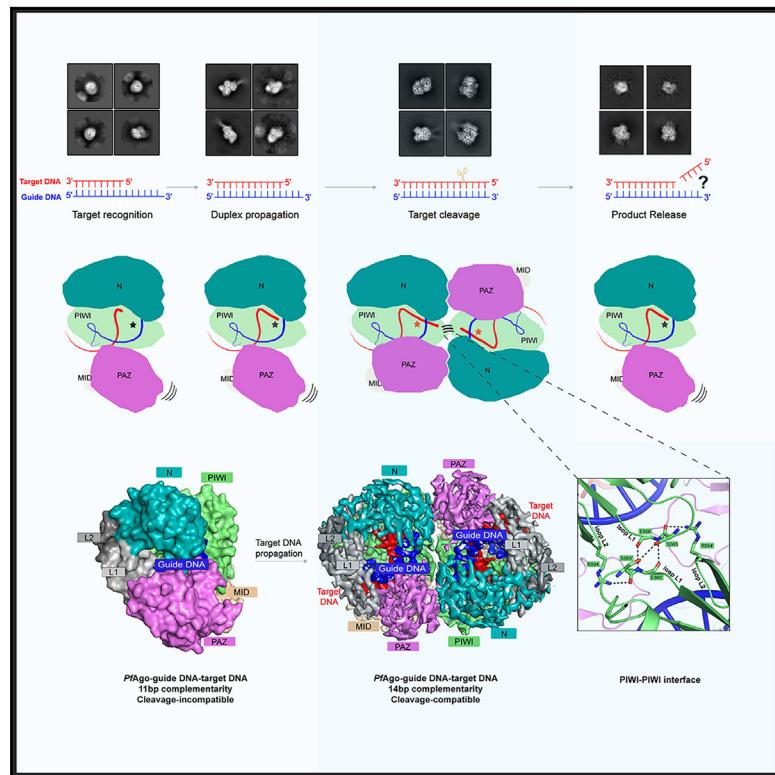


## Molecular mechanism for target recognition, dimerization, and activation of *Pyrococcus furiosus* Argonaute

### Graphical abstract



### Authors

Longyu Wang, Wanping Chen,  
Chendi Zhang, ..., Qiang Wei, Lixin Ma,  
Zhuang Li

### Correspondence

malixing@hubu.edu.cn (L.M.),  
zhuangli@hubu.edu.cn (Z.L.)

### In brief

Wang et al. report a target DNA-induced dimerization mechanism for the Argonaute from thermophilic *Pyrococcus furiosus*. They further observe a similar dimerization configuration in another closely related Argonaute and that the dimerization stabilizes the catalytic loops. This study provides several insights into the working mechanism of long prokaryotic Argonaute.

### Highlights

- The *PfAgo*-guide DNA-target DNA ternary complex is predominantly dimerized
- *PfAgo* dimerization is induced by engagement of target DNA of full complementarity
- Argonaute from *Thermococcus thioreducers* adopts a similar dimerization configuration
- Dimerization stabilizes the catalytic loops in both Argonautes

## Article

# Molecular mechanism for target recognition, dimerization, and activation of *Pyrococcus furiosus* Argonaute

Longyu Wang,<sup>1,2</sup> Wanping Chen,<sup>1,2</sup> Chendi Zhang,<sup>1,2</sup> Xiaochen Xie,<sup>1</sup> Fuyong Huang,<sup>1</sup> Miaomiao Chen,<sup>1</sup> Wuxiang Mao,<sup>1</sup> Na Yu,<sup>1</sup> Qiang Wei,<sup>1</sup> Lixin Ma,<sup>1,\*</sup> and Zhuang Li<sup>1,3,\*</sup>

<sup>1</sup>State Key Laboratory of Biocatalysis and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan, Hubei 430062, China

<sup>2</sup>These authors contributed equally

<sup>3</sup>Lead contact

\*Correspondence: malixing@hubu.edu.cn (L.M.), zhuangli@hubu.edu.cn (Z.L.)

<https://doi.org/10.1016/j.molcel.2024.01.004>

## SUMMARY

The Argonaute nuclease from the thermophilic archaeon *Pyrococcus furiosus* (*PfAgo*) contributes to host defense and represents a promising biotechnology tool. Here, we report the structure of a *PfAgo*-guide DNA-target DNA ternary complex at the cleavage-compatible state. The ternary complex is predominantly dimerized, and the dimerization is solely mediated by *PfAgo* at PIWI-MID, PIWI-PIWI, and PAZ-N interfaces. Additionally, *PfAgo* accommodates a short 14-bp guide-target DNA duplex with a wedge-type N domain and specifically recognizes 5'-phosphorylated guide DNA. In contrast, the *PfAgo*-guide DNA binary complex is monomeric, and the engagement of target DNA with 14-bp complementarity induces sufficient dimerization and activation of *PfAgo*, accompanied by movement of PAZ and N domains. A closely related Argonaute from *Thermococcus thioreducens* adopts a similar dimerization configuration with an additional zinc finger formed at the dimerization interface. Dimerization of both Argonautes stabilizes the catalytic loops, highlighting the important role of Argonaute dimerization in the activation and target cleavage.

## INTRODUCTION

Argonaute proteins are conserved throughout all domains of life.<sup>1</sup> Eukaryotic Argonaute proteins (eAgos) utilize short 5'-phosphorylated guide RNA to target RNA, controlling a wide range of RNA-related cellular processes.<sup>2</sup> In contrast, prokaryotic Argonaute proteins (pAgos) are versatile in choosing guide (5'-P-OH DNA/RNA) and target (DNA/RNA) types.<sup>3,4</sup> Besides the role in host defense,<sup>5,6</sup> pAgos are also involved in a series of cellular activities including DNA replication and DNA processing and repair.<sup>7–9</sup> pAgos are further divided into long pAgos, short pAgos, and PIWI-RE (P-element-induced wimpy testis [PIWI] with conserved R and E residues) proteins.<sup>3,10</sup> Long pAgos and eAgos adopt bilobed architecture,<sup>1</sup> with PAZ lobe (N domain, interdomain linker L1, and PIWI-Argonaute-Zwille [PAZ] domain) connected by interdomain linker L2 to PIWI lobe (Middle [MID] and PIWI domains). In addition, eAgos contain external insertion segments that might contribute to the RNA-induced silencing complex (RISC) assembly.<sup>11</sup> Short pAgos lack the N and PAZ domains but contain additional domains absent in eAgos.<sup>10</sup> The core functions of the four structural domains of Ago proteins are conserved in prokaryotes and eukaryotes.<sup>1</sup>

Extensive biochemical and structural studies have been dedicated to elucidating the working mechanism for guide

and target recognition and activation of long pAgos, including but not restricted to Argonaute from *Thermus thermophilus* (*TtAgo*),<sup>12–15</sup> *Methanocaldococcus jannaschii* (*MjAgo*),<sup>16</sup> *Marinibacteria piezophila* (*MpAgo*),<sup>17,18</sup> *Rhodobacter sphaeroides* (*RsAgo*),<sup>19</sup> *Clostridium butyricum* (*CbAgo*),<sup>20</sup> and *Pseudooceanicola lipolyticus* (*PliAgo*).<sup>21</sup> Generally, MID domain adopts a Rossmann-like fold to bind the 5' end of guide, PIWI domain adopts a RNase H-like fold to accommodate the catalytic tetrad and divalent cations, and PAZ domain adopts an SH3-like barrel fold to bind the 3' end of guide.<sup>1</sup> N domain can be classified into wedge type, which blocks further base pairing between guide and target, and packing type, which stabilizes the extended guide-target duplex.<sup>19</sup> The nucleotide bases from the “seed region” of the guides are pre-oriented in a helical conformation and exposed to the solution, and its pairing with target induces conformational changes for further target recognition.<sup>4,22,23</sup> Upon target loading, N and PAZ domains undergo pivot-like conformational transitions,<sup>12,14</sup> and the release of the 3' end of the guide is accompanied by rotation of the PAZ domain.<sup>14</sup> More importantly, the propagation of guide-target base pairing induces the transition within the nucleic acid-binding surface of the PIWI domain, namely movements in loops L1, L2, and L3.<sup>14,15</sup> Insertion of the glutamic acid from loop L2 (termed Glu finger) in the catalytic pocket results

**Table 1. Cryo-EM data collection, refinement, and validation statistics**

	PfAgo-guide DNA- target DNA (PDB: 8JPX, EMD-36489)	TtdAgo-guide DNA- target DNA (PDB: 8WD8, EMD-37457)
<b>Data collection and processing</b>		
Magnification	105,000	105,000
Voltage (kV)	300	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	54	54
Defocus range (μm)	1.2–2.2	1.2–2.2
Pixel size (Å)	0.85	0.85
Symmetry imposed	C2	C2
Initial particle images (no.)	2,325,621	2,476,975
Final particle images (no.)	621,102	629,194
Map resolution (Å)	2.9	2.9
FSC threshold	0.143	0.143
Map resolution range (Å)	2.8–5.0	2.7–4.5
<b>Refinement</b>		
Initial model used	AlphaFold2	AlphaFold2
Model resolution (Å)	2.9	2.9
FSC threshold	0.5	0.5
Model resolution range (Å)	2.8–5.0	2.7–4.5
Map sharpening B factor (Å <sup>2</sup> )	−162.6	−158.2
<b>Model composition</b>		
Non-hydrogen atoms	14,362	13,667
Protein residues	1,540	1,500
Nucleotides	80	64
Ligands	Mg:6	Zn:1, Mg:6
B factors (Å <sup>2</sup> )		
Protein	61.52	33.08
Nucleotide	45.91	13.42
Ligands	35.64	43.93
RMSDs		
Bond lengths (Å)	0.003	0.005
Bond angles (°)	0.561	1.004
<b>Validation</b>		
MolProbity score	1.84	1.37
Clashscore	6.83	3.54
Poor rotamers (%)	0.42	0.65
<b>Ramachandran plot</b>		
Favored (%)	92.71	96.52
Allowed (%)	7.29	3.48
Disallowed (%)	0.00	0.00

in cleavage of the target strand between the nucleotides that base pair with guide nucleotides 10 and 11.<sup>15</sup>

Argonaute proteins have already found applications in biotechnology for targeted cleavage and detection of nucleic

