HHS Public Access

Author manuscript

Nature. Author manuscript; available in PMC 2022 July 21.

Published in final edited form as:

Nature. 2021 September ; 597(7875): 285–289. doi:10.1038/s41586-021-03856-x.

Structural basis for piRNA-targeting

Todd A. Anzelon[1,†], **Saikat Chowdhury**[1,2,†], **Siobhan M. Hughes**[1,†], **Yao Xiao**[1], **Gabriel C. Lander**[1], **Ian J. MacRae**[1,*]

[1]Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA

[2]Current address: CSIR-Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500007, India

Summary

Piwi proteins use PIWI-interacting RNAs (piRNAs) to identify and silence transposable elements (TEs) and thereby maintain genome integrity between metazoan generations[1]. TE-targeting by Piwi has been compared to mRNA target recognition by Argonaute proteins[2,3], which employ microRNA (miRNA) guides, but the extent to which piRNAs resemble miRNAs is not known. We present cryo-EM structures of a Piwi-piRNA complex from the sponge *Ephydatia fluviatilis* with and without target RNAs and biochemical analysis of target recognition. Mirroring Argonaute, Piwi identifies targets using the piRNA seed-region. However, Piwi creates a much weaker seed so that stable target association requires further piRNA-target pairing, making piRNAs less promiscuous than miRNAs. Beyond the seed, Piwi structure facilitates piRNA-target pairing in a manner tolerant of mismatches, leading to long-lived Piwi-piRNA-target interactions that may accumulate on TE transcripts. Piwi ensures targeting fidelity by physically blocking propagation of piRNA-target interactions in the absence of faithful seed pairing, and by requiring an extended piRNA-target duplex to reach an endonucleolytically active conformation. Piwi proteins thereby minimize off-targeting cellular mRNAs while defending against evolving genomic threats.

Introduction

Most animals produce at least two types of small regulatory RNAs: microRNAs (miRNAs) and Piwi-interacting RNAs (piRNAs)[4]. miRNAs regulate mRNAs during development while piRNAs protect the germline from transposable elements (TEs). On the molecular level, miRNAs and piRNAs function as guides for Argonaute (AGO) and PIWI proteins,

*Correspondence to: macrae@scripps.edu (I.J.M.).
†These authors contributed equally to this work.

Author Contributions

T.A. prepared EfPiwi-piRNA, Siwi-piRNA, and hAgo2-miRNA samples, performed biochemical experiments, built EfPiwi models, and co-wrote the manuscript. S.C. prepared cryo-EM samples, collected data, produced high resolution reconstructions, and assisted with model building. S.M.H. identified and developed EfPiwi as a source of active PIWI protein. Y.X. helped develop EfPiwi and established purification protocols. G.C.L. provided structural insights and guidance in cryo-EM data collection and analysis. I.J.M. provided structural and mechanistic insights and co-wrote the manuscript.

Competing financial interest

The authors declare no competing financial interests.

respectively, which use sequence information in small RNAs to identify transcripts targeted for repression.

Extensive studies have shown miRNA-class AGO proteins tune the binding properties of miRNAs to recognize short segments of complementarity on target mRNAs[5–8]. AGO creates the miRNA seed region by pre-organizing nucleotides 2–7, lowering the entropic cost of target binding[9,10], and adopts a structure that discourages miRNA-target pairing immediately after the seed[11,12]. miRNA-recognition sites are thus short and precisely defined, allowing individual miRNAs to recognize hundreds of mRNAs and collectively target over half of the mRNAs in mammals[13].

piRNAs have been proposed to use a miRNA-like targeting mechanism[2,3]. Indeed, seed complementarity is a feature of piRNA targets[2,3,14–17]. However, the relationship between piRNAs and their targets is fundamentally distinct from that of miRNAs. miRNA targets evolve to be recognized to benefit developmental fitness[18]. In contrast, piRNA targets are parasitic genetic elements under selective pressure to escape recognition, suggesting a benefit for mechanisms adapting piRNAs to evolving threats[19]. Additionally, the piRNA repertoire in animals is usually orders of magnitude larger than that of miRNAs[20], indicating that mechanisms must be in place to avoid silencing the entire germline transcriptome.

Crystal structures of insect PIWI proteins, *Drosophila* Piwi (DmPiwi) and silkworm Siwi, also suggest piRNAs may be distinct from miRNAs[21,22]. These PIWIs have a 3D domain arrangement differing from AGO proteins, indicating piRNA-target interactions are shaped in a unique fashion. Previous studies used PIWI samples containing heterogenous mixtures of co-purifying piRNAs, preventing investigation into piRNA-target interactions. Thus, mechanisms underlying TE recognition by piRNAs, as well as the extent to which piRNA-targeting resembles that of miRNAs, remain unknown.

Source of homogenous PIWI-piRNA complexes

PIWI proteins from natural sources co-purify with heterogenous mixtures of endogenous piRNAs[21,22], which obstruct analysis of targeting. We therefore sought a recombinant system for preparing homogenous PIWI-piRNA complexes. A screen of PIWI protein constructs from various animals revealed that Piwi-A from the freshwater sponge *Ephydatia fluviatilis*[23] (hereafter referred to as EfPiwi) can be expressed recombinantly, loaded with a chemically defined piRNA, and purified as a stable PIWI-piRNA complex. EfPiwi is normally expressed in *E. fluviatilis* totipotent archeocytes[24]. We also identified a construct to reconstitute the Siwi-piRNA complex, albeit at lower levels than EfPiwi (Extended Data Fig. 1a,b).

Unique structural features of the extended PIWI family

EfPiwi belongs to an ancient branch of the PIWI family tree that contains *Drosophila* Ago3, MILI, HILI, and ZILI, and is distinct from the Siwi/DmPiwi branch[25] (Extended Data Fig. 1c). Thus, structural features common to EfPiwi and Siwi/DmPiwi may be broadly conserved in the PIWI family. We used cryo-EM single-particle analysis to determine the structure of the EfPiwi-piRNA complex to ~3.8 Å resolution (Fig. 1a,b, Extended Data

Table 1, and Extended Data Fig. 2). Comparison of metazoan PIWI and AGO structures revealed two features likely to impact targeting by members of the extended PIWI family:

First, although our sample contained a homogenous piRNA, most nucleotides are disordered, including the 3' half of the seed region (guide (g) nucleotides g5–g7) (Fig. 1c). AGO proteins pre-organize the seed 3' end by cradling g5-g6 in a kinked loop that is widely conserved in AGOs. The equivalent loop in PIWIs contains bulky residues that prevent the kink necessary to cradle g5-g6 (Extended Data Fig. 3a,b). Thus, unlike AGOs, PIWIs do not pre-organize the full piRNA seed region.

Second, as noted previously[21] the three-way interface of the L1, L2 and PAZ domains in PIWIs differs from AGOs. An important consequence of this difference is that residues corresponding to the AGO "central-gate", which restricts guide-target interactions in the miRNA central region[11,12], instead form a small α-helix (helix-6) near the piRNA seed region (Fig. 1b, Extended Data Fig. 3c–g). PIWIs thus lack a central-gate, and instead have a widened central cleft and extended "seed-gate" structure (Fig. 1d).

piRNAs are more selective than miRNAs

To determine how structural differences impact target recognition, we compared target RNA binding properties of EfPiwi, Siwi, and human Argonaute2 (hAgo2), a miRNA-class AGO protein. EfPiwi and Siwi bound target RNAs with complementarity limited to the seed region (g2-g7) with >150-fold lower affinity than hAgo2, loaded an equivalent guide RNA (Fig. 2a,b). Similarly, both PIWIs bound a target with extended seed (g2-g8) complementarity with >7-fold lower affinity than hAgo2. Weak seed-pairing appears to arise from higher target off-rates from the PIWIs, which released the g2–g8 matched target RNA ~7-fold faster than hAgo2 (Fig. 2c). Extending target complementarity through the guide central region (g9–g12) had little effect on hAgo2 release rates, consistent with the model that hAgo2 avoids central pairing due to its central-gate (Fig. 2d, Extended Data Fig. 4)[11,12]. In contrast, for EfPiwi and Siwi target release rates dropped to below 1×10^{-3} min^{-1} ($t_{1/2}$ > 12 hours) as guide-target complementarity extended through the central region (Fig. 2d, Extended Data Fig. 4).

We propose differences in PIWI and AGO structure create distinct targeting properties for piRNAs and miRNAs. PIWIs pre-organize a minimal seed, making the piRNA seed far weaker than the miRNA seed. However, unlike AGOs, PIWIs lack a central gate and thus can compensate for weak seed-pairing by extending piRNA-target interactions into the central region. piRNAs are thereby more selective than miRNAs when identifying targets and can remain associated with recognized targets for much longer periods of time.

The seed-gate enforces piRNA target-binding fidelity

Although the piRNA seed is weak, perfect seed complementarity is required for piRNA target recognition *in vivo*[2,3,14–17]. To understand this requirement, we determined the structure of the EfPiwi-piRNA complex bound to a target RNA with complementarity to piRNA nucleotides g2-g16 to ~3.5 Å resolution (Fig. 3a,b, Extended Data Table 1, Extended Data Fig. 5). The reconstruction contains a conspicuous 15 bp piRNA-target RNA duplex (Fig. 3a), which appears to drive EfPiwi into a more opened conformation (Fig. 3c). EfPiwi

opening involves a shift in the seed-gate, which moves ~12 Å to avoid clashing with base pairs 5–8 of the piRNA-target duplex (Fig. 3d) and docks with the duplex at the 3’ end of the seed (Fig. 3e). These contacts, along with contacts from the MID and PIWI domains, probe the piRNA-target duplex back bone and minor grove, likely enforcing ideal pairing to the piRNA seed.

Notably, piRNA-target mismatches in the 3’ half of the seed (g5–g7) nearly abolish target-binding, reducing k_{on} >2000-fold compared to a target with no mismatches (Fig. 3f, Extended Data Fig. 6). Binding to the g5-g7 mismatched target was restored by removing the seed-gate (Δseed-gate, L520–H537 replaced with a Gly_6 linker), which increased k_{on} 14-fold compared to wild-type EfPiwi (Fig. 3f). Δseed-gate EfPiwi bound a target with mismatches at the seed 5’ end (g2–g4) at similar rate to wild-type EfPiwi. In contrast, Δseed-gate EfPiwi bound all targets with intact seed complementarity about 10-fold slower than wild-type. Δseed-gate Siwi behaved similarly (Extended Data Fig. 7a,b). Thus, the seed-gate enforces targeting fidelity by facilitating propagation of piRNA-target pairing initiated at the seed 5’ end, and by inhibiting duplex propagation in the absence of faithful pairing to the seed 3’ end.

Piwi accommodates mismatches beyond the seed

Beyond the seed-paired region, EfPiwi contacts the piRNA-target duplex backbone but does not interrogate the shape of the minor groove (Extended Data Fig. 7c). EfPiwi may thereby recognize the overall helical structure of the duplex in a manner that is tolerant of deviations from A-form geometry and thus insensitive to intermittent mismatches beyond the seed. Indeed, 3 nt mismatched segments after the seed have only minor effects (≤ 2.5-fold) on target binding rates (Fig. 3f) and most bound targets dissociate over the course of hours (Extended Data Fig. 6b and 7d,e). Thus, EfPiwi can engage targets containing ≤3 consecutive non-seed mismatches with exquisite affinity (K_D values in the pM–fM range) (Extended Data Fig. 7f). Tolerance of mismatches beyond the seed may allow piRNAs to recognize related or evolving TEs.

Extensive piRNA 3’ pairing activates target cleavage

PIWIs often endonucleolytically cleave target RNAs[21,22,26]. Cleavage activity of both EfPiwi and Siwi is far weaker than hAgo2 but can be stimulated by Mn^{2+} and increasing temperature (Extended Data Fig. 8a–e), allowing us to characterize target substrate preferences. EfPiwi cleavage activity increased as complementarity extended towards the piRNA 3’ end, reaching a plateau when extended to g18 (Fig. 4a, Extended Data Fig. 9a,b). Siwi cleavage also increased with 3’ complementarity, plateauing at g16 pairing (Extended Data Fig. 9e,f).

We examined how mismatches towards the piRNA 3’ end influence cleavage (Extended Data Fig. 9c,d). A pool of 256 target RNA sequences with combinations of mismatches to g11–g18 was treated with excess EfPiwi-piRNA. RNA-seq of cleavage products revealed the top 10% (26 out of 256 sequences) accounted for 80% of all cleavage events and was devoid of mismatched segments ≥3 bp in length (Fig. 4b, Extended Data Fig. 9f). Thus, EfPiwi

nuclease activity is driven by extended 3’ pairing, optimally with ≤2 mismatches towards the piRNA 3’ end.

