RNA-guided RNA silencing by an Asgard archaeal Argonaute

Bastiaanssen, Carolien; Kim, Kijun; Feng, Yanlei; Anzelon, Todd A.; Tamarit, Daniel; Jinek, Martin; MacRae, Ian J.; Joo, Chirlmin; Swarts, Daan C.; Wu, Fabai

DOI
10.1038/s41467-024-49452-1

Publication date
2024

Document Version
Final published version

Published in
Nature Communications

nature communications

Article

https://doi.org/10.1038/s41467-024-49452-1

RNA-guided RNA silencing by an Asgard archaeal Argonaute

Received: 26 March 2024

Accepted: 4 June 2024

Published online: 29 June 2024

Check for updates

Carolien Bastiaanssen[1,10], Pilar Bobadilla Ugarte[2,10], Kijun Kim[1,10], Giada Finocchio[3,10], Yanlei Feng[4,5,10], Todd A. Anzelon[6], Stephan Köstlbacher[7], Daniel Tamarit[7,8], Thijs J. G. Ettema[7], Martin Jinek[3], Ian J. MacRae[6], Chirlmin Joo[1,9] ✉, Daan C. Swarts[2] ✉ & Fabai Wu[4] ✉

Argonaute proteins are the central effectors of RNA-guided RNA silencing pathways in eukaryotes, playing crucial roles in gene repression and defense against viruses and transposons. Eukaryotic Argonautes are subdivided into two clades: AGOs generally facilitate miRNA- or siRNA-mediated silencing, while PIWIs generally facilitate piRNA-mediated silencing. It is currently unclear when and how Argonaute-based RNA silencing mechanisms arose and diverged during the emergence and early evolution of eukaryotes. Here, we show that in Asgard archaea, the closest prokaryotic relatives of eukaryotes, an evolutionary expansion of Argonaute proteins took place. In particular, a deep-branching PIWI protein (HrAgo1) encoded by the genome of the Lokiarchaeon '*Candidatus* Harpocratesius repetitus' shares a common origin with eukaryotic PIWI proteins. Contrasting known prokaryotic Argonautes that use single-stranded DNA as guides and/or targets, HrAgo1 mediates RNA-guided RNA cleavage, and facilitates gene silencing when expressed in human cells and supplied with miRNA precursors. A cryo-EM structure of HrAgo1, combined with quantitative single-molecule experiments, reveals that the protein displays structural features and target-binding modes that are a mix of those of eukaryotic AGO and PIWI proteins. Thus, this deep-branching archaeal PIWI may have retained an ancestral molecular architecture that preceded the functional and mechanistic divergence of eukaryotic AGOs and PIWIs.

Argonaute proteins facilitate guide oligonucleotide-mediated binding of nucleic acid targets to perform a wide range of functions in prokaryotes and eukaryotes. In eukaryotic RNA silencing pathways, sequence-specific repression of target RNAs is achieved by Argonaute proteins loaded with small guide RNAs[1–5]. Canonical eukaryotic Argonautes (eAgos) can be subdivided into two clades, AGO and PIWI. They are distributed broadly, albeit heterogeneously, across eukaryotic lineages, and thus were thought to have both existed in the last eukaryotic common ancestor (LECA)[6]. AGOs and PIWIs are strictly conserved and arguably best studied in Metazoa (animals), where they rely on various guide generation pathways and carry out distinct physiological functions. Metazoan AGOs use small interfering RNA (siRNA) and/or microRNA (miRNA) guides to post-transcriptionally regulate gene expression[7,8]. In general, base-pairing of a short region at the 5′ end of miRNAs termed the 'seed' (nucleotides 2–8) to a target RNA is sufficient for AGOs to bind target RNA[8]. By contrast, metazoan PIWIs use PIWI-interacting RNA (piRNA) guides to suppress transposable elements (TEs)[9], and generally show lower seed binding strength and target RNA binding requires extended base-pairing in the central region of the guide to achieve stable binding[5]. Contrasting such well-

A full list of affiliations appears at the end of the paper. ✉e-mail: c.joo@tudelft.nl; daan.swarts@wur.nl; fabaiwu@eitech.edu.cn

characterized structural dynamics of Metazoan PIWI and AGO, few homologs across the expansive eukaryotic diversity have been explored to a similar extent. It thus remains unclear whether AGO- and PIWI-based RNA silencing pathways have consistent signatures across the expansive eukaryotic tree of life, and when and how diversification of molecular mechanisms originated.

Prokaryotic homologs of eukaryotic Argonautes (pAgos) are a highly diverse protein family with functions ranging from prokaryotic immunity by neutralizing foreign DNA[10–13] or inducing cell death in invaded cells (abortive infection)[14,15], to aiding in genome replication and recombination[16,17]. They comprise catalytically active Long-A and inactive Long-B pAgos that contain the canonical N-L1-PAZ-L2-MID-PIWI domains, as well as SiAgo-like and short pAgos in which only the MID and PIWI domains are conserved[14,15,18]. While pAgos share their domain architecture and general molecular mechanism (guide-mediated target binding) with eAgos, pAgos characterized to date primarily interact with DNA guides and/or targets; no known pAgo exclusively facilitates eAgo-like guide RNA-mediated RNA targeting. Thermophilic euryarchaeal Argonautes, which have previously been suggested to be most closely related to eukaryotic Argonautes[19], exclusively mediate DNA-guided targeting of invading DNA[13,20]. Furthermore, homologs of proteins involved in eukaryotic guide RNA biogenesis pathways (e.g., Dicer, Drosha, or Zucchini) have not been identified in prokaryotes. Hence, while it is clear that Argonaute proteins have a prokaryotic origin, it remains unclear when and how RNA-silencing functions emerged, and when and how Argonautes diverged into the extant AGO and PIWI subclades.

Here, we describe an Asgard archaeal pAgo that branches deeply in the PIWI-clade, and that mediates RNA-guided RNA silencing. This Asgard archaeal pAgo exhibits hybrid structural and mechanistic properties typically found in either AGO or PIWI, which may reflect an ancestral molecular architecture that laid the foundation for the functional and mechanistic divergence of AGOs and PIWIs in extant eukaryotes.

Results

Asgard archaeal diversification gave rise to eAgo-like Argonautes

Eukaryotes are thought to have evolved from an archaeal lineage belonging to Asgard archaea (or Asgardarchaeota)[21–25]. We thus set out to explore the presence of Argonaute proteins in these organisms using a custom hidden Markov model based on the conserved MID-PIWI domains (see "Methods"). In 496 available Asgard archaeal metagenome-assembled genomes (MAGs), we identified a total of 138 Asgard archaeal Argonaute sequences (asAgos, see Supplementary Dataset 1). Maximum-likelihood phylogenetic analysis shows that asAgos are distributed over 15 clades located across the phylogenetic tree of Argonaute proteins, including groups that branch deeply next to and outside of previously described pAgos clades[15] (Fig. 1a, Fig. S1). Like many other prokaryotic defense systems[26], pAgo genes are present only in a fraction of prokaryotes. We found Argonaute-encoding genes in 21.5% (83/387) of the quality-filtered Asgard archaeal MAGs (Fig. 1b), more prevalent than in any other prokaryotic phylum as classified by the Genome Taxonomic Database[27] (GTDB v207). Such apparent enrichment and the polyphyletic, deep-branching patterns of asAgos, suggests that Asgard archaea played a role in the early diversification of pAgos as well as their spread via horizontal gene transfer (HGT), both as donors and recipients.

'*Candidatus* Harpocratesius repetitus FW102'[24], a deep-sea rock-dwelling Lokiarchaeia archaeon named after the Greek god of silence, encodes two asAgos belonging to Asgard-specific clades distinct from known pAgos. These two proteins are only 23% identical in their MID-PIWI domain sequences. HrAgo1 clusters with and has the same conserved domains as eAgos, while the HrAgo2 subclade comprises a mixture of long and short asAgos and is phylogenetically sister to short pAgos (Fig. 1a). HrAgo2 lacks the N-L1-PAZ domains and is predicted to be catalytically inactive, akin to known short pAgos. Both asAgos are encoded in operon-like gene clusters outside of other genomic defense islands such as CRISPR-Cas and CBASS systems. The HrAgo2 operon encodes various components involved in transposition (InsG and TniQ) and DNA replication (PCNA and TOPRIM). This is different from known short pAgos, which cooperate with immune effectors encoded in their gene neighborhoods to trigger cell death[14,15]. The HrAgo1 operon is also unique in that flanking *hrAgo1* are an *rnc* gene, encoding a protein that comprises an RNaseIII domain fused to a double-stranded RNA binding domain (dsRBD), and a gene encoding a HEDxD/H helicase (Fig. 1c). These domains show homology to the functional domains of eukaryotic Dicer enzymes involved in guide RNA biogenesis[28], and to our knowledge, have not been found previously to be in association with other pAgos.

HrAgo1 shows higher similarity to well-studied PIWIs (25–27% sequence identity) and AGOs (23–24%), than to various other pAgos (16–21%) (Fig. S2). To date, the only eAgo-like HrAgo1 homolog that we could identify is a truncated asAgo sequence found in a Lokiarchaeon assembled from a Siberian soda lake metagenome[29], with 34% sequence identity to HrAgo1 across the obtained L2-MID-PIWI segment. Depending on the parameters used for phylogenetic analyses, HrAgo1 was either sister to the whole eAgo clade, or to the PIWI clade (Fig. S3). To further elucidate the relation between HrAgo1 and eAgos, we expanded the sampling of AGO and PIWI clade homologs across the eukaryotic tree of life and performed Maximum Likelihood analyses using above-found Long-A pAgos/asAgos and the non-canonical Trypanosome-specific TrypAgos as outgroup. When analyzed using the conserved MID-PIWI domains commonly used for Argonaute phylogeny[19], we found that the HrAgo1 subclade is positioned as sister group to the PIWI clade (Fig. 1a, inset). Additionally, we examined the more variable N-L1-PAZ domains as well as the full-length N-L1-PAZ-L2-MID-PIWI domains, which further confirmed that the deep sister position of HrAgo1 was not caused by a rare event of gene recombination between AGO and PIWI (Fig. S4). Our data thus suggest that the Asgard archaeal HrAgo1 and eukaryotic PIWIs evolved from a common ancestor, prompting us to study the molecular mechanism and function of HrAgo1.

HrAgo1 mediates RNA-guided RNA cleavage

The most apparent differences between eAgos and pAgos are their guide and target preferences. We thus analyzed the oligonucleotides that associate with HrAgo1 upon heterologous expression in *E. coli*. 5′-end ^{32}P-labeling of the associated nucleic acids reveals that HrAgo1 associated with 15-25 nt-long small RNAs, but not with DNA (Fig. 2a). Corroborating the ^{32}P-labeling-based detection, small RNA sequencing analysis confirmed that HrAgo1-associated small RNAs are mostly 15-25 nt in length (Fig. 2b). The small RNAs have a bias for uracil (U) at their 5′ end (65%), similar to the guide 5′-end preference observed for most examined PIWIs and AGOs[9,30]. Furthermore, a bias for U is observed to a lesser extent at position 2 (47%) and 3 (49%) of the guide RNA (Fig. 2c). Since previous studies have shown that nucleic acids co-purified with heterologously expressed pAgos generally match the types of their naturally preferred guides[10–12,14,31], our data thus suggest that HrAgo1 utilizes guide RNAs, akin to eAgos.