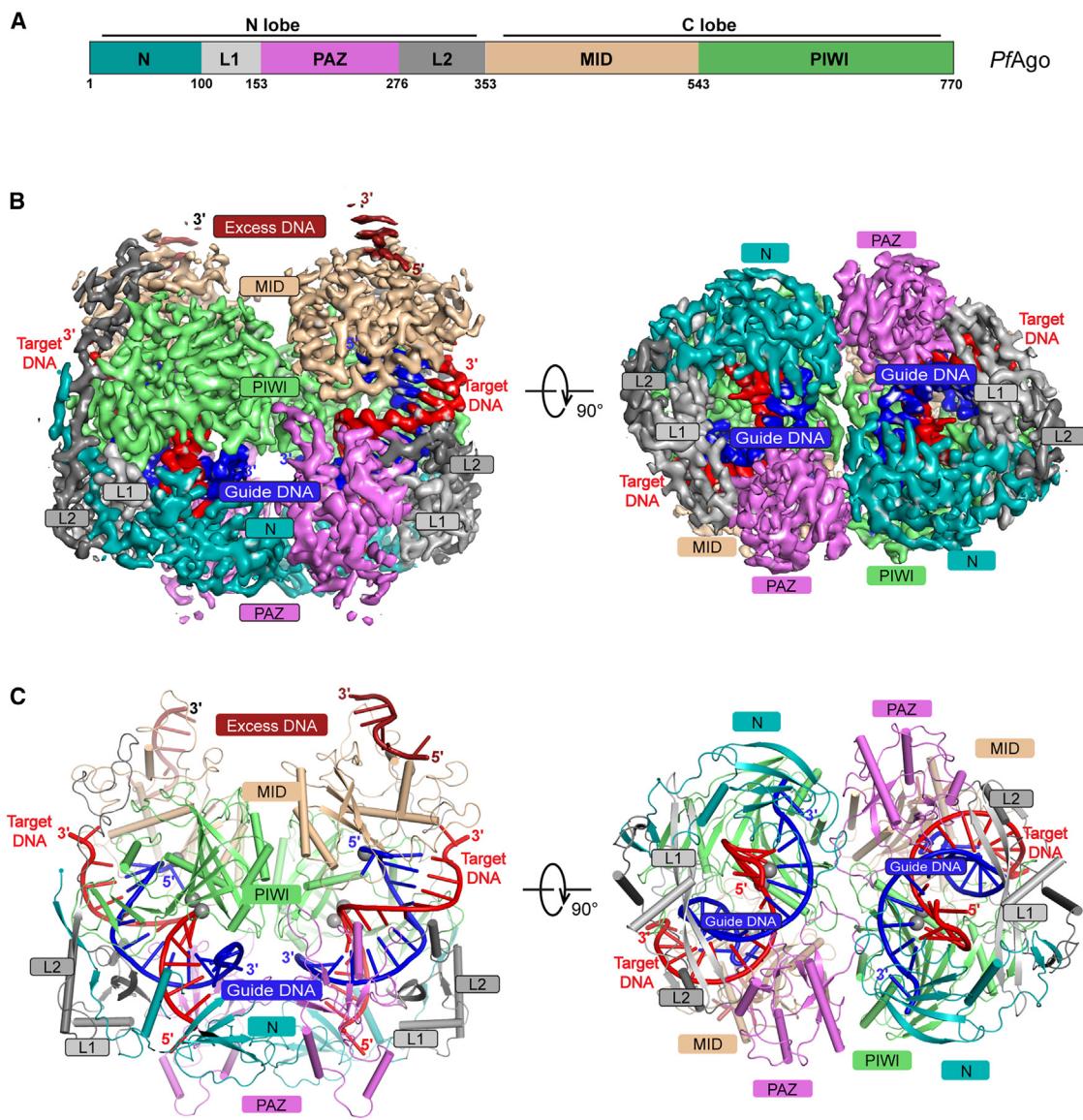
acids.<sup>24</sup> The currently known applications of bacterial Argonaute proteins include detection of nucleic acids using a hairpin structure,<sup>25,26</sup> mutant alleles analysis,<sup>27</sup> rare mutant nucleic acids fraction enrichment for accurate diagnostics,<sup>28</sup> highly sensitive and specific detection of endogenous miRNA,<sup>29</sup> detection of RNA modifications,<sup>30</sup> and RNA secondary structure analysis.<sup>31–33</sup> With fusion to different proteins, catalytically dead Argonaute can be potentially used for imaging, transcriptome editing, and regulation of gene expression and alternative splicing.<sup>24</sup> Moreover, the capability of Argonaute proteins to introduce site-specific cleavage in nucleic acids renders them as potential tools for editing the genome and transcriptome.<sup>24</sup>

The Argonaute nuclease from the thermophilic archaeon *Pyrococcus furiosus* (*PfAgo*) contributes to host defense by interfering with the transformation of plasmid DNA.<sup>34</sup> *PfAgo* utilizes small 5'-phosphorylated DNA guides, without nucleotide preference at the 5' end, to cleave both single- and double-stranded DNA targets and does not utilize RNA as a guide or target.<sup>34</sup> *PfAgo* is most active in the range from 87°C to 99.9°C.<sup>34</sup> Experimental and simulation studies showed that *PfAgo* adopts a loosely packed structure with a partially melted surface at the physiological temperatures.<sup>35</sup> *PfAgo* has been successfully used to assemble genetic constructs<sup>36,37</sup> and for detection of target sequences.<sup>25,26,38–41</sup> Although the structural study of *PfAgo* dates back 20 years ago, the *PfAgo* structure was determined in the absence of guide DNA or target DNA.<sup>42</sup> The underlying mechanism for target recognition and activation remains unknown.

## RESULTS

### Overall structure of PfAgo-guide DNA-target DNA complex

Our previous biochemical characterization indicated that the minimal length of the guide DNA-target DNA duplex for *PfAgo* to exhibit cleavage activity is 15 bp.<sup>26</sup> We assembled the ternary complex by incubating the *PfAgo* with 16-nt 5'-phosphorylated guide DNA (5'-phos-TGAGGTAGTAGGTTGT) and 17-nt target DNA (ACAACCTACTACCTCAT) in the presence of magnesium and determined the structure at 2.9 Å (Table 1), which allows *ab-initio* model building of all the protein residues and magnesium ions around the catalytic site and the 5' end of guide DNA and unambiguous assignment of all the nucleotides (Figure S1). The resulting 2D class average and the cryoelectron microscopy (cryo-EM) map indicated that the ternary complex is predominantly dimerized (~90%, Figure S1B), and the dimerization is solely mediated by *PfAgo* (Figures 1 and S1). Based on the previous research<sup>42</sup> and the secondary structure, the bilobed *PfAgo* is divided into N domain (1–100), interdomain linker L1 (100–153), PAZ domain (153–276), and interdomain linker L2 (276–353), MID domain (353–543), and PIWI domain (543–770). DALI search revealed that the structure of *PfAgo* from the ternary complex resembles those of several long pAgos, including *CbAgo* (root-mean-square deviation [RMSD]: 4.6), *MpAgo* (RMSD: 4.8), *TtAgo* (RMSD: 4.2), and *RsAgo* (RMSD: 3.7). The guide DNA-target DNA duplex is surrounded by L1, PAZ, L2 and PIWI domains, and capped by the MID domain on one end and wedge-type N domain on the other end (Figure 1). Spatial proximity allows sparse interaction between the phosphate groups of the duplex



**Figure 1. Overall structure of *PfAgo*-guide DNA-target DNA ternary complex**

#### (A) Domain architecture of *PfAgo*.

(B and C) Cryo-EM map (B) and atomic model (C) in two views with protein domains colored based on (A), guide DNA colored in blue, target DNA colored in red, magnesium colored in gray, and excess DNA (see also Figure S1l) colored in dark red.

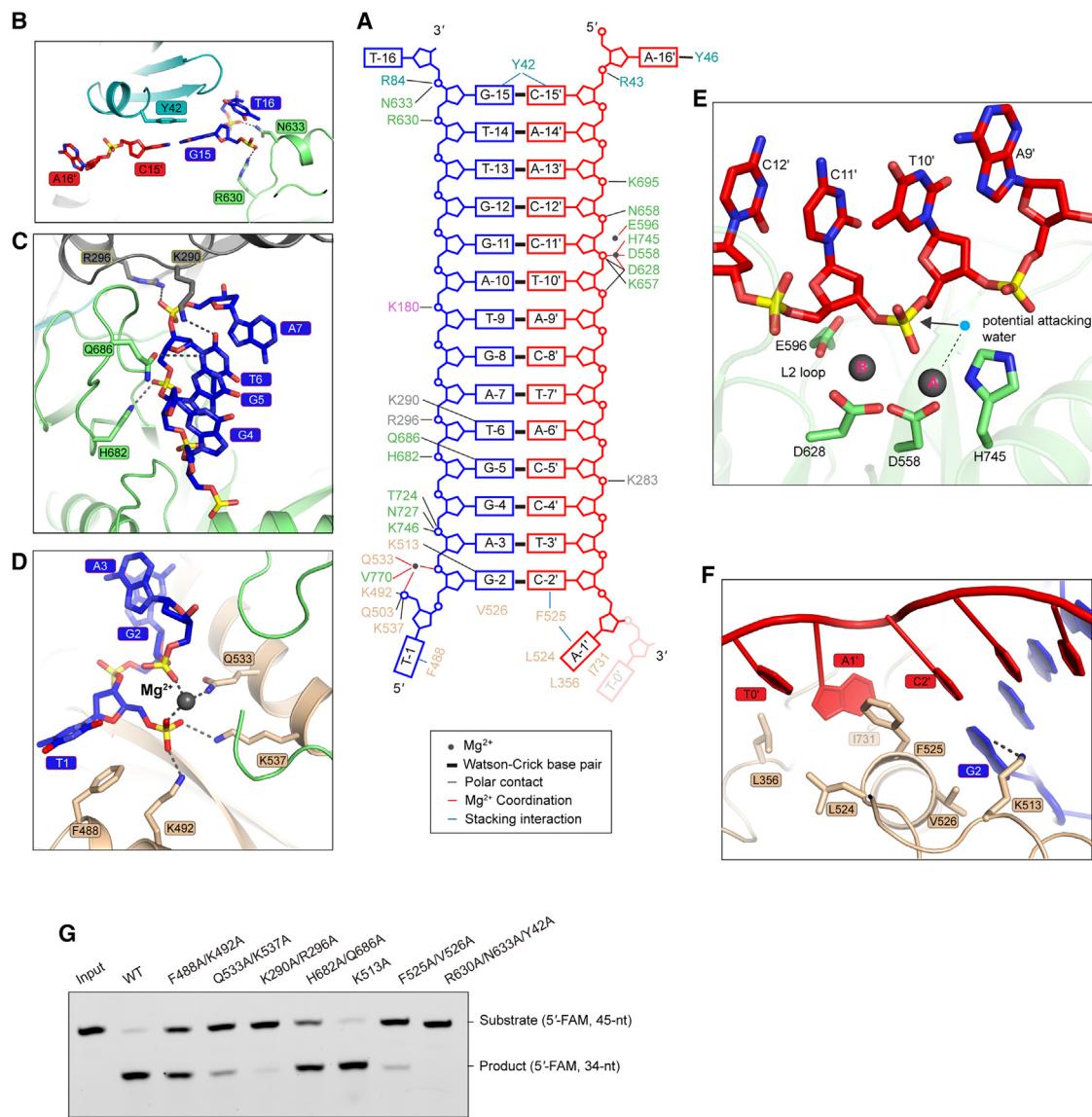
and main chains of the L1 and PAZ domains ([Figures 1](#) and [2A](#)). Notably, *PfAgo* accommodates the shortest guide-target duplex (14 bp), compared with *RsAgo* (17-bp duplex with packing-type N domain), *MpAgo* (19-bp duplex with packing-type N domain), *CbAgo* (15-bp duplex with wedge-type N domain), and *TtAgo* (15-/16-bp duplex with wedge-type N domain) ([Table S1](#)).