Extended piRNA-target pairing drives EfPiwi into an open conformation

We also examined the structure of EfPiwi-piRNA bound to a target RNA with complementarity to g2-g25 by cryo-EM. While the piRNA-target duplex was obvious in 2D averages, EfPiwi appeared smaller than expected (Fig. 4b). An ~7.0 Å resolution 3D reconstruction was consistent with the MID and PIWI domains bound to the end of the piRNA-target duplex (Fig. 4c). Docking the full EfPiwi-piRNA-target model (with pairing to g2-g16) shows that, after position 15, the extended piRNA-target duplex sterically clashes with segments of the N, L1, and PAZ domains (Fig. 4d). Thus, extended piRNA-target pairing drives central cleft opening to such an extent that the entire N-L1-PAZ-L2 lobe becomes conformationally uncoupled from the MID-PIWI lobe.

We suggest that EfPiwi nuclease activity may be activated by widening of the central cleft, driven by formation of an extended piRNA-target duplex. Mismatched segments ≥3 nt in length may impart flexibility to the guide RNA-target duplex, reducing the propensity to drive open the cleft and stimulate cleavage. piRNA cleavage products identified in mouse testes are generally free of mismatched segments ≥3 nt in length[15,17,27], and target cleavage by immuno-purified Miwi also required extended piRNA-target complementarity[26].

Discussion

piRNAs and miRNAs are ancient genetic regulators that may have helped usher in the era of multicellular animal life[4]. Many animals also produce small interfering RNAs (siRNAs), which typically associate with a separate siRNA-class of AGO proteins[25] that remains uncharacterized at the structural level. Our results show PIWI and miRNA-class AGO proteins possess distinct structural features, enabling piRNAs and miRNAs to carry out discrete roles in animal evolution.

The primary piRNA function is to seek out and silence TEs. Considering the immense diversity of the piRNA repertoire, targeting must be stringent enough to avoid inadvertently silencing cellular RNAs. EfPiwi accomplishes this by creating a weak seed and closely monitoring pairing via the seed-gate. Seed-complementarity is thus necessary but insufficient for target recognition. Stringency is further established for target cleavage by the requirement of a piRNA-target duplex strong enough to activate the endonuclease mechanism. Stringency may also be regulated by a recently discovered cleavage stimulating factor[28]. Selectivity at the target cleavage step may help guard against cellular RNAs entering piRNA biogenesis pathways via the ping-pong cycle[1].

On the other hand, flexibility in target recognition would allow piRNAs to respond to evolving TE sequences[19]. Targeting flexibility, in terms of binding, is provided beyond the seed, where the PIWI central cleft is tolerant of helical imperfections arising from piRNA-target mismatches such that high-affinity binding requires less complementarity than cleavage *in vitro*. A striking feature of the Piwi-piRNA complex is that release of bound target molecules is exceptionally slow, on par with the lifetimes of the most

stable mRNAs in ES cells[29]. Thus, once Piwi engages a target it could potentially remain bound for the remainder of the transcript's existence. We propose TE recognition may involve accumulation of multiple Piwi-piRNA complexes on each target transcript, leading to multivalent assemblies that recruit histone/DNA methylation factors to TE loci in the nucleus[30–33], or traffic cytoplasmic transcripts to phase-separated compartments associated with silencing and piRNA production[34]. We suggest that by shaping piRNA-target interactions PIWIs can leverage the vast piRNA pool to restrict TE escape from surveillance while minimizing off-targeting, and have thereby maintained metazoan germlines over the last 800 million years.

Methods

Recombinant PIWI cloning, mutagenesis, and expression

DNA fragments encoding truncations of PIWI proteins from various animals (including human, silk moth, fruit fly, worm, zebra fish, flour beetle, and sponge) cloned into a modified form of pFastBac HTA (Thermo Fisher) to generate expression plasmids for the Bac-to-Bac baculovirus expression system (Thermo Fisher). Resulting baculoviruses were used to infect Sf9 cells and recombinant protein expression levels were assessed 3 days post infection by Ni-NTA purification (Qiagen) and SDS PAGE. The most highly expressed constructs for the top two PIWIs were N-terminal truncations of *Ephydatia fluviatilis* Piwi A (NCIB: AB533505, residues 219–987) and *Bombyx mori* Siwi (NCIB: AB332313, residues 88–899).

Preparation of protein-guide RNA complexes

EfPiwi-guide, Siwi-guide, and hAgo2-guide RNA samples were purified as described previously for hAgo2[35], which is a variation on the Arpon method[36], with exception that the final size-exclusion purification step was omitted from PIWI preparations to maintain sufficient yields. Purified PIWI-guide complexes were stored at −80°C at a protein concentration range of 5–50 μM in 150 mM NaCl, 0.5 mM TCEP, 5–10% glycerol, 20 mM Tris pH 8.

Grid preparation for cryo-EM

EfPiwi-piRNA-target ternary complexes were formed by adding 1.2 molar equivalents of target RNA to purified EfPiwi-guide complex and incubating on ice for 10 minutes in the following buffer: 150 mM NaCl, 0.5 mM TCEP, 50 mM Tris, pH 8. Divalent ions were excluded to prevent target cleavage. Nonetheless, final models include two Mg2+ ions (one bound to the 5' phosphate and one in the endonuclease active site), which appear to have co-purified with EfPiwi. Additionally, for the extended (2–25 paired) target, the procedure used to form the ternary complex appears to have been sub-optimal as many particles were not bound to the target RNA (possibly connected to reduced EfPiwi activity at low temperatures). This turned out to be a stroke of good luck as it also enabled determination the EfPiwi-piRNA binary structure (shown in Fig. 1a) from the same data set. To prepare samples for cryo-EM, 3.5 μl EfPiwi-guide-target complex at 2.5 mg/ml was added onto freshly plasma cleaned (75% nitrogen, 25% oxygen atmosphere at 15 W for 7 seconds in Solarus plasma cleaner, Gatan) 300 mesh holey gold grids (UltrAuFoil R1.2/1.3, Quantifoil).

Excess sample solution was removed from grids by blotting with Whatman No.1 filter paper for 5–7 seconds. Samples were immediately vitrified by plunge freezing in liquid-ethane at −179°C using a manual plunge freezing device. Grid vitrification was performed in a cold room maintained at 4°C with relative humidity between 95–98% to minimize sample evaporation.

Cryo-EM data acquisition

Cryo-EM data were acquired on a 200kV Talos Arctica (Thermo Fisher Scientific) transmission electron microscope. Micrographs were acquired using a K2 Summit (Gatan) direct electron detector, operated in electron-counting mode, using the automated data collection software Leginon[37] by image shift-based movements from the center of four adjacent holes to target the center of each hole for exposures. Each micrograph for the EfPiwi-piRNA-target (2–25) complex (which fortuitously also contained many EfPiwi-piRNA binary complexes) was collected as 48 dose-fractionated movie frames over 12 s and with a cumulative electron exposure of 47.33 $e^{-}/Å^{2}$. For the EfPiwi-piRNA-target (2–16) complex, each micrograph was acquired as 64 dose-fractionated movie frames over 16 s with a cumulative electron exposure of 47.33 $e^{-}/Å^{2}$. Both data sets were collected at a nominal magnification of 36kx, corresponding to 1.15 Å/pixel on the detector, with random nominal defocus values varying between 1 μm and 1.6 μm. 1,765 and micrographs were collected for the EfPiwi-piRNA EfPiwi-piRNA-extended target (pairing to 2–25) complex (these micrographs also contained the EfPiwi-piRNA binary complex). 1,881 micrographs were collected for the EfPiwi-EfPIWI-piRNA-target (pairing to 2–16) complex.

Image processing and 3D reconstruction

Beam-induced motion correction and radiation damage compensation over spatial frequencies (dose-weighting) of the raw movies, was performed using UCSF MotionCor2[38] implemented in the Appion[39] image processing workflow. Motion corrected, summed micrographs were imported into the RELION 2.0[40] data processing pipeline. Contrast Transfer Function (CTF) parameters for these micrographs were estimated using CTFFind4[41]. Laplacian of Gaussian based automated particle picking program in RELION was used for picking 3,280,351 and 2,551,046 particles from the EfPiwi-piRNA and EfPiwi-piRNA-target micrographs, respectively. Picked particles were extracted from the micrographs with a 160 pixel box and subjected to 2D classification in RELION. After discarding particles belonging to classes containing non-particle features, aggregates and low-resolution features, new stacks of particles from 2D classes containing different orientations of the complexes and high-resolution features were created. A subset of 2D classes from the EfPiwi-piRNA-target (g2-g25 paired) dataset resolved features corresponding to a smaller complex with an extended RNA duplex. The 608,488 particles belonging to these classes were isolated into a new particle-stack for further processing. To serve as an initial model for 3D analyses of the EfPiwi-piRNA and EfPiwi-piRNA-target complexes, a 40 Å low pass filtered map was generated from the Siwi crystal structure (PDB ID 5GUH) using the molmap function in UCSF Chimera[42,43]. An initial model for 3D analyses of the smaller subcomplex particles was generated using the Cryosparc v1[44] *ab initio* reconstruction program. The selected particle stacks corresponding to the three distinct complexes were subjected to multiple iterations of 3D classification in RELION and

particles belonging to the most well-resolved 3D class for each complex were selected for downstream 3D processing. After 3D classification, 125,041, 118,493, and 116,655 particles from the best resolved class for the EfPiwi-piRNA, EfPiwi-piRNA-target (g2-g16 paired), and the smaller EfPiwi-piRNA-target (g2-g25 paired), respectively, were re-extracted with 160 pixels box from the respective micrographs with re-centered coordinates. These particles were then subjected to 3D refinement in RELION. 3D binary masks for refinement were generated using 15 Å low-pass filtered selected class volume for each of the complexes with a 5-pixel expansion and 8-pixel Gaussian fall-off in RELION. The final reconstructed map for the EfPiwi-piRNA, EfPiwi-piRNA-target (g2-g16 paired), and EfPiwi-piRNA-target (g2-g26 paired) were at 3.8 Å, 3.5 Å and 8.6 Å (at Fourier Shell Correlation value of 0.143), respectively. Local resolution for these complexes was determined using the local resolution estimation program in RELION and the local-resolution based filtered maps were used for atomic model building. Maps for EfPiwi-piRNA and EfPiwi-piRNA-target complexes were trimmed to box size of 90 pixels for deposition into EM Data Bank. Directional Fourier Shell Correlations (FSCs) were estimated using the 3DFSC server (www.3dfsc.salk.edu) [45].

Model building and refinement

An initial model for the EfPiwi was obtained by threading the EfPiwi primary sequence onto the Siwi crystal structure (PDB ID 5GUH) using SWISS-MODE[46]. Discrete domains from were then docked into the EfPiwi-piRNA reconstruction using UCSF Chimera, followed by manual model building using Coot[47]. The EfPiwi-guide-target model was built in a similar fashion, using the EfPiwi-guide structure as an initial model. Models were refined through iterative rounds of manual building and fixing of geometric and rotameric outliers in Coot and real-space refinement optimizing global minimization, atomic displacement parameters, and local grid search using PHENIX[48]. Base pairing between guide and target RNAs was maintained by including hydrogen atoms. Most residues within the N domain of both models were truncated to alanine at the end of refinement to reflect a lack of supporting cryo-EM density. Model validation were performed using Molprobity (molprobity.biochem.duke.edu)[49] and PDB validation servers (www.wwpdb.org). Structural figures were made using PyMOL (Schrödinger, LLC) and UCSF ChimeraX[50]. Prediction of seed-gate structure in extended PIWI family members was performed by PSIPRED 4.0[51,52]. EfPiwi structure was compared to previous human AGO structures[53–55] in Fig. 1.

Equilibrium target RNA binding assays

Equilibrium dissociation constants for seed-matched target RNAs were determined as described previously[11]. Dissociation constants for targets with extensive complementarity were too small to be accurately measured and were thus calculated as the quotient of measured target dissociation (k_{off}) and association (k_{on}) rates.

Target RNA dissociation assays

A dot-blot apparatus (GE Healthcare) was used to separate protein-RNA complexes from free (unbound) RNA: protein-RNA complexes were immobilized on Protran nitrocellulose membrane (0.45 μm pore size, Whatman, GE Healthcare) and unbound RNA immobilized on Hybond Nylon membrane (Amersham, GE Healthcare). Membranes were stacked so that

sample is first pulled through the protein-binding membrane, and any unbound RNA passes through and binds to the RNA-binding membrane.

Target dissociation rates were determined by incubating guide-loaded EfPiwi, Siwi, or hAgo2 samples with 0.1 nM ^{32}P 5′-radiolabeled target RNA in binding reaction buffer (30 mM Tris pH 8.0, 100 mM potassium acetate, 0.5 mM TCEP, 0.005% (v/v) NP-40, 0.01 mg/mL baker's yeast tRNA) in a single reaction with a volume of 100 μl per time point planned for the experiment (e.g. 1000 μl for 10 time points) at room temperature for 60 minutes. The concentration of hAgo2, EfPiwi, or Siwi guide RNA complex was 5 nM.