Next, we analyzed HrAgo1 guide/target preferences in vitro. Upon incubation of HrAgo1 with 21-nt single-stranded (ss)DNA or ssRNA guide oligonucleotides and complementary 5′ Cy5-labeled ssDNA or ssRNA targets (Fig. 2d, Table S1), HrAgo1 demonstrated ssRNA-guided cleavage of RNA targets in a magnesium-dependent manner, while it was unable to cleave DNA targets (Fig. 2e). Of note, guide ssDNAs also facilitated cleavage of RNA targets, but with lower efficiency compared to guide ssRNAs (Fig. 2e), similar to the in vitro behavior of human AGO2 (hAgo2)[32]. Combined, these results show that, compared to other known pAgos, the prokaryotic HrAgo1 mechanistically acts more similarly to RNA-guided RNA-targeting eAgos.

https://doi.org/10.1038/s41467-024-49452-1

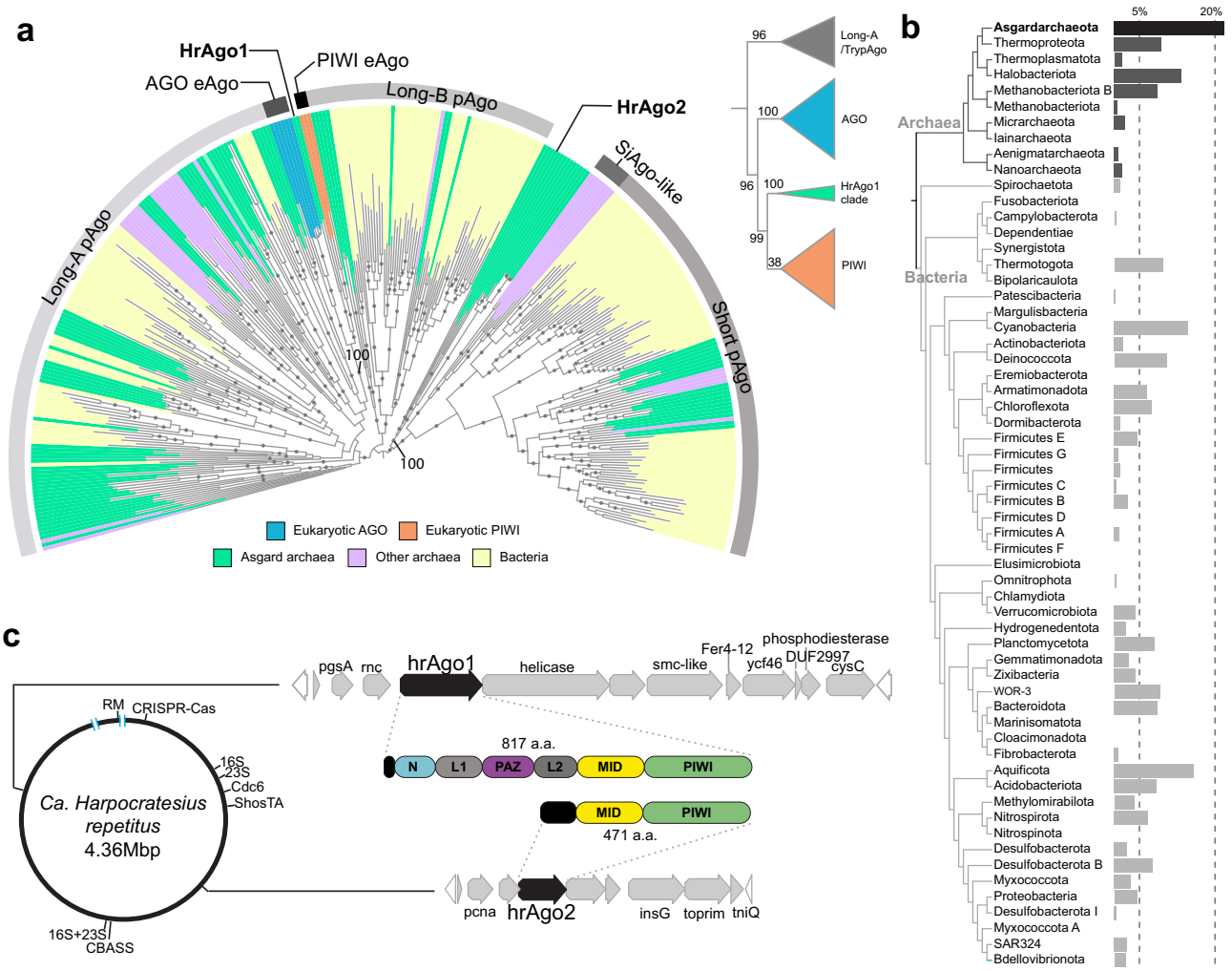

Fig. 1 | Diversity of Asgard archaeal Argonautes and a deep-branching PIWI-clade HrAgo1. a Maximum-likelihood phylogenetic analysis of the MID-PIWI domains of Argonaute proteins showing that asAgos are polyphyletic. HrAgo1 is phylogenetically related to Argonautes in the eukaryotic PIWI clade. 334 representative sequences and 572 sites were analyzed using IQ-TREE based on Q.pfam + C60 + F + G4 model. Color pallets indicate host classification, outer ring indicate pAgo and eAgo classifications. Ultrafast bootstrap 2 (UFBoot2) values above 95, calculated based on 1000 replicates, are shown in gray circles. HrAgo1, HrAgo2, and the UFBoot2 values at the branches are highlighted. Inset, Maximum-likelihood phylogenetic analysis based on the MID-PIWI domains with a broader sampling of AGO and PIWI indicates that HrAgo1 is sister to all eukaryotic PIWIs. UFBoot2 values calculated based on 1000 replicates are indicated. **b** Fraction of Argonaute-encoding genomes in different prokaryotic phyla. The tree topology was adapted from GTDB v207 by removing phyla with less than 40 representatives. Source data are provided in the Source data file. **c** Genomic depiction of Asgard archaeon '*Ca.* H. repetitus', where the genes encoding 16S and 23S rRNA, origin of replication protein Cdc6, and putative immune systems are indicated. The synteny and predicted domain compositions of genes surrounding pAgo-encoding genes are highlighted. RM, restriction-modification system. Blue bars indicate two genome assembly gaps with undetermined sequences.

Structural architecture of HrAgo1

To illuminate the structural basis for RNA-guided RNA cleavage by HrAgo1, we examined HrAgo1 in complex with a 21-nucleotide guide RNA by cryogenic electron microscopy (cryo-EM) and single particle analysis. The resulting reconstruction, determined at a resolution of 3.4 Å, reveals a binary HrAgo1-guide RNA complex (Fig. 3a–d, Fig. S5, Table S2). Resembling eAgos and long pAgos, HrAgo1 adopts a bilobed conformation in which one lobe comprises the N-terminal, linker L1, PAZ, and linker L2 domains, connected to the second lobe comprised of the MID and PIWI domains (Fig. 3c, d). The first six nucleotides of the guide RNA 5′ end (g1–g6) are ordered in the cryo-EM map (Fig. 3b–d). Low-resolution density for four nucleotides at the 3′ end of the guide RNA (g18–g21) is also apparent but uninterpretable, while the remainder of the guide RNA is unstructured (Fig. 3b–d). In accordance with its phylogeny, an all-against-all comparison[33] of experimentally determined structures of Argonaute-family proteins positions HrAgo1 between pAgos and eAgos, and closest to the PIWI-clade Siwi (Fig. 3e).

The catalytic tetrad of HrAgo1 comprises residues Asp585, Glu623, Asp655, and His792 (Fig. 3f). In the structure, all four catalytic residues are ordered and in position to mediate divalent cation binding and catalysis, akin to the catalytic site of AGO structures[2]. This implies that HrAgo1 adopts a catalytically active conformation. The 5′-terminal phosphate group of the guide RNA is sequestered in the MID domain binding pocket through interactions with residues (Phe512, Lys516, Asn528, and Lys555) that are conserved in most Argonautes[34] (Fig. 3g). The negative charge of two phosphates of guide RNA nucleotides 1 and 3, as well as that of the C-terminal carboxyl group of HrAgo1, are neutralized by a Mg^{2+} ion as is observed in pAgos and PIWIs (Fig. 3g). Instead of Mg^{2+}, Metazoan AGOs use another lysine residue in this pocket[35]. A catalytic double mutant (D585A & E623A, $HrAgo1^{DM}$) did not mediate RNA cleavage, confirming that the catalytic DEDH motif in the PIWI domain facilitates target cleavage (Fig. 3h). Corroborating the observed interactions with the 5′-phosphate, HrAgo1 showed higher activity with guide RNAs that are 5′-phosphorylated

https://doi.org/10.1038/s41467-024-49452-1

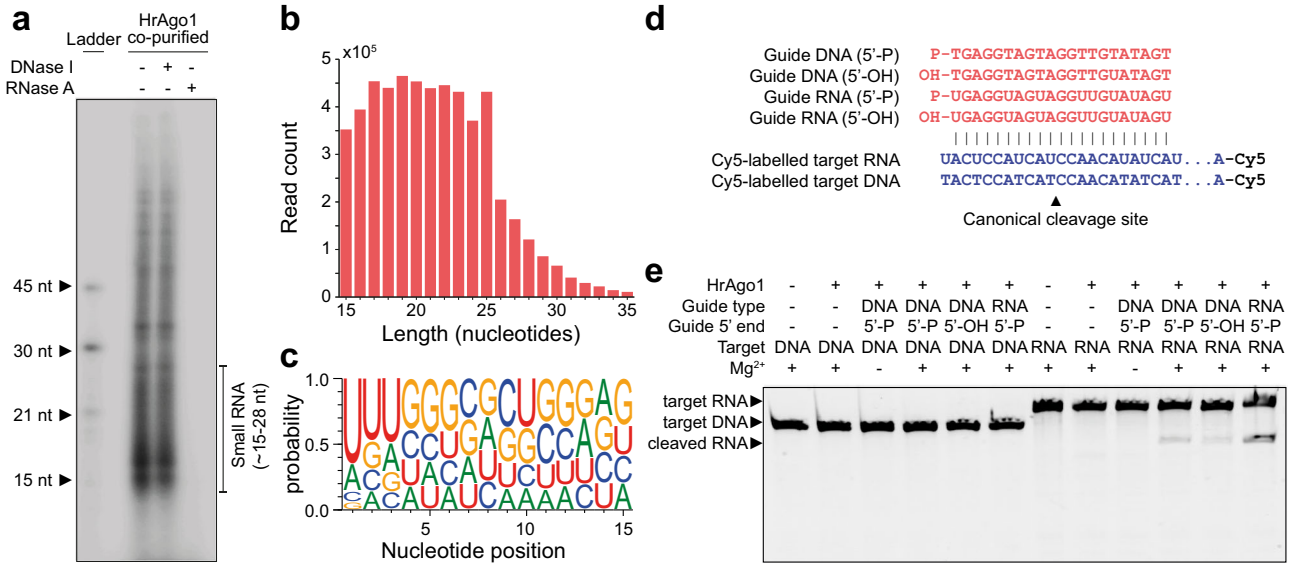

Fig. 2 | HrAgo1 mediates RNA-guided RNA cleavage. a HrAgo1 associates with 5′ phosphorylated (5′ P) small RNAs in vivo from *E. coli*. Nucleic acids that co-purified with HrAgo1 were [γ-^{32}P] labeled, treated with RNase A or DNase I, and resolved on a denaturing gel (15% polyacrylamide 7 M urea). nt: nucleotides. **b** Length distribution of small RNAs associated with HrAgo1 as determined by small RNA sequencing. **c** Small RNAs associated with HrAgo1 have a bias for uracil bases at the 5′ end. **d** Sequences of guide and target oligonucleotides used in in vitro cleavage assays. **e** HrAgo1 cleaves ssRNA (but not ssDNA) targets with ssRNA guides, and ssDNA guides at lower efficiency, in the presence of Mg^{2+}. HrAgo1 was incubated with ssDNA or ssRNA guides and Cy5-labeled ssDNA or RNA targets. Cy5-labeled cleavage products were resolved through denaturing (7 M urea) polyacrylamide gel electrophoresis and visualized by fluorescence imaging. Both ssRNA and ssDNA targets are 45nt. The HrAgo1-bound RNA extraction and digestion was carried out once; the results of the cleavage assays were confirmed by at least three repetitions.

compared to guide RNAs with a 5′-hydroxyl group (Fig. 3h). Remarkably, HrAgo1-mediated RNA-guided RNA cleavage at temperatures ranging from 9 °C to 71 °C (Fig. 3i), coinciding with a steep temperature gradient around the hot hydrothermal vents that '*Ca*. H. repetitus' resides in. Such an extraordinarily broad temperature adaptation apparently places HrAgo1 between the temperature ranges of mesophilic eAgos with those from the euryarchaeal pAgos, which mostly function at temperatures above 75 °C[13,36,37].

