Chapter 2
The microRNA Machinery

Thomas C. Roberts

Abstract MicroRNAs (miRNAs) are short (~22 nucleotides) single-stranded RNA molecules that primarily function to negatively regulate gene expression at the post-transcriptional level. miRNAs have thus been implicated in the regulation of a wide variety of normal cell functions and pathophysiological conditions. The miRNA machinery consists of a series of protein complexes which act to: (1) cleave the precursor-miRNA hairpin from its primary transcript (i.e. DROSHA and DGCR8); (2) traffic the miRNA hairpin between nucleus and cytoplasm (i.e. XPO5); (3) remove the loop sequence of the hairpin by a second nucleolytic cleavage reaction (i.e. DICER1); (4) facilitate loading of the mature miRNA sequence into an Argonaute protein (typically AGO2) as part of the RNA-Induced Silencing Complex (RISC); (5) guide the loaded RISC complex to complementary, or semi-complementary, target transcripts and (6) facilitate gene silencing via one of several possible mechanisms.

Keywords Argonaute • AGO2 • Dicer • DICER1 • Exportin-5 • XPO5 • Drosha • microRNA • DGCR8

Introduction

Microribonucleic acids (microRNAs, miRNAs) are short (21–23 nucleotides), single-stranded, non-coding RNA molecules that are encoded in the genomes of higher organisms. miRNAs primarily function as post-transcriptional gene expression regulators [1] and miRNA-mediated regulation has been implicated in a wide variety of cellular processes and disease conditions. As such, miRNAs are of interest as potential therapeutic targets [2–4] and as disease biomarkers [5–7].

T.C. Roberts (✉)
Development, Aging and Regeneration Program, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA

Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, UK
e-mail: troberts@sbpdiscovery.org; thomas.roberts@dpag.ox.ac.uk

G. Santulli (ed.), *microRNA: Basic Science*, Advances in Experimental Medicine and Biology 887, DOI 10.1007/978-3-319-22380-3_2

miRNAs are embedded within hairpin structures in long (100s–1000s of nucleotides) primary-miRNA (pri-miRNA) transcripts that are transcribed by RNA Polymerase II [8]. pri-miRNAs are capped and polyadenylated transcripts that can be either long non-coding RNAs or conventional protein-coding mRNAs [9] (in which case the hairpin resides in the intronic sequence of the pre-mRNA [10]). The majority of miRNAs are transcribed from independent transcriptional units, although some miRNAs are arranged in polycistronic clusters [11]. For example, the lncRNA gene *MIR17HG* contains the miR-17-92 cluster implicated in tumorigenesis and other diseases [12].

The miRNA hairpin, termed the precursor-miRNA (pre-miRNA), is ~70 nucleotides in length and typically contains multiple bulges at mismatched nucleotides. The mature miRNA species is generated from the pri-miRNA by two sequential endonucleolytic processing reactions. The pre-miRNA is liberated from the primary transcript by the first RNase cleavage reaction, whereas the loop sequence is removed by the second cleavage reaction. These two processing stages are spatially separated such that they occur in the nucleus and cytoplasm, respectively [13]. Subsequently, the mature miRNA is then loaded into RISC (RNA-Induced Silencing Complex) where it acts to guide the complex to target mRNAs and silence their expression [14–18].

This chapter focuses on the canonical mode of miRNA maturation and function, focusing on the protein components which comprise this pathway. Notably, there are numerous caveats and exceptions. These include miRNAs processed from non-standard precursor RNA substrates, miRNAs which bypass various processing stages, and miRNAs which execute non-canonical functions in the nucleus or extracellular space. Such deviations from the canonical paradigm are reviewed elsewhere [19, 20].

While much of the seminal work on miRNAs was performed in model organisms such as *Drosophila melanogaster* and *Caenorhabditis elegans*, this chapter is primarily concerned with the human miRNA machinery for the purpose of medical relevance. HGNC gene symbols are used throughout with commonly used non-standard names in parentheses where applicable.

DROSHA, DGCR8

In human cells, the pre-miRNA hairpin is cleaved from the pri-miRNA transcript in the nucleus by the enzyme DROSHA (Drosha). DROSHA is 145–160 kDa protein [21] with predominantly nuclear localization [22, 23]. The importance of DROSHA is exemplified in the roundworm *Meloidogyne incognita*, where knock-down of Drosha results in embryonic lethality [24]. Similarly, a conditional *Drosha* knockout in postnatal murine testes revealed an essential role for *Drosha* in spermatogenesis [25]. DROSHA belongs to the type III ribonuclease class of enzymes, which characteristically introduce staggered cuts in their RNA substrate duplexes [26]. DROSHA contains two RNase III domains (RIIIDa and RIIIDb). These domains form an intramolecular dimer as they are arranged in close proximity in

three-dimensional space. As such, RIIIDa cleaves the 3′ strand of the hairpin whereas RIIIDb cleaves the 5′ strand [27]. Immunopurified DROSHA is capable of cleaving synthetic pri-miRNAs to generate pre-miRNAs in vitro [23]. Furthermore, silencing of DROSHA by RNA interference (RNAi) resulted in accumulation of pri-miRNAs and depletion of pre-miRNAs [23].

In vivo, DROSHA associates with a multitude of additional proteins including double-stranded RNA binding proteins (dsRBDs), hnRNPs, DEAD-box/DEAH-box RNA helicases and DGCR8 [21]. The *DGCR8* (DiGeorge Syndrome Critical Region 8) gene resides in a region of chromosome 22 that is deleted in sufferers of DiGeorge syndrome, a rare genetic disease with varied symptoms including congenital heart disease, learning difficulties, facial abnormalities and immune dysfunction [28, 29]. Immunoprecipitation of DGCR8 co-precipitated only DROSHA, and the resulting eluates exhibited robust pri-miRNA in vitro cleavage activity [21]. Similarly, recombinant DROSHA and DGCR8 generated in insect cells could reconstitute pri-miRNA cleavage activity when the two proteins were combined. However, either recombinant protein in isolation was insufficient for pri-miRNA processing [21]. Interestingly, DROSHA exhibited some level of non-specific RNase activity which was inhibited when DGCR8 was present. Depletion of DROSHA or DGCR8 by RNAi also abrogated pri-miRNA processing. Together, these studies demonstrated that the DROSHA–DGCR8 complex is necessary and sufficient for pri-miRNA processing [21]. This complex of the DROSHA and DGCR8 proteins is collectively known as the Microprocessor. Similar functional associations between DROSHA and DGCR8 homologues have also been demonstrated in *D. melanogaster* and *C. elegans* [30, 31] (Note: In these organisms, DGCR8 is called Pasha, partner of Drosha, or Pash-1).

