Published in final edited form as:

Curr Opin Struct Biol. 2009 June ; 19(3): 286–293. doi:10.1016/j.sbi.2009.04.006.

Structural and functional modules in RNA interference

Marcin Nowotny¹ and Wei Yang²

¹Laboratory of Protein Structure, International Institute of Molecular and Cell Biology, 4 Ks. Trojdena Street, 02-109, Warsaw, Poland, e-mail: mnowotny@iimcb.gov.pl

²Laboratory of Molecular Biology, NIDDK, NIH, 9000 Rockville Pike, Bethesda, MD 20892, USA, e-mail: wei.yang@nih.gov

Summary

RNA interference (RNAi) uses small RNA molecules to regulate transcriptional and post-transcriptional gene expression. In recent years, a number of structural studies provided insights into the molecular architecture and mechanism of functional modules of RNAi. Mechanisms of nucleic acid recognition and cleavage have been revealed by structural studies of proteins and their nucleic acid complexes involved in RNA biogenesis, for example, Argonaute, PIWI, RNase III, Dicer, Drosha and DGCR8. While quite a few questions remain, an excellent structural and mechanistic overview of RNAi processes has already emerged. In this review, we examine functional modules and their assemblies in RNAi processes.

Introduction

The discovery of RNA interference is a major breakthrough in modern biology. RNAi influences cell proliferation, development, immunity, transposition and tumorigenesis. Functions and processes of RNAi have been reviewed extensively [1–3]. Briefly, two major classes of RNAi are known – small interfering RNAs (siRNA) and micro-RNAs (miRNA) [4,5] (Fig 1A). Both processes require cleavage of dsRNA to generate ~ 20–30 bp dsRNA with 2-nucleotide overhangs at the 3' ends. siRNA fragments are generated by Dicer from dsRNA precursors. The miRNA pathway starts in nucleus from processing of non-coding pri-miRNA by a complex of Drosha and DGCR8 proteins to produce ~70-nt pre-miRNAs. Pre-miRNAs are then transported to the cytoplasm and cleaved by Dicer to make miRNA fragments. One strand of short dsRNA products (known as a guide RNA) is then incorporated into the RISC complex for selecting target mRNAs by base pairing. In siRNA, Argonaute (Ago, the nuclease component of RISC) uses the guide RNA to select complementary mRNA for degradation. In miRNA, base pairing between guide and target mRNA leads to translational inhibition. A recently discovered new class of RNAi, piRNAs (PIWI interacting RNA, PIWI and Ago are closely related), prevents the spreading of selfish genetic elements by methyl-dependent epigenetic silencing [6,7] and cleavage of transposon mRNA [8] (Fig. 1B).

The following functions are repeatedly required during RNAi: recognition of the 3' and 5' ends of RNA, binding of dsRNA, and cleavage of one or both strands of dsRNA at a defined distance from one end. Remarkably, nature uses a limited number of protein modules to accomplish these tasks (Fig 1C). PAZ domain recognizes the 3' end of RNA in both miRNA and siRNA, the Mid domain of PIWI and Ago binds the 5'-phosphate, and dsRNA-binding domains

(dsRBD) are found in Drosha, DGCR8 and Dicer. An RNase III-like endonuclease domain (endoND) carries out all dsRNA cleavage in RNAi. The PIWI domain in PIWI and Ago resembles RNase H and can nick mRNA (passenger or target RNA) that forms base pairs with the guide RNA. This is often referred to as the “slicer” activity. Some PIWI and Ago do not possess the slicer activity due to active site mutations, and their function is most likely to bind dsRNA. In this review, we briefly summarize the structural and functional modules in RNAi processes and how they are assembled into various complexes.

PAZ domain and 3′-end recognition

The PAZ domain contains around 130 residues and binds the 3′ end of ssRNA or a 3′ end protruding from dsRNA [9,10]. The core of the PAZ domain consists of a short six-stranded β -barrel that resembles an OB fold and a β -hairpin and α -helix on the side of the barrel (Fig. 2A). In some PAZ domains two α -helices cover one end of the barrel (Fig. 2B) [11–14]. Structures of the *Drosophila* Argonaute PAZ domain bound to ssRNA or ssDNA [11], human Ago PAZ bound to siRNA-like duplex [10] and *T. thermophilus* Ago (Tt-Ago) bound to a ssDNA guide strand and RNA/DNA duplex [15,16] reveal a conserved mechanism for nucleic acid recognition. PAZ approaches the 3′ end of RNA or DNA with its open β barrel end (Fig. 2B), and the interactions with the last two nucleotides at the 3′ end are conserved among these structures (Fig 2B). The orientation and structures of the remaining ds or ss nucleic acids are variable apparently due to limited and non-specific interactions with PAZ.

