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

Lidiya Lisitskaya, Alexei A. Aravin & Andrey Kulbachinskiy

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 unconventional 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.
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. 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). 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 builds 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 adjourn 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 contain the APAZ ("analog of PAZ") domain of unknown functions.

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 (AfAgo14–17, TtAgo15–18, KpAgo38, hAgo238, 40–42, and hAgo343) and RNA-guided (RsAgo34,35 and MpAgo36,37) pAgos (see Table 1 for pAgo abbreviations). The most complete structural information was obtained for TiAgo that was crystalized 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 KpAgo38 (yeast *Kluyveromyces polysporus*), hAgo139, hAgo238,40–42, and hAgo343 (human) from the AGO-clade and SIWI44 (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. 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. 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, resulting in mature miRNAs, found in Metazoa, that 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. miRNAs and siRNAs associate with AGO-clade Ago proteins. In some species like *Drosophila* miRNAs and siRNAs are predominantly sorted into different Ago proteins, while in others including mammals they can be loaded into the same protein. piRNAs associate with the PIWI clade Ago proteins. The resulting effector complexes then regulate the expression of host genes, suppress transposons or combat viral infection.
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 hAgo\(^{21}\), and RsAgo\(^{31}\)).

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\(^{28,40,52,53}\). 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)\(^{29,30,34}\). 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)\(^{22,28,36,38,40-50}\). 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\(^{18-20}\) and of DNA targets by analyzed pAgos (e.g., MpAgo\(^{37}\) and RsAgo\(^{31}\)). 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\(^{21,36,57}\) 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\(^{33,51}\), MpAgo\(^{36}\) or completely

### 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 | AI-Ago (short pAgo) | DNA (RNA) | 5′-P; Unknown            | DNA (RNA) | Inactive           | -                   | 14-17      |
| Marinitoga piezophila | TpAgo     | RNA   | 5′-OH, None                   | DNA (RNA) | Guide-dependent     | Reduced plasmid content and transformation efficiency | 27-23      |
| Methanocaldococcus jannaschii | MP-Ago | DNA   | 5′-P; Purines                 | DNA     | Guide-dependent; Chopping | Reduced plasmid content and transformation efficiency | 24-26      |
| Pyrococcus furiosus | PI-Ago    | DNA   | 5′-P; None                    | DNA     | Guide-dependent; Guide-independent | Reduced transcription of reporter genes and plasmid content | 34,35      |
| Rhodobacter sphaeroides | RsAgo     | RNA\(^a\) | 5′-P; g1U                    | DNA\(^a\) | Inactive           | Reduced transcription of reporter genes and plasmid content | 27-33      |
| Thermus thermophilus | TtAgo    | DNA\(^a\) | 5′-P; g1C/g1Y                | DNA\(^a\) | 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         |
| Klyuyveromyces polysporus | KpAgo    | RNA   | 5′-P; g1U                    | RNA     | Guide-dependent    | miRNA pathway      | 38,40-42   |
| Bombyx mori | SIWI      | RNA   | 5′-P; g1U                    | RNA     | Guide-dependent    | piRNA pathway      | 44         |

\(^{a}\)Both in vitro and in vivo

NATURE COMMUNICATIONS | DOI: 10.1038/s41467-018-07449-7

**Note:** The table shows a comparison of the functional properties of analyzed pAgo proteins in comparison with eAgos. The table includes information about the host, argonaute, guide, 5′-end nucleotide preference, target, catalytic activity, functional activity, and references.
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 ions26,29,31,36. 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 site9,10. In addition, RsAgo remains in the unplugged conformation even after ternary complex formation, which also prevents catalytic metal binding (Supplementary Fig. 3)35,31. 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 pAgos34,58 (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)18,22,27,29,36. Intriguingly, more than one cleavage site was observed for MjAgo21–23 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 cleavage25,48,49,59,60. Mismatches in both the seed region and the 3′-supplementary guide site increase the enzyme turnover, although at the cost of decreased target binding48,49,60. At the same time, target release is not rate limiting for catalysis by the thermophilic TtAgo protein29, 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 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 3′-part of the target strand has dissociated from TtAgo and the corresponding 3′-portion of...
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 RNA. 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 have the most deleterious effects on the efficiency of silencing in eukaryotes. Similarly, mismatches and bulges within the seed region significantly impair target binding and cleavage by studied pAgos. 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). 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. 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. 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. Mismatches and bulges around the active site greatly decrease the efficiency of target cleavage by most studied eAgos and pAgos alike. 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). Thus, formation of the perfect guide-target duplex in the active site is a critical checkpoint in the specific target
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. 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. When purified from bacterial cells, TtAgo is associated exclusively with short DNA molecules. 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. 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. 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.

