ELSEVIER

Novel modes of protein–RNA recognition in the RNAi pathway

Andreas Lingel and Michael Sattler

Gene silencing mediated by RNA interference (RNAi) depends on short interfering RNAs (siRNAs) and micro RNAs (miRNAs). These RNAs have unique features, namely a defined size of 19–21 base pairs, and characteristic two-nucleotide single-stranded 3′ overhangs and 5′ monophosphate groups. These molecular features of siRNAs and miRNAs are produced by RNase III enzymes, which are a hallmark of gene silencing induced by double-stranded RNA. Recent structural studies of components of the RNAi pathway, including PAZ, Piwi and RNase III domains, as well as full-length Argonaute and viral p19 proteins, have revealed distinct and novel modes of sequence-independent recognition of the characteristic features of siRNAs and miRNAs in the RNAi pathway.

Addresses
EMBL Heidelberg, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Corresponding author: Sattler, Michael (sattler@embl.de)

Current Opinion in Structural Biology 2005, **15**:107–115

This review comes from a themed issue on
Protein–nucleic acid interactions
Edited by Aneel K Aggarwal and Jennifer A Doudna

Available online 26th January 2005

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2005.01.010

Abbreviations

ds	double-stranded
dsRBD	dsRNA-binding domain
K_d	dissociation constant
miRNA	micro RNA
OB-fold	oligonucleotide/oligosaccharide-binding fold
PAZ	Piwi, Argonaute, Zwille
PTGS	post-transcriptional gene silencing
RISC	RNA-induced silencing complex
RNAi	RNA interference
siRNA	short interfering RNA

Introduction

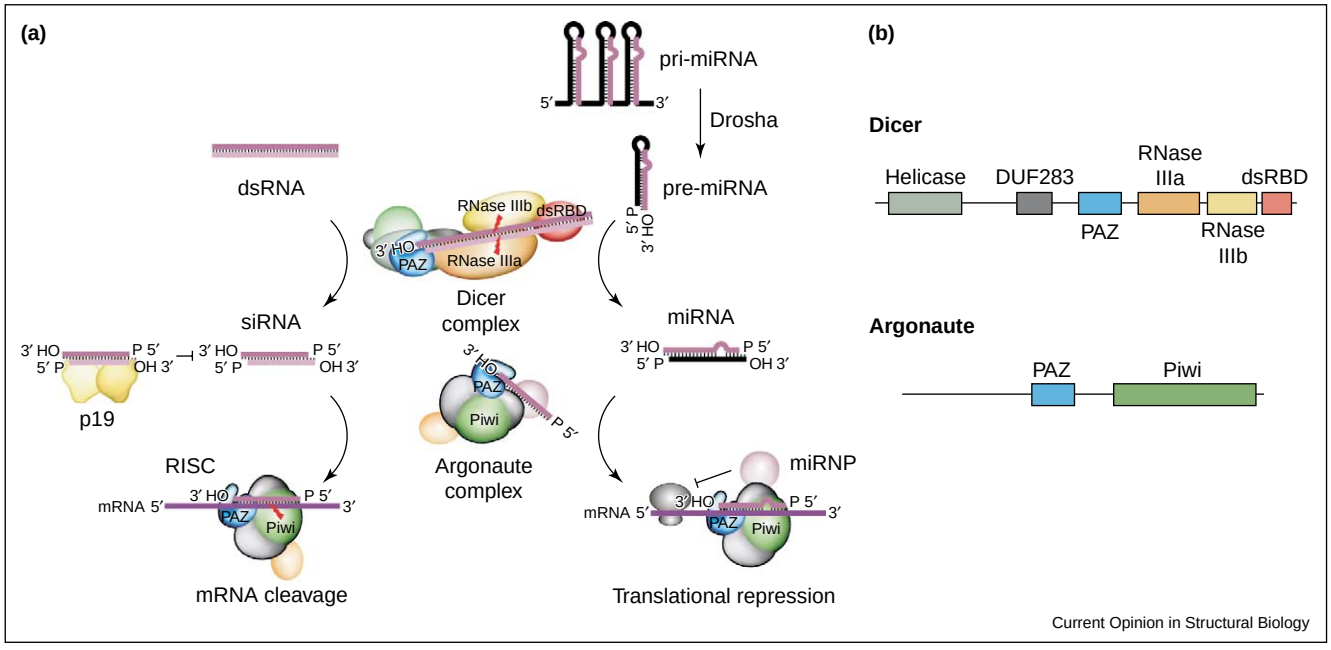
RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) is an evolutionarily conserved cellular response to the presence of double-stranded (ds) RNA [1•,2•,3–5]. RNAi is thought to originate from an ancient endogenous defense mechanism against viral and other heterologous dsRNAs [6,7]. Gene silencing induced by dsRNA is widely used as a tool to study gene function in higher eukaryotes [8]. In the cell, the dsRNA is first cleaved by the RNase III enzyme Dicer into short interfering RNA duplexes (siRNAs) of 21–23 nucleotides (nt) with 3′ overhangs of 2 nt (Figure 1a). In a subsequent step, one strand of the siRNA is incorporated into a multimeric RNA-induced silencing complex (RISC), where it guides the selection of a complementary mRNA. The evolutionarily conserved family of Argonaute proteins are essential components of the RISC. An endonuclease in the RISC effector complex (termed Slicer) cleaves the mRNA in the region that has sequence complementarity to the siRNA guide. Subsequently, the mRNA initially cleaved by Slicer is degraded by exonucleases and thus silenced.