The last 2 nts at the 3′ end are stacked in the A conformation and inserted into a large binding pocket between the side of the β -barrel and the hairpin/helix element forming extensive polar and hydrophobic interactions with PAZ (Fig. 2B). Several conserved residues including key aromatic side chains line the pocket and interact mainly with the phosphosugar backbone. In particular, a characteristically conserved tyrosine stabilizes the phosphate between the last two nucleotides with a hydrogen bond. The 3′-end is surrounded by hydrophobic and aromatic residues, and the 3′-OH is hydrogen bonded with a main-chain carbonyl oxygen. The W–C edges of these unpaired bases are exposed to solvent, thus nucleic acid-binding is non-sequence specific. There is no interaction with 2′-OH of the RNA, which agrees with the observation that PAZ binds either RNA or DNA.

dsRBD domain

The ~70-residue dsRNA-binding domain (dsRBD) is present in many proteins involved in transcription, RNA processing, mRNA localization and translation [17]. Structures of several dsRBDs have been determined. They all contain a central three-stranded anti-parallel β -sheet and two helices in the topological order of α - β - β - α [18–20]. The two helices are roughly parallel and located on one side of the β -sheet. Structures of different dsRBDs in complex with RNAs (*S. cerevisiae* Rnt1p [21], *X. leavis* RNA-binding protein A [22] and RNase III from *A. aeolicus* [23]) reveal two alternative binding modes. Most often a dsRBD binds dsRNA without sequence preference and interacts with the phosphosugar backbones in the A conformation, which is characterized by a wide minor groove and narrow major groove (Fig. 2C). Charge-charge interactions are formed between the helical dipoles (N-termini) of the α -helices and the phosphate groups surrounding the narrow major groove. Recognition of riboses instead of deoxyriboses is achieved through hydrogen bonding between protein side chains and 2′-OH groups along the minor groove (Fig. 2C). Alternatively, a dsRBD can recognize hairpin-loop structures, mismatches and bulges in a sequence-specific manner by interacting with unpaired RNA or a loop region. For example the dsRBD from Rnt1p (yeast RNase III) binds a tetraloop (Fig. 2D) and directs cleavage of adjacent dsRNA at a specific location by endoND [24].

Recognition of the 5' end by Mid domain

The 5' end of a guide RNA is recognized by RISC complex, and the cleavage site of target mRNA is determined by the number of base pairs from the 5' end [25]. The Mid domain found in PIWI and Ago contains a four-stranded parallel β -sheet flanked with a pair of α -helices on each side (Fig. 2E). The fold is similar to the sugar-binding domain of Lac repressor [26]. Structures of Mid domain associated with dsRNA and ssDNA (*A. fulgidus* PIWI protein bound to a small siRNA-like duplex [27,28] and *T. thermophilus* Ago with ssDNA and RNA/DNA [15,16]) reveal that it recognizes the 5'-phosphate specifically. Even with dsRNA most contacts are formed with the strand presenting the 5' end. The 5' end of either RNA or DNA makes a U-turn. The first base is flipped out and stacked with an Arg or Tyr side chain, and the 5' phosphate is placed in a highly conserved pocket in the Mid domain interacting with invariant Tyr, Lys, and Gln residues (Fig. 2E). Interestingly, the 5'-phosphate is juxtaposed to the 3rd phosphate, and a metal ion coordinated by the carboxylate group of a C-terminal valine or leucine stabilizes these phosphates (Fig. 2F).

dsRNA cleavage by RNase III endonuclease domain (endoND)

All RNase III-like endonucleases are dimeric and cleave dsRNA. The cleavage products of 5'-phosphate and 3'-OH with a 2nt overhang are characteristic of all RNAi pathways [29]. Bacterial RNases III and yeast Rnt1p are true dimers with one endoND in each subunit (Fig. 3A). Dicer and Drosha are pseudo-dimeric, and the two endoNDs are tandem in one polypeptide chain (Fig. 1C, Fig. 3B) [30].

