Argonaute and GW182 proteins: an effective alliance in gene silencing

Janina Pfaff† and Gunter Meister*†

†Biochemistry Center Regensburg (BZR), Laboratory for RNA Biology, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, and Group Intracellular Transport and RNA Biology, Helmholtz-Zentrum München – German Research Center for Environmental Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany

Abstract

Argonaute proteins interact with small RNAs and facilitate small RNA-guided gene-silencing processes. Small RNAs guide Argonaute proteins to distinct target sites on mRNAs where Argonaute proteins interact with members of the GW182 protein family (also known as GW proteins). In subsequent steps, GW182 proteins mediate the downstream steps of gene silencing. The present mini-review summarizes and discusses our current knowledge of the molecular basis of Argonaute–GW182 protein interactions.

Introduction

Small RNAs such as miRNAs (microRNAs) or siRNAs (short interfering RNAs) have emerged as key regulators of many cellular pathways. Both small RNA species derive from double-stranded RNA precursors and are processed to short ~21-nt-long RNAs [1]. miRNAs are endogenously expressed as stem–loop-structured precursors and are processed by the consecutive action of the two RNase III enzymes Drosha and Dicer (Figure 1). siRNAs and miRNAs are loaded into the RISC (RNA-induced silencing complex) (often distinguished as siRISC or miRISC depending on the small RNA that is carried). Within RISC, the small RNA directly interacts with a member of the Argonaute protein family [2,3]. miRNAs and siRNAs target RISC to complementary sequences on mRNAs and, in the case of full complementarity, the target RNA is cleaved by a cleavage-competent Ago protein. Animal miRNAs typically bind to partially complementary target sites located in 3′-UTRs (untranslated regions) of mRNAs. For efficient silencing, perfect base-pairing of nucleotides 2–8 (known as the seed sequence) of the miRNA with the target is required. However, miRNA-target pairs with mismatches in the seed sequence have also been reported [4–6]. As a consequence of miRISC binding, the target RNA is deadenylated and subsequently degraded by cellular exonucleases [7]. For mediating miRNA-guided gene silencing, Argonaute proteins interact with members of the GW182 protein family.

Argonaute proteins

The protein family was first described in *Arabidopsis thaliana* and is named after a mutant plant, which phenotypically resembled the small squid *Argonauta* [8]. Argonaute proteins are highly conserved and are found in archaea, bacteria and eukaryotes. Copy numbers of Argonaute genes range from eight in humans to five in *Drosophila melanogaster* and 26 in *Caenorhabditis elegans* with different expression profiles, small RNA preferences and substrate specificities [2,3,9].

Argonaute proteins cluster into three clades: the germ-line-specific Piwi clade that binds piRNAs (Piwi-interacting RNAs) (a germ-line-specific class of small RNAs that repress mobile genetic elements [10]), the ubiquitously expressed Ago clade interacting with siRNAs and miRNAs and the WAGO subfamily only found in nematodes [11]. Four of the human Argonaute proteins belong to the Piwi subfamily (HIWI1, HIWI2, HIWI3 and HIWI4), whereas the other four belong to the Ago subfamily (Ago1, Ago2, Ago3 and Ago4). The functional difference between the four Ago proteins is unclear [2,3]. All four Ago proteins are capable of repressing miRNA target genes. In addition, Ago2 possesses endonucleolytic activity and cleaves target RNAs that are fully complementary to the bound small RNA [12,13].

Functional domains of Argonaute proteins

Ago proteins typically have a molecular mass of approximately 100 kDa. Initial structural studies on prokaryotic Ago proteins revealed a bilobal architecture comprising four globular domains: the N-terminal, PAZ (PWW/Piwi/Ago2/Zwille), MID (middle) and PIWI domains. The guide RNA strands of siRNAs or miRNAs span between both lobes with its 3′ end anchored in the PAZ domain and the 5′ end bound to the MID domain. Besides recognition of the small RNA ends, Ago proteins offer extensive interactions with the phosphate backbone, thereby positioning the seed region (nucleotides...
miRNA genes are transcribed by RNA polymerase II or III as stem-loop-structured primary miRNAs. These transcripts are processed to miRNA precursors by the Microprocessor complex containing the RNase III enzyme Drosha and its cofactor DGCR8 (DiGeorge syndrome critical region 8). After export to the cytoplasm by Exportin5, the RNase III Dicer together with its cofactor TRBP (TAR DNA-binding protein) processes the precursor to a double-stranded intermediate. One strand of the duplex is selected and incorporated into Ago protein containing RISC complexes. Depending on the degree of complementarity, the target is either cleaved directly by RISC or translationally repressed followed by exonuclease-mediated decay. N, nucleus; C, cytoplasm; AAA, poly(A) tail.

