

# Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA

Daan C. Swarts<sup>1</sup>, Jorrit W. Hegge<sup>1</sup>, Ismael Hinojo<sup>1</sup>, Masami Shiimori<sup>2</sup>, Michael A. Ellis<sup>2</sup>, Justin Dumrongkulraksa<sup>2</sup>, Rebecca M. Terns<sup>2</sup>, Michael P. Terns<sup>2</sup> and John van der Oost<sup>1,\*</sup>

<sup>1</sup>Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, 6703 HB Wageningen, The Netherlands and <sup>2</sup>Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia 30602, USA

Received March 2, 2015; Revised April 13, 2015; Accepted April 16, 2015

## ABSTRACT

Functions of prokaryotic Argonautes (pAgo) have long remained elusive. Recently, Argonautes of the bacteria *Rhodobacter sphaeroides* and *Thermus thermophilus* were demonstrated to be involved in host defense. The Argonaute of the archaeon *Pyrococcus furiosus* (PfAgo) belongs to a different branch in the phylogenetic tree, which is most closely related to that of RNA interference-mediating eukaryotic Argonautes. Here we describe a functional and mechanistic characterization of PfAgo. Like the bacterial counterparts, archaeal PfAgo contributes to host defense by interfering with the uptake of plasmid DNA. PfAgo utilizes small 5'-phosphorylated DNA guides to cleave both single stranded and double stranded DNA targets, and does not utilize RNA as guide or target. Thus, with respect to function and specificity, the archaeal PfAgo resembles bacterial Argonautes much more than eukaryotic Argonautes. These findings demonstrate that the role of Argonautes is conserved through the bacterial and archaeal domains of life and suggests that eukaryotic Argonautes are derived from DNA-guided DNA-interfering host defense systems.

## INTRODUCTION

Eukaryotic Argonaute proteins (eAgos) are the key players in RNA interference (RNAi) pathways (reviewed in (1–3)). During RNAi, eAgos are loaded with small 5'-phosphorylated RNAs, ranging from 20 to 30 nt, in a pathway-specific ribonuclease-dependent process. The eAgos sometimes form the core of a multiprotein RNA-induced silencing complex (RISC; reviewed in (4)). Depending on which proteins associate with eAgos, Argonaute-mediated target binding can be specifically adjusted to differentially control gene expression. RNA interference

generally results in silenced expression of the target gene, via functional variations that include decreased transcription (heterochromatin formation), decreased translation (mRNA binding) and decreased mRNA half-life (mRNA cleavage, mRNA de-adenylation).

Whereas prokaryotes also possess Argonaute proteins (pAgos), they appear to lack the accessory proteins involved in eukaryotic RNA interference pathways (3,5,6). Initially, pAgo variants from *Aquifex aeolicus* and *Thermus thermophilus* were characterized biochemically, revealing that they could use DNA guides for RNA and DNA target cleavage (7–9). Recently, it has become evident that at least some bacterial Agos play a role in host defense by interfering with invading nucleic acids (10–12). *Rhodobacter sphaeroides* Argonaute (RsAgo) acquires small RNA guides that allow interference with plasmid DNA (10). *T. thermophilus* Argonaute (TtAgo) acquires 5'-phosphorylated single stranded DNA (ssDNA) guides termed small interfering DNAs (siDNAs). These siDNAs are utilized by TtAgo to cleave ssDNA and double stranded DNA (dsDNA) targets, the latter by cleaving each of the strands individually (11).

Archaeal pAgos initially have been explored to get insights in structural organization of Argonaute proteins (13,14). Along with the structure of *Archaeoglobus fulgidus* Argonaute (AfAgo), which is a truncated pAgo, the binding affinity of AfAgo for various nucleic acids has been described (14). The apo-AfAgo has a higher affinity for ssDNA and dsDNA molecules than for ssRNA and dsRNA molecules (14). The enzyme of *Methanocaldococcus jannaschii* (MjAgo) is the only archaeal Argonaute of which *in vitro* activity assays have been reported, revealing that this protein can utilize DNA guides to cleave DNA targets (15). Like AfAgo, MjAgo binds ssDNA and dsDNA molecules with a much higher affinity than ssRNA and dsRNA molecules (15). Furthermore, MjAgo is unable to cleave RNA target strands. The physiological role of MjAgo has not been characterized. MjAgo belongs to a clade of euryarchaeal thermophilic Argonautes, which is the pAgo clade that is most closely related to the eukaryotic Arg-

\*To whom correspondence should be addressed. Tel: +31 317483108; Fax: +31 317483829 Email: john.vanderoost@wur.nl

onautes (3,6) (Figure 1A). This clade also contains the *Pyrococcus furiosus* Argonaute (*PfAgo*). *P. furiosus* is a hyperthermophilic archaeon, like *A. fulgidus* and *M. jannaschii*, and grows optimally at temperatures between 80 and 100°C (16). *PfAgo* was the first Ago whose complete three-dimensional structure was determined (13). However, the physiological role and the molecular mechanism of *PfAgo* have not yet been reported. Despite belonging to the same clade, *PfAgo* has only 28% shared identity with *MjAgo* (BLASTp), which makes it difficult to predict the functionality of *PfAgo* based on sequence homology alone. Furthermore, the closest related pAgo of which the function is determined is bacterial *TtAgo*, that has a query cover of only 38%, and only 24% shared identities with *PfAgo* (BLASTp). In the present study, a combination of *in vivo* and *in vitro* analyses demonstrates that archaeal *PfAgo* is involved in host defense by mediating DNA-guided DNA interference.

## MATERIALS AND METHODS

### Strains and cultivation

For *in vivo* experiments, *P. furiosus* strain JFW02 (17) was used, which is referred to in this chapter as Pfu or wild-type. Furthermore two genomic variants of strain JFW02, *PfuΔago* (*ago* knockout strain) and *Pfu-ago-O/E* (*ago* over-expression strain) were used (Figure 1; Supplementary Table S1). Strains were cultivated anaerobically in a defined medium with cellobiose as the carbon source (18) at 90°C in anaerobic culture bottles or on medium solidified with 1% (w/v) Gelrite (Research product international). For growth of uracil auxotrophic strains, the defined medium was supplemented with uracil to a final concentration of 20 μM.

### Genomic mutants

*Pfu-ago-O/E* was generated by transforming an *Nru*I-linearized pHSG298 *ago* plasmid (Supplementary Table S2) into the wild-type Pfu strain. The plasmid contains the gene encoding *PfAgo* with an upstream *Thermococcus kodakaraensis* *csg* promoter, flanked by PF1223 and PF1224 gene sequences for homologous recombination. Furthermore, the plasmid encodes *TrpA* and *TrpB* with an upstream *gdh* promoter. The plasmid was constructed by overlap polymerase chain reaction and oligonucleotides used to generate this plasmid are shown in Supplementary Table S3. Two rounds of colony purification was performed by plating 10<sup>-3</sup> dilutions of transformant cultures onto selective plate medium (without tryptophan) and picking isolated colonies into selective liquid medium. The *PfuΔago* was created by pop-out marker replacement strategy as described previously (17). The sequences of oligonucleotide used are shown in Supplementary Table S3.

