in prokaryotic cell may directly interfere with gene transcription, similarly to DNA-binding transcription repressors (Fig. 6b). At present, no studies of the effects of RNA-guided pAgos on the expression of chromosomal genes were reported, but RsAgo was shown to repress transcription of plasmid genes[34]. Intriguingly, TtAgo stimulates (directly or indirectly) expression of certain chromosomal genes, including the CRISPR-Cas locus, in *T. thermophilus* strains containing plasmid DNA, suggesting a functional interplay between the pAgo and CRISPR systems[32]. Efficient transcription inhibition in bacterial cells was previously reported for a catalytically inactive variant of

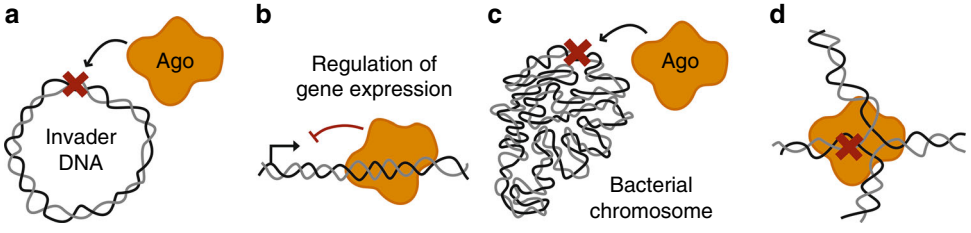

Fig. 6 Possible functions of pAgos. In addition to their function in cell defense against invader DNA (or RNA) (**a**), pAgo proteins might hypothetically be involved in the regulation of gene expression (**b**), function as suicide systems (**c**), or participate in the processing of noncanonical DNA structures and DNA repair (**d**)

the Cas9 nuclease loaded with gene-specific RNA guides[96]. It will be important to explore if pAgos might also be adopted for synthetic regulation of gene expression.

Beyond repression of foreign genetic elements and host genes, pAgos might act as a suicide system similar to abortive infection systems (reviewed in ref. [97]) that kill a bacterial cell under stress conditions (Fig. 6c). A similar function was also proposed for CRISPR-Cas systems[98,99]. In this scenario, environmental stress, extensive DNA repair or phage infection result in the appearance of partially melted DNA regions, which may be a preferable substrate for pAgo action, resulting in pAgo loading with small DNA fragments corresponding to genomic sequences. ssDNA-guided pAgos can then effectively destroy DNA, thus resulting in cell death and preventing phage multiplication.

Finally, we hypothesize that pAgos might act as components of an ancient DNA repair pathway, by inducing DNA cleavage at the sites of noncanonical DNA structures, such as broken replication forks, 5′-flaps, Holliday junctions, and R-loops (Fig. 6d). Previously, a DNA repair function was proposed for CRISPR-Cas systems[100], and CRISPR-associated nucleases have indeed been shown to play various roles beyond interference (reviewed in ref. [99,101]). In particular, the Cas1 protein from *E. coli* can process a variety of noncanonical DNA substrates in vitro[102], beyond the canonical DNA integration intermediates recognized by the Cas1-Cas2 complex[103,104]. Cas1 also physically and genetically interacts with DNA recombination factors in vivo, and its deletion renders the cells more sensitive to DNA damage[102]. Furthermore, the CRISPR-system was shown to attack noncanonical DNA substrates—mostly, damaged replication forks—and cooperate with cellular DNA repair pathways during spacer acquisition[105]. Recently, partially complementary regions were shown to promote guide-independent DNA cleavage by TtAgo[33]. Thus, we speculate that the nuclease activity of pAgos towards unusual DNA structures might stimulate their processing by other cellular nucleases and repair proteins.

eAgo proteins have been implicated in double-strand break (DSB) repair in plant and human cells, in a process that requires transcription[85,106,107]. Small RNA-loaded Ago2 was proposed to recognize the sites of DSBs through pairing with complementary DNA sequences or nascent RNA transcripts, followed by recruitment of other DSB repair proteins[85]. Moreover, Ago1 in plants was shown to interact with DNA damage-binding protein 2 (DDB2) and, possibly, facilitate recognition of the sites of UV-damage through direct base-pairing with the DNA substrate[84]. Stress-induced DNA targeting by pAgos, possibly coupled to transcription, might also play a role in DNA repair and in stress response in prokaryotic cells.

