

Published in final edited form as:

*Nature*. 2005 March 31; 434(7033): 663–666. doi:10.1038/nature03462.

## Structural insights into mRNA recognition from a PIWI domain – siRNA guide complex

James S. Parker, S. Mark Roe, and David Barford

Section of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London, SW3 6JB, UK

### Abstract

RNA interference (RNAi) and related RNA silencing phenomena use short antisense guide RNA molecules to repress expression of target genes<sup>1,2</sup>. Argonaute proteins<sup>3</sup>, containing N-terminal PAZ domains and C-terminal PIWI domains, are core components of these mechanisms. Here we present the crystal structure of a PIWI protein from *Archaeoglobus fulgidus* (AfPiwi) in complex with a small interfering (si)RNA-like duplex, that mimics the 5' end of a guide RNA strand bound to an overhanging target mRNA. The structure reveals a highly conserved metal-binding site that anchors the 5' nucleotide of the guide RNA. The first base pair of the duplex is unwound, separating the 5' nucleotide of the guide from the complementary nucleotide on the target strand, which exits with the 3' overhang through a short channel. The remaining base-paired nucleotides assume an A-form helix, accommodated within a channel in the PIWI domain, which can be extended to place the scissile phosphate of the target strand adjacent to the putative slicer catalytic site. This study provides insights into mechanisms of target mRNA recognition and cleavage by an Argonaute-siRNA guide complex.

---

RNA silencing is mediated by the effector complexes RISC (RNA-Induced Silencing Complex) and RITS (RNA-Induced initiation of Transcriptional gene Silencing), functioning in post-transcriptional and transcriptional silencing respectively<sup>4–6</sup>.

Argonautes, comprising Ago and Piwi subfamilies<sup>3</sup>, are the only proteins common to these complexes. Recent structural and biochemical studies implicate the PAZ domain in RNA 3' overhang recognition<sup>7–10</sup>. The PIWI domain is believed to incorporate the slicer catalytic site, responsible for cleavage of target mRNA during RNAi<sup>11–13</sup>. However, these results are insufficient to explain the requirement for a 5' phosphate group on the first nucleotide of the guide strand<sup>13–16</sup>, and the striking observation that the identity of the scissile phosphate in the target mRNA is measured relative to the 5' end of the guide strand, irrespective of modification or the length (within limits) of the 3' end<sup>15–18</sup>.

AfPiwi, an isolated archaeal PIWI domain protein, which serves as a model for understanding eukaryotic Argonaute PIWI domains, forms a single discrete complex with a 16 nucleotide siRNA-like duplex<sup>12</sup>. With the aim of providing a molecular picture of PIWI domain – RNA interactions, we have crystallized and determined the structure of this complex at 2.2 Å resolution. We assign the two RNA strands as guide and target, denoting the guide nucleotides G1 to G16 (5' – 3') and complementary nucleotides T1 to T14 (Fig. 1a). The 5' end of the duplex refers to the 5' end of the guide.

---

Correspondence and requests for materials should be addressed to D.B. (e-mail: david.barford@icr.ac.uk). The coordinates and structure factors have been deposited with the Protein Data Bank with accession numbers, 2bgg and r2bggsf, respectively..

**Supplementary Information** accompanies the paper on *Nature's* website (<http://www.nature.com>).

AfPiwi binds to the 5' end of the siRNA-like duplex, engaging the RNA in an L-shaped channel (Fig. 1b). The RNA electron density is strongest at this end, allowing us to model nucleotides G1 – G9 of the guide strand and T(-2) – T9 of the target (Fig. 1c). Nucleotides 2 – 9 on each strand base pair, adopting an A-form helix that extends along the channel (Fig. 1b and Supplementary Fig. 1). Strikingly, the 5' end of the RNA duplex is significantly distorted, resulting in the unwinding of G1 (uridine) and T1 (adenosine). G1, which carries the obligatory 5' phosphate of the guide, is displaced into a pocket surrounding the C-terminus of the PIWI domain (G1 binding pocket). T1 passes through the short “exit” portion of the channel, allowing the two 3' overhanging nucleotides, T(-1) and T(-2), to fold back along the protein surface. G1 and T1 are divided by the  $\alpha$ D helix, which contributes two aromatic residues (Phe 151 and Tyr 152) capping the first 5' base pair of the RNA helix (Fig. 1b).