## Interaction between *PfAgo* and guide DNA-target DNA duplex

In the structure of the ternary complex, the A-form guide-target DNA duplex only contains 14 base pairs ranging from position 2 to position 15, with T1 and A1' (' denotes the nucleotide from target strand) splayed out by MID domain, and T16 and A16'

splayed out by N domain. Extensive intermolecular interactions were observed between *PfAgo* and the guide-target DNA duplex (**Figure 2A**), especially around the 5'-region of the guide DNA, the terminal nucleotides (T1, A1', T16, and A16') and the scissile phosphate group of target DNA.

The 5'-phosphate group of guide DNA is anchored within the PIWI pocket to form extensive contacts, supporting the observation that *PfAgo* favors 5'-phosphorylated guide DNA for target cleavage. Specifically, a magnesium ion coordinates the 5'-phosphate group, the phosphate group of A3 and Q533 from MID domain (Figure 2D). The 5'-phosphate group also interacts with K492 and K537 from the MID domain (Figure 2D). The nucleobase of T1 is stabilized via stacking against F488 of the



**Figure 2. Interactions between PfAgo and guide DNA-target DNA duplex**

- (A) Schematic of all the possible interactions between PfAgo and guide DNA-target DNA duplex. The nucleotides and amino acids are colored based on Figure 1.
- (B) Interaction details between terminal nucleotides (G15, C15', T16, and A16') and N and PIWI domains.
- (C) Interaction details between the 5'-region of guide DNA (G4, G5, T6, and A7) and the PIWI domain of PfAgo.
- (D) Interaction details between the 5'-phosphorylated end of guide DNA and the MID domain (see also Figure S2B).
- (E) Active site of PfAgo. Two magnesium ions (A and B) are shown as gray sphere (see also Figure S1J). The hypothetical water to perform nucleophilic in-line attack is shown in cyan sphere. The catalytic residues (D558, E596 from loop L2, D628 and H745) and the nucleotides flanking the scissile phosphate group were shown in stick representation. The dashed line indicates the potential hydrogen bond. The arrow indicates the direction of the in-line nucleophilic water attack.
- (F) Interaction details between terminal nucleotides (T0', A1', G2, and C2'), and the MID domain of PfAgo (see also Figure S2C).
- (G) Target DNA cleavage assay with wild-type PfAgo and the mutants of the duplex interacting residues. The 5'-FAM labeled 45-nt DNA was used as substrate, and the PfAgo cleavage generated 34-nt 5'-FAM products. The results shown are the representative of three experiments.

MID domain (Figure 2D). Similar residues (Figure S2B) were also found to stabilize the 5' end nucleotides in *CbAgo* (Y472), *MpAgo* (I383), *TtAgo* (R418), and *RsAgo* (Y463). However, unlike *TtAgo* and *RsAgo*, which encode “specificity loops” to discriminate different terminal nucleotides (Figure S2B), *PfAgo* has no further contact with the nucleobase of T1, explaining why *PfAgo* has no nucleotide preference at the 5' end of the guide. Amino acids that

play similar roles in interacting with the guide-target DNA duplex were grouped for mutagenesis studies (Figure 2G). Alanine substitution of F488/K492 and Q533/K537 impaired the cleavage activity (Figure 2G). The guide DNA segment ranging from G4 to A7 forms polar contact with K290, R296, Q686, and H682 (Figure 2C). Alanine substitution of K290/R296 abolished the target DNA cleavage activity, whereas alanine substitution

# Molecular Cell

## Article



of H682/Q686 mildly impaired the target DNA cleavage activity (**Figure 2G**).

Around the 5' end of guide DNA, the base pairing between guide DNA and target DNA is capped by MID domain, and further base pairing between A1' and T1 is therefore blocked (**Figure 2F**). The terminal nucleotide A1' of target DNA inserts into a hydrophobic pocket formed by L524, L356, F525, and I731 (**Figure 2F**). Similarly, the terminal nucleotide T1' of target DNA inserts into a pocket formed by Y412 and Y507 in *CbAgo*,<sup>20</sup> and the target G1' forms base-specific interaction with MID domain in *TtAgo* (**Figure S2C**), explaining why *in vivo* *TtAgo* is preferentially loaded with guides with a 5'-C.<sup>43,44</sup> Whether the accommodation of the target A1' in *PfAgo* is physiologically relevant, such as generating bias in target acquisition, needs further *in vivo* study. F525 and V526 from MID domain act as “wedges” to cap the base pairing between guide DNA and target DNA, reminiscent of the WED domain of Cas12 nuclease.<sup>45</sup> Alanine substitution of F525/V526 impaired the cleavage activity (**Figure 2G**). At the 5' end of target DNA, the base pairing between guide DNA and target DNA is capped by Y42 from N domain (**Figure 2B**). The phosphate group of the terminal nucleotides T16 is stabilized by R630 and N633 (**Figure 2B**). Alanine substitution of R630/N633/Y42 abolished the cleavage activity (**Figure 2G**).

At the active site, we observed two magnesium ions (termed A and B) around the catalytic residues (H745, D558, D628, and E596) and the scissile phosphate group between T10' and C11' (**Figures 2E** and **S1J**). Mg<sup>2+</sup> A and Mg<sup>2+</sup> B measure a distance of 4.4 Å and form orthogonal geometry with lined catalytic residues and potential coordinating waters, consistent with the configuration of the magnesium ions that mediate cleavage in *TtAgo*.<sup>15</sup> Mg<sup>2+</sup> cation A is coordinated by catalytic residues H745, D558, and D628. The loop (termed L2) containing the fourth catalytic residue E596 (termed Glu finger) points toward the active site and coordinates Mg<sup>2+</sup> cation B. The insertion of the loop L2 that harbors the fourth catalytic residue is considered the hallmark of the activation of target cleavage, and such configuration has been reported in *TtAgo* in which Glu finger E512 within loop L2 inserts into the catalytic pocket to form the catalytic tetrad.<sup>15</sup> Altogether, these observations established that the structure represents the cleavage-compatible state.

### Dimerization of *PfAgo*

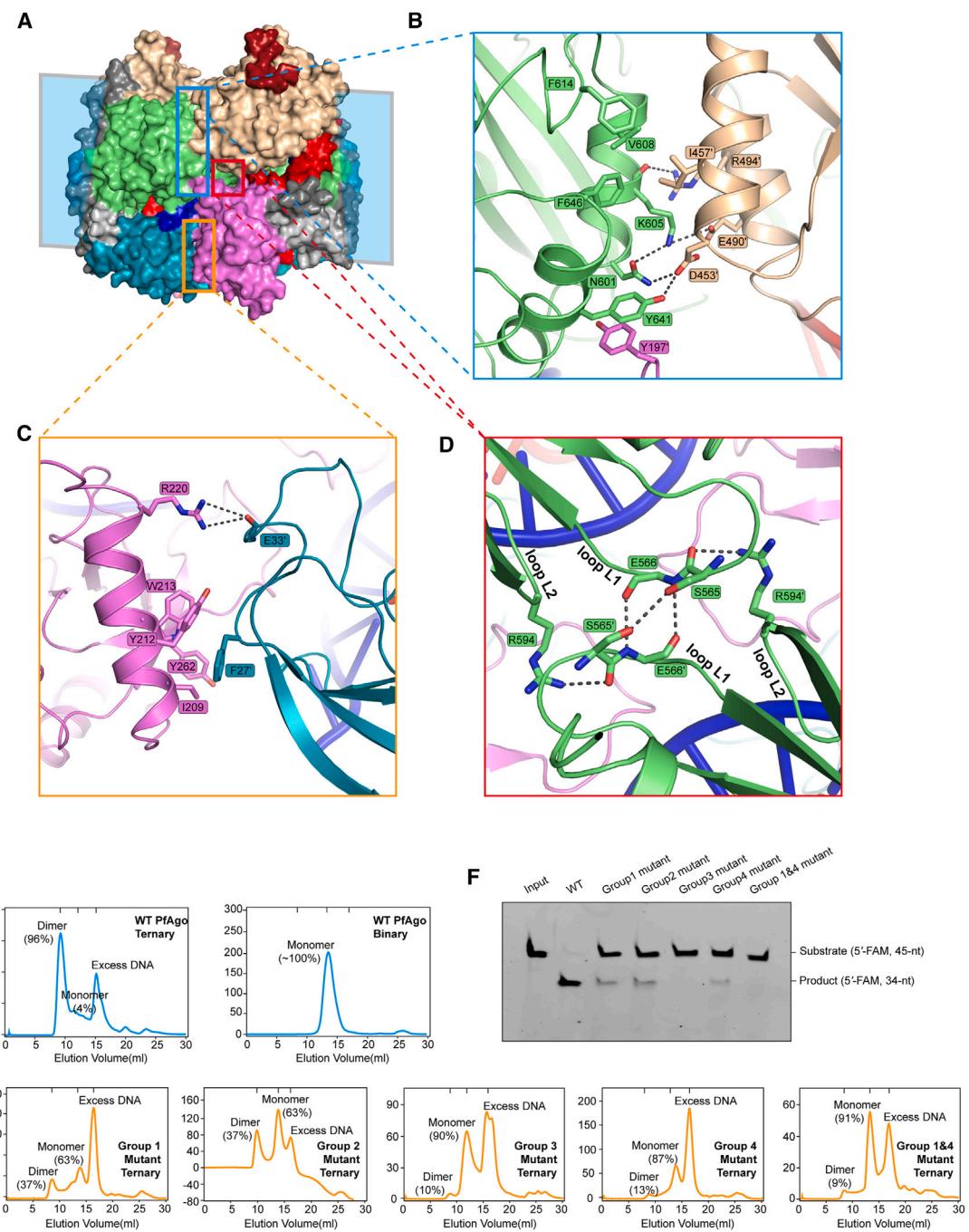
With a large buried solvent accessible area (14,885 Å<sup>2</sup>) on the *PfAgo* dimerization interface, extensive intermolecular contact was observed (**Figure 3**). The interaction for dimerization happens at five distinct patches, including two PIWI-MID interfaces, one PIWI-PIWI interface, and two PAZ-N interfaces (**Figure 3A**). At the PIWI-MID interface (**Figure 3B**), K605, N601, and Y641 form a polar contact network with D453', E490', and R494' (' was used to denote the amino acid from the other *PfAgo* molecule). I457' forms hydrophobic interaction with V608, F614, F646, and K605 (**Figure 3B**). Moreover, Y641 from PIWI domain forms π-π stacking interaction with Y197' from the PAZ domain, possibly stabilizing the relatively flexible PAZ domain (**Figure 3B**). Given the 2-fold symmetry, the same interaction mode was observed in the other PIWI-MID interface. Group alanine substitution of the aforementioned amino acids (group 1 mutation: D453A, I457A, E490A, R494A, N601A, K605A, V608A, F614A,