Future directions in pAgo studies

Many functional features of the proposed bacterial DNA/RNA interference systems, as well as possible regulatory pathways involving pAgos, remain to be established. The experimental evidence for their role in host defense is still very limited; for example, nothing is known about their possible effects on the replication of bacteriophages, the most abundant bacteria-targeting genetic elements. The three principal questions that have to be answered about pAgos are (1) how the nucleic acid guides associated with pAgos are generated, (2) what are the natural targets of the pAgo/guide complexes and how are they selected, and (3) what happens with the target upon its recognition by these complexes. Some specific problems that need to be addressed about pAgos are briefly outlined below.

Guide biogenesis. The molecular pathways of guide biogenesis are certainly different for RNA-guided and DNA-guided pAgos, and it remains to be known how the nucleic acid substrates are selected for initial processing. While DNA chopping was shown to be a route for guide generation in vitro[23,33], not all pAgos show this activity, and it still remains a question how the nucleic acid guides are generated in vivo. Since DNA chopping requires DNA pre-melting[33], partially single-stranded DNA that appears during invasion and replication of mobile genetic elements might be first attacked by non-guided pAgos. In the case of CRISPR/Cas systems, the RecBCD exonuclease was shown to process DNA for spacer generation during the adaptation step of CRISPR/Cas-interference[105]. The same system might contribute to preferable processing of foreign DNA into DNA guides utilized by pAgo proteins.

The RNA guide biogenesis may depend on the transcription-translation coupling (not existing in eukaryotes), which may drive RNA processing and guide loading into pAgos. The features that might make an mRNA a preferable source of guide molecules include its inefficient translation (which makes RNA unprotected by the ribosomes)[74], or specific secondary structure. The nucleases involved in RNA cleavage are unknown but likely candidates include pAgo-associated proteins encoded in the same operons. It remains to be known whether Cas nucleases may participate in guide RNA processing in the specific case of CRISPR-Cas-associated pAgos (MpAgo)[36]. It will be also interesting to test whether pAgos can also perform guide-independent cleavage of (partially double-stranded) RNA precursors, similarly to the processing of a subclass of miRNAs by eAgo2[61,62,81].

Target selection. Almost nothing is known about the mechanisms that may target pAgos to specific genomic loci or foreign replicons, such as extrachromosomal DNA, transposons, plasmids or phages. Unusual replication properties of these elements can lead to the formation of partially single-stranded DNA intermediates that may be preferably recognized by guide-loaded pAgos[31,33,34]. Single-stranded DNA regions can appear in the cell during DNA repair and transposition, or as a result of perturbed transcription. Single-stranded DNA can also enter the cell during the processes of conjugation and natural transformation, thus making horizontally acquired DNA more susceptible to the pAgo action. The multicopy nature of plasmids and transposable elements can rise the number of produced guide molecules and may induce silencing when this number exceeds a threshold level. For MjAgo, DNA coverage by archaeal histone proteins was proposed to protect genomic DNA from cleavage thus making plasmids more susceptible for Ago action[23]. Architectural DNA binding proteins may introduce a similar bias in bacteria.

Gene-specific differences in the transcription and translation levels may also affect target selection. In prokaryotes, foreign DNA sequences are less efficiently translated because of suboptimal codon bias[74]. Decreased translation results in lower rates of transcription due to inefficient transcription-translation coupling and increased RNA polymerase backtracking[92], which may in turn affect DNA replication and repair[75,108], and co-transcriptional pAgo loading.