The G1 binding pocket is part of a highly conserved region in the PIWI domain, designated CRI (Conserved Region I)<sup>12</sup>. A prominent structural feature is a divalent metal ion coordinated to the carboxylate group of the C-terminal residue, Leu 427 (Fig. 2a). This metal ion plays a crucial role in the distortion and anchoring of the guide RNA, contacting the phosphates of G1 and G3. Without exception, all Argonaute sequences share three conserved hydrophobic residues at their extreme C-terminus (Supplementary Fig. 2) indicating that an exposed carboxylate is an invariant structural feature.

Within the G1 binding pocket, the side chains of four invariant residues in Argonaute proteins (Tyr 123, Lys 127, Gln 137 and Lys 163; Supplementary Fig. 2) contact the 5' phosphate group of G1 (Fig. 2a). Tyr 123 further stabilises the flipped conformation of G1 by stacking with the unpaired uracil base. In addition to interactions with these invariant side chains and the metal ion, the 5' phosphate contacts the backbone amide of Phe 138, fixed in position by the adjacent Gln 137. No other amino acids contact G1. Only five other residues are invariant in the PIWI domains of Argonaute proteins (Supplementary Fig. 2). Four of these appear to maintain the PIWI fold, and the fifth is a putative slicer catalytic site residue<sup>11,12</sup>. The G1 binding pocket is therefore the most conserved structural feature of the PIWI domain. Consistent with our structural data, mutation of this pocket significantly reduced the affinity of AfPiwi for the siRNA-like duplex<sup>12</sup>.

Interactions between AfPiwi and the guide are also mediated by the phosphate groups of G2 – G5 (Fig. 2). A second key anchor site is the phosphate group of G3. In addition to its coordination by the metal, the phosphate interacts with Gln 159 and Arg 380, residues conserved in the Piwi and Ago subfamilies, respectively (Supplementary Fig. 2). The phosphate of G2 contacts Tyr 152, conserved as Tyr or Thr in the Ago subfamily, and the main chain amide of Arg 140. The phosphate of G4 also contacts Arg 380, and the phosphate of G5 interacts with Arg 385, via a water molecule. The only protein-base contact occurs between Asn 155, conserved in the Ago subfamily, and the 2' oxygen of the G2 uracil. Hydrogen bond acceptors occupy equivalent positions in cytosine and purines. Therefore, as anticipated, none of the interactions provide RNA sequence specificity. Interactions with G1 – G5 account for all conserved residues of CRI (Fig. 3b), strongly indicating a conserved mode of RNA 5' end binding within the Argonaute family. In contrast to the extensive interactions between PIWI and the guide, few contacts are formed to the target strand, consistent with its requirement to dissociate after mRNA cleavage<sup>19,20</sup>.

With G1 distorted into the conserved binding pocket, the corresponding base pair (G1 -T1) is unwound (Fig. 1b). The finding that a single guide – target mismatch at this position can enhance slicer activity<sup>19</sup>, demonstrates that the AfPiwi-RNA complex corresponds to a cleavage-competent conformation. It also suggests an equilibrium between the preferred “active” RNA duplex conformation we observe, and an “inhibitory” conformation where G1

pairs with T1. The mode of guide RNA binding to PIWI - a highly-ordered anchored 5' region, with bases exposed apart from G1 - is consistent with the observation that, for animal miRNAs, only nucleotides 2 – 8 of the miRNA guide (the “seed”) are critical for target recognition and pairing<sup>21–24</sup>. Similarly, in RISC-mediated mRNA cleavage, bases at the 5' end of the siRNA guide dominate its affinity for target RNA<sup>19</sup>.

The G1 binding pocket and the putative slicer catalytic region<sup>11,12</sup> lie at either end of the PIWI domain channel. Nucleotides G10 – G16 of the guide and T10 – T14 of the target are less clearly defined in our electron density maps. To visualise the trajectory of a full-length (19 nucleotide) siRNA guide bound to a mRNA target, we modelled the remaining nucleotides, assuming an A-form helix contiguous with the visible duplex (Fig. 3 and Supplementary Fig. 1). The weaker electron density for nucleotides G10 and G11 is compatible with this conformation (Fig. 1c). This model places the scissile phosphate of the target strand (linking T10 and T11)<sup>17,18</sup> adjacent to the putative slicer catalytic region<sup>11,12</sup> (Fig. 3 and Supplementary Fig. 3). The PIWI-RNA binding conformation is therefore consistent with our functional interpretation of the complex, being representative of a guide-target interaction, and provides insight into how the site-specificity of mRNA cleavage is determined with reference to the 5' end of the guide<sup>17,18</sup>.

