Published online 18 February 2011 *Nucleic Acids Research, 2011, Vol. 39, No. 11* ***4827–4835***
doi:10.1093/nar/gkr077

The conformation of microRNA seed regions in native microRNPs is prearranged for presentation to mRNA targets

Nicole J. Lambert, Sam G. Gu and Alan M. Zahler*

Department of MCD Biology and the Center for Molecular Biology of RNA, University of California, Santa Cruz, CA 95064, USA

Received July 22, 2010; Revised January 30, 2011; Accepted January 31, 2011

ABSTRACT

MicroRNAs control gene expression by post-transcriptional down-regulation of their target mRNAs. Complementarity between the seed region (nucleotides 2–8) of a microRNA and the 3′-UTR of its target mRNA is the key determinant in recognition. However, the structural basis of the ability of the seed region to dominate target recognition in eukaryotic argonaute complexes has not been directly demonstrated. To better understand this problem, we performed chemical probing of microRNAs held in native argonaute-containing complexes isolated from *Caenorhabditis elegans*. Direct probing of the RNA backbone in isolated native microRNP complexes shows that the conformation of the seed region is uniquely constrained, while the rest of the microRNA structure is conformationally flexible. Probing the Watson–Crick edges of the bases shows that bases 2–4 are largely inaccessible to solvent, while seed region bases 5–8 are readily modified; collectively our probing results suggest a model in which these bases are primed for initiating base pairing with the target mRNA. In addition, an unusual DMS reactivity with U at position 6 is observed. We propose that interaction of miRNAs with argonaute proteins pre-organizes the structure of the seed sequence for specific recognition of target mRNAs.

INTRODUCTION

Gene silencing by microRNAs (miRNAs) is a global post-transcriptional regulatory process affecting a wide array of cellular functions in eukaryotic cells. Since the initial discovery of *lin-4* and *let-7*, two miRNAs that regulate developmental timing in *Caenorhabditis elegans* (1–3), thousands of miRNAs have been discovered in plants and animals (4). miRNAs are targeted to specific messenger RNAs through base pairing complementarity (5). This interaction ultimately leads to mRNA destabilization (6–13) and/or translational inhibition (1,14–17).

Initial computational methods designed to identify the targets of miRNAs discovered that Watson–Crick base pairing between nucleotides 2 and 8 of miRNAs and their targets was an important determinant in target recognition (18–20). Biochemical and biophysical evidence has shown that this region contributes an increased binding affinity for target sequences relative to other parts of the miRNA (21,22). Additionally, systematic reporter studies have shown that functional regulation by miRNAs is highly sensitive to base pair mismatches within nucleotides 2–8 of the miRNA, which have been defined as the seed region (23–25). Although the functional importance of seed region complementarity as the major determinant of miRNA targeting is well established (19–21,26–28), obtaining direct evidence of the structural basis underlying the ability of the seed region to dominate target recognition in eukaryotes is still warranted.

Argonaute proteins are a large family of PIWI and PAZ domain-containing proteins that associate with and act as effectors for a range of small RNAs, including silencing RNAs, PIWI RNAs and microRNAs (29,30). Crystal structures of eukaryotic MID domains bound to a 5′-terminal nucleotide have been solved, illustrating how argonaute can have 5′-base and monophosphate specificity (31,32). To date, no crystal structure of a eukaryotic microRNP complex has been reported, and there is a need to further explore the biochemical nature of the seed region in the native complex. To address this, we have isolated native argonaute-containing miRNP complexes from *C. elegans* and subjected them to chemical probing strategies designed to reveal the backbone flexibility and base pairing potential of the bound miRNAs. Based on our findings, interactions between the argonaute protein and the microRNA seed region specifically pre-organize

*To whom correspondence should be addressed. Tel +1 831 459 5131; Fax +1 831 459 3139; Email: amzahler@ucsc.edu
Present address:
Sam G. Gu, Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305-5324, USA.

and optimize the conformation of these bases for base pairing with its mRNA target sequence. Conversely, we propose that the lack of direct contact of argonaute with the 3′-portion of miRNAs, deemphasizes the role of these bases in target discrimination.