Y641A, and F646A) impaired the dimerization and cleavage activity, without disrupting the structural integrity (**Figures 3E**, **3F**, and **S4G**), highlighting the importance of these amino acids in the dimerization and target cleavage. At the PAZ-N interface, F27' forms hydrophobic contact with Y212, W213, Y262, and I209, and the amide group of E33' interacts with the side chain of R220 (**Figure 3C**). The same interaction was also observed in the other PAZ-N interface due to symmetry. Group alanine substitution of the aforementioned amino acids (group 4 mutation: F27A, E33A, I209A, Y262A, Y212A, W213A, and R220A) impaired the dimerization without disrupting the structural integrity and impaired the DNA cleavage activity (**Figures 3E**, **3F**, and **S4G**). Furthermore, alanine substitution of the amino acids on PAZ-N and PIWI-MID interface (group 1 and 4 mutations) abolished the dimerization and the DNA cleavage activity without disrupting the structural integrity (**Figures 3E**, **3F**, and **S4G**). Polar contacts that stabilize the catalytic loops L1 and L2 were observed at the PIWI-PIWI interface, including the contact between the main chain of S565' from loop L1 and the side chain of R594 from loop L2, the contact between the side chains of S565 and S565' from loop L1, the contact between the main chains of E566 and E566' from loop L1, and the contact between the side chain of R594' from loop L2 and the main chain of S565 from loop L1 (**Figure 3D**). Given the involvement of main chain contacts on PIWI-PIWI interface, group glycine substitutions (group 2 mutation: P561G, M562G, K563G, R564G, S565G, E566G, and Y568G; group 3 mutation: K589G, I590G, E592G, Q593G, R594G, S597G, and V598G) were introduced to increase the loop flexibility to disrupt any potential local main chain interactions between the catalytic loops. Group 2 and 3 mutations disrupted the dimerization and impaired the cleavage activity without obviously impairing the structural integrity (**Figures 3E**, **3F**, and **S4G**), highlighting the important roles of dimerization-mediated stabilization of catalytic loops in the target cleavage of *PfAgo*.

### Target DNA engagement induces the dimerization of *PfAgo*

In stark contrast, the *PfAgo*-guide DNA binary complex is monomeric (**Figures 3E** and **4A**), indicating that the dimerization of *PfAgo* is induced via target DNA engagement. This observation raised two questions regarding the underlying conformational changes required for *PfAgo* dimerization and the minimal length of target DNA to induce *PfAgo* dimerization. To address these questions, the binary complex, the ternary complex containing 16-nt guide DNA, and target DNA of varying lengths (the complexes with 7-, 9-, 10-, 11-, 12-, and 13-bp complementarity were termed 7-, 9-, 10-, 11-, 12-, and 13-bp complex, respectively) to mimic the intermediate state during duplex propagation, and the post-cleavage complex were subjected to cryo-EM analysis.

As evidenced by cryo-EM raw images and 2D class averages (**Figures 4A** and **S3**), the 7-, 9-, 10-, 11-, and 12-bp complexes are exclusively monomeric, the 13-bp complex shows a small percentage of dimerization (6.8%, 10,138 dimer particles out of 147,912 particles), the 14-bp complex is predominantly dimeric (90%) as mentioned before, and the post-cleavage complex is mixed with dimer (58%) and monomer (42%). These results



**Figure 3. Dimerization of *PfAgo***

(A) Surface representation of *PfAgo*-guide DNA-target DNA ternary complex. The rectangles with different colors represent different dimerization interfaces. The symmetry plane is colored in cyan.

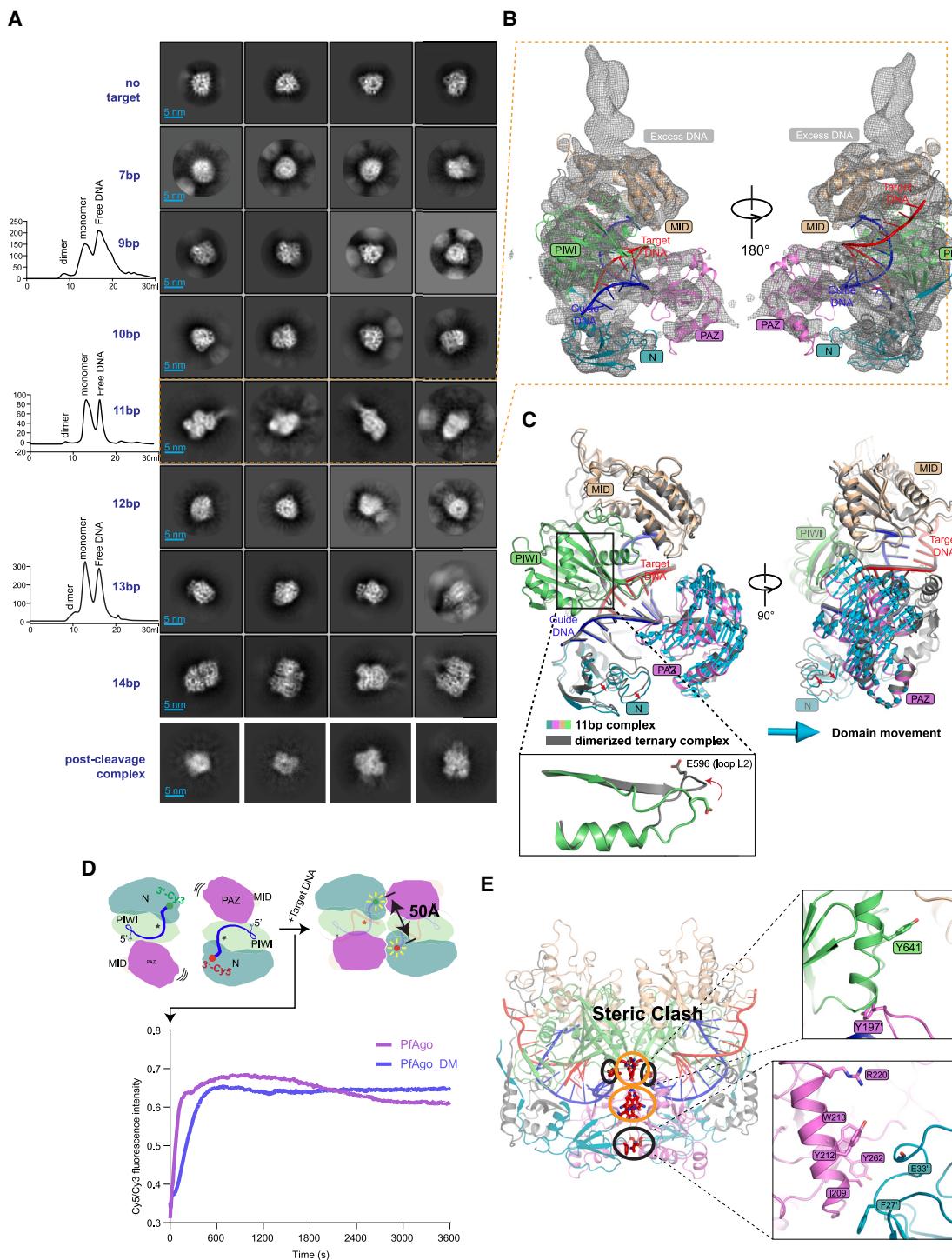
(B-D) Interaction details at PIWI-MID interface (B), PAZ-N interface (C), and PIWI-PIWI interface (D).

(E) Size exclusion chromatography of *PfAgo* binary complex, wild-type *PfAgo* ternary complex, and mutant *PfAgo* ternary complexes (group 1–4 and group 1 and 4 mutations). Group 1 mutation includes D453A, I457A, E490A, R494A, N601A, K605A, V608A, F614A, Y641A, and F646A. Group 2 mutation includes P561G, M562G, K563G, R564G, S565G, E566G, and Y568G. Group 3 mutation includes K589G, I590G, E592G, Q593G, R594G, S597G, and V598G. Group 4 mutation includes F27A, E33A, I209A, Y212A, W213A, R220A, and Y262A. Group 1 and 4 mutation is a combination of group 1 and group 4 mutations. The three vertical lines above the elution profile indicate dimer, monomer, and excess DNA, respectively. The dimer/monomer ratios were calculated based on the area of the respective peaks.

(F) Target DNA cleavage assay with wild-type *PfAgo* and mutant *PfAgo* (group 1–4 and group 1 and 4 mutations). The 5'-FAM labeled 45-nt DNA was used as substrate, and the *PfAgo* cleavage generated 34-nt 5'-FAM products. The results shown are the representative of three experiments.

# Molecular Cell

## Article



**Figure 4. Target DNA induces the dimerization of PfAgo**

(A) Analysis of the binary complex, the ternary complex containing 16-nt guide DNA and target DNA of varying lengths, and the post-cleavage complex via size exclusion chromatography and cryo-EM 2D classification (see also Figure S3).

(B) Cryo-EM map of the 11-bp complex with the individual domain of PfAgo fitted in.

(C) The structural difference between dimerized PfAgo-guide DNA-target DNA complex (colored in gray) and 11-bp complex (color coded based on Figure 1A) shown in two views. The vector length corresponds to the movement distance. The rectangle with close-up view shows the conformational difference of loops L2 and Glu finger E596 (see also Figure S4H).

(legend continued on next page)

indicate that the engagement of target DNA with full complementarity induces sufficient dimerization of *PfAgo*. Furthermore, ensemble fluorescence resonance energy transfer (FRET) experiment indicated the dimer formation upon adding target DNA and dimer dissociation during cleavage, confirming the dimer dissociation of *PfAgo* observed from the cryo-EM analysis (Figure 4D).

The 11-bp complex was determined at 7 Å resolution with high map quality (Figure S4), allowing us to faithfully dock individual domains into the cryo-EM map with an overall cross-correlation of 0.7, and trace the main chains of the catalytic loops (Figures 4B and S4H). Due to intrinsic flexibility, cryo-EM reconstitution of the binary complex and the rest of the ternary complex cannot achieve intermediate resolution to allow resolving secondary structure features (Figure 4A). Structural comparison between the 11-bp complex and the dimerized ternary complex revealed that the biggest movement lies in the PAZ domain, with the N domain to a lesser extent (Figure 4C). In contrast to the dimerized ternary complex in which the loop L2 inserts into the catalytic site to complete the catalytic tetrad (Figure 2E), the loop L2 is unplugged in the 11-bp complex (Figures 4C and S4H), consistent with the observation from *TtAgo*.<sup>15</sup> Superposition of two copies of 11-bp complex molecules into the structure of the dimerized ternary complex caused severe steric clash (Figure 4E). Besides, some of the aforementioned interactions cannot be observed from the superposed structure. Specifically, Y197' of PAZ domain and Y641 of PIWI domain locate too far away to form π-π interaction, and PAZ misaligns with N domain so that the hydrophobic interaction among F27', Y212, W213, Y262, and I209 and the charged interaction between R220 and E33' cannot be formed (Figure 4E), indicating that proper orientation of PAZ domain plays important roles in mediating *PfAgo* dimerization.