Our structure might also reflect the initial phase of RISC association with an siRNA duplex; here the target strand would represent the passenger strand. The relative stabilities of the ends of an siRNA duplex determine strand fate: the strand whose 5' end is located at the less stable end becomes the guide, the other (passenger) strand is discarded<sup>25,26</sup>. Specific recognition of the siRNA-like 5' end, and partial unwinding of the duplex revealed from the AfPiwi-RNA complex, suggests that binding to PIWI may be the first step in siRNA duplex unwinding in RISC; PIWI could therefore participate in strand selection. Consistent with this notion is evidence supporting a role for Argonaute 2 in siRNA unwinding in *Drosophila*<sup>27</sup>. R2D2 has also been implicated in guide strand selection, binding with a preference to the more stable end of an siRNA duplex<sup>28</sup>.

A role for the PIWI domain in mediating interactions with the 5' end of the guide RNA complements the function of the PAZ domain in recognising RNA 3' overhangs<sup>7–10</sup>. In *Pyrococcus furiosus* Argonaute, the PAZ domain is positioned over the channel linking the G1 binding pocket and the putative slicer site<sup>11</sup>. The PAZ domain may function as a mobile unit that locks down over the guide-target duplex, providing additional constraints for slicing<sup>11</sup>. However, when bound to a target mRNA, it is unlikely that the recessed guide 3' end would interact with the PAZ domain<sup>10</sup>; instead, we visualise the duplex emerging into solvent or the immediate environment of RISC.

Our structure rationalises much molecular and biochemical data for mechanisms of RNA silencing, which together with the universal conservation of the G1 binding pocket, indicates that the AfPiwi-RNA complex provides a representative picture of PIWI domain-RNA interactions. The view now emerges that Argonaute functions as a dynamic molecule mediating interactions with both the 5' and 3' ends of short guide RNA molecules, presenting them for recognition by cognate mRNA.

## Methods

### Complex formation and crystallization

AfPiwi was purified as described<sup>12</sup>, except that the protein was eluted at the final step in 5 mM Hepes pH 7.5, 800 mM NaCl, 2 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>, 8 mM DTT. The RNA oligonucleotide was purchased purified, deprotected and desalted from Dharmacon, and resolubilized in 10 mM Hepes pH 7.5, 100 mM NaCl, 2 mM MgCl<sub>2</sub>. Annealing was

performed by incubation at 90°C for 1 min followed by gradual cooling (1°C/5 min) from 65°C to 4°C. To form the complex, equal volumes of AfPwi (260 μM) and annealed RNA (310 μM) were mixed and incubated at room temperature for 30 mins. Crystallization was performed using the hanging drop method at 14°C. Crystal clusters were obtained overnight by mixing 1 μl of complex with 1 μl 200 mM KCl, 10 mM CaCl<sub>2</sub>, 50 mM Na Cacodylate pH 5.5, 15% PEG 4000, 5 mM DTT. Single crystals were obtained by streak seeding into a similar drop containing 12% PEG 4000 after an overnight incubation. For cryoprotection, crystals were passed through stabilisation solutions containing increasing amounts of MPD to 17% and then flash frozen at 100K.

### Structure determination

The structure (with two AfPwi-RNA complexes per P1 unit cell) was solved by molecular replacement with PHASER<sup>29</sup> using the native protein (PDB code 1W9H) as search model<sup>12</sup>. Some loops and all the RNA were rebuilt manually in COOT<sup>30</sup> and the complex was refined with REFMAC<sup>29</sup>. 508 waters were added in the final stages of refinement. The final R and R<sub>free</sub> were 0.18 and 0.21 respectively (Supplementary Table 1). Terminal nucleotides of the target strands in the two RNA duplexes per unit cell adopt slightly different conformations (Supplementary Fig. 1). Based on crystallization conditions and refinement, we have assigned the metal ion within the G1 binding pocket as manganese.