After sample equilibration, a zero-time point was taken by applying 100 μl of the reaction to the dot-blot apparatus under vacuum, followed by 100 μl of ice-cold wash buffer (30 mM Tris pH 8.0, 0.1 M potassium acetate, 0.5 mM TCEP). The dissociation time course was started by addition of 300 nM (final concentration) unlabeled target RNA. Aliquots of 100 μl were taken at various times and immediately applied to a dot-blot apparatus under vacuum, followed by 100 μl of ice-cold wash buffer. Time points ranged from 0.25–100 minutes. Membranes were air-dried and visualized by phosphorimaging. Quantification of ^{32}P signal was performed using ImageQuant TL (GE Healthcare). The fraction of target RNA bound was calculated as the ratio of bound to total (bound + free) target RNA for various concentrations of Ago2-guide, EfPiwi-guide, or Siwi-guide complexes. Dissociation rates were calculated by plotting data as fraction bound versus time and fitting to a one-phase exponential curve using Prism version 8.0 (GraphPad Software, Inc.).

Target RNA association assays

Target association rates were determined by incubating 5 pM ^{32}P 5′-radiolabeled target RNA in binding reaction buffer (30 mM Tris pH 8.0, 100 mM potassium acetate, 0.5 mM TCEP, 0.005% (v/v) NP-40, 0.01 mg/mL baker's yeast tRNA) in a single large reaction with a volume of 100 μl per time point planned for the experiment (e.g. 1000 ul for 10 time points) at room temperature for 15 minutes.

A zero-time point was taken by applying 100 μl of the reaction to a dot-blot apparatus under vacuum, followed by 100 μl of ice-cold wash buffer (30 mM Tris pH 8.0, 0.1 M potassium acetate, 0.5 mM TCEP). The association reaction was started by addition of guide-loaded EfPiwi, or Siwi at concentrations ranging from 0.05–15 nM. Aliquots of 100 μl were taken at various time points and applied to a dot-blot apparatus under vacuum, followed by 100 μl of ice-cold wash buffer as before. Time points ranged from 0.25–15 minutes. Membranes were air-dried and visualized, signal quantified, and bound fraction calculated as described above. Because target dissociation was very slow compared to association, and the target RNA concentrations were always at least 10-fold less than total EfPiwi-piRNA or Siwi-piRNA concentration, association was treated as an irreversible pseudo first-order system. Data were fit to a one-phase exponential curve using Prism version 8.0 (GraphPad Software, Inc.) to determine an observed binding rate constant (k_{obs}). Observed k_{obs} values were proportional to protein concentration over the range used in our experiments. The association rate (k_{on}) was calculated by dividing k_{obs} by the concentration of the EfPiwi-piRNA or Siwi-piRNA complex used in the experiment.

Target RNA cleavage assays

Purified EfPiwi-guide, Siwi-guide, or hAgo2-guide complexes (100 nM, final concentration) were incubated at 37°C with complementary 5′-^{32}P-labeled target RNAs (10 nM, final concentrations) in reaction buffer composed of 30 mM Tris pH 8.0, 2 mM $MgCl_2$, 0.5 mM TCEP, and 0.01 mg/mL baker's yeast tRNA. Target cleavage was stopped at various times by mixing aliquots of each reaction with an equal volume of denaturing gel loading buffer (98% w/v formamide, 0.025% xylene cyanol, 0.025% w/v bromophenol blue, 10 mM EDTA pH 8.0). Intact and cleaved target RNAs were resolved by denaturing PAGE (15%) and visualized by phosphorimaging. Quantification of signal was performed using ImageQuant TL (GE Healthcare).

Adjustments were made to this protocol to determine which divalent cations were catalytic with EfPiwi or Siwi, by replacing 2 mM $MgCl_2$ with 2 mM $MnCl_2$, $CoCl_2$, $CaCl_2$, or $NiCl_2$. Similar adjustments were made to titrate in $MgCl_2$ or $MnCl_2$.

To determine the piRNA-target pairing length requirement for cleavage, 100 nM protein-guide complex and 1 nM radiolabeled target RNA of complementarity g2-g15, g2-g16, g2-g17, g2-g18, g2-g19, g2-g20, or g2-g21 were combined in buffer with a mixture of 2 mM $MgCl_2$ + 2 mM $MnCl_2$. Reactions proceeded for 1 hour at 37 °C before an equal volume of denaturing gel loading buffer was added to stop the reaction. Reactions were run on denaturing PAGE (15%) and analyzed as described above.

To determine the effects of the mismatch triplets on slicing, 100 nM protein-guide complex and 1 nM radiolabeled target RNA with mismatches from g2-g4, g5-g7, g8-g10, g11-g13, g14-g16, g17-g19, or g20-g22 were combined in buffer with 2 mM $MgCl_2$ + 2 mM $MnCl_2$. Reactions proceeded for 1 hour at 37 °C before an equal volume of denaturing gel loading buffer was added to stop the reaction. Reactions were run on denaturing PAGE (15%) and analyzed as described above. Uncropped and unprocessed scans of gels and quantification of replicate data can be found in the Source Data file.

Pooled target cleavage assay

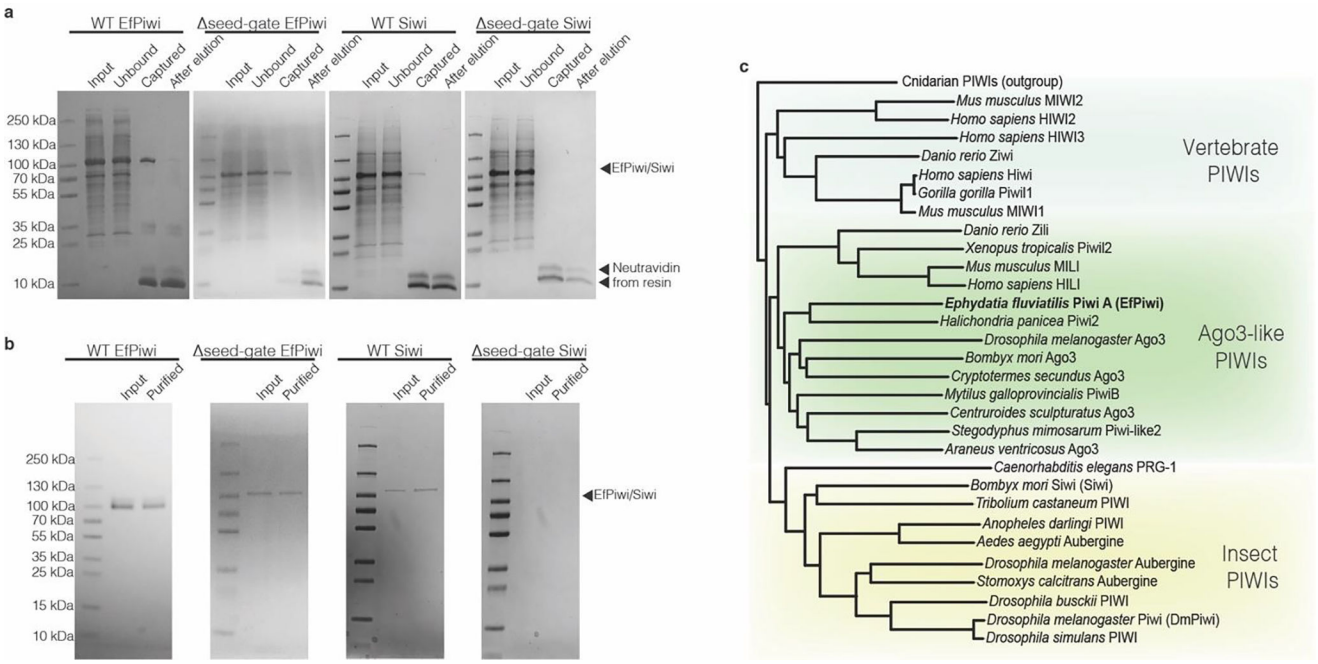
A pool of 256 target RNAs containing molecules with combinations of mismatched pairing to g11-g18 of our standard piRNA sequence was synthesized by Integrated DNA Technologies. Synthesis included piRNA-matched and mismatched phosphoramidites mixed in a 1:1 ratio at steps generating t11–t18, resulting in 256 unique target sequences. All mismatches are expected to reside on the 5' cleavage product because EfPiwi cleaves targets between t10-t11.

The target RNA pool (27 nM) was incubated with EfPiwi-miR122 25mer complex (133 nM) in cleavage buffer (100 mM NaCl, 20 mM Tris pH 8, 2 mM $MgCl_2$, 2mM $MnCl_2$, 0.05 mg/ml tRNA, 0.5 mM TCEP) for 60 minutes at 37°C. Cleavage products from the 5' end of the target (complementary to the piRNA 3' end) were isolated by denaturing gel electrophoresis, cut out of the gel, solubilized, and purified by ethanol precipitation followed by passage through an Oligo Clean and Concentrator column (Zymo Research). Following this step, cleavage products had 5' and 3' adapters ligated with the T4 RNA Ligase and T4 RNA Ligase 2 Truncated enzymes (New England Biolabs), respectively and were subjected

to a reverse transcriptase reaction using the SuperScript II reverse transcriptase (Lifetech). Resulting cDNA was sequenced on a NextSeq 500 (Illumina). To control for variations in 3' adapter ligation efficiency to the cleaved sequences, a "pre-cleaved" control was used, which was a pool of 256 RNAs synthesized to mimic all possible 5' cleavage products, but with a three-base barcode distinct from the intact target sequences. Experimental and pre-cleaved target pools were combined before adapter ligation, reverse transcription, and sequencing.

Following sequencing, unique reads were trimmed of the adapter sequences and those of the correct read length were retained using cutadapt[56]. Remaining reads were sorted into two subsets, experimental or ligation ("pre-cleaved") control, using the three-base barcodes. A read count analysis was performed on each subset, and counts were normalized by dividing the experimental counts by the pre-cleaved counts and multiplying by 1000 so that total counts for each observed sequence was ≥1. Targets were then sorted by abundance, resulting in the histogram in Fig. 4c.

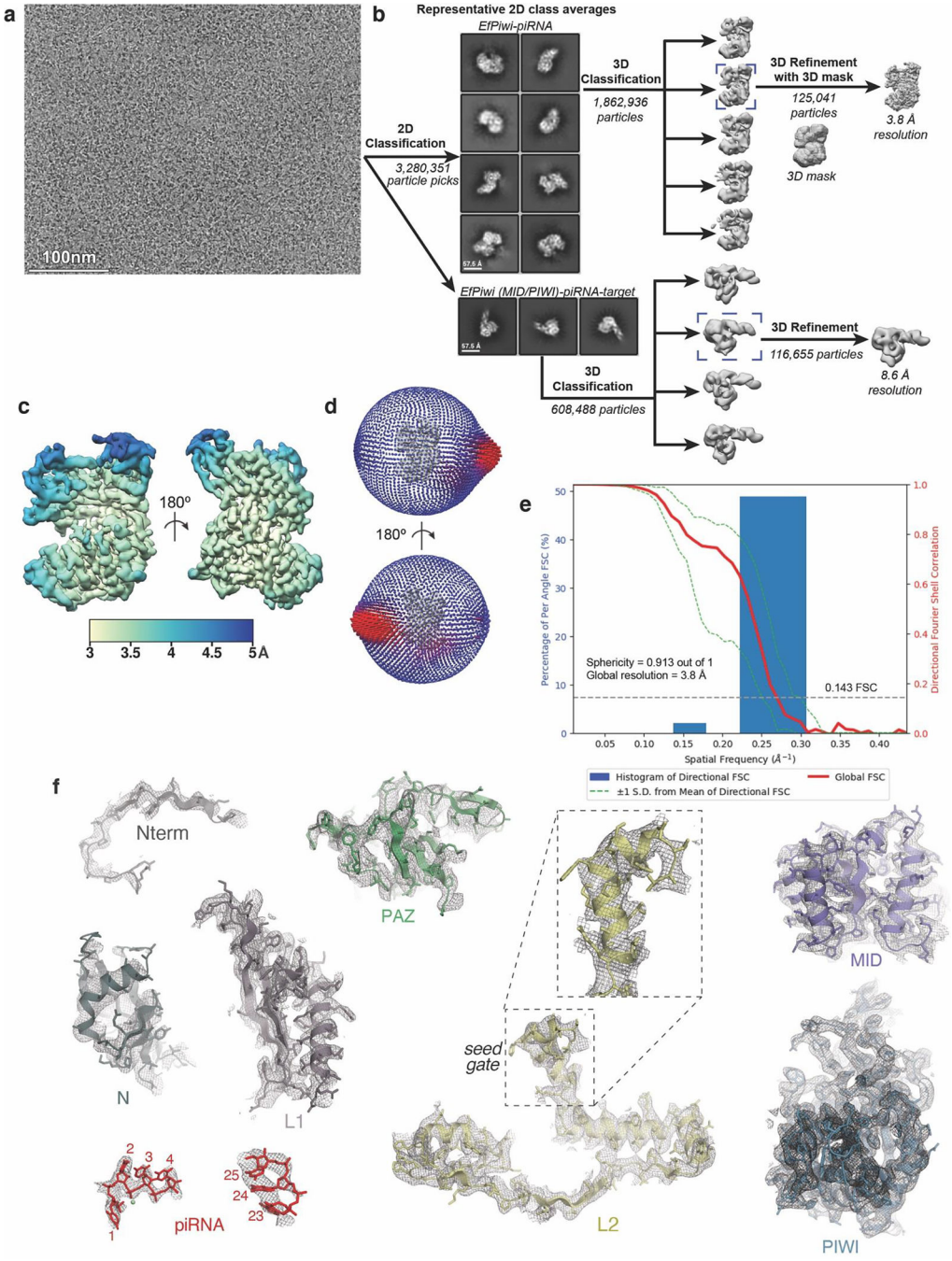
Extended Data



Extended Data Fig. 1. PIWI protein purification and extended PIWI family tree.