Our structural data combined with biochemical experiments thus illuminate the mechanistic adaptation of the archaeal HrAgo1 as an eAgo-like RNA-guided RNA-cleaving enzyme.

HrAgo1 structure bridges AGO and PIWI

The ability to rapidly find complementary mRNA targets in a crowded cellular environment and to effectively distinguish them from other transcripts is critical for the functions of eAgos. Argonaute proteins achieve efficient target binding by ordering the 'seed' region of the guide strand (nucleotides 2–8) in an A-form helical conformation[2,38,39]. AGO-clade eAgos achieve full seed pre-organization using a loop that binds the g5–g6 backbone and a helix-7 which organizes g7–g8[40], facilitating strong target association at short matching lengths[41]. By contrast, known PIWI-clade eAgos only pre-organize the first few nucleotides of the seed (typically g2–g4), and require further guide-target pairing beyond the seed to achieve stable target binding[3,5]. In HrAgo1, guide RNA nucleotides g2–g6 are pre-ordered in an A-form-like helical conformation (Fig. 4a). Remarkably, HrAgo1 simultaneously possesses structural features related to guide RNA ordering that are typically observed in either AGOs or in PIWIs (Fig. 4b): HrAgo1 is AGO-like in that it uses contacts in the g5–g6-binding loop (H738 and the main chain amide of R746) to organize guide nucleotides g5–g6, and is also PIWI-like in that helix-7 is tilted away from the seed, leaving g7–g8 disordered. Tilting of helix-7 in HrAgo1 may be attributed to the presence of a loose "seed-gate" structure (residues Phe333-Gln354), also present typically found in PIWIs, but more stretched in AGOs[3,5]. In PIWIs, the seed-gate structure has been proposed to enable extended target probing at positions g5–g9[5]. Overall, HrAgo1 structure pre-organizes guide RNA in a manner more similar to AGOs, while showing features similar to PIWIs that may influence its target binding.

HrAgo1 displays a unique hybrid mode of target binding

To investigate the target RNA binding kinetics of HrAgo1, we performed a single-molecule fluorescence resonance energy transfer (FRET) binding assay (Fig. 4c). Guide and target RNAs were labeled with Cy5 and Cy3 dyes respectively so that binding of the HrAgo1-guide complex to the target gives rise to a high FRET signal (Fig. 4d–f, Fig. S6a, Table S3). We quantitatively investigated the binding of the HrAgo1-guide RNA complex to target RNAs with varying guide-target complementarity and compared that to the same experiments performed with EfPiwi and to hAgo2 data from literature[41] (Fig. 4g, Fig. S7). The interactions between HrAgo1 and the target became observable when the latter matches the nt 2–4 (N3) positions of the guide RNA, and the dwell time increases drastically with the increase in guide-target match length (Figs. 4f, g, Fig. S6). At these short match lengths, the dwell time distribution follows a simple exponential decay, similar to previous observations of hAgo2[41]. Starting from N6, the majority of the guide-target association events of HrAgo1 and hAgo2 persist beyond the experimental time limit of 200 s (Fig. 4g). The overall binding kinetics of HrAgo1 are thus similar to the behavior of hAgo2. This contrasts EfPiwi, which only shows observable interactions with the target at a match length of N6, and shows stable binding only at N15, in agreement with structural predictions[5].

While HrAgo1 facilitates prolonged binding for most of the guide-target pairs between N6 and N8, a notable sub-population remains only transiently bound, resembling the behavior of EfPiwi (Fig. 4g). The appearance of a second population has been occasionally observed previously when the binding pocket of Argonaute interacts with a specific species of nucleotide in the first position of the target, e.g., deoxyguanosine by TtAgo[42] and deoxyadenosine by hAgo2[43]. However, the two-population behavior we observe here is independent of the identity of the first target nucleotide (Fig. S6b, c), suggesting that HrAgo1 intrinsically utilizes two modes of target search, i.e., an overall strong seed binding mode as observed for AGOs, and a second mode

https://doi.org/10.1038/s41467-024-49452-1

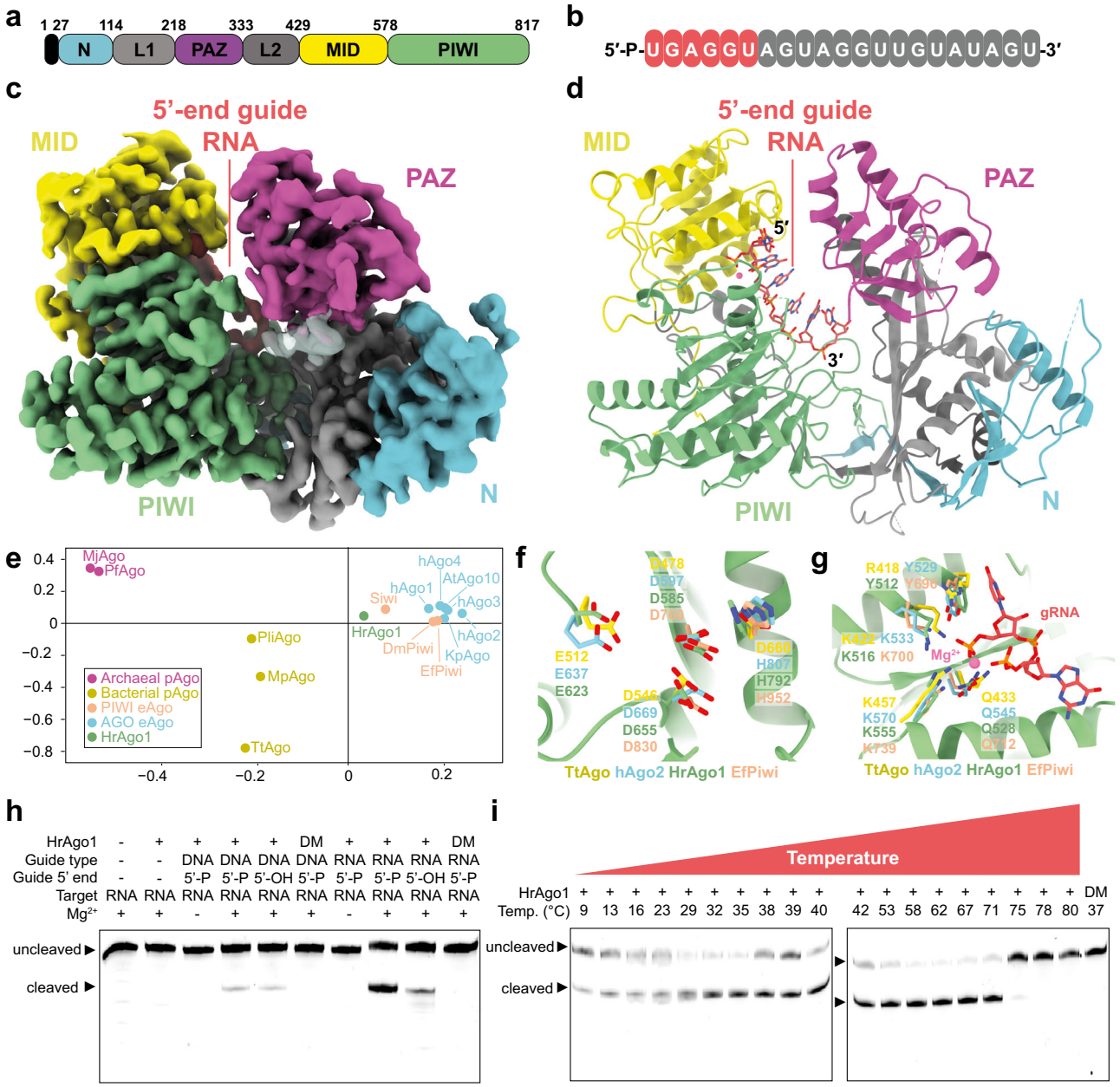

Fig. 3 | Molecular architecture of HrAgo1 bound to a guide RNA. a Schematic diagram of the domain organization of HrAgo1. N, N-terminal domain; L1 and L2, linker domains; PAZ, PIWI-ARGONAUTE-ZWILLE domain; MID, Middle domain; PIWI, P-element induces wimpy testis domain. **b** Schematic representation of the HrAgo1-bound guide RNA. Structurally ordered residues are colored red, while disordered residues are colored gray. **c** Cryo-electron microscopic (cryo-EM) density map of HrAgo1 bound to a guide RNA Cartoon, colored according to individual domains. The unmodeled 3′-end guide RNA density is represented as a transparent surface. **d** Cartoon representation of the overall structure of the HrAgo1-guide RNA complex. **e** All-against-all structure comparison of selected Argonaute proteins. Source data are provided in the Source data file. **f** Close-up view of the HrAgo1 catalytic site aligned to that of other representative Argonaute proteins. **g** Close-up view of the HrAgo1 guide RNA 5′-end binding site in the MID domain aligned to that of other representative Argonaute proteins. **h** Efficient HrAgo1-mediated RNA cleavage requires a guide RNA with a 5′ phosphate and an intact catalytic site. HrAgo1 was incubated with ssDNA or ssRNA guides and Cy5-labeled ssRNA targets. DM: HrAgo1 catalytic mutant with D585A and E623A substitutions. **i** HrAgo1 mediates RNA-guided RNA cleavage at temperatures ranging from 9 °C to 71 °C. HrAgo1 was incubated with ssRNA guides and Cy5-labeled ssRNA targets. For (**h**) and (**i**), Cy5-labeled cleavage products were resolved on a denaturing (7 M urea) polyacrylamide gel and visualized by fluorescence imaging. In (**h**) and (**i**), both uncleaved ssRNA targets are 45nt. The results of the cleavage assays were confirmed by at least three repetitions.

of transient seed binding akin to PIWIs. Consistent with its hybrid structural features, HrAgo1 thus facilitates a unique hybrid mode of guide RNA-mediated target RNA binding.

