LETTERS

Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2

Filipp Frank[1,2,3], Nahum Sonenberg[1,2] & Bhushan Nagar[1,3]

MicroRNAs (miRNAs) mediate post-transcriptional gene regulation through association with Argonaute proteins (AGOs)[1]. Crystal structures of archaeal and bacterial homologues of AGOs have shown that the MID (middle) domain mediates the interaction with the phosphorylated 5′ end of the miRNA guide strand and this interaction is thought to be independent of the identity of the 5′ nucleotide in these systems[2,3]. However, analysis of the known sequences of eukaryotic miRNAs and co-immunoprecipitation experiments indicate that there is a clear bias for U or A at the 5′ position[4–7]. Here we report the crystal structure of a MID domain from a eukaryotic AGO protein, human AGO2. The structure, in complex with nucleoside monophosphates (AMP, CMP, GMP, and UMP) mimicking the 5′ end of miRNAs, shows that there are specific contacts made between the base of UMP or AMP and a rigid loop in the MID domain. Notably, the structure of the loop discriminates against CMP and GMP and dissociation constants calculated from NMR titration experiments confirm these results, showing that AMP (0.26 mM) and UMP (0.12 mM) bind with up to 30-fold higher affinity than either CMP (3.6 mM) or GMP (3.3 mM). This study provides structural evidence for nucleotide-specific interactions in the MID domain of eukaryotic AGO proteins and explains the observed preference for U or A at the 5′ end of miRNAs.

Small RNAs (miRNAs and short interfering RNAs (siRNAs)) are short (21–24 nucleotides in length), double-stranded regulatory RNAs[8] that contain two signature characteristics: a two-nucleotide 3′ end overhang and a phosphorylated 5′ end. The functional strand, referred to as the guide strand, interacts directly with the multi-domain AGO proteins to form the RNA-induced silencing complex (Supplementary Fig. 1a). The passenger strand, miRNA*, is usually non-functional and released or degraded. The RNA-induced silencing complex recognizes messenger RNAs that contain sequences complementary to the 'seed region' (nucleotides two to eight) of the guide strand and targets them for post-transcriptional regulation[9].

A number of recent discoveries point towards a critical role for the phosphorylated 5′ nucleotide of small RNAs in the function of these RNAs and possibly in the biogenesis of RNA-induced silencing complexes. Initially, the termini of miRNAs are determined by cleavage of a precursor RNA with the ribonuclease III (RNaseIII) enzymes Drosha and Dicer, a process which can be subject to varying degrees of precision[6,10]. Variation in the site of cleavage at the 5′ end of the miRNA will affect the identity of the 5′ nucleotide as well as the sequence of the seed region, which is the main determinant of target recognition. Therefore, a further mechanism for reducing the variation of the position of the 5′ nucleotide must be in place and this seems to be achieved by the process of loading miRNAs onto AGOs[10]. Indeed, sequence analyses of nematode, plant and fly miRNAs show that there is a strong bias to having a U or an A at the 5′ position of the miRNA guide strand[4–7] and, importantly, a further enrichment of these bases is observed after loading onto AGOs[5]. We also carried out a bioinformatic analysis of conserved human miRNAs and found a similar bias towards U or A at the 5′ end (Supplementary Fig. 1c). The significance of the identity of the 5′ nucleotide is highlighted by its ability to influence which strand of the miRNA duplex gets incorporated into the RNA-induced silencing complex[5,11–13]. Collectively, these observations indicate that the interaction of eukaryotic miRNAs with AGOs is somehow responsible for the observed bias at the 5′-nucleotide position of the guide strand, which influences both the integrity of the seed sequence and miRNA strand selection.