MATERIALS AND METHODS

Purification of miRNPs

Endogenous miRNP complexes were enriched by fractionating mixed stage *C. elegans* cellular extract as described (33). Native miRNPs were fractionated using successive chromatographic steps of anion (monoQ) and cation exchange (monoS) columns. miRNA-containing fractions and naked miRNAs extracted from these fractions, were collected and used for chemical probing experiments. To measure the concentration of total miRNAs, the extracted miRNAs were treated with calf intestine alkaline phosphatase and subsequently labeled using T4 polynucleotide kinase (NEB) with [γ-^{32}P]ATP and separated on 10% PAGE. The radioactive signals were compared to a standard of labeled 23-nt RNA oligos.

microRNP affinity chromatography

A control RNA, not complementary to known miRNAs (5′-CAAUCCCUUAGAGGCAGAAUGGUUGUAUA AAGAACCAUUG) or a bulged RNA imperfectly complementary to mir-58 (5′-AAAAAAUUGCCGACGAUC UCAAAAAA) were transcribed *in vitro* from DNA oligonucleotides using T7 RNA polymerase. Control or mir-58 target RNA of 500 pmols were covalently immobilized to beads, as previously described in ref. (34). *Caenorhabditis elegans* miRNPs (40 pmols) from an ALG-1::GFP expressing strain were incubated with RNA-coupled beads at 4°C for 1 h (WM84—a gift from Craig Mello) (35). Proteins that bound to RNA-coupled beads were detected by immunoblots using anti-TSN-1 (a gift from Ron Plasterk) or anti-GFP antibodies (Roche).

For DNA affinity experiments, either 800 pmols of biotinylated negative control DNA (AATCCCTTAGAGGC AGAATGGTTGTATAAAGAACCATT) or a mixture of biotinylated DNA oligos perfectly complementary to the four most abundant *C. elegans* microRNAs as determined by Gu *et al.* (33) (mir-58 AAAAAATTGCC GTACTGAACGATCTCAAAAAA, mir-52 AAAAAAG CACGGAAACATATGTACGGGTGAAAAA, mir-66 AAAAATCACATCCCTAATCAGTGTCATGAAAAA AAAA and mir-71 AAAAATCACTACCCATGTCTTT CAAAAAA) were bound to streptavidin beads (Promega). *Caenorhabditis elegans* miRNPs were incubated with DNA-bound streptavidin beads for 30 min at room temperature. Enriched proteins were precipitated with TCA, resolved on 10% SDS–PAGE and silver stained using SilverSNAP Stain for Mass Spectroscopy (Thermo Scientific). microRNP-specific proteins bands were cut out of the SDS–PAGE gel and peptide fragments were generated by an in-gel trypsin digest. ALG-1 and ALG-2 were identified by LC/MS/MS (Thermo Finnigan).

RNA modification with kethoxal, dimethyl sulfate and NMIA

For the kethoxal modification using the enriched miRNPs, 10 µl of 250 mM kethoxal were added to 300 µl of the miRNP fraction (40 pmols total miRNAs, 100 mM HEPES–NaOH, pH 7.6, 140 mM KCl, 5 mM $MgCl_2$). The reaction mixture was incubated at 30°C for 10 min. To stop the reaction, 30 µl of 250 mM sodium borate, 30 µl of 3 M sodium acetate pH 5.0, and 750 µl of ethanol were added. The procedure for dimethyl sulfate (DMS) modification of miRNAs was adapted from Moazed *et al.* (36). The reaction mixtures were incubated at 30°C for 10 or 20 min. To stop the reactions, 200 µl of 1 M 2-mercaptoethanol were added. To monitor backbone flexibility, 1/10 v of either DMSO, 65mM (+) or 130mM NMIA (++) were added to 40 pmols of total miRNPs. The reaction mixtures were incubated at 37°C for 45 min.

Where indicated, 100 pmols of target DNA oligos perfectly complementary to mir-58 (GCATGGGCTACTCT CATTGCCGTACTGAACGATCTCAAAAAA), perfectly complementary to the seed region (AAAAATAAC GGCATGACTTCGATCTCAAAAA), complementary to mir-58 outside of the seed region (AAAAAATTGCC GTACTGAAGCTAGAGAAAAA), a bulged mir-58 target (AAAAATTGCCGTACCGATCTCAAAAA) or a control DNA oligo (CCCTTAGAGGCAGAATGGT TGTATAAAGAACCA) were added to miRNPs and incubated at room temperature for 30 min prior to DMS or NMIA treatment. In the case of probing synthetic mir-58 RNA in the presence of DNA targets, 0.5 pmols of mir-58 RNA was first incubated with 2.5 pmols of the target DNA oligo. The mixture was incubated at 90°C for 1 min, 65°C for 2 min and 37°C for 10 min and then chemically modified.

Splinted ligation

Purified, chemically-modified miRNAs were splint-ligated to an extension RNA sequence (pGAGAGUAGCCCAU GC). The splint ligation method was adapted from previously described work (37). Two picomoles of extension RNA, 1.5 pmols of a mir-58 DNA splint oligo (GCATG GGCTACTCTCATTGCCGTACTGAACGATCTCA) and modified RNA were incubated at 90°C for 1 min, 65°C for 2 min and 37°C for 10 min. Splint ligation reactions proceeded with T4 RNA ligase 2 (NEB) at 37°C for 30 min and were subsequently treated with 2 U of RQ1 RNase-free DNase (Promega) for 10 min.