HrAgo1 mediates RNA silencing in human cells

The physiological function of HrAgo1 can provide clues to the emergence and diversification of RNA silencing pathways. However, Asgard archaea are notoriously slow-growing, largely uncultivated, and not genetically accessible. Furthermore, '*Ca*. H. repetitus' was enriched from undetectable to only 1% of the community on a low-biomass hydrothermal rock[24], and is therefore not a suitable host for physiological characterization of HrAgo1. Given the structural and mechanistic resemblance of HrAgo1 to eAgos, particularly its main binding characteristics resembling that of the human Ago2, we

https://doi.org/10.1038/s41467-024-49452-1

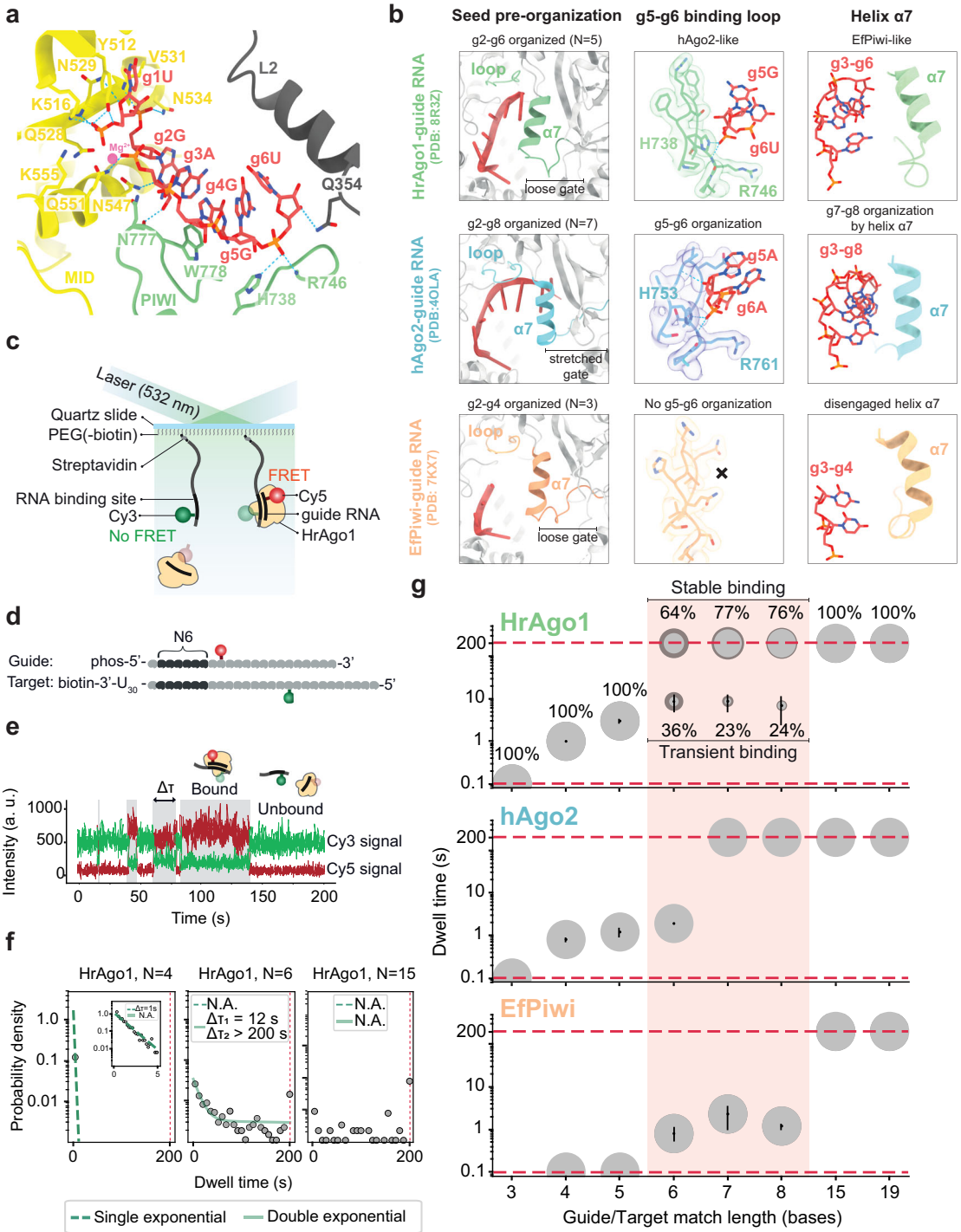

Fig. 4 | HrAgo1 displays a unique hybrid mode of guide organization and target binding. a Close-up view of guide RNA organization by HrAgo1. **b** Comparison of structural features involved in guide RNA seed segment organization in HrAgo1, EfPiwi, and hAgo2. **c** Schematic of the single-molecule binding assay. Only when the HrAgo1-guide complex binds to the target, FRET will occur. **d** Schematic representation of a guide and target used in the single-molecule binding assay. Complementary nucleotides are indicated in dark gray and mismatched nucleotides are shown in light gray. N6 indicates base pairing with nucleotides 2–7 of the guide. **e** A representative time trace with four binding events, of which the dwell time (Δτ) of one is indicated. **f** Dwell time distributions for N4, N6, and N15. N4 and N6 are best fit with a single and double exponential, respectively. N15 cannot be fit and shows stable binding. The distributions and fits for the other match lengths and representative time traces can be found in Fig. S6. **g** Bubble plots showing the increase of dwell times for increasing complementarity between the guide and target for HrAgo1, EfPiwi and hAgo2. The area of the bubbles corresponds to the percentage of the total population belonging to this sub-population. The dashed lines indicate the time resolution (0.1 s) and the observation time limit (200 s). In the bubbles, the mean ± SD are indicated, and the darker shaded area indicates the SD of the fractional values of at least three independent experiments. The dwell times for EfPiwi were obtained in a similar way as for HrAgo1. For hAgo2, previously published dwell times were used[41]. Source data are provided in the Source data file.

https://doi.org/10.1038/s41467-024-49452-1

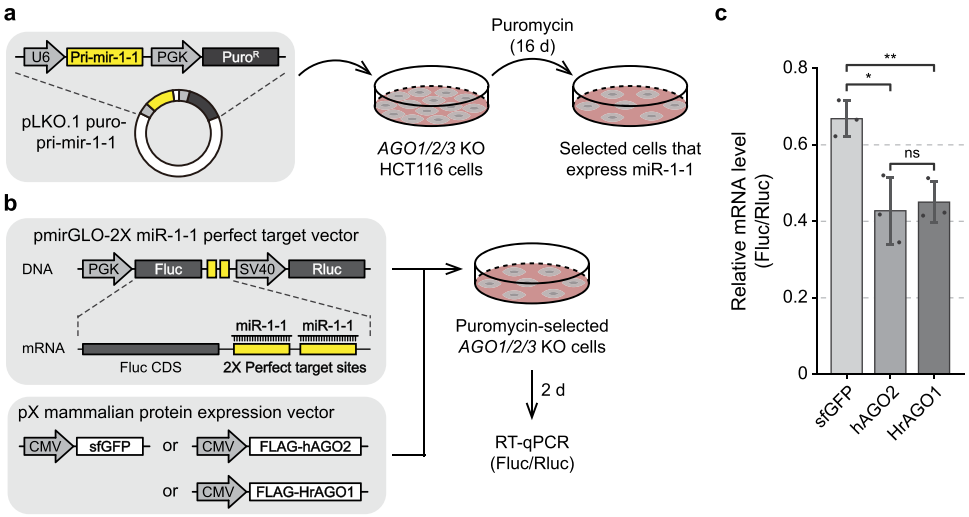

Fig. 5 | HrAgo1 mediates RNA silencing in human cells. a A schematic diagram for stable transfection of miR-1-1. pLKO.1 puro-pri-mir-1-1 vector was transfected into *AGO1/2/3 KO* HCT116 cells, which were subsequently subjected to puromycin selection for 16 days to generate cells that stably express miR-1-1. **b** A schematic diagram for the RNAi rescue experiment. Puromycin-selected cells were co-transfected with a dual-luciferase expression vector containing two perfect target sites for miR-1-1 in the 3′ UTR of the firefly luciferase gene (Fluc) and a protein expression vector encoding sfGFP or hAgo2 or HrAgo1. Two days after the transfection, total RNA was isolated and subjected to RT-qPCR. **c** qPCR results for relative mRNA expression levels between firefly luciferase (Fluc) and *Renilla* luciferase (Rluc). Bars indicate mean ± SD ($n = 3$, biological replicates). ns, not significant; $^{**}p < 0.01$; $^{*}p < 0.05$ by independent two-sided t-test. The p values are: 0.725 (hAGO2 vs. HrAGO1), 0.014 (sfGFP vs. hAGO2), and 0.0064 (sfGFP vs. HrAGO1). Source data are provided in the Source data file.

examined whether HrAgo1 can perform RNA silencing in a human cell line.

To exclude any endogenous RNA interference (RNAi) activity, we adopted an HCT116 cell line in which *hAgo1/2/3* genes are knocked out (*AGO1/2/3* KO HCT116)[44]. We first performed stable transfection of the pLKO.1 puro-pri-mir-1-1 vector, which encodes puromycin *N*-acetyltransferase that confers resistance to puromycin and a primary hairpin transcript (pri-mir-1-1) that acts as a precursor for mature miR-1-1 whose expression is suppressed in the parental cells[45] (Fig. 5a, Fig. S8a). Puromycin-selected cells were then co-transfected with a dual-expression vector that encodes firefly luciferase (Fluc) and *Renilla* luciferase (Rluc), as well as with an expression vector encoding FLAG-tagged HrAgo1 (FLAG-HrAgo1) (Fig. 5b, Fig. S8b). In addition, vectors expressing superfolder GFP (sfGFP) and FLAG-tagged hAgo2 (FLAG-hAgo2) were used as negative and positive controls, respectively. The 3′ UTR of the Fluc gene has two binding sites with perfect complementarity to miR-1-1, which allows Ago-mediated silencing of Fluc expression. To monitor miR-1-1-guided Fluc silencing, we performed qPCR to measure the relative expression level between target Fluc mRNA and the control Rluc mRNA. Remarkably, cells in which miR-1-1 and HrAgo1 were co-expressed showed a significant ($p < 0.01$) decrease in the Fluc/Rluc mRNA ratio compared to cells in which the sfGFP control was co-expressed with miR-1-1 (Fig. 5c). Moreover, the level of post-transcriptional repression by HrAgo1 was comparable to that of hAgo2 without significant difference. This demonstrates that HrAgo1 is capable of RNA silencing in human cells.

Discussion

In this study, we identified and characterized HrAgo1 as a prokaryotic Argonaute capable of RNA-guided RNA silencing both in vitro and in vivo. This is characteristically different from all pAgos studied thus far, and is consistent with its phylogenetic position sister to all PIWI-clade eAgos. Remarkably, this PIWI-clade HrAgo1 exhibited AGO-clade hAgo2-like gene silencing function in human cells when co-expressed with a dsRNA hairpin that is a precursor of a canonical miRNA. Although it is currently unclear whether HrAgo1 can also fulfill PIWI-like behavior, the observed AGO-like behavior of the PIWI-clade HrAgo1 suggests that the Metazoa-deduced AGO-PIWI dichotomy may have been a result of a more recent specialization; deeper-branching eukaryotic PIWIs and AGOs may be more flexible in their use of guide-generating mechanisms and functions. Indeed, miRNAs and their processing by Dicer-like ribonucleases were found in social amoeba *Dictystelium discoideum*, which encodes only PIWI but not AGO[46]. Since prokaryotic Argonautes often form partnerships with their gene neighbors[14,15] and AGO-Dicer-like gene neighbors have occasionally been observed in eukaryotes[47], HrAgo1's flanking genes encoding ribonuclease III and helicase (Fig. 1b) should also be experimentally investigated in the future for their potential ability to generate guides for HrAgo1. Future studies unveiling functional characteristics and associated guide-generating mechanisms of other deep-branching AGOs and PIWIs, be them archaeal or eukaryotic, would shed new light on the evolutionary conservation and diversifications of RNA silencing.

HrAgo1 as a deep-branching PIWI also provided insights into the diversification of AGO and PIWI at the molecular level. Metazoan AGOs require strong seed binding to enable targeting of a wide range of mRNAs: in humans, miRNAs sculpt the transcriptome by modestly repressing nearly half of the cellular mRNAs[8]. By contrast, PIWIs display weak seed binding and require further guide/target base pairing to achieve potent and specific silencing of TEs[48]. In agreement with its structural features, HrAgo1 exhibits both a primary strong seed-binding mode akin to AGO as well as a secondary mode of weak seed-binding akin to PIWI; this is in agreement with its structural features. Based on our comparative analyses of structural and smFRET data between HrAgo1 and different eAgos, we hypothesize that the common ancestor of AGO- and PIWI-clade eAgos had a g5–g6 pre-organizing loop, while its helix-7 did not embrace g7–g8. Metazoan AGOs kept and further refined the g5–g6 organizing loop, while repositioning helix-7 to enable g7–g8 pre-organization, allowing strong target association at short matching lengths to facilitate post-transcriptional silencing of a multitude of genes[2,7,8,15,41]. The structure of HrAgo1 suggests that ancestral PIWIs had the g5–g6 loop, which was lost in metazoan evolution, giving rise to the more relaxed targeting preferences of metazoan piRNAs that enable defense against evolving genomic threats[5,49]. So far, structural studies of AGO and PIWI have focused on Metazoa, fission yeast, and land plants[1,2,34,39,50], and PIWIs only from Metazoa[3–5]. Future structural characterizations of other deep-branching AGOs and PIWIs will further refine

the above hypothesis on the evolutionary processes leading to their diversification.