### Transformation experiments

Plasmid transformation was performed as described previously using 2.5 ng plasmid DNA μl<sup>-1</sup> of culture (18). pJFW18 plasmid (19) and a modified pYS3 plasmid (20) were used (Supplementary Table S2). The modified pYS3 plasmid was generated by replacement of the Sim<sup>R</sup> cassette

of pYS3 by the *Pgdh pyrF* cassette from pJFW18. The sequences of oligonucleotides used are shown in Supplementary Table S3. The transformation efficiencies reported were calculated as the number of transformed colonies per μg of DNA added.

### *PfAgo* antibodies and *PfAgo* immunodetection

Polyclonal antibodies were raised in chickens against purified (nickel chromatography) N-terminal, 6X His-tagged recombinant *PfAgo* or *PfCsa2* (loading control) proteins as previously described (21). Western blotting was performed by standard procedures. The blots were incubated with polyclonal IgY immune antibodies and HRP-conjugated Anti-IgY secondary antibody (Gallus Immunotech). The protein bands on the blots were detected using an enhanced chemiluminescent substrate for HRP (horse radish peroxidase) activity (GE Life Sciences) and exposure to autoradiography film.

### Sequence and structural alignments

For the active site residue sequence alignment, previously published alignments were used (3). For structural alignments, structures from *PfAgo* (PDB: 1Z25) and *TtAgo* (PDB: 4N47) were loaded in PyMOL. All residues but the MID and PIWI domains were deleted and the remaining residues were aligned with the PyMOL software. For clarity, only residues forming the active site of *PfAgo*, and active site residues of *TtAgo* were displayed.

### *PfAgo* expression and purification

A synthetic codon-optimized gene encoding *PfAgo* with an N-terminal Strep(II)-tag was ordered from GenScript USA Inc. and was directionally cloned into expression vector pCDF-1b as indicated in Supplementary Table S2 (pWUR790). Plasmids pWUR791-pWUR796 (Supplementary Table S2) were generated by introducing mutations according to an adapted QuikChange Site-Directed mutagenesis Kit instruction manual (Stratagene) using primers described in Supplementary Table S3. These plasmids were transformed into *Escherichia coli* KRX (Promega) according to the protocol provided by the manufacturer. Strains were cultivated in LB medium containing 50 μg ml<sup>-1</sup> streptomycin and 0.4% (w/v) glucose in a shaker incubator at 37°C. After overnight incubation, cultures were centrifuged for 5 min at 4700 × *g*, after which the supernatant was removed. Cell pellets were resuspended in LB medium containing 50 μg ml<sup>-1</sup> streptomycin and incubated in shaker incubator at 37°C until an OD<sub>600 nm</sub> of 0.6–0.8 was reached. Cultures were cold-shocked by incubation in an ice bath for 15 min. *PfAgo* expression was induced by adding isopropyl-β-D-thiogalactoside (IPTG) and L-Rhamnose to a final concentration of 1 mM and 0.1% (w/v), respectively. Expression was continued in a shaker incubator at 20 or 37°C for 16 h. Cells were collected by centrifugation for 15 min at 6000 × *g*, after which the supernatant was removed. Cells were resuspended in Buffer I (20 mM Tris/HCl pH 8, 1 M NaCl, 2 mM MnCl<sub>2</sub>) and disrupted by sonication with a Branson Sonifier B-12 and Branson Converter with a 5 mm

tip (ten 30 s pulses at 30% power with 30 s pause between pulses). The solution was centrifuged for 30 min at  $32\ 913 \times g$  at  $4^\circ\text{C}$ , after which the supernatant was used for purification by Strep-Tactin affinity chromatography (IBA, Germany) with an adapted protocol. Before loading of the supernatant, the Strep-Tactin Sepharose column was equilibrated with Buffer I. After loading, the column was washed 16 CV (column volumes) Buffer I. N-terminally Strep(II)-tagged *PfAgo* was eluted in Buffer II (buffer I supplemented with 2.5 mM biotin (Sigma-Aldrich)). For purification of *PfAgo* used in cation preference and cation gradient experiments, no  $\text{MnCl}_2$  was added to the purification buffers.

### Activity assays

For activity assays, elution fractions containing *PfAgo* or *PfAgo* mutants were diluted with Buffer II to a final protein concentration of  $5\ \mu\text{M}$ . A total of  $5\ \mu\text{l}$  protein sample was mixed with synthetic ssDNA or ssRNA guides (Supplementary Table S3) in a 5:1 ratio (protein:guide) in reaction buffer (20 mM Tris/HCl, pH 8) and incubated for 15 min at  $95^\circ\text{C}$ . After pre-incubation, ssDNA or ssRNA targets (Supplementary Table S3) were added to a final 5:1:1 ratio (protein:guide:target) and incubated for 1 h at  $95^\circ\text{C}$ . Final reaction concentrations were 15 mM Tris/HCl pH 8, 250 mM NaCl,  $0.5\ \mu\text{M}$   $\text{MnCl}_2$ ,  $1.25\ \mu\text{M}$  protein,  $0.25\ \mu\text{M}$  guide and  $0.25\ \mu\text{M}$  target. Note that in different experiments incubation temperature, incubation time and salt concentration were varied (indicated in figures of corresponding experiments). For cation preference experiments, *PfAgo* or *PfAgoDM* and ssDNA guides (Supplementary Table S3) were mixed in a 5:1 ratio (protein:guide) in reaction buffer to which different cations were added, and incubated for 15 min at  $95^\circ\text{C}$ . After pre-incubation, ssDNA targets (Supplementary Table S3) were added to a final 5:1:1 ratio (protein:guide:target) and incubated for 1 h at  $95^\circ\text{C}$ . Final reaction concentrations were 15 mM Tris/HCl pH 8, 250 mM NaCl,  $0.5\ \text{mM}$  metal- $\text{Cl}_2$  ( $\text{FeCl}_2$ ,  $\text{CoCl}_2$ ,  $\text{CuCl}_2$ ,  $\text{NiCl}_2$ ,  $\text{MgCl}_2$ ,  $\text{MnCl}_2$ ,  $\text{ZnCl}_2$  or  $\text{CaCl}_2$ ),  $1.25\ \mu\text{M}$  protein,  $0.25\ \mu\text{M}$  guide,  $0.25\ \mu\text{M}$  target. After incubation Loading Buffer (95% (deionized) formamide, 5 mM ethylenediaminetetraacetic acid, 0.025% sodium dodecyl sulphate, 0.025% Bromophenol blue and 0.025% xylene cyanol) was added in a 1:1 ratio and samples were incubated for 10 min at  $95^\circ\text{C}$  before resolving on 15 or 20% denaturing polyacrylamide gels. Nucleic acids were stained using SYBR gold Nucleic Acid Gel Stain (Invitrogen) and visualized using a G:BOX Chemi imager (Syngene). For plasmid assays, elution fractions containing *PfAgo* or *PfAgo* mutants were diluted with Buffer II to a final protein concentration of  $5\ \mu\text{M}$ . A total of  $5\text{-}\mu\text{l}$  protein sample was mixed with synthetic ssDNA guides (Supplementary Table S3) in a 5:1 ratio (protein:guide) in reaction buffer (20 mM Tris/HCl, pH 8 and varying NaCl concentrations) and incubated for 15 min at  $95^\circ\text{C}$ . After pre-incubation,  $\sim 200\text{--}300\ \text{ng}$  plasmid pWUR704 or pWUR790 (Supplementary Table S2) was added, and the samples (final reaction concentrations were 15 mM Tris/HCl pH 8,  $0.5\ \mu\text{M}$   $\text{MnCl}_2$ , and 250 mM or 500 mM NaCl) were incubated for 16 h at  $75^\circ\text{C}$ . Reactions were stopped by adding Proteinase K solution (Ambion) and  $\text{CaCl}_2$  (final concentration 5 mM) and samples were