2–8) of the miRNA suitable for target binding [14,15]. Recently, full-length human Ago2 was crystallized, providing further structural insights into the mechanism of eukaryotic Ago function [16,17]. The overall architecture resembles prokaryotic structures with a bilobal shape and a central cleft for miRNA/siRNA binding. The individual domains align well with the available structures; however, the relative positions of the lobes differ from those of their prokaryotic counterparts.

The PAZ domain resembles the oligosaccharide/oligonucleotide-binding domain or Sm-fold domain. As shown previously, conserved residues of a hydrophobic pocket specifically recognize the 3′ dinucleotide of the miRNA/siRNA, which result from Dicer cleavage. The last base stacks against a conserved phenylalanine residue (Phe296) and remaining interactions are mainly mediated by the phosphate backbone and are thus sequence-independent [16–20].

The MID domain reveals a fold similar to that of the sugar-binding domain of the lac-repressor. Similarly to prokaryotic Ago proteins [21,22], the MID domain of human Ago2 forms a pocket for tethering the 5′ phosphate of the miRNA/siRNA. Several residues (Tyr529, Lys533, Gln545, Cys546, Lys566, Lys570, Arg812 and Ala859) were identified that are essential for binding [16,17]. Deep sequencing of small RNAs revealed that human miRNAs show a clear uracil bias at their 5′ ends. More rarely, adenine is also found as the 5′ terminal nucleotide [5]. Structural work has demonstrated that adenine and uracil nucleotides are indeed preferred. This is due to a loop in the MID domain that prevents cytosine and guanine binding [23]. Thus the MID domain not only binds the 5′ end, but also provides specificity for the 5′ nucleotide.

The PIWI domain comprises an RNase H fold. RNase H usually catalyses the hydrolysis of RNA in RNA–DNA duplexes. Accordingly, prokaryotic Ago proteins were found to be DNA-guided RNases. In contrast, some eukaryotic Ago proteins possess cleavage activity towards RNA–RNA duplexes. The catalytic centre of RNase H contains an Asp-Asp-Glu triad with two co-ordinated divalent metal ions required for cleavage activity. Similarly, Ago-mediated cleavage is dependent on the presence of divalent metal ions. However, the active site comprises an Asp-Asp-His triad [24,25]. Recent structural work on a yeast Ago protein extended the catalytic triad to a tetrad composed of Asp-Glu-Asp-His (DEDH) [26]. This catalytic tetrad is a requirement, but it is not sufficient for cleavage activity. In humans, Ago2 and Ago3 offer the Asp-Asp-His motif, but only Ago2 has cleavage activity towards fully complementary targets and is therefore referred to as ‘slicer’ [12,13]. Biochemical studies revealed that the N domain contains an unstructured loop that is critical for cleavage. Ago3 containing the loop of Ago2 retains cleavage activity [27].

In addition to its cleavage activity, the PIWI domain is the binding platform for GW182 proteins. Two highly specific pockets on the surface of the PIWI domain accommodate two tryptophan residues of GW182 proteins, allowing for efficient interactions [16].
Figure 2 | Mechanism of miRNA-guided gene silencing
miRNAs guide Ago proteins to target sites on mRNAs. Ago proteins recruit GW182 proteins to the mRNA, which facilitate all downstream effects of gene silencing. These proteins establish interactions with the poly(A)-binding protein PABPC1 as well as deadenylases such as PAN2/3 or the CCR4/NOT complex. The mRNA is subsequently deadenylated, decapped by DCAP1/2 and degraded by the 5′ to 3′ exonuclease XRN1.