Biochemical studies have dissected the substrate preference of DROSHA. A study by Zeng et al. suggested that DROSHA has a strong preference for pre-miRNA hairpins with large (≥10 nucleotide) unstructured terminal loop sequences [32]. However, Han et al. found that the loop sequence is dispensable for DROSHA processing, and that synthetic "hairpins" in which the loop is replaced by an extended duplex with open termini are easily processed [33]. Both studies identified the importance of single-stranded RNA regions flanking the base of the lower stem for recognition of the pre-miRNA by the Microprocessor complex [32, 33]. Thermodynamic analysis of hundreds of human and *D. melanogaster* pri-miRNA hairpins revealed that the DROSHA cleavage site is typically ~11 nucleotides (~1 helical turn) from the ssRNA–dsRNA stem junction and ~22 nucleotides (~2 helical turns) from the terminal loop, suggesting that DROSHA measures the distance from either the base of the stem or the loop in order to determine the cleavage site. Interestingly, pre-miR-30a does not effectively compete with pri-miR-30a, suggesting that the major site of pri-miRNA recognition by DGCR8 resides outside of the pre-miRNA hairpin structure [33].

Structures of intact DROSHA are currently lacking, although the C-terminal dsRBD has been solved by solution NMR [34]. This structure consists of the αββββα fold typical of dsRBDs (Fig. 2.1a). However, biochemical assays (EMSA and immunoprecipitation with radiolabelled synthetic pri-miRNA) failed to show RNA binding by DROSHA alone [33]. Furthermore, competition assays show that

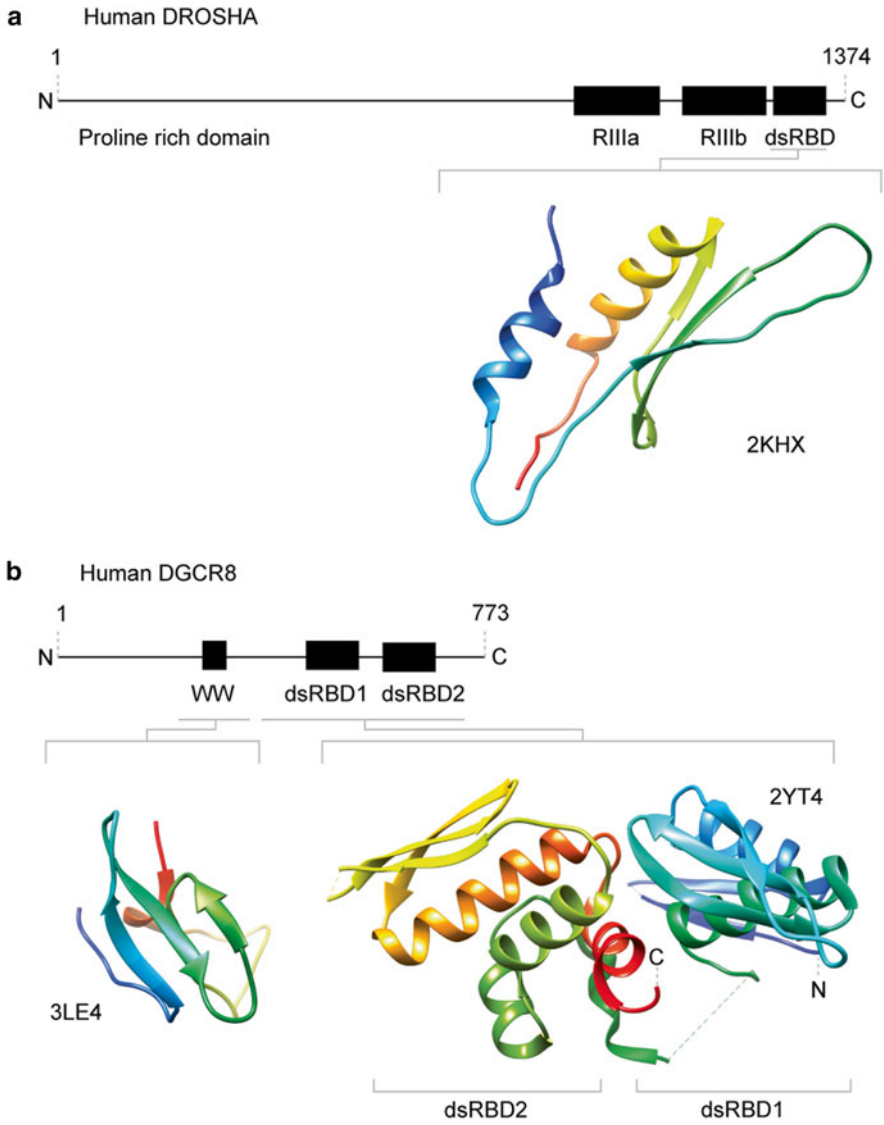

Fig. 2.1 Structures of DROSHA and DGCR8. (**a**) Domain structure of human DROSHA protein. A depiction of a solution NMR structure of the dsRNA binding domain (dsRBD) is shown (2KHX) [34]. (**b**) Domain structure of human DGCR8 protein and depictions of available crystal structures for the dimerization domain (3LE4) [36] and core domain (2YT4) [35]

processed pre-miRNAs rapidly dissociate from DGCR8, and that DROSHA exhibits low RNA binding activity [33]. As a result, it is likely that DGCR8 docks with the pri-miRNA by anchoring to the ssRNA–dsRNA stem junction in order to locate the cleavage site for DROSHA. DROSHA then associates with hairpin only transiently in order to perform catalytic pri-miRNA cleavage [33].