A related pathway of PTGS involves endogenously expressed non-coding RNAs of 21–25 nt called micro RNAs (miRNAs). These RNAs are derived from long primary transcripts termed pri-miRNAs, which are cleaved by the RNase III enzyme Drosha into precursor miRNA (pre-miRNA) hairpins of about 70 nt with a 3′ overhang of 2 nt (Figure 1a) [9]. Double-stranded miRNAs are excised from pre-miRNAs by Dicer. As products of RNase III cleavage, the resulting double-stranded miRNAs have 2-nt 3′ overhangs and 5′ monophosphate groups, similar to siRNAs. In contrast to siRNAs, which guide endonucleolytic cleavage, the single-stranded miRNA guide shows imperfect sequence complementarity to a target mRNA (usually in 3′ untranslated regions) and leads to the inhibition of translation. However, siRNAs with imperfect sequence complementarity have also been shown to effect translational inhibition and, conversely, miRNAs with full sequence complementarity can induce mRNA degradation (reviewed in [1•,2•]). Thus, it is thought that the siRNA and miRNA pathways lead to the assembly of related effector complexes and that the mode of gene silencing (mRNA degradation or inhibition of translation) depends on the degree of sequence complementarity between the siRNA or miRNA guide and the mRNA.

A third but less well characterized dsRNA-induced pathway mediates chromatin-based gene silencing at the level of transcription in plants, yeast and metazoa [10,11]. Finally, viruses have developed mechanisms to counteract RNAi by providing proteins that sequester siRNAs targeted against viral RNAs [6,12,13].

The Dicer and Argonaute protein families constitute crucial components of the RNAi pathway. Members of the Dicer family are multidomain proteins of about 200 kDa that comprise two RNase III domains, a PAZ domain [named after the proteins Piwi, Argonaute and Zwille (also known as Pinhead)], a dsRNA-binding

Figure 1



Gene silencing pathways in RNAi. **(a)** Pathways of RNAi originating from dsRNA or endogenous hairpin pre-miRNAs. Current models of the structural and topological features of the protein–RNA complexes involved in RNAi are shown. The color coding of the different domains is the same in all figures. Red arrows indicate endonucleolytic cleavage. RISC and miRNP (micro ribonucleoprotein particle) are the effector complexes of the siRNA and miRNA pathways, respectively. **(b)** Domain structure of Dicer and Argonaute proteins. Dicer comprises a DEXH helicase, a dsRBD, a PAZ domain and a domain of unknown function (DUF283), in addition to two RNase III domains. Argonaute proteins share a PAZ domain and a Piwi domain.

domain (dsRBD), a domain of unknown function (DUF283) and a helicase domain (Figure 1b) [14,15••]. The Argonaute family comprises highly basic proteins of about 100 kDa that are characterized by a central PAZ domain and a C-terminal Piwi domain [16] (Figure 1b). The genomes of higher eukaryotic organisms encode several paralogs that are implicated in the maturation and function of effector complexes in the pathways mediated by siRNAs and miRNAs.

The cellular machinery triggered by siRNAs or miRNAs must be highly selective and must distinguish siRNAs and miRNAs from other RNAs to avoid erroneous gene silencing. In this respect, the non-coding RNAs that induce RNAi have unique molecular features: a defined size — namely, a dsRNA helix of 19–21 bp; characteristic 2-nt single-stranded 3′ overhangs with a 3′ hydroxyl group; and 5′ monophosphates at both ends. These features are produced by RNase III enzymes, which play a unique role in RNA-induced gene silencing.

In this review, we survey the structural biology of components of the RNAi pathway. These components have evolved distinct and novel modes of RNA recognition by exploiting the characteristic molecular features of siRNAs and miRNAs.

Nucleic acid 3′-end recognition by the PAZ domain

The PAZ domain is found exclusively in the Argonaute and Dicer protein families [16,17], and was initially hypothesized to mediate protein–protein interactions because Argonaute and Dicer proteins co-immunoprecipitate. Recent three-dimensional structures of the *Drosophila* Ago1 and Ago2 PAZ domains have revealed a central five-stranded β barrel flanked by two α helices, and a conserved inserted module of about 35 residues comprising a β hairpin and α helix (Figure 2) [18••–20••]. Unexpectedly, these PAZ domains have been found to bind nucleic acids *in vitro* [18••–20••]. ssRNAs and dsRNAs with single-stranded 3′ overhangs bind strongly, whereas blunt-ended dsRNA molecules show a reduced affinity for PAZ domains [19••,20••].

Using NMR chemical shift perturbation and mutational analysis, the nucleic-acid-binding surface of the PAZ domain has been mapped to a hydrophobic cleft between the central β barrel and the conserved β hairpin/α helix module [18••–20••]. The conservation of residues in this region suggests that it performs a similar function in all PAZ domains. Indeed, conserved nucleic acid binding has been observed for PAZ domains from the *Drosophila* Ago1 (*Dm*Ago1), Ago2 (*Dm*Ago2), Piwi