ARTICLES

Structure of the guide-strand-containing argonaute silencing complex

Yanli Wang[1], Gang Sheng[1], Stefan Juranek[2], Thomas Tuschl[2] & Dinshaw J. Patel[1]

The slicer activity of the RNA-induced silencing complex is associated with argonaute, the RNase H-like PIWI domain of which catalyses guide-strand-mediated sequence-specific cleavage of target messenger RNA. Here we report on the crystal structure of *Thermus thermophilus* argonaute bound to a 5′-phosphorylated 21-base DNA guide strand, thereby identifying the nucleic-acid-binding channel positioned between the PAZ- and PIWI-containing lobes, as well as the pivot-like conformational changes associated with complex formation. The bound guide strand is anchored at both of its ends, with the solvent-exposed Watson–Crick edges of stacked bases 2 to 6 positioned for nucleation with the mRNA target, whereas two critically positioned arginines lock bases 10 and 11 at the cleavage site into an unanticipated orthogonal alignment. Biochemical studies indicate that key amino acid residues at the active site and those lining the 5′-phosphate-binding pocket made up of the Mid domain are critical for cleavage activity, whereas alterations of residues lining the 2-nucleotide 3′-end-binding pocket made up of the PAZ domain show little effect.

RNA interference is routinely used in biological research to effectively silence genes of interest[1–3] and is being increasingly investigated as a therapeutic modality against diverse human diseases[4,5]. The argonaute (Ago) protein[6–11], as a key catalytic component of the RNA-induced silencing complex, has a central role in the RNA interference pathway by mediating the maturation of small interfering RNA (siRNA) through initial degradation of the passenger strand, followed by guide-strand-mediated sequence-specific cleavage of target mRNA.

X-ray crystal structures of archaeal *Pyrococcus furiosus* Ago in its free state[12,13] and eubacterial *Aquifex aeolicus* Ago in its free state[14] and when externally bound to siRNA[15] have defined the folds and relative positioning of protein domains and connecting linker segments (see Supplementary Fig. 1 for sequence alignments of prokaryotic and eukaryotic Agos). A mechanistic understanding of guide-strand-mediated mRNA recognition and cleavage chemistry mediated by Ago requires additional structural information on the binary complex of Ago with bound guide strand substrate and, in the longer term, on the ternary complex with added mRNA strand.

Given that binding and cleavage studies have identified eubacterial *A. aeolicus* Ago to be a site-specific DNA-guided endoRNase[14], we attempted to crystallize the binary complex of an eubacterial *T. thermophilus* Ago with bound 5′-phosphorylated DNA guide strands ranging in length from 8–12 to 14, 16, 18–24, 26 and 28 bases. Diffraction quality crystals were obtained solely for binary complexes with 21-base and 10-base DNA guide strands. No crystals were obtained for the complex with the corresponding 5′-phosphorylated 21-base RNA guide strand.

Structure of the 21-base binary complex

The 3.0 Å crystal structure of *T. thermophilus* Ago complexed with a 5′-phosphorylated (5′-phos-$T_1GAGG_5TAGTA_{10}GGTTG_{15}TATAG_{20}T$) 21-base DNA guide strand is shown in Fig. 1a, with crystallographic statistics listed in Supplementary Table 1. The protein in a ribbon representation is colour-coded by domains (N, PAZ, Mid and PIWI) and linkers (L1 and L2). The bound DNA is shown in red, with the chain traceable for nucleotides 1–11 and 18–21, including two phosphates within the remaining disordered segment. The corresponding view to that in Fig. 1a, with the protein in a surface representation, colour-coded according to its electrostatic potential, is shown in Fig. 1b. The bound DNA guide strand threads its way within a central basic channel between the PAZ-containing (N, L1 and PAZ) and PIWI-containing (Mid and PIWI) lobes of the bilobal scaffold of Ago, thereby contacting all domains and linkers and defining the nucleic-acid-binding channel within the protein (Fig. 1a, b).

Base stacking within the seed segment

An alternate view of the complex, with the protein in a ribbon representation, is shown in stereo in Fig. 1c. A rotating view of Fig. 1c can be accessed as Movie 3DLH in the Supplementary Information. Within its traceable 1 to 11 segment, the bound DNA is continuously stacked from residues 2 to 10, with distinct breaks at the 1–2 and 10–11 steps (Fig. 1c, Supplementary Fig. 2 and $2F_o - F_c$ density map of bound DNA in Supplementary Fig. 3a). The observed integrity of stacked residues in the 5′-half (residues 2 to 10) but not the 3′-half of the 21-base DNA in the complex (Supplementary Fig. 2) is consistent with conclusions of functional experiments that established a toleration for mismatches within the 3′-segment but not the 5′-segment of the guide strand in its complex with mRNA[16,17].