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. 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 (gIC) but in vitro analysis demonstrated that TtAgo rather recognizes complementary guanosine residue in the target DNA strand (tIG). 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. Indeed, prolonged incubation of guide-free TtAgo with double-stranded substrates, but not ssDNA, resulted in their cleavage. 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. Thermophilic AaAgo, MjAgo and PfAgo exhibited efficient guide-dependent cleavage of single-stranded or supercoiled plasmid substrates. At elevated temperatures (275 °C), MjAgo and PfAgo also cleaved linear or plasmid double-stranded DNA substrates without the addition of guide molecules. Although hyperthermophiles (such as *P. furiosus* 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. 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
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 residues in the active site (Supplementary Fig. 3)\(^3^{4,5}\). 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\(^3^{4}\). 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\(^2^{3,4}\); these residues are specifically recognized by RsAgo in vitro\(^3^{5,5}\).

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\(^3^{4}\). 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\(^3^{4}\). 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\(^3^{4,5}\). The mechanism of DNA processing remains unknown; however, since RsAgo lacks catalytic activity and small RNAs are processed outside of the region of complementarity to guide RNAs, the involvement of other cellular DNases was proposed\(^3^{4}\). 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\(^3^{4}\).

Overall, these studies suggested the model of RNA-guided interference by RsAgo shown in Fig. 5\(^3^{4}\). 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\(^3^{3}\). 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 nucleus74–78. In contrast, most studied pAgos, including archaela proteins that likely served as predecessors of eAgos33, 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 systems22. 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 targets18,27,36. 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 species31,34, and initial DNA processing and guide loading by TtAgo and MiAgo in vitro does not require any accessory factors23,33. 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 pAgos61. Small DNA or RNA guide loading into pAgos does not require any accessory factors23,33. 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.

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-number8. Small DNAs found in association with RsAgo in vivo84 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-translationally bind its gene 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 backtracking92. Intriguingly, recent studies suggested that, similarly to RsAgo, the plant AGO4 protein may directly recognize DNA targets and prevent their further transcription through heterochromatinization83.

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 locus56,91,93–95. 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 genes34. 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 systems32. Efficient transcription inhibition in bacterial cells was previously reported for a catalytically inactive variant of

![Image](https://example.com/image.png)

**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. 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. In this scenario, environmental stress, conditions (Fig. 6c). A similar function was also proposed for CRISPR-Cas systems. In this scenario, environmental stress, conditions (Fig. 6c).

Previously, a DNA repair function was proposed for CRISPR-Cas systems (DSB) repair in plant and human cells, in a process that requires nucleases and repair proteins. DNA structures might stimulate their processing by other cellular 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 (DSB) repair in plant and human cells, in a process that requires nucleases and repair proteins. DNA structures might stimulate their processing by other cellular systems (reviewed in ref. 97) that kill a bacterial cell under stress conditions (Fig. 6c).

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, S' flaps, Holliday junctions, and R-loops (Fig. 6d). Previously, a DNA repair function was proposed for CRISPR-Cas systems 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, beyond the canonical DNA integration intermediates recognized by the Cas1-Cas2 complex. Cas1 also physiologically and genetically interacts with DNA recombination factors in vivo, and its deletion renders the cells more sensitive to DNA damage. 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. Recently, partially complementary regions were shown to promote guide-independent DNA cleavage by TtAgo. 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. 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. 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. 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, 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, 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. 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), 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). 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 eAgo.

**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. 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 raise 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. 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. Decreased translation results in lower rates of transcription due to inefficient transcription-translation coupling and increased RNA polymerase backtracking, which may in turn affect DNA replication and repair, 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; 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). 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\(^ {11}\).

**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\(^ {10,11}\) but analysis of diverse pAgos found in prokaryotes and eukaryotes: common ancestry vs convergence. Comprehensive analysis of the structure, evolution and functions of pAgos and associated proteins is presented.

**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 RNA. *Science* 293, 1146–1150 (2001).
3. Martinez, J., Patkanwiska, A., Urlaub, H., Lührmann, 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 antiviral defense systems in prokaryotes and eukaryotes: common ancestry vs convergence. *Biol. Direct.* 12, 5 (2017). This review discus 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. 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 RNA-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 protein-potential 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-guided 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. *Nature Biol.* 11, 45–56 (2014).
22. Willkomm, S. et al. Structural and mechanistic insights into an archaeal RNA-guided Argonaute protein. *Nat. Microbiol.* 2, 17035 (2017). A detailed structural and mutational analysis of catalytically active MgAgo 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 Methanothermobacter jannaschii. *Nat. Microbiol.* 2, 17034 (2017). This study compares two principal modes of MgAgo action, guide-dependent 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 archaean 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., Juraneck, 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 Ago complexes by T-Ago, illustrating stepwise extension of the guide-target duplex and release of the 3′-end guide from the PAZ pocket.