a, Coomassie-stained SDS PAGE of piRNA-loaded PIWI proteins captured using an immobilized complementary oligonucleotide. *Input* shows partially purified protein samples that were incubated with capture resin. *Unbound* shows protein that did not bind the resin. *Captured* shows protein retained on the resin after washing (eluted by boiling in SDS). *After elution* shows protein retained on the resin after incubation with the competitor oligonucleotide (eluted by boiling in SDS). **b,** capture-purified PIWI proteins before and after anion exchange purification. Input fraction shows samples after elution by competitor oligonucleotide in capture-purification step. Purified indicates the final purification products. Note: Δseed-gate Siwi was captured at such low levels that it was unclear whether any

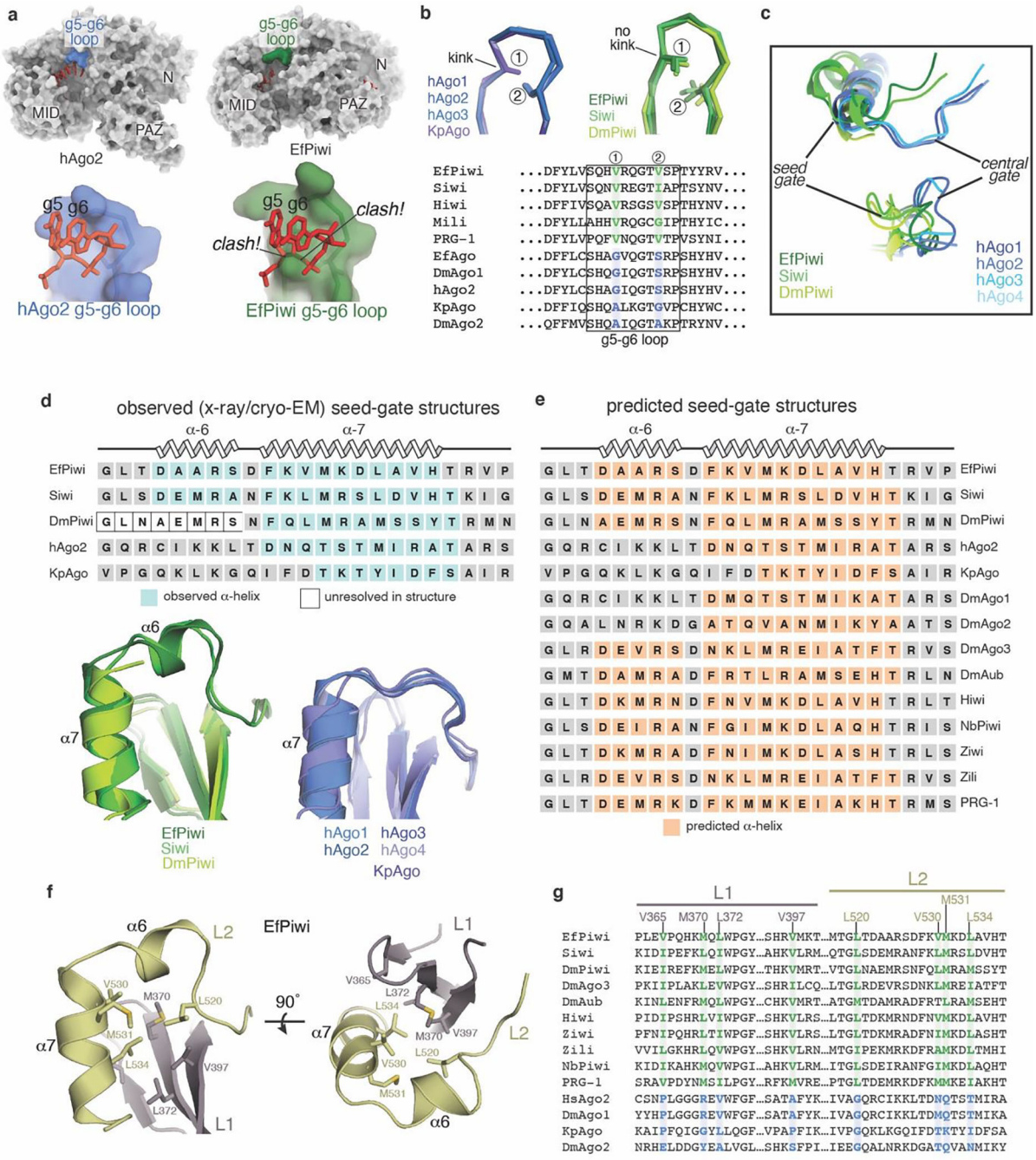
active Siwi was obtained until observing the sample's ability to specifically bind ^{32}P-labeled target RNAs. **c,** phylogenetic tree of PIWI proteins shows EfPiwi belongs to the ancient *Drosophila* AGO3-like branch.



Extended Data Fig. 2. Imaging and processing of the EfPiwi-piRNA complex (and EfPiwi-piRNA-long target complex).

a, Representative cryo-EM micrograph (1,765 micrographs collected in total). Input sample contained EfPiwi-piRNA and a long target RNA (complementary to piRNA nucleotides g2–g25). **b**, Cryo-EM data processing workflow. The data set contained two populations

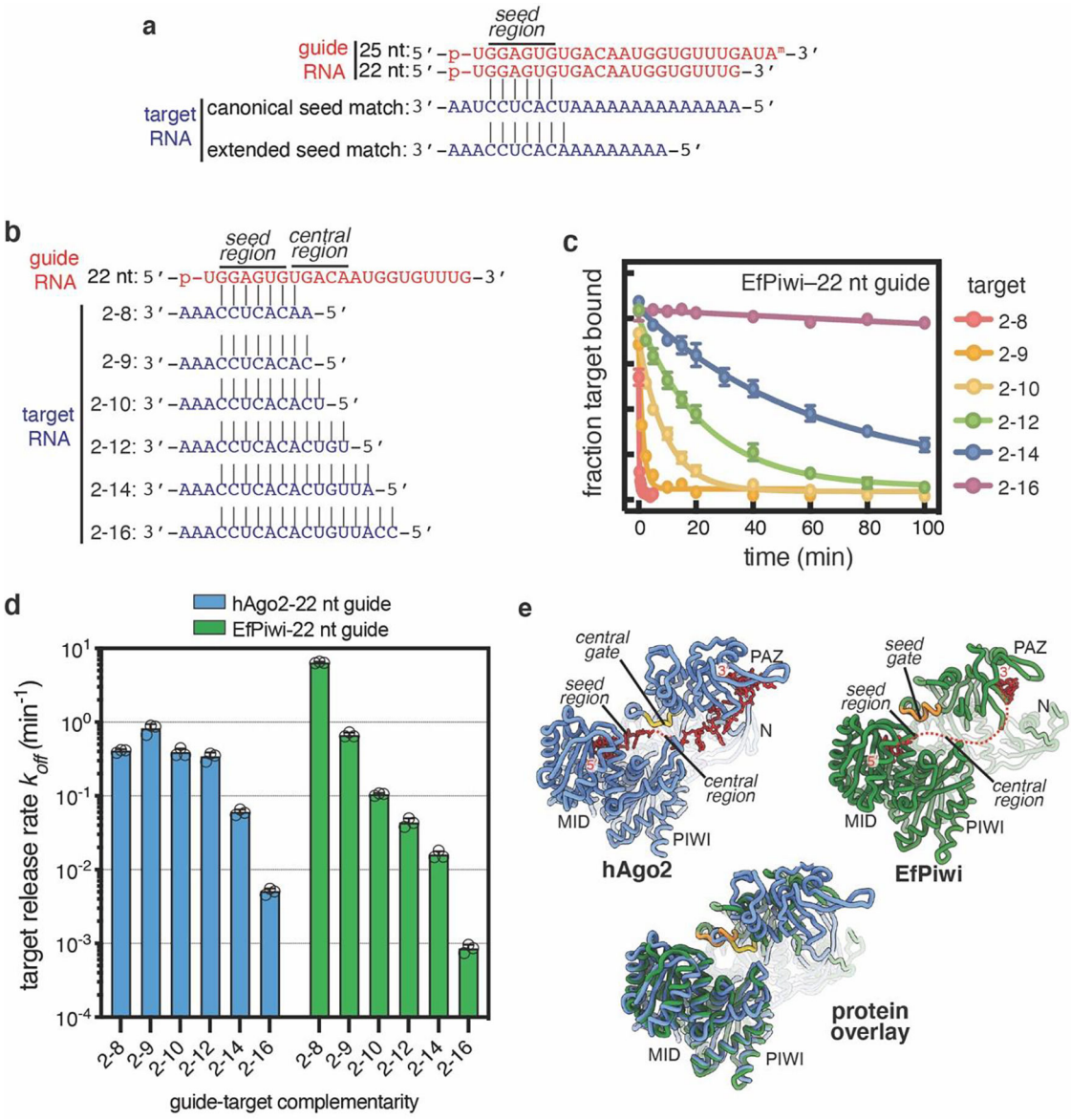
of well resolved particles, one for the binary EfPiwi-piRNA complex and another for the ternary EfPiwi-piRNA-long target complex. Particles isolated from micrographs were sorted by reference-free 2D classification. Only particles containing high-resolution features for the intact complex were selected for downstream processing. 3D classification was used to further remove low-resolution or damaged particles, and the remaining particles were refined to obtain a 3.8 Å reconstruction for the EfPiwi-guide complex, and 8.6 Å for the ternary EfPiwi-piRNA-long target complex. **c**, The final 3D map for the EfPiwi-piRNA complex colored by local resolution values, where the majority of the map was resolved between 3.5 Å and 4 Å with the flexible PAZ and N domains having lower resolution. **d**, Angular distribution plot showing the Euler angle distribution of the EfPiwi-piRNA particles in the final reconstruction. The position of each cylinder corresponds to the 3D angular assignments and their height and color (blue to red) corresponds to the number of particles in that angular orientation. **e**, Directional Fourier Shell Correlation (FSC) plot representing 3D resolution anisotropy in the reconstructed map, with the red line showing the global FSC, green dashed lines correspond to ±1 standard deviation from mean of directional resolutions, and the blue histograms correspond to percentage of directional resolution over the 3D FSC. F. EM density quality of EfPiwi-piRNA complex. Individual domains of EfPiwi fit into the EM density, EM density shown in mesh; molecular models (colored as in Fig. 2) shown in cartoon representation with side chains shown as sticks; piRNA shown in stick representation.