Primer extension assays

In order to detect NMIA, DMS and kethoxal modifications, primer extension assays were used. miRNAs (extracted from each chemical probing reaction) were mixed with 1 µl 0.5 µM ^{32}P 5′-end-labeled DNA oligo complementary to either the 3′-end of specific miRNAs (mir-58 ATTGCCGTAC, mir-66 TCACATCCCT, mir-80 TCGGCTTTCA) or, if a splint ligation was performed, a labeled DNA oligonucleotide complementary to the extension RNA sequence (GCATGGGCTACTCTCA). miRNAs were reverse transcribed at 42°C for 30 min

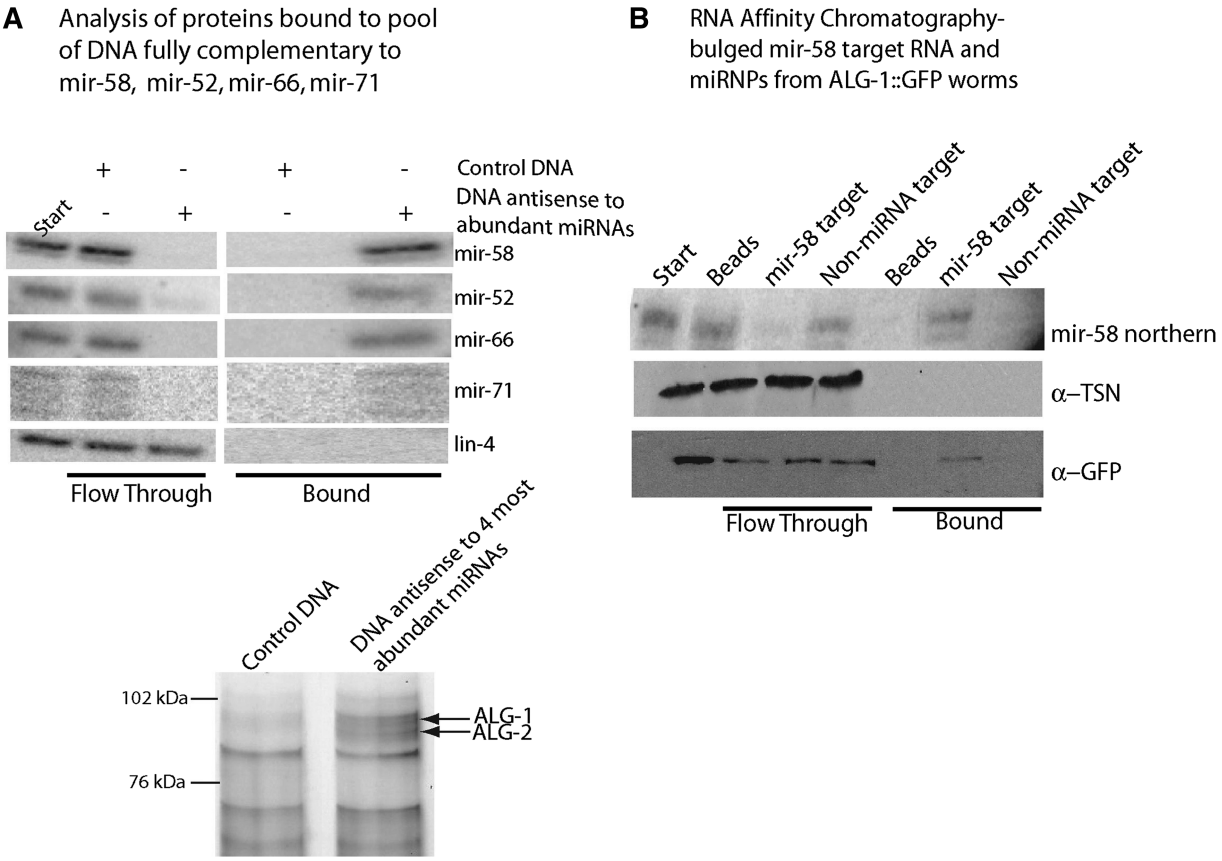

Figure 1. Native *C. elegans* microRNPs contain ALG-1/ALG-2 and bind to complementary targets. (**A**) A biotinylated control DNA sequence, or a mixture of biotinylated DNA oligos perfectly complementary to four abundant miRNAs (mir-58, mir-52, mir-66, mir-71) were bound to streptavidin beads and incubated with enriched miRNPs from N2 worms. Northern blots probed with mir-58, mir-52, mir-66, mir-71 and negative control lin-4 show the enrichment. Silver stain of SDS–PAGE shows co-purified proteins. ALG-1 and ALG-2 were identified by mass spectrometry. (**B**) An *in vitro* transcribed control RNA and an imperfectly complementary synthetic RNA target for mir-58 were covalently immobilized to beads. RNA-bound beads and a beads-only control were incubated with enriched miRNP fractions from an ALG-1::GFP expressing strain, WM84. The enrichment of mir-58 was detected by northern blot; ALG-1::GFP and TSN-1 were detected by immunoblots.

using SuperScript III (Invitrogen). Samples were ethanol precipitated, resolved on a 10% polyacrylamide gel and detected by PhosphorImager (Molecular Dynamics).