Both endoND and the dimeric interface are highly conserved. A single endoND typically comprises seven α -helices. A pair of long perpendicular α -helices forms the active site as well as the dimer interface. (Fig. 3A, B). The active site is at the intersection of the two helices and consists of two pairs of carboxylates, one pair from each helix. Dimerization results in a valley-shaped dsRNA-binding surface along the length of the dimer [23,31], and two active sites ~ 20 Å apart are symmetrically located at the bottom of the valley. The placement of the active sites matches the two scissile bonds 2 base pairs apart across the minor groove (Figure 3C).

Divalent metal ions, preferably Mg^{2+} , are essential for catalysis. Although in many RNase III structures only one metal is observed, biochemical data [32–34] and crystallographic studies of *A. aeolicus* RNase III complexed with RNA [31] and *G. intestinalis* Dicer with Er^{3+} ions [12] indicate that dsRNA cleavage occurs via the two-metal ion mechanism (see later discussion on RNase H). Structures of bacterial RNase III-dsRNA complexes (Fig 3C) [23, 31] indicate that the specificity for dsRNA cleavage and not dsDNA or RNA/DNA hybrid is due to specific recognition of A-form duplex and 2'-OH groups.

Processing dsRNA by Dicer and Drosha-DGCR8

The only crystal structure of full-length Dicer to date is that of *G. intestinalis* protein (Gi-Dicer), which consists of two endoNDs, PAZ and platform domains (Fig. 3D) [12]. The structure of Gi-Dicer has been likened to a hatchet with the dimeric endoNDs forming the blade and the PAZ at the end of the handle. Gi-Dicer cleaves dsRNA to 25 bps with 2 nt overhang at the 3' end. Using structures of bacterial RNase III-dsRNA and PAZ-nucleic acid complexes as templates, a model of Gi-Dicer - dsRNA complex was proposed (Fig. 3D) [35]. The platform domain between the endoND and PAZ appears to have a positively charged flat surface for dsRNA binding and separates the cleavage site (endoND) from the 3' end (PAZ) by ~ 25 -bps. One can imagine that by slightly changing the size and shape of the platform domain, Dicer may generate 20 to 30 bp products.

Gi-Dicer may be the minimal Dicer that is functional in vitro and in vivo [35]. As depicted in Fig. 1C, Drosha-DGCR8 microprocessor complex and other Dicers contain dsRBD in addition to endoND and PAZ domains. DUF283 in Dicer has been suggested to resemble dsRBD based on sensitive sequence comparisons [36]. In bacterial RNase III, each dsRBD domain is tethered to an endoND and contributes significantly to dsRNA binding (Fig. 3C). The core of DGCR8 (DiGeorge syndrome critical region gene 8 in human, Pasha in *C. elegans*) contains two tandem dsRBDs and complements Drosha in pri-miRNA processing as efficiently as the full-length DGCR8 [37–39]. In the structure of the DGCR8 core domain, the dsRBDs interact extensively to form a compact structure [40]. If these dsRBDs bind a single dsRNA, the RNA has to be significantly bent. Alternatively they may interact with different segments or even separate molecules of pri-miRNA [40]. How dsRBDs increase the catalytic activity of Dicer awaits future experiments. The absence of dsRBD in Gi-Dicer may be compensated by the platform domain (Fig. 3D).

Both Drosha and Dicer can process imperfectly base-paired dsRNA (Fig. 1A). Bacterial RNases III can nick bulges present in dsRNA [41]. By analogy, Dicers may accommodate unpaired bulges and bubbles. Indeed Dicer is reported to cleave substrates with bulges or bubbles 6–10 bps away from the endoND active sites and located between the endoND and PAZ [35]. In addition the dsRBD of Dicer may facilitate the binding to a hairpin loop region as shown in the structure of dsRBD from Rnt1p (Fig. 2D).