DGCR8 consists of a core region containing two dsRBDs near the C-terminus and a central WW motif which serves to mediate the interaction between DGCR8 and the proline-rich region within the N-terminal region of DROSHA. High-resolution crystal structures have been published for both of these domains [35, 36] (Fig. 2.1b). In the core domain structure (2.6 Å resolution), the two dsRBDs form a butterfly-like shape whereby the dsRBDs exhibit pseudo twofold symmetry [35]. Each dsRBD adopts an αββββα fold structure. A C-terminal helix and a flexible linker region act to bridge the two dsRBDs. Using FRET, the authors showed that pri-miRNA binding is unlikely to induce large-scale changes in DGCR8 conformation [35]. The structure of the dimerization domain (1.7 Å resolution) shows that the WW motif resides within a heme binding domain and forms a domain-swapped dimer structure [36]. DGCR8 has also been shown to bind to heme, which promotes self-dimerization [37]. Furthermore, in the absence of heme, the heme binding region inhibits the pri-miRNA processing activity of DGCR8, thereby acting as an autoinhibitory domain [37].

XPO5

The nuclear membrane which separates the nucleoplasm from the cytoplasm is punctuated by nuclear pores. The nuclear pore complex is a multi-protein structure that facilitates regulated trafficking of macromolecules through the nuclear pore. While small molecules can passively diffuse between compartments, larger macromolecules (e.g. the pre-miRNA hairpin) require active transport.

In human cells, the Karyopherin XPO5 (Exportin-5) [38] facilitates the transport of pre-miRNA hairpins though the nuclear pore complex and into the cytoplasm in a RAN-GTP-dependent manner [39–41]. XPO5 binds to dsRNA in a sequence-independent manner [41], but recognizes a minihelix motif that is common to pre-miRNAs and several other substrates including tRNA, Y1 RNA and the adenoviral VA1 RNA [42, 43]. Disruption of XPO5 function by RNAi, inhibition with antibodies or competition with VA1 RNA leads to a reduction in the levels of mature miRNAs [39–41, 44].

A common feature of the Exportins is that they take advantage of the gradient of RAN-GTP that exists across the nuclear envelope. This gradient is the result of the differential nucleocytoplasmic location of protein factors which regulate the GTP binding status of RAN. Specifically, RANGAP1 (RanGAP, RAN-GTPase Activating Protein) is cytoplasmic [45], and RCC1 (an RAN exchange factor) is nuclear [46]. In the nucleus, the XPO5 forms a trimeric complex with its pre-miRNA cargo and RAN-GTP. The complex is subsequently translocated through the nuclear pore complex. Once in the cytoplasm, the GTP is hydrolyzed to GDP which induces a conformational change in RAN with concomitant dissociation of the complex and release of the pre-miRNA hairpin. Binding of XPO5 to the pre-miRNA is dependent on the binding of the RAN-GTP in complex with GTP, as determined by electrophoretic mobility shift assay [40]. Furthermore, depletion of RAN-GTP by nuclear microinjection of Ran-GAP in *Xenopus* oocytes also inhibited miRNA export [41].

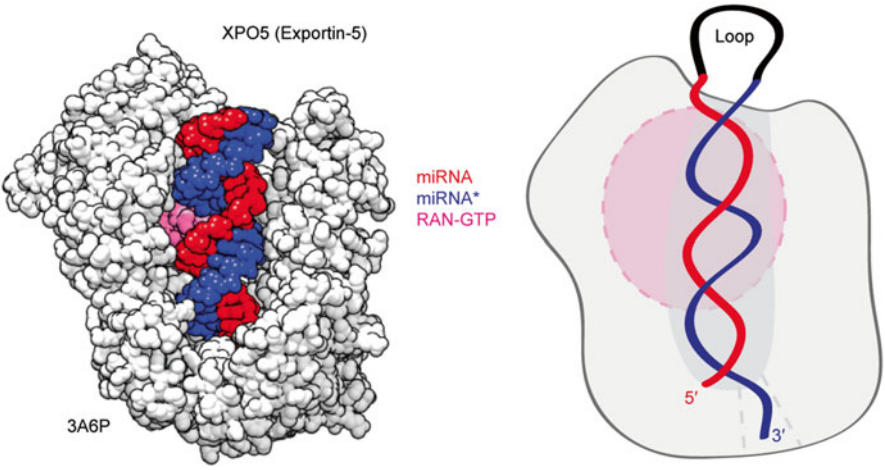

Fig. 2.2 Structure of XPO5. Depiction of the XPO5 (Exportin-5) structure. Structure (3A6P) [48] and a cartoon schematic are shown side by side. XPO5 protein is shown in *white* and RAN-GTP in *pink*. The pre-miRNA hairpin is shown bound to XPO5 with the 5′ strand in *red* and the 3′ strand in *blue*. The loop sequence (not visible in the crystal structure) is in *black*

The major determinants of pre-miRNA recognition by XPO5 are a helical stretch (~16 base pairs) and 3′ terminal overhangs (produced by DROSHA processing) [42, 43, 47]. As a result, XPO5 interacts with the majority of the pre-miRNA hairpin (excluding the terminal loop). Interestingly, the binding of XPO5 to a pre-miRNA has the additional effect of reducing degradation of the hairpin in the nucleus [47].

A high-resolution (2.9 Å) crystal structure of human XPO5 has been reported in complex with a fragment of canine Ran-GTP and the pre-miR-30a hairpin [48] (Fig. 2.2). XPO5 comprises multiple HEAT repeats and forms an overall "baseball mitt" shaped structure with Ran-GTP binding towards the top of the mitt. At the base of the mitt is a tunnel-like structure. The pre-miRNA hairpin forms an A-form helix that sits in the interior of the mitt (as if being held in the palm of the hand). The terminal nucleotides are oriented towards the base of the mitt with the 3′ overhang nucleotides inserted into the tunnel. The interior of the tunnel consists of several basic residues which make a number of close contact hydrogen bonds and salt bridges to the sugar-phosphate backbone of the overhang. Furthermore, an Arginine residue sidechain (Arg602) acts to sterically inhibit the double-stranded portion of the pre-miRNA from entering the tunnel. In silico modelling showed that a putative hairpin with a 5′ overhang would sterically clash with neighbouring HEAT repeat domains [48]. These observations explain the selectivity of XPO5 for cargo molecules with 3′ overhangs. In addition, the close contacts of XPO5 with the terminal nucleotides also explain why XPO5-bound pre-miRNAs are protected from nucleolytic degradation, as the terminal nucleotides are shielded by the surrounding protein.