The presence of a PIWI-clade Argonaute in the Asgard archaea opened an avenue for exploring the evolutionary origin of RNA silencing pathway by studying extant organisms. A recent phylogenomic study suggested that the *Hodarchaeales* is likely the closest lineage to eukaryotes[25], while HrAgo1 clade that likely share a common ancestor with eukaryotic PIWI was so far only found in the *Lokiarchaeia*, which is sister to *Hodarchaeales*. This is consistent with previous observations that most eukaryotic signature proteins (ESPs) are patchily distributed across Asgard archaeal lineages[21–25], and the proposal that they became imported into the protoeukaryotes, an elusive state prior to LECA, via HGT[24]. Since AGOs and PIWIs most likely have already diverged in the LECA[6], HrAgo1 could have been horizontally transferred between an ancestral Lokiarchaeon and a protoeukaryote. The timing and directionality of the HGT is yet difficult to ascertain, as we have not yet found a pAgo/asAgo clade sister to the eukaryotic AGO clade, or RNA-silencing pAgos/asAgos at the base of both AGO and PIWI clade. On the other hand, taking into account that *Hodarchaeales* is so far much less sampled than *Lokiarchaeia* (25 genomes v.s. 167 genomes in our dataset) and pAgo genes undergo frequent loss, which can cause the same distribution patterns as HGT[51], we cannot exclude a possibility that RNA-silencing Argonautes were inherited vertically from the prokaryotic common ancestor of *Hodarchaeales* and eukaryotes. Further improving the taxonomic sampling and quality of Asgard archaeal genomes may provide new evidence to resolve this evolutionary question.

Finally, the arms race against TEs and global gene silencing are critical drivers of eukaryotic genome evolution[24,52–54], recent analyses found Argonaute to be one of the most abundant defense genes in the Asgard archaea[55,56], and our analysis found Argonaute to be more enriched in the Asgard archaea than in any other prokaryotic phyla. Hence, further resolving how Argonaute functions diverged in the Asgard archaea and how RNA silencing pathways diverged in early (proto)eukaryotes might also shed light on how eukaryotes emerged and expanded.

Methods

Identification and selection of Argonautes encoded by Asgard archaeal genomes

A custom-built Hidden Markov Model (HMM) encompassing MID-PIWI domain representatives from all known prokaryotic and eukaryotic Argonaute types is provided as Supplementary Datasets 2–4. This HMM was used to search across 496 Asgard archaea MAGs from NCBI, yielding 138 putative asAgo sequences. Since some sequences are truncated due to fragmented genome assembly, we identified their gene position and the presence of start codon and stop codon to determine the completeness of the genes. The information is provided in Supplementary Dataset 1. Incomplete sequences were excluded from phylogenetic analyses except for ASG308_00888, which is the only close homolog of HrAgo1 found in this study but truncated at its N terminus due to contig break.

Phylogenetic analysis of Asgard archaeal Argonaute

To examine the phylogenetic relation between asAgos and known Argonaute proteins, previously identified pAgos[18] were first clustered at 60% identity using CD-HIT[57] v4.8.1. This set was aligned using MAFFT[58] v7.475 option auto, and sequences with clear N-terminal or C-terminal truncations were removed. The alignment was trimmed using trimAl[59] v1.4.1 option gappyout, and phylogenetically analyzed using IQ-TREE[60] v2.1.12 model LG + R9 with 2000 ultrafast bootstrap replicates. The tree was reduced using Treemmer[61] v0.3 to represent the diversity with fewer related sequences, and well-studied pAgo representatives (highlighted in Fig. S1) were manually added back if were removed by Treemmer. Next, the well-studied, structurally characterized canonical PIWI and AGO clade proteins were selected to comprise 9 eAgo representatives. The reference Argonaute proteins highlighted in Fig. S1 are PIWI from *Ephydatia flauviatilis* Piwi (EfPIWI), AGO from *Homo sapiens* (hAGO2), archaeal Argonautes from *Pyrococcus furiosus* (PfAgo), *Methanocaldococcus jannaschii* (MjAgo), and *Natronobacterium gregoryi* (NgAgo), *Archaeoglobus fulgidus* (AfAgo), *Sulfolobus islandicus* (SiAgo), and bacterial Argonautes from *Aquifex aeolicus* (AaAgo), *Thermus thermophilus* (TtAgo), *Clostridium butyricum* (CbAgo), *Marinotoga piezophila* (MpAgo), *Rhodobacter sphaeroides* (RsAgo), *Pseudooceanicola lipolyticus* (PliAgo), *Runella slithyformis* (RslAgo), *Crenotalea thermophila* (CrtAgo), *Kordia jejudonensis* (KjAgo), *Xanthomonas vesicatoria* (XavAgo), and *Joostella marina* (JomAgo). 109 asAgos, quality-filtered as described above, were used. The final set comprises a total of 334 Argonautes. These proteins were aligned using MAFFT option linsi, and the MID-PIWI section was retained using the amino acid positions in the HrAgo1 structure as reference. The cropped alignment was then trimmed using trimAl option gt 0.1 to remove the most highly variable regions and used for phylogenetic analysis. Maximum likelihood phylogenetic analysis was carried out using IQ-TREE v2.1.12. The best fitting model was identified using ModelFinder[62] among all combinations of the LG, WAG, and Q.pfam models combined with the empirical profile mixture model C60[63], and with modeled rate heterogeneity (either +R4 and +G4). The Q.pfam + C60 + F + R4 was selected by the ModelFinder. Statistical support was evaluated using 1,000 replicates via ultrafast boostrap 2 (UFBoot2)[64]. The phylogenetic tree was visualized using iTOL[65], where ultrafast bootstrap values above 95 were indicated in Fig. 1a.

To examine the stability of the Long-A pAgo branches sister to the eAgo clade, we used two different alignment combinations and three different models. Besides the MID-PIWI domains of all Ago types described above, we omitted the short pAgo clade and made a full-length alignment encompassing the N-L1-PAZ-L2-MID-PIWI domains. In addition to the Q.pfam + C60 + F + R4, we also used LG + C60 + F + R4 and WAG + C60 + F + R4. Statistical support was evaluated using 1000 replicates via UFBoot2. Branches closest to the eAgo clade were shown in Fig. S3.

Diverse eukaryotic AGO and PIWI full-length sequences were used to create HMM profiles via HMMER v3.3.2 (http://hmmer.org/). To ensure the full recruitment of evolutionary intermediates between AGO and PIWI, the medium bitscore of AGO members was used as cutoff for PIWI HMM searches, and vice versa. These profiles and bitscore cutoffs were used to recruit eukaryotic Argonaute proteins from the EukProt v3 database[66]. After quality filtering by removing truncated sequences lacking the major domains of Argonaute, 1312 putative AGOs and 454 putative PIWIs were aligned using MAFFT option auto and phylogenetically analyzed using FastTree[67] v2.1.10 model LG. The AGO clade and PIWI clade of the trees were pruned using Treemmer down to 100 branches each, where each eukaryotic supergroup was forced to keep at least 3 sequences if possible. 201 eukaryotic Argonaute representatives were combined with HrAgo1 and ASG308_00888 (the truncated homolog of HrAgo1), TrypAgos, and LongA pAgo sequences, aligned using MAFFT option linsi, trimmed using trimAl option gt 0.1, and analyzed using IQ-TREE v2.1.12. The best fitting model was identified using ModelFinder among all combinations of the LG, WAG, and Q.pfam models combined with the empirical profile mixture model C60, and with modeled rate heterogeneity (either +R4 and +G4). Statistical support was evaluated using 1000 replicates via UFBoot2. The phylogenetic tree was visualized using iTOL.

Identification of various features in the '*Ca*. H. repetitus' genome

The present '*Ca*. H. repetitus FW102' genome assembly is a single scaffold with two gaps (GenBank accession: JAIZWK010000001.1). The basic features including the origin of replication protein Cdc6 and 16S and 23S rRNA subunits was annotated as described previously[24], using eggNOG mapper[68] v.2 and RNAmmer[69] v.1.2 (embedded in RASTtk[70]),

respectively. The CRISPR-Cas operon was annotated using CCTyper[71] v1.8.0. Other defense systems were identified using the Defense-Finder online tool[26], which also identified HrAgo2 and the CRISPR-Cas system, but did not identify HrAgo1.

Presence of Argonaute homologs across prokaryotic lineages

The custom MID-PIWI HMM profile was used to search for Argonaute homologs in the GTDB database v207 (for all prokaryotic phyla except Asgard archaea) and an Asgard archaea database (387 genomes after quality filtering using the same standard as GTDB). Prokaryotic phyla with less than 40 representatives were removed for comparison.

Sequence similarity between HrAgo1 with various eAgos and Long pAgos

Representative sequences were each aligned with HrAgo1, the number of aligned sites with the same identity was divided by the total number of amino acids in HrAgo1 as metrics for sequence similarity.

Plasmid construction

The *HrAgo1* gene, codon-optimized for *E. coli* and synthesized by Genscript, Inc., was inserted under the T7 promoter in the expression plasmid pET28a to yield pFWC01 (*Pt7::HrAgo1*). A plasmid suitable for expression of a HrAgo1 catalytic double-mutant (D585A & E623A; HrAgo1^{DM}) was generated by Quikchange Site-Directed Mutagenesis using primers oPB199 and oPB201 for D585A and oPB200 and oPB198 for E623A, using *E. coli* strain NEB 5-alpha (New England Biolabs) (Table S4).

pX-sfGFP vector was a kind gift from Prof. Jae-Sung Woo (Korea University, South Korea). Linear pX vector backbone was prepared by PCR with primers bypassing sfGFP coding region and then subjected to gel purification. Insert DNA fragments with human codon-optimized coding sequences for FLAG-hAgo2 and FLAG-HrAgo1, flanked by pX vector homology regions, were synthesized commercially (Twist Bioscience). Insert DNA fragments were cloned into the linear pX vector backbone by Gibson assembly (in lab). Competent *E. coli* cells were transformed with the Gibson assembly products, and plasmids (pX-FLAG-hAgo2 and pX-FLAG-HrAgo1) were purified using PureYield Plasmid Miniprep System (Promega).

pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega) was linearized by PCR with primers that insert two fully complementary binding sites ("perfect target sites") for human miR-1-1 3p in the 3′ UTR of the firefly luciferase gene. Competent *E. coli* cells were transformed with the linearized vectors, and plasmids (pmirGLO-2X miR-1-1 perfect target site) were purified by miniprep.

pLKO.1 puro was a gift from Bob Weinberg (Addgene plasmid # 8453; http://n2t.net/addgene:8453; RRID:Addgene_8453)[72]. pLKO.1 puro vector was linearized by PCR with primers that insert human pri-mir-1-1 sequence in the downstream of the U6 promoter. Competent *E. coli* cells were transformed with linearized vectors, and plasmids (pLKO.1 puro-pri-mir-1-1) were purified by miniprep. All plasmids were verified by Sanger sequencing (Macrogen). The cloning primers are listed in Table S5.