The corresponding expanded view of the complex with the protein in an electrostatic surface representation is shown in Supplementary Fig. 4. The Watson–Crick edges of bases 2 to 6 are exposed, identifying this segment of the guide strand as the nucleation site for pairing with mRNA. Our structural results are consistent with previous functional studies that identified a 'seed' segment (residues 2 to 8) within the guide strand as critical for mRNA pairing[18,19]. In contrast, bases 7 to 11 are threaded inside the Ago scaffold in the binary complex (Supplementary Fig. 4).

Anchoring both ends of bound guide strand

The 21-base DNA is anchored at both its 5′ and 3′ ends in the complex, thereby constraining the trajectory of the bound guide

[1]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [2]Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, New York, New York 10065, USA.

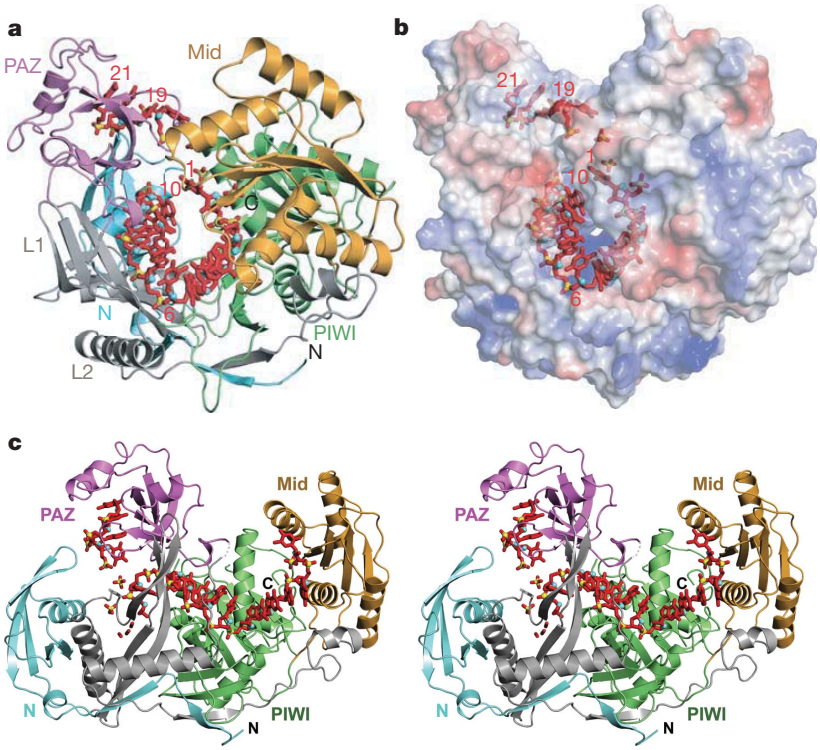

Figure 1 | Crystal structure of *T. thermophilus* argonaute bound to a 5′-phosphorylated 21-base DNA guide strand. **a**, Stick (DNA) and ribbon (Ago) view of the 3.0 Å structure of the complex. Individual domains and linkers are colour-coded. The bound 21-base DNA guide strand is coloured red (with phosphorus atoms in yellow), and can be traced for nucleotides 1 to 11 and 18 to 21, together with phosphates at positions 16 and 17. **b**, Corresponding stick (DNA) and electrostatically colour-coded surface (Ago) view of the complex. **c**, An alternate stereo view of the complex from that shown in **a**. For rotation of **c**, see Supplementary Movie 2DLH.

strand. The 5′-phosphate is anchored within the binding pocket in the Mid domain, with its oxygens hydrogen-bonded to the side chains of highly conserved residues (R418, K422, S432, Q433, K457) and a bound magnesium ion (Fig. 2a) as previously observed in the *A. fulgidus* Piwi protein–siRNA complex[20,21]. The magnesium is coordinated to the first and third phosphates from the 5′ end, as well as the carboxy-terminal carboxylate end (V685) of the PIWI domain (Fig. 2a). This result reinforces earlier functional conclusions that the integrity of the 5′-phosphate-binding pocket within the Mid domain is critical for slicing fidelity[21].

The bases are splayed apart at the 1–2 step, with base 1 stacked over the side chain of an arginine (R418, Fig. 2a). This distortion at the 5′ end of the guide strand reinforces an earlier structural observation in the *A. fulgidus* Piwi–siRNA complex[20,21], as well as functional experiments[16–18], which indicated that non-canonical pairing with mRNA is preferred at the 1 position in the guide strand.