Nicking of dsRNA by PIWI

Crystal structures of PIWI and Ago, which contains PIWI, reveal that the PIWI domain resembles RNase H, an endonuclease that nicks RNA strands in an RNA/DNA hybrid (Fig. 4A). Further biochemical analyses have confirmed the intrinsic endonucleolytic activity of Ago [42]. Two catalytically essential carboxylates of RNase H are conserved in Ago and PIWIs that possess the slicer activity, and the third carboxylate, which could be Asn in some RNases H, is replaced by His and occasionally Lys or Asp in Ago (Fig. 4A) [43]. RNase H cleaves RNA by the two-metal ion mechanism [44], the same as employed by RNase III. An important feature of two-metal ion catalysis is that metal-ion coordination in the active site is substrate dependent. This ensures substrate specificity – only when the proper substrate is bound can the metal ions be coordinated to promote catalysis [45]. This is probably why RNase H binds dsRNA and RNA/DNA hybrids but cleaves only RNA/DNA hybrids [46]. Similar to the RNase H specificity, PIWI domains have high affinities for ssDNA, medium affinities for RNA/DNA and dsDNA, and the lowest for dsRNA and ssRNA [27,47]. Yet the most efficient slicing activity has been observed with RNA/DNA hybrids followed by dsRNA [47].

RISC assembly

Ago is the core of the RISC complex, which uses ~22nt ssRNA processed as siRNA or miRNA to target and cleave mRNA. Functionally Ago recognizes 3' and 5' ends of a guide RNA, dsRNA formed between guide and target RNA, and cleaves target RNA at the 10th bp from the 5' end of guide RNA. Structurally, Ago can be divided into four domains: N-terminal, PAZ, Mid and RNase H-like PIWI domains. Except for the N domain, each has been ascribed with a specific function. The structures of Ago proteins with and without nucleic acid fragments [15,16,26,27,47] clearly show that each domain can move relative to the others (Fig. 4B). Recently, structures of *T. thermophilus* Ago (Tt-Ago) bound to a ssDNA mimicking a guide RNA [16] and RNA/DNA duplex [15] have been reported. In both structures the guide strand anchors the 5' phosphate in the Mid domain and the 3' -OH in PAZ (Fig 4C). In the RNA/DNA complex (with DNA as the guide strand) the duplex can only be traced for the first 10–11 base pairs from the 5' end of the guide strand. Nucleotides 10 and 11 of the target strand are in the vicinity of the active site and are partially disordered due to the mismatch introduced

there to prevent cleavage. Also visible in the electron density are the last 3 nucleotides of the guide strand bound to the PAZ domain. To accommodate the RNA/DNA duplex the central groove is significantly wider compared to other Ago structures. This is a result of large movements of the N-ter and PAZ domains which are flexible and can be compared to two lobster claws, while the Mid and PIWI (RNase H) domains of Ago may be compared to the lobster tail and head. When fully stretched, it can accommodate >10 bp of dsRNA or RNA/DNA from the 5'-end of a guide RNA. The structure demonstrates that this so-called "seed region" is key for target recognition through base pairing, which is also confirmed by biochemical data. Guide strands as short as 9 nt can support cleavage and introduction of bulges or mismatches in the seed region leads to a decreased cleavage, while even up to six contiguous mismatches were tolerated in the 3'-end region of the guide after the active site [15].

Conclusions and prospects

Tremendous progress has been made in understanding of the mechanism of RNAi from the structural perspective, but structures of key protein-RNA complexes remain to be determined, e.g. Drosha, the N-terminal domains in Droshas and Dicers, and Ago complexed with dsRNA substrates longer than 10 bps. In addition, new RNAi pathways are being discovered [8,48] and new Dicer and Slicer variations may emerge with different substrate specificity. Structural biology will likely continue to play a central role in enlightening RNAi pathways.

Acknowledgement

W. Yang is supported by NIDDK, NIH intramural research program. M. Nowotny is supported by the Wellcome Trust International Senior Research Fellowship and the EMBO Installation Grant. We thank Dr. R. Craigie for editing the manuscript.