The GW182 protein family
In animals, miRNAs guide Ago proteins to target sites on mRNAs. Ago proteins, in turn, interact with a member of the GW182 protein family (also referred to as GW182 protein family), which mediate all further downstream processes. GW182 proteins were initially described when researchers observed a 182 kDa protein containing multiple GW (Gly-Trp)-repeats in autoimmune serum from a patient suffering from motor and sensory neuropathy [28]. The protein was shown to accumulate in discrete cytoplasmic speckles termed GW bodies associated with mRNA and a role in post-transcriptional regulation of gene expression was suggested. A subsequent study gave evidence that GW-bodies coincide with cytoplasmic structures containing markers of the mRNA-degradation pathway [29]. GW-bodies are also known as cytoplasmic processing bodies or P-bodies. The link to miRNA-mediated gene silencing was established when miRNAs and Ago proteins were also found as components of P-bodies and the C. elegans GW182 homologue AIN-1 was shown to interact with the miRISC in P-bodies [30,31]. A short time later, GW182 proteins were demonstrated to be essential for miRNA-mediated gene silencing, and a physical interaction between Ago and GW182 proteins was proven [32–36].

In the last few years, a huge amount of work in various organisms has led to a model of GW182 protein function in miRNA-guided gene silencing. These proteins bind to Ago proteins with their N-terminal domain. The C-terminal part establishes interactions with the poly(A)-binding protein on the poly(A) tail of the mRNA. In addition, direct interactions of the C-terminal part with cellular deadenylases such as the CCR4/NOT complex leads to deadenylation of the mRNA. As a consequence, the mRNA is decapped and subsequently degraded [7,37] (Figure 2).

Domain organization of GW182 proteins
Three GW182 paralogues were identified in mammals termed TNRC6 (trinucleotide repeat-containing protein 6) A (corresponding to GW182), B and C. Insects comprise a single orthologue (GW182 or Gawky). The C. elegans GW182 proteins are known as AIN1 and AIN2. Vertebrates and insect proteins share a similar domain organization. All proteins feature a N-terminal Ago-binding region, a UBA (ubiquitin-associated)-like domain, a Q (glutamine)-rich region, a PAM2 {PABP [poly(A)-binding protein]-interacting motif 2} region and a C-terminal RRM (RNA-recognition motif) (Figure 3).

GW182 proteins have a molecular mass of approximately 180 kDa and are mainly disordered except for two globular domains: the UBA-like domain and an RRM [37].

UBA domains are common sequence motifs of ∼45 amino acids usually occurring in proteins involved in the ubiquitin–proteasome pathway, DNA repair or cell growth. They fold into a three-helix bundle with large hydrophobic patches, which form the interface for binding mono- or polyubiquitinated substrates [38].

RRMs usually bind single-stranded RNA and adopt a fold containing four antiparallel β-strands and two α-helices [39]. The structure of D. melanogaster GW182 RRM revealed a canonical fold with an additional C-terminal helix. The lack of RNA-binding features suggests that the RRM is rather involved in protein–protein interactions [40]. The exact function of the RRM domain has not been characterized.

PAM2 is responsible for the interaction with PABPC1 and is embedded in a disordered part of the protein’s C-terminus between the UBA-like domain and the RRM domain. N-terminally to PAM2, GW182 proteins possess a Q-rich region, which is necessary for P-body localization, but dispensable for the silencing activity [41,42].
Tryptophan motifs in the C-terminal, middle and N-terminal region of GW182 have been implicated in recruiting effector proteins as the CCR4/NOT complex responsible for mRNA repression [43–45]. One study identified two CCR4/NOT-interaction motifs (CIM1 and CIM2) in the M1 and C-terminal region that are conserved in mammals [46]. In parallel, it was published that CCR4/NOT binds to several tryptophan repeats in the M2 and C-terminal region [43]. The authors suggested that the tryptophan motifs form a bipartite domain with autonomous silencing activity. In contrast with the C-terminal tryptophan motifs, several studies demonstrated that GW-repeats in the N-terminus of GW182 are important for the interaction with Ago [47–50].