The interaction between miRNAs and AGOs occurs through several contact points on the protein, mediated by distinct domains (Supplementary Fig. 1a). The 3′ end of the guide strand is recognized by the PAZ (Piwi-Argonaute-Zwille) domain, which was the first crystal structure determined from a eukaryotic AGO domain[14]. The 5′ nucleotide is anchored to the MID domain, as shown by several crystal structures of archaeal and bacterial AGO homologues in complex with nucleic acids[2,3,15–17] that serve as models for understanding eukaryotic AGOs. However, these structures failed to explain the observed 5′-nucleotide bias in eukaryotic miRNAs as these proteins preferentially interact with DNA instead of RNA and their sequence identities with eukaryotic AGOs are low (Supplementary Fig. 2). To understand the molecular basis for the 5′-nucleotide bias that is observed in sequence profiles of miRNAs from eukaryotic systems, we determined the crystal structure of the MID domain from human AGO2 (hAGO2) alone and in complex with UMP, AMP, CMP and GMP, which mimic the 5′ end of the guide RNA.

The structure of the hAGO2 MID domain resembles a Rossmann fold, and is characterized by a total of four alternating β-strands and α-helices (Fig. 1a and Supplementary Fig. 1b). The β-strands form an extended parallel β-sheet, which constitutes the core of the protein flanked by the α-helices. The hAGO2 MID domain has the same fold as that found in archaeal and bacterial homologues[2,3,15–19], although the positions of the secondary structural elements and loops vary considerably among the species (Supplementary Fig. 4).

The structures of the MID domain in the nucleotide co-complexes are essentially the same as that of the nucleotide-free state (root mean squared deviation (r.m.s.d.) for all Cα-atoms: AMP complex, 0.18 Å; CMP complex, 0.14 Å; GMP complex, 0.23 Å; UMP complex, 0.19 Å). The phosphate groups of the nucleotides are bound to a conserved, basic 5′-phosphate-binding pocket, which constitutes the most extensive positively charged region on the surface of the MID domain (Fig. 1b, c). The phosphate is hydrogen bonded to the side chains of four highly conserved residues, Y529, K533, Q545 and K570 (Supplementary Fig. 2), and the backbone of C546. Consistent with our structure is that mutation of the homologous residues in *Archaeoglobus fulgidus* Piwi (Y123A, K127A, Q137A and K163A)

[1]Department of Biochemistry, McGill University, Montréal, Québec H3G 0B1, Canada. [2]Goodman Cancer Center, McGill University, Montréal, Québec H3G 0B1, Canada. [3]Groupe de Recherche Axe sur la Structure des Proteines, Montréal, Québec H3G 0B1, Canada.

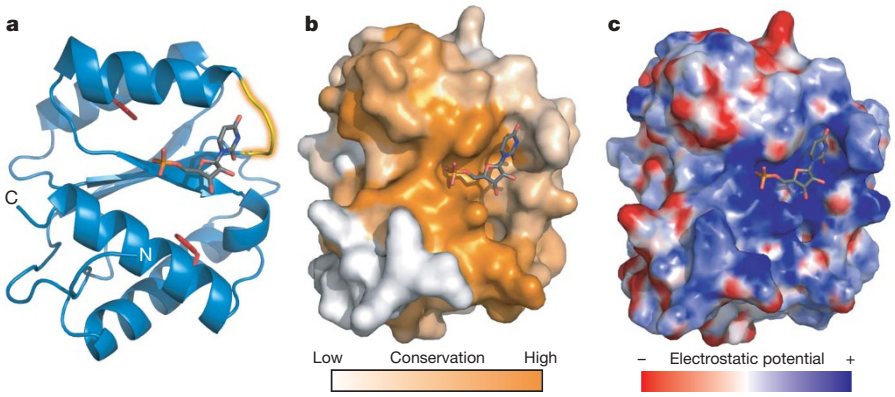

Figure 1 | Overall structure of the hAGO2 MID domain. a, Ribbon representation of the MID domain, with UMP depicted in stick representation. Highlighted in yellow is the nucleotide specificity loop. The phenylalanine residues, F470 and F505 (shown as sticks in red), proposed to be involved in cap binding[24] are both buried within the hydrophobic core of the protein (Supplementary Fig. 3). **b**, Mapping of surface conservation of selected eukaryotic AGOs (see Supplementary Fig. 2). **c**, Electrostatic potential surface representation[25]. All molecular figures were generated using PyMol (http://www.pymol.org).

resulted in decreased affinity for 5′-phosphorylated RNA, and that mutation of the phosphate-binding residues in hAGO2 (K533A, Q545A and K570A) led to decreased cleavage activity[2].

