A long look at short prokaryotic Argonautes

Trends in Cell Biology

Koopal, Balwina; Mutte, Sumanth K.; Swarts, Daan C.

https://doi.org/10.1016/j.tcb.2022.10.005

Trends in Cell Biology

CellPress

Review

A long look at short prokaryotic Argonautes

Balwina Koopal,[1] Sumanth K. Mutte,[1] and Daan C. Swarts [1,*]

Argonaute proteins (Agos) use small 15–30 nucleotide-long guides to bind and/or cleave complementary target nucleic acids. Eukaryotic Agos mediate RNA-guided RNA silencing, while 'long' prokaryotic Agos (pAgos) use RNA or DNA guides to interfere with invading plasmid and viral DNA. Here, we review the function and mechanisms of truncated and highly divergent 'short' pAgos, which, until recently, remained functionally uncharacterized. Short pAgos have retained the Middle (MID) and P-element-Induced Wimpy Testis (PIWI) domains important for guide-mediated target binding, but lack the ability to cleave their targets. Instead, emerging insights reveal that various short pAgos interact with distinct accessory 'effector' enzymes. Upon guide-mediated detection of invading DNA by short pAgos, their associated effector enzymes kill the host cell and, consequentially, prevent spread of the invader.

Highlights

In all domains of life, Argonaute proteins (Agos) use short nucleic acid guides to bind complementary target nucleic acids in a sequence-specific manner.

Whereas eukaryotic Agos mediate RNA-guided RNA targeting, prokaryotic Agos (pAgos) show distinct guide/target preferences (DNA/RNA), have varied domain compositions, and associate with a wide range of auxiliary proteins.

Not only have full-length pAgos (long pAgos) been shown to target and degrade invading DNA, but also the role and mechanisms of the more abundant truncated 'short' pAgos have recently been elucidated.

Short pAgos have lost the ability to cleave targets; instead they genetically associate with a wide variety of putative effector proteins, including nucleases, NAD(P)ases, and depolarization-inducing membrane enzymes.

Rather than directly degrading their targets, characterized short pAgos form complexes with their associated effector enzymes to trigger cell death upon guide-mediated recognition of invading DNA and, thus, function as abortive infection systems.

Argonaute proteins are found in all domains of life

In all domains of life, Agos use small (15–30 nucleotides) oligonucleotides as guides to bind complementary nucleic acid targets. Eukaryotic Agos (eAgos) are the key effector enzymes in RNA silencing pathways and can be subdivided in two main clades: eAGO and ePIWI (lineage-specific eAgos are not discussed here). eAGOs generally bind small interfering RNA (siRNA) or microRNA (miRNA) guides generated by Dicer and/or Drosha nucleases [1]. Together with accessory proteins, eAGO forms an RNA-induced silencing complex (RISC), which silences mRNA targets to regulate gene expression [1]. ePIWIs bind PIWI-interacting RNA (piRNA) guides generated from longer genomic transcripts [2]. ePIWIs mainly silence transposons by cleaving their transcripts or by recruiting accessory proteins that induce heterochromatin formation [2–4].

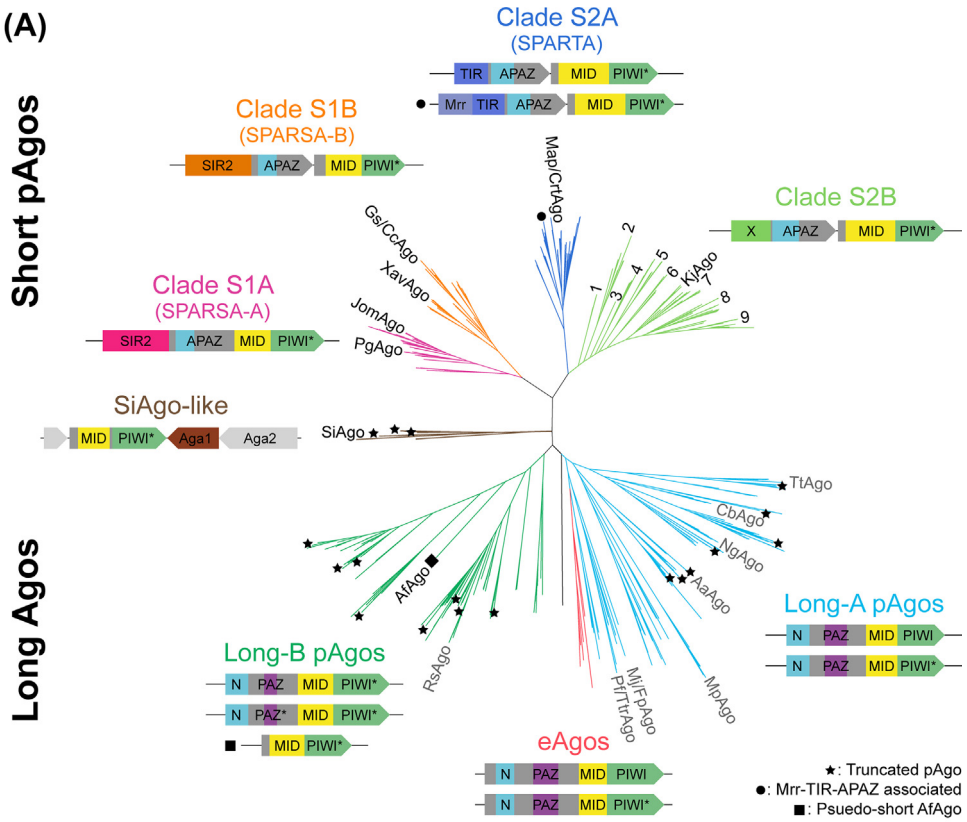
Compared with eAgos, pAgos are highly diversified in sequence and domain composition [5–7]. They show a patchy distribution over the bacterial and archaeal phyla [5] and are often associated with other host defense genes in so-called 'defense islands': clusters of genes related to prokaryotic defense [7,8]. Based on their genetic association with other defense genes, their (predicted) nuclease activity, and because they are frequently subjected to horizontal gene transfer, it was hypothesized that pAgos have a role in host defense against invading DNA, such as plasmids and viruses [7]. Based on their phylogeny, pAgos can be subdivided in long-A, long-B, and short pAgos (Figure 1A). Long-A and long-B pAgos have a canonical bilobed domain composition comprising the N terminal (N), PIWI-Argonaute-Zwille (PAZ), MID, and PIWI domains similar to eAgos [9–12] (Figures 1 and 2A–C; Box 1). While this suggests that Agos in all domains of life rely on similar mechanisms, proteins that are typically required in eukaryotic RNA silencing pathways (e.g., Dicer) have not been identified in prokaryotes [13]. Therefore, it is likely that pAgos execute different functions. By contrast, short pAgos lack the N- and PAZ domains and, thus, are comprise the MID and PIWI domains only (Figures 1 and 2D–I).

Of the long-A pAgos, 79% have an intact catalytic DEDX tetrad (Box 1) in the PIWI domain, and most characterized long-A pAgos are capable of DNA guide-mediated target DNA cleavage [12,14–18]. However, at least *in vitro*, one long-A pAgo uses RNA guides to target DNA [19], others use DNA guides to target RNA [20,21], and some do not have a clear preference for a

[1]Laboratory of Biochemistry, Wageningen University, 6708, WE, Wageningen, The Netherlands

*Correspondence: daan.swarts@wur.nl (D.C. Swarts).

specific type of guide and target [22,23]. In accordance with earlier hypotheses [7], catalytically active long-A pAgos have been implicated in host defense against invading nucleic acids, such as plasmids and viruses [14,15,24,25]. In addition, long-A pAgos function beyond host defense by stimulating homologous recombination [26,27] and assisting DNA replication [28]. No catalytically inactive long-A pAgos have been characterized to date. Less is known about long-B pAgos.



(See figure legend at the bottom of the next page.)

All long-B pAgos lack the catalytic DEDX tetrad and the PAZ domain is often partially truncated (PAZ*) [5] (Box 1). The only canonical long-B pAgo characterized to date is that of *Rhodobacter sphaeroides* (RsAgo), which uses RNA guides to bind DNA targets [29]. Despite its lack of nuclease activity, RsAgo reduces the expression of plasmid-encoded genes and stimulates plasmid degradation in *Escherichia coli* [29]. While the underlying mechanism is unclear, it was hypothesized that binding of RsAgo to target plasmids may cause their transcriptional silencing and/or recruitment of nucleases that degrade the plasmids [29]. Another (non-canonical) long-B pAgo from *Archaeoglobus fulgidus* (AfAgo) has been used as a structural model for Ago and Ago–nucleic acid interactions [30,31]. However, AfAgo is truncated: it comprises the MID and PIWI domains only (Figure 2I). Given that it can be considered to be a short pAgo, it is discussed in more detail below.