HrAgo1 expression and purification

HrAgo1 was heterologously expressed in *Escherichia coli* BL21-Gold (DE3). Expression cultures were shaken at 120 rpm in an incubator at 37 °C in LB supplemented with 50 mg/ml kanamycin until an optical density at 600 nm ($OD_{600\ nm}$) of 0.4 was reached. The incubation temperature was then decreased to 18 °C. When the $OD_{600\ nm}$ reached 0.6, expression of HrAgo1 was induced by adding isopropyl-b-D-thiogalactoside (IPTG) to a final concentration of 0.2 mM. Expression of HrAgo1 took place at 18 °C for 20 h. Cells were harvested by centrifugation at 4000 × *g* at 4 °C for 30 min and were lysed by sonication (QSONICA Q700A-220 sonicator with ½" tip, amp 35%, 1 s ON/2 s OFF for 4 min) in Lysis Buffer (1 M NaCl, 5 mM Imidazole, 20 mM Tris-HCl pH 8) supplemented with protease inhibitors (100 µg/ml AEBSF and 1 µg/ml Pepstatin A). After centrifugation at 40,000 × *g* at 4 °C for 45 min, the cell free extract was loaded on 5 ml HisTrap HP column (Cytiva Life Sciences) which was subsequently washed with 25 ml of Washing Buffer I (1 M NaCl, 20 mM Imidazole, 20 mM Tris-HCl pH 8). Bound protein was eluted with Elution Buffer I (1 M NaCl, 250 mM Imidazole, 20 mM Tris-HCl pH 8). The eluted protein was loaded on a custom 20 ml amylose resin column and was washed with Washing Buffer II (1 M NaCl, 20 mM Tris-HCl pH 8, 1 mM DTT). The protein was eluted with Elution Buffer II (1 M NaCl, 20 mM Tris-HCl pH 8, 10 mM Maltose, 1 mM DTT). TEV protease was added in a 1:50 (w/w) ratio (TEV:total protein), and the mixture was dialyzed overnight in Snake-Skin dialysis tubing (30 kDa MWCO, Thermo Scientific) against 2 l dialysis buffer (1 M KCl, 20 mM HEPES-KOH pH 7.5, 1 mM DTT, 2 mM EDTA) at 4 °C for 16 h. TEV-mediated removal of the His-MBP tag was confirmed by SDS-PAGE analysis. The sample was concentrated to a volume of 1 ml using 30 K centrifugal filter units (Amicon). After concentrating, the sample was centrifuged for 10 min at 16,000 × *g* at 4 °C to remove aggregates and the supernatant was loaded on a custom 200 ml Superdex 200 resin column which was pre-equilibrated with SEC buffer (1 M KCl, 20 mM HEPES-KOH pH 7.5, 1 mM DTT). The peak fractions were analyzed by SDS-PAGE and fractions containing HrAgo1 were combined and concentrated, aliquoted and flash frozen in liquid nitrogen before storage at −70 °C until further use.

HrAgo1^{DM} was expressed and purified as HrAgo1 with minor modifications: For expression *E. coli* BL21 Star (DE3) was used. Furthermore, expression was performed in TB medium containing 20 µg/ml kanamycin.

For HrAgo1-bound small RNA analysis, 6xHis-MBP-tagged HrAgo1 was purified. *E. coli* culture was inoculated into Terrific medium supplemented with 50 µg/ml Kanamycin and were shaken at 180 rpm at 37 °C in an incubator until an $OD_{600\ nm}$ of 0.4. IPTG was added to reach a final concentration of 0.1 mM, and the culture was moved to 24 °C and kept shaking for 8 h. Cells were harvested by centrifugation at 4250 × *g* at 4 °C for 20 min and were lysed by sonication (Daxluot Multichannel sonicator with 6 mm tip, 5 s ON/8 s OFF for 1 h under constant cooling circulation) in Lysis Buffer (500 mM NaCl, 5 mM Imidazole, 50 mM Tris-HCl pH 8) supplemented with 1 mM PMSF and a protease inhibitor cocktail (Roche). TCEP was added immediately after sonication to a concentration of 1 mM. After centrifugation at 6000 × *g* for 30 min at 4 °C, the cell-free extract was loaded on 5 ml HisTrap HP column (Cytiva Life Sciences) which was subsequently washed with Washing Buffer I (500 mM NaCl, 20 mM Imidazole, 1 mM TCEP, 50 mM Tris-HCl pH 8). Bound protein was eluted with 15 ml Elution Buffer (500 M NaCl, 500 mM Imidazole, 5% glycerol, 50 mM Tris-HCl pH 8). The sample was concentrated to a volume of 2 ml using 30 kDa Amicon ultracentrifugal filter. Imidazole-free buffer (500 M NaCl, 5% glycerol, 50 mM Tris-HCl pH 8) was added, well-mixed, and re-concentrated three times to replace the original buffer. The final round of centrifugation was done to result in ~1 ml solution, moved to a fresh 1.5 ml tube and was centrifuged for 30 min at 16,000 × *g* at 4 °C to remove aggregates. The supernatant was loaded on a Superdex 200 Increase 10/300 GL column (Cytiva) for size exclusion. The peak fractions were analyzed by SDS-PAGE and fractions containing HrAgo1 were combined and concentrated, aliquoted and flash frozen in liquid nitrogen before storage at −80 °C until further use.

Cleavage activity assays

HrAgo1 activity assays were performed in reactions with a final volume of 20 µl with the following final concentrations: 0.4 µM HrAgo1, 0.4 µM guide oligonucleotide (ogDS001, ogDS002, ogDS003, or oBK458 (Table S1)), 0.1 µM Cy5-labeled target oligonucleotide (oDS401 or oDS403; Table S1), 5 mM HEPES-KOH, 125 mM KCl, and 2 mM divalent metal salt ($MnCl_2$ or $MgCl_2$)). Prior to addition of the target, HrAgo1 and the guide were incubated for 15 min at 37 °C. After addition of the

target, HrAgo1:guide:target ratios were 4:4:1. The mixture was incubated for 1 h at 37 °C. The reaction was stopped by adding 2X RNA Loading Dye (250 mM EDTA, 5% v/v glycerol, 95% v/v formamide) and further incubation at 95 °C for 10 min. For the activity assays at different temperatures, pre-incubation (15 min) and incubation after addition of the target (60 min) took place at the indicated temperatures. The samples were resolved on a 20% denaturing (7 M Urea) polyacrylamide gel. The gels were imaged on an Ettan DIGE Imager (GE Healthcare (480/530 nm)).

Small RNA extraction and analysis

Two nanomoles of purified HrAgo1 were incubated with 250 μg/ml Proteinase K (Thermo Scientific) for 4 h at 65 °C. Next, phenol:chloroform:IAA 25:24:1 pH 7.9 (Invitrogen) was added in a 1:1 ratio. The sample was vortexed and centrifuged at 16,000 × *g* in a table top centrifuge for 10 min. The upper layer containing the nucleic acids was transferred to a clean tube and the nucleic acids were precipitated through ethanol precipitation. To this end, 99% cold ethanol and 3 M sodium acetate pH 5.2 were added to the sample in a 2:1 and 1:9 ratio, respectively. The sample was incubated overnight at −80 °C, after which it was centrifuged at 16,000 × *g* in a table top centrifuge for 1 h. The pellet was washed with 70% ethanol and subsequently dissolved in nuclease-free water.

Purified nucleic acids were [γ-^{32}P]-ATP labeled with T4 polynucleotide kinase (PNK; Thermo Scientific) in an exchange-labeling reaction. After stopping the reaction by incubation at 75 °C for 10 min, the labeled oligonucleotides were separated from free [γ-^{32}P] ATP using a custom Sephadex G-25 column (GE Healthcare). Labeled nucleic acids were incubated with nucleases (Rnase A, Dnase and protease-free (Thermo Scientific), or Dnase I, Rnase-free (Thermo Scientific) for 30 min at 37 °C. After nuclease treatment, samples were mixed with Loading Buffer (95% (deionized) formamide, 5 mM EDTA, 0.025% SDS, 0.025% bromophenol blue and 0.025% xylene cyanol), heated for 5 min at 95 °C and resolved on 15% denaturing (7 M Urea) polyacrylamide gels. Radioactivity was captured from gels using phosphor screens and imaged using a Typhoon FLA 7000 laser-scanner, GE Healthcare).

Small RNA sequencing libraries were prepared and sequenced by GenomeScan (Leiden, The Netherlands) using Illumina NovaSeq6000 sequencing with paired-end reads and 150 bp read length. Paired-end small RNA reads were merged, adapter sequences were trimmed, and length was trimmed to 35 nucleotides using Bbtools[73] v38.90. Processed reads of all sequencing libraries were aligned to the genome of *E. coli* BL21 (GenBank: CP053602.1) and to the expression plasmid (pFWC01) using HISAT2[74] v2.1.0. Length, sequence distribution, and abundance of specific small RNAs were analyzed using FastQC v0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) after extracting uniquely mapped reads using HISAT2 and Samtools v1.2[75].

Cryo-EM sample preparation and data collection

Purified HrAgo1 was mixed with a 5′-phosphorylated RNA guide (5′-UGAGGUAGUAGGUUGUAUAGU-3′) in assembly buffer (5 mM HEPES pH 7.5, 250 mM KCl, 5 mM $MgCl_2$). The final sample contained 8.6 μM HrAgo1 and 8.6 μM of g-RNA in a total volume of 60 μL. The volume was incubated at 37 °C for 15 min and centrifuged at 21,130 × *g* for 10 min at room temperature. After adding CHAPSO (Sigma-Aldrich) to a final concentration of 0.8 mM, the sample was used for cryo-EM grid preparation.

2.5 μL of the above sample was applied to a freshly glow discharged 300-mesh UltrAuF R1.2/1.3 grid (Quantifoil Micro Tools), blotted for 5 s at 100% humidity, 4 °C, plunge frozen in liquid ethane (using a Vitrobot Mark IV plunger, FEI) and stored in liquid nitrogen. Cryo-EM data collection was performed on a FEI Titan Krios G3i microscope (University of Zurich, Switzerland) operated at 300 kV and equipped with a Gatan K3 direct electron detector in super-resolution counting mode. A total of 8977 movies were recorded at ×130,000 magnification, resulting in a super-resolution pixel size of 0.325 Å. Each movie comprised 47 subframes with a total dose of 56.81 $e^-/Å^2$. Data acquisition was performed with EPU Automated Data Acquisition Software for Single Particle Analysis (ThermoFisher Scientific) with three shots per hole at −1.0 mm to −2.4 mm defocus (0.2 mm steps).

CryoEM data processing and model building

The collected exposures were processed in cryoSPARC[76] v4.2. Patch Motion Correction and Patch CTF Correction were used to align and correct the imported 8977 movies. Movies with CTF resolution higher than 20 Å were discarded, resulting in a total of accepted 8275 movies. Template picker (particle diameter 140 Å; templates were selected from a previous data collection on the same sample) was used to select particles, which were included for further processing based on their NCC and power score. Particles were extracted (extraction box size 360 pix; Fourier-cropped to box size 120 pix) and classified in 50 classes using 2D Classification. 22 classes (2188198 particles) were selected and given as input to a 2-classes Ab-Initio Reconstruction. The 1299949 particles corresponding to one of the two reconstructions were further sorted in 100 classes using 2D Classification. 28 classes (533275 particles) were used for a 2-classes Ab-Initio Reconstruction (maximum resolution 6 Å; initial resolution 20 Å; initial minibatch size 300; final minibatch size 2000). The particles of one of the two reconstructions were assigned to 80 classes using 2D classification, 57 of which (283659 particles) were extracted to full resolution and selected for non-uniform refinement (initial lowpass resolution 20 Å; per-particle CTF parameters and defocus optimization). A final round of non-uniform refinement (dynamic mask start resolution 1 Å; initial lowpass resolution 20 Å; per-particle CTF parameters and defocus optimization) resulted in a 3.40 Å (GSFSC resolution, FSC cutoff 0.143) density. A detailed processing workflow is shown in Fig. S5.