Extended Data Fig. 3. Conserved structural features in extended PIWI family.

a, Surface of hAgo2 (left) and EfPiwi (right), highlighting g5-g6 nucleotide-binding loops. Superimposing g5-g6 nucleotides (red sticks) from hAgo2 onto EfPiwi results in steric clashes. **b**, g5-g6 loop in AGO structures (left) is kinked, enabling pre-organization of seed 3’ end. Equivalent loop in PIWI structures (right) cannot kink due to bulky residues (labeled positions 1 and 2), conserved in PIWI family. **c**, Close up superposition of central-gate and seed-gate structures in AGO and PIWI proteins, respectively. **d**, Superposition of seed-gate regions from all known PIWI (left) and AGO (right) structures, with secondary structure

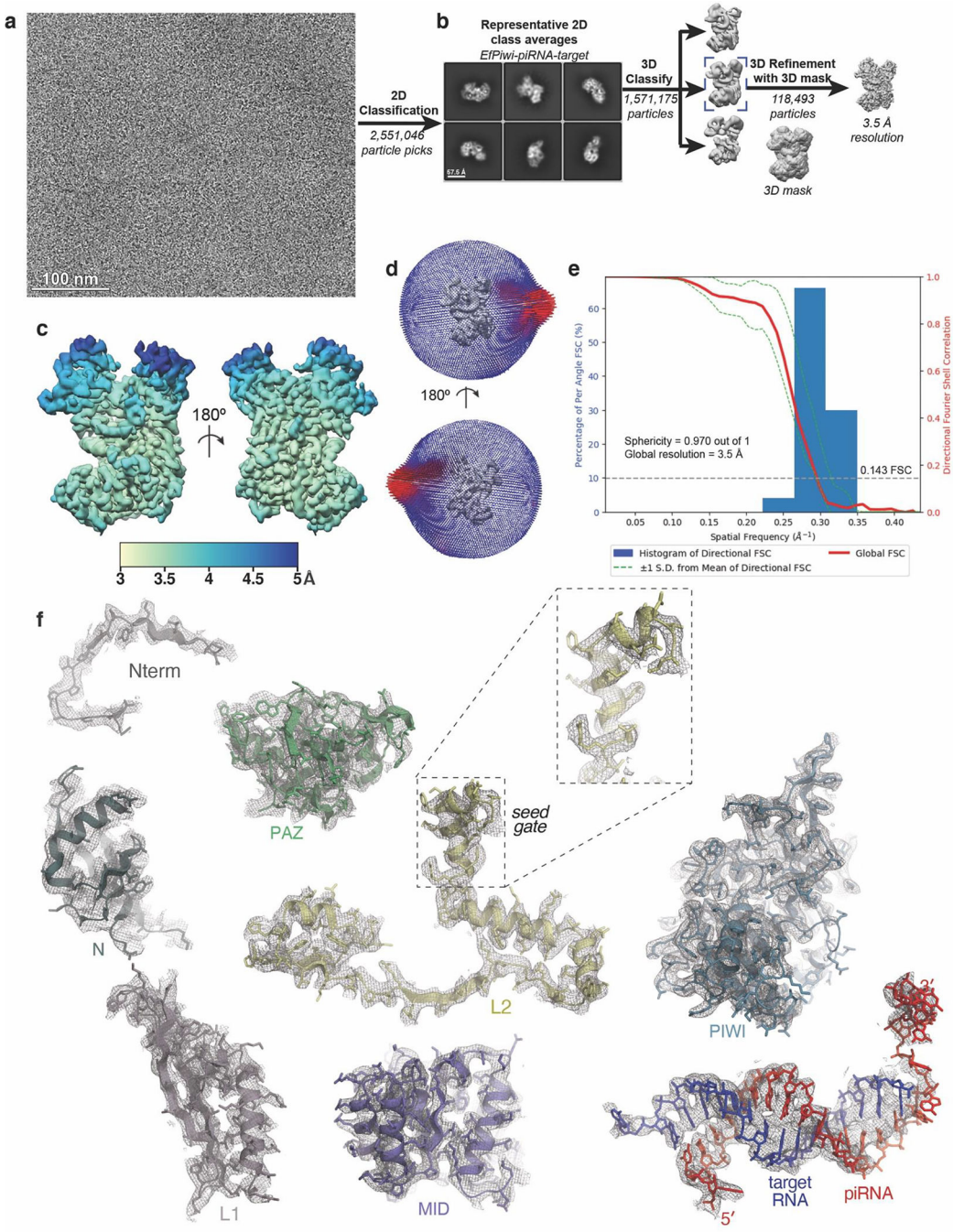
schematics shown above. **e,** Secondary structure predictions indicate the α6 extension is a defining feature of the PIWI family. Predictions were by PSIPRED 4.0. **f,** L1-L2 interface near seed-gate in EfPiwi. Hydrophobic residues buried at the L1-L2 interface are shown. **g,** Sequence alignment shows L1-L2 interface residues in EfPiwi are broadly conserved in PIWIs (green) and distinct from the equivalent residues in AGOs (blue).



Extended Data Fig. 4. Target release from hAgo2 and EfPiwi loaded with identical guides.

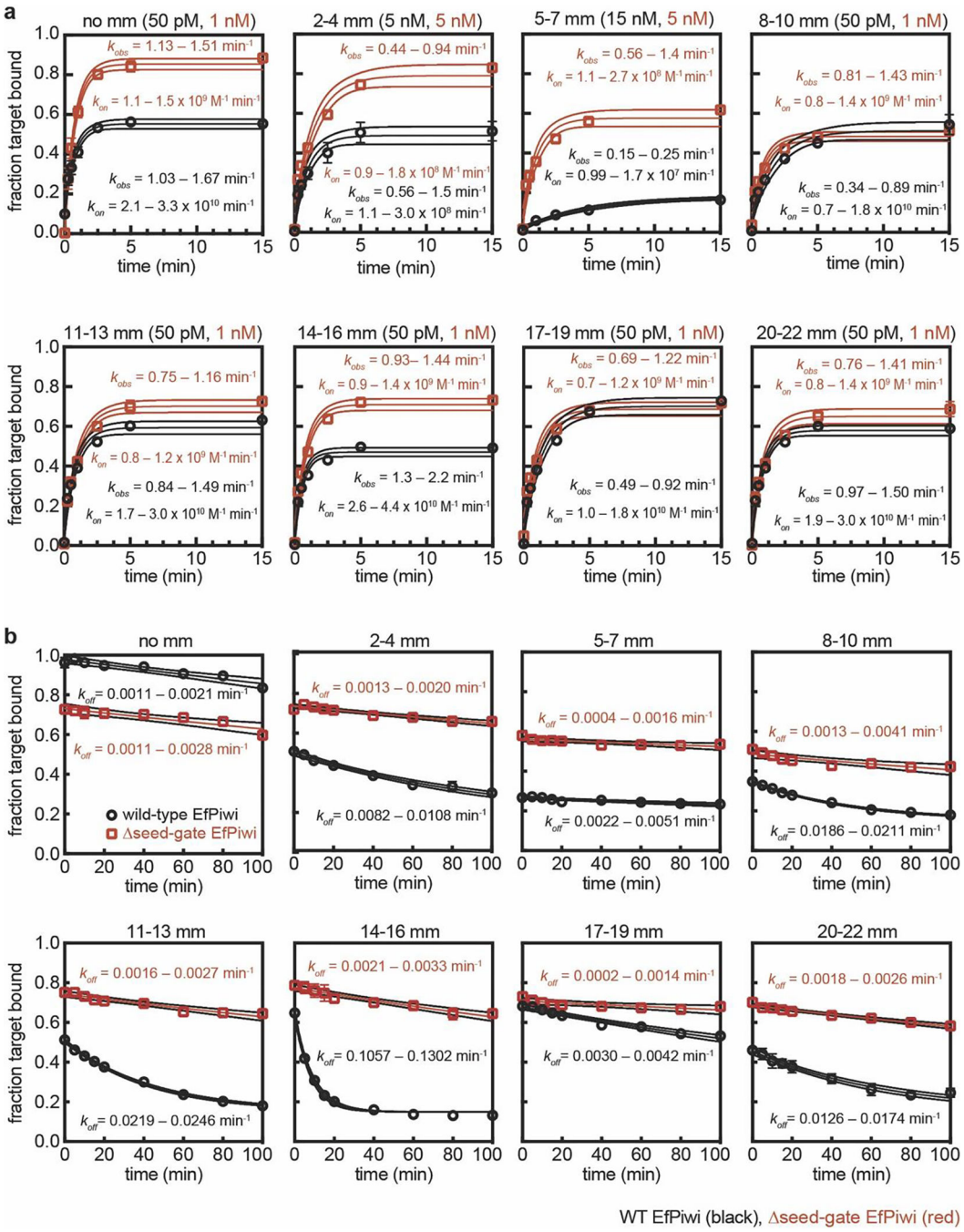
a, Schematic of pairing between guide RNAs and seed-matched target RNAs used in main text Fig. 2b,c. **b**, Schematic of pairing between 22 nt guide RNA and target RNAs spanning the seed and central regions. **c.** Release of ^{32}P-labeled target RNAs from EfPiwi-22 nt guide in the presence of excess unlabeled target RNA over time. **d.** Release rates of target RNAs from hAgo2–22 nt guide (data from main text Fig. 2d, left panel) and EfPiwi-22nt guide (panel c). Results show hAgo2 and EfPiwi create distinct binding properties for the same

guide RNA. All plotted data are the mean values of triplicate measurements. Error bars indicate SD. **e**, Ribbon representation of hAgo2, EfPiwi, and an overlay illustrating relative positions of the central-gate and seed-gate. In **c-d**, n=3 independent experiments, data are mean ± s.d.



Extended Data Fig. 5. Imaging and processing of the EfPiwi-piRNA-target complex.
a, Representative cryo-EM micrograph of EfPiwi-piRNA-target complex (1,881 micrographs collected in total). **b**, Workflow for processing EfPiwi-piRNA-target complex

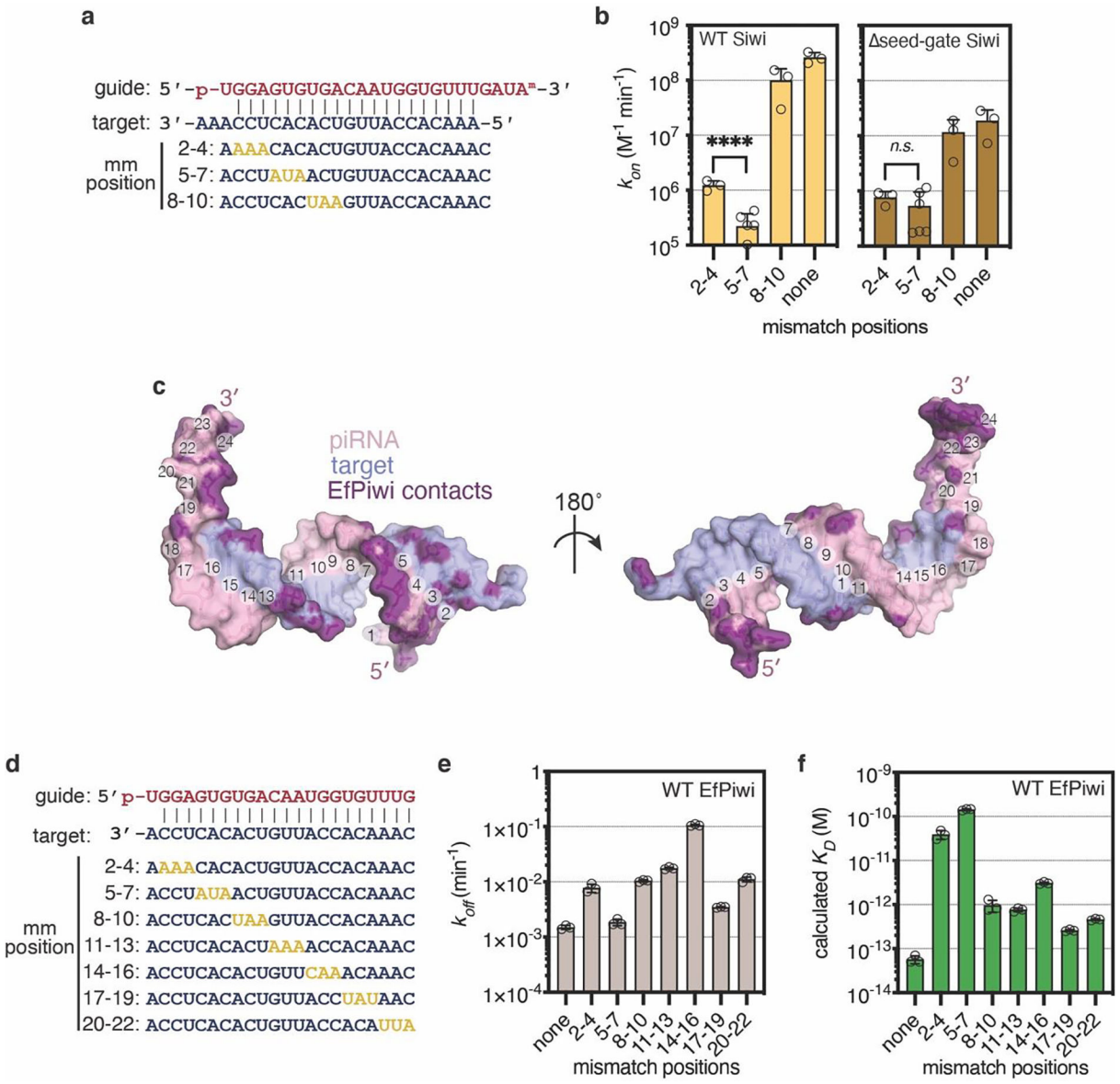
data set. Particles isolated from micrographs were sorted by reference-free 2D classification. Only particles containing high-resolution features for the intact complex were selected for downstream processing. 3D classification was used to further remove low-resolution or damaged particles, and the remaining particles were refined to obtain a 3.5 Å map. **c**, The EfPiwi-piRNA-target complex map colored by local resolution. **d**, Euler angle distribution plot for the EfPiwi-piRNA-target complex particles. **e**, Directional Fourier Shell Correlation (FSC) plot representing 3D resolution anisotropy in the reconstructed map. Red line shows global FSC; green dashed lines ±1 standard deviation from mean of directional resolutions; blue histograms indicate percentage of directional resolution over the 3D FSC. **f**, EM density quality of EfPiwi-piRNA-target complex. Individual domains of EfPiwi and RNAs fit into the EM density; EM density shown in mesh; protein models shown in cartoon representation (colored as in Fig. 5) with side chains shown as sticks; RNAs shown in stick representation.