Modification analysis

SHAPE chemical probing data were quantified using semi-automated footprinting analysis (SAFA) (38). The quantitations for each nucleotide position for either miRNP or miRNA complexes are simply normalized to the background in the DMSO (no NMIA) control (Figure 2C) or are displayed as enhancement of chemical reactivity in miRNPs (reactivity in lanes ± NMIA) over that of miRNAs (reactivity in lanes ± NMIA) (Figure 2D). For quantitation of DMS reactivity in Figure 3, ImageJ software (39) was used to determine the area under the peak for each band in the unmodified control and DMS-modifed lanes for the different bands in .tiff images.

RESULTS

To determine the structural basis of the seed region's importance, we isolated native *C. elegans* microRNP complexes for structural probing experiments. We previously reported the purification of native miRNPs from *C. elegans* using a combination of size exclusion, anion (monoQ) and cation (monoS) exchange chromatographic steps; >95% of the RNAs purified from this biochemically enriched miRNP fraction consisted of known miRNAs (33). No target mRNAs were detected in these high-throughput sequencing experiments, as all of the RNAs in the enriched miRNP fraction were 21–25 nt in length. To further characterize the protein content of these purified miRNPs, we prepared a mixture of biotinylated DNA oligonucleotides complementary to the four most abundant *C. elegans* miRNAs (mir-58, mir-52, mir-62 and mir-71), which together account for 43% of the pool of *C. elegans* miRNAs (33). The oligonucleotides were immobilized on streptavidin beads and used to specifically enrich their miRNP protein partners. Figure 1A shows specific enrichment of the targeted miRNAs on the immobilized oligonucleotides. Specific proteins that co-purified with these abundant miRNAs were analyzed by mass spectrometry. The argonautes ALG-1 and ALG-2, homologs of the human argonaute AGO2 that is found associated with miRNAs (40), were the only

proteins from the miRNP-enriched fraction that specifically co-purified with the four abundant miRNAs (Figure 1A).

To confirm that miRNAs in our purified miRNP fraction are associated with argonaute proteins, miRNPs were isolated from a *C. elegans* strain that expressed a GFP-tagged ALG-1 protein. mir-58 was specifically purified from the miRNP pool using an immobilized bulged target RNA. A GFP-tagged ALG-1 protein specifically co-purifies with mir-58 (Figure 1B), but TSN-1, another protein that can be detected in our miRNP-enriched fraction, does not. These results demonstrate that native miRNPs are argonaute-containing complexes that can perform the first essential step of miRNA-mediated gene regulation, binding to a specific target RNA sequence.

Native *C. elegans* miRNP complexes isolated by anion and cation exchange chromatography were used for chemical probing experiments. Abundant miRNAs in these microRNP pools were probed with chemical reagents to elucidate the structural properties of the seed region of microRNAs complexed with argonaute proteins.

microRNA sugar–phosphate backbone probing

It has been hypothesized that argonaute proteins can somehow pre-organize the seed region of miRNAs, explaining the seed region's distinct role in target recognition (5,21,22,41). In order to test this, we performed selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) assays. In this procedure, nucleotide reactivity to the chemical N-methylisatoic anhydride (NMIA) is used to monitor RNA backbone flexibility (42). NMIA reactivity is sensitive to the sugar pucker conformation of ribose; it only reacts with the 2′-OH of ribose when in the C2′-endo conformation. If the backbone of the miRNA is constrained by RNA–RNA interactions in an A-form helix, or by RNA–protein interactions that limit flexibility and prevent transition into the C2′-endo conformation, the 2′-OH group will not be reactive with NMIA. In addition, protein contact with a 2′-OH group could also confer protection from NMIA reactivity to a specific nucleotide position.

In our chemical probing assays, the purified microRNP fraction, or a control of naked miRNA extracted from the miRNP fraction, was modified by NMIA. Then, the chemically modified miRNAs were recovered and a short RNA oligonucleotide was ligated to the 3′-end of a specific miRNA in the pool using a variation of the splint-ligation protocol of Maroney *et al.* (37) (Figure 2A). Chemically-modified bases were detected as pauses in reverse transcription from a primer sequence that was complementary to the RNA oligo used in the splint ligation reaction. The specificity of this method was confirmed by dideoxy sequencing the splint-ligated miRNA (Figure 2B). Given that we are probing a pool of all native *C. elegans* miRNPs or the miRNAs extracted from them, this experimental strategy allowed detection of the modification of specific miRNAs from the pool.