incubated for 1 h at  $65^\circ\text{C}$ . Samples were mixed with  $6\times$  loading dye (Thermo Scientific) before they were resolved on 0.8% agarose gels. As marker, either a 1 kb Generuler Marker (Thermo Scientific) or 1 kb DNA ladder (New England Biolabs) and additionally a custom plasmid marker, were used. The custom plasmid marker consisted of non-treated pWUR704 (mostly in supercoiled conformation), Nb.BSMI (New England Biolabs) nicked pWUR704 (open circular conformation) and BcuI (Thermo Scientific) linearized pWUR704. Agarose gels were stained with SYBR-gold Nucleic Acid Gel Stain (Invitrogen) and visualized using a G:BOX Chemi imager (Syngene).

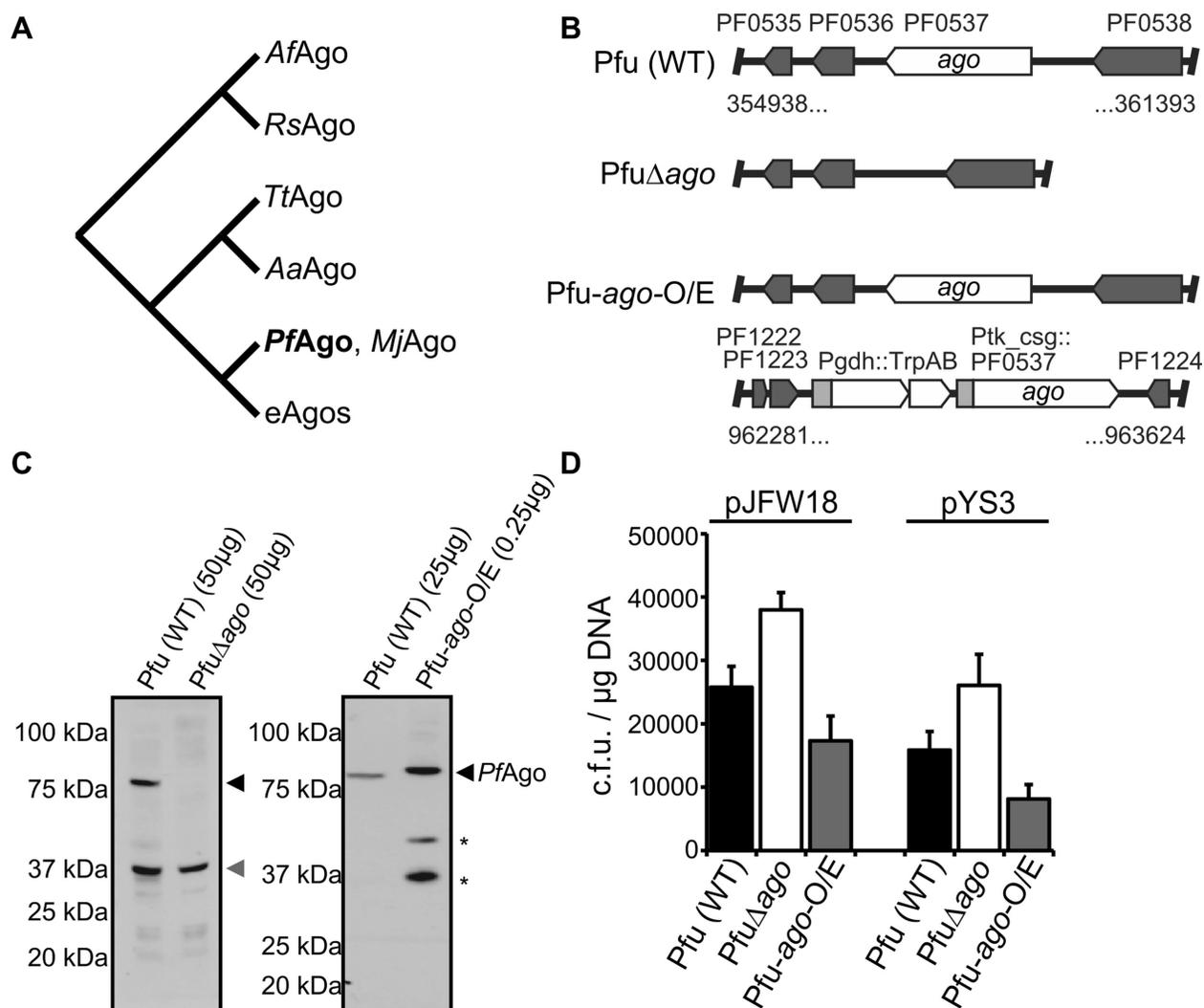
### Guide co-purification

About 1/10 volume Proteinase K (Ambion) and 1/10 volume  $\text{CaCl}_2$  (50 mM) were added to 500 pmol purified proteins in Buffer II and samples were incubated for 4 h at  $65^\circ\text{C}$ . Nucleic acids were separated from protein content using Roti(R) Phenol/Chloroform/Isoamyl alcohol (Carl Roth GmbH). About 1/10 volume 3 M NaAc (pH 5.2) and  $10\ \mu\text{l}$  0.5% linear polyacrylamide was added, before addition of 96% ethanol in a 2:1 ratio (v/v, ethanol:sample). Samples were incubated at  $-20^\circ\text{C}$  for 2 days and centrifuged for 30 min at  $4^\circ\text{C}$  at  $20\ 000 \times \text{rpm}$  in a table top centrifuge. Supernatant was removed and the pellet was dried for 10 min at  $50^\circ\text{C}$ . The pellet was washed in 70% ethanol, followed by centrifugation for 30 min at  $4^\circ\text{C}$  at  $20\ 000 \times \text{rpm}$  in a table top centrifuge. Supernatant was removed and the pellet was dried for 10 min at  $50^\circ\text{C}$ . The pellet was resuspended in  $50\ \mu\text{l}$  MilliQ  $\text{H}_2\text{O}$ . Purified nucleic acids were [ $\gamma\text{-}^{32}\text{P}$ ] ATP labeled with T4 PNK (Thermo Scientific) in exchange or forward labeling reactions, in which 5' phosphates or 5' OH groups are labeled, respectively. Nucleic acids were separated from free [ $\gamma\text{-}^{32}\text{P}$ ] ATP using a Sephadex G-25 column (GE). Nucleic acids were incubated with DNase-free RNase A (Thermo Scientific) or RQ1 RNase-free DNase I (Promega) for 1 h at  $37^\circ\text{C}$ . Samples were mixed with Loading Buffer in a 1:1 ratio and samples were incubated for 10 min at  $95^\circ\text{C}$ . [ $\gamma\text{-}^{32}\text{P}$ ] ATP labeled nucleic acids were resolved on 20% denaturing polyacrylamide gels. Radioactivity was captured from gels using phosphor screens.