The positions of the nucleotides seen here are essentially the same as that of the 5′ nucleotide in previous structures of archaeal and bacterial AGO homologues bound to nucleic acids[2,3,15–18] (Supplementary Fig. 5), which showed that the 5′ nucleotide flips out of the helical duplex and engages the MID domain with several hydrogen bonds and van der Waals interactions. This suggests that our structures of the isolated MID domain bound to 5′-end miRNA-mimics reflect the physiological binding mode observed in intact AGOs bound to guide RNA. In full-length AGO, the phosphate group of the 5′ nucleotide also interacts with the carboxy terminus of AGO via a magnesium ion. Superposition of full-length *Thermus thermophilus* AGO onto the hAGO2 MID domain indicates that the C terminus interacts only with the phosphate moiety of the nucleotide and not the base (Supplementary Fig. 6).

The structures of the UMP and AMP complexes show clear, well-defined electron density for all components of the nucleotide (Fig. 2a, b). The bases of both AMP and UMP stack up against Y529 of the phosphate-binding pocket, which contributes non-specific recognition of the 5′ nucleotide in hAGO2. Similar non-specific interactions were observed in complexes of *A. fulgidus* Piwi and *T. thermophilus* AGO with nucleic acids[2,3,15–18]. However, in our structures we observe further specific interactions made by the bases of AMP and UMP with backbone atoms from within the loop connecting strand 3 and helix 3, which we term the nucleotide specificity loop (Fig. 2a, b). The larger purine moiety of AMP is slightly tilted relative to that of UMP such that the hydrogen bond acceptors in their pyrimidine rings (O4 and N1 from AMP and UMP, respectively) coincide in space, allowing them to make hydrogen bonds with the backbone amide group of T526 (Supplementary Fig. 7). Similarly, hydrogen bond donors in the pyrimidine ring (N3 and C2 from UMP and AMP, respectively) make contact with the carbonyl of G524. C2 from the aromatic pyrimidine ring of AMP can act as a hydrogen bond donor owing to the δ+ charge on the H2 atom[20].

For the CMP and GMP co-crystals, we observe residual electron density for the phosphate and ribose sugar but, notably, electron density for the bases is missing (Fig. 2c, d). CMP and GMP have hydrogen-bonding patterns opposite to those in UMP and AMP, respectively, resulting in charge repulsion from backbone atoms in the nucleotide specificity loop (Supplementary Fig. 8). Furthermore, GMP contains an amine group in its pyrimidine ring that would clash with the carbonyl of G524 (Fig. 2d). These structural observations are in agreement with the bias against G or C at the 5′ position of miRNAs (Supplementary Fig. 1c). It should be noted that the region corresponding to the nucleotide specificity loop in *T. thermophilus* AGO also makes backbone hydrogen bonds with the 5′-nucleotide base; however, whether these interactions confer selectivity was not further investigated[16] (Supplementary Fig. 9).

In the absence of bound nucleotide, the conformation of the nucleotide specificity loop is unchanged (Supplementary Fig. 10). P523 and P527 located at either end of this loop help to maintain its rigid conformation and G524 places a sharp kink in its trajectory (Supplementary Fig. 11a). The distinct conformation of this loop seems to be an important aspect of base recognition in hAGOs as the corresponding region in *A. fulgidus* Piwi is pulled away from the nucleotide base and does not make similar specific interactions[3,18] (Supplementary Fig. 12). The nucleotide specificity loop is absolutely conserved in all four human AGOs, as well as the miRNA-associated *Drosophila melanogaster* AGO1 and *Caenorhabditis elegans* ALG-1 and ALG-2 (Supplementary Fig. 11b). Conversely, the loop and indeed the majority of the protein are not highly conserved in AGOs that act in other small RNA pathways, such as the siRNA associated *D. melanogaster* AGO2 (Supplementary Fig. 11b).