Robust NMIA reactivity of bases outside the seed region was observed for mir-58 (Figure 2B, lanes 6 and 7) and mir-52 (Figure 2B, lanes 12 and 13) native miRNPs. In contrast, the seed regions of these miRNP complexes (with the exception of position 8) appear completely unreactive to NMIA (Figure 2B). mir-58 position 8, the last base of the seed region, is strongly reactive with NMIA, and mir-52 position 8 is moderately reactive with NMIA. mir-58 position 9, just outside the seed region, is strongly reactive with this chemical, but due to background stops by reverse transcriptase seen in the DMSO only control, we cannot assay whether position 9 of the mir-52 miRNP is protected. NMIA protections are completely limited to nucleotides 1–7 of the miRNAs, illustrating that sugar–phosphate backbone rigidity is a hallmark of seed nucleotide positions and/or that every 2′-OH group in this region has contacts with the argonaute protein. Interestingly, the high level of NMIA reactivity outside of the seed region is significantly greater in the miRNP than in the naked miRNA (Figure 2B, lanes 9 and 15), indicating that the miRNP complex not only organizes the seed region but perhaps also helps to induce increased flexibility of the other bases (Figure 2B–D).

Naked mir-58 and mir-52 in the total miRNA pool are not uniformly modified by NMIA, and more weakly modified overall than mir-58 and mir-52 in miRNPs (Figure 2B–D). This is likely due to the formation of secondary structure and/or intermolecular interactions with other RNAs in the pool. The SHAPE probing results for the naked miRNAs are consistent with the structure prediction for mir-58 from m-fold (43,44) where bases 6–9 and 15–18 of mir-58 can potentially form base pairs (Supplementary Figure S1). This is confirmed by the lack of reactivity of A7, G8, A15, C16 and G17 in naked miRNA probed with DMS and kethoxal (Figure 3B). For mir-52 miRNA, bases 6–9 and 20–23 are predicted to form interactions (Supplementary Figure S1). In the native miRNP, both mir-52 and mir-58 are heavily modified by NMIA outside the seed region, showing phosphate backbone flexibility. Because predicted secondary structures for mir-58 and mir-52 both involve bases 6–9 which include parts of the seed region, it is difficult to use naked miRNAs as a control for NMIA reactivity in these miRNPs. However, the nucleotide positions outside of the seed region for mir-58 and mir-52, including those that form foldback structures with positions 6–9 in the naked miRNA, exhibit much stronger reactivity when loaded into miRNPs, while the seed region positions 2–7 do not. This is consistent with the argonaute protein inducing conformational constraints to this part of the seed region. These results also illustrate that the argonaute functions to prevent promiscuous inter- and intramolecular interactions that the naked miRNA might typically form.

Base solvent accessibility

To test if endogenous eukaryotic miRNP complexes differentially hold the seed region so that the Watson–Crick edges are accessible for base pairing, we chemically modified miRNP complexes with DMS and kethoxal. DMS covalently modifies A and C residues and kethoxal modifies G residues; these modifications inhibit the