#### TtdAgo adopts a similar dimerization configuration

To investigate whether the Argonaute dimerization observed from *Pyrococcus furiosus* is conserved in other pAgos, *Thermococcus thioreducers* Argonaute (*TtdAgo*), with 54.3% sequence identity to *PfAgo* and a similar domain architecture (Figures 5A and S2A), was subjected to cryo-EM and biochemical studies. The 2.9 Å cryo-EM map (Figure S5) allows building models of all the amino acids, nucleotides, and zinc and magnesium ions. Comparison between *TtdAgo* and *PfAgo* ternary complex revealed several similarities and differences. Similarly, *TtdAgo*, with an overall RMSD of 1.384 against *PfAgo*, is also predominantly dimerized (93%, Figure S5C) in the presence of guide DNA and target DNA (Figure 5B). The dimerization interfaces include two PIWI-MID interfaces (Figure 5C), two PAZ-N interfaces (Figure 5D), and one PIWI-PIWI interface (Figure 5E). The interaction modes at PIWI-MID and PAZ-N interfaces of *TtdAgo* are similar to those of *PfAgo*. In contrast, a zinc finger, consisting of two H249 from

PAZ domain, two H546 from loop L1, and the chelating zinc ion, is formed at the dimerization interface (Figure 5E). Alanine substitution of H249 and H546 mildly impaired the target DNA cleavage activity of *TtdAgo* (Figure 5F), highlighting the important roles of dimerization-mediated stabilization of the catalytic loop in the activation and target cleavage of *TtdAgo*.

#### Working model of *PfAgo*

Altogether, these presented data lead to a working model of *PfAgo* (Figure 6). Specifically, during the process of target DNA propagation, the *PfAgo* is monomeric when target DNA pairs with the 5'-region of guide DNA (7-bp complex) and target DNA surpasses the active site (10-bp complex). Upon full-length base pairing between guide DNA and target DNA, the *PfAgo* gets both dimerized and activated, and the dimerization-mediated interaction stabilizes the catalytic loops. The dimerization of *PfAgo* is accompanied by the insertion of the Glu finger in the process of target DNA propagation (Figure 4C). After cleavage, *PfAgo* returns to monomeric as evidenced by the cryo-EM and FRET analysis (Figures 4A and 4D). In summary, *PfAgo* undergoes a “monomer-dimer-monomer” transition during a cleavage cycle.

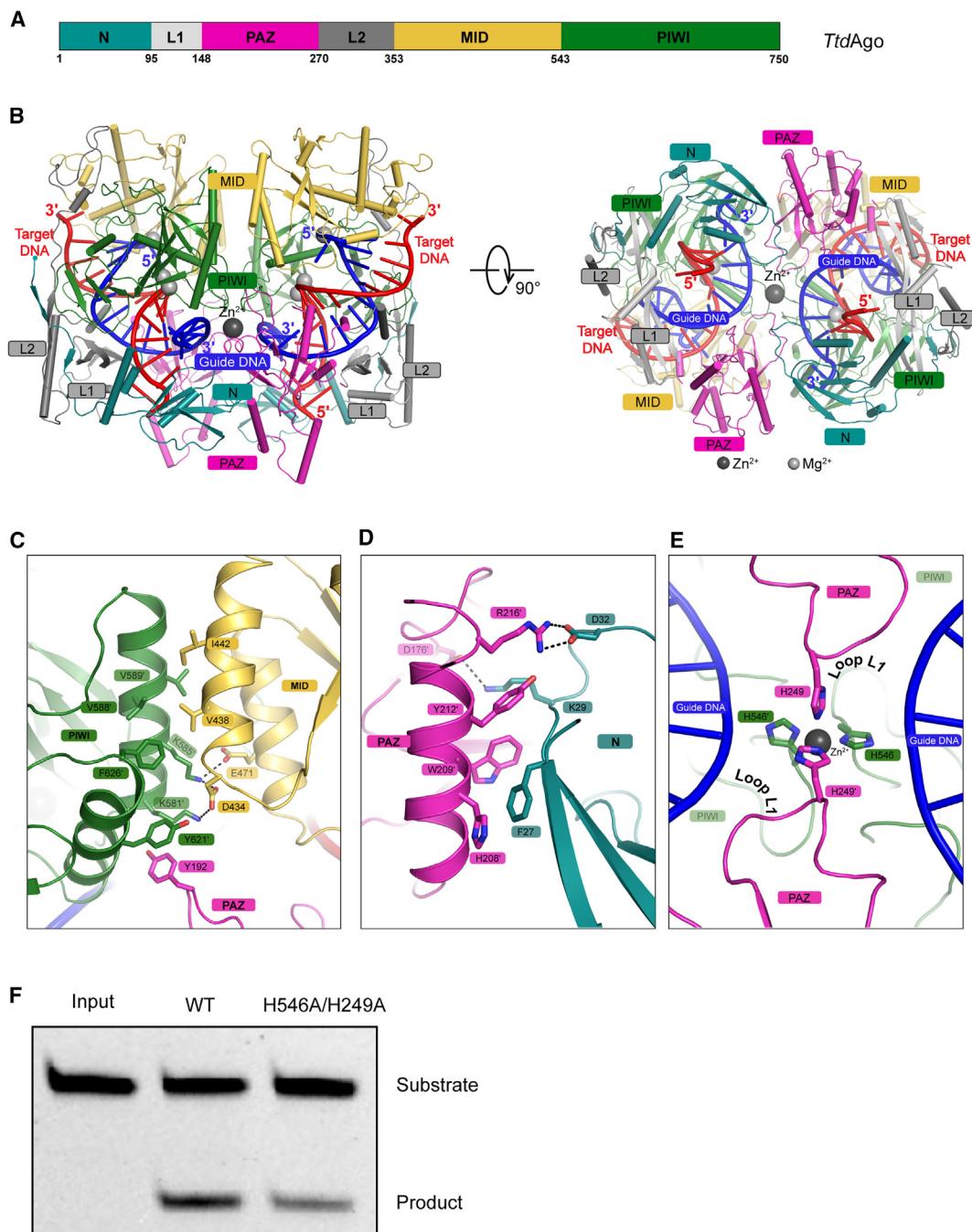
## DISCUSSION

*PfAgo* contributes to host defense by interfering with the uptake of plasmid DNA and represents a promising tool for genetic engineering and nucleic acid detection. The most remarkable finding of this study lies in the observation of target DNA-induced dimerization of *PfAgo*. Besides the *PfAgo* and *TtdAgo*, we found that another Argonaute from thermophilic *Thermogladus calderae* shows dimerization upon target DNA engagement (data not shown), suggesting that dimerization is probably a common feature for Argonaute from thermophilic prokaryotes. Since the multimerization was frequently reported to correlate with thermostability,<sup>46–48</sup> it is tempting to speculate that the dimerization of these pAgos facilitates their physiological function in high living temperature.

With target DNA of different lengths, we found that only target DNA with a length of 14 bp complementarity can induce sufficient dimerization of *PfAgo*. Restricted by the resolution, we can only clarify the conformational changes in the process of *PfAgo* dimerization at the secondary structure level, with the movement of PAZ and N domains playing essential roles. Since the N domain forms π-π and charged interactions with the end nucleotide (G15 and C15'), it would be tempting to speculate that the full-length (14 bp complementarity) guide DNA-target DNA duplex functions as an anchor to stabilize both N and PAZ domains, facilitating the contact formation at the PAZ-N interface. Of all the ternary complexes we have tried

(D) Ensemble FRET experiment to measure the dimer formation upon adding target DNA and dimer dissociation after cleavage. The cartoon model is used to indicate the spatial proximity between Cy3 and Cy5. *PfAgo\_DM*, D558A, and D628A. The ratio of Cy5 and Cy3 fluorescence intensity is plotted to time. The FRET experiment is performed at 60°C, which is the maximum temperature allowed by the fluorescence spectrophotometer and much lower than the optimum temperature for *PfAgo* cleavage (80°C–99°C).

(E) Superposition of two copies of 11-bp complex structure into the dimerized *PfAgo*-guide DNA-target DNA structure. The amino acids that cause steric clash are shown in stick representation and enclosed with orange circles. The amino acids with missing interaction are enclosed with black circles and shown with close-up views.



**Figure 5. Structure of *TtdAgo*-guide DNA-target DNA ternary complex**

(A) Domain architecture of *Thermococcus thioreducens* Argonaute (*TtdAgo*).

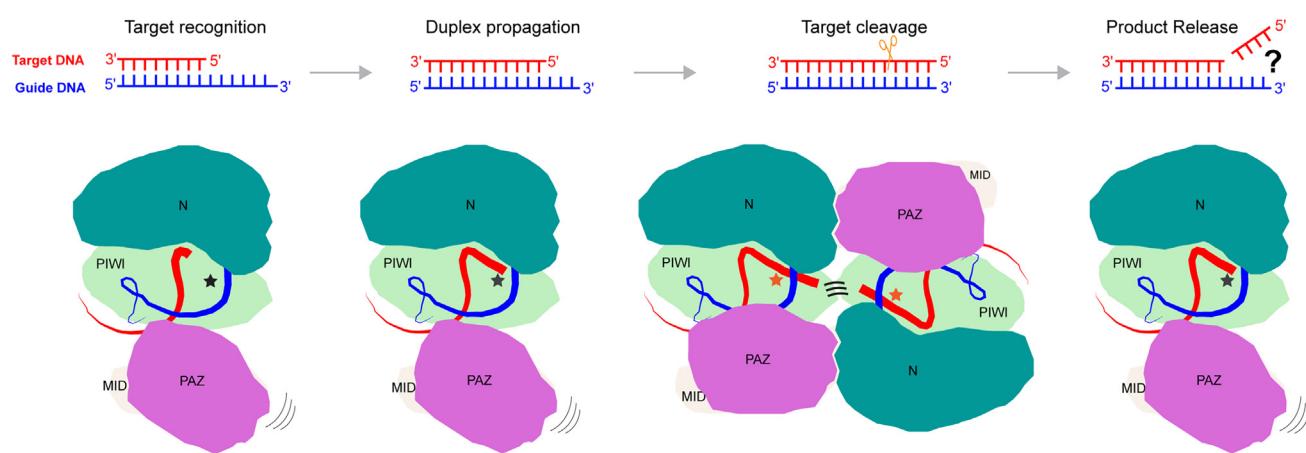
(B) Cartoon representation of *TtdAgo*-guide DNA-target DNA atomic model in two views, with protein domains colored based on (A), the guide DNA colored in blue, the target DNA colored in red, the magnesium ion colored in gray, and the zinc ion colored in dark gray.

(C–E) Interaction details on PIWI-MID (C), PAZ-N (D), and PIWI-PIWI (E) interfaces.

(F) Target DNA cleavage assay with wild-type *TtdAgo* and mutant *TtdAgo*. The 5'-FAM labeled 45-nt DNA was used as a substrate, and the *TtdAgo* cleavage generated the 34-nt 5'-FAM product. The results shown are the representative of three experiments.

with cryo-EM analysis, only the 11-bp complex can achieve intermediate resolution. Rather than the 11-bp complex being more stable than the rest, this distinction should be attributed

to the presence of excess DNA protruding from the MID pocket, which might help the particle alignment during the cryo-EM reconstruction.