While pAgo research has predominantly focused on the eAgo-like long-A pAgos, many pAgos (59%) are 'short pAgos' (Figure 1A) [5]. Similar to long-B pAgos, all short pAgos lack the DEDX tetrad required for target cleavage. Instead, they genetically associate in operons with putative enzymes previously predicted to be nucleases (Figures 1 and 2D–G) [5–7]. Given that the MID-PIWI domains of eAgos were shown to be sufficient for guide-mediated target binding [32,33], it has been hypothesized that short pAgos function in a modular host defense system, in which they act as guide-mediated target binders, while relying on the associated enzymatic domains for target degradation [6]. In this review, we discuss recent studies that show that the functional mechanisms of short pAgos and their associated effector enzymes are fundamentally distinct from those of long pAgos and eAgos [34–37].

Phylogeny of short pAgos

Most short pAgos form a distinct phylogenetic clade and strictly associate with proteins containing an 'analog of PAZ' (APAZ) domain [5,6,36,38] (Figure 1A and Box 1). However, we also find truncated pAgos scattered over different branches of the long-A and long-B clades (Figure 1A) [5]. This implies that loss of the N- and PAZ domains occurred multiple times during the evolution of pAgos [5]. Consequentially, not all short pAgos are phylogenetically related: the short pAgo from *Sulfolobus islandicus* (SiAgo) and homologs thereof form a clade of pAgos that does not cluster with either long or short pAgos (the SiAgo-like clade; Figure 1A), and the short pAgo AfAgo clusters with long-B clade pAgos (Figure 1A). SiAgo and AfAgo are not associated with APAZ domains, and rely on distinct functional mechanisms (described below). Therefore, from here onward, we refer to the phylogenetically clustered short pAgos associated with APAZ domains as 'short pAgos' and to other truncated pAgos as 'pseudo-short pAgos'.

Of the short pAgos, 94% are encoded by bacteria [proteobacteria (pseudomonadota): 54%, Bacteroidetes: 22%, Actinobacteria: 8%, other bacterial phyla combined: 10%] and only 6%

Figure 1. Phylogeny of prokaryotic Argonaute (Ago) proteins. (A) Maximum likelihood-based unrooted phylogenetic tree containing all prokaryotic (bacterial and archaeal) Ago (pAgo) homologs identified in the RefSeq database containing at least scaffold-level assemblies, as well as selected experimentally characterized pAgos, eukaryotic Agos (eAgos), and *Sulfolobus islandicus* pAgo (SiAgo) homologs. Analysis performed as previously reported [36]. Clade S1A, sirtuin (SIR)2–analog of PIWI-Argonaute-Zwille (APAZ)–pAgo fusion (SPARSA-A); Clade S1B, operon with pAgo and SIR2-APAZ (SPARSA-B); Clade S2A, operon with pAgo and Toll-interleukin receptor (TIR)-APAZ (SPARTA); Clade S2B, operon with pAgo and APAZ fused to one of various domains (denoted 'X'), including Mrr, DUF4365, and DHS-like domains [see also (B)]. ★, truncated pAgos in Long-A, Long-B, and SiAgo-like clades; ●, short pAgos associated with Mrr-TIR-APAZ. ■, pseudo-short AfAgo. The fully annotated tree can be downloaded from Mendeley Data (https://dx.doi.org/10.17632/mpgn9by7z2.1). (B) Typically associated effector domains and distribution of 297 short pAgos from (A). Typically putative associated effectors domains as identified by InterPro (identified in >50% of short pAgo systems in this specific clade). Schlafen/Alba, IPR038461; RecG, IPR038475; DHS-like, IPR029035; DUF4365, IPR025375; ?, effector domain not identified by InterPro; *, the Mrr-like domain of certain S2B effector proteins is not identified by InterPro but is based on earlier studies [42] and AlphaFold2 predictions (see Figure 2 in the main text) [41,102]. Note that the short pAgo (sub) clades contain a limited number of representatives, and might be found in prokayotes beyond the indicated distribution.

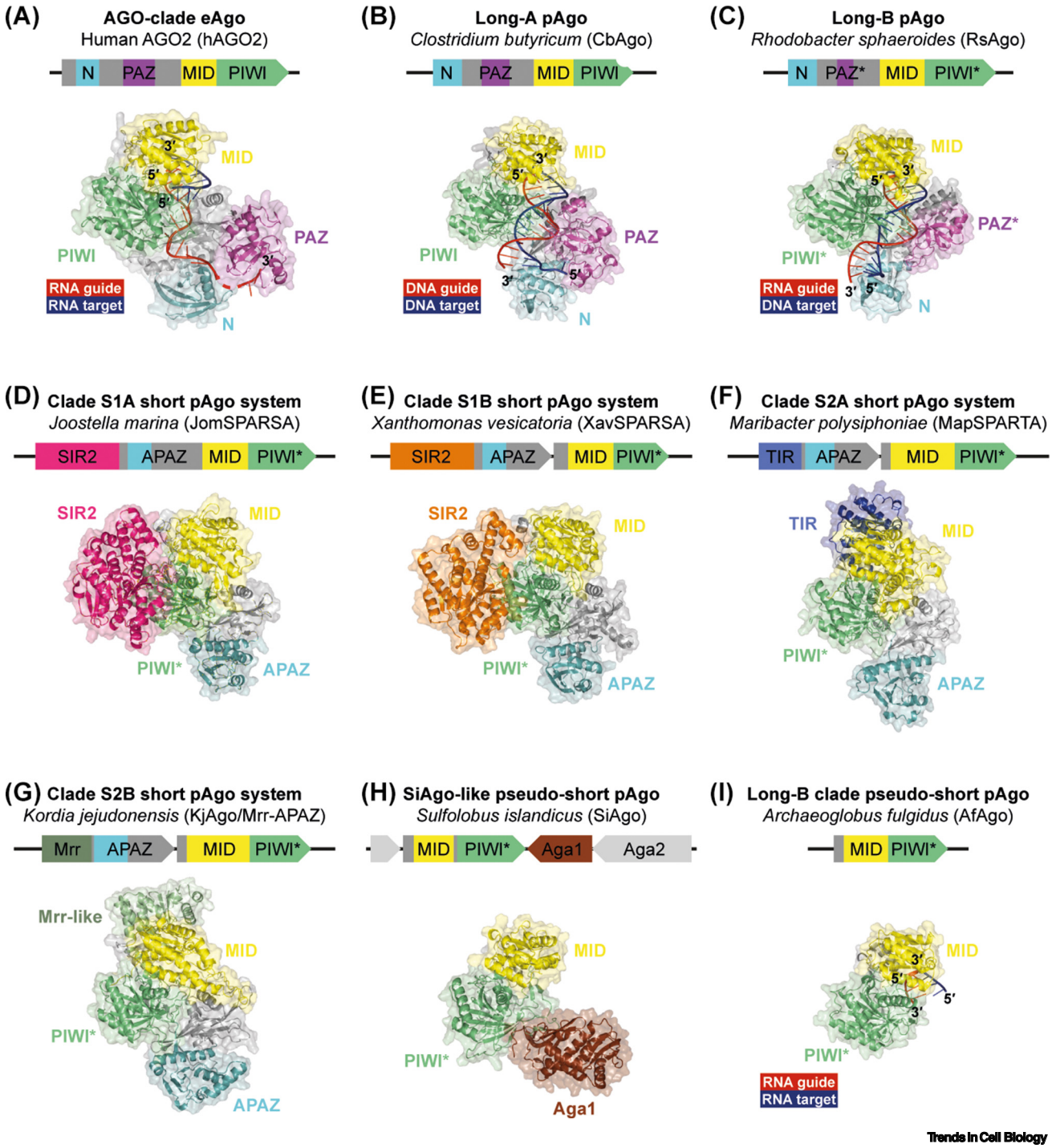


Figure 2. Operon structure, domain composition, and (predicted) structural architecture of prokaryotic Argonaute (pAgo) proteins. Crystal structures of (A) human AGO2 [Protein Data Bank identifier (PDB): 4W5O], (B) *Clostridium butyricum* long-A pAgo (CbAgo; PDB: 6QZK), (C) *Rhodobacter sphaeroides* long-B pAgo (PDB: 5AWH), and (I) *Archaeoglobus fulgidus* truncated long-B pAgo (PDB: 1YTU) in complex with guide and target strands, and Alphafold2-predicted models [41] for (pseudo-)short pAgo systems from distinct phylogenetic clades, including (D) clade S1A *Joostella marina* SPARSA (JomSPARSA), (E) clade S2B *Xanthomonas vesicatoria* SPARSA (XavSPARSA), (F) clade S2A *Maribacter polysiphoniae* SPARTA (MapSPARTA), (G) the clade S2B *Kordia jejudonensis* (KjAgo/Mrr-APAZ) system, and (H) the pseudo-short *Sulfolobus islandicus* (SiAgo-SiAga1-SiAga2) system. The related (predicted) domain composition (with corresponding colors) and operon structure are given above each structural model. Explanation of domain functions are given in Box 1. Alphafold2 predictions can be downloaded from Mendeley Data (https://dx.doi.org/10.17632/mpgn9by7z2.1).

