Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein

Andreas Boland, Felix Tritschler, Susanne Heimstädt, Elisa Izaurralde+ & Oliver Weichenrieder++

INTRODUCTION

Proteins of the Argonaute (AGO) family have essential roles in RNA-mediated gene silencing mechanisms throughout the eukaryotic lineage. They associate with small non-coding RNAs to repress the translation of target messenger RNAs (mRNAs); the middle (MID) domain that provides a binding pocket for the 5′-phosphate of guide RNAs; and the P-element-induced whimpy testes (PIWI) domain that adopts an RNaseH fold and has endonucleolytic activity in some but not all AGOs (Parker et al, 2004, 2005, 2009; Song et al, 2004; Ma et al, 2005; Yuan et al, 2005; Wang et al, 2008a,b, 2009). Thus far, structural information on full-length AGOs is only available for the homologous proteins from archaea and eubacteria (Song et al, 2004; Ma et al, 2005; Parker et al, 2005; Yuan et al, 2005; Wang et al, 2008a,b, 2009), which can use DNA instead of RNA as guide molecules (Ma et al, 2005; Parker et al, 2005; Yuan et al, 2005; Wang et al, 2008a,b, 2009).

The function of AGO, however, has only been studied in eukaryotic systems; and, in the microRNA (miRNA) pathway, the precise mechanism by which AGO proteins mediate translational repression of target messenger RNAs (mRNAs) remains controversial (Eulalio et al, 2008a). Understanding such silencing mechanisms requires detailed structural information about the domains of eukaryotic AGOs; however, so far, this level of information is only available for the PAZ domain (Lingel et al, 2003; Song et al, 2003; Yan et al, 2003).

In this study, we focused on the structure of a eukaryotic AGO MID domain, because it is thought to perform essential and distinct functions in the miRNA pathway and because a considerable amount of discussion has focused on the MID domain fold (Kiriakidou et al, 2007; Kinch & Grishin, 2009; Djuranovic et al, 2010). Furthermore, it has been proposed that certain eukaryotic AGO MID domains sequester the mRNA 5′-cap structure (m^7GpppN) of target mRNAs, thereby repressing translation. Initially, this idea was based on the observation that the human AGO2 binds to m^7GTP Sepharose beads (Kiriakidou et al, 2007). Recent data by Djuranovic et al (2010) supported this concept, indicating that, in addition to the 5′-phosphate-binding pocket for the guide RNA, Drosophila melanogaster AGO1 has a second miRNA-dependent site that can bind to nucleotides such as the 5′-cap.

Argonaute (AGO) proteins are core components of RNA-induced silencing complexes and have essential roles in RNA-mediated gene silencing. They are characterized by a bilobal architecture, consisting of one lobe containing the amino-terminal and PAZ domains and another containing the MID and PIWI domains. Except for the PAZ domain, structural information on eukaryotic AGO domains is not yet available. In this study, we report the crystal structure of the MID domain of the eukaryotic AGO protein QDE-2 from Neurospora crassa. This domain adopts a Rossmann-like fold and recognizes the 5′-terminal nucleotide of a guide RNA in a manner similar to its prokaryotic counterparts. The 5′-nucleotide-binding site shares common residues with a second, adjacent ligand-binding site, suggesting a mechanism for the cooperative binding of ligands to the MID domain of eukaryotic AGOs.

Keywords: GW182; miRNAs; RNAi; silencing; siRNAs

EMBO reports (2010) 11, 522–527. doi:10.1038/embor.2010.81

Received 30 March 2010; revised 10 May 2010; accepted 12 May 2010; published online 11 June 2010
**RESULTS**