The remainder of hairpin stem makes numerous contacts with the basic surface of the "mitt" interior [48]. This suggests that XPO5 primarily interacts with the

pre-miRNA hairpin via ionic interactions, which account for the sequence-independence of cargo recognition by XPO5. In further support, high ionic strength buffers promote dissociation of the pre-miRNA from XPO5 [48].

DICER1

In the cytoplasm, the pre-miRNA hairpin is subjected to a second cleavage reaction catalyzed by a second RNase III family member DICER1 (Dicer) which removes the terminal loop sequence [49]. DICER1 acts as a "molecular ruler" that cleaves the duplex at a defined distance from the terminus. (Similarly, human DICER1 will also progressively process long dsRNA molecules by cleaving ~22 nucleotide duplex fragments sequentially from their termini [50]).

Accuracy of DICER1 processing is important, as shifts in the site of cleavage will generate mature miRNAs with non-canonical seed sequences. This will lead to retargeting of the miRNA to a new set of mRNAs and/or potentially alter the strand selection preference [51]. Such an eventuality would effectively diminish the cellular concentration of correctly processed miRNAs and generate non-physiological miRNA–target interactions—with both outcomes being potentially detrimental, or at the very least wasteful, to the cell.

Dicer was initially discovered in *Drosophila* cell extracts where it was shown to cleave long dsRNA substrates into ~22 nucleotide siRNA duplexes [26]. Additionally, further studies showed that Dicer also generates mature miRNA species from let-7 precursors, thereby demonstrating its role in miRNA processing [52–54]. Genetic ablation of *Dicer1* in mice is lethal at the early stages of embryonic development on account of stem cell depletion [55]. (miRNAs are required for correct control of stem cell division [56].)

Human DICER1 is a ~218 kDa protein that is predominantly cytoplasmic and colocalizes with the endoplasmic reticulum marker calreticulin [49]. DICER1 consists of a number of conserved globular domains. The N-terminus of the protein contains three helicase domains (HEL1, HEL2 and HEL2i). The central region contains the platform domain and PAZ (Piwi/Argonaute/Zwille) domain. The PAZ domain is important for recognizing the RNA termini of the substrate duplex [57]. The C-terminal end contains a dsRBD and two catalytic RNase III-like domains (RIIIa and RIIIb) [49]. As with DROSHA, the two RNase III domains (RIIIa and RIIIb) of DICER1 are oriented in space such that they form an intramolecular dimer [58, 59]. As a result, DICER1 has a single catalytic centre at which both phosphodiester cleavage reactions occur (one on each strand of the substrate duplex). This configuration is responsible for the generation of the characteristic 2 nucleotide 3′ overhangs [58–60].

Human DICER1 requires the presence of Mg^{2+} for substrate cleavage but not binding, and catalytic activity is sensitive to ionic strength suggesting that electrostatic substrate–enzyme interactions are important for DICER1 functionality [49]. In contrast to *D. melanogaster* Dicer (Dcr-2) [61], human DICER1 does not require ATP for RNA cleavage to occur [49, 50].

There are currently no crystal structures of intact human DICER1 available, although a high-resolution (3.3 Å) structure of an intact Dicer from the protozoan *Giardia intestinalis* was published in 2006 [59]. This structure revealed that the PAZ domain is separated from the catalytic site by ~65 Å which corresponds with the length of dsRNA products (25–27 base pairs) generated by the G. *intestinalis* Dicer. Deletion of the PAZ domain in *G. intestinalis* results in loss of the size specificity of cleavage products [62] suggesting that the PAZ domain is required for molecular ruler functionality. Furthermore, the surface connecting the PAZ and RNase III domains contains a number of positively charged amino acid residues which, when substituted with alanine by site-directed mutagenesis, reduce the activity of Dicer [62]. These findings suggest that this region makes important electrostatic interactions between the substrate phosphodiester backbone and positively charged surface residues.

Several studies have utilized electron microscopy in order to determine structures for human DICER1 [63–65]. Most notably, Lau et al. showed that human DICER1 forms an "L"-shaped structure, and mapped the known globular domain structures to this general shape [65]. As such, the PAZ and platform domains were found to be tightly associated at the head of the "L," the RNase III domain was located centrally, and the helical domains formed the base, or "arm" of the L shape. Importantly, in the human DICER1 structure, the PAZ-RNase III domain distance is reduced relative to *G. intestinalis*, consistent with shorter human dsRNA cleavage products and "molecular ruler" functionality [65]. Furthermore, the helicase domains formed distinct lobes which the authors described as a "clamp" like structure. This configuration suggested that the substrate RNA duplex resides in a groove that runs vertically along the "L" shape [65].

Studies on reconstituted Dicer in *G. intestinalis* led to the so-called 3′ counting rule whereby the location of the Dicer cleavage event is determined by "counting" ~22 nucleotides along the duplex starting at the 3′ terminal nucleotide anchored in the PAZ domain. However, it was subsequently shown in human cells that the phosphorylated 5′ terminus of the pre-miRNA hairpin is also anchored in a binding pocket within the PAZ domain, and that a "5′ counting rule" predominates in humans [66]. (Interestingly this binding pocket is not conserved in *G. intestinalis*.) Additionally, a "loop counting rule" is also operative in human cells whereby DICER1 "counts back" from the loop (or bulged sequence) at the closed end of the hairpin [67]. Consequently, human DICER1 utilizes the 3′ counting, 5′ counting, and loop counting mechanisms in order to measure the length of the substrate duplex and precisely determine the site of cleavage.

AGO2, RISC

Following DICER1 processing, the miRNA duplex is passed to the RNA Induced Silencing Complex (RISC, or miRISC) and specifically to an Argonaute protein which constitutes the core of the complex [68]. Subsequently, one strand of the duplex is discarded to leave only a single-stranded mature ~23 nucleotide miRNA

bound to RISC. RISC loading is asymmetric such that the Argonaute protein shows a preference for the strand with the least thermodynamically stable (and therefore most easily unwound) 5′ terminus [69, 70]. However, loading is probabilistic, and the opposite "passenger" strand (often denoted as miRNA*) may be incorporated in a minority of cases. Furthermore, the termini of some miRNA duplexes are equally stable, leading to a mixture of 5′ arm- and 3′ arm-loaded RISC complexes.