Extended Data Fig. 6. EfPiwi target binding data.

a, Raw data for k_{on} values shown in Fig. 3f. Plots of target RNAs with mismatches (sequences shown in Extended Data Fig. 6c) binding to EfPiwi-piRNA complexes over time. Protein concentrations used in each experiment are indicated at top of each graph. 95% confidence limits of observed association rates (k_{obs}) and k_{on} values indicated. WT EfPiwi (black), Δseed-gate EfPiwi (red). **b**, Raw data for k_{off} values shown in Extended Data Fig. 7e. Plots of target RNAs with mismatches (mm) dissociating from EfPiwi-piRNA complexes over time. All data were fit to a plateau value of 0.15. 95% confidence limits of k_{on} values

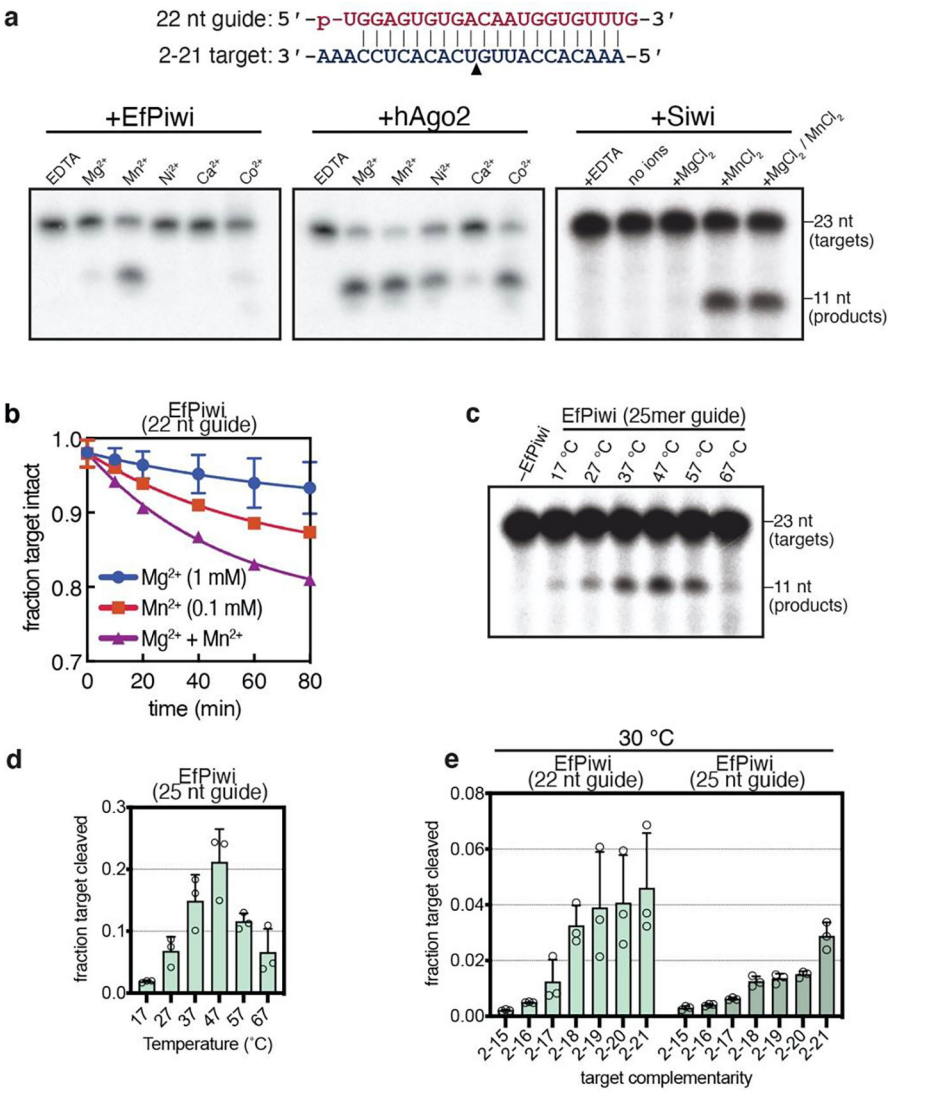
indicated. All data points were measured three times. Error bars indicate SEM. Center line indicates best fit to data. Surrounding lines indicate 95% confidence limits. In all panels, n=3 independent experiments, data are mean ± s.d.



Extended Data Fig. 7. Target binding with mismatches

a, Guide-target pairing schematic for select mismatched targets binding Siwi-piRNA complexes. Mismatches colored gold. **b**, Association rates of target RNAs (shown in panel **a**) with wild-type Siwi (left) and Δseed-gate Siwi. Indicated p-values from two-sided t-test are 6.11×10^{-5} and 0.205 for wild-type and Δseed-gate Siwi, respectively. **c**, Guide-target pairing schematic for mismatched targets used in main text Fig. 3f, and panels d and e here. Mismatches colored gold. **d**, Dissociation rates of ^{32}P-labeled target RNAs with three consecutive mismatches from wild-type EfPiwi. Most mismatched segments had moderate (~10-fold) effects on k_{off}, except 14–16 mismatches, which increased k_{off} ~70 fold. **e**,

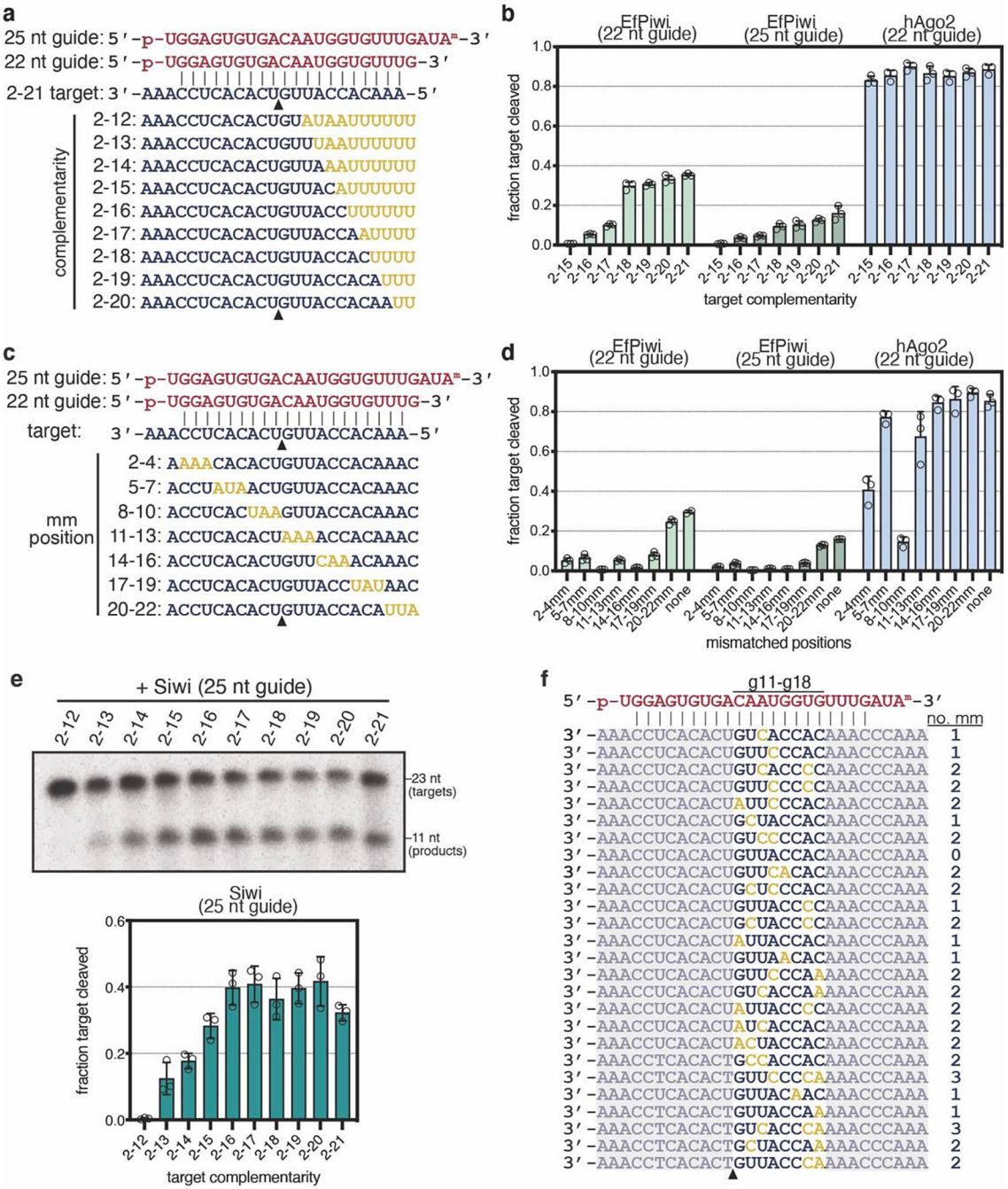
Dissociation constants (K_D) calculated from k_{on} and k_{off} values for target RNAs binding wild-type EfPiwi-guide complex. **f**, Surface representation of the modeled piRNA-target duplex. piRNA nucleotides numbered at the Watson-Crick face. Non-hydrogen RNA atoms positioned ≤ 4 Å from an EfPiwi atom colored purple. In **b,d**, and **e**, n=3 independent experiments, data are mean ± s.d.



Extended Data Fig. 8. Target cleavage by EfPiwi and Siwi (part I).

a, Denaturing gels showing cleavage of g2-g21 matched ^{32}P-labeled target RNA by EfPiwi, hAgo2 or Siwi in the presence of various divalent cations (2 mM each). Schematic of piRNA-target pairing shown (top). Gels are representative results for n=3 independent experiments for EfPiwi and hAgo2, and n=2 independent experiments for Siwi. **b**, Time course showing cleavage of g2-g21 paired ^{32}P-labeled target RNA by EfPiwi in the presence of Mg^{2+}, Mn^{2+}, or both at approximate physiological divalent cation concentrations. **c**, Cleavage of g2-g21 matched ^{32}P-labeled target RNA by EfPiwi at various temperatures shows activity over the full physiological range (17–30°C). Gel is representative of n=3

independent experiments. **d**, Quantitation of results (and replicates) in c. **e**, Cleavage of target RNAs with varying degrees of 3’ complementarity by EfPiwi at 30°C. In **b, d** and **e**, n=3 independent experiments, data are mean ± s.d.



Extended Data Fig. 9. Target cleavage by EfPiwi and Siwi (part II).

a, Guide-target pairing schematic for targets with varying degrees of complementarity to piRNA 3’ end used in Fig. 4a (main text) and panels b and e (here). **b**, Quantification target RNAs (1 nM) cleaved after treating with excess (100 nM) EfPiwi loaded with a 22 or 25

nt guide or hAgo2 loaded with a 22 nt guide for 1 hour. *n*=3 independent experiments, data are mean ± s.d. **c**, Guide-target pairing schematic for targets with 3 nt mismatched regions. **d**, Quantification mismatched target RNAs (1 nM) cleaved after treating with excess (100 nM) EfPiwi loaded with a 22 or 25 nt guide or hAgo2 loaded with a 22 nt guide for 1 hour. *n*=3 independent experiments, data are mean ± s.d. **e,** Cleavage of targets with varying degrees of complementarity to piRNA 3' end (shown panel a) by Siwi (2 mM $MnCl_2$, 37°C). Gel is representative of *n*=3 independent experiments, with data plotted as mean ± s.d. shown below. **f**, From Fig. 4c: sequences of the 26 target RNAs, with 0–8 mismatches opposite piRNA nucleotides g11–g18, that were most readily cleaved by EfPiwi (listed in order of cleavage product abundance). Mismatched nucleotides colored yellow. Grayed out sequences indicate constant regions shared by all target RNAs. Triangle indicates cleavage site.