**Structure of a eukaryotic AGO MID domain**

The 2.2 Å structure of the *N. crassa* AGO MID domain (*Nc* QDE-2 MID; residues Lys 514–Gly 640) was solved by experimental phasing using a selenomethionine-substituted protein and was refined with excellent stereochemistry (supplementary Table S1 online). The protein adopts a Rossmann-like fold with a central four-stranded parallel β-sheet sandwiched between α-helices in a prokaryotic βαβαβαα topology (Fig 1A–D). The structure superimposes well with the structures of previously determined archaeal and eubacterial AGO MID domains, despite the low sequence identity (Fig 1C,D). Among the best-scoring structural relatives is the AGO MID domain from *Archeoglobus fulgidus* (*Af* AGO1; Fig 1D). The structure of the *Neurospora crassa* QDE-2 MID domain has no similarity to eubacterial AGO MID domains, despite the low sequence identity of 9%. For comparison, the r.m.s.d./sequence identity is 2.2 Å/16% for *Aquilegulus asciifolius* (As); PDB ID: Iyvu), 2.7 Å/12% for *Pyrococcus furiosus* (Pf); PDB ID: Iu04) and 3.3 Å/14% for *Thermus thermophilus* (Tt; PDB ID: 3h01). By contrast, the sequence identities for eukaryotic homologues are generally greater than 20% (24% for *Homo sapiens* (Hs) AGO2, 26% for *D. melanogaster* (Dm) AGO1; Fig 1D). The structure of the *Nc* QDE-2 MID domain thus serves as a prototype for eukaryotic AGO MID domains.

The *Nc* QDE-2 MID domain structure has no similarity to eukaryotic elf4E. Consequently, it is unlikely that eukaryotic AGO MID domains sandwich the m7GpppG mRNA 5′-cap between two aromatic amino-acid side chains in an elf4E-like manner, sandwiching the m7GpppG cap between two phenylalanine residues (Marchi et al, 1997). By contrast, Djuranovic et al (2010) modelled the MID domain onto the Rossmann-like fold of the archaeal and eubacterial homologues. This latter modelling leads the authors to propose allosteric control of cap binding by the presence of the guide RNA 5′-nucleotide in a conserved binding pocket corresponding to that observed in the prokaryotic MID domains.

In this study, we present the crys stal structure of the *N. crassa* quelling deficient 2 (QDE-2) MID domain (Fulci & Macino, 2007) as an example of a eukaryotic AGO MID domain. This structure convincingly refutes the idea that eukaryotic AGO MID domains adopt an elf4E-like fold and shows the guide RNA 5′-phosphate-binding site is conserved. Furthermore, we describe a second, previously unreported ligand-binding site that allows a simple mechanistic explanation for the allosteric effects reported by Djuranovic et al (2010).
manner. In our structure, the two residues (Ala 535 and Phe 570) corresponding to those proposed to be important for cap binding (Kirikidou et al., 2007) lie more than 25 Å apart (Fig 1A, purple). Earlier sequence analyses and homology modelling reached similar conclusions (Kinch & Grishin, 2009; Djuranovic et al., 2010). Furthermore, in previous studies, we showed that double valine substitutions of these residues abolished the silencing activity of Dm AGO1 and Hs AGO2 by preventing the interaction with both miRNAs and the GW182 protein, which is essential for miRNA-mediated silencing in animal cells (Eulalio et al., 2008b).

The 5′-phosphate-binding pocket is highly conserved

Several archaeal AGO proteins have been crystallized in complex with a guide RNA/DNA (Ma et al., 2005; Parker et al., 2005; Yuan et al., 2005; Wang et al., 2008a,b, 2009). The respective structures reveal that the 5′-terminal nucleotide of the guide molecule is not available for base pairing with the RNA target but is strongly bent and accommodated in a preformed pocket of the MID domain (Fig 1B). In each case, the base stacks onto an aromatic or arginine side chain (Fig 1D, red asterisk). Most importantly, the 5′-terminal phosphate (a hallmark of small interfering RNAs and miRNAs) is coordinated precisely by several side-chain and main-chain contacts to residues from helix α3, strand β4 and helix α4 (supplementary Fig S1A,B online; Parker et al., 2005; Wang et al., 2009). Furthermore, in the context of full-length proteins, the carboxyl terminus of the PIWI domain contacts the 5′-phosphate through a metal ion (Parker et al., 2005; Wang et al., 2009).