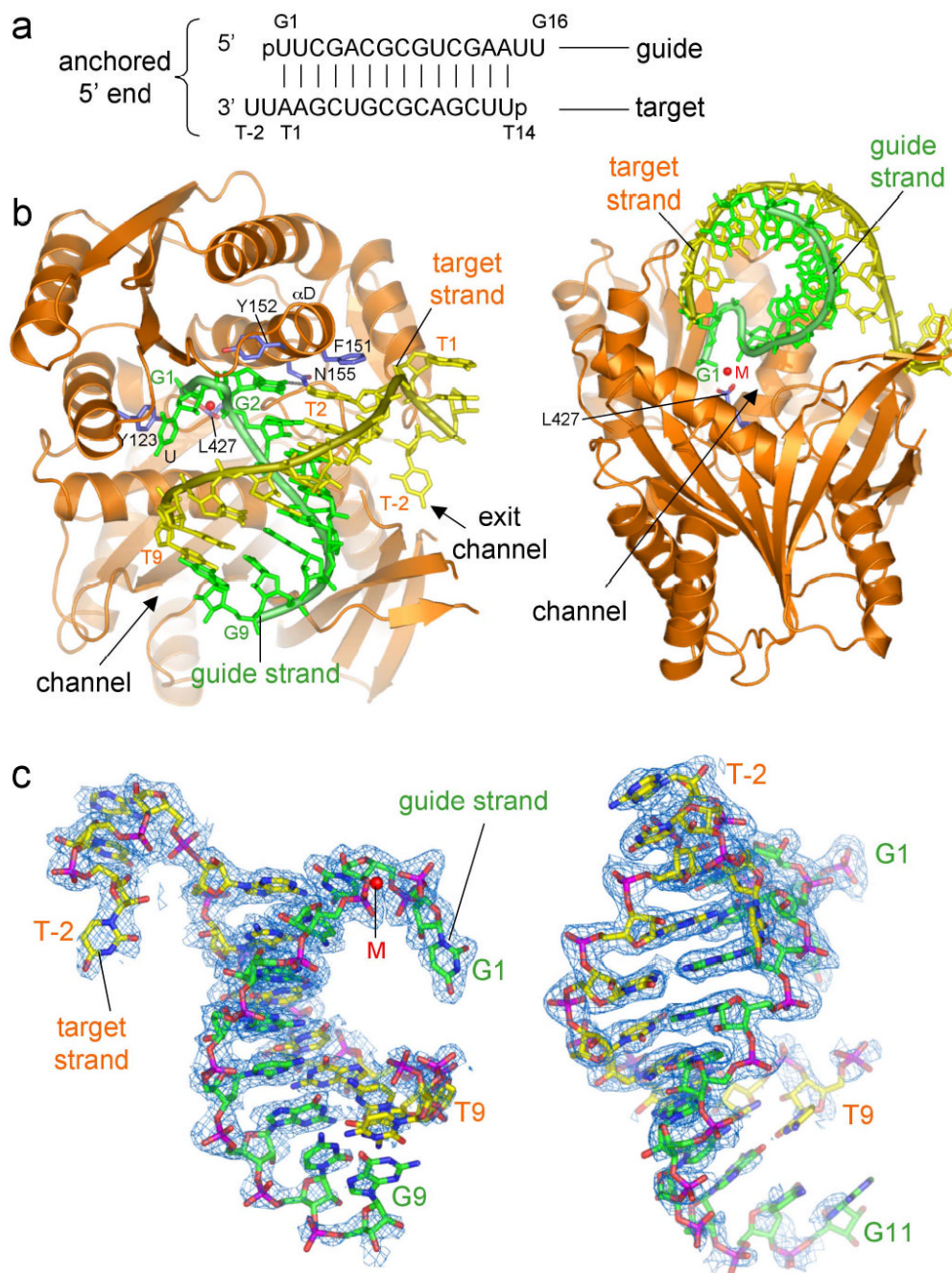
### Acknowledgments

The work was funded by Cancer Research-UK, Wellcome Trust and ICR. We are grateful to staff at ESRF for help with data collection. The authors declare no competing interests.