Box 1. Characteristics of pAgo domains and other structural features of Agos

MID

The MID domain forms a pocket that, in most pAgos, contains four conserved residues (Y/R, K, Q, and K [5]) and coordinates a divalent cation [75] which together anchor the 5′-phosphate (5′-P) group of the guide. Certain pAgos preferentially bind 5′-hydroxylated (5′-OH) guides instead, which is attributed to a more hydrophobic binding pocket and absence of the divalent cation [5,11], while other pAgos show no clear preference for 5′-P or 5′-OH guides [21,22]. The 5′-nucleotide of the guide is also bound in the MID domain pocket (sometimes in a sequence-specific manner [88]) and, therefore, is unavailable for base-pairing with the target.

PIWI(*)

The PIWI domain is homologous to RNAse H, which cleaves the RNA strand of RNA/DNA substrates [89]. In Agos, the PIWI domain coordinates and preorders the 'seed' nucleotides of the guide (nucleotides 2–7 or 2–8) in a helical conformation, with bases exposed to the solvent [90]. This promotes guide/target base pairing by lowering the entropic costs of duplex formation. PIWI is also the domain responsible for target cleavage (slicing): slicing Agos contain a DEDX motif (see below) in the PIWI domain, while this motif is mutated in Agos that rely on target binding for their function. PIWI domains lacking the DEDX tetrad are referred to as PIWI*.

DEDX catalytic tetrad

The DEDX motif (where X denotes D, H, or K) found in the PIWI domain of slicing Agos facilitates target strand cleavage. The glutamic acid residue is located on a structural feature termed the glutamate finger [91], which completes the catalytic tetrad upon target binding. The DEDX tetrad coordinates two divalent cations that catalyze hydrolysis of the target strand phosphate backbone between nucleotides 10 and 11 [92]. Most eAgos from the AGO clade, some long-A pAgos, and long-B and (pseudo-)short pAgos have mutations in the DEDX motif that abolish cleaving activity [5]. However, even in slicing Agos, target cleavage can be affected by mismatches between the guide and the target [93,94].

MID-PIWI (PIWI lobe)

The MID and PIWI domains together form the PIWI lobe, which is sufficient for guide-mediated target binding and cleavage [32,33]. In line with their functional importance, these domains show the highest conservation over Agos from all domains of life, and are the only domains found in all Ago family proteins [5,6]. Short pAgos generally comprise the MID-PIWI domains only.

N domain

The N domain is the least conserved domain in Agos and functions in guide loading and target cleavage: It acts as a wedge to facilitate removal of the passenger strand during guide loading and removal of cleaved target strands in slicing Agos [95,96].

PAZ(*)

The PAZ domain interacts with the 3′ end of the guide, protecting it from degradation [33]. It usually displays no sequence preference, although PAZ domains of ePIWIs specifically recognize methylated 3′ ends [97]. Upon guide-mediated target binding, the 3′-end of the guide is released from the PAZ domain [98,99]. The associated conformational changes result in target cleavage. All long-B pAgos and some long-A pAgos have a truncated PAZ domain, which is referred to as PAZ* [5]. In PAZ*, the guide 3′ end binding pocket is at least partially lost [100,101]. This might allow for more extensive duplex formation between the guide and the target [100,101].

N-PAZ (N lobe)

The N and PAZ domains together form the N lobe, which is not required for guide-mediated target binding but enhances on-target cleavage specificity [32,33]. All (pseudo-)short pAgos lack the N lobe [5].

APAZ

Short pAgos lack both the N and PAZ domains, but are generally co-encoded in an operon or fused to an analog of PAZ (APAZ) domain-containing protein. The APAZ domain is usually fused to a variable catalytic domain (see Figure 2 in the main text) [5–7]. The APAZ domain was originally thought to be a functional analog of the PAZ domain [7]. However, other predictions [39,40] as well as Alphafold2 modeling (see Figure 2 in the main text) show that at least part of the APAZ domain is instead homologous to the N domain. Furthermore, the Alphafold2 models suggest that APAZ assumes the same position with respect to the MID-PIWI domains as the N-PAZ domains in long pAgos. The function of APAZ is currently unknown: it could have a 'wedging' function, similar to the N domain, control short pAgo-mediated activation of the effector domain, or it could have another unknown function.

by archaea (Euryarchaeota: 6%, TACK group archaea: 1%) (Figure 1B). Short pAgos can be divided into four phylogenetic clades: S1A, S1B, S2A, and S2B [36], and, based on their phylogeny, clade S2B short pAgos are further subdivided into nine subclades (S2B-1 to S2B-9; Figure 1). The different (sub)clades of short pAgos are typically found in only a couple of prokaryotic phyla (Figure 1B).

In all clades, short pAgos are encoded in operons that also encode an APAZ domain (Box 1). Initially, APAZ was predicted to functionally replace the PAZ domain in long pAgos [7], but later studies suggested that it is homologous to the N domain of long Agos [39,40]. AlphaFold2-generated models [41] of short pAgo systems corroborate that the N domain and APAZ are homologous and assume the same position with respect to the MID-PIWI lobe (Figure 2; see also [35]). The N terminus of APAZ is generally fused to a (putative) catalytic domain [5] (Figure 2D–G). In clades S1A and S1B, APAZ is fused to a Silent information regulator 2 (SIR2, also known as Sirtuin) domain (Figure 2D,E). In clade S2A, APAZ is fused to a Toll-interleukin receptor (TIR) domain (Figure 2F). In the different S2B subclades, APAZ is fused to one of various domains (e.g., Schlafen/Alba, Mrr-like, DUF4365, or RecG/DHS-like) (Figures 1B and 2G) [5,36,42]. In clade S1A, the APAZ domain-containing protein is fused to the N terminus of short pAgo (Figure 2D). In other clades, the APAZ domain-containing proteins are encoded by a separate gene upstream of short pAgo (Figure 2E–G), which suggests functional interdependence. We refer to short pAgo and its associated APAZ domain-containing protein as a 'short pAgo system'.

Similar to short pAgos, pseudo-short pAgos genetically colocalize with other proteins. For example, SiAgo is encoded in an operon with a predicted transcriptional regulator, while two other proteins are encoded on the opposite DNA strand [34] (Figure 2H). The presence of these three genes in close proximity to SiAgo is conserved for SiAgo homologs across different strains [34]. AfAgo has so far only been characterized as a standalone protein [43], despite it being encoded in an operon with two hypothetical proteins. Combined, these data suggest both short and pseudo-short pAgos function in conjunction with proteins encoded in their genomic context.

Indeed, recent studies confirm that (pseudo-)short pAgos form complexes with their associated proteins to protect their host against invading DNA, such as plasmids and viruses [34–36]. Rather than triggering invader DNA degradation akin to long pAgos, these (pseudo-)short pAgo systems function as abortive infection systems [44] that kill their host to prevent replication and spread of the invader to other cells [34–37] (Figure 3). The diversified mechanisms on which these (pseudo-) short pAgo systems rely are detailed below.