To characterize the interaction of the various nucleotides with the MID domain directly in solution, we carried out ^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) NMR titration experiments. Both AMP and UMP caused substantial chemical shift changes in at least 12 peaks of the spectra corresponding mainly to backbone amides (Fig. 3a and Supplementary Fig. 13). In contrast, titration with GMP or CMP led to only half as many chemical shift differences. We next took the magnitudes of these chemical shift differences at increasing concentrations of nucleotide titrant to construct binding curves from which dissociation constants could be extracted (Fig. 3b and Table 1). In agreement with our crystallographic results, the MID domain has the highest affinity for UMP (0.12 mM) followed by AMP (0.26 mM), whereas the affinities for CMP (3.6 mM) and GMP (3.3 mM) are considerably weaker.

The interactions made between the nucleotide specificity loop and the nucleotide base are mediated by protein backbone atoms only; therefore, choosing mutations that can modulate the specificity of the loop to verify its role is difficult. As noted above, the region corresponding to the nucleotide specificity loop in *A. fulgidus* Piwi is too far from the base of the 5′ nucleotide to make specific interactions and its position varies depending on the nucleotide present (Supplementary Fig. 12). Comparison of the loop sequences of hAGO2 and *A. fulgidus* Piwi suggests that this may be due to a one-residue insertion in the latter (Supplementary Fig. 2). Hence, to test the function of the loop in imparting base specificity, we chose to introduce an extra residue (G) between K525 and T526. NMR titration analysis of this loop-insertion mutant of hAGO2 with nucleoside monophosphates shows that base