Target processing. The mechanisms of target degradation by pAgos in vivo remain poorly understood. For TtAgo, short DNAs are uniformly distributed along a target plasmid, arguing against sequence-dependent or ordered DNA cleavage[31]; nothing is known about in vivo DNA processing by other catalytically active pAgos. It is plausible that other cellular nucleases, such as homologous recombination machinery, may contribute to dsDNA processing (similarly to the CRISPR-Cas interference[105]). The RecBCD system might participate in plasmid degradation after its initial cleavage by pAgo proteins, resulting in its preferable processing resulting from the absence of Chi-sites.

Recently, it was shown that in vitro cleavage of double-stranded DNA by TtAgo can also be promoted by the UvrD helicase and the SSB protein[109]; however, it remains to be established whether these or other factors also facilitate DNA processing in vivo.

Catalytically inactive pAgos, such as RsAgo, process target DNA by an unknown mechanism that may involve the action of pAgo-associated nucleases. Furthermore, it remains unknown whether DNA cleavage is an essential step in the action of these type of pAgos, since their strong association with DNA may by itself affect target replication, transcription and repair, as discussed above[34].

Functional activities of short pAgos. While short pAgos constitute a large part of all pAgos, their functional activities and the ability to interact with nucleic acids in vivo were never tested (and hence their DNA/RNA specificity remains unknown). Short pAgos lack the N-terminal half of the protein, including the PAZ and MID domains involved in guide binding and target recognition (Fig. 1, AfAgo), and contain inactivated catalytic site. Furthermore, the path of DNA and RNA duplexes bound by AfAgo in reported structures (Fig. 1) significantly differs from long pAgos, suggesting that other (APAZ-containing) proteins encoded in the same operons may participate in DNA/RNA binding and processing.

Noncanonical pAgo functions. As we argue in this review, protection against invader DNA may not be the only cellular function of pAgo proteins (Fig. 6). To date, detailed in vivo studies have been performed for only two proteins (TtAgo and RsAgo) from the highly divergent evolutionary tree of pAgos. The detailed understanding of possible pAgo roles in genetic regulation, stress response and DNA repair will therefore require study of new bacterial and archaeal pAgos, selected on the basis of their evolutionary and functional diversity[10,11], and the availability of convenient genetic systems for their analysis.

The use of pAgos in genetic engineering. In addition to understanding pAgo function in their host prokaryotic cells, it is worth exploring the possibility to use pAgos as tools for transcription regulation, genome editing and epigenome rewriting[13]. Initial attempts to use an archaeal Ago protein for genome editing were irreproducible[110–112] but analysis of diverse pAgos found in various bacterial and archaeal species may help to select better candidates for genome manipulations. Further studies may help to find efficient RNA-targeting pAgos, which, in contrast to eAgos, will not interfere with the cellular RNAi pathways. Several studied pAgos (AaAgo, MpAgo, TtAgo) are able recognize and cleave RNA in vitro[18,27,36,80], and MpAgo was recently adopted for detection of specific RNA species from complex mixtures[80]. The main problems that need to be solved include the directing of pAgos to desired genomic locations or mRNA targets and avoiding off-target effects. For this purpose, pAgos can be fused with additional domains for specific loading of RNA or DNA guides and chromatin modification[110].

Data availability

No datasets were generated or analysed during the current study.

Received: 2 February 2018 Accepted: 26 October 2018
Published online: 04 December 2018