Functions and mechanisms of short pAgos

SPARSA systems

Short prokaryotic Argonaute/SIR2-APAZ (SPARSA, also known as Sir2/Ago) systems found in clades S1A (fused SIR2-APAZ-pAgo) and S1B (co-encoded SIR2-APAZ and pAgo) are typified by the fusion of APAZ to a SIR2 domain (Figures 1 and 2D,E) [5,36]. In eukaryotes, SIR2 proteins are involved in NAD^+-dependent protein or histone deacetylation or ADP ribosylation, which can have implications for chromatin formation [45], DNA repair [46], and programmed cell death [47]. In prokaryotes, SIR2 domains are NAD^+-dependent deacetylases with a role in stress resistance [48] and that alter the immune response of their hosts through histone deacetylation [49]. In addition, it was shown that several prokaryotic abortive infection systems use SIR2 to deplete NAD^+ upon detection of invading DNA [37,50]. This causes cell death of invaded cells, thereby preventing spread of the invader and providing population-based immunity (Figure 3).

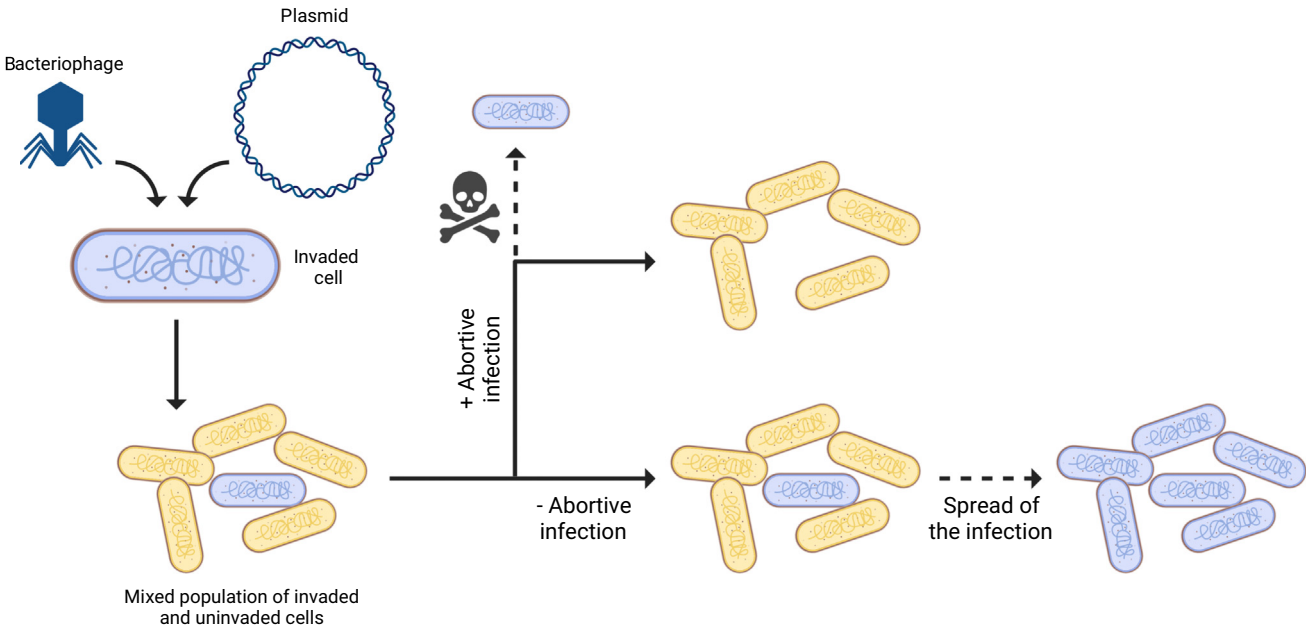


Figure 3. General mechanism of abortive infection systems. Invading nucleic acid, such as phages or plasmids, can enter prokaryotic cells and, in the absence of abortive infection systems, may either persist in the population, providing a metabolic burden, or spread to neighboring cells through lysis and infection of neighboring cells (phages) or through conjugation (conjugative plasmids). Abortive infection systems prevent the spread of invading nucleic acids by sensing the invader and subsequently killing their host cell. This removes invaded cells from the population. Figure created using BioRender (https://biorender.com/).

Fused SPARSA systems from clade S1A [*Paraburkholderia graminis* (PgSPARSA) and *Joostella marina* (JomSPARSA)] as well as co-encoded clade S1B SPARSA systems [*Geobacter sulfurreducens* (GsSPARSA), *Caballeronia cordobensis* (CcSPARSA) and *Xanthomonas vesicatoria* (XavSPARSA)] deplete NAD^+ upon detection of invader DNA [35–37]. PgSPARSA, GsSPARSA, and CcSPARSA provide protection against double-stranded (ds)DNA phage lambda-vir, while GsSPARSA and PgSPARSA also protect against dsDNA phage SECphi27. GsSPARSA and CcSPARSA also provide protection against transformation of plasmids containing a CloDF13 origin of replication (*ori*), but not against plasmids containing other *oris* (ColA, p15A, and RSF1030) [35]. For GsSPARSA, it was shown that invader interference critically relies on both guide binding by GsAgo and GsSIR2-APAZ NADase activity [35].

While GsSIR2-APAZ and GsAgo are encoded separately, the proteins form a heterodimeric complex [35]. When provided with an RNA guide and complementary single-stranded (ss)DNA target *in vitro*, the complex is activated and degrades NAD^+ (Figure 4). *In vivo*, GsSPARSA associates with small 20-nucleotide-long guide RNAs with a 5′-AU sequence. While most guides are derived from genome-encoded genes, guides derived from plasmid-encoded transcripts are mostly obtained from their *oris*, suggesting that guide acquisition relies on RNA-dependent priming plasmid replication. Replication of the phage lambda-vir also relies on RNA-dependent priming [51], which might provide clues to how SPARSA obtains guides that facilitate specific detection of plasmid and viral invaders. Although other viruses and plasmids relying on the same replication mechanism are not affected by SPARSA [35], many prokaryotic immune systems only work against a subset of viruses and under specific conditions [50,52]. Combined, this shows that SPARSA provides population-based immunity by triggering cell death through NAD^+ depletion upon RNA-guided invader DNA detection (Figure 4, Key figure).

Key figure

Schematic of host defense against invading DNA by (pseudo-) short prokaryotic Argonaute (pAgo) systems