The miRNA-loaded RISC complex is then able to scan the pool of cytoplasmic mRNAs for potential complementary targets. miRNAs primarily target the 3′ untranslated region (UTR) regions of mRNAs where they typically bind with imperfect complementarity [71, 72]. It has been estimated that >60% of human protein-coding transcripts have been under selective pressure to maintain miRNA binding sites [73], suggesting that miRNA-mediated gene regulation is a widespread phenomenon. The degree of miRNA-target base-pairing complementarity determines the fate of the target transcript. Perfect complementarity leads to target cleavage via the "slicer" activity of AGO2, analogous to small interfering RNA-induced RNAi [74, 75]. Slicer cleavage occurs in the target mRNA at the intervening scissile phosphate group opposite to nucleotides 10 and 11 in the mature miRNA sequence. In contrast, incomplete complementarity triggers mRNA silencing by distinct mechanisms which may involve translational repression, slicer-independent mRNA degradation and/or sequestration in cytoplasmic processing bodies (P-bodies) [76]. In this case, there is typically complete complementarity between the seed region (nucleotides 2–8) and the target mRNA with scattered base-pairing and bulged nucleotides in the remainder of the duplex [15]. Importantly, translational repression via sequestration in P-bodies is a reversible process suggesting that mRNAs can be stored in a translationally inactive form and then released in response to cell stress [77]. In humans, slicer-independent mechanisms of mRNA repression dominate on account of the majority of miRNAs lacking complementarity with their cognate targets around position 10/11. (The resulting bulge structure inhibits the slicer activity of AGO2 [78].) A detailed discussion of the mechanisms of RISC-mediated gene silencing is beyond the scope of this chapter and has been discussed elsewhere [1, 17, 79–81].

miRNAs can exert complex, combinatorial control over gene expression as one miRNA can target multiple mRNAs [82]. In this manner, one miRNA can target a family of transcripts with related functions in order to regulate a cellular process. For example, miR-29 family miRNAs repress a plethora of fibrosis-associated transcripts (collagens, fibrillins, elastin, fibronectin, etc.) [4]. Down-regulation of miR-29 is therefore often a feature of pathogenic fibrotic processes. On the other hand, a typical mRNA 3′ UTR contains binding sites for multiple miRNAs. This enables the transcript to integrate signals from different miRNAs, or to fine-tune expression (with the expression of the transcript inversely proportional to the number of miRNAs which are regulating it).

Transfer of the processed miRNA duplex to an Argonaute protein is accomplished by the RISC-loading complex (RLC), which consists primarily of an Argonaute protein, DICER1 (as discussed above) and TARBP2 (trans-activation response (TAR) RNA Binding Protein, also known as TRBP) [64, 83]. The direct association of DICER1 with an Argonaute protein promotes the transfer of the

miRNA duplex between the two proteins (TARBP2 also contributes to this process) [83, 84]. It has been suggested that binding by TARBP2 may allow the siRNA intermediate to stay associated with the RLC after release from DICER1 and may also help optimize the orientation of the siRNA for AGO2 loading [64]. Depletion of TARBP2 leads to a reduction in pre-miRNA processing [84], although TARBP2 is not required for DICER1 activity [83]. Instead, TARBP2 appears to be required for recruitment of AGO2 to the DICER1-miRNA complex [83]. Complexes of human DICER1 with TARBP2 [63] and with both TARBP2 and AGO2 have been solved by electron microscopy [64]. These studies suggest that AGO2 contacts the C-terminal region of DICER1, whereas TARBP2 interacts with the DExH/D domain [64].

In humans, there are four Argonaute proteins (AGO1–4). These homologous proteins appear to execute redundant functions in terms of miRNA-mediated gene silencing and have similar preferences for endogenous miRNAs or exogenous synthetic siRNAs [85]. Each Argonaute protein consists of four major domains: N-terminal (N), PAZ, MID and PIWI [86]. However, only AGO2 can mediate target mRNA "slicing" on account of an Asp-Glu-Asp-His (DEDH) catalytic tetrad at its active site [87, 88]. AGO2 has thus been called "the catalytic engine of RNAi" [78, 85]. The miRNA/siRNA-AGO2 complex is a multiple turnover enzyme such that after target cleavage, the loaded RISC complex can bind to another target and thereby induce multiple further gene silencing events [74]. In contrast, AGO1, AGO3 and AGO4, which lack the catalytic residues required for slicer functionality, mediate gene silencing via slicer-independent mechanisms only [85].

Ago2 knockout mice exhibit embryonic lethality [85, 89, 90], and transgenic mice that are homozygous for a catalytically deficient Ago2 die shortly after birth as a result of anaemia [91]. The slicer activity of Ago2 is uniquely required for Dicer-independent maturation of miR-451, an miRNA which is essential for haematopoiesis [91, 92].

In 2012, a high-resolution (2.3 Å) crystal structure of full length human AGO2 in complex with a heterogeneous mixture of guide RNAs was published by the McRae lab that revealed new insights into its function [93] (Fig. 2.3). AGO2 forms a bilobular structure with a central groove which accommodates the mature miRNA-target mRNA duplex. The first seven nucleotides are held in a well-defined, uniform conformation [93]. (The structure of AGO2 has also been likened to that of a duck, with the MID, PIWI and N domains forming the "body" and the PAZ domain the "head" of the "duck" [94].) AGO2 interacts with the guide RNA in a sequence-independent manner on account of multiple electrostatic interactions with the phosphate backbone and van der Waals interactions with the ribose sugar [93]. The guide RNA adopts an A-form conformation with nucleotides 2–6 "splayed out" such that their base-pairing surface is solvent accessible and available to interact with a target mRNA [93]. The 3′ terminal nucleotide was found to bind in a pocket residing within the PAZ domain. Interestingly, the guide RNA was found to kink after nucleotide 6, after which the A-form conformation of the guide was disrupted and the remainder of the nucleotides were less well ordered [93]. These observations provide a structural basis for the importance of seed sequence base-pairing as the primary determinant of miRNA-target recognition [15, 95]. The AGO2 structure also provided evidence of two tryptophan binding pockets within the PIWI domain which are a likely docking site for other RISC cofactors such as TNRC6A [93].