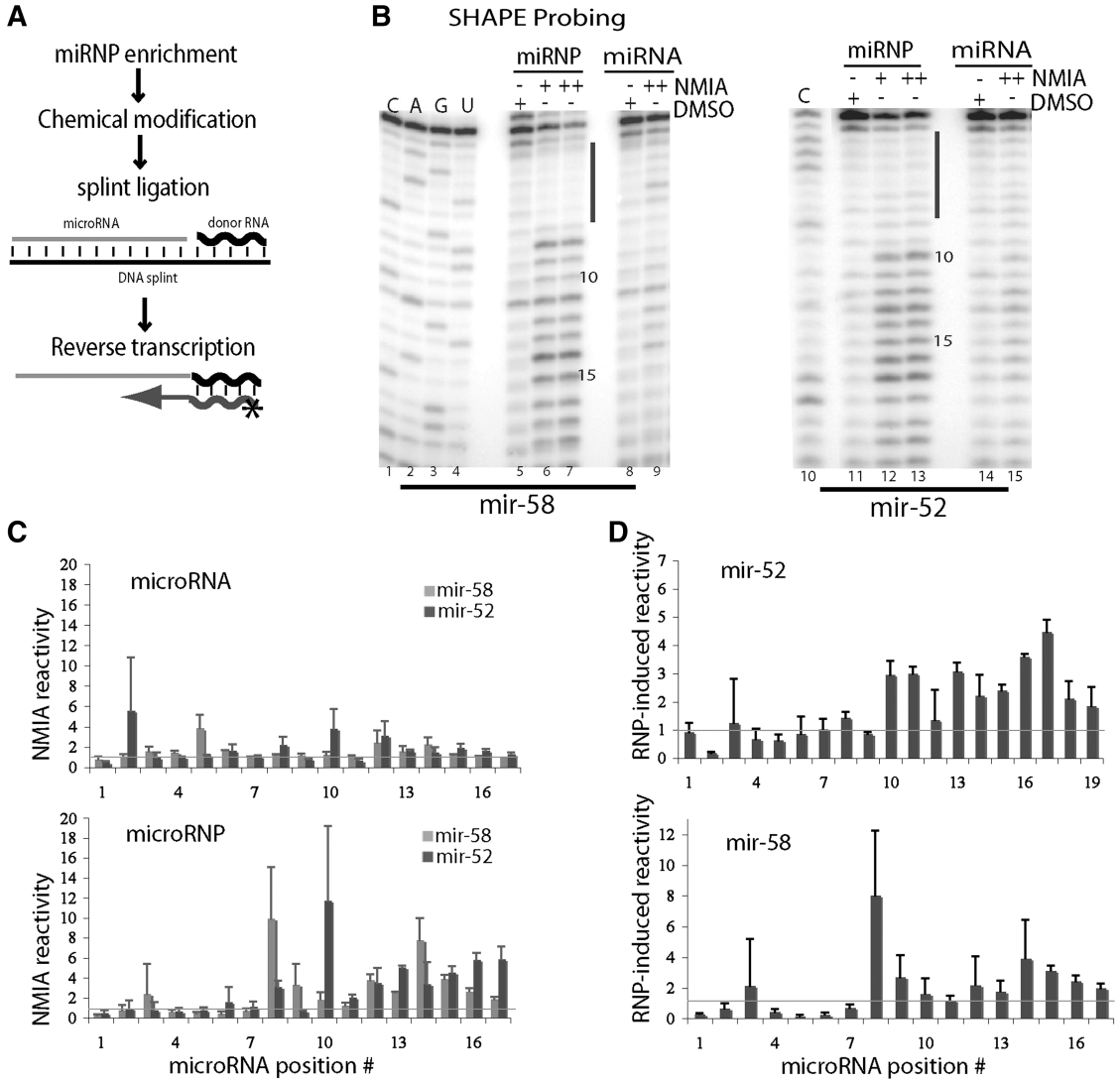

Figure 2. miRNA seed regions are defined by structural backbone rigidity. (**A**) Schematic for splinted ligation. (**B**) Probing of mir-58 and mir-52 miRNPs by SHAPE. DMSO (control) or NMIA were added to miRNPs. The miRNAs were extracted, ligated to a donor sequence and modifications detected. (**C**) Quantitation of NMIA reactivity of mir-58 and mir-52 in naked RNA or in miRNPs. For each nucleotide position, the intensity in the NMIA-modified lane relative to the DMSO control was determined. Results from three independent experiments are plotted. (**D**) Measurement of the miRNP induced reactivity, NMIA chemical reactivity of mir-58 and mir-52 containing miRNPs relative to the chemical reactivity of the corresponding naked miRNA. For each position, the reactivity of the modified miRNP backbone position determined in (C) was divided by reactivity of the naked miRNA at that position (also determined in (C).

progress of reverse transcriptase, allowing the detection of solvent-exposed nucleotides in primer extension assays (36,45). Modifications were detected by reverse transcription from a 10 nt ^{32}P 5′-end-labeled DNA oligonucleotide primer (Figure 3A and C) or by the splinted ligation method described earlier (Figure 3B). For mir-58 miRNPs, modification by DMS of A and C residues in the seed region (A3, A5, C7) is detected, with A3 being the weakest substrate (Figure 3A, lanes 15 and 16; Figure 3B, lanes 2 and 3). Interestingly, there is also an uncharacteristic DMS-reactive uracil at position 6 in mir-58 miRNPs; naked mir-58 RNA does not show this unusual reactivity (Figure 3A, lane 18, Figure 3B, lanes 5 and 6). mir-66-containing miRNPs are also highly modified in seed residues 5–8 (A5, C6, A7, C8) (Figure 3A, lanes 6 and 7). Consistent with results for mir-58, DMS hyperreactivity is also observed at position 6.