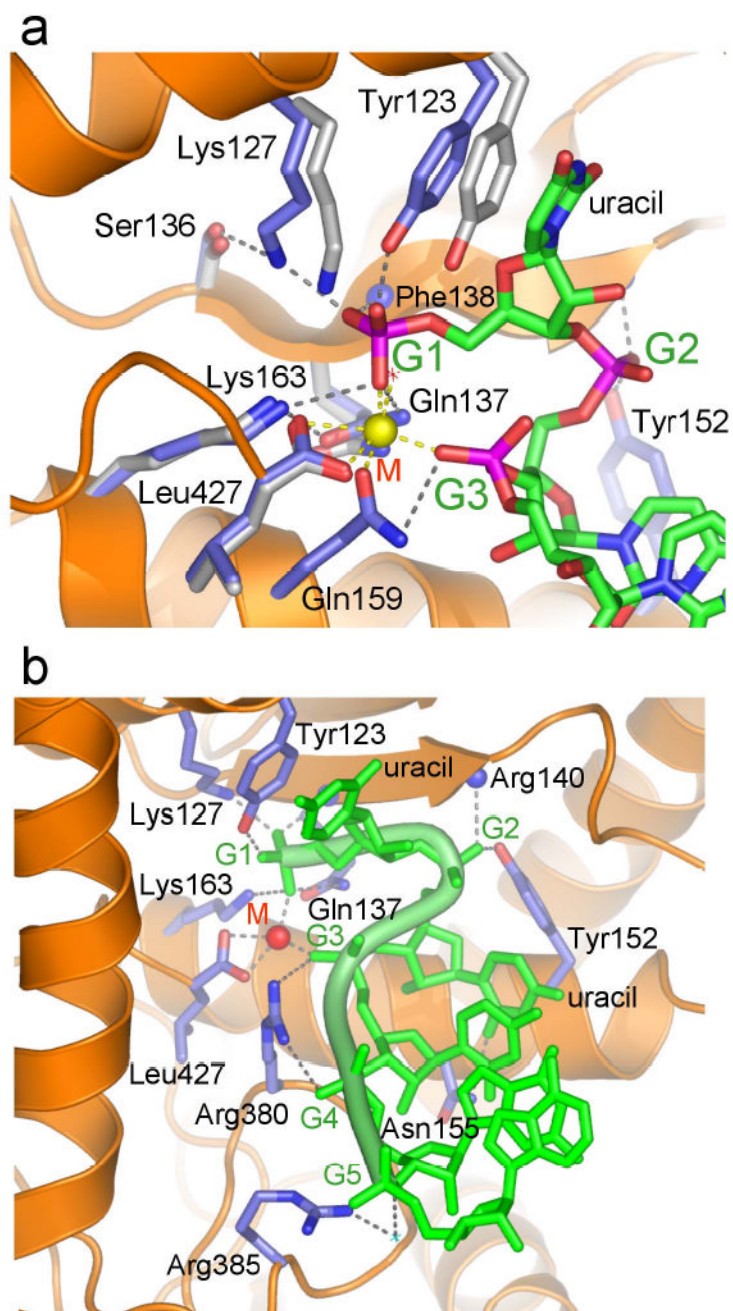
### References

1. Mello CC, Conte D Jr. Revealing the world of RNA interference. *Nature*. 2004; 431:338–42. [PubMed: 15372040]
2. Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. *Nature*. 2004; 431:343–9. [PubMed: 15372041]
3. Carmell MA, Xuan Z, Zhang MQ, Hannon GJ. The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis. *Genes Dev*. 2002; 16:2733–42. [PubMed: 12414724]
4. Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science*. 2001; 293:1146–50. [PubMed: 11498593]
5. Verdell, A. et al. RNAi-mediated Targeting of Heterochromatin by the RITS Complex. *Science* (2004).
6. Sontheimer, E. J. Assembly and function of RNA silencing complexes. *Nat Rev Mol Cell Biol* (2005).
7. Yan KS, et al. Structure and conserved RNA binding of the PAZ domain. *Nature*. 2003; 426:468–74. [PubMed: 14615802]
8. Song JJ, et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol*. 2003; 10:1026–32. [PubMed: 14625589]
9. Lingel A, Simon B, Izaurralde E, Sattler M. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol*. 2004; 11:576–7. [PubMed: 15156196]
10. Ma JB, Ye K, Patel DJ. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature*. 2004; 429:318–22. [PubMed: 15152257]
11. Song JJ, Smith SK, Hannon GJ, Joshua-Tor L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*. 2004; 305:1434–7. [PubMed: 15284453]
12. Parker JS, Roe SM, Barford D. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J*. 2004; 23:4727–37. [PubMed: 15565169]