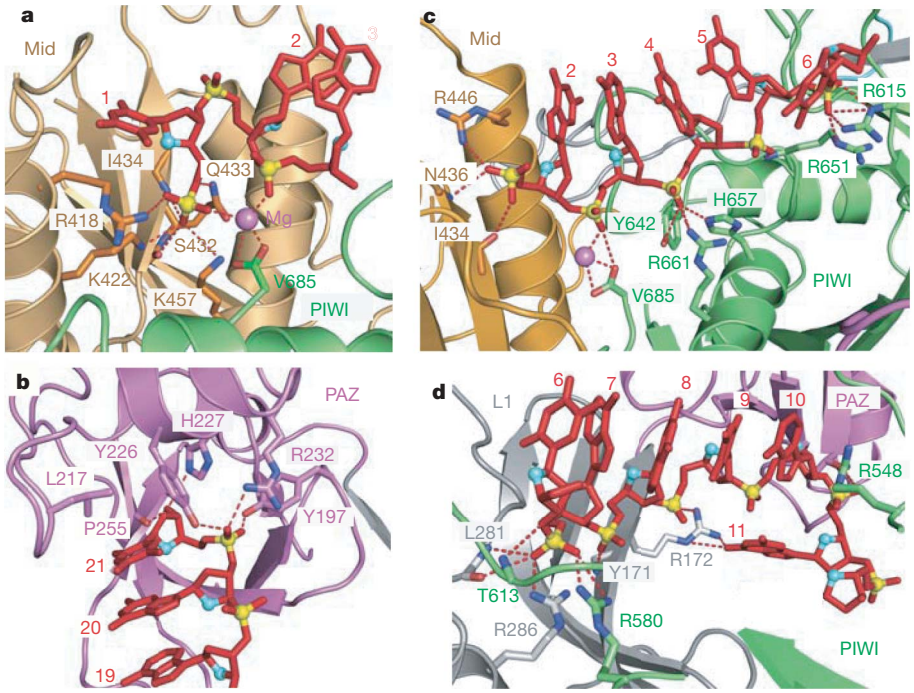

Figure 2 | Intermolecular hydrogen-bonding alignments in the *T. thermophilus* Ago bound to a 21-base DNA guide strand. **a**, Insertion of the 5′-phosphate of the DNA guide strand into the binding pocket in the Mid domain. **b**, Positioning of stacked 3′-end residues 20 and 21 of the DNA guide strand into the binding pocket in the PAZ domain. **c**, Positioning of stacked residues 2 to 6 of the DNA guide strand, with emphasis on intermolecular interactions. **d**, Positioning of stacked residues 6 to 10 and 11 of the DNA guide strand, with emphasis on intermolecular interactions.

Residues 20 and 21 at the 3′ end of the complex are anchored within the binding pocket in the PAZ domain (Fig. 2b and $2F_o - F_c$ density map of bound DNA in Supplementary Fig. 3b), with the oxygens of the phosphate linking residues 20 and 21 and the 3′-OH of residue 21 hydrogen-bonded to aromatic (Y197, Y226, H227) and basic (R232) side chains, as previously observed in PAZ–siRNA[22] and PAZ–single-stranded RNA[23] complexes.

The positioning of the RNase H-type catalytic residues D546, D478 (on adjacent β-strands) and D660 of the PIWI domain of *T. thermophilus* Ago relative to segment 1–11 of the bound 21-base DNA guide strand is outlined in Supplementary Fig. 2.

There is extensive hydrogen bonding and salt bridge formation between the backbone phosphates of the DNA guide strand and basic Arg side chains spanning the various domains in the protein (Fig. 2c, d). These extensive intermolecular contacts are consistent with related interactions located towards the 5′ end of the guide strand in the earlier *A. fulgidus* Piwi–siRNA structure[20,21] and proposed models of the catalytic cycle[2,7,14], where the anchored guide strand is retained while the cleaved message dissociates during each cycle of the multiple turnover catalytic process[10,11].