Extended Data Table 1

Cryo-EM data collection, refinement, and validation statistics

Sample name	**EfPiwi-piRNA-target (2–16)**	**EfPiwi-piRNA**	**EfPiwi (MID/PIWI)-piRNA-long-target (2–25)**
EMDB ID	EMD-23063	EMD-23061	EMD-23062
PDB ID	7KX9	7KX7	
Microscope	Talos Arctica	Talos Arctica	
Detector (Mode)	Gatan K2 Summit, counting mode	Gatan K2 Summit, counting mode	
Voltage (kV)	200	200	
Magnification (nominal / at detector)	36,000× / 43,478×	36,000× / 43,478×	
Total electron fluence ($e^-/Å^2$)	47.33	47.33	
Electron flux (e^-/pixel/sec)	3.91	5.22	
Defocus range (μm)	−1.0 to −1.6	−1.0 to −1.6	
Pixel size (Å)	1.15	1.15	
Total exposure time (sec)	16	12	
Total frames/micrograph	64	48	
Exposure per frame ($e^-/Å^2$/frame)	0.74	0.99	
Micrographs collected (no.)	1,881	1,765	
Total extracted particles (no.)	2,551,046	3,280,351	
Particles used for 3D analyses (no.)	1,571,175	1,862,936	608,488

Sample name	EfPiwi-piRNA-target (2–16)	EfPiwi-piRNA	EfPiwi (MID/PIWI)-piRNA-long-target (2–25)
Final refined particles (no.)	118,493	125,041	116,655
Accuracy of rotations (deg.) / translations (pixels)	1.972 / 0.581	2.748 / 0.785	8.374 / 1.739
Symmetry imposed	C1	C1	C1
Global Resolution (Å)			
FSC 0.5 (masked/unmasked)	3.8/4.0	4.3/4.5	8.6 / 10.8
FSC 0.143 (masked/unmasked)	3.4/3.6	3.8/3.9	8.6 / 10.8
FSC Sphericity	0.97	0.91	
Local resolution range (Å)	3.5 – 6.0	3.5 – 6.0	
Map Sharpening *B* factors (Å^2)	−80	−80	
Model composition			
Non-hydrogen atoms	6578	5828	
Protein residues	734	729	
RNA residues	40	7	
Mg2+ ions	2	2	
Refinement			
Refinement package (s)	Phenix	Phenix	
Map Correlation Coefficient			
Global	0.79	0.82	
Local	0.78	0.83	
R.m.s. deviations			
Bond lengths (Å)	0.008	0.008	
Bond angles (°)	1.183	1.092	
Validation			
EMRinger score	2.84	2.98	
MolProbity score	2.22	2.18	
Clashscore	13.11	13.87	
Rotamer outliers (%)	0.17	0.34	

Sample name	EfPiwi-piRNA-target (2–16)	EfPiwi-piRNA	EfPiwi (MID/PIWI)-piRNA-long-target (2–25)
Cβ deviations (%)	0	0	
Ramachandran plot			
Favored (%)	88.46	90.57	
Allowed (%)	11.54	9.43	
Disallowed (%)	0	0	
CaBLAM outliers (%)	3.74	3.8	

Acknowledgments

We are grateful to Noriko Funayama for the *Ephydatia fluviatilis piwi-a* cDNA clone, Yukihide Tomari for the Siwi cDNA clone, and to Irwin H. Segel for advice about measuring binding reactions with very slow off-rates. Research of G.C.L. is supported by NIH grant R21AG067594 and an Amgen Young Investigator Award. Research of I.J.M. is supported by NIH grant R35GM127090.

Data Availability

Maps for the EfPiwi-piRNA and EfPiwi-piRNA-target complexes were deposited in the Electron Microscopy Data Bank under accession IDs EMD-23061 and EMD-23063, respectively. Corresponding atomic models were deposited in the Protein Data Bank under accession IDs 7KX7 and 7KX9. The EfPiwi(MID/PIWI)-piRNA-long-target complex map was deposited in the Electron Microscopy Data Bank under accession ID EMD-23062.

References

1. Ozata DM, Gainetdinov I, Zoch A, O'Carroll D & Zamore PD PIWI-interacting RNAs: small RNAs with big functions. Nature reviews. Genetics 20, 89–108, doi:10.1038/s41576-018-0073-3 (2019).
2. Shen EZ et al. Identification of piRNA Binding Sites Reveals the Argonaute Regulatory Landscape of the C. elegans Germline. Cell 172, 937–951 e918, doi:10.1016/j.cell.2018.02.002 (2018). [PubMed: 29456082]
3. Gou LT et al. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res 24, 680–700, doi:10.1038/cr.2014.41 (2014). [PubMed: 24787618]
4. Grimson A et al. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature 455, 1193–1197, doi:10.1038/nature07415 (2008). [PubMed: 18830242]
5. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP & Burge CB Prediction of mammalian microRNA targets. Cell 115, 787–798 (2003). [PubMed: 14697198]
6. Wee LM, Flores-Jasso CF, Salomon WE & Zamore PD Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. Cell 151, 1055–1067, doi:10.1016/j.cell.2012.10.036 (2012). [PubMed: 23178124]
7. Salomon WE, Jolly SM, Moore MJ, Zamore PD & Serebrov V Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. Cell 162, 84–95, doi:10.1016/j.cell.2015.06.029 (2015). [PubMed: 26140592]
8. Chandradoss SD, Schirle NT, Szczepaniak M, MacRae IJ & Joo C A Dynamic Search Process Underlies MicroRNA Targeting. Cell 162, 96–107, doi:10.1016/j.cell.2015.06.032 (2015). [PubMed: 26140593]

9. Parker JS, Parizotto EA, Wang M, Roe SM & Barford D Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. Molecular cell 33, 204–214, doi:10.1016/j.molcel.2008.12.012 (2009). [PubMed: 19187762]
10. Schirle NT & MacRae IJ The crystal structure of human Argonaute2. Science 336, 1037–1040, doi:10.1126/science.1221551 (2012). [PubMed: 22539551]
11. Schirle NT, Sheu-Gruttadauria J & MacRae IJ Structural basis for microRNA targeting. Science 346, 608–613, doi:10.1126/science.1258040 (2014). [PubMed: 25359968]
12. Sheu-Gruttadauria J, Xiao Y, Gebert LF & MacRae IJ Beyond the seed: structural basis for supplementary microRNA targeting by human Argonaute2. The EMBO journal, doi:10.15252/embj.2018101153 (2019).
13. Friedman RC, Farh KK, Burge CB & Bartel DP Most mammalian mRNAs are conserved targets of microRNAs. Genome research 19, 92–105, doi:10.1101/gr.082701.108 (2009). [PubMed: 18955434]
14. Zhang D et al. The piRNA targeting rules and the resistance to piRNA silencing in endogenous genes. Science 359, 587–592, doi:10.1126/science.aao2840 (2018). [PubMed: 29420292]
15. Goh WS et al. piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. Genes & development 29, 1032–1044 (2015). [PubMed: 25995188]
16. Halbach R et al. A satellite repeat-derived piRNA controls embryonic development of Aedes. Nature 580, 274–277, doi:10.1038/s41586-020-2159-2 (2020). [PubMed: 32269344]
17. Zhang P et al. MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. Cell Res 25, 193–207, doi:10.1038/cr.2015.4 (2015). [PubMed: 25582079]
18. Nozawa M et al. Evolutionary Transitions of MicroRNA-Target Pairs. Genome Biol Evol 8, 1621–1633, doi:10.1093/gbe/evw092 (2016). [PubMed: 27189995]
19. Brennecke J et al. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128, 1089–1103, doi:10.1016/j.cell.2007.01.043 (2007). [PubMed: 17346786]
20. Wang J et al. piRBase: a comprehensive database of piRNA sequences. Nucleic acids research 47, D175–D180, doi:10.1093/nar/gky1043 (2019). [PubMed: 30371818]
21. Matsumoto N et al. Crystal Structure of Silkworm PIWI-Clade Argonaute Siwi Bound to piRNA. Cell 167, 484–497 e489, doi:10.1016/j.cell.2016.09.002 (2016). [PubMed: 27693359]
22. Yamaguchi S et al. Crystal structure of Drosophila Piwi. Nature communications 11, 858, doi:10.1038/s41467-020-14687-1 (2020).
23. Funayama N, Nakatsukasa M, Mohri K, Masuda Y & Agata K Piwi expression in archeocytes and choanocytes in demosponges: insights into the stem cell system in demosponges. Evol Dev 12, 275–287, doi:10.1111/j.1525-142X.2010.00413.x (2010). [PubMed: 20565538]
24. Alie A et al. The ancestral gene repertoire of animal stem cells. Proceedings of the National Academy of Sciences of the United States of America 112, E7093–7100, doi:10.1073/pnas.1514789112 (2015). [PubMed: 26644562]
25. Wynant N, Santos D & Vanden Broeck J The evolution of animal Argonautes: evidence for the absence of antiviral AGO Argonautes in vertebrates. Sci Rep 7, 9230, doi:10.1038/s41598-017-08043-5 (2017). [PubMed: 28835645]
26. Reuter M et al. Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. Nature 480, 264–267, doi:10.1038/nature10672 (2011). [PubMed: 22121019]
27. Wu PH et al. The evolutionarily conserved piRNA-producing locus pi6 is required for male mouse fertility. Nature genetics 52, 728–739, doi:10.1038/s41588-020-0657-7 (2020). [PubMed: 32601478]
28. Arif A et al. The tiny, conserved zinc-finger protein GTSF1 helps PIWI proteins achieve their full catalytic potential. bioRxiv, 2021.2005.2004.442675, doi:10.1101/2021.05.04.442675 (2021).
29. Herzog VA et al. Thiol-linked alkylation of RNA to assess expression dynamics. Nature methods 14, 1198–1204, doi:10.1038/nmeth.4435 (2017). [PubMed: 28945705]
30. Sienski G, Donertas D & Brennecke J Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression. Cell 151, 964–980, doi:10.1016/j.cell.2012.10.040 (2012). [PubMed: 23159368]

31. Le Thomas A et al. Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes & development 27, 390–399, doi:10.1101/gad.209841.112 (2013). [PubMed: 23392610]
32. Aravin AA et al. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. Molecular cell 31, 785–799, doi:10.1016/j.molcel.2008.09.003 (2008). [PubMed: 18922463]
33. Kuramochi-Miyagawa S et al. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes & development 22, 908–917, doi:10.1101/gad.1640708 (2008). [PubMed: 18381894]
34. Nott TJ et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. Molecular cell 57, 936–947, doi:10.1016/j.molcel.2015.01.013 (2015). [PubMed: 25747659]

Methods References

35. Sheu-Gruttadauria J & MacRae IJ Phase Transitions in the Assembly and Function of Human miRISC. Cell 173, 946–957 e916, doi:10.1016/j.cell.2018.02.051 (2018). [PubMed: 29576456]
36. Flores-Jasso CF, Salomon WE & Zamore PD Rapid and specific purification of Argonaute-small RNA complexes from crude cell lysates. RNA 19, 271–279, doi:10.1261/rna.036921.112 (2013). [PubMed: 23249751]
37. Suloway C et al. Automated molecular microscopy: the new Leginon system. J Struct Biol 151, 41–60, doi:10.1016/j.jsb.2005.03.010 (2005). [PubMed: 15890530]
38. Zheng SQ et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature methods 14, 331–332, doi:10.1038/nmeth.4193 (2017). [PubMed: 28250466]
39. Lander GC et al. Appion: an integrated, database-driven pipeline to facilitate EM image processing. J Struct Biol 166, 95–102, doi:10.1016/j.jsb.2009.01.002 (2009). [PubMed: 19263523]
40. Kimanius D, Forsberg BO, Scheres SH & Lindahl E Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. Elife 5, doi:10.7554/eLife.18722 (2016).
41. Rohou A & Grigorieff N CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol 192, 216–221, doi:10.1016/j.jsb.2015.08.008 (2015). [PubMed: 26278980]
42. Goddard TD, Huang CC & Ferrin TE Visualizing density maps with UCSF Chimera. J Struct Biol 157, 281–287, doi:10.1016/j.jsb.2006.06.010 (2007). [PubMed: 16963278]
43. Pettersen EF et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605–1612, doi:10.1002/jcc.20084 (2004). [PubMed: 15264254]
44. Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature methods 14, 290–296, doi:10.1038/nmeth.4169 (2017). [PubMed: 28165473]
45. Tan YZ et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. Nature methods 14, 793–796, doi:10.1038/nmeth.4347 (2017). [PubMed: 28671674]
46. Waterhouse A et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research 46, W296–W303, doi:10.1093/nar/gky427 (2018). [PubMed: 29788355]
47. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of Coot. Acta crystallographica. Section D, Biological crystallography 66, 486–501, doi:10.1107/S0907444910007493 (2010). [PubMed: 20383002]
48. Liebschner D et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr D Struct Biol 75, 861–877, doi:10.1107/S2059798319011471 (2019). [PubMed: 31588918]
49. Williams CJ et al. MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci 27, 293–315, doi:10.1002/pro.3330 (2018). [PubMed: 29067766]
50. Goddard TD et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci 27, 14–25, doi:10.1002/pro.3235 (2018). [PubMed: 28710774]