An initial model of HrAgo1 was generated using AlphaFold2 ColabFold[77]. The model was manually docked as rigid body in the cryoEM density map using UCSF ChimeraX[78] v1.7.1, followed by real space fitting with the Fit in Map function. The model was subjected to manual refinement against the corresponding cryoEM map using the software Coot[79] v0.9.8.92 and real space refine in Phenix[80] v1.21.1-5286. Secondary structure restraints, side chain rotamer restraints and Ramachandran restraints were used. The final model comprises one copy of HrAgo1(27–99,103–193,198–271,282–307,322–589,595–817), one copy of the gRNA (1–6) and 1 Mg^{2+} ion. Low-resolution density for the RNA 3′ end was visible in the map, but not confidently interpretable, therefore it was not built in the final model. Figures preparation of model and map was performed using UCSF ChimeraX.

Single-molecule experimental set-up

All single-molecule experiments were performed on a custom-built microscope setup. An inverted microscope (IX73, Olympus) with prism-based total internal reflection was used in combination with a 532 nm diode-pumped solid-state laser (Compass 215 M/50 mW, Coherent). Photons are collected with a 60x water immersion objective (UPLSAPO60XW, Olympus), after which a 532 nm long pass filter (LDP01-532RU-25, Semrock) blocks the excitation light. A dichroic mirror (635 dcxr, Chroma) separates the fluorescence signal which is then projected onto an EM-CCD camera (iXon Ultra, DU-897U-CS0-#BV, Andor Technology).

Single-molecule sample preparation

Synthetic RNA was purchased from Horizon Discovery (United Kingdom). The guide and target strands (Table S3) were labeled with Cy5 Mono NHS Ester and Cy3 Mono NHS Ester (Sigma-Aldrich), respectively. 5 μl of 200 μM RNA, 1 μl of 0.5 M freshly prepared 0.5 M sodium bicarbonate and 1 μl of 20 mM dye in DMSO were mixed and incubated

overnight at 4 °C in the dark, followed by ethanol precipitation. The labeling efficiency was ~100%. The target strands were subsequently ligated with a biotinylated polyuridine strand (U_{30}-biotin). To this end, 200 pmol of target RNA strand was mixed with U_{30}-biotin and a DNA splint in a 1:1:3 ratio in TE buffer with 100 mM NaCl. The mixture was annealed in a thermal cycler by rapidly heating it to 80 °C for 4 min and then slowly cooling it down with 1 °C every 4 min. The annealed constructs were ligated using 2 μL T4 RNA ligase2 (NEB, 10 U/μL), 3 μL 0.1% BSA (Ambion), 3 μL 10x reaction buffer (NEB), 0.25 μl 1 M $MgCl_2$ and 0.3 μl Rnasin ribonuclease inhibitor (Promega, 0.4 u/μl) in a final volume of 30 μl at 25 °C overnight. After acidic phenol-chloroform extraction and ethanol precipitation, the ligated RNA strands were purified on a 10% denaturing (7 M urea) polyacrylamide gel.

For the t1-target assays, the RNA target strands were produced through in vitro transcription of DNA templates (Table S3). All synthetic DNA was purchased from Ella Biotech (Germany). First, an annealing mix was prepared with template DNA and IVT T7 promotor oligos at a final concentration of 40 μM each in a 10 μl reaction with 1x annealing buffer (50 mM NaCl and 10 mM Tris-HCl pH 8.0). The annealing mix was heated to 90 °C for 3 min and then slowly cooled with 1 °C every min to 4 °C. Next, in vitro transcription was performed using the TranscriptAid T7 High Yield kit (Thermo Scientific) for 4 h at 37 °C according to the manufacturer's instructions. After acidic phenol-chloroform extraction and ethanol precipitation, the RNA strands were purified on a 10% denaturing (7 M urea) polyacrylamide gel. Finally, the purified RNA target strands (2 μM in 10 μl) were annealed to the immobilization strand and imager strand in a 2:1:5 ratio in annealing buffer by heating to 90 °C for 3 min and then slowly cooling with 1 °C every min to 4 °C.

Microfluidic chambers with a polymer(PEG)-coated quartz surface were prepared as described previously[81]. Each chamber was incubated with 20 μl 0.1 mg/ml Streptavidin (Sigma) for 30 s. Unbound Streptavidin was flushed out with 100 μl T50 (10 mM Tris-HCl pH 8.0, 50 mM NaCl). Next, 50 μl 50 pM Cy3-labeled target RNA was introduced into the chamber and incubated for 1 min. Unbound target RNA was flushed out with 100 μl T50 and 100 μl imaging buffer (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM Trolox (Sigma), 0.8% glucose, 0.5 mg/mL glucose oxidase (Sigma), 85 μg/mL catalase (Merck), 0.4 u/μl Rnasin ribonuclease inhibitor (Promega)) was introduced into the chamber. EfPiwi was purified as previously described[5]. For EfPiwi, binding was much weaker so to enable observation of these events within our time resolution 50 mM instead of 500 mM NaCl was used in the imaging buffer. The PIWI binary complex was formed by incubating 15 nM purified PIWI in imaging buffer (minus the glucose oxidase and catalase which were added after incubation) with 1 nM Cy5-labeled guide RNA at 37 °C for 10 min. The binary complex was introduced in the chamber, after which 200 s long movies were recorded. The experiments were performed at room temperature (22 ± 2 °C).

Single-molecule data acquisition and analysis

CCD movies of time resolution 0.1 s were acquired using Andor Solis software v4.32. Co-localization between the Cy3 and Cy5 signal and time trace extraction were carried out using Python. The extracted time traces were processed using FRETboard[82] v0.0.3. The dissociation rate was estimated by measuring the dwell times off all binding events. The dwell time distributions were fit with an exponential decay curve ($Ae^{-t/\Delta\tau}$) or with the sum of two exponential decay curves ($A1e^{-t/\Delta\tau1} + A2e^{-t/\Delta\tau2}$).

Mammalian cell culture and transfection

HCT116 *AGO1/2/3* knockout (KO) cells were obtained from the Corey lab (UT Southwestern, USA). Cells were grown and maintained in McCoy's 5A Modified Medium (Thermo Fisher Scientific) supplemented with 9% (v/v) fetal bovine serum (Cytiva) in an incubator at 37 °C and 5% CO_2.

For stable transfection, 2E6 KO cells were seeded in a 9 ml medium on a 100 mm culture dish 1 day before transfection. Transfection was performed with 5 μg of pLKO.1 puro-pri-mir-1-1 vector using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's instructions. Two days after the transfection, culture medium was replaced by a medium containing 2 μg/ml puromycin (Thermo Fisher Scientific) ("selection medium"). The selection proceeded for 16 days, and the selection medium was replaced every 4 days.

For transient transfection, 2E6 puromycin-selected cells were seeded in a 9 ml medium on a 100 mm culture dish 1 day before transfection. The cells were co-transfected with 4 μg of pX vector (pX-sfGFP or pX-FLAG-hAGO2 or pX-FLAG-HrAGO1) and 1 μg of pmirGLO-2X miR-1-1 perfect target site using Lipofectamine 3000. Cells were harvested 2 days after transfection, snap-frozen by liquid nitrogen, and then stored at −80 °C.

RT-qPCR

Total RNAs were isolated using TRIzol (Thermo Fisher Scientific), treated with RQ1 RNase-Free DNase (Promega), and then phenol-extracted. Complementary DNAs (cDNAs) were synthesized from 2 μg of total RNAs using SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific) and random hexamer according to the manufacturer's instructions. qPCR was performed using Power SYBR Green PCR Master Mix (Thermo Fisher Scientific) and QuantStudio 5 Real-Time PCR systems. The qPCR primers are listed in Table S5.

Western blotting

Cells were lysed by re-suspension in ice-cold lysis buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 1% Triton X-100, 0.5% Sodium Deoxycholate, 0.1% SDS) supplemented with Protease Inhibitor Cocktail Set III, EDTA-Free (Millipore). The lysed cells were centrifuged at 16,100 × *g*, 4 °C, 15 min, and the supernatant ("total protein lysate") was transferred to a fresh e-tube. The concentration of the total protein lysates was measured by Pierce BCA Assay (Thermo Fisher Scientific). 50 μg of total protein lysates were boiled with 4x Laemmli Sample Buffer (Bio-Rad), run on NuPAGE 4–12% Bis-Tris gel (Thermo Fisher Scientific) with PageRuler Plus Prestained Protein Ladder (Thermo Fisher Scientific), and then transferred onto methanol-activated Immun-Blot PVDF membrane (Bio-Rad) using Mini Blot Module (Thermo Fisher Scientific). The membrane was blocked in PBS-T (PBS (PanReac AppliChem) + 0.1% Tween-20 (Sigma)) containing 5% skim milk, probed with primary antibodies at 4 °C, ON, and then washed three times with PBS-T. Rabbit polyclonal FLAG antibody (1:1000, Sigma, F7425) was used to probe ectopically expressed FLAG-hAGO2 or FLAG-HrAGO1, and rat monoclonal Tubulin antibody (1:1000, Invitrogen, MA1-80017) was used to probe loading control. The washed membranes were probed with secondary antibodies at RT for 1 h, and then washed three times with PBS-T. Alexa Fluor 647-conjugated donkey anti-rabbit IgG (1:2000, Jackson ImmunoResearch, 711-605-152) and Alexa Fluor 546-conjugated goat anti-rat IgG (1:2000, Invitrogen, A-11081) were used as secondary antibodies. The protein bands were detected by fluorescence using the Typhoon laser-scanner platform system (Cytiva).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The HMM profiles, protein sequence alignment files, and phylogenetic trees generated in this study are provided in the Supplementary Datasets 2–18. The small RNA sequencing data is available at the Gene Expression Omnibus database under accession number GSE267550. Atomic coordinates and cryo-EM maps have been deposited in the Protein Data Bank (PDB) under accession code 8R3Z and Electron

Microscopy Data Bank (EMDB) under accession code EMD-18878. Source data are provided with this paper.