**Figure 6. A working model for the dimerization and activation of PfAgo**

Guide DNA is colored in blue, target DNA is colored in red, and the protein is colored based on Figure 1A. The cartoon of the ternary complex is drawn in the same view angle of the right panel of Figure 1B. The curved lines in the dimerized cartoon between PIWI domain are used to represent the stabilizing interaction between catalytic loops. The pentagrams denote the active site, and gray and orange color indicate inactive and cleavage-compatible states, respectively. The curved lines around PAZ domain indicate its flexibility. The question mark indicates that the experimental evidence to determine whether the 5'-region of target DNA is first released after cleavage in PfAgo is still lacking.

An interesting finding from the PfAgo and *TtdAgo* ternary complex is the short guide-target duplex. It is not surprising that the duplex length is longer for *MpAgo* and *RsAgo* that encode packing-type N domains. Given the fact that the guide DNA-target DNA duplex of PfAgo, *TtdAgo*, *CbAgo*, and *TtAgo* are all A form with the same helix rise per base pair, the difference of duplex length should be attributed to the length of the duplex-accommodating channels capped by MID and N domains. The physiological significance of accommodating duplexes of different lengths needs further study.

A surprising discovery lies in the fact that we found excess DNA in an unexpected nucleotide-binding pocket of the MID domain in PfAgo, but not in *TtdAgo* (Figures 1B, 4B, and S1I). To our knowledge, this is the first time the MID domain has been found to contain a second nucleotide-binding pocket; however, we cannot understand the physiological role of this pocket. Given that the 5' end of the excess DNA is in close proximity to the 3' end of target DNA, it would be tempting to speculate that it acts as a loading site for endogenous target DNA that gets stabilized by the MID domain before being scanned by guide DNA.

More importantly, in both PfAgo and *TtdAgo*, the dimerization-mediated stabilization of catalytic loops is important for the activation. Altogether, these *in vitro* data justified the importance of PfAgo and *TtdAgo* dimerization in their activation and target cleavage, and the physiological significance of the dimerization needs delicate *in vivo* study in the future.

#### Limitations of the study

First, hindered by intrinsic flexibility, we failed to determine the structures of PfAgo binary complex, PfAgo ternary complex with short target DNA, and post-cleavage complex. The 11-bp complex was determined at a medium resolution, only allowing us to analyze the conformational change caused by target engagement at the secondary structure level. Second, although

we observed that the dimerization of PfAgo is accompanied by the insertion of the Glu finger in the process of target DNA propagation, whether there is a causal relationship between these two events remains unknown and needs further delicate study. Third, we cannot understand the physiological role of the second nucleotide-binding pocket found in the MID domain of PfAgo. Fourth, although we observed the dimer dissociation of PfAgo after cleavage, the experimental evidence to determine whether 5'-region of the target DNA is first released after cleavage is still lacking.

#### STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
- **METHOD DETAILS**
  - Molecular cloning
  - Protein expression and purification
  - Complex assembly
  - Circular dichroism spectroscopy
  - Target DNA cleavage assay
  - Ensemble fluorescence resonance energy transfer assay
  - Electron microscopy
  - Image processing
  - Model building, refinement, and visualization
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

# Molecular Cell

## Article



### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.molcel.2024.01.004>.

### ACKNOWLEDGMENTS

We thank the staff members from Cryo-Electron Microscopy Facility of Hubei University for help on data collection and computation. We thank the reviewers for the thoughtful comments and detailed suggestions on the paper. This work was supported by China National Key Research and Development (R&D) Program (2021YFC2100100), the National Natural Science Foundation of China (32201004 to Z.L. and 21708006 to W.M.), the Hubei Provincial Natural Science Foundation (ZRMS2022000096), the Key Science and Technology Innovation Project of Hubei Province (2021BAD001), and the Scientific Research Program of Hubei Provincial Department of Education.

### AUTHOR CONTRIBUTIONS

L.M. and Z.L. supervised the study. L.W., W.C., X.X., F.H., and M.C. prepared the samples. C.Z. and Z.L. collected and processed the cryo-EM data and built the atomic model. L.W., W.C., X.X., F.H., M.C., W.M., N.Y., and Q.W. performed the biochemical analysis. Z.L., L.W., W.C., and C.Z. prepared the manuscript and figures with input from L.M., X.X., F.H., and M.C. All authors analyzed the data.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 3, 2023

Revised: November 15, 2023

Accepted: January 5, 2024

Published: January 30, 2024

### REFERENCES

1. 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.
2. Ketting, R.F. (2011). The many faces of RNAi. *Dev. Cell* 20, 148–161.
3. Hegge, J.W., Swarts, D.C., and van der Oost, J. (2018). Prokaryotic Argonaute proteins: novel genome-editing tools? *Nat. Rev. Microbiol.* 16, 5–11.
4. Lisitskaya, L., Aravin, A.A., and Kulbachinskiy, A. (2018). DNA interference and beyond: structure and functions of prokaryotic Argonaute proteins. *Nat. Commun.* 9, 5165.
5. 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.
6. 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.
7. Olina, A., Kuzmenko, A., Ninova, M., Aravin, A.A., Kulbachinskiy, A., and Esyunina, D. (2020). Genome-wide DNA sampling by Ago nuclease from the cyanobacterium *Synechococcus elongatus*. *RNA Biol.* 17, 677–688.
8. Jolly, S.M., Gainetdinov, I., Jouravleva, K., Zhang, H., Strittmatter, L., Bailey, S.M., Hendricks, G.M., Dhabaria, A., Ueberheide, B., and Zamore, P.D. (2020). *Thermus thermophilus* Argonaute functions in the completion of DNA replication. *Cell* 182, 1545–1559.e18.
9. Kuzmenko, A., Ogienko, A., Esyunina, D., Yudin, D., Petrova, M., Kudinova, A., Maslova, O., Ninova, M., Ryazansky, S., Leach, D., et al. (2020). DNA targeting and interference by a bacterial Argonaute nuclease. *Nature* 587, 632–637.
10. Ryazansky, S., Kulbachinskiy, A., and Aravin, A.A. (2018). The Expanded Universe of Prokaryotic Argonaute Proteins. *mBio* 9, e01935–e01918.
11. Nakanishi, K., Weinberg, D.E., Bartel, D.P., and Patel, D.J. (2012). Structure of yeast Argonaute with guide RNA. *Nature* 486, 368–374.
12. Wang, Y., Sheng, G., Juranek, S., Tuschl, T., and Patel, D.J. (2008). Structure of the guide-strand-containing argonaute silencing complex. *Nature* 456, 209–213.
13. Wang, Y., Juranek, S., Li, H., 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, 921–926.
14. Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G.S., Tuschl, T., and Patel, D.J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* 461, 754–761.
15. 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. USA* 111, 652–657.
16. Willkomm, S., Oellig, C.A., Zander, A., Restle, T., Keegan, R., Grohmann, D., and Schneider, S. (2017). Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein. *Nat. Microbiol.* 2, 17035.
17. Doxzen, K.W., and Doudna, J.A. (2017). DNA recognition by an RNA-guided bacterial Argonaute. *PLoS One* 12, e0177097.
18. Kaya, E., Doxzen, K.W., Knoll, K.R., Wilson, R.C., Strutt, S.C., Kranzusch, P.J., and Doudna, J.A. (2016). A bacterial Argonaute with noncanonical guide RNA specificity. *Proc. Natl. Acad. Sci. USA* 113, 4057–4062.
19. Miyoshi, T., Ito, K., Murakami, R., and Uchiimi, T. (2016). Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat. Commun.* 7, 11846.
20. Hegge, J.W., Swarts, D.C., Chandradoss, S.D., Cui, T.J., Kneppers, J., Jinek, M., Joo, C., and van der Oost, J. (2019). DNA-guided DNA cleavage at moderate temperatures by *Clostridium butyricum* Argonaute. *Nucleic Acids Res.* 47, 5809–5821.
21. Lisitskaya, L., Shin, Y., Agapov, A., Olina, A., Kropocheva, E., Ryazansky, S., Aravin, A.A., Esyunina, D., Murakami, K.S., and Kulbachinskiy, A. (2022). Programmable RNA targeting by bacterial Argonaute nucleases with unconventional guide binding and cleavage specificity. *Nat. Commun.* 13, 4624.
22. Grosswendt, S., Filipchyk, A., Manzano, M., Kironomos, F., Schilling, M., Herzog, M., Gottwein, E., and Rajewsky, N. (2014). Unambiguous identification of miRNA:target site interactions by different types of ligation reactions. *Mol. Cell* 54, 1042–1054.
23. Filipowicz, W., Bhattacharyya, S.N., and Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* 9, 102–114.
24. Kropocheva, E.V., Lisitskaya, L.A., Agapov, A.A., Musabirov, A.A., Kulbachinskiy, A.V., and Esyunina, D.M. (2022). Prokaryotic Argonaute Proteins as a Tool for Biotechnology. *Mol. Biol.* 56, 854–873.
25. He, R., Wang, L., Wang, F., Li, W., Liu, Y., Li, A., Wang, Y., Mao, W., Zhai, C., and Ma, L. (2019). *Pyrococcus furiosus* Argonaute-mediated nucleic acid detection. *Chem. Commun. (Camb)* 55, 13219–13222.
26. Wang, L., He, R., Lv, B., Yu, X., Liu, Y., Yang, J., Li, W., Wang, Y., Zhang, H., Yan, G., et al. (2021). *Pyrococcus furiosus* Argonaute coupled with modified ligase chain reaction for detection of SARS-CoV-2 and HPV. *Talanta* 227, 122154.
27. Xiao, G., Fu, X., Zhang, J., Liu, S., Wang, Z., Ye, T., and Zhang, G. (2021). Rapid and cost-effective screening of CRISPR/Cas9-induced mutants by DNA-guided Argonaute nuclease. *Biotechnol. Lett.* 43, 2105–2110.
28. Song, J., Hegge, J.W., Mauk, M.G., Chen, J., Till, J.E., Bhagwat, N., Azink, L.T., Peng, J., Sen, M., Mays, J., et al. (2020). Highly specific enrichment of rare nucleic acid fractions using *Thermus thermophilus* argonaute with applications in cancer diagnostics. *Nucleic Acids Res.* 48, e19.