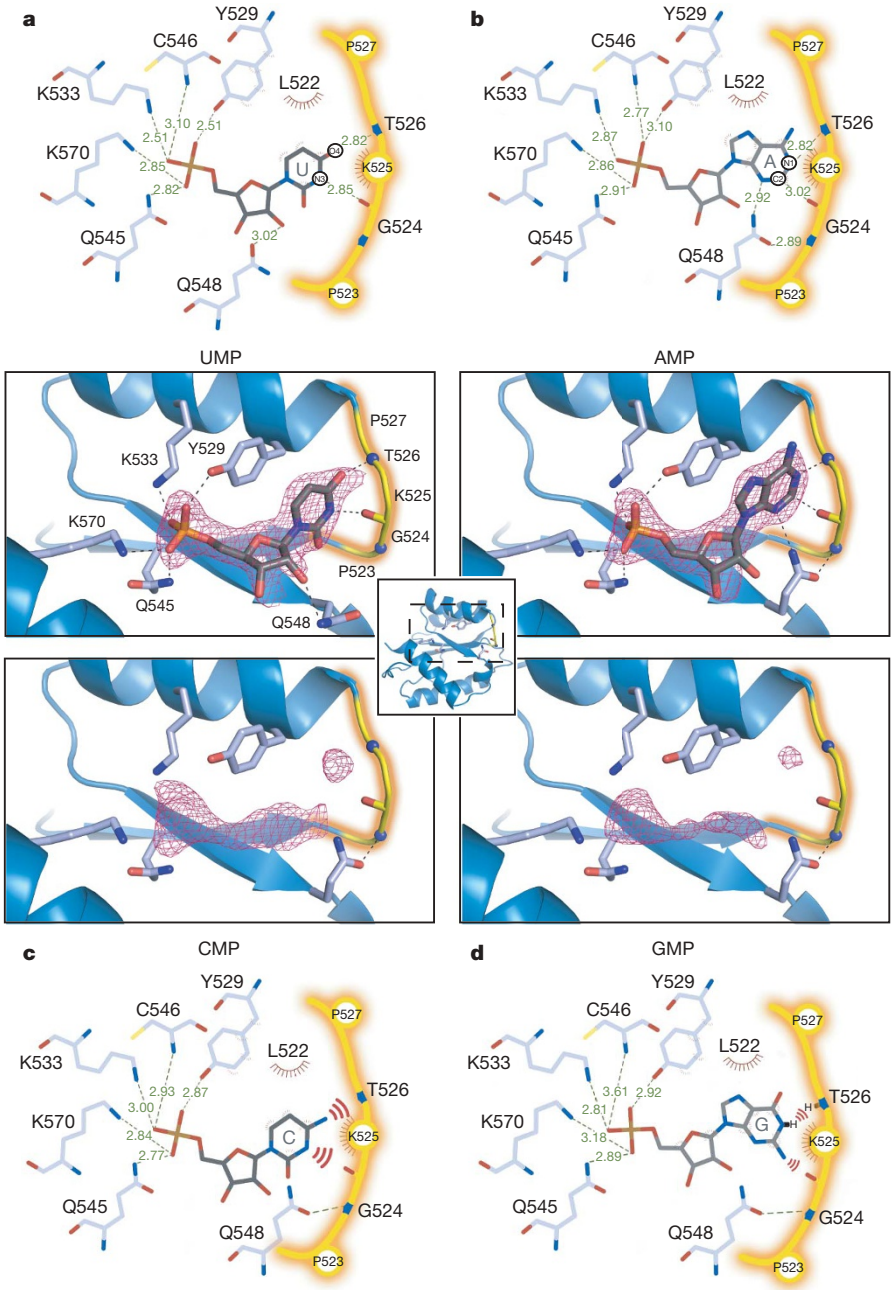

Figure 2 | Crystal structures of hAGO2 MID domain in complex with UMP, AMP, CMP and GMP. **a**–**d**, Human AGO2 MID domain in complex with UMP (**a**), AMP (**b**), CMP (**c**) and GMP (**d**). UMP and AMP are modelled and shown in stick representation. Only the phosphate groups of GMP and CMP were modelled (not shown). The nucleotide specificity loop is shown without side chains and highlighted in yellow. Relevant backbone atoms in the loop are indicated with blue spheres (nitrogen) and red sticks (oxygen). Difference electron density contoured at 2.5σ is shown before inclusion of any nucleotide in the model. Dotted black lines indicate hydrogen bonds. Adjacent to each structure are two-dimensional representations of the contacts between NMPs and the MID domain. The 5′-nucleotide specificity loop is highlighted in yellow and orange. In green are hydrogen bonds with distances in Å. 'Eyelashes' represent van der Waals contacts. Atoms of UMP and AMP that interact with the nucleotide specificity loop are labelled and circled. Curved red lines mark repulsive interactions between the base of GMP or CMP and the nucleotide specificity loop. Clashing hydrogen atoms for GMP are shown to highlight the repulsive interactions.

selectivity is abolished (Table 1 and Supplementary Figs 14 and 15). This confirms that the nucleotide specificity loop has an important role in selecting the identity of the base bound to the hAGO2 MID domain.

It is becoming apparent that the 5′ end of small RNAs is a crucial factor for various functional aspects of RNA silencing. Crystallographic studies carried out with *T. thermophilus* AGO have established that the 3′ end of the guide strand is released from its binding site in the PAZ domain on duplex formation with target RNA, whereas the 5′ end remains anchored to the MID domain[16]. Furthermore, a study using recombinant hAGO2 found that changing the 5′ U of let-7 miRNA abolishes cleavage of the target mRNA in an *in vitro* system, but the interaction between the single-stranded miRNA and hAGO2 was maintained[21]. Together with our structural data, this indicates that a tight interaction between the 5′ nucleotide and the MID domain, provided only by U or A, is required for maintaining the complex once the 3′ end is released on duplex formation and hence for proper target processing.