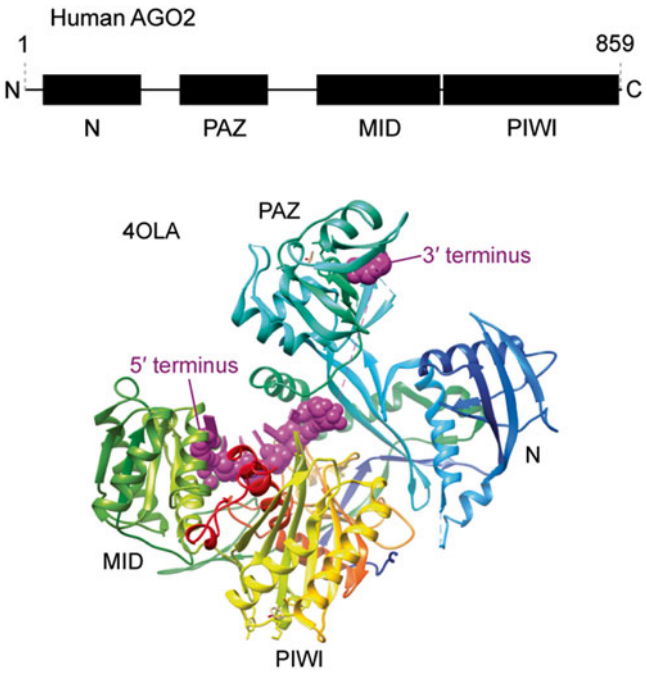

Fig. 2.3 Structure of AGO2. Domain organization for human AGO2 and depiction of a crystal structure of the full length protein (4OLA) [93]. The guide RNA is shown in *purple* (the middle portion of the guide RNA was too disordered to be resolved)

Subsequently, a similar structure (2.2 Å resolution) of human AGO2 in complex with miR-20a was published [94]. This structure clearly showed that the miRNA contacts multiple AGO2 subunits with the 5′ miRNA terminus bound in a pocket within the MID domain (consistent with previous studies [96, 97]) and that the 3′ terminal nucleotide bound in the PAZ domain [94]. Interestingly, this study also showed that miRNA binding stabilizes AGO2 such that it becomes resistant to proteolytic degradation [94].

A complete description of RISC components is currently lacking, although immunoprecipitation-mass spectrometry studies have identified a multitude of Argonaute-binding proteins [98–100]. Notably, a trinucleotide repeat containing protein partner, TNRC6A (also known as GW182), is required for miRNA function and is localized to P-bodies [100].

Conclusion

In summary, mature miRNAs are generated in a multi-step process requiring several enzymatic cleavage reactions and multiple protein cofactors. Biochemical and structural studies have revealed much of the mechanistic detail of how these

proteins operate. However, a complete understanding of how the miRNAs are loaded into the RISC complex, how RISC mediates gene silencing, and the regulatory control operating at each step is currently lacking. Future studies will likely shed light on these issues and potentially reveal new details of disease miRNA-associated pathophysiology, or improved methods for modulating miRNA function for therapeutic purposes.

References

1. Filipowicz W, Jaskiewicz L, Kolb FA, Pillai RS. Post-transcriptional gene silencing by siRNAs and miRNAs. Curr Opin Struct Biol. 2005;15:331–41.
2. Roberts TC, Wood MJA. Therapeutic targeting of non-coding RNAs. Essays Biochem. 2013; 54:127–45.
3. Cacchiarelli D, Incitti T, Martone J, Cesana M, Cazzella V, Santini T, Sthandier O, Bozzoni I. miR-31 modulates dystrophin expression: new implications for Duchenne muscular dystrophy therapy. EMBO Rep. 2011;12:136–41.
4. Wang L, Zhou L, Jiang P, Lu L, Chen X, Lan H, Guttridge DC, Sun H, Wang H. Loss of miR-29 in myoblasts contributes to dystrophic muscle pathogenesis. Mol Ther. 2012;20:1222–33.
5. Roberts TC, Godfrey C, McClorey G, Vader P, Briggs D, Gardiner C, Aoki Y, Sargent I, Morgan JE, Wood MJA. Extracellular microRNAs are dynamic non-vesicular biomarkers of muscle turnover. Nucleic Acids Res. 2013;41:9500–13.
6. Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A. 2008;105:10513–8.
7. Gilad S, Meiri E, Yogev Y, et al. Serum microRNAs are promising novel biomarkers. PLoS One. 2008;3, e3148.
8. Lee Y, Kim M, Han J, Yeom K-H, Lee S, Baek SH, Kim VN. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004;23:4051–60.
9. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA. 2004;10:1957–66.
10. Kim Y-K, Kim VN. Processing of intronic microRNAs. EMBO J. 2007;26:775–83.
11. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T. Identification of novel genes coding for small expressed RNAs. Science. 2001;294:853–8.
12. Mogilyansky E, Rigoutsos I. The miR-17/92 cluster: a comprehensive update on its genomics, genetics, functions and increasingly important and numerous roles in health and disease. Cell Death Differ. 2013;20:1603–14.
13. Lee Y, Jeon K, Lee J-T, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. EMBO J. 2002;21:4663–70.
14. Olsen PH, Ambros V. The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. Dev Biol. 1999; 216:671–80.
15. Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. Genes Dev. 2004;18:504–11.
16. Ohler U, Yekta S, Lim LP, Bartel DP, Burge CB. Patterns of flanking sequence conservation and a characteristic upstream motif for microRNA gene identification. RNA. 2004;10: 1309–22.
17. Guo H, Ingolia NT, Weissman JS, Bartel DP. Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature. 2010;466:835–40.
18. Gregory RI, Chendrimada TP, Cooch N, Shiekhattar R. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell. 2005;123:631–40.
19. Roberts TC. The microRNA biology of the mammalian nucleus. Mol Ther Nucleic Acids. 2014;3:e188