29. Shin, S., Jung, Y., Uhm, H., Song, M., Son, S., Goo, J., Jeong, C., Song, J.J., Kim, V.N., and Hohng, S. (2020). Quantification of purified endogenous miRNAs with high sensitivity and specificity. *Nat. Commun.* **11**, 6033.
30. Lapinaite, A., Doudna, J.A., and Cate, J.H.D. (2018). Programmable RNA recognition using a CRISPR-associated Argonaute. *Proc. Natl. Acad. Sci. USA* **115**, 3368–3373.
31. Kropocheva, E., Kuzmenko, A., Aravin, A.A., Esvunina, D., and Kulbachinskiy, A. (2021). A programmable pAgO nuclease with universal guide and target specificity from the mesophilic bacterium Kurthia massiliensis. *Nucleic Acids Res.* **49**, 4054–4065.
32. Liu, Y., Li, W., Jiang, X., Wang, Y., Zhang, Z., Liu, Q., He, R., Chen, Q., Yang, J., Wang, L., et al. (2021). A programmable omnipotent Argonaute nuclease from mesophilic bacteria Kurthia massiliensis. *Nucleic Acids Res.* **49**, 1597–1608.
33. Dayeh, D.M., Cantara, W.A., Kitzrow, J.P., Musier-Forsyth, K., and Nakanishi, K. (2018). Argonaute-based programmable RNase as a tool for cleavage of highly-structured RNA. *Nucleic Acids Res.* **46**, e98.
34. Swarts, D.C., Hegge, J.W., Hinojo, I., Shiiomori, M., Ellis, M.A., Dumrongkulraksa, J., Terns, R.M., Terns, M.P., and van der Oost, J. (2015). Argonaute of the archaeon *Pyrococcus furiosus* is a DNA-guided nuclease that targets cognate DNA. *Nucleic Acids Res.* **43**, 5120–5129.
35. Zheng, L., Lu, H., Zan, B., Li, S., Liu, H., Liu, Z., Huang, J., Liu, Y., Jiang, F., Liu, Q., et al. (2022). Loosely-packed dynamical structures with partially-melted surface being the key for thermophilic argonaute proteins achieving high DNA-cleavage activity. *Nucleic Acids Res.* **50**, 7529–7544.
36. Enghiad, B., and Zhao, H. (2017). Programmable DNA-Guided Artificial Restriction Enzymes. *ACS Synth. Biol.* **6**, 752–757.
37. Enghiad, B., Xue, P., Singh, N., Boob, A.G., Shi, C., Petrov, V.A., Liu, R., Peri, S.S., Lane, S.T., Gaither, E.D., and Zhao, H. (2022). PlasmidMaker is a versatile, automated, and high throughput end-to-end platform for plasmid construction. *Nat. Commun.* **13**, 2697.
38. Wang, F., Yang, J., He, R., Yu, X., Chen, S., Liu, Y., Wang, L., Li, A., Liu, L., Zhai, C., and Ma, L. (2021). PfAgO-based detection of SARS-CoV-2. Biosens. *Bioelectron.* **177**, 112932.
39. Liu, Q., Guo, X., Xun, G., Li, Z., Chong, Y., Yang, L., Wang, H., Zhang, F., Luo, S., Cui, L., et al. (2021). Argonaute integrated single-tube PCR system enables supersensitive detection of rare mutations. *Nucleic Acids Res.* **49**, e75.
40. Xun, G., Lane, S.T., Petrov, V.A., Pepa, B.E., and Zhao, H. (2021). A rapid, accurate, scalable, and portable testing system for COVID-19 diagnosis. *Nat. Commun.* **12**, 2905.
41. Li, Y., Kou, J., Han, X., Qiao, J., Zhang, W., Man, S., and Ma, L. (2023). Argonaute-triggered visual and rebuilding-free foodborne pathogenic bacteria detection. *J. Hazard. Mater.* **454**, 131485.
42. 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.
43. Swarts, D.C., Szczepaniak, M., Sheng, G., Chandradoss, S.D., Zhu, Y., Timmers, E.M., Zhang, Y., Zhao, H., Lou, J., Wang, Y., et al. (2017). Autonomous Generation and Loading of DNA Guides by Bacterial Argonaute. *Mol. Cell* **65**, 985–998.e6.
44. 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.J., and van der Oost, J. (2014). DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **507**, 258–261.
45. Yamano, T., Nishimatsu, H., Zetsche, B., Hirano, H., Slaymaker, I.M., Li, Y., Fedorova, I., Nakane, T., Makarova, K.S., and Koonin, E.V. (2016). Crystal structure of Cpf1 in complex with guide RNA and target DNA. *Cell* **165**, 949–962.
46. Zeng, H., Zhu, G., Zhang, S., Li, X., Martin, J., Morgner, N., Sun, F., Peng, G., Xie, H., and Michel, H. (2020). Isolated Heme A Synthase from *Aquifex aeolicus* Is a Trimer. *mBio* **11**, e02615–e02619.
47. Zhang, L., Li, L., Yan, L., Ming, Z., Jia, Z., Lou, Z., and Rao, Z. (2018). Structural and biochemical characterization of endoribonuclease Nsp15 encoded by middle east respiratory syndrome coronavirus. *J. Virol.* **92**, e00893–18.
48. Fredslund, F., Otten, H., Gemperlein, S., Poulsen, J.C., Carius, Y., Kohring, G.W., and Lo Leggio, L. (2016). Structural characterization of the thermo-stable *Bradyrhizobium japonicum*D-sorbitol dehydrogenase. *Acta Crystallogr. F Struct. Biol. Commun.* **72**, 846–852.
49. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332.
50. Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12.
51. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166.
52. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296.
53. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.
54. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501.
55. Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczki, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877.
56. Thompson, R.F., Iadanza, M.G., Hesketh, E.L., Rawson, S., and Ranson, N.A. (2019). Collection, pre-processing and on-the-fly analysis of data for high-resolution, single-particle cryo-electron microscopy. *Nat. Protoc.* **14**, 100–118.
57. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.

**STAR★METHODS**

**KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and virus strains</b>		
<i>Escherichia coli</i> strain BL21(DE3)	Tsingke Biotechnology Co., Ltd.	TSC-E01
<i>Escherichia coli</i> strain DH5 $\alpha$	TransGen Biotech	CD201-01
<b>Chemicals, peptides, and recombinant proteins</b>		
NaCl	Sinopharm Chemical ReagentCo., Ltd	7647-14-5
Tryptone	OXOID	LP0042B
Yeast extract	OXOID	LP0021B
kanamycin	Solarbio	25389-94-0
isopropyl $\beta$ -D-1-thiogalactopyranoside	BioFroxx	367-93-1
protease inhibitor cocktail tablet	Roche	11873580001
Tris-base	Biosharp	BS083
HEPES	Biosharp	7365-45-9
Glycerol	Sinopharm Chemical ReagentCo., Ltd	56-81-5
DTT	Solarbio	3483-12-3
MnCl <sub>2</sub>	Sinopharm Chemical ReagentCo., Ltd	13446-34-9
MgCl <sub>2</sub>	Sinopharm Chemical ReagentCo., Ltd	7791-18-6
formamide	Macklin	75-12-7
EDTA	Biosharp	60-00-4
SDS	BioFroxx	151-21-3
bromophenol blue	Amresco	115-39-9
T5 Exonuclease	NEB	M0663S
DpnI	NEB	R0176S
PrimeSTAR® Max DNA Polymerase	Takara	R045A
Cutsmart buffer	NEB	B7204
Buffer 4	NEB	B7004S
PfAgo	Song et al. <sup>42</sup>	N/A
PfAgo, various mutants	This paper	N/A
TtdAgo	This paper	N/A
TtdAgo mutant	This paper	N/A
<b>Deposited data</b>		
PfAgo ternary complex coordinates	This paper	PDB: 8JPX
PfAgo ternary complex map	This paper	EMD-36489
TtdAgo ternary complex coordinates	This paper	PDB: 8WD8
TtdAgo ternary complex map	This paper	EMD-37457
Unprocessed images	This paper	<a href="https://doi.org/10.17632/pjgkgbvrfr.1">https://doi.org/10.17632/pjgkgbvrfr.1</a>
<b>Oligonucleotides</b>		
16 nt-guide DNA	5'P/TGAGGTAGTAGGTTGT3'	N/A
3'-Cy3 16 nt-guide DNA	5'P/TGAGGTAGTAGGTTGT-Cy3-3'	N/A
3'-Cy5 16 nt-guide DNA	5'P/TGAGGTAGTAGGTTGT-Cy5-3'	N/A
17 nt-target DNA	5'ACAAACCTACTACCTCAT3'	N/A
16 nt-target DNA	5'CAACCTACTACCTCAT3'	N/A
15 nt-target DNA	5'AACCTACTACCTCAT3'	N/A
14 nt-target DNA	5'ACCTACTACCTCAT3'	N/A
13 nt-target DNA	5'CCTACTACCTCAT3'	N/A
12 nt-target DNA	5'CTACTACCTCAT3'	N/A

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
11 nt-target DNA	5'TACTACCTCAT3'	N/A
9 nt-target DNA	5'CTACCTCAT3'	N/A
45 nt-target DNA	5'FAM/AAACGACGCCAGTGCAA GCTTACTATACAACCTACTACCTCTT3'	N/A
<b>Recombinant DNA</b>		
Plasmid pET28a-PfAgo	This paper	N/A
Plasmid pET28a-PfAgo_DM	This paper	N/A
Plasmid pET28a-PfAgo_F488A/K492A	This paper	N/A
Plasmid pET28a-PfAgo_Q533A/K537A	This paper	N/A
Plasmid pET28a-PfAgo_K290A/R296A	This paper	N/A
Plasmid pET28a-PfAgo_H682A/Q686A	This paper	N/A
Plasmid pET28a-PfAgo_K513A	This paper	N/A
Plasmid pET28a-PfAgo_F525A/V526A	This paper	N/A
Plasmid pET28a-PfAgo_R630A/N633A/Y42A	This paper	N/A
Plasmid pET28a-PfAgo_Group1	This paper	N/A
Plasmid pET28a-PfAgo_Group2	This paper	N/A
Plasmid pET28a-PfAgo_Group3	This paper	N/A
Plasmid pET28a-PfAgo_Group4	This paper	N/A
Plasmid pET28a-TtdAgo	This paper	N/A
Plasmid pET28a-TtdAgo_H546A/H249A	This paper	N/A
<b>Software and algorithms</b>		
PyMOL	Schrödinger	<a href="https://pymol.org">https://pymol.org</a>
Image J	NIH	N/A
Prism 8	GraphPad Software LLC	<a href="https://www.graphpad.com/">https://www.graphpad.com/</a>
MotionCor2	Zheng et al. <sup>49</sup>	<a href="https://emcore.ucsf.edu/ucsf-software">https://emcore.ucsf.edu/ucsf-software</a>
Gctf	Zhang <sup>50</sup>	<a href="https://sbgrid.org/software/titles/gctf">https://sbgrid.org/software/titles/gctf</a>
Relion	Zivanov et al. <sup>51</sup>	<a href="https://www3.mrc-lmb.cam.ac.uk/relion/index.php?title&gt;Main_Page">https://www3.mrc-lmb.cam.ac.uk/relion/index.php?title&gt;Main_Page</a>
CryoSPARC	Punjani et al. <sup>52</sup>	<a href="https://cryosparc.com/">https://cryosparc.com/</a>
UCSF Chimera	Pettersen et al. <sup>53</sup>	<a href="https://www.cgl.ucsf.edu/chimera/">https://www.cgl.ucsf.edu/chimera/</a>
Coot	Emsley et al. <sup>54</sup>	<a href="https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/">https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/</a>
Phenix	Liebschner et al. <sup>55</sup>	<a href="https://www.phenix-online.org/">https://www.phenix-online.org/</a>
<b>Other</b>		
Ni-charged Resin	BIO-RAD	Cat#1560131
Superdex 200 Increase 10/300	GE Healthcare	28990944

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Zhuang Li ([zhuangli@hubu.edu.cn](mailto:zhuangli@hubu.edu.cn)).

