Argonaute MID domain takes centre stage

Christopher R. Faehnle & Leemor Joshua-Tor

Argonaute (AGO) proteins are the central component of small RNA-mediated gene silencing in eukaryotes. Functional AGO complexes are loaded with single-stranded small RNAs, which guide AGO to a messenger RNA (mRNA) target through base pairing. Although the structure of a full-length eukaryotic AGO has yet to be described, insights into the mechanism of guide RNA binding and target recognition have been revealed by the structures of distantly related AGO homologues from archaea and eubacteria (Song *et al.*, 2004; Wang *et al.*, 2008, 2009). These studies show that AGO proteins are composed of amino-terminal, PAZ (PIWI/Argonaute/Zwille), MID (middle) and PIWI (P-element-induced Wimpy Testes) domains. The phosphorylated 5'-end of the guide strand RNA is localized in the MID-PIWI domain interface with the 3'-end anchored to the PAZ domain. On binding to mRNA the catalytic RNase H-like active site located in the PIWI domain is in position to cleave the targeted mRNA.

Two recent papers, one in *EMBO reports* (Boland *et al.*, 2010) and one in *Nature* (Frank *et al.*, 2010), give us the first eukaryotic structures of AGO MID domains. The human AGO MID domain structure provides a structural basis for the 5'-nucleotide recognition of the guide strand observed in eukaryotic AGOs, and the structure of the MID domain of QDE-2 from *Neurospora crassa* published in this journal offers a possible explanation for the allosteric regulation of RNA binding discovered earlier this year by Rachel Green's group (Djuranovic *et al.*, 2010).

The two structures have a similar topology resembling a Rossmann-like fold with four β -strands forming a central β -sheet flanked by α -helices. Superposition of the two structures, which are 24% identical in sequence, shows that they are also similar in three dimensions, with a root mean square deviation (r.m.s.d.) of 2.1 Å. The archaeal and eubacterial AGO MID domains solved previously share less than 20% sequence identity and have a

greater than 2.5 Å r.m.s.d. for backbone atoms from both QDE-2 and human AGO2, despite having a similar overall fold.

Crystals of the QDE-2 MID domain contain two sulphate ions. The first sulphate (sulphate I) is coordinated by the highly conserved amino acids Tyr 595, Lys 599 and Lys 638, and is in the same position as the 5'-phosphate of UMP observed in the human AGO2 MID domain structure (Fig 1A). These interactions are similar to those observed for the 5'-phosphate of the guide strand of the previously solved archaeal and eubacterial structures (Ma *et al.*, 2005; Parker *et al.*, 2005; Wang *et al.*, 2008). Thus, sulphate I bound in the QDE-2 MID domain structure likely represents the 5'-nucleotide-binding site. Most intriguing is the position of the second sulphate (sulphate II), located in an adjacent but partly overlapping binding site with sulphate I. Sulphate II is 6.3 Å from sulphate I, shares coordination with Lys 599 and Lys 638, and is further coordinated by Thr 610. Sulphate II can be excluded from representing the phosphate backbone of a microRNA (miRNA) or target because it is bound in the side of the MID domain opposite from where the guide RNA extends from the 5'-nucleotide-binding site. Although the presence of sulphate II does not guarantee a biologically relevant ligand-binding site, it is tempting to speculate, in the light of a recent study by Djuranovic *et al.*, that sulphate II occupies an allosteric ligand-binding site.

Djuranovic *et al.* describe a second ligand-binding site in the *Drosophila melanogaster* AGO1 MID domain that is separate and distinct from the 5'-nucleotide-binding site. They demonstrate that free nucleotides, including the cap analogue m⁷GpppG, bind to an allosteric site, which in turn enhances the binding of miRNA. Cap binding was reported previously for human AGO2 (Kiriakidou *et al.*, 2007), leading to the proposal that two phenylalanine residues in the MID domain make stacking interactions with the m⁷GpppG cap structure, analogous

to eukaryotic initiation factor 4E. However, in the human AGO2 MID domain structure it is clear that these phenylalanine residues are on opposite sides of the MID domain and are located in the hydrophobic core. In the QDE-2 MID domain only one of these phenylalanines is conserved, but a similar conclusion is drawn on the basis of the positions of the two residues being more than 25 Å apart. This strongly argues that an alternative mechanism exists for cap binding by eukaryotic AGO proteins.

The data presented by Djuranovic *et al.* might be explained by the structure of the QDE-2 MID domain, with sulphate I representing the 5'-binding site of a miRNA and sulphate II representing the allosteric site. This argument is strengthened by the fact that the two binding sites are partly shared, namely by interactions with the side chains of Lys 599 and Lys 638, so it would not be surprising that binding of a ligand to one site would have a positive effect on ligand binding at the other site. To identify the location of the potential allosteric site, Djuranovic *et al.* mutated Asp 627—a conserved residue located in a loop 15 Å away from the 5'-nucleotide binding site—to a lysine in *D. melanogaster* AGO1. The D627K mutant failed to bind cap analogues, indicating the importance of Asp 627 for binding ligands in the allosteric site. When mapped onto the new structures of the QDE-2 and human AGO2 MID domains, this loop and Asp 627 (Asp 603 in QDE-2 and Asp 537 in human AGO2) are in the vicinity of sulphate II, thus Asp 627 is probably a part of the allosteric binding site (Fig 1A,B). The most significant finding in the Djuranovic *et al.* study is that the D627K mutant located in the allosteric site fails to bind to miRNA in the 5'-nucleotide-binding site and no longer associates with GW182, an essential factor in miRNA-induced gene silencing.