References

1. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* **10**, 94–108 (2009).
2. Hammond, S. M., Boettcher, S., Caudy, A. A., Kobayashi, R. & Hannon, G. J. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**, 1146–1150 (2001).
3. Martinez, J., Patkaniowska, A., Urlaub, H., Luhrmann, R. & Tuschl, T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563–574 (2002).
4. Meister, G. et al. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15**, 185–197 (2004).
5. Joshua-Tor, L. The Argonautes. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 67–72 (2006).
6. Peters, L. & Meister, G. Argonaute proteins: mediators of RNA silencing. *Mol. Cell* **26**, 611–623 (2007).
7. Pratt, A. J. & MacRae, I. J. The RNA-induced silencing complex: a versatile gene-silencing machine. *J. Biol. Chem.* **284**, 17897–17901 (2009).
8. Moazed, D. Small RNAs in transcriptional gene silencing and genome defence. *Nature* **457**, 413–420 (2009).
9. 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* **4**, 29 (2009). ***The hypothesis is presented that pAgos are components of a novel prokaryotic immune system that targets foreign nucleic acids.***
10. Swarts, D. C. et al. The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.* **21**, 743–753 (2014). ***Comprehensive analysis of the structure, evolution and functions of pAgos and associated proteins is presented.***
11. Koonin, E. V. Evolution of RNA- and DNA-guided antivirus defense systems in prokaryotes and eukaryotes: common ancestry vs convergence. *Biol. Direct.* **12**, 5 (2017). ***This review discuss general principles in the evolution and action of RNA- and DNA-guided immune systems in prokaryotes and eukaryotes, including Ago-centered and CRISPR/Cas pathways.***
12. Willkomm, S., Makarova, K. & Grohmann, D. DNA-silencing by prokaryotic Argonaute proteins adds a new layer of defence against invading nucleic acids. *FEMS Microbiol. Rev.* **42**, 376–387 (2018).
13. Hegge, J. W., Swarts, D. C. & van der Oost, J. Prokaryotic Argonaute proteins: novel genome-editing tools? *Nat. Rev. Microbiol.* **16**, 5–11 (2017).
14. Parker, J. S., Roe, S. M. & Barford, D. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J.* **23**, 4727–4737 (2004).
15. Ma, J. B. et al. Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature* **434**, 666–670 (2005).
16. Parker, J. S., Roe, S. M. & Barford, D. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* **434**, 663–666 (2005).
17. 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* **33**, 204–214 (2009).
18. Yuan, Y. R. et al. Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell* **19**, 405–419 (2005).
19. Yuan, Y. R., Pei, Y., Chen, H. Y., Tuschl, T. & Patel, D. J. A potential protein-RNA recognition event along the RISC-loading pathway from the structure of A. aeolicus Argonaute with externally bound siRNA. *Structure* **14**, 1557–1565 (2006).
20. Rashid, U. J. et al. Structure of Aquifex aeolicus argonaute highlights conformational flexibility of the PAZ domain as a potential regulator of RNA-induced silencing complex function. *J. Biol. Chem.* **282**, 13824–13832 (2007).
21. Zander, A., Holzmeister, P., Klose, D., Tinnefeld, P. & Grohmann, D. Single-molecule FRET supports the two-state model of Argonaute action. *RNA Biol.* **11**, 45–56 (2014).
22. Willkomm, S. et al. Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein. *Nat. Microbiol.* **2**, 17035 (2017). ***A detailed structural and mutational analysis of catalytically active MjAgo is presented, including characterization of its structural plasticity and specificity during guide binding.***
23. Zander, A. et al. Guide-independent DNA cleavage by archaeal Argonaute from Methanocaldococcus jannaschii. *Nat. Microbiol.* **2**, 17034 (2017). ***This study compares two principal modes of MjAgo action, guide-depended target DNA cleavage and DNA chopping, and demonstrates that both activities depend on the same structural elements of pAgo.***
24. Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434–1437 (2004).
25. Rivas, F. V. et al. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat. Struct. Mol. Biol.* **12**, 340–349 (2005).
26. Swarts, D. C. et al. Argonaute of the archaeon Pyrococcus furiosus is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.* **43**, 5120–5129 (2015).
27. Wang, Y. et al. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921–926 (2008).
28. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209–213 (2008).