20. Roberts TC, Wood MJ. Non-canonical microRNA biogenesis and function. In: Arbuthnot P, Weinberg M, editors. Applied RNAi: from fundamental research to therapeutic applications. Norfolk: Caister Academic Press; 2014. p. 19–42.
21. Gregory RI, Yan K-P, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. The Microprocessor complex mediates the genesis of microRNAs. Nature. 2004;432:235–40.
22. Wu H, Xu H, Miraglia LJ, Crooke ST. Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J Biol Chem. 2000;275:36957–65.
23. Lee Y, Ahn C, Han J, et al. The nuclear RNase III Drosha initiates microRNA processing. Nature. 2003;425:415–9.
24. Dalzell JJ, Warnock ND, Stevenson MA, Mousley A, Fleming CC, Maule AG. Short interfering RNA-mediated knockdown of drosha and pasha in undifferentiated Meloidogyne incognita eggs leads to irregular growth and embryonic lethality. Int J Parasitol. 2010;40: 1303–10.
25. Wu Q, Song R, Ortogero N, et al. The RNase III enzyme DROSHA is essential for microRNA production and spermatogenesis. J Biol Chem. 2012;287:25173–90.
26. Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature. 2001;409:363–6.
27. Han J, Lee Y, Yeom K-H, Kim Y-K, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. 2004;18:3016–27.
28. Wilson DI, Burn J, Scambler P, Goodship J. DiGeorge syndrome: part of CATCH 22. J Med Genet. 1993;30:852–6.
29. Shiohama A, Sasaki T, Noda S, Minoshima S, Shimizu N. Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region. Biochem Biophys Res Commun. 2003;304:184–90.
30. Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. Processing of primary microRNAs by the microprocessor complex. Nature. 2004;432:231–5.
31. Landthaler M, Yalcin A, Tuschl T. The human DiGeorge syndrome critical region gene 8 and its D. melanogaster homolog are required for miRNA biogenesis. Curr Biol. 2004;14:2162–7.
32. Zeng Y, Yi R, Cullen BR. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. EMBO J. 2005;24:138–48.
33. Han J, Lee Y, Yeom K-H, Nam J-W, Heo I, Rhee J-K, Sohn SY, Cho Y, Zhang B-T, Kim VN. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell. 2006;125:887–901.
34. Mueller GA, Miller MT, Derose EF, Ghosh M, London RE, Hall TMT. Solution structure of the Drosha double-stranded RNA-binding domain. Silence. 2010;1:2.
35. Sohn SY, Bae WJ, Kim JJ, Yeom K-H, Kim VN, Cho Y. Crystal structure of human DGCR8 core. Nat Struct Mol Biol. 2007;14:847–53.
36. Senturia R, Faller M, Yin S, Loo JA, Cascio D, Sawaya MR, Hwang D, Clubb RT, Guo F. Structure of the dimerization domain of DiGeorge critical region 8. Protein Sci. 2010;19: 1354–65.
37. Faller M, Matsunaga M, Yin S, Loo JA, Guo F. Heme is involved in microRNA processing. Nat Struct Mol Biol. 2007;14:23–9.
38. Brownawell AM, Macara IG. Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. J Cell Biol. 2002;156:53–64.
39. Lund E, Güttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear export of microRNA precursors. Science. 2004;303:95–8.
40. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. 2003;17:3011–6.
41. Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA. 2004;10:185–91.
42. Gwizdek C, Ossareh-Nazari B, Brownawell AM, Doglio A, Bertrand E, Macara IG, Dargemont C. Exportin-5 mediates nuclear export of minihelix-containing RNAs. J Biol Chem. 2003;278: 5505–8.

43. Gwizdek C, Bertrand E, Dargemont C, Lefebvre JC, Blanchard JM, Singer RH, Doglio A. Terminal minihelix, a novel RNA motif that directs polymerase III transcripts to the cell cytoplasm. Terminal minihelix and RNA export. J Biol Chem. 2001;276:25910–8.
44. Lu S, Cullen BR. Adenovirus VA1 noncoding RNA can inhibit small interfering RNA and MicroRNA biogenesis. J Virol. 2004;78:12868–76.
45. Bischoff FR, Krebber H, Kempf T, Hermes I, Ponstingl H. Human RanGTPase-activating protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport. Proc Natl Acad Sci U S A. 1995;92:1749–53.
46. Bischoff FR, Ponstingl H. Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature. 1991;354:80–2.
47. Zeng Y, Cullen BR. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. Nucleic Acids Res. 2004;32:4776–85.
48. Okada C, Yamashita E, Lee SJ, Shibata S, Katahira J, Nakagawa A, Yoneda Y, Tsukihara T. A high-resolution structure of the pre-microRNA nuclear export machinery. Science. 2009;326: 1275–9.
49. Provost P, Dishart D, Doucet J, Frendewey D, Samuelsson B, Rådmark O. Ribonuclease activity and RNA binding of recombinant human Dicer. EMBO J. 2002;21:5864–74.
50. Zhang H, Kolb FA, Brondani V, Billy E, Filipowicz W. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J. 2002;21:5875–85.
51. Humphreys DT, Hynes CJ, Patel HR, Wei GH, Cannon L, Fatkin D, Suter CM, Clancy JL, Preiss T. Complexity of murine cardiomyocyte miRNA biogenesis, sequence variant expression and function. PLoS One. 2012;7, e30933.
52. Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. Cell. 2001;106:23–34.
53. Hutvágner G, McLachlan J, Pasquinelli AE, Bálint E, Tuschl T, Zamore PD. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science. 2001;293:834–8.
54. Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. 2001;15:2654–9.
55. Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KV, Hannon GJ. Dicer is essential for mouse development. Nat Genet. 2003;35:215–7.
56. Hatfield SD, Shcherbata HR, Fischer KA, Nakahara K, Carthew RW, Ruohola-Baker H. Stem cell division is regulated by the microRNA pathway. Nature. 2005;435:974–8.
57. Ma J-B, Ye K, Patel DJ. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature. 2004;429:318–22.
58. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. Cell. 2004;118:57–68.
59. 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;311:195–8.
60. Basyuk E, Suavet F, Doglio A, Bordonné R, Bertrand E. Human let-7 stem-loop precursors harbor features of RNase III cleavage products. Nucleic Acids Res. 2003;31:6593–7.
61. Cenik ES, Fukunaga R, Lu G, Dutcher R, Wang Y, Tanaka Hall TM, Zamore PD. Phosphate and R2D2 restrict the substrate specificity of Dicer-2, an ATP-driven ribonuclease. Mol Cell. 2011;42:172–84.
62. MacRae IJ, Zhou K, Doudna JA. Structural determinants of RNA recognition and cleavage by Dicer. Nat Struct Mol Biol. 2007;14:934–40.
63. Lau P-W, Potter CS, Carragher B, MacRae IJ. Structure of the human Dicer-TRBP complex by electron microscopy. Structure. 2009;17:1326–32.
64. Wang H-W, Noland C, Siridechadilok B, Taylor DW, Ma E, Felderer K, Doudna JA, Nogales E. Structural insights into RNA processing by the human RISC-loading complex. Nat Struct Mol Biol. 2009;16:1148–53.