The 10–11 cleavage site step

The guanidinium group of R548 (conserved amongst the thermophilic Agos, whereas Val occupies this position in human AGOs) stacks on the base at position 10, with the resulting block enforcing an unanticipated orthogonal alignment of bases 10 and 11 (Fig. 2d). In addition, the guanidinium group of R172 (also observed in *A. aeolicus* Ago, but otherwise not conserved) has a bridging role through hydrogen bonding both to the phosphate group of residue 9 and to the base edge of residue 11 (Fig. 2d and $2F_o - F_c$ density map in Supplementary Fig. 3c). We anticipate that Arg residues R172 and R548, which enforce the pronounced kink at the 10–11 step in the Ago–guide-strand binary complex (Fig. 2d), will probably undergo a conformational switch on mRNA pairing during the propagation step of ternary complex formation, so as to generate a catalytically competent undistorted helical conformation at the 10–11 step associated with the mRNA cleavage site[17,24,25].

The helical trajectory of the stacked 2 to 10 segment of the bound 21-base DNA guide strand in the binary Ago complex (in red, Supplementary Fig. 5a, b) deviates from the helical trajectory of both A-form (in cyan, Supplementary Fig. 5a) and B-form (in green, Supplementary Fig. 5b) helices, following best-fit superposition of residues 2 to 6. This probably implies that either anchoring of both ends (Figs 1c and 2a, b) and/or intermolecular interactions (Fig. 2c, d) perturb the helical trajectory of the bound DNA guide strand in the complex. Nevertheless, it is significant that the 5′ end of the bound DNA is ordered in the absence of pairing with message.

Conformational transitions

We have been unable to grow crystals of *T. thermophilus* Ago in the free state, but have succeeded in solving the 2.7 Å structure with a bound 5′-phosphorylated 10-base DNA guide strand (5′-phos-$T_1GAGG_5TAGTA_{10}$, Fig. 3a), with crystallographic statistics listed in Supplementary Table 1. We can trace 3′-end residues 6 to 10 in the bound 10-base DNA sequence (Supplementary Fig. 6), with residues 9 and 10 inserted into the PAZ pocket (Fig. 3a). The C terminus of the *T. thermophilus* Ago protein is disordered in the 10-base DNA guide strand complex, and hence the 5′-phosphate-binding pocket in the Mid domain must be disrupted in the absence of an inserted C-terminal carboxylate group. The relative positioning of the 3′ ends of the bound 21-base (residues 18 to 21, in magenta) and 10-base (residues 6 to 10, in green) DNAs in their respective complexes after superposition of their PAZ domains are shown in Fig. 3b. The two 3′-end nucleotides that are inserted within their PAZ pockets superposition quite well, but the remaining nucleotides (18 and 19 in the 21-base and 6 to 8 in the 10-base complexes) follow different trajectories (Fig. 3b). This reflects differences in the relative orientations of the N and PAZ domains between the two complexes (Supplementary Fig. 7a), as well as intermolecular contacts (Supplementary Fig. 7b, c), suggestive of at least two nucleic-acid-binding channels leading from the PAZ pocket.

The relative domain alignments of Ago in complex with bound 10-base and 21-base DNA guide strands, after superposition of their PAZ domains (boxed region), is shown in Fig. 3c and in stereo in Supplementary Fig. 8. There are significant conformational changes