29. Wang, Y. et al. Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754–761 (2009). ***This study shows successive steps in the formation of ternary complexes by TtAgo, illustrating stepwise extention of the guide-target duplex and release of the 3'-guide end from the PAZ pocket***.
30. Sheng, G. et al. Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl Acad. Sci. USA* **111**, 652–657 (2014). ***This study describes structural transitions in the active site of TtAgo during interactions with its native substrates, guide DNA and target DNA, illuminating the closure of the active site, catalytic metal binding and target cleavage***.
31. Swarts, D. C. et al. DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261 (2014). ***It is shown here that DNA-guided TtAgo preferably cuts plasmid DNA in vivo and can act as a barrier for the uptake and propagation of foreign DNA in bacterial cells***.
32. Swarts, D. C., Koehorst, J. J., Westra, E. R., Schaap, P. J. & van der Oost, J. Effects of Argonaute on gene expression in thermus thermophilus. *PLoS One* **10**, e0124880 (2015).
33. Swarts, D. C. et al. Autonomous generation and loading of DNA guides by bacterial Argonaute. *Mol. Cell* **65**, 985–998 (2017). ***The choppping activity of TtAgo toward double-stranded DNA is described in vitro, revealing its preferences for partially-melted or noncanonical DNA substrates***.
34. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. Bacterial argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell* **51**, 594–605 (2013). ***The first study to show that pAgo proteins can preferentially target foreign nucleic acids in bacteria; RsAgo is demonstrated to associate with small nucleic acids corresponding to plasmids, transposons and phage genes and to suppress plasmid transcription***.
35. Miyoshi, T., Ito, K., Murakami, R. & Uchiumi, T. Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.* **7**, 11846 (2016). ***The first structure of an inactive long pAgo (RsAgo), accompanied by functional analysis of mutations in its structural elements involved in guide and target interactions***.
36. Kaya, E. et al. A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl Acad. Sci. USA* **113**, 4057–4062 (2016).
37. Doxzen, K. W. & Doudna, J. A. DNA recognition by an RNA-guided bacterial Argonaute. *PLoS One* **12**, e0177097 (2017). ***References 36 and 37 describe an unusual CRISPR-associated pAgo (MpAgo) that recognizes 5'-OH guide RNAs and preferentially cleaves DNA targets, thus suggesting a possible interplay between the pAgo and CRISPR systems***.
38. Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. Structure of yeast Argonaute with guide RNA. *Nature* **486**, 368–374 (2012).
39. 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.* **3**, 1901–1909 (2013).
40. Elkayam, E. et al. The structure of human argonaute-2 in complex with miR-20a. *Cell* **150**, 100–110 (2012).
41. Schirle, N. T., Sheu-Gruttadauria, J., Chandradoss, S. D., Joo, C. & MacRae, I. J. Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. *eLife* **4**, e07646 (2015).
42. Willkomm, S., Zander, A., Grohmann, D. & Restle, T. Mechanistic insights into Archaeal and human Argonaute substrate binding and cleavage properties. *PLoS One* **11**, e0164695 (2016).
43. Park, M. S. et al. Human Argonaute3 has slicer activity. *Nucleic Acids Res.* **45**, 11867–11877 (2017).
44. Matsumoto, N. et al. Crystal structure of silkworm PIWI-clade Argonaute Siwi bound to piRNA. *Cell* **167**, 484–497 (2016).
45. Willkomm, S., Zander, A., Gust, A. & Grohmann, D. A prokaryotic twist on argonaute function. *Life* **5**, 538–553 (2015).
46. Sheu-Gruttadauria, J. & MacRae, I. J. Structural foundations of RNA silencing by Argonaute. *J. Mol. Biol.* **429**, 2619–2639 (2017).
47. Globyte, V., Kim, S. H. & Joo, C. Single-molecule view of small RNA-guided target search and recognition. *Annu Rev. Biophys.* **47**, 569–593 (2018).
48. Wee, L. M., Flores-Jasso, C. F., Salomon, W. E. & Zamore, P. D. Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell* **151**, 1055–1067 (2012).
49. Salomon, W. E., Jolly, S. M., Moore, M. J., Zamore, P. D. & Serebrov, V. Single-molecule imaging reveals that Argonaute reshapes the binding properties of its nucleic acid guides. *Cell* **162**, 84–95 (2015). ***A detailed kinetic study that compares fine details of target binding, cleavage and dissociation by eukaryotic and prokaryotic Ago proteins at the single-molecule level***.
50. Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20–51 (2018).
51. Liu, Y. et al. Accommodation of helical imperfections in *Rhodobacter sphaeroides* Argonaute ternary complexes with guide RNA and target DNA. *Cell Rep.* **24**, 453–462 (2018). ***This study reveals the effects of mismatches and bulges in the ternary complexes of RsAgo on target binding and Ago recycling; it is shown that imperfect targets can promote dissociation of guide molecules from RsAgo***.