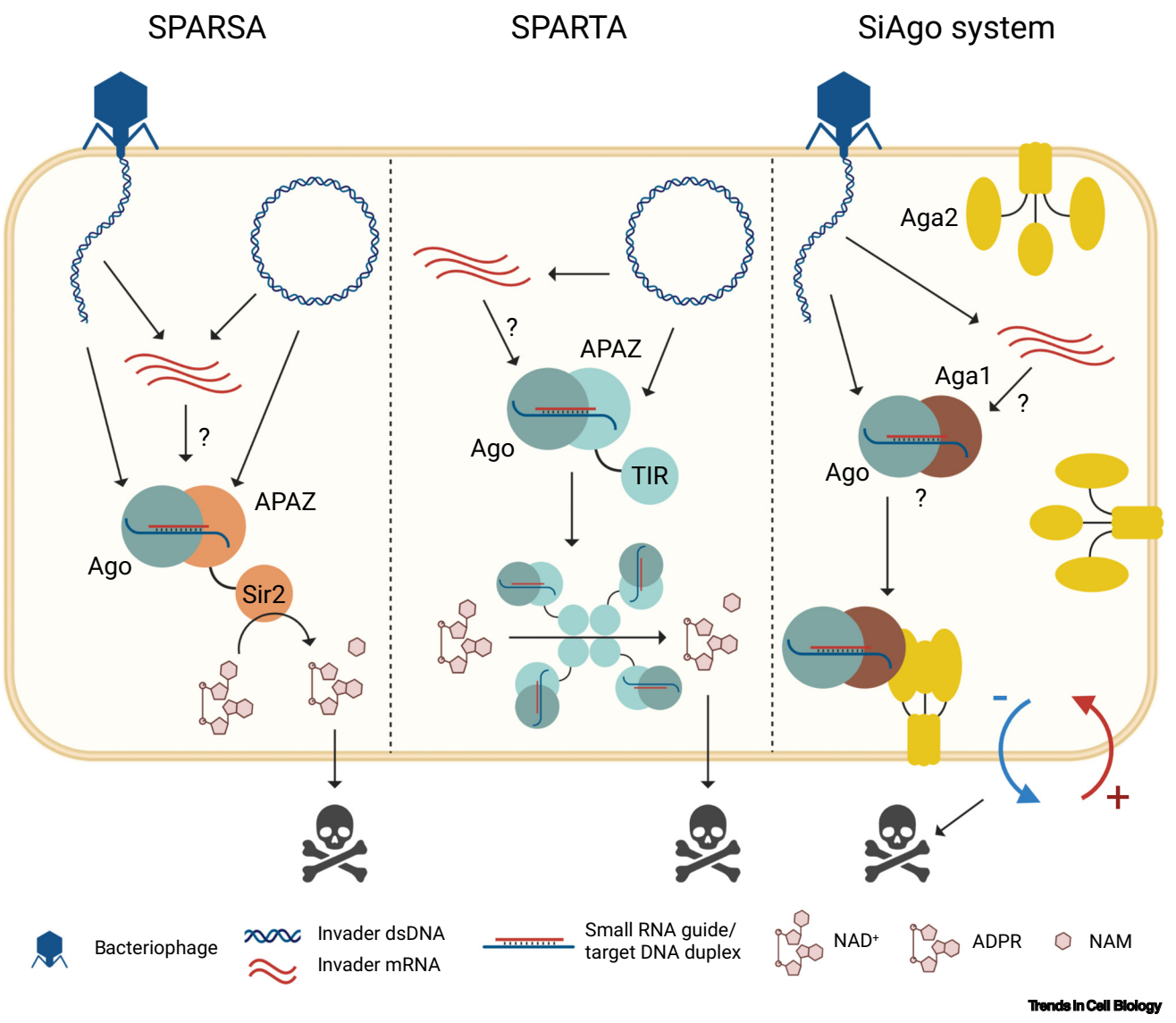


Figure 4. Short pAgos from sirtuin (SIR)2–analog of PIWI-Argonaute-Zwille (APAZ) (SPARSA) systems (left), Toll-interleukin receptor (TIR)-APAZ (SPARTA) systems (middle), and pseudo-short *Sulfolobus islandicus* (Si)Ago systems (right) form heterodimeric complexes with their accessory effector proteins. Invading viral and plasmid double-stranded (ds)DNA (blue) enters the cell, after which the (pseudo-)short pAgo system acquires small guide RNAs from invader RNA transcripts (red) by an unknown mechanism. The guide RNA facilitates sequence-specific recognition of invading DNA strands (blue), which results in catalytic activation of the effector domains: SIR2 (SPARSA), TIR (SPARTA), or Aga2 (SiAgo system). In SPARSA and SPARTA, SIR2/TIR activation leads to conversion of $NAD(P)^+$ to NAM and ADPR(P). In SPARTA, catalytic activation requires tetramerization of four guide/target-bound heterodimeric SPARTA complexes. Upon invader detection, the SiAgo-SiAga1 complex is recruited to the membrane protein SiAga2, which induces membrane depolarization. Both $NAD(P)^+$ depletion and membrane depolarization induce cell death, which protects neighboring cells and cures metabolically costly plasmids from the population (see also Figure 3 in the main text). Figure created using BioRender (https://biorender.com/).

SPARTA systems

Short prokaryotic Argonaute/TIR-APAZ (SPARTA) systems comprise clade S2A and are typified by the fusion of APAZ to a TIR domain (Figures 1 and 2F) [5,36]. TIR domains were originally identified as scaffolding proteins associated with eukaryotic receptor proteins [53], but later they were found to have NADase activity [54], which is important for their immune function in both eukaryotes [55–60] and prokaryotes [50,61–63]. While some TIR domains deplete cellular NAD^+ and function as abortive infection systems (Figure 3), others generate signaling molecules (e.g., cyclic ADPR and v-ADPR) to trigger downstream effects [50,54,55,64].

Akin to SPARSA, SPARTA systems from *Crenotalea thermophila* (CrtSPARTA) and *Maribacter polysiphoniae* (MapSPARTA) degrade NAD^+ (and $NADP^+$) in the presence of plasmid DNA [36]. In contrast to SPARSA, however, no direct interference with plasmid transformation was observed; instead, SPARTA decreases the viability of plasmid-invaded cells through $NAD(P)^+$ depletion, removing invaded cells from bacterial cultures. In addition, no SPARTA-mediated immunity against bacteriophages was observed. While both SPARSA and SPARTA are short pAgo systems that degrade NAD^+, these differences suggest that SPARSA and SPARTA rely on different mechanisms to detect invaders. However, the different experimental conditions under which SPARSA and SPARTA have been functionality characterized could also have had a role in the observed differences. Further research is required to establish differences and similarities between SPARSA and SPARTA mechanisms.

As expected, the NAD(P)ase activity of SPARTA is attributed to the TIR domain of TIR-APAZ, which is catalytically active in absence of the short pAgo partner [36]. In contrast to TIR domains, which generate signaling molecules c-ADPR or v-ADPR [50,54,55,64], SPARTA converts $NAD(P)^+$ to noncyclic ADPR(P) and NAM. In heterodimeric short pAgo/TIR-APAZ complexes, this activity is quenched, implying that short pAgo controls the activity of TIR-APAZ. Through *in vitro* experiments, it was determined that guide RNA-mediated target ssDNA binding by SPARTA induces tetramerization of four guide/target-bound SPARTA heterodimers, which reinstates the NAD(P)ase activity of the TIR domain (Figure 4). Oligomerization-dependent activation of TIR domains is a general mechanism conserved from prokaryotic to eukaryotic immune systems [61,62,65,66].

In vivo, SPARTA associates with 15–25-nucleotide-long guide RNAs with a 5′-A that are mostly derived from highly transcribed genes encoded on multicopy plasmids [36]. In line with this observation, plasmids with a high copy number and/or encoding highly transcribed genes activate SPARTA, whereas low-copy number plasmids lacking highly transcribed genes do not. Invading DNAs exploiting their host cell for propagation are often present in high copy numbers and highly transcribed. Combined, this implies that SPARTA senses invader activity through the high abundance of both their DNA and their RNA transcripts, which triggers its NAD(P)ase activity and, consequentially, cell death, thereby removing invaded cells from the bacterial population (Figure 4).

S2B-clade short pAgo system from *Kordia jejudonensis*

S2B clade short pAgo systems are the most diverse in terms of distinct protein domains that are fused to APAZ, but they are also the least explored: only the S2B system of *Kordia jejudonensis* has been studied *in vitro* [42] (Figures 1 and 2G). The KjAgo-associated APAZ protein is fused to an Mrr-like domain [36], which is homologous to *E. coli* Mrr, which acts as a methylation-dependent DNA nuclease [67] (Figure 2). Akin to short pAgos from other clades [35,36], KjAgo forms a heterodimeric complex with Mrr-APAZ [42]. *In vitro*, this complex catalyzes not only RNA/DNA-guided cleavage of ssDNA targets, but also nonspecific cleavage of ssDNA and dsDNA [42]. No functional role of the KjAgo/Mrr-APAZ system has yet been determined, but it is conceivable that the KjAgo system mediates prokaryotic immunity akin to other short pAgo systems. Yet, the

results suggest that the KjAgo system relies on mechanisms that are clearly different from SPARSA and SPARTA: upon its activation, the KjAgo/Mrr-APAZ system might indiscriminately degrade nucleic acids to shut down the cell, an abortive infection strategy that has been described for other prokaryotic immune systems (e.g., CBASS, CRISPR-Cas13, and type III CRISPR systems) [68–70]. Alternatively, the Mrr-like domain might compensate for loss of the catalytic activity of the PIWI domain and specifically cleave guide-bound (invader) nucleic acids analogous to long-A pAgos and most CRISPR-Cas systems [14,24,71–73].

Taken together, these studies show that short pAgos and their associated APAZ domain-containing proteins are fused or form heterodimeric complexes (Figures 2D–G and 4). The fact that fusion of short pAgos with their associated APAZ effectors has occurred multiple times in evolution underscores the importance of their complexation. In short pAgo systems, the short pAgo acts as a 'sensor' that facilitates guide-mediated recognition and binding of nucleic acid invaders. Upon target binding, the APAZ-fused domain is catalytically activated and acts as an 'effector'. In SPARSA and SPARTA systems, target detection results in unleashed SIR2- or TIR-mediated NAD(P)ase activity that consequentially kills the host of the invaded cells (Figures 3 and 4). For clade S2B short pAgos, the exact catalytic mechanism activated and the consequences thereof remain to be determined.