30. Sheng, G. et al. Structure-based cleavage mechanism of Thermus thermophilus Argonaute 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 bacterial Argonaute. *Cell* 65, 985–998 (2017). The chopping activity of TtAgo toward double-stranded DNA is described in vitro, revealing its preferences for partially-melted or noncanonical DNA substrates.

32. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. Bacterial argonaute samples the transcriptome to identify foreign DNA. *Nature Commun.* 8, 256 (2017).

33. Swarts, D. C. et al. Autonomous generation and loading of DNA guides by bacterial Argonaute. *Nat. Commun.* 7, 11067 (2016).

34. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, D. K. & Aravin, A. A. Bacterial argonaute samples the transcriptome to identify foreign DNA. *Nature Commun.* 8, 256 (2017).

35. Miyoshi, T., Ito, K., Murakami, R. & Uchiiumi, T. Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.* 7, 11067 (2016). The first structure of an inactive long Ago (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. *Proc. Natl Acad. Sci. USA* 114, 8217–8222 (2017). References 36 and 37 describe an unusual CRISPR-associated Pago (MpaGog) that recognizes 5′-OH guide RNAs and preferentially cleaves DNA targets, thus suggesting a possible interplay between the Pago and CRISPR systems.

38. Nakashima, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. Structure of yeast Argonaute with guide RNA. *Cell* 170, 1391–1405 (2017). This study describes structural changes 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.

39. Swarts, D. C. et al. Autonomous generation and loading of DNA guides by bacterial Argonaute. *Nat. Commun.* 7, 11067 (2016). The first structure of an inactive long Ago (RsAgo), accompanied by functional analysis of mutations in its structural elements involved in guide and target interactions.

40. Kaya, E. et al. A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl Acad. Sci. USA* 113, 4057–4062 (2016).

41. Schirle, N. T., Sheu-Gruttadauria, J., Chandradoss, S. D., Joo, C. & MacRae, I. J. Structural basis for ARGONAUTE slicer activity. *Proc. Natl Acad. Sci. USA* 111, 761–766 (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.

42. Swarts, D. C. et al. Autonomous generation and loading of DNA guides by bacterial Argonaute. *Nat. Commun.* 7, 11067 (2016). The first structure of an inactive long Ago (RsAgo), accompanied by functional analysis of mutations in its structural elements involved in guide and target interactions.

43. Kaya, E. et al. A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl Acad. Sci. USA* 113, 4057–4062 (2016).

44. Matsumoto, N. et al. Crystal structure of silkworm PIWI-clade Argonaute Siwi in its catalytic core. *Genes Dev.* 27, 1342–1353 (2013). The crystal structure of human Argonaute2. *Science* 336, 1037–1040 (2012).

45. Willkomm, S., Zander, A., Gust, A. & Grohmann, D. A prokaryotic twist on Argonaute. *Cell Rep.* 22, 1438–1445 (2017). The crystal structure of human Argonaute2. *Science* 336, 1037–1040 (2012).

46. Sheu-Gruttadauria, J. & MacRae, I. J. Structural basis for ARGONAUTE slicer activity. *Proc. Natl Acad. Sci. USA* 111, 761–766 (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.

47. Globyte, V., Kim, S. H. & Joo, C. Single-molecule view of small RNA-guided Argonaute slicing complexes. *Cell Rep.* 23, 453–462 (2018). This study reveals the effects of mismatches and bulges on the structure of complexes of RsAgo on target binding and Ago recycling; it shows that imperfect targets can promote dissociation of guide molecules from RsAgo.
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 RNA. PLoS Pathog. 10, e004492 (2016).
87. Miesen, P., Joosten, J. & van Rij, R. P. piRNAs go viral: arborica-derived piRNAs in vector mosquitos. 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 mRNA methylation in mice. Mol. Cell. 31, 785–799 (2008).
91. Sienski, G., Donertas, D. & Brennecke, J. Transcriptional silencing of transposons by Pwii and maelfrom and its impact on chromatin state and gene expression. Cell 151, 964–980 (2012).
92. Proshkin, S., Rahmouni, A. R., Mironov, A. & Nadler, 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 2 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 genomics-based approach to the classification of bacterial CRISPR systems. J. Mol. Biol. 404, 331–431 (2010).
101. 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 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.