52. Nakanishi, K. et al. Eukaryote-specific insertion elements control human ARGONAUTE slicer activity. *Cell Rep.* **3**, 1893–1900 (2013).
53. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336**, 1037–1040 (2012).
54. Schirle, N. T., Sheu-Gruttadauria, J. & MacRae, I. J. Structural basis for microRNA targeting. *Science* **346**, 608–613 (2014).
55. Klum, S. M., Chandradoss, S. D., Schirle, N. T., Joo, C. & MacRae, I. J. Helix-7 in Argonaute2 shapes the microRNA seed region for rapid target recognition. *EMBO J.* **37**, 75–88 (2018).
56. Jung, S. R. et al. Dynamic anchoring of the 3'-end of the guide strand controls the target dissociation of Argonaute-guide complex. *J. Am. Chem. Soc.* **135**, 16865–16871 (2013).
57. Hur, J. K., Zinchenko, M. K., Djuranovic, S. & Green, R. Regulation of Argonaute slicer activity by guide RNA 3' end interactions with the N-terminal lobe. *J. Biol. Chem.* **288**, 7829–7840 (2013).
58. Sakakibara, K. & Siomi, M. C. The PIWI-interacting RNA molecular pathway: insights from cultured silkworm germline cells. *Bioessays* **40**, 1700068 (2018).
59. Ameres, S. L., Martinez, J. & Schroeder, R. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130**, 101–112 (2007).
60. Haley, B. & Zamore, P. D. Kinetic analysis of the RNAi enzyme complex. *Nat. Struct. Mol. Biol.* **11**, 599–606 (2004).
61. Cheloufi, S., Dos Santos, C. O., Chong, M. M. & Hannon, G. J. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**, 584–589 (2010).
62. Cifuentes, D. et al. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**, 1694–1698 (2010).
63. Sun, G. et al. Differences in silencing of mismatched targets by sliced versus diced siRNAs. *Nucleic Acids Res.* **46**, 6806–6822 (2018).
64. Klenov, M. S. et al. Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline. *Nucleic Acids Res.* **35**, 5430–5438 (2007).
65. Kuramochi-Miyagawa, S. et al. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* **22**, 908–917 (2008).
66. Le Thomas, A. et al. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev.* **27**, 390–399 (2013).
67. Rozhkov, N. V., Hammell, M. & Hannon, G. J. Multiple roles for Piwi in silencing Drosophila transposons. *Genes Dev.* **27**, 400–412 (2013).
68. Dahlgren, C. et al. Analysis of siRNA specificity on targets with double-nucleotide mismatches. *Nucleic Acids Res.* **36**, e53 (2008).
69. Doench, J. G. & Sharp, P. A. Specificity of microRNA target selection in translational repression. *Genes Dev.* **18**, 504–511 (2004).
70. Du, Q., Thonberg, H., Wang, J., Wahlestedt, C. & Liang, Z. A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Res.* **33**, 1671–1677 (2005).
71. Sheng, G. et al. Structure/cleavage-based insights into helical perturbations at bulge sites within T. thermophilus Argonaute silencing complexes. *Nucleic Acids Res.* **45**, 9149–9163 (2017). ***A series of structures of ternary complexes of TtAgo with bulges in the guide or target strands is reported, together with the description of their effects on the activity of pAgo***.
72. Park, J. H., Shin, S. Y. & Shin, C. Non-canonical targets destabilize microRNAs in human Argonautes. *Nucleic Acids Res.* **45**, 1569–1583 (2017).
73. Wu, H. Y., Shyy, S. H., Wang, J. C. & Liu, L. F. Transcription generates positively and negatively supercoiled domains in the template. *Cell* **53**, 433–440 (1988).
74. Quax, T. E., Claassens, N. J., Soll, D. & van der Oost, J. Codon bias as a means to fine-tune gene expression. *Mol. Cell* **59**, 149–161 (2015).
75. Nudler, E. RNA polymerase backtracking in gene regulation and genome instability. *Cell* **149**, 1438–1445 (2012).
76. Vagin, V. V. et al. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**, 320–324 (2006).
77. Aravin, A. A. et al. Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line. *Mol. Cell. Biol.* **24**, 6742–6750 (2004).
78. Brennecke, J. et al. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**, 1089–1103 (2007).
79. Abudayyeh, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).
80. Lapinaite, A., Doudna, J. A. & Cate, J. H. D. Programmable RNA recognition using a CRISPR-associated Argonaute. *Proc. Natl Acad. Sci. USA* **115**, 3368–3373 (2018).
81. Yang, J. S. et al. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc. Natl Acad. Sci. USA* **107**, 15163–15168 (2010).
82. Chen, G. R., Sive, H. & Bartel, D. P. A seed mismatch enhances Argonaute2-catalyzed cleavage and partially rescues severely impaired cleavage found in fish. *Mol. Cell* **68**, 1095–1107 (2017).
83. Lahmy, S. et al. Evidence for Argonaute4-DNA interactions in RNA-directed DNA methylation in plants. *Genes Dev.* **30**, 2565–2570 (2016).