The importance of the 5′-nucleotide interaction with the MID domain is further emphasized by recent studies that analysed small RNAs associating with AGO proteins in *Arabidopsis thaliana*, which have identified a strong preference for a particular 5′ nucleotide in

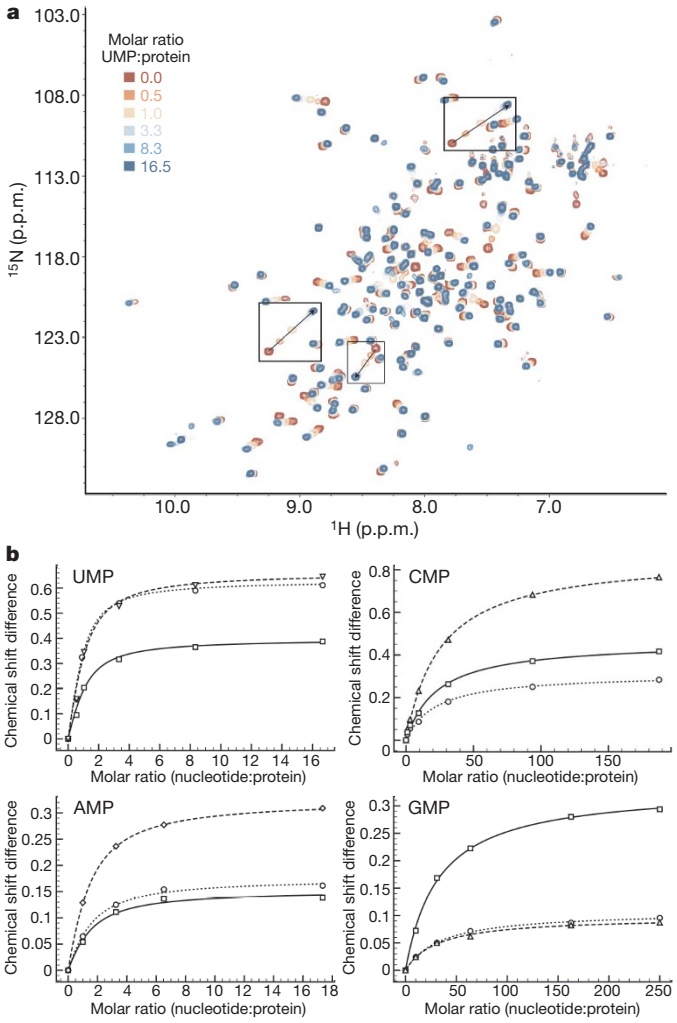

Figure 3 | Determination of dissociation constants using NMR titration experiments. a, Representative ^{15}N HSQC NMR spectra from the MID domain of hAGO2 with increasing amounts of UMP added. Significantly shifting peaks used for the determination of the dissociation constant of UMP are boxed and marked by arrows. **b**, Chemical shift differences were calculated as $\Delta\delta = [(\Delta\delta_H)^2 + (0.2*\Delta\delta_N)^2]^{1/2}$, where $\Delta\delta_H$ and $\Delta\delta_N$ are the observed chemical shift changes for ^{1}H and ^{15}N, respectively. For determination of dissociation constants, $\Delta\delta$ was plotted as a function of the molar ratio (nucleotide:protein) and the data for multiple peaks were fitted using the maximum shift and dissociation constant as adjustable parameters.

the small RNA to interact with specific members of the plant AGO protein family[11–13,22]. A similar observation was also made in the small RNAs that associate with *D. melanogaster* AGO1 and AGO2[5,23]. *A. thaliana* AGOs as well as *D. melanogaster* AGO2 share little overall sequence similarity with hAGO2 and therefore direct extrapolations of the results presented here are not easily made. However, our results indicate that the region corresponding to the nucleotide specificity loop in these proteins will have an important role in determining the sorting of small RNAs via the 5′ nucleotide by distinct AGO proteins, although the exact manner in which it does so awaits further structural analyses. Taken together, the crystal structures of the hAGO2–nucleoside monophosphate (NMP) complexes and the binding assays establish that the MID domain of eukaryotic AGOs specifically recognizes the 5′-nucleotide base of small RNAs and this interaction probably has a key role in RNA silencing.