Low chemical reactivity for the first 4 nt of miRNPs, is exemplified by the weak reactivity of mir-66 nucleotides C1 and A2 with DMS (Figure 3A, lanes 6 and 7). Protection of the first nucleotide in miRNPs from DMS and kethoxal modification is not unexpected; structural studies suggest that argonaute has a 5′-phosphate binding pocket (46,47). Additionally, complementarity between first nucleotide of miRNAs and the cognate target RNA is not required for functional miRNA-mediated down-regulation (22). Kethoxal probing of mir-58 miRNPs surprisingly shows complete protection

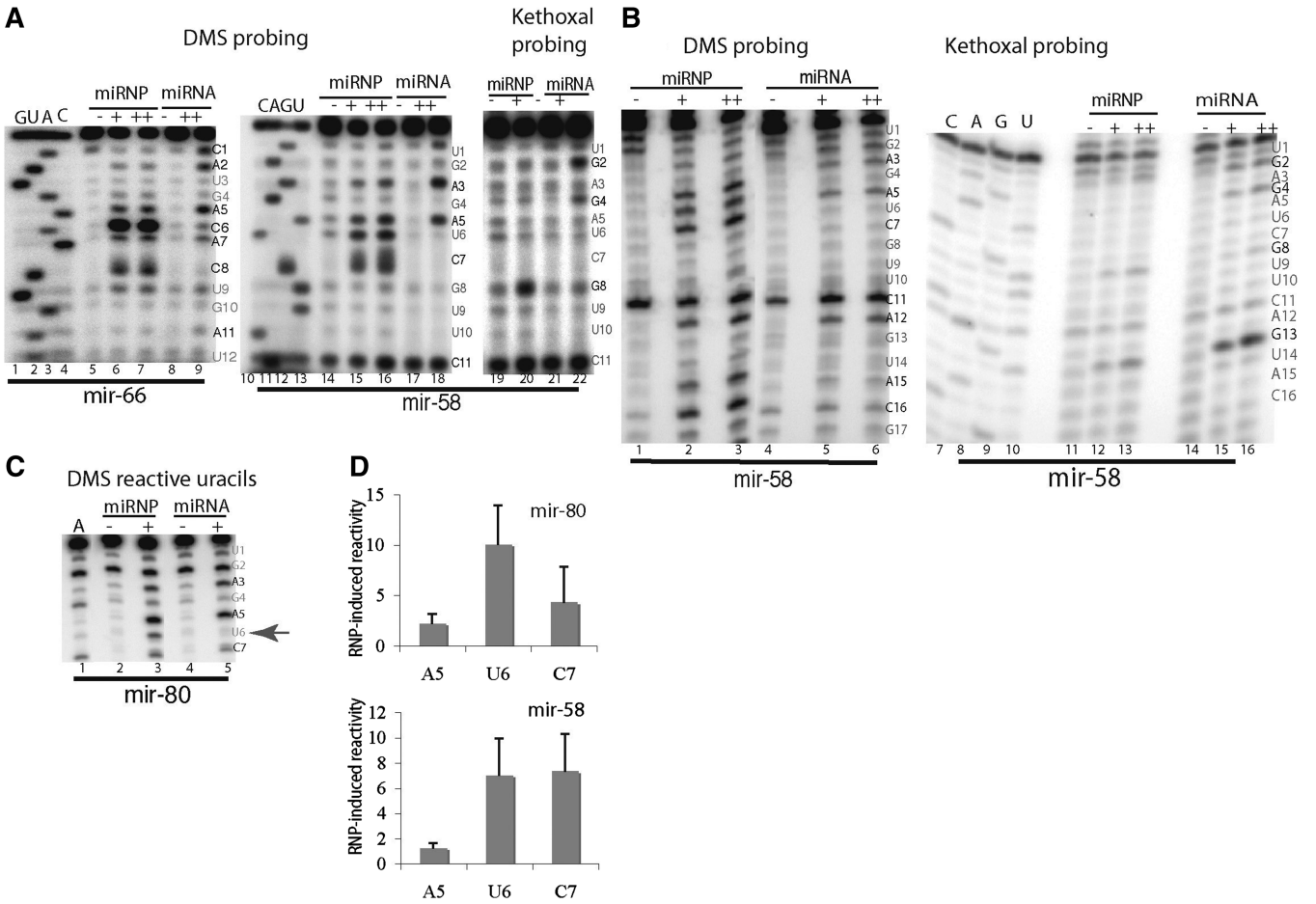

Figure 3. Seed region nucleotides exhibit unusual chemical reactivity to DMS and show protections from chemical modification (**A–C**). DMS or kethoxal were used to probe accessibility of bases of enriched *C. elegans* miRNPs, or miRNAs extracted from the enriched miRNPs. For mir-58 in (B), modifications were detected by splinted ligation. For the others, a ^{32}P labeled oligonucleotide complementary to the terminal 10 nt of each indicated miRNA was used as primer. '+' and '++' indicate DMS incubation times of 10 and 20 min. Note the DMS reactive uracils at position 6 of mir-58 and (C) mir-80. (**D**) RNP-induced reactivity is plotted for three adjacent seed region bases. Naked miRNA and miRNP DMS reactivity was quantified. Reactivity of the modified position in miRNPs relative to the DMS reactivity in naked miRNAs is the RNP-induced reactivity.