Functions and mechanisms of pseudo-short pAgos

Sulfolobus islandicus pAgo

The archaeal SiAgo and homologs thereof do not cluster with either long or short pAgos, but form a separate branch in the phylogenetic pAgo tree (Figure 1). SiAgo functionality relies on two genetically associated proteins (Figures 1 and 2H): Ago associated protein 1 (SiAga1) and Ago associated protein 2 (SiAga2) [34]. These three proteins function together as a prokaryotic immune system that protects its host against infection by the dsDNA virus SMV1 [34]. Akin to short pAgos, SiAgo and SiAga1 form heterodimeric complexes upon their co-expression [34]. However, SiAgo and SiAga1 alone do not confer defense against bacteriophages: they additionally rely on the effector Aga2. Aga2 is a membrane protein that forms large oligomeric complexes and binds phospholipids. While SiAgo-Aga1 complexes reside in the cytoplasm, Aga2 mostly localizes in the membrane. Upon viral infection, the SiAgo-Aga1 complex is directed to the membrane, triggering activation of Aga2. This results in loss of membrane polarity and consequentially triggers cell death in invaded cells [34].

While *in vitro* SiAgo appears to bind RNA guides, the SiAgo-SiAga1 complex associates with all guide/target type (DNA/RNA) combinations [34]. However, it is unknown what combination of guides/target type activates the system [34]. Nevertheless, a model is proposed wherein the SiAgo-SiAga1 complex recognizes invading nucleic acids through guide-target binding, consequentially activating Aga2 to initiate abortive infection [34]. Contrary to short pAgo systems, the SiAgo system does not appear to target plasmids in the reported experimental setup: overexpression of genes from plasmids does not cause substantial Aga2 activation. In conclusion, although it relies on different accessory proteins than short pAgos, the SiAgo system also mediates population-based immunity by abortive infection upon detection of invading DNA (Figure 4).

Archaeoglobus fulgidus pAgo

AfAgo is a truncated long-B pAgo comprising the MID and PIWI domains only [74] (Figures 1 and 2I). AfAgo was one of the first Ago proteins to be crystalized and served as a model to study Ago–nucleic acid interactions [75]. In these structures, AfAgo crystallizes as a homodimer, but the functional relevance of AfAgo dimerization remained unknown. A recent study showed that AfAgo also forms dimers in solution, and that dimerization is stabilized upon binding of dsDNA ends [43]. Given that each AfAgo component in the dimer can accommodate a separate

dsDNA end, AfAgo dimers can stimulate formation of DNA loops. This suggests that AfAgo is involved in DNA repair or integration of mobile genetic elements. In addition, long-A pAgos have been implicated in homologous recombination [26,27], but there are no indications that dimerization is required in this process. Although its natural preferences for DNA/RNA have not been studied, AfAgo interacts with both RNA and DNA *in vitro* [75]. It is also unknown whether AfAgo, similar to other pAgos, acts as a host defense system. Akin to other (pseudo-) short pAgos, AfAgo is encoded in an operon that also encodes hypothetical proteins. Further experimentation is required to determine the biological function of AfAgo.

Concluding remarks

Short pAgos and pseudo-short pAgos are prokaryotic Agos comprising the MID and PIWI domains only, lack catalytic activity, and can have different phylogenetic origins. Given that they retain their ability to use guides to bind complementary targets, they serve as 'sensors' that detect invading DNA. All (pseudo-)short pAgos functionally characterized thus far mediate abortive infection (Figure 3): they form a complex with 'effector' proteins that are activated upon guide-mediated invader detection, thereby killing their host to protect neighboring cells (Figure 4).

Despite the major advances in understanding of (pseudo-)short pAgo functionality, many aspects of their functionality remain mysterious (see Outstanding questions). It is unknown how (pseudo-)short pAgos interact with their effector partners. Given that the APAZ domain associated with short pAgos is homologous to the N domain found in long pAgos and eAgos, it might facilitate interactions with short pAgo and/or have a role in guide loading and/or target release. However, it is also conceivable that APAZ serves to activate the effector domain upon guide-mediated target binding by Ago, likely through a conformational change. Likewise, in the SiAgo system, SiAga1 may connect SiAgo to SiAga2 in response to guide-mediated target binding, which induces membrane depolarization by Aga2. In addition, the SiAgo/SiAga1 complex showed increased affinity for nucleic acids compared with the SiAgo monomer, which suggests that SiAga1 either aids in binding guide and/or target nucleic acids or induces a conformation of SiAgo to favor nucleic acid binding. Structure-derived insights into protein–protein and protein–nucleic acid interactions are required to reveal the role of APAZ in short pAgo systems and Aga1 in the SiAgo system, and could illuminate the mechanism by which (pseudo-)short pAgos control the activity of their associated effectors.

Short pAgos characterized to date utilize small RNA guides to DNA, but it remains largely unknown how (pseudo-)short pAgos distinguish self (genomic DNA) from nonself (invader DNA). Potentially, guide RNA biogenesis specifically yields invader-targeting guide RNAs. In contrast to that hypothesis, both short and long-B pAgos acquire guide RNAs from the entire transcriptome [29,35,36], which results in the loading of self-targeting guide RNAs. Yet, these observations were made while overexpressing the pAgos in heterologous hosts, and it cannot be ruled out that guide RNA biogenesis and loading is more specific in the natural host under native expression conditions. Eukaryotes have dedicated pathways for guide RNA biogenesis: in general eAGOs are loaded with miRNA or siRNA guides that are generated from dsRNA substrates, such as RNA hairpins or bidirectional transcripts, by the RNase III family proteins Dicer and/or Drosha [76–78]. PIWIs bind piRNA guides that are derived from transcripts of piRNA clusters to target transposon transcripts, or secondary piRNAs that are generated from targeted transposon transcripts in an ePIWI-dependent manner [79]. While prokaryotes lack Dicer and Drosha homologs, bacteria encode simpler RNase III family proteins [13], as well as an array of other housekeeping RNases that could have a role in guide RNA biogenesis. Analogous to the ePIWI pathway, CRISPR-Cas systems rely on CRISPR RNA guides (crRNAs) derived from transcripts of CRISPR loci to target nucleic acid invaders [80]. Although some pAgos are encoded in the context of Cas genes [5,6,38], no association has been found between crRNA guides and pAgos thus far. While short pAgos do not strictly rely on genome-encoded small RNAs [35,36], it can also not be

Outstanding questions

Does NAD(P)$^+$ depletion by SPARSA and SPARTA only trigger cell death or potentially also dormancy?

What metabolic and/or signaling pathways result in cell death after NAD(P)$^+$ depletion by SPARSA and SPARTA?

Many phages use counter-defense systems that disable prokaryotic host defense systems. Do anti-(pseudo-)(short) pAgo mechanisms exist, and how do they function?

Are other pseudo-short pAgos (i.e., truncated long-A and long-B pAgos) correctly annotated, are they functional, and do they genetically cluster with potential accessory proteins? What are their functions and mechanisms?

Certain TIR-APAZ proteins of SPARTA systems are additionally fused to an Mrr-like domain. What is the functional relevance of this domain?

In SPARSA systems of clade S1A (and in one SPARTA system), APAZ is fused to short pAgo rather than being co-encoded in an operon. Are there functional advantages and/or disadvantages of co-encoding versus fusion of short pAgos and their effector enzymes?

The S2B clade KjAgo has only been investigated *in vitro*. Furthermore, in the S2B clade, short pAgo-associated APAZ domains are fused to a large variety of putative effector domains. What are the function and mechanisms of S2B clade pAgo systems?

How do (pseudo-)short pAgos control the catalytic activity of their accessory effector enzymes?

Many long pAgos genetically associate with putative effector enzymes, some of which are shared between long and short pAgos. Do these long pAgos, similar to short pAgos, function as sensors, while their associated catalytic domains act as effectors? If so, are there advantages and/or disadvantages of having a long pAgo versus a short pAgo for guide-mediated target detection?

ruled out that certain pAgo hosts encode precursors of invader-targeting guide RNAs on the genome.

It is conceivable that the first pAgo was catalytically active. At the same time, catalytically inactive short pAgos are structurally less complex than long pAgos. What evolved first, catalytically inactive short pAgos or catalytically active long pAgos?

Besides preferential guide generation, the availability of target DNA might have a role in differentiating self and nonself DNA. All (pseudo-)short pAgos reported to date target ssDNA *in vitro*, but act on dsDNA invaders *in vivo* [34–36]. Being unable to unwind dsDNA targets, pAgos might rely on other enzymes or processes (e.g., replication or transcription) to make DNA susceptible for detection. If such process or enzymes are different for genomic and invading DNA, they could contribute to invader-specific targeting by pAgos. Therefore, resolving the mechanisms facilitating pAgo-mediated dsDNA recognition can contribute to understanding how pAgos distinguish self from nonself.