Table 1 | Dissociation constants of NMPs and the MID domain of hAGO2

	UMP (mM)	AMP (mM)	CMP (mM)	GMP (mM)
Wild type (PGKTP)	0.12 ± 0.01	0.26 ± 0.01	3.63 ± 0.25	3.32 ± 0.14
Loop insertion (PGK**G**TP)	1.18 ± 0.10	1.21 ± 0.08	0.74 ± 0.04	0.9 ± 0.12

METHODS SUMMARY

Protein preparation. The hAGO2 MID domain was cloned into the BamHI and NotI sites of a pSMT3 vector[26], which contains an amino-terminal Ulp1 cleavable His_6-Sumo tag. The protein was bacterially expressed using standard protocols and purified by Ni-affinity chromatography, cleavage of the tag and Superdex-75 size-exclusion chromatography in 25 mM Tris buffer, pH 8.0, 150 mM NaCl, 5% glycerol, and 3 mM dithiothreitol (DTT) (for crystallization), or in 25 mM MES buffer, pH 6.5, 200 mM NaCl, and 3 mM DTT (for NMR). Site-directed mutagenesis was performed using the Quickchange kit (Stratagene).

Crystallization, data collection and structure determination. Crystals (approximately 0.2 mm × 0.2 mm × 0.6 mm in size) of native and SeMet-substituted hAGO2 MID domain (439–578) were grown by hanging-drop vapour diffusion at 4 °C. Protein at 10–15 mg ml^{-1} was mixed 1:1 with a well solution containing 0.1 M imidazole, pH 8.0, 0.2 M NaCl, 0.46 M NaH_2PO_4 and 1.84 M K_2HPO_4. Protein-nucleotide crystals were obtained by soaking native crystals in a drop containing 0.2 M $(NH_4)_2SO_4$, 0.1 M Na Cacodylate buffer, pH 6.5, 15% PEG 8,000, 20% glycerol, and 20 mM NMP. Diffraction data for flash-cooled SeMet-substituted crystals (cryoprotected with 15% glycerol) were collected at the Cornell High Energy Synchrotron Source (CHESS) beamline A1 and processed using HKL2000[27] (Supplementary Table 1). Diffraction data for nucleotide-containing crystals were collected on a Rigaku rotating copper-anode generator. The structure was solved by single-wavelength anomalous diffraction (SAD) using the Solve/Resolve program[28] and refined with Arp/Warp[29] and Phenix[30]. The nucleotide complexes were solved by difference Fourier analysis using the native structure.

NMR titration experiments. All NMR experiments were performed at 293 K using a Bruker 600 MHz spectrometer. NMR titrations were carried out by acquiring ^{1}H–^{15}N HSQC spectra on samples of 0.10–0.25 mM ^{15}N-labelled hAGO2 MID domain (432–578) with the addition of increasing amounts of unlabelled ligand.

Received 16 December 2009; accepted 19 March 2010.
Published online 26 May 2010.

1. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nature Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
2. 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).
3. 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).
4. Ghildiyal, M. *et al.* Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* **320**, 1077–1081 (2008).
5. Ghildiyal, M., Xu, J., Seitz, H., Weng, Z. & Zamore, P. Sorting of *Drosophila* small silencing RNAs partitions microRNA* strands into the RNA interference pathway. *RNA* **16**, 43–56 (2009).
6. Hu, H. Y. *et al.* Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* **10**, doi:10.1186/1471-2164-10-413 (2009).
7. Lau, N. C., Lim, L. P., Weinstein, E. G. & Bartel, D. P. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**, 858–862 (2001).
8. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nature Rev. Genet.* **10**, 94–108 (2009).
9. Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Rev. Genet.* **9**, 102–114 (2008).
10. Seitz, H., Ghildiyal, M. & Zamore, P. D. Argonaute loading improves the 5′ precision of both microRNAs and their miRNA strands in flies. *Curr. Biol.* **18**, 147–151 (2008).
11. Mi, S. *et al.* Sorting of small RNAs into *Arabidopsis* Argonaute complexes is directed by the 5′ terminal nucleotide. *Cell* **133**, 116–127 (2008).
12. Montgomery, T. A. *et al.* Specificity of ARGONAUTE7-miR390 interaction and dual functionality in *TAS3 trans*-acting siRNA formation. *Cell* **133**, 128–141 (2008).
13. Takeda, A., Iwasaki, S., Watanabe, T., Utsumi, M. & Watanabe, Y. The mechanism selecting the guide strand from small RNA duplexes is different among *Argonaute* proteins. *Plant Cell Physiol.* **49**, 493–500 (2008).
14. Ma, J.-B., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* **429**, 318–322 (2004).
15. Wang, Y. *et al.* Structure of an Argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921–926 (2008).
16. Wang, Y. *et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754–761 (2009).
17. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing Argonaute silencing complex. *Nature* **456**, 209–213 (2008).
18. Parker, J. S., Parizotto, E. A., Wang, M., Roe, S. M. & Barford, D. Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Mol. Cell* **33**, 204–214 (2009).

19. 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).
20. Wahl, M. C. & Sundaralingam, M. C–H...O hydrogen bonding in biology. *Trends Biochem. Sci.* **22,** 97–102 (1997).
21. Felice, K. M., Salzman, D. W., Shubert-Coleman, J., Jensen, K. P. & Furneaux, H. M. The 5′ terminal uracil of let-7a is critical for the recruitment of mRNA to Argonaute2. *Biochem. J.* **422,** 329–341 (2009).
22. Wu, L. *et al.* Rice microRNA effector complexes and targets. *Plant Cell* **21,** 3421–3435 (2009).
23. Czech, B. *et al.* An endogenous small interfering RNA pathway in *Drosophila*. *Nature* **453,** 798–802 (2008).
24. Kiriakidou, M. *et al.* An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129,** 1141–1151 (2007).
25. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98,** 10037–10041 (2001).
26. Mossessova, E. & Lima, C. D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell* **5,** 865–876 (2000).
27. Otwinowski, Z. Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* **276,** 307–326 (1997).
28. Terwilliger, T. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59,** 38–44 (2002).
29. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nature Protoc.* **3,** 1171–1179 (2008).
30. Adams, P., Grosse-Kunstleve, R. & Hung, L. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58,** 1948–1954 (2002).

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

Acknowledgements We thank O. Larsson for help with bioinformatic analysis; J.-F. Trempe, G. Kozlov, S. Azeroual and D. Rodionov for help with X-ray and NMR data collection at the QANUC NMR facility; and K. Gehring and A. Berghuis for sharing resources and critical reading of the manuscript. We also thank M. Fabian, T. Sundermeier and T. Duchaine for comments on the manuscript; W. Filipowicz for discussions; R. Szittner, K. Illes and G. Virgili for technical assistance. B.N. is supported by a Canada Research Chair, a Career Development Award from the Human Frontiers Science Program (CDA 0018/2006-C/1) and an operating grant from the Canadian Institutes of Health Research (CIHR grant MOP-82929). N.S. is funded by a CIHR grant. F.F. is supported by a Boehringer Ingelheim Fonds PhD Fellowship.

Author Contributions B.N., N.S. and F.F. designed the project. B.N. and F.F. wrote the manuscript. F.F. performed all of the NMR and crystallographic work.

Author Information Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 3LUC, 3LUD, 3LUG, 3LUH, 3LUJ and 3LUK. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to B.N. (bhushan.nagar@mcgill.ca).