When a binary complex containing the Af MID and PIWI domains plus a guide RNA (PDB ID: 2bgg; Parker et al., 2005) is superposed onto the Nc QDE-2 MID domain, the 5′-terminal uridine of the guide RNA is positioned favourably into its positively charged binding pocket, with the base stacking on the aromatic ring of Tyr 595 of the Nc QDE-2 MID domain and the 5′-phosphate being coordinated by the Tyr 595 hydroxyl group (Fig 2A,B). Most importantly, the 5′-phosphate superimposes almost perfectly with a sulphate ion (ion I; from the crystallization conditions) that is coordinated additionally by the invariant lysines Lys 599 (helix α3) and Lys 638 (helix α4), as well as by the main-chain nitrogen of Cys 612 (strand β4) and the poorly conserved Lys 634 (helix α4; Fig 2A,B). Conversely, a superposition of the Nc QDE-2 MID domain onto the structures of Af or Tt MID domains places the sulphate ion (I) at the position of the 5′-phosphate of the guide RNA or DNA, respectively; this illustrates that the ligand-binding residues are conserved structurally (supplementary Fig S1A,B online). We conclude that eukaryotic AGO MID domains accommodate the 5′-terminal nucleotide of the guide RNA in a manner similar to their archaeal and eubacterial counterparts.

A second ligand-binding pocket

The structure of the Nc QDE-2 MID domain contains a second sulphate ion (II), only 6.3 Å from the first ion; this second ion is also well coordinated (Fig 3A,B). This second ligand-binding pocket is separated from the first 5′-phosphate-binding pocket by only two invariant lysines, Lys 599 and Lys 638, which thus participate simultaneously in coordinating both sulphate ions (Fig 3A,B). The second sulphate ion is coordinated further by the main-chain and side-chain atoms from Thr 610 (strand β4), which is invariantly a serine or threonine, and by the non-conserved side-chain of His 609 (strand β4; Fig 3A,B). As a consequence, the two sulphate-binding pockets are not independent of each other, but are likely to display positive cooperativity (that is, a ligand binding to either would orient the lysines favourably and promote binding of a ligand to the other site).

The presence of two positively coupled ligand-binding sites on the Nc AGO MID domain is particularly interesting in the context of a recent report by Djuranovic et al (2010), indicating...
that miRNA binding to the MID domain of Dm AGO1 might be under allosteric control (Djuramovic et al, 2010). These authors suggested that the MID domains of certain eukaryotic AGOs (that is, those involved in miRNA-mediated gene regulation) contain a second nucleotide-binding site in addition to the 5′-phosphate-binding pocket, which gains affinity for m7GpppG cap analogues (or eventually for other ligands) only in the presence of the guide-strand RNA (and vice versa). It is thus tempting to speculate that the second sulphate ion observed in the context of the Nc QDE-2 MID domain could occupy such a second ligand-binding site.

Ligand binding by the two sites is expected to be cooperative. Accordingly, Dm AGO1b binding to miRNAs, m7GpppG cap analogue and GW182 protein was abrogated by substituting the invariant tyrosine residue in the first binding site (Dm AGO1b residue Tyr 619, corresponding to Nc QDE-2 Tyr 595) or swapping the charge of an aspartic residue in the second ligand-binding site (Dm AGO1b residue Asp 627, corresponding to Nc QDE-2 Asp 603; Djuramovic et al, 2010). The inhibitory effect of the D627K substitution on miRNA binding is surprising because this residue is distant from the miRNA-binding site (Djuramovic et al, 2010). The equivalent residue in the Nc QDE-2 MID domain Asp 603 is located less than 7 Å away from the second sulphate (Fig 3A,B). Consequently, the Asp 627 side chain of Dm AGO1b could indeed participate in binding a ligand that superimposes with the second sulphate.

DISCUSSION

The structure of the Nc QDE-2 MID domain suggests that AGO proteins could be regulated by positive cooperativity of distinct ligands. Indeed, the structure reveals two precisely coordinated sulphate ions in adjacent binding pockets. Two highly conserved lysines (Lys 599 and Lys 638) coordinate both sulphates simultaneously and are thus shared by both binding pockets. Clearly, the first of the two sulphates on the Nc QDE-2 MID domain occupies the binding pocket for the 5′-terminal phosphate of the miRNA/short interfering RNA. The second sulphate is coordinated directly by three of the six most highly conserved side chains in AGO MID domains (that is, TiS610, Lys 599 and Lys 638; Fig 1D). These residues are conserved not only in sequence but also in their structural orientation in bacterial and archaeal MID domains (supplementary Fig S1A,B online). Thus, the second sulphate might mark an ancient second (allosteric) binding site that has evolved different specificities in different AGO homologues (for example, for the cap structure, GW182 or other ligands).