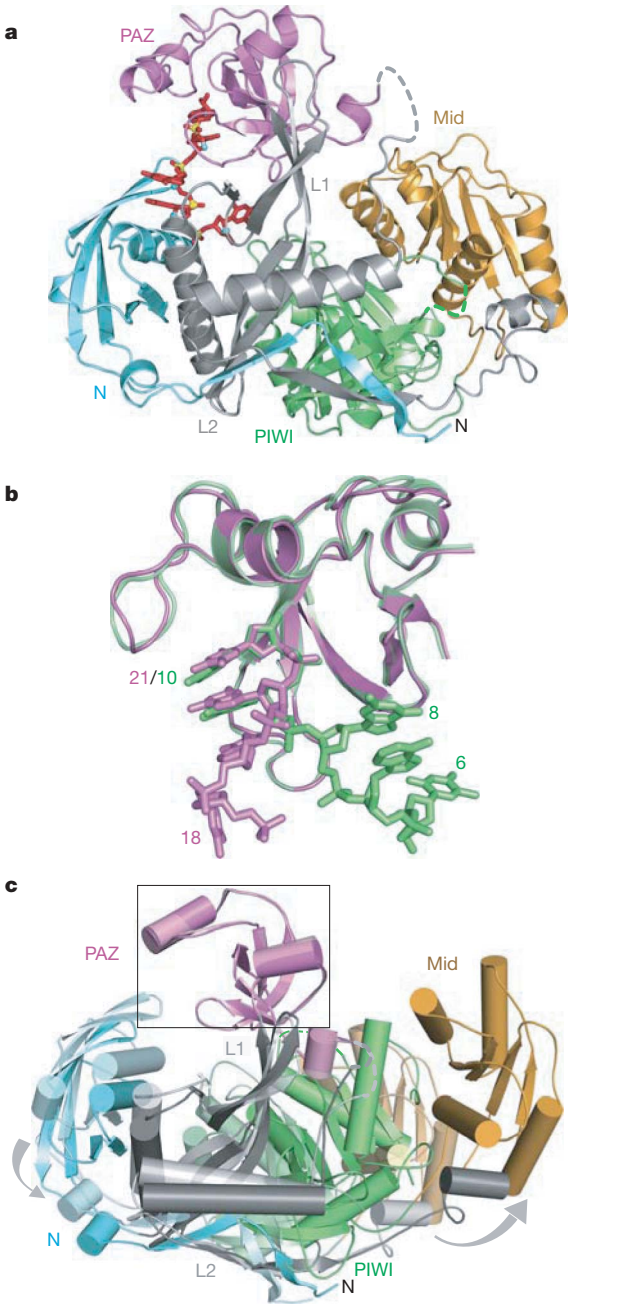

Figure 3 | Conformational changes in *T. thermophilus* Ago on formation of the 21-base DNA guide strand complex. **a**, The 2.7 Å structure of *T. thermophilus* Ago bound to a 5′-phosphorylated 10-base DNA guide strand. The bound 10-base DNA can be traced for nucleotides 6 to 10. **b**, Superpositioning of the PAZ domains together with observable 3′ ends in DNA 21-base (in magenta) and 10-base (in green) guide strand complexes. **c**, Alignment of Agos in complexes with bound 10-base (in lighter shades) and 21-base (in darker shades) DNA guide strands, after superpositioning of their PAZ domains (boxed segment). The grey arrows indicate the magnitude of the conformational changes on proceeding from the 10-base to the 21-base complexes. For transition between **a** and Fig. 1c, see Supplementary Movie 3DLB–3DLH.

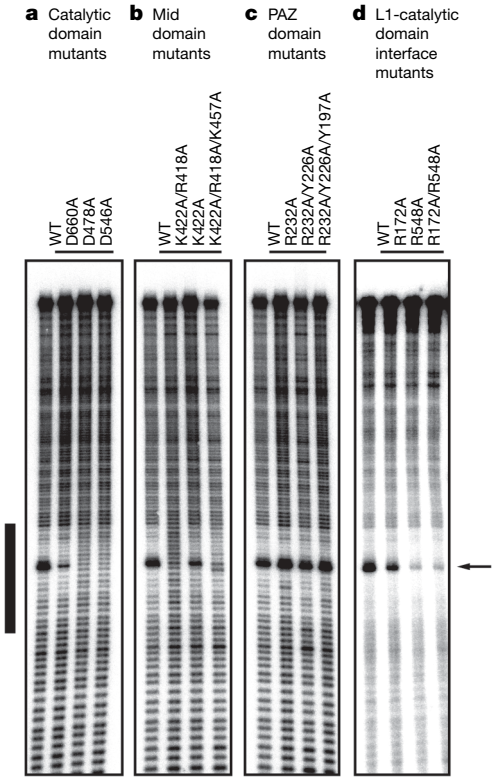

Figure 4 | DNA-guide-dependent RNA cleavage activity of wild-type and mutant *T. thermophilus* Ago. A 5′-phosphorylated 21-nucleotide DNA oligonucleotide sequence was pre-incubated with the indicated Ago proteins followed by the addition of 5′-radiolabelled RNA cleavage substrate. Cleavage products were resolved using denaturing polyacrylamide gel electrophoresis gels and visualized by phosphoimaging. The black bar indicates the region of the cleavage substrate covered by the guide DNA, with the cleavage site shown by an arrow. **a**, Catalytic Asp mutants (left panel); **b**, Mid domain mutants (central left panel); **c**, PAZ domain mutants (central right panel); and **d**, the L1-catalytic domain interface mutants (right panel).

on proceeding from the 10-base complex with an anchored 3′ end (lighter colours, Fig. 3c) to the 21-base complex with both 5′- and 3′-anchored ends (darker colours in Fig. 3c), with grey arrows highlighting the magnitude of the transitions. Specifically, the Mid domain rotates by 22° (Supplementary Fig. 9a) and the PAZ-containing lobe rotates by 25° (Supplementary Fig. 9b), both with respect to the PIWI domain. These movements can be tracked in Movie 3DLB–3DLH provided in the Supplementary Information. The net result of rotating the Mid domain relative to the PIWI domain is to increase the length of the nucleic-acid-binding channel towards the 5′-binding site by 8 Å, thereby accommodating the full length of the bound 21-base DNA guide strand. By contrast, the conformational change in the PAZ-containing lobe relative to the PIWI domain reflects a closing movement that results in the generation of a narrower nucleic acid-binding channel necessary to hold and orient the bound 21-base DNA guide strand.