84. Schalk, C. et al. Small RNA-mediated repair of UV-induced DNA lesions by the DNA damage-binding protein 2 and Argonaute 1. *Proc. Natl Acad. Sci. USA* **114**, E2965–E2974 (2017).
85. Gao, M. et al. Ago2 facilitates Rad51 recruitment and DNA double-strand break repair by homologous recombination. *Cell Res.* **24**, 532–541 (2014).
86. Miesen, P., Ivens, A., Buck, A. H. & van Rij, R. P. Small RNA profiling in dengue virus 2-infected aedes mosquito cells reveals viral piRNAs and novel host miRNAs. *PLoS Negl. Trop. Dis.* **10**, e0004452 (2016).
87. Miesen, P., Joosten, J. & van Rij, R. P. PIWIs go viral: arbovirus-derived piRNAs in vector mosquitoes. *PLoS Pathog.* **12**, e1006017 (2016).
88. Malone, C. D. & Hannon, G. J. Small RNAs as guardians of the genome. *Cell* **136**, 656–668 (2009).
89. Siomi, M. C., Sato, K., Pezic, D. & Aravin, A. A. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat. Rev. Mol. Cell Biol.* **12**, 246–258 (2011).
90. Aravin, A. A. et al. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol. Cell* **31**, 785–799 (2008).
91. Sienski, G., Donertas, D. & Brennecke, J. Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**, 964–980 (2012).
92. Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**, 504–508 (2010).
93. Verdel, A. et al. RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672–676 (2004).
94. Liu, W. et al. RNA-directed DNA methylation involves co-transcriptional small-RNA-guided slicing of polymerase V transcripts in Arabidopsis. *Nat. Plants* **4**, 181–188 (2018).
95. Pezic, D., Manakov, S. A., Sachidanandam, R. & Aravin, A. A. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev.* **28**, 1410–1428 (2014).
96. Qi, L. S. et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183 (2013).
97. Dy, R. L., Richter, C., Salmond, G. P. & Fineran, P. C. Remarkable mechanisms in microbes to resist phage infections. *Annu Rev. Virol.* **1**, 307–331 (2014).
98. Strotskaya, A. et al. The action of Escherichia coli CRISPR-Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic Acids Res.* **45**, 1946–1957 (2017).
99. Westra, E. R., Buckling, A. & Fineran, P. C. CRISPR-Cas systems: beyond adaptive immunity. *Nat. Rev. Microbiol.* **12**, 317–326 (2014).
100. Makarova, K. S., Aravind, L., Grishin, N. V., Rogozin, I. B. & Koonin, E. V. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res.* **30**, 482–496 (2002).
101. Faure, G., Makarova, K. S. & Koonin, E. V. CRISPR-Cas: complex functional networks and multiple roles beyond adaptive immunity. *J. Mol. Biol.*, https://doi.org/10.1016/j.jmb.2018.08.030 (2018).
102. Babu, M. et al. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* **79**, 484–502 (2011).
103. Nunez, J. K. et al. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* **21**, 528–534 (2014).
104. Wright, A. V. et al. Structures of the CRISPR genome integration complex. *Science* **357**, 1113–1118 (2017).
105. Levy, A. et al. CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* **520**, 505–510 (2015).
106. Wei, W. et al. A role for small RNAs in DNA double-strand break repair. *Cell* **149**, 101–112 (2012).
107. Hawley, B. R., Lu, W. T., Wilczynska, A. & Bushell, M. The emerging role of RNAs in DNA damage repair. *Cell Death Differ.* **24**, 580–587 (2017).
108. Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M. E. & Nudler, E. Linking RNA polymerase backtracking to genome instability in E. coli. *Cell* **146**, 533–543 (2011).
109. Hunt, E. A., Evans, T. C. Jr. & Tanner, N. A. Single-stranded binding proteins and helicase enhance the activity of prokaryotic argonautes in vitro. *PLoS One* **13**, e0203073 (2018).
110. Lee, S. H. et al. Failure to detect DNA-guided genome editing using *Natronobacterium gregoryi* Argonaute. *Nat. Biotechnol.* **35**, 17–18 (2017).
111. Javidi-Parsijani, P. et al. No evidence of genome editing activity from Natronobacterium gregoryi Argonaute (NgAgo) in human cells. *PLoS One* **12**, e0177444 (2017).
112. Khin, N. C., Lowe, J. L., Jensen, L. M. & Burgio, G. No evidence for genome editing in mouse zygotes and HEK293T human cell line using the DNA-guided Natronobacterium gregoryi Argonaute (NgAgo). *PLoS One* **12**, e0178768 (2017).
113. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
114. Tolia, N. H. & Joshua-Tor, L. Slicer and the argonautes. *Nat. Chem. Biol.* **3**, 36–43 (2007).
115. Vaucheret, H. Plant Argonautes. *Trends Plant Sci.* **13**, 350–358 (2008).
116. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
117. Carmell, M. A. & Hannon, G. J. RNase III enzymes and the initiation of gene silencing. *Nat. Struct. Mol. Biol.* **11**, 214–218 (2004).
118. Ketting, R. F. et al. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. *Genes Dev.* **15**, 2654–2659 (2001).
119. Lee, Y. et al. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
120. Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the microprocessor complex. *Nature* **432**, 231–235 (2004).
121. Huang, X., Fejes Toth, K. & Aravin, A. A. piRNA biogenesis in Drosophila melanogaster. *Trends Genet.* **33**, 882–894 (2017).
122. Forstemann, K., Horwich, M. D., Wee, L., Tomari, Y. & Zamore, P. D. Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell* **130**, 287–297 (2007).
123. Tomari, Y., Du, T. & Zamore, P. D. Sorting of Drosophila small silencing RNAs. *Cell* **130**, 299–308 (2007).
124. Aravin, A. et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203–207 (2006).
125. Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199–202 (2006).
126. Lau, N. C. et al. Characterization of the piRNA complex from rat testes. *Science* **313**, 363–367 (2006).
127. Baulcombe, D. RNA silencing in plants. *Nature* **431**, 356–363 (2004).
128. Mussabekova, A., Daeffler, L. & Imler, J. L. Innate and intrinsic antiviral immunity in Drosophila. *Cell. Mol. Life Sci.* **74**, 2039–2054 (2017).
129. Pumplin, N. & Voinnet, O. RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. *Nat. Rev. Microbiol.* **11**, 745–760 (2013).

Acknowledgements

We apologize to many colleagues whose work is not cited due to space limitations. This work was supported by the Grant of the Ministry of Education and Science of Russian Federation 14.W03.31.0007.

Author contributions

L.L., A.A.A. and A.K. analyzed the data. L.L. and A.K. prepared the figures. A.A.A. and A. K. wrote the manuscript.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-07449-7.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2018