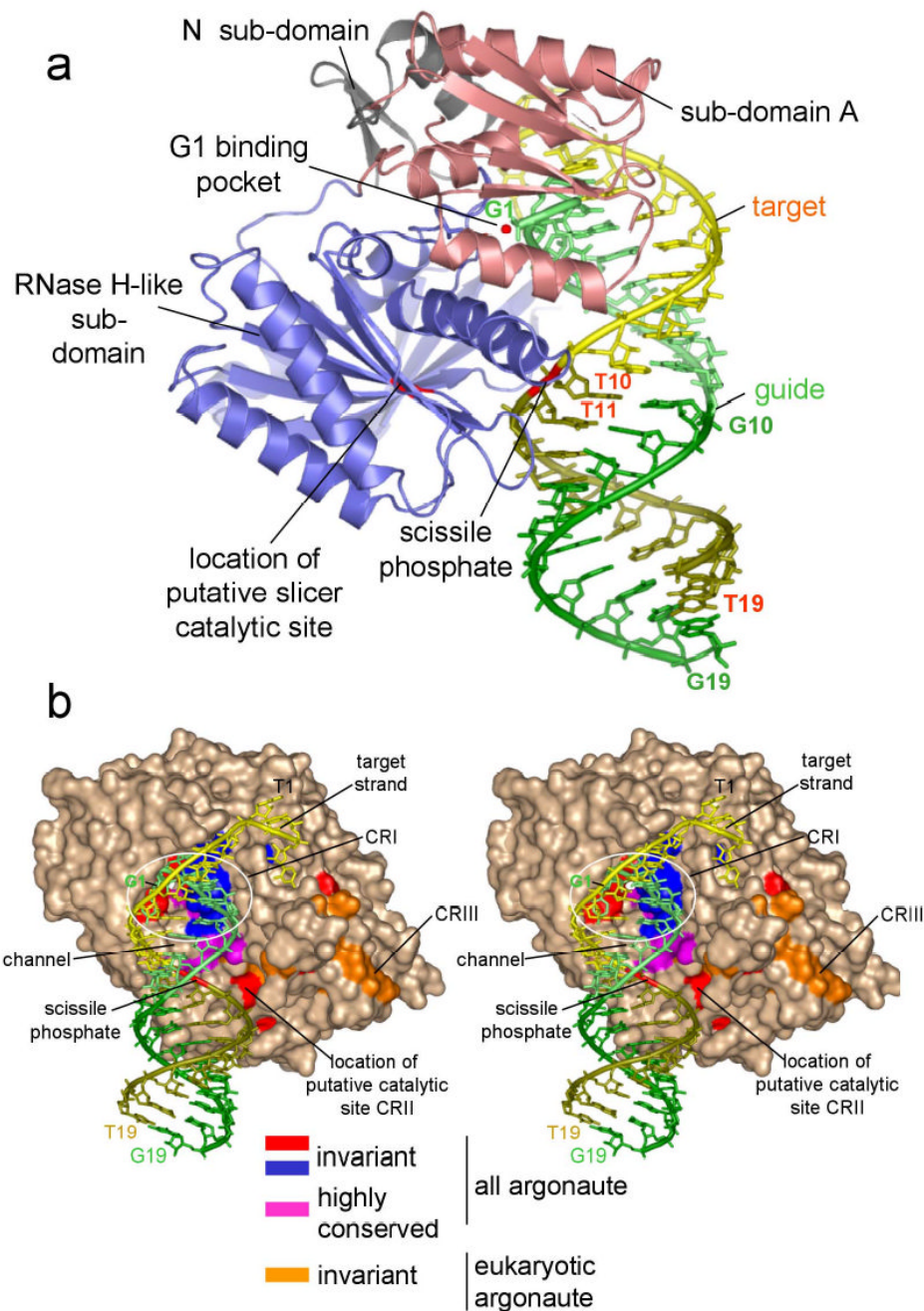
13. Liu J, et al. Argonaute2 is the catalytic engine of mammalian RNAi. *Science*. 2004; 305:1437–41. [PubMed: 15284456]
14. Nykanen A, Haley B, Zamore PD. ATP requirements and small interfering RNA structure in the RNA interference pathway. *Cell*. 2001; 107:309–21. [PubMed: 11701122]
15. Chiu YL, Rana TM. RNAi in human cells: basic structural and functional features of small interfering RNA. *Mol Cell*. 2002; 10:549–61. [PubMed: 12408823]
16. Martinez J, Patkaniowska A, Urlaub H, Luhrmann R, Tuschl T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*. 2002; 110:563–74. [PubMed: 12230974]
17. Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T. Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *EMBO J*. 2001; 20:6877–88. [PubMed: 11726523]
18. Elbashir SM, Lendeckel W, Tuschl T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev*. 2001; 15:188–200. [PubMed: 11157775]
19. Haley B, Zamore PD. Kinetic analysis of the RNAi enzyme complex. *Nat Struct Mol Biol*. 2004; 11:599–606. [PubMed: 15170178]
20. Hutvagner G, Zamore PD. A microRNA in a multiple-turnover RNAi enzyme complex. *Science*. 2002; 297:2056–60. [PubMed: 12154197]
21. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell*. 2003; 115:787–98. [PubMed: 14697198]
22. Stark A, Brennecke J, Russell RB, Cohen SM. Identification of *Drosophila* MicroRNA Targets. *PLoS Biology*. 2003; 1:397–409.
23. Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes Dev*. 2004; 18:504–11. [PubMed: 15014042]
24. Lewis BP, Burge CB, Bartel DP. Conserved Seed Pairing, Often Flanked by Adenosines, Indicates that Thousands of Human Genes are MicroRNA Targets. *Cell*. 2005; 120:15–20. [PubMed: 15652477]
25. Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. *Cell*. 2003; 115:209–16. [PubMed: 14567918]
26. Schwarz DS, et al. Asymmetry in the assembly of the RNAi enzyme complex. *Cell*. 2003; 115:199–208. [PubMed: 14567917]
27. Okamura K, Ishizuka A, Siomi H, Siomi MC. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev*. 2004; 18:1655–66. [PubMed: 15231716]
28. Tomari Y, Matranga C, Haley B, Martinez N, Zamore PD. A protein sensor for siRNA asymmetry. *Science*. 2004; 306:1377–80. [PubMed: 15550672]
29. Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*. 1994; 50:760–3. [PubMed: 15299374]
30. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:2126–2132. [PubMed: 15572765]