Molecular mechanism of the binding of GW-repeats to Ago proteins

GW182 proteins are known to serve as universal Ago-binding platforms. In post-transcriptional mRNA regulation, SLiMs (short linear interaction motifs), like the GW-repeats, are common sequence features in a variety of pathways [51–53]. SLiMs have a length of ~6–11 amino acids and are typically found in disordered protein regions. To characterize the binding mode of GW-repeats, it might help to examine the results obtained for related SLiMs besides investigation of GW182–Ago interaction data.

Several studies have provided evidence that tryptophan and its neighbour glycine are crucial for efficient interaction of GW-repeats and Ago [47,48,54,55]. Moreover, investigation of different GW-repeat-containing proteins showed little sequence conservation except for the presence of GW-repeats. Owing to the low conservation score and high variability of GW-repeats in number and length, it is very difficult to predict GW-repeat-containing proteins. Using bioinformatic approaches, tools were developed to detect GW motifs in proteins [56]. It was found that the amino acids embedding the tryptophan are mainly of small, hydrophilic and charged nature and suggest that this might facilitate surface accessibility of the tryptophan.

GW-repeats as conserved Ago-binding hooks

The importance of GW-repeats in miRNA-mediated gene silencing was soon recognized and GW182 proteins quickly emerged as key components in the miRNA pathway recruiting effector proteins to the mRNA by binding Ago through GW-repeats. However, GW182 proteins are not the only proteins using GW-repeats to associate with Ago proteins.

Proteins involved in TGS (transcriptional gene silencing) by the RITS (RNA-induced transcriptional silencing) complex in fission yeast were shown to interact with Ago in a similar way [57]. RITS mediates transcriptional silencing and contains an Ago protein, the chromodomain protein Chp1 and Tas3. An interaction module within Tas3 was reported with high sequence conservation to the GW-repeat-containing Ago hook of GW182 proteins. This motif was essential for TGS and mutation of a single glycine or tryptophan residue abolished binding [47,58].

In plants, a RNAi (RNA interference)-related mechanism known as RdDM (RNA-directed DNA methylation) silences target DNAs by methylation and involves specific nuclear RNA polymerases that synthesize or bind scaffold RNAs [59]. RNA Pol V (polymerase V) largest subunit NRPE1 comprises a specific GW-repeat that interacts with plant Ago4 and is required for Pol V function [55]. In the same pathway, an SPT5 homologue (KTF1) was shown to link Pol V transcription with Ago4 and the interaction is mediated by GW-repeat interactions [60]. Furthermore, the Tetrahymena Ago protein was shown to interact with two GW-repeat proteins, thereby promoting genome rearrangement [61].

Recently, the interaction between the GW-repeat-containing prion protein PrPc was established which promoted accumulation of miRNA effector complexes [62]. In addition, GW182 proteins were also implicated in pathogen defence against the host RNAi machinery in plants. Two independent studies provide evidence that viruses use proteins containing Ago-hooks to hijack the host Ago-silencing system and are essential for VSR (viral suppressor of RNA silencing) function [63,64].

Overall, those reports indicate that GW-repeats function in numerous proteins of divergent species as evolutionarily conserved Ago-binding platforms. Although considerable progress on how Ago and GW182 proteins interact has been made, many questions are still unsolved. For example, how is specificity achieved and how are the correct tryptophan residues selected among the many others that are not engaged in Ago binding? Further biochemical and structural studies will help to elucidate these interesting functional aspects on small RNA-guided gene silencing.

Acknowledgements

We are grateful to Dierk Niessing for support and comments.

Funding

Our research is supported by the Deutsche Forschungsgemeinschaft [grant numbers SFB 960 and FOR855 (to G.M.)], the European Research Council (ERC grant ‘sRNAs’ to G.M.), the Bavarian Genome Research Network (BayGene to G.M.), the Bavarian Systems-Biology Network (BioSysNet to G.M.) and the European Union (grant ‘ONCOMIRs’ to G.M.).

References


Received 10 April 2013
doi:10.1042/BST20130047