## RESULTS

### *PfAgo* interferes with plasmid DNA transformation

To elucidate the physiological role of *PfAgo* in *P. furiosus* JFW02 (Pfu) (17), an *ago* knockout strain (Pfu $\Delta$ *ago*) and an *ago* overexpressing strain (Pfu-*ago*-O/E) were generated (Figure 1B). Immunoblots demonstrate that *PfAgo* is expressed in Pfu, while no *PfAgo* is detected in Pfu $\Delta$ *ago* (Figure 1C). Approximately 200-fold increased levels of full-length *PfAgo* were detected in the Pfu-*ago*-O/E strain (Figure 1C). Bacterial *TiAgo* has previously been described to be involved in host defense by interfering with plasmid transformation (11,12). In the latter studies, transformations were performed using the natural competence system of *T. thermophilus*, which transports extracellular DNA into the cell (22). *P. furiosus* JFW02 also has a natural competence system (18). This allowed comparing natural transformation efficiencies of strains Pfu, Pfu $\Delta$ *ago* and Pfu-*ago*-



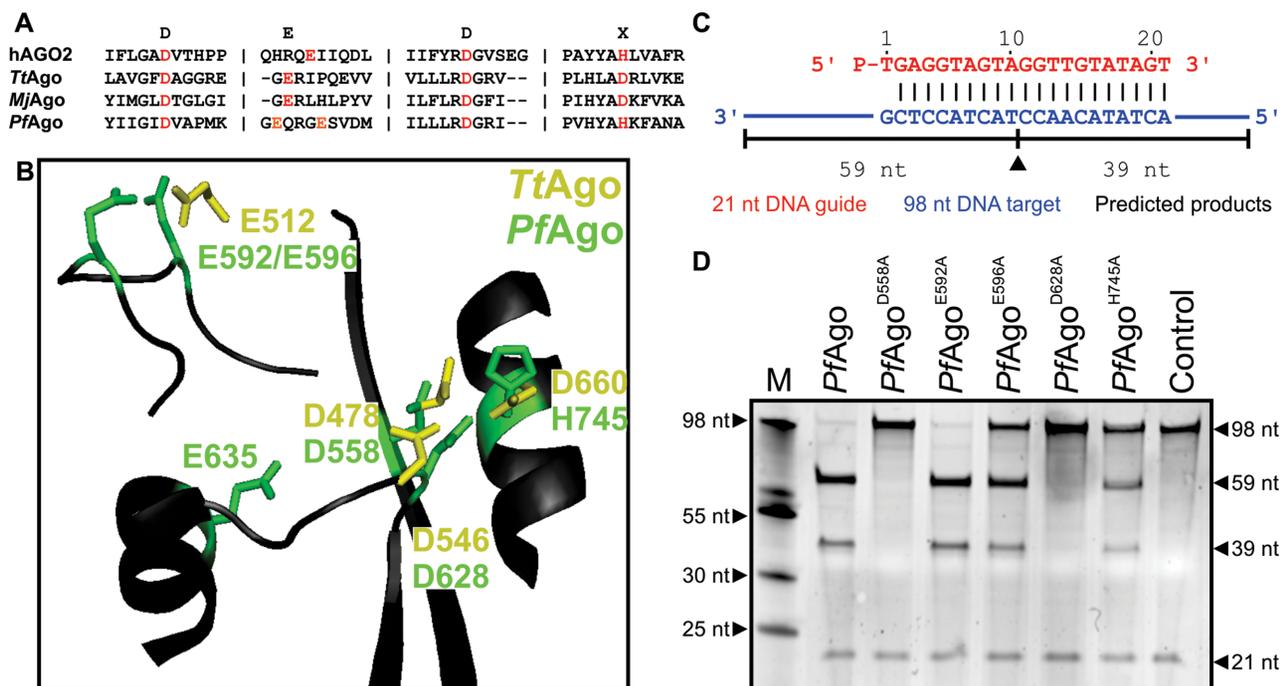
**Figure 1.** *PfAgo* interferes with plasmid transformation. (A) Schematic phylogenetic tree of Argonaute proteins, adapted from (3). *AfAgo*: *Archaeoglobus fulgidus* Ago. *RsAgo*: *Rhodobacter sphaeroides* Ago. *TtAgo*: *Thermus thermophilus* Ago. *AaAgo*: *Aquifex aeolicus* Ago. *PfAgo*: *Pyrococcus furiosus* Ago. *MjAgo*: *Methanocaldococcus jannaschii* Ago. eAgos: eukaryotic Agos. (B) Overview of *ago* gene loci of *P. furiosus* strains Pfu (wild-type), *PfuΔago* (knock-out) and *Pfu-ago-O/E* (*PfAgo* overexpression strain). (C) Immunoblot analysis of *PfAgo* (indicated with a black triangle) content in Pfu and *PfuΔago* with Csa2 protein (indicated with a gray triangle) serving as the internal standard (left panel) or Pfu and *Pfu-ago-O/E* (right panel). The amount of lysate analyzed is indicated and the asterisk (\*) denotes apparent breakdown products observed when *PfAgo* is overexpressed. (D) Plasmid transformation efficiencies of the *P. furiosus* strains. Error bars indicate standard deviations of biological triplicates.

O/E. Each strain was transformed with plasmid pJFW18 (19) or with plasmid pYS3 (20). For both plasmids, transformation efficiency is ~1.5 to 2.5-fold higher in *PfuΔago* compared to wild-type Pfu (Figure 1D,  $P < 0.001$ ). When *PfAgo* is overexpressed, the transformation efficiency is lowered even further (Figure 1D,  $P < 0.001$ ). This demonstrates that archaeal *PfAgo* lowers competence by interfering with plasmid transformation.