REFERENCES

1. Bushati N, Cohen SM. microRNA functions. *Annu Rev Cell Dev Biol* 2007;23:175–205. [PubMed: 17506695]
2. Filipowicz W, Bhattacharyya SN, sonenberg N. Mechanisms of posttranscriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* 2008;9:102–114. [PubMed: 18197166]
3. Rana TM. Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* 2007;8:23–36. [PubMed: 17183358]
4. Chapman EJ, Carrington JC. Specialization and evolution of endogenous small RNA pathways. *Nat Rev Genet* 2007;8:884–896. [PubMed: 17943195]
5. Zamore PD, Haley B. the big world of small RNAs. *Science* 2005;309:1519–1524. [PubMed: 16141061]
6. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 2008;322:1387–1392. [PubMed: 19039138]
7. Klattenhoff C, Theurkauf W. Biogenesis and germline functions of piRNAs. *Development* 2008;135:3–9. [PubMed: 18032451]
8. Aravin AA, Hannon GJ, Brennecke J. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* 2007;318:761–764. [PubMed: 17975059]
9. Lingel A, Simon B, Izaurralde E, Sattler M. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* 2003;426:465–469. [PubMed: 14615801]
10. Ma JB, Ye K, Patel DJ. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature* 2004;429:318–322. [PubMed: 15152257]
11. Lingel A, Simon B, Izaurralde E, Sattler M. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* 2004;11:576–577. [PubMed: 15156196]
12. Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA. Structural basis for double-stranded RNA processing by Dicer. *Science* 2006;311:195–198. [PubMed: 16410517]

** The structure of a full-length and functional Dicer. It reveals an intramolecular dimeric mimic of RNase III and spatial arrangement of PAZ, spacer and catalytic RNase III domains.

13. Song JJ, Liu J, Tolia NH, Schneiderman J, Smith SK, Martienssen RA, Hannon GJ, Joshua-Tor L. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol* 2003;10:1026–1032. [PubMed: 14625589]
14. Yan KS, Yan S, Farooq A, Han A, Zeng L, Zhou MM. Structure and conserved RNA binding of the PAZ domain. *Nature* 2003;426:468–474. [PubMed: 14615802]
15. Wang Y, Juranek S, Li H, Sheng G, Tuschl T, Patel DJ. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* 2008;456:921–926. [PubMed: 19092929] ** This reports the most complete Ago-RNA substrate complex structure to date. It also shows the seed region of a nucleic acid substrate recognized by Ago.
16. Wang Y, Sheng G, Juranek S, Tuschl T, Patel DJ. Structure of the guide- strand-containing argonaute silencing complex. *Nature* 2008;456:209–213. [PubMed: 18754009]
17. Tian B, Bevilacqua PC, Diegelman-Parente A, Mathews MB. The doublestranded-RNA-binding motif: interference and much more. *Nat Rev Mol Cell Biol* 2004;5:1013–1023. [PubMed: 15573138]
18. Bycroft M, Grunert S, Murzin AG, Proctor M, St Johnston D. NMR solution structure of a dsRNA binding domain from *Drosophila* stau protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J* 1995;14:3563–3571. [PubMed: 7628456]
19. Kharrat A, Macias MJ, Gibson TJ, Nilges M, Pastore A. Structure of the dsRNA binding domain of *E. coli* RNase III. *EMBO J* 1995;14:3572–3584. [PubMed: 7628457]
20. Nanduri S, Carpick BW, Yang Y, Williams BR, Qin J. Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *EMBO J* 1998;17:5458–5465. [PubMed: 9736623]
21. Wu H, Henras A, Chanfreau G, Feigon J. Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc Natl Acad Sci U S A* 2004;101:8307–8312. [PubMed: 15150409] * This structure reveals a different mode of RNA recognition by dsRBD.
22. Ryter JM, Schultz SC. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J* 1998;17:7505–7513. [PubMed: 9857205]
23. Gan J, Tropea JE, Austin BP, Court DL, Waugh DS, Ji X. Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* 2006;124:355–366. [PubMed: 16439209] ** This gives the prototype of a catalytic Dicer-dsRNA complex.
24. Chanfreau G, Buckle M, Jacquier A. Recognition of a conserved class of RNA tetraloops by *Saccharomyces cerevisiae* RNase III. *Proc Natl Acad Sci U S A* 2000;97:3142–3147. [PubMed: 10716739]
25. Elbashir SM, Lendeckel W, Tuschl T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev* 2001;15:188–200. [PubMed: 11157775]
26. Song JJ, Smith SK, Hannon GJ, Joshua-Tor L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* 2004;305:1434–1437. [PubMed: 15284453] * This paper reports the first structure of a full-length and functional argonaute.
27. Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ. Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* 2005;434:666–670. [PubMed: 15800629]
28. Parker JS, Roe SM, Barford D. Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* 2005;434:663–666. [PubMed: 15800628] * These two papers give the first structural glimpse of the PIWI and Ago family.
29. Ohmichi T, Karimata H, Sugimoto N. Effect of secondary structure of short double-stranded RNA on RNAi efficiency. *Nucleic Acids Res Suppl* 2002;63–64. [PubMed: 12903106]
30. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 2004;18:3016–3027. [PubMed: 15574589]
31. Gan J, Shaw G, Tropea JE, Waugh DS, Court DL, Ji X. A stepwise model for double-stranded RNA processing by ribonuclease III. *Mol Microbiol* 2008;67:143–154. [PubMed: 18047582] * This paper gives details of the active site configuration and catalytic mechanism of RNase III.
32. Campbell FE Jr, Cassano AG, Anderson VE, Harris ME. insights into mechanism and conformational changes associated with binding and catalysis. *J Mol Biol* 2002;317:21–40. [PubMed: 11916377]