65. Lau P-W, Guiley KZ, De N, Potter CS, Carragher B, MacRae IJ. The molecular architecture of human Dicer. Nat Struct Mol Biol. 2012;19:436–40.
66. Park J-E, Heo I, Tian Y, Simanshu DK, Chang H, Jee D, Patel DJ, Kim VN. Dicer recognizes the 5′ end of RNA for efficient and accurate processing. Nature. 2011;475:201–5.
67. Gu S, Jin L, Zhang Y, Huang Y, Zhang F, Valdmanis PN, Kay MA. The loop position of shRNAs and pre-miRNAs is critical for the accuracy of dicer processing in vivo. Cell. 2012;151:900–11.
68. Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ. Argonaute2, a link between genetic and biochemical analyses of RNAi. Science. 2001;293:1146–50.
69. Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. Cell. 2003;115:209–16.
70. Schwarz DS, Hutvágner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell. 2003;115:199–208.
71. Tokumaru S, Suzuki M, Yamada H, Nagino M, Takahashi T. let-7 regulates Dicer expression and constitutes a negative feedback loop. Carcinogenesis. 2008;29:2073–7.
72. Lai EC. Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet. 2002;30:363–4.
73. Friedman RC, Farh KK-H, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. 2009;19:92–105.
74. Hutvágner G, Zamore PD. A microRNA in a multiple-turnover RNAi enzyme complex. Science. 2002;297:2056–60.
75. Yekta S, Shih I-H, Bartel DP. MicroRNA-directed cleavage of HOXB8 mRNA. Science. 2004;304:594–6.
76. Zeng Y, Yi R, Cullen BR. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. Proc Natl Acad Sci U S A. 2003;100:9779–84.
77. Bhattacharyya SN, Habermacher R, Martine U, Closs EI, Filipowicz W. Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell. 2006;125: 1111–24.
78. Meister G, Landthaler M, Patkaniowska A, Dorsett Y, Teng G, Tuschl T. Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. Mol Cell. 2004;15:185–97.
79. Pillai RS, Bhattacharyya SN, Filipowicz W. Repression of protein synthesis by miRNAs: how many mechanisms? Trends Cell Biol. 2007;17:118–26.
80. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet. 2008;9:102–14.
81. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004;116: 281–97.
82. Krek A, Grün D, Poy MN, et al. Combinatorial microRNA target predictions. Nat Genet. 2005; 37:495–500.
83. Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature. 2005;436:740–4.
84. Haase AD, Jaskiewicz L, Zhang H, Lainé S, Sack R, Gatignol A, Filipowicz W. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep. 2005;6:961–7.
85. Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song J-J, Hammond SM, Joshua-Tor L, Hannon GJ. Argonaute2 is the catalytic engine of mammalian RNAi. Science. 2004;305: 1437–41.
86. Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV, Patel DJ, van der Oost J. The evolutionary journey of Argonaute proteins. Nat Struct Mol Biol. 2014;21: 743–53.
87. Rivas FV, Tolia NH, Song J-J, Aragon JP, Liu J, Hannon GJ, Joshua-Tor L. Purified Argonaute2 and an siRNA form recombinant human RISC. Nat Struct Mol Biol. 2005;12: 340–9.

88. Schürmann N, Trabuco LG, Bender C, Russell RB, Grimm D. Molecular dissection of human Argonaute proteins by DNA shuffling. Nat Struct Mol Biol. 2013;20:818–26.
89. Alisch RS, Jin P, Epstein M, Caspary T, Warren ST. Argonaute2 is essential for mammalian gastrulation and proper mesoderm formation. PLoS Genet. 2007;3, e227.
90. Morita S, Horii T, Kimura M, Goto Y, Ochiya T, Hatada I. One Argonaute family member, Eif2c2 (Ago2), is essential for development and appears not to be involved in DNA methylation. Genomics. 2007;89:687–96.
91. Cheloufi S, Dos Santos CO, Chong MMW, Hannon GJ. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature. 2010;465:584–9.
92. Yang J-S, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. Proc Natl Acad Sci U S A. 2010;107: 15163–8.
93. Schirle NT, MacRae IJ. The crystal structure of human Argonaute2. Science. 2012;336: 1037–40.
94. Elkayam E, Kuhn C-D, Tocilj A, Haase AD, Greene EM, Hannon GJ, Joshua-Tor L. The structure of human Argonaute-2 in complex with miR-20a. Cell. 2012;150:100–10.
95. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005;120:15–20.
96. Ma J-B, Yuan Y-R, 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;434:666–70.
97. Frank F, Sonenberg N, Nagar B. Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. Nature. 2010;465:818–22.
98. Landthaler M, Gaidatzis D, Rothballer A, Chen PY, Soll SJ, Dinic L, Ojo T, Hafner M, Zavolan M, Tuschl T. Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. RNA. 2008;14:2580–96.
99. Höck J, Weinmann L, Ender C, Rüdel S, Kremmer E, Raabe M, Urlaub H, Meister G. Proteomic and functional analysis of Argonaute-containing mRNA-protein complexes in human cells. EMBO Rep. 2007;8:1052–60.
100. Liu J, Rivas FV, Wohlschlegel J, Yates JR, Parker R, Hannon GJ. A role for the P-body component GW182 in microRNA function. Nat Cell Biol. 2005;7:1261–6.