#### *PfAgo* has a DEDH catalytic tetrad

Target cleavage by Agos is mediated by a conserved DEDX triad, of which the X can be a histidine or an aspartic acid (3) (Figure 2A). Whereas the DDX residues are positioned close together in the available pAgo structures, the glutamic acid (E) is located on a structural sub-domain termed the 'glutamate finger' (23). In *TtAgo*, this finger is located at

a distance of 12.8 Å from the catalytic site when no target nucleic acid is bound (unplugged conformation), but it is inserted into the catalytic site (plugged-in conformation) upon target binding (24). To identify the catalytic residues of *PfAgo*, a sequence alignment (3) and a structural alignment of *TtAgo* and *PfAgo* were analyzed (Figure 2A, B). Previously, it was predicted that the catalytic site of *PfAgo* includes residues D558, D628 and E635 (13). However, E635 is located away from the catalytic site residues of other Agos both in sequence and structural alignments (Figure 2). In contrast, H745 perfectly aligns with catalytic residues from other Agos in both sequences and structures. Like in the structure of unplugged *TtAgo*, the glutamate finger of *PfAgo* is unplugged in the structure of *PfAgo* (Figure 2B). However, the *PfAgo* glutamate finger encompasses two glutamic acids (E592 and E596). Which of these two gluta-



**Figure 2.** DEDH catalytic site of *PfAgo*. (A) Sequence alignment of human AGO2 (hAGO2), *TtAgo*, *MjAgo* and *PfAgo*, adapted from (3). Only regions containing the DEDX catalytic residues (indicated in red) are shown. *PfAgo* catalytic residues E592 and E596 are colored orange. (B) *TtAgo* catalytic residues DEDD (yellow; PDB: 4N47) aligned to *PfAgo* catalytic site (black; PDB: 1Z25). Predicted catalytic residues of *PfAgo* are colored green. (C) Synthetic 21 nucleotide siDNA (red) and 98 nucleotide ssDNA target (blue) used for *in vitro* activity assays. The black triangle indicates the predicted cleavage site, black lines indicate the predicted 59 and 39 nucleotide cleavage products. (D) *PfAgo* and mutants were loaded with a 21 nucleotide long siDNA and were incubated with a 98 nucleotide ssDNA target in a 5:1:1 molar ratio (*PfAgo*:guide:target). Products were resolved on a 15% denaturing polyacrylamide gel. M: ssDNA marker. nt: nucleotide. The 'Control' sample contains no protein.

mate residues is involved in target cleavage is impossible to deduce from the sequence alignment or from the available *PfAgo* structure (unplugged conformation).

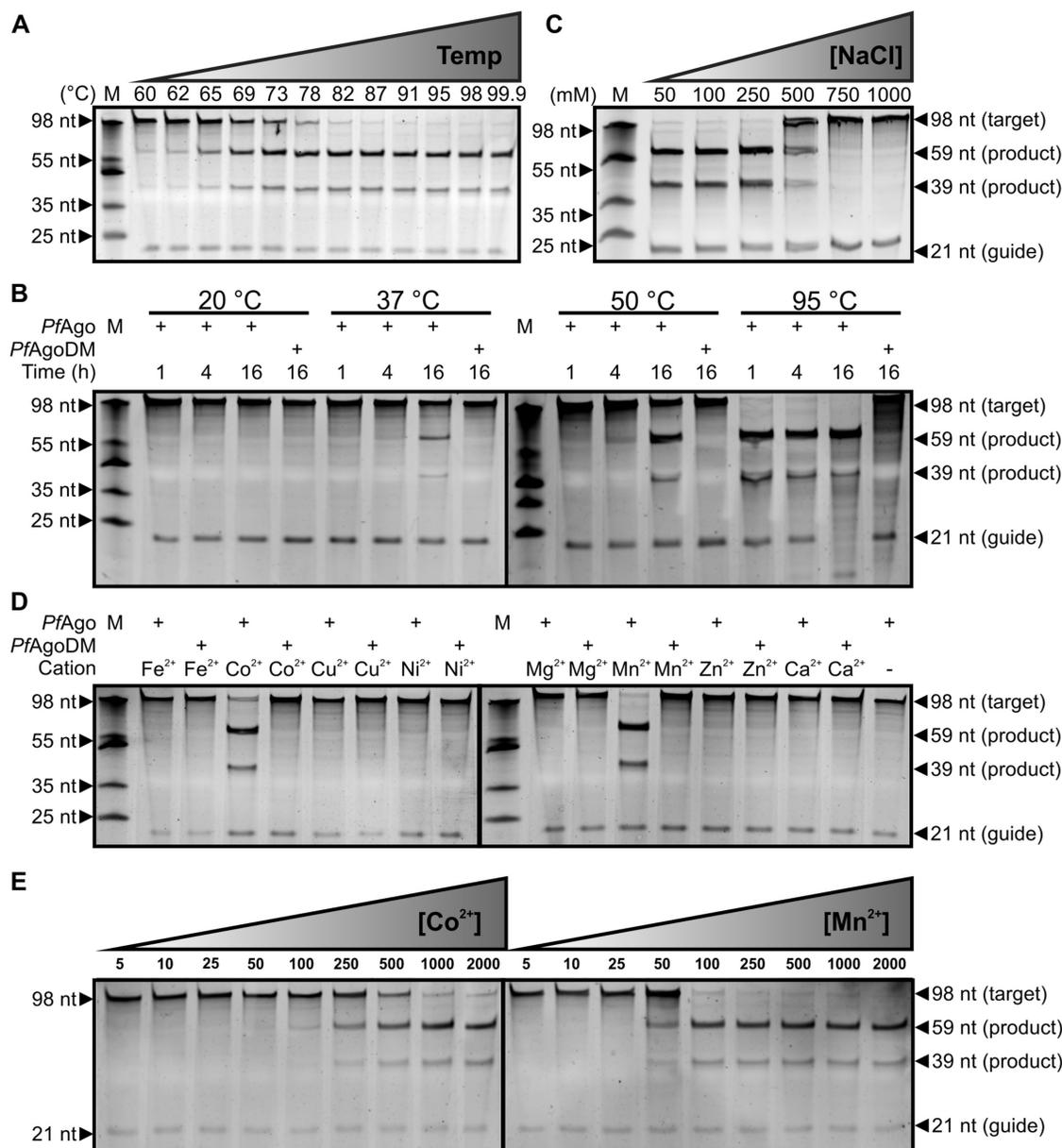
To experimentally identify the residues involved in *PfAgo* activity, Strep(II)-tagged *PfAgo* and five predicted catalytic mutants were heterologously produced in *E. coli* KRX. After affinity purification, expressed *PfAgo* variants were tested for activity. As both *TtAgo* and *MjAgo* (which is closely related to *PfAgo*) can utilize DNA guides to cleave ssDNA targets, we incubated *PfAgo* with synthetic 5'-phosphorylated, 21 nt long siDNAs. Subsequently, a synthetic 98 nt ssDNA target was added (Figure 2C and D). After 1 h incubation at 95°C with *PfAgo* and with *PfAgo*<sup>E592A</sup>, all ssDNA target is cleaved (Figure 2D). This indicates that *PfAgo* can utilize siDNAs to cleave DNA targets. In contrast, *PfAgo* mutants D558A, E596A, D628A and H745A show impaired activity (Figure 2D), indicating that the latter residues form the DEDH catalytic tetrad of *PfAgo*.