The majority of the vast (pseudo-)short pAgo-diversity awaits exploration. For example, certain S2A clade TIR-APAZ domains are additionally fused to an Mrr domain (Figure 2), and many S2B-clade short pAgos associate with domains of which the functions are not clear. For most pseudo-short pAgos, the genetic context is yet to be investigated. In addition, more distant pAgos homologs have been identified (e.g., PIWI-RE [6,81] and DdmE [82]). With the rapidly expanding metagenomics data available, it is not unlikely that additional clades containing distinct (pseudo-)short or long pAgos or more distantly related homologs thereof will be discovered in the near future. Characterization of pAgo variants will determine what functional roles they can fulfil, which mechanism they convey, and what evolutionary turns pAgos have taken.

Finally, short pAgo systems can be isolated and programmed with short (synthetic) RNA guides with sequences of choice, akin to CRISPR-Cas systems [83–86] and long pAgos [87]. Given that target binding triggers a specific catalytic activity, they could be repurposed for a range of sequence-specific applications. For example, SPARTA has been repurposed for the detection of ssDNA and dsDNA sequences [36]. Possibly, other pAgo systems with a distinct catalytic activity could be repurposed for targeted nucleic acid modification or genome editing. As such, the characterization (pseudo-)short pAgos not only uncovered a novel, highly diverse class of immune systems, but could also inspire a new generation of programmable molecular tools.

Acknowledgments

This work was supported by a Netherlands Organization for Scientific Research (NWO) VENI grant (016.Veni.192.072) and a European Research Council (ERC) starting grant (ERC-2020-STG 948783) to D.C.S.

Declaration of interests

D.C.S. and B.K. (together with Ana Potocnik) have submitted a patent application regarding the utilization of short pAgo systems for nucleic acid detection.

References

1. Meister, G. (2013) Argonaute proteins: functional insights and emerging roles. *Nat. Rev. Genet.* 14, 447
2. Ozata, D.M. *et al.* (2019) PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* 20, 89–108
3. Vaucheret, H. (2008) Plant ARGONAUTES. *Trends Plant Sci.* 13, 350–358
4. Sheu-Gruttadauria, J. and MacRae, I.J. (2017) Structural foundations of RNA silencing by argonaute. *J. Mol. Biol.* 429, 2619–2639
5. Ryazansky, S. *et al.* (2018) The expanded universe of prokaryotic argonaute proteins. *MBio* 9, e01935-18
6. Swarts, D.C. *et al.* (2014) The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.* 21, 743–753
7. Makarova, K.S. *et al.* (2009) Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol. Direct* 4, 29
8. Makarova, K.S. *et al.* (2011) Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. *J. Bacteriol.* 193, 6039
9. Sheng, G. *et al.* (2014) Structure-based cleavage mechanism of *Thermusthermophilus* argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl. Acad. Sci. U. S. A.* 111, 652–657
10. Rashid, U.J. *et al.* (2007) 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
11. Kaya, E. *et al.* (2016) A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl. Acad. Sci. U. S. A.* 113, 4057–4062
12. Hegge, J.W. *et al.* (2019) DNA-guided DNA cleavage at moderate temperatures by *Clostridium butyricum* Argonaute. *Nucleic Acids Res.* 47, 5809–5821

13. Shabalina, S.A. and Koonin, E.V. (2008) Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.* 23, 578–587
14. Swarts, D.C. *et al.* (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* 507, 258–261
15. Swarts, D.C. *et al.* (2015) Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.* 43, 5120–5129
16. Zander, A. *et al.* (2017) Guide-independent DNA cleavage by archaeal Argonaute from *Methanocaldococcus jannaschii*. *Nat. Microbiol.* 2, 17034
17. Lee, K.Z. *et al.* (2021) NgAgo possesses guided DNA nicking activity. *Nucleic Acids Res.* 49, 9926–9937
18. Ishino, Y. *et al.* (2021) Hyperthermophilic Argonaute from *Ferroglobus placidus* with specificity on guide binding pattern. *Microbiology* 12, 654345
19. Doxzen, K.W. and Doudna, J.A. (2017) DNA recognition by an RNA-guided bacterial Argonaute. *PLoS ONE* 12, 1–14
20. Lisitskaya, L. *et al.* (2022) Programmable RNA targeting by bacterial Argonaute nucleases with unconventional guide binding and cleavage specificity. *Nat. Commun.* 2, 1–15
21. Li, W. *et al.* (2022) A programmable pAgo nuclease with RNA target preference from the psychrotolerant bacterium *Mucilaginibacter paludism*. *Nucleic Acids Res.* 50, 5226–5238
22. Wang, L. *et al.* (2022) A bacterial Argonaute with efficient DNA and RNA cleavage activity guided by small DNA and RNA. *Cell Rep.* 41, 111533
23. Kropocheva, E. *et al.* (2021) A programmable pAgo nuclease with universal guide and target specificity from the mesophilic bacterium *Kurthia massiliensis*. *Nucleic Acids Res.* 49, 4054–4065
24. Kuzmenko, A. *et al.* (2020) DNA targeting and interference by a bacterial Argonaute nuclease. *Nature* 587, 632–637
25. Swarts, D.C. *et al.* (2017) Autonomous generation and loading of DNA guides by bacterial Argonaute. *Mol. Cell* 65, 985–998.e6
26. Fu, L. *et al.* (2019) The prokaryotic Argonaute proteins enhance homology sequence-directed recombination in bacteria. *Nucleic Acids Res.* 47, 3568–3579
27. Esyunina, D. *et al.* (2022) Specific targeting of plasmids with Argonaute enables genome editing. *bioRxiv* Published online April 14, 2022. https://doi.org/10.1101/2022.04.14.488398
28. Jolly, S.M. *et al.* (2020) Thermus thermophilus Argonaute Functions in the Completion of DNA Replication. *Cell* 182, 1545–1559.e18
29. Olovnikov, I. *et al.* (2013) Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell* 51, 594–605
30. Parker, J.S. *et al.* (2009) Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Mol. Cell* 33, 204–214
31. Ma, J.B. *et al.* (2004) Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* 429, 318–322
32. Dayeh, D.M. *et al.* (2018) Structural and functional analyses reveal the contributions of the C- and N-lobes of Argonaute protein to selectivity of RNA target cleavage. *J. Biol. Chem.* 293, 6308–6325
33. Hur, J.K. *et al.* (2013) Regulation of Argonaute slicer activity by guide RNA 3′ end interactions with the N-terminal lobe. *J. Biol. Chem.* 288, 7829–7840
34. Zeng, Z. *et al.* (2022) A short prokaryotic argonaute activates membrane effector to confer antiviral defense. *Cell Host Microbe* 30, 930–943
35. Zaremba, M. *et al.* (2022) Short prokaryotic Argonautes provide defence against incoming mobile genetic elements through NAD+ depletion. *Nat. Microbiol.* 7, 1857–1869
36. Koopal, B. *et al.* (2022) Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA. *Cell* 185, 1471–1486
37. Garb, J. *et al.* (2022) Multiple phage resistance systems inhibit infection via SIR2-dependent NAD+ depletion. *Nat. Microbiol.* 7, 1849–1856
38. Willkomm, S. *et al.* (2018) DNA silencing by prokaryotic Argonaute proteins adds a new layer of defense against invading nucleic acids. *FEMS Microbiol. Rev.* 42, 376–387
39. Burroughs, A.M. *et al.* (2013) Two novel PIWI families: roles in inter-genomic conflicts in bacteria and Mediator-dependent modulation of transcription in eukaryotes. *Biol. Direct* 8, 1–15
40. Willkomm, S. *et al.* (2017) Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein. *Nat. Microbiol.* 2, 17035
41. Bryant, P. *et al.* (2022) Improved prediction of protein-protein interactions using AlphaFold2. *Nat. Commun.* 13, 1–11
42. Kim, S. *et al.* (2020) Argonaute system of *Kordia jejudonensis* is a heterodimeric nucleic acid-guided nuclease. *Biochem. Biophys. Res. Commun.* 525, 755–758
43. Golovinas, E. *et al.* (2021) Prokaryotic Argonaute from *Archaeoglobus fulgidus* interacts with DNA as a homodimer. *Sci. Rep.* 11, 1–14
44. Lopatina, A. *et al.* (2020) Abortive infection: bacterial suicide as an antiviral immune strategy. *Annu. Rev. Virol.* 7, 371–384
45. Yeom, S. *et al.* (2022) Spreading-dependent or independent Sir2-mediated gene silencing in budding yeast. *Genes Genomics* 44, 359–367
46. Onn, L. *et al.* (2020) SIRT6 is a DNA double-strand break sensor. *eLife* 9, 1–26
47. Vendrell, A. *et al.* (2011) Sir2 histone deacetylase prevents programmed cell death caused by sustained activation of the Hog1 stress-activated protein kinase. *EMBO Rep.* 12, 1062–1068
48. Ma, Q. and Wood, T.K. (2011) Protein acetylation in prokaryotes increases stress resistance. *Biochem. Biophys. Res. Commun.* 410, 846–851
49. Mishra, P. *et al.* (2022) vp1524, a *Vibrio parahaemolyticus* NAD +-dependent deacetylase, regulates host response during infection by induction of host histone deacetylation. *J. Biochem.* 171, 673–693
50. Ofir, G. *et al.* (2021) Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature* 600, 116–120
51. Yoda, K.-Y. *et al.* (1988) RNA-primed intitiation sites of DNA replication in the origin region of bacteriophage lambda genome. *Nucleic Acids Res.* 16, 6531–6546
52. Doron, S. *et al.* (2018) Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* 359, eaar4120
53. Pawson, T. and Nash, P. (2003) Assembly of cell regulatory systems through protein interaction domains. *Science* 300, 445–452
54. Essuman, K. *et al.* (2018) TIR domain proteins are an ancient family of NAD+-consuming enzymes. *Curr. Biol.* 28, 421–430
55. Essuman, K. *et al.* (2017) The SARM1 Toll/Interleukin-1 receptor domain possesses intrinsic NAD+ cleavage activity that promotes pathological axonal degeneration. *Neuron* 93, 1334–1343
56. Hopkins, E.L. *et al.* (2021) A novel NAD signaling mechanism in axon degeneration and its relationship to innate immunity. *Front. Mol. Biosci.* 8, 1–16
57. Horsefield, S. *et al.* (2019) NAD + cleavage activity by animal and plant TIR domains in cell death pathways. *Science* 365, 793–799
58. Killackey, S.A. *et al.* (2019) The mitochondrial Nod-like receptor NLRX1 modifies apoptosis through SARM1. *Mol. Cell. Biochem.* 453, 187–196
59. Wan, L. *et al.* (2019) TIR domains of plant immune receptors are NAD+-cleaving enzymes that promote cell death. *Science* 365, 799–803
60. Duxbury, Z. *et al.* (2020) Induced proximity of a TIR signaling domain on a plant-mammalian NLR chimera activates defense in plants. *Proc. Natl. Acad. Sci. U. S. A.* 117, 18832–18839
61. Morehouse, B.R. *et al.* (2020) STING cyclic dinucleotide sensing originated in bacteria. *Nature* 586, 429–433
62. Tal, N. *et al.* (2021) Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell* 184, 5728–5739.e16
63. Hogrel, G. *et al.* (2022) Cyclic nucleotide-induced superhelical structure activates a bacterial TIR immune effector. *Nature* 1 2022.05.04.490601