We cannot predict what the ligand for the second binding site might be for different AGOs, because the properties and accessibility of this second binding site would depend on the orientation and the nature of the side chains provided by the PIWI domain. Indeed, in the current structures of the archaeal and eu bacterial proteins, the PIWI domain restricts access to the second ligand-binding site. Nevertheless, the situation might be different for the eukaryotic AGOs, as Djuramovic et al (2010) observed an allosteric effect between the two binding sites using protein constructs that contained the PIWI domain. For AGO proteins involved in the miRNA pathway, the second ligand-binding site was suggested to bind to the m7GpppN cap structure, or residues from the GW182 protein (Djuramovic et al, 2010). In agreement with this hypothesis, substituting residues in either of the two binding pockets simultaneously abrogated Dm AGO1 binding to miRNAs, GW182 and m7GpppG cap analogue (Djuramovic et al, 2010). However, other AGO proteins, such as Nc QDE-2, might have evolved to be regulated by distinct ligands or might no longer be regulated by a second ligand. Nevertheless, for those AGO proteins that show ligand-dependent regulation, the coupled participation of the invariant lysines, Lys 599 and Lys 638, in both ligand-binding sites would be an attractive and mechanistically simple explanation for the ligand-dependent regulation of AGO protein function.

METHODS

Cloning, protein expression, purification and crystallization. The sequence encoding the MID domain of Nc QDE-2 was amplified by PCR from a pBluescript-5K + plasmid containing the genomic sequence of the qde-2 gene (provided by G. Macino) and subsequently inserted into the NcoI and NolI sites of
the pETM60 vector (derived from pET24d; Novagen) using the primer pairs: 5′-CATGGCATGCTAAGGTCGCTCGAAGCCCTT3′ (forward) and 5′-ATAAGAAATCAGGCGCCGAGTGATGTGAAGCGC3′ (reverse).

The resulting NusA–6 × His–MID protein fusion was expressed in the Escherichia coli BL21 Star (DE3) strain (Invitrogen) at 20 °C overnight. Protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside at an 

The fusion protein was purified first by a Ni²⁺-affinity chromatography and subsequent gel filtration (HiLoad 26/60 Superdex 75 pg; GE Healthcare). Finally, the QDE-2 MID domain was concentrated to 25 mg per ml in 10 mM HEPES (pH 7.2), 150 mM NaCl and 1 mM DTT.

Crystalline clusters of native and selenomethionine-substituted QDE-2 MID domain were grown by hanging drop vapour diffusion over a 500 μl reservoir at 18 °C. The protein solution described above was mixed 1:1 with various solutions containing 16–30% polyethylene glycol 4000 and 150–250 mM ammonium sulphate. Crystals were optimized by microseeding and flash-frozen in liquid nitrogen in the respective reservoir solution supplemented with 10 or 15% glycerol for cryoprotection.

Data collection, structure solution and refinement. Diffraction data were recorded on a PILATUS 6M detector (Broennimann et al, 2006) at the beamline PXII of the Swiss Light Source at a wavelength of 1.0 Å for the native data set and at the absorption peak of the selenomethionine derivative at 0.9792 Å. Diffraction images were processed with XDS (Kabsch, 1993). The crystals belong to space group C222₁, with one molecule per asymmetric unit and 52% solvent. The structure was solved from the selenomethionine data by single anomalous dispersion. AutoSHARP (Vonrhein et al, 2007) was used to search for five selenium sites per molecule. The assignment of the correct hand and solvent flattening was done automatically. ARP/wARP (Cohen et al, 2008) built a partial model comprising 99% of the backbone and 93% of the side chains docked correctly. The model was completed manually in COOT (Emsley & Cowtan, 2004) by using the native data. Refinement was done using Phenix (Adams et al, 2010) and COOT iteratively. The structure was validated with MOLPROBITY (Davis et al, 2007) and WHATCHECK (Hooft et al, 1996).

Accession codes. The coordinates of the MID domain have been deposited at the PDB with the accession code 2xdy.

Supplementary information is available at EMBO reports online (http://www.emboreports.org).