33. Meng W, Nicholson AW. Heterodimer-based analysis of subunit and domain contributions to double-stranded RNA processing by *Escherichia coli* RNase III in vitro. *Biochem J* 2008;410:39–48. [PubMed: 17953512]
34. Sun W, Pertzev A, Nicholson AW. kinetic and inhibitor evidence for the involvement of two magnesium ions in RNA phosphodiester hydrolysis. *Nucleic Acids Res* 2005;33:807–815. [PubMed: 15699182]
35. MacRae IJ, Zhou K, Doudna JA. Structural determinants of RNA recognition and cleavage by Dicer. *Nat Struct Mol Biol* 2007;14:934–940.940 [PubMed: 17873886] ** This paper shows how dicer “measures” dsRNA length for cleavage.
36. Dlakic M. DUF283 domain of Dicer proteins has a double-stranded RNA-binding fold. *Bioinformatics* 2006;22:2711–2714. [PubMed: 16954143]
37. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. *Nature* 2004;432:231–235. [PubMed: 15531879]
38. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. The Microprocessor complex mediates the genesis of microRNAs. *Nature* 2004;432:235–240. [PubMed: 15531877]
39. Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, Sohn SY, Cho Y, Zhang BT, Kim VN. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 2006;125:887–901.901 [PubMed: 16751099] * This contains excellent biochemical analysis of Drosha-DGCR8 activity.
40. Sohn SY, Bae WJ, Kim JJ, Yeom KH, Kim VN, Cho Y. Crystal structure of human DGCR8 core. *Nat Struct Mol Biol* 2007;14:847–853. [PubMed: 17704815]
41. Nicholson AW, Niebling KR, McOsker PL, Robertson HD. Accurate in vitro cleavage by RNase III of phosphorothioate-substituted RNA processing signals in bacteriophage T7 early mRNA. *Nucleic Acids Res* 1988;16:1577–1591. [PubMed: 3279395]
42. Rivas FV, Tolia NH, Song JJ, Aragon JP, Liu J, Hannon GJ, Joshua-Tor L. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nat Struct Mol Biol* 2005;12:340–349. [PubMed: 15800637]
43. Tolia NH, Joshua-Tor L. Slicer and the argonautes. *Nat Chem Biol* 2007;3:36–43. [PubMed: 17173028]
44. Nowotny M, Gaidamakov SA, Crouch RJ, Yang W. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* 2005;121:1005–1016. [PubMed: 15989951]
45. Yang W, Lee JY, Nowotny M. Making and breaking nucleic acids: two- Mg^{2+} -ion catalysis and substrate specificity. *Mol Cell* 2006;22:5–13. [PubMed: 16600865]
46. Nowotny M, Gaidamakov SA, Ghirlando R, Cerritelli SM, Crouch RJ, Yang W. insight into HIV reverse transcription. *Mol Cell* 2007;28:264–276. [PubMed: 17964265]
47. Yuan YR, Pei Y, Ma JB, Kuryavyy V, Zhadina M, Meister G, Chen HY, Dauter Z, Tuschl T, Patel DJ. Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol Cell* 2005;19:405–419. [PubMed: 16061186]
48. Faehnle CR, Joshua-Tor L. Argonautes confront new small RNAs. *Curr Opin Chem Biol* 2007;11:569–577. [PubMed: 17928262]