### Requirements for target cleavage by *PfAgo*

To further determine the prerequisites for *PfAgo*-mediated target cleavage, we tested the influence of temperature, salt concentration and divalent cation type on siDNA-guided ssDNA cleavage. As a negative control, we used a catalytic double mutant, *PfAgo*DM (*PfAgo*<sup>D558A,D628A</sup>). For all assays, the guide and target shown in Figure 2C were used. *PfAgo* is most active in the range from 87 °C to 99.9 °C (Figure 3A; higher temperatures were not tested). Longer

(16 h) incubations show that *PfAgo* exhibits some activity at 37°C, but not at 20°C (Figure 3B). *PfAgo* is active in reactions with a NaCl concentration of 50 to 250 mM, whereas at NaCl concentrations of 500 to 1000 mM the activity is lowered or absent (Figure 3C; Supplementary Figure S1). To investigate if *PfAgo* functions as a multi-turnover protein *PfAgo*, siDNAs and ssDNA targets were incubated in a 2.5:1:20 ratio (*PfAgo*:siDNA:target). *PfAgo*-siDNA complexes cleave >95% of the 20-fold excess of target DNA within 30 min in buffer with 250 mM NaCl (Supplementary Figure S1). To investigate if higher NaCl concentrations inhibit multi-turnover reactions (for example by restricting the release of cleaved target strands) we additionally performed the same experiment at 500 mM and 1 M NaCl (Supplementary Figure S1). Even after 16 h, no activity is observed at 1M NaCl. At 500 mM NaCl the activity of *PfAgo* is severely lowered, but it still depletes the 20-fold excess of target DNA after 16 h, indicating that high NaCl concentrations do not inhibit the multi-turnover characteristic of *PfAgo*.

Next, we investigated which divalent cations *PfAgo* can utilize to mediate siDNA-guided DNA target cleavage. *PfAgo* is able to utilize Mn<sup>2+</sup> and Co<sup>2+</sup> as cation, with Mn<sup>2+</sup> being a better cation than Co<sup>2+</sup> (Figure 3D and E). In contrast to the phylogenetically closely related *MjAgo* (15), *TtAgo* (9,11,24) and eAgos (25), *PfAgo* is unable to use Mg<sup>2+</sup> as cation for its activity (Figure 3D). Also Fe<sup>2+</sup>, Cu<sup>2+</sup>, Ni<sup>2+</sup> and Ca<sup>2+</sup> do not allow siDNA-guided *PfAgo* cleavage of ssDNA (Figure 3D). As eAgos have been shown



**Figure 3.** Effect of temperature, salt concentration and cation on *PfAgo* activity. *PfAgo* loaded with a 21 nt long siDNA was incubated with a 98 nt ssDNA target (see Figure 2C) in a 5:1:1 molar ratio (*PfAgo*:guide:target) under various conditions. Unless otherwise indicated, target cleavage took place at 95°C for 1 h, with 0.5 mM Mn<sup>2+</sup> as cation. Nucleic acids are resolved on denaturing polyacrylamide gels. M: ssDNA marker. nt: nucleotide. (A) *PfAgo* activity is highest at temperatures between 90 and 99.9°C. (B) *PfAgo* shows activity at temperatures  $\geq 37^\circ\text{C}$  if incubation is extended. (C) NaCl concentrations  $\geq 500$  mM interfere with *PfAgo* activity. (D) *PfAgo*-guide complexes show Co<sup>2+</sup> and Mn<sup>2+</sup> mediated ssDNA target cleavage. (E) Mn<sup>2+</sup> is preferred above Co<sup>2+</sup> as cation for *PfAgo*-guide mediated ssDNA target cleavage.

to preferentially bind guides with specific 5'-end nucleotides (26,27), we tested if *PfAgo* has a preference for a specific 5'-end nucleotide on the siDNA. Like *TtAgo*, *PfAgo* is able to utilize siDNAs with different 5'-end nucleotides equally well (Supplementary Figure S2).

#### *PfAgo* is a DNA-guided protein that cleaves DNA targets

To investigate whether *PfAgo* acquires specific guides *in vivo*, we analyzed nucleic acids that co-purify with *PfAgo* after expression in *E. coli*. No DNA was observed in RNase A-treated samples, whereas in DNase I-treated samples it

was observed that RNAs of undefined length co-purify with *PfAgo* and *PfAgoDM* (Supplementary Figure S3). The latter is suspected to be non-specifically bound RNA, as has previously been described for purification of *TtAgo* and *TtAgoDM* (3). This is supported by the observation that RNAs associate with *PfAgo* purified in presence and absence Mn<sup>2+</sup>, whereas a divalent cation is required for specific binding of the 5' end of the guide by pAgos (28). Unfortunately, attempts to identify guides associated with *PfAgo* expressed in *P. furiosus* were not successful (data not shown).

To investigate if besides ssDNA, ssRNA can also guide *PfAgo* activity, *PfAgo* was incubated with 21 nt DNA or RNA guides after which 45 nt ssDNA or ssRNA targets were added (Figure 4A and B). As expected, *PfAgo* catalyzes siDNA-guided cleavage of complementary DNA targets (Figure 4B). In contrast to *TtAgo*, *PfAgo* does not show siDNA-guided cleavage of RNA targets (Figure 4B). With the RNA guide, we did not observe any *PfAgo*-mediated cleavage, neither of ssDNA nor of ssRNA targets. This implies that *PfAgo* only mediates DNA-guided DNA interference. In addition, the DNA guide length range of *PfAgo* was tested. *PfAgo*-mediated cleavage of DNA targets is only observed with siDNAs with a length ranging from 15 to at least 31 nt (longer not tested; Figure 4C and D). *TtAgo* utilizes siDNAs with a length ranging from 9 to at least 36 nt (longer not tested) (29). As guides with a length of 31 nt are too long for canonical guide binding by Ago, it has been hypothesized that the guide adopts an alternative trajectory to allow 3' end insertion into the PAZ binding pocket (29). Alternatively, the 3' end of guide sticks out of the protein, as the nucleic acid-binding channel of *TtAgo* is open to the outside (29). Most likely, the same is true for *PfAgo*.