Cleavage assays

To characterize the catalytic activity of the *T. thermophilus* Ago protein, we performed target RNA cleavage assays (Fig. 4). Similar to the observations made previously for *A. aeolicus* Ago protein[14], the cleavage activity was enhanced by addition of divalent manganese ions and the cleavage position was located 10 nucleotides downstream from the residue paired with the 5′ end of the DNA guide strand (see also Supplementary Fig. 10). Mutation of any one of the Asp residues of the catalytic triad abolished or substantially reduced target RNA cleavage (Fig. 4a). Alteration of K422 in the 5′-phosphate-binding Mid domain pocket also reduced activity, and double or triple mutations including R418 and K457 completely abolished activity (Fig. 4b). In contrast, alterations of up to three conserved residues of the 2-nucleotide 3′-end-binding PAZ domain pocket showed no reduction of cleavage activity (Fig. 4c). We also evaluated the impact on cleavage after mutation of the two arginines whose interactions promote the orthogonal alignment of bases 10 and 11 in the complex. The R172A mutant (linker L1) had a small, but noticeable, impact on cleavage, while significantly reduced cleavage was observed for the R548A mutant (PIWI domain), as well as for the R172A/R548A dual mutant (Fig. 4d).

Our long-term goal is to capture Ago–nucleic acid complexes during distinct assembly and functional steps of the RNA interference catalytic cycle[26,27]. The earlier structures of the *A. fulgidus* Piwi protein[28] bound to siRNA duplexes[20,21] together with functional implications[10,11] and our current structure of the *T. thermophilus* Ago bound to a 21-base DNA guide strand represent initial but critical contributions to the success of this endeavour. We expect that the structural insights thus derived will be portable to eukaryotic AGOs that function as RNA-guide-strand-mediated endoRNases, thereby providing the molecular underpinnings defining guide-strand-mediated recognition and processing of mRNA[29,30], with eventual impact on design and delivery strategies for silencing human disease using RNA interference[4,5].

METHODS SUMMARY

Wild-type and mutant *T. thermophilus* Agos were overexpressed from *Escherichia coli* and purified by column chromatography. The structure of Ago complexed with the 5′-phosphorylated 10-base DNA was determined by multi-wavelength anomalous diffraction on the selenomethionine (SeMet)-modified protein. The structure of Ago complexed with 5′-phosphorylated 21-base DNA was determined by molecular replacement using the domains of the Ago–10-base DNA complex structure as search models. Details of all biochemical and crystallographic procedures are listed in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 April; accepted 6 August 2008.
Published online 27 August 2008.

1. Baulcombe, D. RNA silencing in plants. *Nature* **431,** 356–363 (2004).
2. Filipowicz, W. The nuts and bolts of the RISC machine. *Cell* **122,** 17–20 (2005).
3. Rana, T. M. Illuminating the silence: understanding the structure and function of small RNAs. *Nature Rev. Mol. Cell Biol.* **8,** 23–36 (2007).
4. Kim, D. H. & Rossi, J. J. Strategies for silencing human disease using RNA interference. *Nature Rev. Genet.* **8,** 173–184 (2007).
5. De Fougerolles, A., Vornlocher, H.-P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nature Rev. Drug. Discov.* **6,** 443–453 (2007).
6. Hall, T. M. Structure and function of Argonaute proteins. *Structure* **13,** 1403–1408 (2005).
7. Tomari, Y. & Zamore, P. D. Perspective: machines for RNAi. *Genes Dev.* **19,** 517–529 (2005).
8. Tolia, N. H. & Joshua-Tor, L. Slicer and the argonautes. *Nature Chem. Biol.* **3,** 36–43 (2007).
9. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nature Rev. Mol. Cell Biol.* **9,** 22–32 (2008).
10. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends Biochem. Sci.* **31,** 622–630 (2006).
11. Patel, D. J. *et al.* Structural biology of RNA silencing and its functional implications. *Cold Spring Harb. Symp. Quant. Biol.* **71,** 81–93 (2006).
12. Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305,** 1434–1437 (2004).
13. Rivas, F. V. *et al.* Purified Ago2 and an siRNA form recombinant human RISC. *Nature Struct. Mol. Biol.* **12,** 340–349 (2005).
14. Yuan, Y. R. *et al.* Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell* **19,** 405–419 (2005).
15. Yuan, Y. R. *et al.* A potential protein–RNA recognition event along the RISC loading pathway from the structure of *A. aeolicus* Ago with externally bound siRNA. *Structure* **14,** 1557–1565 (2006).
16. Doench, J. G. & Sharp, P. A. Specificity of miRNA target selection in translational repression. *Genes Dev.* **18,** 504–511 (2004).
17. Haley, B. & Zamore, P. D. Kinetic analysis of the RNAi enzyme complex. *Nature Struct. Mol. Biol.* **11,** 599–606 (2004).