### Materials availability

All unique/stable reagents generated in this study are available from the [lead contact](#) with a completed Materials Transfer Agreement.

### Data and code availability

- Cryo-EM reconstruction of *PfAgo*-guide DNA-target DNA and *TtdAgo*-guide DNA-target DNA have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD-36489 and EMD-37457 respectively. Coordinates for atomic models of *PfAgo*-guide DNA-target DNA and *TtdAgo*-guide DNA-target DNA have been deposited in the Protein Data Bank

under the accession numbers 8JPX and 8WD8 respectively. The data of unprocessed image files have been deposited in the Mendeley Data repository (<https://doi.org/10.17632/pjgkgbvfr.1>). The deposited data are publicly accessible as the date of publication.

- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

*Escherichia coli* cells were cultured at 37 °C in LB medium (containing 50 mg/l kanamycin) for plasmid and protein preparation.

## METHOD DETAILS

### Molecular cloning

The *PfAgo/TtdAgo* gene was derived from the plasmid pET23a-*PfAgo/TtdAgo* constructed in our laboratory. The gene fragment of *PfAgo/TtdAgo* was obtained by PCR and loaded into the pET28a expression vector to obtain the pET28a-*PfAgo/TtdAgo* expression vector in frame with the N-terminal His tag. For the construction of *PfAgo/TtdAgo* mutants, the primers were designed according to the mutation sites. The recombinant plasmids were constructed by DNA cloning method based on T5 exonuclease I and low temperature operation.

### Protein expression and purification

*Escherichia coli* strain BL21(DE3) was used as a host for expression of *PfAgo/TtdAgo*. Single colony was selected and cultured in 5 mL LK liquid medium overnight at 37°C in a shaking bed, and then transferred to 1 L TB medium containing 50 mg/mL Kana at a ratio of 1:5. When OD<sub>600</sub> value was 0.6-0.8, IPTG with final concentration of 0.5 mM was added to induce protein expression at 18°C for 16-20 h. The bacteria were collected by centrifuge at 6000 rpm for 10 min, and the supernatant was removed. The bacteria harvested were suspended in *PfAgo/TtdAgo* purification buffer (20 mM Tris-HCl pH 7.5, 400 mM NaCl, 5% glycerol), and PMSF with a final concentration of 1 mM was added. Then the cells were lysed using a high-pressure homogenizer and centrifuged at 18000 rpm for 50 min. The supernatant was filtered by 0.45 µm filter membrane. An appropriate amount of Ni-NTA resin was added into the gravity purification column, and the Ni beads were washed with *PfAgo/TtdAgo* purification buffer and activated with *PfAgo/TtdAgo* purification buffer containing 10 mM imidazole. The filtered supernatant was added to the gravity purification column containing Ni-NTA resin and combined for 50 min at 4°C. First, *PfAgo/TtdAgo* purification buffer containing 20 mM and 50 mM imidazole was used to wash impurity proteins, and then *PfAgo/TtdAgo* purification buffer containing 75 mM, 100 mM, 200 mM, 300 mM, 400 mM and 500 mM imidazole was used to elute target proteins. The Ni beads affinity chromatography results were analyzed by SDS-PAGE, and high purity proteins were selected for ultrafiltration concentration using Amicon 50K filter units (Millipore). Final proteins concentration was determined using Bradford Protein Assay Kit, the concentrated proteins were stored in separate packages after gel filtration, frozen quickly with liquid nitrogen, and stored at -80°C.

### Complex assembly

Guides and targets were dissolved with *PfAgo/TtdAgo* purification buffer. The *PfAgo/TtdAgo* protein was incubated with 16 nt long 5'-phosphorylated guide DNA in *PfAgo/TtdAgo* purification buffer containing 1 mM Mg<sup>2+</sup>. Then the 17 nt long target DNA was added. The molar ratio of *PfAgo/TtdAgo* protein, guides, and targets was 1:1.1:1.1. The incubation temperature was 80°C. The final volume was 500 µL and the amount of *PfAgo/TtdAgo* protein was 2 mg. The acquired *PfAgo* ternary complexes were applied to a size exclusion column (Superdex 200 Increase 10/300 GL, Cytiva). The purification buffer (20 mM Tris-HCl pH 7.5, 250 mM NaCl) was used for elution. The proteins eluted at different time were analyzed by SDS-PAGE and A<sub>260/A<sub>280</sub></sub> determination. Finally, the protein at the peak of the chromatogram was selected for subsequent structural analysis.

### Circular dichroism spectroscopy

Circular dichroism spectra of wild type and mutants of *PfAgo* were measured at room temperature with a Chirascan spectropolarimeter (Applied Photophysics) using a quartz cuvette with a 10 mm path length. The samples (0.2 mg/ml) were measured in PBS buffer. The measurements were recorded at wavelengths from 260 to 195 nm (0.2 nm/step) with a total of three scans for each. The spectra were corrected by subtracting the spectrum from the buffer background.

### Target DNA cleavage assay

The 5'-P-labeled DNA guides (18 nt in length) and 5'-FAM-labelled DNA targets (45 nt in length) were synthesized by Shanghai Sangon. The purified *TtdAgo*, guides, and targets were mixed with molar ratio of 5:10:2 (500 nM *TtdAgo*: 1 µM guide: 200 nM target) in 1×Reaction Buffer (10 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 5 mM MnCl<sub>2</sub>, 5% glycerol). Then the samples were incubated for 10 min at 95°C for target cleavage. The purified *PfAgo*, guides, and targets were mixed with molar ratio of 5:10:2 (500 nM *PfAgo*: 1 µM guide: 200 nM target) in 1×Reaction Buffer (10 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 5 mM MgCl<sub>2</sub>, 5% glycerol). Then the

samples were incubated for 25 min at 80°C for target cleavage. All reactions were terminated by adding an equal volume of RNA loading dye (95% formamide, 18 mM EDTA, 0.025% SDS and 0.025% bromophenol blue) and heating for 5 min at 95°C. The cleavage products were resolved by 20% denaturing polyacrylamide gel and visualized using Gel Doc™ XR+ (Bio-Rad).

#### Ensemble fluorescence resonance energy transfer assay

The 5'-P-labeled and 3'-Cy3/Cy5-labeled guide DNA (gDNA) of 16 nt length were synthesized by Sangon. The 3'-Cy3-labeled and 3'-Cy5-labeled gDNA were separately premixed with PfAgo for 10 min at 70°C for formation of Cy3-labeled gDNA/PfAgo (C1) and Cy5-labeled gDNA/PfAgo (C2) complexes. Then C1, C2 and 17 nt-target were quickly mixed and added to the 200 μL fluorescent colorimetric cuvette. PfAgo, guides and targets were mixed with 1:1.1:1.1 ratio (100 nM PfAgo:110 nM guide:110 nM target) in Reaction Buffer (10 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 0.5 mM MnCl<sub>2</sub>). The Cy3 (donor) and Cy5 (acceptor) signals were recorded at 1 s intervals using a fluorescence spectrophotometer (F4600, HITACHI) for 1 h at 60°C. The excitation wavelength and emission wavelength of Cy3 were 550 nm and 570 nm, respectively. The excitation wavelength and emission wavelength of Cy5 were 650 nm and 670 nm, respectively. The ratio of Cy5 and Cy3 fluorescence intensity was plotted to time. A catalytically inactive double mutant of PfAgo, named PfAgo\_DM (D558A/D628A), was constructed and conducted the same fluorescence resonance energy transfer experiments.

#### Electron microscopy

Aliquots of 4 μL samples (~0.6 mg/ml for sample of PfAgo-guide DNA with target DNA of different length, and ~0.7 mg/ml for TtdAgo-guide DNA-target DNA) were applied to glow-discharged Quantifoil holey carbon grids (Au, R1.2/1.3, 200 mesh). The grids were blotted with force 2 for 10 seconds and plunged into liquid ethane using a Vitrobot. Cryo-EM data were collected with a Titan Krios microscope (FEI) operated at 300 kV and images were collected using EPU<sup>56</sup> at a nominal magnification of 105,000x (resulting in a calibrated physical pixel size of 0.85 Å/pixel) with a defocus range from -1.2 μm to -2.2 μm. The images were recorded on a K3 sum-mit electron direct detector in super-resolution mode at the end of a GIF-Quantum energy filter operated with a slit width of 20 eV. A dose rate of 15 electrons per pixel per second and an exposure time of 2.5 seconds were used, generating 40 movie frames with a total dose of ~ 54 electrons per Å<sup>2</sup>. A total of 2404, 1474 and 1806 movie stacks were collected for PfAgo-guide DNA-target DNA, PfAgo-guide DNA-target DNA 11bp complex and TtdAgo-guide DNA-target DNA, respectively (Table 1).

#### Image processing

The movie frames were imported to RELION-3,<sup>51</sup> aligned using MotionCor2<sup>49</sup> with bin 2, and subjected to contrast transfer function (CTF) estimation using Gctf<sup>50</sup> on the fly.

For PfAgo-guide DNA-target DNA dataset, 2,325,621 particles were blob picked and extracted from the dose weighted micrographs. 911,385 particles were selected from 2D classification. Particles from different views were used to generate initial models in cryoSPARC.<sup>52</sup> Heterogeneous refinement (3D classification) was performed to distinguish different conformational states. 621,102 particles were used for final 3D refinement, converging at 2.9 Å resolution. Details of the cryo-EM image processing are summarized in Table 1.

For dataset of PfAgo-guide DNA-target DNA 11bp complex, 610,079 particles were blob picked and extracted from the dose weighted micrographs. 32,913 particles were selected from 2D classification. 9000 articles from different views were used to generate initial models in cryoSPARC. 3D classification was performed to distinguish different conformational states. 16,186 particles were used for final 3D refinement, converging at ~7 Å resolution.

For TtdAgo-guide DNA-target DNA dataset, 2,476,975 particles were auto-picked and extracted from the dose weighted micrographs. 2D classification was performed to exclude false and bad particles that fall into 2D averages with poor features. 676,552 particles were selected for further processing. Particles from different views were used to generate initial models in cryoSPARC. Heterogeneous refinement (3D classification) was performed to distinguish different conformational states. 629,194 particles were used for final 3D refinement, converging at 2.9 Å resolution. Details of the cryo-EM image processing are summarized in Table 1.

#### Model building, refinement, and visualization

For high-resolution model building, AlphaFold2-predicted<sup>57</sup> PfAgo and TtdAgo was manually adjusted against the corresponding cryo-EM maps using Coot.<sup>54</sup> Guide DNA and target DNA was built *de novo* using Coot. Model refinement was performed using *phenix.real\_space\_refine* tool in Phenix.<sup>55</sup> The refinement statistics are summarized in Table 1. For the structure of 11bp complex, PfAgo protein structure was split into individual domains and fit into the map using rigid body fitting tool in UCSF Chimera.<sup>53</sup> Figures were generated via PyMOL and UCSF Chimera.

#### QUANTIFICATION AND STATISTICAL ANALYSIS

*In vitro* experiments were performed at least three times.