References

1. Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. Structure of yeast Argonaute with guide RNA. *Nature* **486**, 368–374 (2012).
2. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336**, 1037–1040 (2012).
3. Matsumoto, N. et al. Crystal structure of silkworm PIWI-clade Argonaute Siwi bound to piRNA. *Cell* **167**, 484–497.e489 (2016).
4. Yamaguchi, S. et al. Crystal structure of Drosophila Piwi. *Nat. Commun.* **11**, 858 (2020).
5. Anzelon, T. A. et al. Structural basis for piRNA targeting. *Nature* **597**, 285–289 (2021).
6. Cerutti, H. & Casas-Mollano, J. A. On the origin and functions of RNA-mediated silencing: from protists to man. *Curr. Genet.* **50**, 81–99 (2006).
7. Kim, V. N., Han, J. & Siomi, M. C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **10**, 126–139 (2009).
8. Bartel, D. P. Metazoan microRNAs. *Cell* **173**, 20–51 (2018).
9. Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. & Zamore, P. D. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **20**, 89–108 (2019).
10. Swarts, D. C. et al. DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261 (2014).
11. Olovnikov, I., Chan, K., Sachidanandam, R., Newman, DianneK. & Aravin, AlexeiA. Bacterial Argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell* **51**, 594–605 (2013).
12. Kuzmenko, A. et al. DNA targeting and interference by a bacterial Argonaute nuclease. *Nature* **587**, 632–637 (2020).
13. Swarts, D. C. et al. Argonaute of the archaeon Pyrococcus furiosus is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.* **43**, 5120–5129 (2015).
14. Koopal, B. et al. Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA. *Cell* **185**, 1471–1486.e1419 (2022).
15. Zeng, Z. et al. A short prokaryotic Argonaute activates membrane effector to confer antiviral defense. *Cell Host Microbe* **30**, 930–943.e936 (2022).
16. Jolly, S. M. et al. Thermus thermophilus argonaute functions in the completion of DNA replication. *Cell* **182**, 1545–1559.e1518 (2020).
17. Fu, L. et al. The prokaryotic Argonaute proteins enhance homology sequence-directed recombination in bacteria. *Nucleic Acids Res.* **47**, 3568–3579 (2019).
18. Ryazansky, S., Kulbachinskiy, A. & Aravin, A. A. The Expanded Universe of Prokaryotic Argonaute Proteins. *mBio* **9**, https://doi.org/10.1128/mbio.01935-18 (2018).
19. Swarts, D. C. et al. The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.* **21**, 743–753 (2014).
20. Zander, A. et al. Guide-independent DNA cleavage by archaeal Argonaute from Methanocaldococcus jannaschii. *Nat. Microbiol.* **2**, 17034 (2017).
21. Spang, A. et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
22. Zaremba-Niedzwiedzka, K. et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353–358 (2017).
23. Liu, Y. et al. Expanded diversity of Asgard archaea and their relationships with eukaryotes. *Nature* **593**, 553–557 (2021).
24. Wu, F. et al. Unique mobile elements and scalable gene flow at the prokaryote–eukaryote boundary revealed by circularized Asgard archaea genomes. *Nat. Microbiol.* **7**, 200–212 (2022).
25. Eme, L. et al. Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes. *Nature* **618**, 992–999 (2023).
26. Tesson, F. et al. Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* **13**, 2561 (2022).
27. Parks, D. H. et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* **50**, D785–D794 (2022).
28. Mukherjee, K., Campos, H. & Kolaczkowski, B. Evolution of animal and plant dicers: early parallel duplications and recurrent adaptation of antiviral RNA binding in plants. *Mol. Biol. Evol.* **30**, 627–641 (2013).
29. Vavourakis, C. D. et al. Metagenomes and metatranscriptomes shed new light on the microbial-mediated sulfur cycle in a Siberian soda lake. *BMC Biol.* **17**, 69 (2019).
30. Medley, J. C., Panzade, G. & Zinovyeva, A. Y. microRNA strand selection: unwinding the rules. *Wiley Interdiscip. Rev. RNA* **12**, e1627 (2021).
31. Hegge, J. W. et al. DNA-guided DNA cleavage at moderate temperatures by Clostridium butyricum Argonaute. *Nucleic Acids Res.* **47**, 5809–5821 (2019).
32. Willkomm, S., Zander, A., Grohmann, D. & Restle, T. Mechanistic insights into archaeal and human Argonaute substrate binding and cleavage properties. *PLoS ONE* **11**, e0164695 (2016).
33. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
34. Boland, A., Tritschler, F., Heimstädt, S., Izaurralde, E. & Weichenrieder, O. Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. *EMBO Rep.* **11**, 522–527 (2010).
35. Kuhn, C.-D. & Joshua-Tor, L. Eukaryotic Argonautes come into focus. *Trends Biochem. Sci.* **38**, 263–271 (2013).
36. Fang, M. et al. Characterization and application of a thermophilic Argonaute from archaeon Thermococcus thioreducens. *Biotechnol. Bioeng.* https://doi.org/10.1002/bit.28153 (2022).
37. Willkomm, S. et al. Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein. *Nat. Microbiol.* **2**, 17035 (2017).
38. 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).
39. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209–213 (2008).
40. Klum, S. M., Chandradoss, S. D., Schirle, N. T., Joo, C. & MacRae, I. J. Helix-7 in Argonaute2 shapes the microRNA seed region for rapid target recognition. *EMBO J.* **37**, 75–88 (2018).
41. Chandradoss, S. D., Schirle, N. T., Szczepaniak, M., MacRae, I. J. & Joo, C. A dynamic search process underlies microRNA targeting. *Cell* **162**, 96–107 (2015).
42. Swarts, D. C. et al. Autonomous generation and loading of DNA guides by bacterial Argonaute. *Mol. Cell* **65**, 985–998.e986 (2017).
43. Schirle, N. T., Sheu-Gruttadauria, J., Chandradoss, S. D., Joo, C. & MacRae, I. J. Water-mediated recognition of t1-adenosine anchors Argonaute2 to microRNA targets. *eLife* **4**, e07646 (2015).
44. Chu, Y. et al. Argonaute binding within 3′-untranslated regions poorly predicts gene repression. *Nucleic Acids Res.* **48**, 7439–7453 (2020).
45. Suzuki, H. et al. Genome-wide profiling of chromatin signatures reveals epigenetic regulation of microRNA Genes in Colorectal Cancer. *Cancer Res.* **71**, 5646–5658 (2011).
46. Liao, Z., Kjellin, J., Hoeppner, M. P., Grabherr, M. & Söderbom, F. Global characterization of the Dicer-like protein DrnB roles in miRNA biogenesis in the social amoeba Dictyostelium discoideum. *RNA Biol.* **15**, 937–954 (2018).
47. Casas-Mollano, J. A. et al. Diversification of the core RNA interference machinery in Chlamydomonas reinhardtii and the role of DCL1 in transposon silencing. *Genetics* **179**, 69–81 (2008).

48. Czech, B. & Hannon, G. J. One loop to rule them all: the ping-pong cycle and piRNA-guided silencing. *Trends Biochem. Sci.* **41**, 324–337 (2016).
49. Gainetdinov, I. et al. Relaxed targeting rules help PIWI proteins silence transposons. *Nature* **619**, 394–402 (2023).
50. Xiao, Y., Maeda, S., Otomo, T. & MacRae, I. J. Structural basis for RNA slicing by a plant Argonaute. *Nat. Struct. Mol. Biol.* **30**, 778–784 (2023).
51. Snel, B., Bork, P. & Huynen, M. A. Genomes in flux: the evolution of archaeal and proteobacterial gene content. *Genome Res.* **12**, 17–25 (2002).
52. Lane, N. & Martin, W. The energetics of genome complexity. *Nature* **467**, 929–934 (2010).
53. Koonin, E. V. Viruses and mobile elements as drivers of evolutionary transitions. *Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci.* **371**, 20150442 (2016).
54. Kapusta, A., Suh, A. & Feschotte, C. Dynamics of genome size evolution in birds and mammals. *Proc. Natl Acad. Sci. USA* **114**, E1460 (2017).
55. Shomar, H. et al. Viperin immunity evolved across the tree of life through serial innovations on a conserved scaffold. Preprint at *bioRxiv* https://doi.org/10.1101/2023.09.13.557418 (2023).
56. Leão, P. et al. Asgard archaea defense systems and their roles in the origin of eukaryotic immunity. Preprint at *bioRxiv* https://doi.org/10.1101/2023.09.13.557551 (2024).
57. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).
58. Nakamura, T., Yamada, K. D., Tomii, K. & Katoh, K. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* **34**, 2490–2492 (2018).
59. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
60. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
61. Menardo, F. et al. Treemmer: a tool to reduce large phylogenetic datasets with minimal loss of diversity. *BMC Bioinforma.* **19**, 164 (2018).
62. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
63. Si Quang, L., Gascuel, O. & Lartillot, N. Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics* **24**, 2317–2323 (2008).
64. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
65. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
66. Richter, D. J. et al. EukProt: A database of genome-scale predicted proteins across the diversity of eukaryotes. *Peer Community Journal* **2**, e56 (2022).
67. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
68. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* **47**, D309–D314 (2019).
69. Lagesen, K. et al. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108 (2007).
70. Brettin, T. et al. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **5**, 8365–8365 (2015).
71. Russel, J., Pinilla-Redondo, R., Mayo-Muñoz, D., Shah, S. A. & Sørensen, S. J. CRISPRCasTyper: automated identification, annotation, and classification of CRISPR-Cas Loci. *CRISPR J.* **3**, 462–469 (2020).
72. Stewart, S. A. et al. Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9**, 493–501 (2003).
73. Bushnell, B., Rood, J. & Singer, E. BBMerge—accurate paired shotgun read merging via overlap. *PLOS ONE* **12**, e0185056 (2017).
74. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
75. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
76. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
77. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
78. Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).
79. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D.* **66**, 486–501 (2010).
80. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D.* **66**, 213–221 (2010).
81. Chandradoss, S. D. et al. Surface passivation for single-molecule protein studies. *JoVE* https://doi.org/10.3791/50549 (2014).
82. de Lannoy, C. V., Filius, M., Kim, S. H., Joo, C. & de Ridder, D. FRETboard: semisupervised classification of FRET traces. *Biophys. J.* **120**, 3253–3260 (2021).

Acknowledgements

HCT116 knockout cells were a kind gift from Prof. David R. Corey (UT Southwestern, USA). F.W. thanks Claudio Nunes-Alves for critical comments on the paper, Woodward Fischer and Diaoqiong Zheng for valuable discussions, and Danxi Cui for technical support. D.C.S. and P.B.U. thank members of the Swarts lab for valuable discussions and technical support. C.B. and C.J. thank Martin Depken and Hidde Offerhaus for valuable discussions on the data analysis. F.W. was supported by National Natural Science Foundation of China grant (32370003). C.J. was supported by ERC Consolidator grant (819299) of the European Research Council and Frontier 10-10 (Ewha Womans University). D.C.S. was supported by grants from the European Research Council (ERC-2020-STG 948783) and the Dutch Research Council (016.Veni.192.072). P.B.U. was supported by Consejo Nacional de Ciencia y Technology (CVU No. 682509). K.K. was supported by an EMBO Postdoctoral Fellowship (ALTF 76-2022). T.J.G.E. was supported by ERC Consolidator grant (817834) of the European Research Council, Vici grant (VI.C.192.016) of the Dutch Research Council and Life grant (96725) of the Volkswagen Foundation.

Author contributions

F.W. conceived the project and supervised it with C.J. and D.C.S. Y.F. and F.W. performed protein identification and phylogenetic analyses. P.B.U. and F.W. constructed plasmids and purified HrAgo1 proteins. T.A.A. purified EfPiwi proteins. C.B. isolated and sequenced small RNAs. D.C.S. performed HrAgo1-associated small RNA analyses. P.B.U. performed in vitro cleavage assays. C.B. performed single-molecule assay and analyzed the data with C.J. G.F. performed CryoEM and analyzed data with D.C.S., I.J.M., and M.J. I.J.M. performed structural comparisons with

eukaryotic Argonautes. K.K. constructed plasmids and performed RNA silencing experiments in human cell lines. S.K., D.T., and T.J.G.E. provided additional asAgo sequences and phylogenetic pipelines. F.W., D.C.S., C.J., C.B., P.B.U., K.K., G.F., and I.J.M. wrote the manuscript. All authors commented on the manuscript.

Competing interests

F.W., C.J., K.K., D.C.S., and P.B.U. applied for a patent based on the use of HrAgo1 for RNA silencing. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-49452-1.

Correspondence and requests for materials should be addressed to Chirlmin Joo, Daan C. Swarts or Fabai Wu.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

[1]Department of BioNanoScience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands. [2]Laboratory of Biochemistry, Wageningen University, Wageningen, The Netherlands. [3]Department of Biochemistry, University of Zurich, Zurich, Switzerland. [4]School of Life Sciences, College of Science, Eastern Institute of Technology, Ningbo, China. [5]Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou, China. [6]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA. [7]Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands. [8]Theoretical Biology and Bioinformatics, Department of Biology, Faculty of Science, Utrecht University, Utrecht, The Netherlands. [9]Department of Physics, Ewha Womans University, Seoul, Korea. [10]These authors contributed equally: Carolien Bastiaanssen, Pilar Bobadilla Ugarte, Kijun Kim, Giada Finocchio, Yanlei Feng. ✉e-mail: c.joo@tudelft.nl; daan.swarts@wur.nl; fabaiwu@eitech.edu.cn