#### ***PfAgo* mediates guide-free and siDNA-guided cleavage of dsDNA plasmids**

To test whether *PfAgo* cleaves plasmid DNA, *PfAgo* was incubated with its expression plasmid pWUR790 (Figure 5A). As incubation of this plasmid at 95°C in the presence of Mn<sup>2+</sup> results in degradation of the plasmid (even in the absence of *PfAgo*), we incubated the reaction mixtures at 75°C for 16 h. Still, the majority of the plasmid DNA is turned to the open circular confirmation under these conditions, even in absence of *PfAgo* (Figure 5B). Strikingly, pWUR790 is linearized when incubated with *PfAgo* in absence of guides (Figure 5B). We observed this suspected guide-free *PfAgo*-mediated cleavage of pWUR790 in buffer with 250 mM NaCl, but not in buffer with 500 mM NaCl (Figure 5B). These findings suggest either that *PfAgo* cleaves plasmids independently of co-purified DNAs or that DNA guides are utilized but the concentration of such co-purified DNA is below the detection limit of the assay. It has previously been demonstrated that *in vivo*, *TtAgo* acquires guides that allow for targeting its expression vector (11). To rule out that this *PfAgo* activity was guided by co-purified RNA or by co-purified DNA guides that we were unable to detect, we used pWUR704 as target plasmid (Figure 5C). pWUR704 has minimal sequence similarity to the *PfAgo* expression vector (pWUR790; no sequences longer than 13 consecutive identical base pairs between the two plasmids). After incubation without *PfAgo*, the plasmid is present both in open circular and supercoiled configuration (Figure 5E, lane 1). When *PfAgo* is added, supercoiled pWUR704 is linearized even in absence of guides (Figure 5E, lane 2). Like pWUR790 cleavage, this activity is more pronounced at 250 mM NaCl compared to 500 mM NaCl. In contrast, *TtAgo*, which co-purified with detectable levels of siDNA, was unable to cleave pWUR704, unless synthetic siDNAs targeting pWUR790 were added (3). Low levels of guide-free *PfAgo*-mediated nicking of pWUR704 takes place within 1 h incu-

bation at 75°C (Supplementary Figure S4). As *PfAgo* does not mediate RNA-guided activity within the same time span (Figure 4), it is unlikely that the activity on plasmid DNA is mediated by RNA that co-purified with *PfAgo*. Instead, these findings suggest that *PfAgo* has the potential to cleave the plasmid independently of guides.

We next investigated whether siDNAs can guide *PfAgo*-mediated plasmid cleavage. We loaded *PfAgo* either with a FW siDNA targeting the (–) strand of pWUR704 or with a RV siDNA targeting the (+) strand of pWUR704 (Figure 5D). *PfAgo* and *PfAgo*–siDNA complexes were incubated with plasmid pWUR704 as target, in buffers containing 250 or 500 mM NaCl. Additionally, both *PfAgo* complexes with FW guides and with RV guides were mixed and incubated with pWUR704. *PfAgo* loaded with a single siDNA nicks the plasmid DNA, generating open circular plasmids (Figure 5E, lane 3 and 4). Interestingly, reactions to which FW or RV siDNAs are added contain more open circular plasmid compared to the sample to which no siDNA is added, in which plasmids are mostly turned to the linear state. This suggests siDNA loading hinders guide-free *PfAgo* activity or alternatively that guide-free cleavage of plasmid DNA by *PfAgo* requires plasmid supercoiling. Incubation of the plasmids with *PfAgo*–siDNA complexes targeting both strands of the plasmid results in plasmid linearization (Figure 5E, lane 5). Both processes are performed more efficiently in reactions with 250 mM NaCl as compared to reactions with 500 mM NaCl. More efficient cleavage at 250 mM NaCl has also been observed for ssDNA targets (Figure 3C). These findings demonstrate that *PfAgo*–siDNA complexes can target dsDNA plasmids, resulting in a dsDNA breaks if both strands of the plasmid DNA are targeted.

## **DISCUSSION**

This work describes the first combined *in vivo* and *in vitro* characterization of an archaeal Argonaute. *PfAgo* functions as a DNA-guided DNA endonuclease that requires divalent cations such as Mn<sup>2+</sup> or Co<sup>2+</sup> for its activity. It utilizes 5'-phosphorylated ssDNA guides that are at least 15 and up to at least 31 nt long. As such long guides cannot be bound canonically, we predict that the 3'-end of the guide 'sticks out' a gap in the PAZ domain. Whether this gap exists in the PAZ domain of *PfAgo* and the function of this predicted feature remain to be determined. Like the bacterial *TtAgo*, archaeal *PfAgo*s can utilize a single siDNA to nick plasmid DNA, while two *PfAgo*–siDNA complexes (each targeting a single strand of the plasmid) together can generate dsDNA breaks. Both pAgo mediate host defense by interfering with invading plasmid DNA (Figure 1). As *TtAgo* and *PfAgo* do not belong to the same branch in the phylogenetic tree, these findings suggest a broad conservation of pAgo functions. Nevertheless, *PfAgo* has some interesting characteristics: *PfAgo* cannot use Mg<sup>2+</sup> as cation (while *TtAgo* and *MjAgo* can (11,15)) and *PfAgo*–siDNA complexes do not mediate RNA target cleavage (while *TtAgo*–siDNA complexes do (9)). The preference for DNA or RNA guides or targets of eAgo, *TtAgo* and *PfAgo* most likely is determined on a structural level. The guide-target duplex binding channel and guide 5'-end



binding pockets of eAgos are strongly positively charged, allowing RNA guide and target binding. In contrast, the duplex binding channels of *TtAgo* and *PfAgo* are much less positively charged and their guide 5'-end binding pockets are hydrophobic (30), possibly excluding the possibility to bind RNA guides. However, this hypothesis requires further investigation before any claims can be made.

We were unable to co-purify siDNA guides from *PfAgo* heterologously expressed in *E. coli* KRX. Possibly *PfAgo* is not able to acquire guides at the low temperatures at which it is expressed (20 or 37°C) or it requires host factors for guide generation and/or loading. Future research should focus on guides associated with *PfAgo* expressed in *P. furiosus*. Interestingly, under some conditions, *PfAgo* appears to cleave dsDNA plasmids in absence of siDNAs, suggesting that *PfAgo* shows unguided, non-specific nuclease activity (Figure 5). We have observed similar activity for *TtAgo* under specific conditions (Swarts *et al.*, unpublished results) and hypothesize that unguided dsDNA cleavage might be related to the generation of siDNAs.

We chose to characterize *PfAgo* as it belongs to the clade of euryarchaeal pAgos, which is the clade of pAgos that is most closely related to eAgos (Figure 1A). eAgos are best known for their RNA-guided RNA cleaving role in RNA interference. Therefore, we predicted that *PfAgo* possibly was involved in both RNA interference and DNA interference. In contrast, *PfAgo* solely mediates DNA cleavage in a DNA guide-dependent manner. Strikingly, a recent comment paper cites numerous papers in which DNA binding by eAgo has been described (31). It was suggested that some eAgos might be involved in DNA binding, a hypothesis that requires further investigation. As *PfAgo* is one of the pAgos most closely related to eAgos, these findings suggest that eAgos were derived from the same ancestors as *PfAgo*, which might have utilized DNA guides and/or DNA targets as well. Alternatively, the pAgos changed their guide and target specificity multiple times during evolution. Combined, these findings indicate that bacterial and archaeal Argonautes and possibly even ancient eukaryotic Argonautes, mediate host defense by DNA-guided DNA interference.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank Stan J. J. Brouns for critical reading of the manuscript.