51. Buchan DWA & Jones DT The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic acids research 47, W402–W407, doi:10.1093/nar/gkz297 (2019). [PubMed: 31251384]

52. Jones DT Protein secondary structure prediction based on position-specific scoring matrices. Journal of molecular biology 292, 195–202, doi:10.1006/jmbi.1999.3091 (1999). [PubMed: 10493868]

53. Nakanishi K et al. Eukaryote-specific insertion elements control human ARGONAUTE slicer activity. Cell Rep 3, 1893–1900, doi:10.1016/j.celrep.2013.06.010 (2013). [PubMed: 23809764]

54. Park MS et al. Human Argonaute3 has slicer activity. Nucleic acids research 45, 11867–11877, doi:10.1093/nar/gkx916 (2017). [PubMed: 29040713]

55. Park MS et al. Multidomain Convergence of Argonaute during RISC Assembly Correlates with the Formation of Internal Water Clusters. Molecular cell 75, 725–740 e726, doi:10.1016/j.molcel.2019.06.011 (2019). [PubMed: 31324450]

56. Martin M Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011 17, 3, doi:10.14806/ej.17.1.200 (2011).

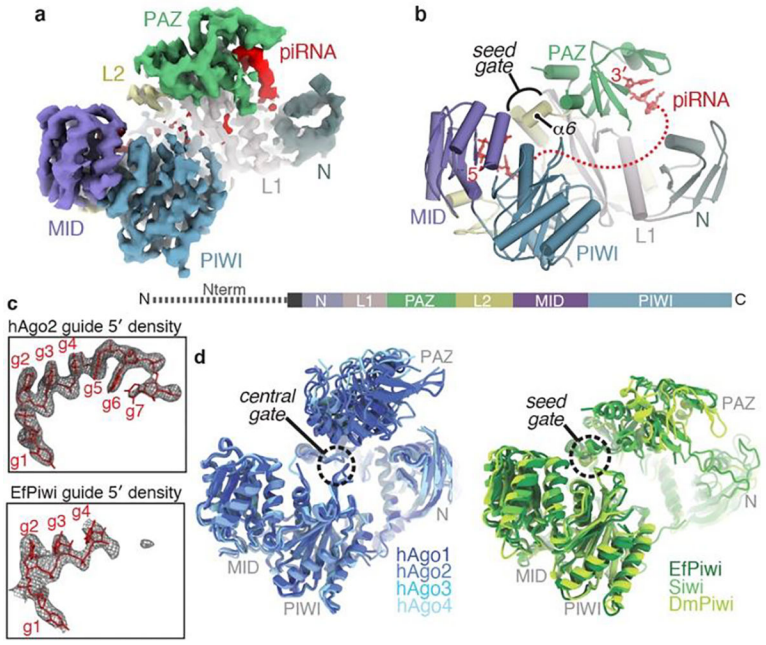


Fig. 1. Structural features unique to PIWIs.

a, Segmented EfPiwi-piRNA cryo-EM reconstruction, colored by domain. **b,** Cartoon representation of EfPiwi model. Disordered piRNA nucleotides, dashed red line. Linear domain schematic shown below. **c,** Density maps (grey mesh) in space occupied by seed regions of hAgo2 (PDB 4OLA) and EfPiwi. Modeled guide RNAs as red sticks. **d,** Superposition of miRNA-class AGO structures (left, PDB: 4KXT, 4OLA, 5VM9, 6OON) and known PIWI structures (right, PDB: 7KX7, 5GUH, 6KR6). AGO central-gate and PIWI seed-gate structures indicated.

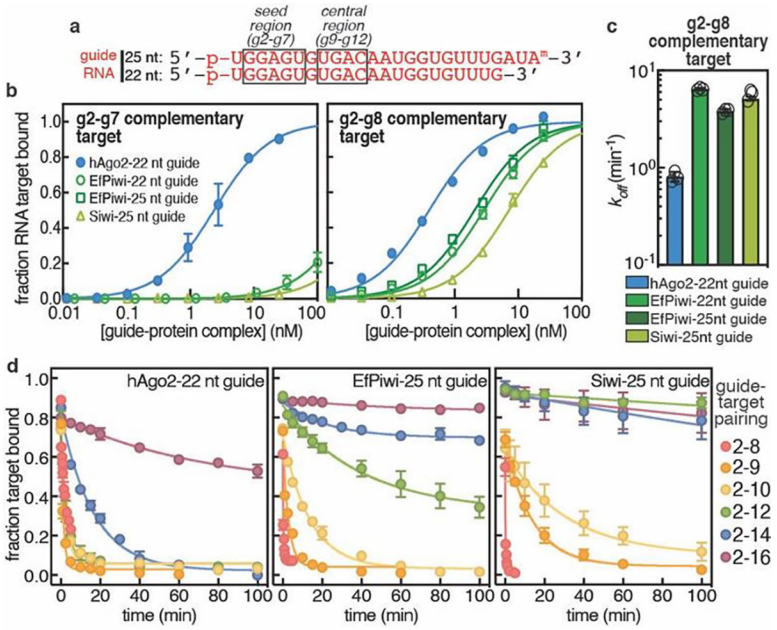


Fig. 2. piRNAs are more selective than miRNAs.

a, Schematic of guide RNAs used. **b,** Fraction of target RNAs bound by hAgo2-guide, EfPiwi-guide, and Siwi-guide RNA complexes at equilibrium versus protein-guide concentration. **c,** Release rates of g2-g8 complementary target protein-guide RNA complexes. **d,** Release of ^{32}P-labeled target RNAs from indicated protein-guide complexes in the presence of excess unlabeled target RNA over time. In **b-c** n=3 independent experiments, data are mean ± s.d.

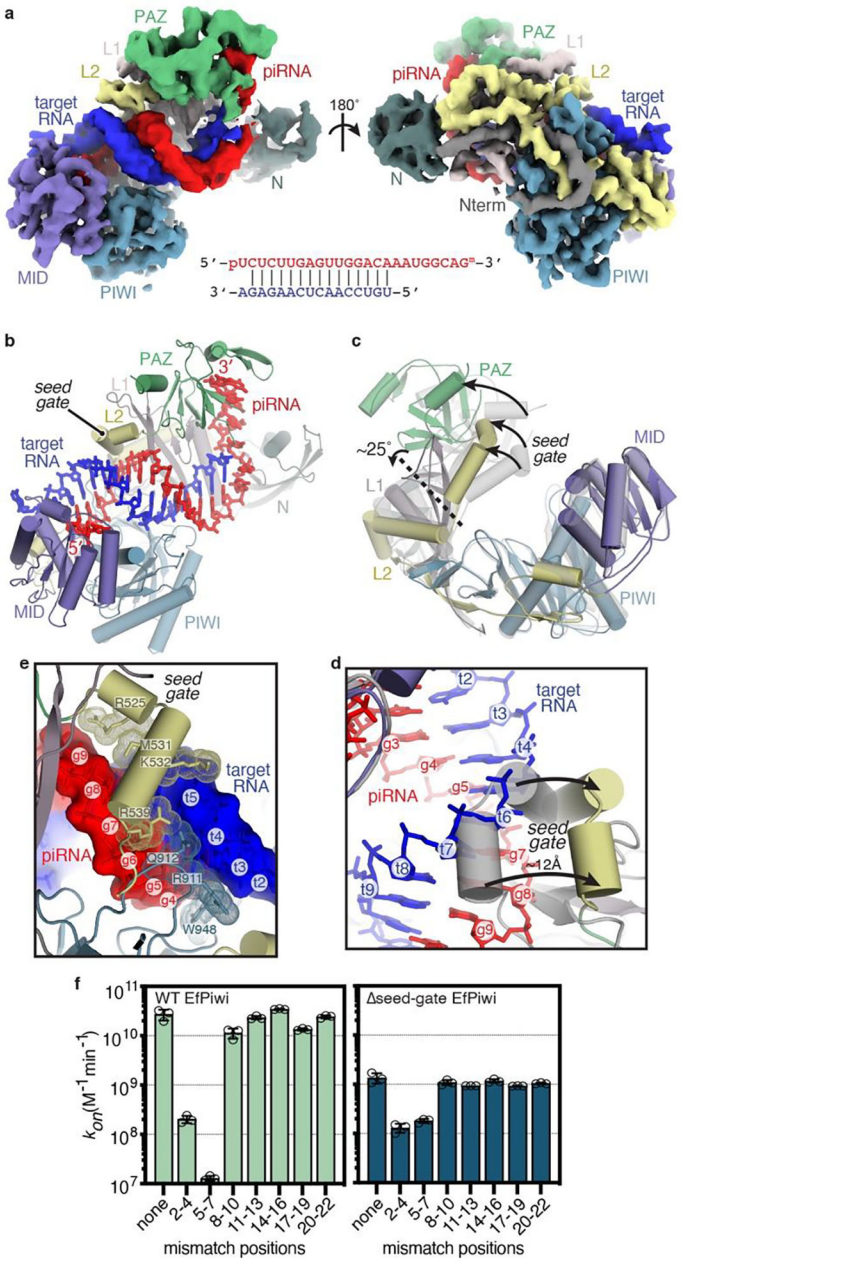


Fig. 3. Structural basis for piRNA target binding.

a, Segmented EfPiwi-piRNA-target RNA cryo-EM reconstruction, colored as in Fig 1. piRNA-target base pairing schematic shown below. **b,** Cartoon representation of EfPiwi-piRNA-target RNA model. **c,** Superposition of EfPiwi-piRNA (gray, semi-transparent) and EfPiwi-piRNA-target (colored and solid) structures. Arrows indicate movements to target-bound structure. Dashed line indicates hinge in L1 stalk. **d,** Closeup view showing the seed-gate shifts ~12 Å upon target binding. Arrows indicate direction of movement from guide-only to target-bound conformation. **e,** Interactions between EfPiwi and the piRNA-target duplex around the seed region. **f,** Association rates of target RNAs with

three consecutive mismatches binding wild-type (left) and Δseed-gate (right) EfPiwi. *n*=3 independent experiments, data are mean ± s.d.

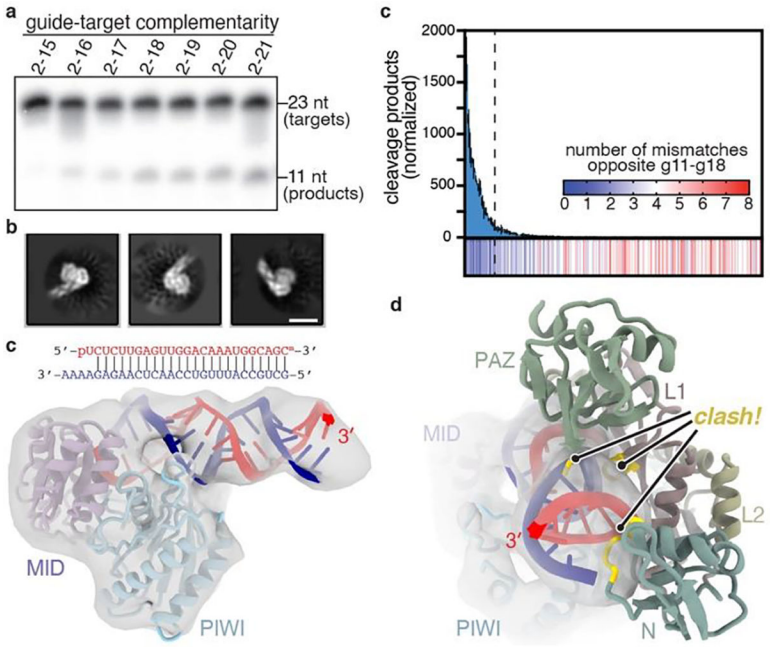


Fig. 4. Extensive pairing activates piRNA-target cleavage.

a, Representative denaturing gel showing intact and cleaved target RNAs, with increasing amounts of complementarity to a piRNA, after incubation with excess EfPiwi-piRNA for one hour (results typical of 3 replicates). **b,** relative number of RNA-seq reads of cleavage products after treating a pool of 256 target RNAs with excess EfPiwi-piRNA. Heatmap indicates number of mismatches opposite piRNA nucleotides g11–g18 in each sequence. Dashed line denotes 26 most abundant products. *n*=3 independent experiments, data are mean ± s.d. **d,** Three examples of 2D class averages of EfPiwi-piRNA complex bound to a target with complementarity from g2-g25. Scale bar is 5 nm. **e,** 3D cryo-EM reconstruction fit with the MID/PIWI lobe and extended piRNA-target duplex. piRNA 3’ end indicated. **g,** Docking the EfPiwi-piRNA-target model with pairing limited to g2-g16 (see Fig. 3) reveals the N, L1, and PAZ domains sterically clash (gold highlights) with the extended piRNA-target duplex.