ACKNOWLEDGEMENTS

We thank G. Macino for providing Neurospora crassa QDE-2 plasmid DNA and R. Böttner for technical assistance. We thank the staff at the PX beamlines of the Swiss Light Source for assistance with data collection. This study was supported by the Max Planck Society, by grants from the Deutsche Forschungsgemeinschaft (FOR855 and the Gottfried Wilhelm Leibniz Program awarded to E.I.), and by the Sixth Framework Programme of the European Commission through the Silencing RNAs: Organisers and Coordinators of Complexity in Eukaryotic Organisms (SIROCCO) Integrated Project LSHG-CT-2006-037900.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Adams PD et al (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66: 213–221

Broennimann C et al (2006) The PILATUS 1M detector. J Synchrotron Radiat 13: 120–130

Cohen SX et al (2008) ARP/wARP and molecular replacement: the next generation. Acta Crystallogr D Biol Crystallogr 64: 49–60

Davis IW et al (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35: W573–W583

Djuranovic S, Zinchenko MK, Hur JK, Nahvi A, Brunelle JL, Rogers EJ, Green R (2010) Allosteric regulation of Argonaute proteins by miRNAs. Nat Struct Mol Biol 17: 144–150

Emsley P, Cowtan K (2004) COOT: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126–2132

Eulalio A, Hurtzinger E, Izaurralde E (2008a) Getting to the root of miRNA-mediated gene silencing. Cell 132: 9–14

Eulalio A, Hurtzinger E, Izaurralde E (2008b) GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and miRNA decay. Nature Struct Mol Biol 15: 346–353

Fulci V, Macino G (2007) Quelling: post-transcriptional gene silencing guided by small RNAs in Neurospora crassa. Curr Opin Microbiol 10: 199–203

Hooft RW, Friend G, Sander C, Abola EE (1996) Errors in protein structures. Nature 381: 272

Jinek M, Doudera JA (2009) A three-dimensional view of the molecular machinery of RNA interference. Nature 457: 405–412

Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Crystallogr 26: 795–800

Kinch LN, Grinich NV (2009) The human Ago2 MC region does not contain an eIF4E-like mRNA cap binding motif. Biol Direct 4: 2

Kirikidou M, Tan GS, Laimprinaki S, De Planell-Saguer M, Nelson PT, Morelaut Z (2007) An mRNA m7G cap binding-like motif within human Ago2 represses translation. Cell 129: 1141–1151

Lingel A, Simon B, Izaurralde E, Sattler M (2003) Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. Nature 426: 465–469

Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ (2005) Structural basis for 5′-end-specific recognition of guide RNA by the A. fulgidus PIWI protein. Nature 434: 666–670

Marnocotrigiano J, Gringas AC, Sonenberg N, Burley SK (2007) Cocrystal structure of the messenger RNA 5′ cap-binding protein (eIF4E) bound to 7-methyl-cGDP. Cell 99: 951–961

Parker JS, Roe SM, Barford D (2004) Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. EMBO J 23: 4727–4737

Parker JS, Roe SM, Barford D (2005) Structural insights into mRNA recognition from a PIWI domain–siRNA guide complex. Nature 434: 663–666

Parker JS, Parizotto EA, Wang M, Roe SM, Barford D (2009) Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. Mol Cell 33: 204–214

Song JJ, Liu J, Tolia NH, Schneiderman J, Smith SK, Martienssen RA, Hannon GJ, Joshua-Tor L (2003) The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat Struct Mol Biol 10: 1026–1032

Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305: 1434–1437

Tolia NH, Joshua-Tor L (2007) Slicer and the Argonautes. Nat Chem Biol 3: 36–43

Vonrhein C, Blanc E, Roversi P, Bricegogne G (2007) Automated structure solution with autoSHARP. Methods Mol Biol 364: 215–230

Wang Y, Sheng C, Juranić S, Tuschl T, Patel DJ (2008a) Structure of the guide-strand-containing Argonaute silencing complex. Nature 456: 209–213

Wang Y, Juranić S, Li H, Sheng G, Tuschl T, Patel DJ (2008b) Structure of an Argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature 456: 921–926

Wang Y, Juranić S, Li H, Sheng G, Wardle GS, Tuschl T, Patel DJ (2009) Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature 461: 754–761
Yan KS, Yan S, Farooq A, Han A, Zeng L, Zhou MM (2003) Structure and conserved RNA binding of the PAZ domain. Nature 426: 468–474

Yuan YR, Pei Y, Ma JB, Kuryavyi V, Zhadina M, Meister G, Chen HY, Dauter Z, Tuschi T, Patel DJ (2005) Crystal structure of A. aeolicus Argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. Mol Cell 19: 405–419