## FUNDING

TOP grant from the Netherlands Organization of Scientific Research [854-10-003 to J.vd.O.]; National Institutes of Health (NIH) [R01 GM54682 to M.P.T. and R.M.T.]. Funding for open access charge: TOP grant from the Netherlands Organization of Scientific Research [854-10-003 to J.vd.O.]; NIH [R01 GM54682 to M.P.T. and R.M.T.]. *Conflict of interest statement.* None declared.

## REFERENCES

- Ketting,R.F. (2011) The many faces of RNAi. *Dev. Cell*, **20**, 148–161.
- Meister,G. (2013) Argonaute proteins: functional insights and emerging roles. *Nat. Rev. Genet.*, **14**, 447–459.
- Swarts,D.C., Makarova,K., Wang,Y., Nakanishi,K., Ketting,R.F., Koonin,E.V., Patel,D.J. and van der Oost,J. (2014) The evolutionary journey of Argonaute proteins. *Nat. Struct. Mol. Biol.*, **21**, 743–753.
- Pratt,A.J. and MacRae,I.J. (2009) The RNA-induced silencing complex: a versatile gene-silencing machine. *J. Biol. Chem.*, **284**, 17897–17901.
- Shabalina,S.A. and Koonin,E.V. (2008) Origins and evolution of eukaryotic RNA interference. *Trends Ecol. Evol.*, **23**, 578–587.
- Makarova,K.S., Wolf,Y.I., van der Oost,J. and Koonin,E.V. (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.
- Yuan,Y.R., Pei,Y., Ma,J.B., Kuryavyi,V., Zhadina,M., Meister,G., Chen,H.Y., Dauter,Z., Tuschl,T. and Patel,D.J. (2005) Crystal structure of A-aeolicus Argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell*, **19**, 405–419.
- Rashid,U.J., Paterok,D., Koglin,A., Gohlke,H., Piehler,J. and Chen,J.C. (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.
- Wang,Y.L., Sheng,G., Juraneck,S., Tuschl,T. and Patel,D.J. (2008) Structure of the guide-strand-containing argonaute silencing complex. *Nature*, **456**, U209–U234.
- Olovnikov,I., Chan,K., Sachidanandam,R., Newman,D.K. and Aravin,A.A. (2013) Bacterial argonaute samples the transcriptome to identify foreign DNA. *Mol. Cell*, **51**, 594–605.
- Swarts,D.C., Jore,M.M., Westra,E.R., Zhu,Y., Janssen,J.H., Snijders,A.P., Wang,Y., Patel,D.J., Berenguer,J., Brouns,S.J. *et al.* (2014) DNA-guided DNA interference by a prokaryotic Argonaute. *Nature*, **507**, 258–261.
- Blesa,A., Cesar,C.E., Averhoff,B. and Berenguer,J. (2015) Noncanonical cell-to-cell DNA transfer in *Thermus* spp. is insensitive to argonaute-mediated interference. *J. Bacteriol.*, **197**, 138–146.
- Song,J.J., Smith,S.K., Hannon,G.J. and Joshua-Tor,L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*, **305**, 1434–1437.
- Ma,J.B., Yuan,Y.R., Meister,G., Pei,Y., Tuschl,T. and Patel,D.J. (2005) Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature*, **434**, 666–670.
- Zander,A., Holzmeister,P., Klose,D., Tinnefeld,P. and Grohmann,D. (2014) Single-molecule FRET supports the two-state model of Argonaute action. *RNA Biol.*, **11**, 45–56.
- Fiala,G. and Stetter,K.O. (1986) *Pyrococcus-furiosus* Sp-Nov represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100-degrees C. *Arch. Microbiol.*, **145**, 56–61.
- Farkas,J., Stirrett,K., Lipscomb,G.L., Nixon,W., Scott,R.A., Adams,M.W. and Westpheling,J. (2012) Recombinogenic properties of *Pyrococcus furiosus* strain COM1 enable rapid selection of targeted mutants. *Appl. Environ. Microbiol.*, **78**, 4669–4676.
- Lipscomb,G.L., Stirrett,K., Schut,G.J., Yang,F., Jenney,F.E. Jr, Scott,R.A., Adams,M.W. and Westpheling,J. (2011) Natural competence in the hyperthermophilic archaeon *Pyrococcus furiosus* facilitates genetic manipulation: construction of markerless deletions of genes encoding the two cytoplasmic hydrogenases. *Appl. Environ. Microbiol.*, **77**, 2232–2238.
- Farkas,J., Chung,D., DeBarry,M., Adams,M.W. and Westpheling,J. (2011) Defining components of the chromosomal origin of replication of the hyperthermophilic archaeon *Pyrococcus furiosus* needed for construction of a stable replicating shuttle vector. *Appl. Environ. Microbiol.*, **77**, 6343–6349.
- Waegel,I., Schmid,G., Thumann,S., Thomm,M. and Hausner,W. (2010) Shuttle vector-based transformation system for *Pyrococcus furiosus*. *Appl. Environ. Microbiol.*, **76**, 3308–3313.
- Carte,J., Pfister,N.T., Compton,M.M., Terns,R.M. and Terns,M.P. (2010) Binding and cleavage of CRISPR RNA by Cas6. *RNA*, **16**, 2181–2188.

22. Averhoff, B. (2009) Shuffling genes around in hot environments: the unique DNA transporter of *Thermus thermophilus*. *FEMS Microbiol. Rev.*, **33**, 611–626.
23. Nakanishi, K., Weinberg, D.E., Bartel, D.P. and Patel, D.J. (2012) Structure of yeast Argonaute with guide RNA. *Nature*, **486**, 368–374.
24. Sheng, G., Zhao, H., Wang, J., Rao, Y., Tian, W., Swarts, D.C., van der Oost, J., Patel, D.J. and Wang, Y. (2014) Structure-based cleavage mechanism of *Thermus thermophilus* Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 652–657.
25. Schwarz, D.S., Tomari, Y. and Zamore, P.D. (2004) The RNA-induced silencing complex is a Mg<sup>2+</sup>-dependent endonuclease. *Curr. Biol.*, **14**, 787–791.
26. Frank, F., Sonenberg, N. and Nagar, B. (2010) Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*, **465**, 818–822.
27. Frank, F., Hauver, J., Sonenberg, N. and Nagar, B. (2012) Arabidopsis Argonaute MID domains use their nucleotide specificity loop to sort small RNAs. *EMBO J.*, **31**, 3588–3595.
28. Parker, J.S. (2010) How to slice: snapshots of Argonaute in action. *Silence*, **1**, 3.
29. Wang, Y.L., Juranek, S., Li, H.T., Sheng, G., Tuschl, T. and Patel, D.J. (2008) Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*, **456**, U921–U972.
30. Willkomm, S., Zander, A., Gust, A. and Grohmann, D. (2015) A prokaryotic twist on argonaute function. *Life*, **5**, 538–553.
31. Smalheiser, N.R. and Gomes, O.L. (2014) Mammalian Argonaute-DNA binding? *Biol. Direct.*, **10**, 27.