**Figure 1.** General features of AfPiwi-siRNA interactions and distortion of the 5' terminus of the RNA duplex. (a) Schematic of the RNA duplex co-crystallized with AfPiwi, indicating nomenclature for guide and target strands. (b) Orthogonal views of the AfPiwi-RNA complex. The divalent metal in the G1 binding pocket is shown as a red sphere (M). (c) Two orthogonal views of the RNA 2Fo-Fc electron density map contoured at  $1\sigma$ . Left: view of the RNA as seen by the protein. G1 to G9 (G1 to G11 right) and T(-2) to T9 are placed in electron density. Figure generated using PYMOL (<http://www.pymol.com>).



**Figure 2.** Interactions between AfPiwi and the guide RNA strand. (a) Detailed view of the G1 binding pocket. Side chains coloured grey show conformations of selected residues in the absence of RNA12. The concerted shift of Tyr 123 and Lys 127 suggests an induced fit mechanism for phosphate binding. The divalent metal (M) is shown as a yellow sphere. (b) View of interactions with the phosphate groups of G1 to G5. The target strand is omitted for clarity.



**Figure 3.**

Proposed model for a 19 nucleotide guide - target duplex bound to AfPiwi. (a) View showing the positioning of the scissile phosphate of the target strand (between T10 and T11, coloured red) with respect to the location of the putative slicer catalytic site. The modelled RNA region is coloured darker for both strands (G10 to G19 and T10 to T19). PIWI sub-domains are shown<sup>12</sup>. Catalytic Asp residues of the DDE motif<sup>11,12</sup> are absent in AfPiwi. (b) Stereoview with the molecular surface of AfPiwi colour-coded according to structural conservation (Supplementary Fig. 2). CRI, CRII and CRIII (Conserved Regions I, II and III) are indicated<sup>12</sup>.