64. Eastman, S. *et al.* (2022) A phytobacterial TIR domain effector manipulates NAD+ to promote virulence. *New Phytol.* 233, 890–904
65. Ma, S. *et al.* (2020) Direct pathogen-induced assembly of an NLR immune receptor complex to form a holoenzyme. *Science* 370, eabe3069
66. Martin, R. *et al.* (2020) Structure of the activated ROQ1 resistosome directly recognizing the pathogen effector XopQ. *Science* 370, eabd9993
67. Waite-Rees, P.A. *et al.* (1991) Characterization and expression of the *Escherichia coli* Mrr restriction system. *J. Bacteriol.* 173, 5207–5219
68. Lowey, B. *et al.* (2020) CBASS immunity uses CARF-related effectors to sense 3′–5′- and 2′–5′-linked cyclic oligonucleotide signals and protect bacteria from phage infection. *Cell* 182, 38–49
69. Meeske, A.J. *et al.* (2019) Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* 570, 241–245
70. Mayo-Muñoz, D. *et al.* Type III CRISPR-Cas provides resistance against nucleus-forming jumbo phages via abortive infection. *BioRxiv* Published online June 20, 2022. https://doi.org/10.1101/2022.06.20.496707
71. Garneau, J.E. *et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71
72. Zetsche, B. *et al.* (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771
73. Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–822
74. Parker, J.S. *et al.* (2004) Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J.* 23, 4727–4737
75. Ma, J.-B. *et al.* (2005) Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* 434, 666–670
76. Carthew, R.W. and Sontheimer, E.J. (2009) Origins and mechanisms of miRNAs and siRNAs. *Cell* 136, 642–655
77. Liu, X. *et al.* (2008) MicroRNAs: biogenesis and molecular functions. *Brain Pathol.* 18, 113–121
78. Ha, M. and Narry Kim, V. (2014) Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15, 509–524
79. Gainetdinov, I. *et al.* (2018) A single mechanism of biogenesis, initiated and directed by PIWI proteins, explains piRNA production in most animals. *Mol. Cell* 71, 775–790
80. Makarova, K.S. *et al.* (2019) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18, 67–83
81. Huang, F. *et al.* (2022) Catalytic properties and biological function of a PIWI - RE nuclease from *Pseudomonas stutzeri*. *Bioresour. Bioprocess* 9, 57
82. Jaskólska, M. *et al.* (2022) Two defence systems eliminate plasmids from seventh pandemic *Vibrio cholerae*. *Nature* 604, 323–329
83. Chen, J.S. *et al.* (2018) CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science* 360, 436–439
84. Gootenberg, J.S. *et al.* (2017) Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science* 356, 438–442
85. Li, L. *et al.* (2019) HOLMESv2: A CRISPR-Cas12b-assisted platform for nucleic acid detection and DNA methylation quantitation. *ACS Synth. Biol.* 8, 2228–2237
86. Steens, J.A. *et al.* (2021) SCOPE enables type III CRISPR-Cas diagnostics using flexible targeting and stringent CARF ribonuclease activation. *Nat. Commun.* 12, 1–12
87. Qin, Y. *et al.* (2022) Emerging Argonaute-based nucleic acid biosensors. *Trends Biotechnol.* 40, 910–914
88. Frank, F. *et al.* (2010) Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2/. *Nature* 465, 818–822
89. Nowotny, M. *et al.* (2005) Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* 121, 1005–1016
90. Bartel, D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297
91. Nakanishi, K. *et al.* (2012) Structure of yeast Argonaute with guide RNA. *Nature* 486, 368–374
92. Parker, J.S. (2010) How to slice: snapshots of Argonaute in action. *Silence* 1, 1–10
93. Chen, G.R. *et al.* (2017) A seed mismatch enhances Argonaute2-catalyzed cleavage and partially rescues severely impaired cleavage found in fish. *Mol. Cell* 68, 1095–1107
94. Sheu-Gruttadauria, J. *et al.* (2019) Structural basis for target-directed microRNA degradation. *Mol. Cell* 75, 1243–1255.e7
95. Kwak, P.B. and Tomari, Y. (2012) The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nat. Struct. Mol. Biol.* 19, 145–151
96. Kawamata, T. *et al.* (2009) Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat. Struct. Mol. Biol.* 16, 953–960
97. Tian, Y. *et al.* (2011) Structural basis for piRNA 2′-O-methylated 3′-end recognition by Piwi PAZ (Piwi/Argonaute/Zwille) domains. *Proc. Natl. Acad. Sci. U. S. A.* 108, 903–910
98. Wang, Y. *et al.* (2009) Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* 461, 754–761
99. Meister, G. *et al.* (2022) Single-molecule FRET uncovers hidden conformations and dynamics of human Argonaute 2. *Nat. Commun.* 13, 3825
100. Liu, Y. *et al.* (2018) Accommodation of helical imperfections in *Rhodobacter sphaeroides* Argonaute ternary complexes with guide RNA and target DNA. *Cell Rep.* 24, 453–462
101. Miyoshi, T. *et al.* (2016) Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.* 7, 1–12
102. Holm, L. (2022) Dali server: structural unification of protein families. *Nucleic Acids Res.* 50, W210–W215