Figure 1. Major RNA interference pathways and domain composition of proteins involved in them (A) Processing of RNAs in miRNA and siRNA. Pri-miRNA transcripts are processed by Drosha/DGCR8 in the nucleus and Dicer in the cytoplasm. The cleavage sites are indicated by orange triangles. siRNAs start as dsRNA and are processed by Dicer. In both cases the end product 20–30 nt RNA is incorporated into the RISC complex. (B) Model of piRNA pathway. Transposon mRNA is cleaved by PIWI using a piRNA guide strand. A fragment of cleaved transposon RNA is incorporated into Ago3 RISC and used to cleave the piRNA cluster transcript. Its fragment is incorporated into PIWI protein to start another round of the cycle. (C) Domain composition of proteins involved in RNA interference.

Figure 2. Domains involved in nucleic acid binding in RNAi

(A) PAZ domain of human Argonaute eIF2c1 (PDB: 1SI3). (B) Comparison of nucleic acid binding by PAZ domains from eIF2c1 (green) and *T. thermophilus* Ago (PDB: 3DLH) (blue). The nucleic acid from each structure is shown in cartoon representation (light green for eIF2c1 structure and purple blue for *T. thermophilus* Ago structure). The conserved tyrosine that binds the nucleic acid backbone is shown in ball-and-stick representation. (C) dsRBD (ds-RNA-binding domain) from *X. leavis* RNA-binding protein A bound to a dsRNA duplex (PDB: 1DI2). (D) dsRBD from Rnt1 bound to a tetraloop-stem RNA (PDB: 1T4L). (E) Mid domain from *A. fulgidus* PIWI protein bound to ssRNA (PDB: 1YTU). The C-terminal carboxylate is shown in red and the metal ion coordinating the 5'-phosphate of the RNA in green. Conserved

residues stabilizing the 5' end of the RNA are shown in balls-and-sticks. (F) Close-up view of the interactions stabilizing the 5' end of the RNA.

Figure 3. RNases III

(A) *A. aeolicus* RNase III – intermolecular homodimer. The two endoNDs from two subunits are shown in orange and pink with the central perpendicular helices shown in darker colors. The residues forming the active site are shown in colored balls-and-sticks and those forming the hydrophobic core of the dimer interface in white balls-and-sticks. (B) *G. intestinalis* Dicer – the intramolecular dimerization of two endoND domains. The linker domain inserted between them is shown in light grey. (C) The structure of *A. aeolicus* RNase III in complex with dsRNA (PDB: 2NUF). dsRBDs from each dimer subunit are in purple and blue. The RNA is shown in red. (D) The structure of *G. intestinalis* Dicer (PDB: 2FFL). The PAZ domain is shown in

green and the platform domain in light blue. dsRNA modeled to interact with the endoND and PAZ domains is shown in yellow.

Figure 4. Argonautes

(A) The RNase H-like PIWI domain of *P. furiosus* Argonaute. The central β -sheet is shown in red and the catalytic residues – two aspartates and a histidine as balls-and-sticks. The site of an ~80 residue insertion into the RNase H fold, which is mainly composed of β -strands, is indicated with a gray dashed line. (B) Comparison of apo *P. furiosus* Ago structure (left, PDB: 1Z25) and *T. thermophilus* Ago in complex with 21-nt ss DNA (right, PDB: 3DLH). The protein structures were superimposed using the positions of the active site residues in PIWI (RNase H-like) domain. The central β -sheets of nucleic acid binding domains are shown in color (N-ter - blue, PAZ - green, Mid - yellow, and RNase H-like in red) and the differences in their positions between the two structures are indicated with arrows. (C) Structure of *T. thermophilus* Ago complexed with an RNA/DNA hybrid (PDB: 3F73). The nucleic acid is

shown in cartoon representation (RNA in pink and DNA in light blue). The central β -sheets of each domain are colored as in B. The active site carboxylates are shown in ball-and-sticks.