18. Lewis, B. P. *et al.* Prediction of mammalian microRNA targets. *Cell* **115,** 787–798 (2003).
19. Stark, A., Brennecke, J., Russel, R. B. & Cohen, S. M. Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1,** 397–409 (2003).
20. Parker, J. S., Roe, S. M. & Barford, D. Structural insights into mRNA recognition from a PIWI domain–siRNA guide complex. *Nature* **434,** 663–666 (2005).
21. Ma, J. B. *et al.* Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* **434,** 666–670 (2005).
22. Ma, J., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the Paz domain. *Nature* **429,** 318–322 (2004).
23. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3′-end recognition by the Argonaute2 Paz domain. *Nature Struct. Mol. Biol.* **11,** 576–577 (2004).
24. Nowotny, M., Gaidamakov, S. A., Crouch, R. J. & Yang, W. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **121,** 1005–1016 (2005).
25. Chiu, Y. L. & Rana, T. M. RNAi in human cells: Basic structural and functional features of small interfering RNA. *Mol. Cell* **10,** 549–561 (2002).
26. Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15,** 188–200 (2001).
27. Martinez, J. & Tuschl, T. RISC is a 5′-phosphomonoester-producing RNA endonuclease. *Genes Dev.* **18,** 975–980 (2004).
28. Parker, J. S., Roe, S. & Barford, D. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J.* **23,** 4727–4737 (2004).
29. Liu, J. *et al.* Argonaute2 is the catalytic engine of RNAi. *Science* **305,** 1437–1441 (2004).
30. Meister, G. *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15,** 185–197 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The research was supported by funds from the NIH to D.J.P. and T.T. We thank the staff of NE-CAT beam line at the Advanced Photon Source, Argonne National Laboratory, supported by the US Department of Energy, for assistance with data collection.

Author Contributions Y.W. and G.S. expressed and purified *T. thermophilus* Ago and its mutants, and grew crystals of the complex. Y.W. collected X-ray diffraction data and solved the structure of the complex. The structural studies were undertaken under the supervision of D.J.P. S.J. was responsible for the cleavage assays on the wild-type and mutant Agos under the supervision of T.T. All authors read and approved the submitted manuscript.

Author Information The structures of *T. thermophilus* Ago bound to 5′-phosphorylated 21-base and 10-base DNAs have been deposited to the Protein Data Bank under accession codes 3DLH and 3DLB, respectively. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.J.P. (pateld@mskcc.org).

METHODS

Protein expression and purification. An expression vector (PET-SUMO, Invitrogen) was constructed to produce full-length *T. thermophilus* Ago. The *E. coli* cells were collected for 8 h after incubation with 0.1 mM IPTG at 20 °C. The cells were lysed in 20 mM Tris-HCl buffer, 1 M NaCl, 2 mM $MgCl_2$, pH 7.5, at 25 °C. The protein was first purified using His-select nickel agarose resin (GE Healthcare). The protein was digested by ubiquitin-like protein 1 (Ulp1) SUMO protease (Invitrogen) and dialysed against 20 mM Tris-HCl, 0.5 M NaCl, 2 mM $MgCl_2$, pH 7.5, for 16 h at room temperature. The protein was then heated at 55 °C for 15 min and next purified by His-select nickel agarose and gel filtration (Superdex 200, Amersham) chromatography. Finally, it was concentrated to 30 mg ml^{-1} in 20 mM Tris-HCl, 0.5 M NaCl and 2 mM $MgCl_2$.

The Ago mutants were made with the Quik-Change kit (Stratagene) and verified by sequencing. We expressed and purified the mutant proteins following the same protocol as that used for the wild-type protein.