Almost all miRNA sequences and RNA sequencing data obtained from immunopurified AGO proteins show a marked bias

Fig 1 | Structural comparison of *Neurospora crassa* QDE-2 MID domain and human Argonaute 2 MID domain. (A) The *N. crassa* QDE-2 MID domain structure (green ribbon). UMP is modelled from a superposition of the human AGO2 MID domain structure in a complex with UMP (Protein Data Bank code 3LUJ). Sulphate I and II (S I and S II) are shown as observed in the QDE-2 structure. Conserved QDE-2 amino acids involved in binding sulphate I and II are shown as sticks. (B) Human AGO2 MID domain structure (blue ribbon) in complex with UMP (sticks). The nucleotide specificity loop is coloured in yellow. Sulphate I and II are modelled from a superposition of the *N. crassa* QDE-2 MID domain. AGO, Argonaute; MID, middle.

for uridine and adenosine nucleotides at the 5'-end of miRNA guide strands. The structure of human AGO2 MID domain alone and in a complex with UMP, AMP, GMP and CMP provides the first explanation for the observed 5'-nucleotide bias in eukaryotic AGO proteins (Frank *et al*, 2010). There is little movement induced on nucleotide binding in the overall fold of the MID domain. Electron density is observed for the entire nucleotide in the case of UMP and AMP. The 5'-phosphate in the UMP and AMP complexes is hydrogen bonded to the highly conserved side chains of Tyr529, Lys533, Gln545 and Lys570. The base of each nucleotide stacks with Tyr529, completing a nonspecific recognition pocket for the 5'-nucleotide (Fig 1B). A similar pocket is formed in the *N. crassa* MID domain structure to recognize the 5'-nucleotide. Interestingly, clear electron density for only the phosphate and ribose is observed for GMP and CMP, with density for the GMP and CMP bases missing. These results are consistent with the preference for UMP and AMP binding to human AGO2, but where does this specificity originate?

A closer look at the UMP and AMP complex structures show that base-specific contacts are formed with backbone atoms of a loop spanning residues Pro523 through Pro527, appropriately termed the nucleotide specificity loop. In the case of GMP and

CMP, the hydrogen-bonding partners are in the opposite orientation, resulting in charge repulsion from backbone atoms in the nucleotide specificity loop, thus explaining the observed bias in the 5'-position of the guide strand. In the nucleotide-free structure, the conformation of the nucleotide specificity loop is merely unchanged, suggesting that the loop is rather rigid. When the length of the nucleotide specificity loop is increased by the insertion of a single glycine residue, the specificity for uridine and adenosine nucleotides is lost, further endorsing the idea that the particular conformation and rigid nature of the loop is essential for specific base recognition. Interestingly, the QDE-2 MID domain deviates from the human AGO2 MID domain in the nucleotide specificity loop. An insertion of an aspartate residue in QDE-2 makes the nucleotide specificity loop one amino acid longer, suggesting that QDE-2 might have lost its specificity for nucleotides at the 5'-end of miRNAs, although this is yet to be tested.

A complete understanding of miRNA loading and the allosteric mechanism will have to await structures of full-length eukaryotic AGO proteins, as the PIWI domain contributes numerous contacts with the MID domain, encompassing both the 5'-nucleotide-binding site and the putative allosteric site. However, the structures of

N. crassa QDE-2 MID and human AGO2 MID domains together are important pieces of the puzzle in our understanding of the mechanism of RNA interference. Specific recognition of the 5'-nucleotide of the guide strand might be a quality control mechanism for some eukaryotic AGOs, ensuring that after primary processing the correct miRNA guide sequence is loaded. Once loaded with a proper guide strand, AGO might trigger the adjacent allosteric site to bind to m⁷GpppG-capped mRNA, GW182 or other unknown ligands. Together, these events ultimately lead to effective gene silencing.

REFERENCES

- Boland A *et al* (2010) *EMBO Rep* **11**: 522–527
- Djuranovic S *et al* (2010) *Nat Struct Mol Biol* **17**: 144–150
- Frank F *et al* (2010) *Nature* **465**: 818–822
- Kiriakidou M *et al* (2007) *Cell* **129**: 1141–1151
- Ma JB *et al* (2005) *Nature* **434**: 666–670
- Parker JS *et al* (2005) *Nature* **434**: 663–666
- Song JJ *et al* (2004) *Science* **305**: 1434–1437
- Wang Y *et al* (2008) *Nature* **456**: 209–213
- Wang Y *et al* (2009) *Nature* **461**: 754–761

Christopher R. Faehnle & Leemor Joshua-Tor
are at the Cold Spring Harbor Laboratory,
Cold Spring Harbor, NY, USA.

**L.J. is also a Howard Hughes Medical
Institute Investigator.**

E-mail: leemor@cshl.edu

Published online 16 July 2010

EMBO reports (2010) **11**, 564–565.
doi:10.1038/embor.2010.110