Crystallization and data collection. DNA oligonucleotides were purchased from Invitrogen. All crystals were grown by the hanging-drop vapour diffusion method. Native Ago or selenomethionine (SeMet)-labelled Ago was mixed with 5′-phosphorylated 10-base DNA at 1:2 molar ratio. The SeMet–Ago–10-base DNA crystals were grown from 3–5% (v/v) polyethylene glycerol 550 monomethyl ether (PEGMME550), 50 mM Tris-HCl, pH 7.5, 0.1 M KCl and 10 mM $MgCl_2$ at 35 °C. Crystals were flash-frozen in cryoprotecting solutions containing 3% PEGMME550, 50 mM Tris-HCl, pH 7.5, 0.1 M KCl, 10 mM $MgCl_2$ and 25% (v/v) glycerol.

The Ago–21-base DNA complex was prepared for crystallization by mixing the purified Ago with 1:1 molar ratio of 5′-phosphorylated 21-base DNA. The crystals were grown from 18% (v/v) polyethylene glycerol (PEG)8000, 50 mM MES, pH 6.5, 200 mM KCl and 80 mM $MgAc_2$ at 40 °C. They were flash-frozen in cryoprotecting buffer containing 16% PEG8000, 50 mM MES, pH 6.5, 200 mM KCl, 80 mM $MgAc_2$ and 25% (v/v) glycerol.

Diffraction data were collected at 100 K at beam line NE-CAT 24-IDC, the Advanced Photon Source (APS), Argonne National Laboratory. All data were processed with the HKL2000 suite[31] and data processing statistics are summarized in Supplementary Table 1. The Ago–10-base DNA complex belonged to the $P2_1$ space group and diffracted to 2.7 Å. The Ago–21-base DNA complex belonged to the $P2_12_12_1$ space group and diffracted to 3.0 Å.

Structure determination and refinement. The structure of the Ago–10-base DNA complex was solved by the multi-wavelength anomalous diffraction method using the SeMet-labelled complex. Phases were calculated with programs SOLVE and RESOLVE[32–34]. The model was built with program COOT[35] and was improved by several cycles of manual rebuilding and refinement with CNS[36] and REFMAC[37]. There are two molecules in the asymmetric unit. For the Ago–10-base DNA complex, 43 residues at the C terminus could not be traced, nor could we trace DNA residues 1 to 5 in the complex.

The structure of the Ago–21-base DNA complex was solved by molecular replacement with the program PHASER[38], using the domains of the Ago–10-base DNA complex structure as a search model. The final model of the Ago–21-base DNA complex contains two molecules in the asymmetric unit. Almost all protein residues and nucleotides 1–11 and 18–21 of the bound DNA in one molecule and nucleotides 1–10 and 19–21 of the other molecule in the asymmetric unit could be traced in the complex.

Cleavage activity assay of *T. thermophilus* Ago. 5′-phosphorylated 21-nucleotide guide oligodeoxynucleotides were prepared by solid-phase synthesis using standard DNA amidites (Sigma-Proligo) and chemical phosphorylating reagent (Glen Research) on an ABI 3400 DNA synthesizer. The 177-nucleotide cleavage substrate was prepared by *in vitro* transcription from a PCR template as described previously[39]. The transcript was dephosphorylated using alkaline phosphatase (Roche) and then 5′-radiolabelled using T4 polynucleotide kinase (Fermentas) and γ-^{32}P-ATP.

Recombinant *T. thermophilus* Ago (0.5 μM) was incubated with a reaction mixture containing 10 mM HEPES-KOH, pH 7.5, 100 mM NaCl, 5 mM $MnCl_2$, 8 U of RNasin and 0.1 μM guide strand for 90 min at 55 °C in a final volume of 15 μl. Next, ATP and GTP were added to a final concentration of 1 mM and 0.2 mM, respectively, together with 0.1 μM 5′-^{32}P-labelled RNA substrate at a final concentration of 0.5 μM. The incubation was continued for 60 min. The reaction was stopped by addition of 185 μl proteinase K solution (1 mg ml^{-1} proteinase K, 20 mM HEPES-KOH, pH 7.5, 1.5 mM EDTA, 100 mM NaCl, 1.5 mM $CaCl_2$, 1.5% SDS), and incubated at 55 °C for 10 min, followed by phenol-chloroform extraction and ethanol precipitation. The cleavage products were resolved by 8% denaturing PAGE, and radioactivity was detected by phosphorimaging.

31. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
32. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55**, 849–861 (1999).
33. Terwilliger, T. C. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59**, 38–44 (2003).
34. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
35. Emsley, P. and Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
36. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
37. The CCCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
38. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
39. Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R. & Tuschl, T. Single-stranded anti-sense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563–574 (2002).