of seed region positions G2 and G4, but strong reactivity at G8 (Figure 3A, lane 20; and Figure 3B, lanes 12 and 13). A3 of mir-80 (Figure 3C, lane 3) and A3 of mir-58 (Figures 3A, lane 15 and 16) in microRNPs are moderately reactive with DMS, but relatively weakly modified compared to A5 in both. This indicates some variability in solvent accessibility of the third position in miRNPs. If chemical probing is performed following splinted ligation allowing for the detection of modifications in the 3′-regions of miRNAs, the bases downstream of the seed region are robustly reactive with DMS (A12, A15, C16; Figure 3B, lanes 2 and 3) and kethoxal (G13), indicating that bases downstream of the seed region are accessible to solvent. The protection of mir-58 bases C7, G8, A15, C16 from DMS and kethoxal probing in naked RNA (Figure 3B) is consistent with the predicted secondary structure of mir-58 (Supplementary Figure S1). The fact that these bases become strongly reactive in the miRNP suggests that the argonaute protein prevents intramolecular interactions from occurring.

microRNPs induce conformational change in seed regions

Illustrating the unique holding of the seed region by miRNP complexes is the robust U6 reactivity reproducibly observed in DMS probing of mir-58 miRNPs (Figure 3A, lanes 15 and 16). U6 is also strongly reactive in mir-80 miRNPs (Figure 3C, lane 3), and unreactive in their respective naked RNAs (Figure 3C, lane 5). Uracils at positions 3, 5, 7, 9 and 10 in other miRNPs tested are not reactive with DMS, demonstrating that uracil DMS reactivity in miRNPs is position-dependent. These results demonstrate that miRNPs are capable of specifically altering base chemistry at a particular position of the seed region.

Changes in chemical probing upon binding of targets

To insure that all of the chemical modifications detected in this study were specific to miRNP complexes capable of binding to target sequences, we tested whether the binding of miRNPs to complementary DNA substrates excludes

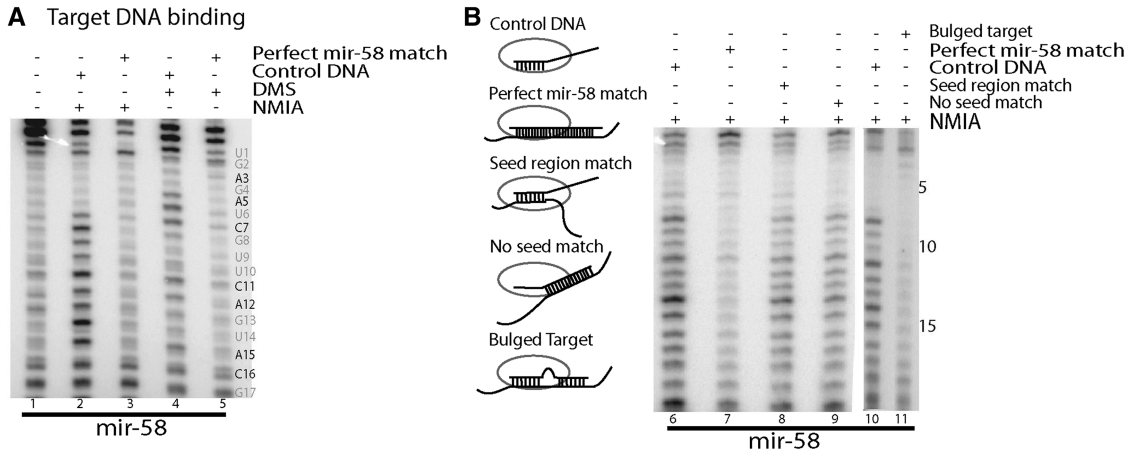

Figure 4. microRNPs require seed sequence complementarity to target sequences for substrate recognition. (**A** and **B**) A perfectly complementary mir-58 target DNA, a control DNA sequence, a target that is only complementary to the seed region, a bulged complementary DNA target, or a target DNA that is complementary to all residues outside of the seed region were incubated with miRNPs and then probed with either DMS or NMIA. The modifications of mir-58 were detected after splinted ligation.

miRNAs from DMS and NMIA chemical reactivity. Although miRNPs recognize RNA substrates, they can also bind to DNA targets in a sequence-specific manner (Figure 1A). DNA targets were used because they can be digested by DNase after the chemical modification of microRNPs, so they can be removed and not interfere with the subsequent splinted ligation reaction. Annealing target DNA oligos perfectly complementary to mir-58 protects mir-58-containing miRNPs from DMS and NMIA modifications, including at U6 (Figure 4A, lanes 3 and 5), while a non-complementary control DNA does not (Figure 4A, lanes 2 and 4). This illustrates that the microRNPs are capable of base pairing for the full two helical turns to a perfectly complementary target sequence. Synthetic mir-58 RNA can also be modified by NMIA (Supplementary Data S2, lane 15), all modified positions are also inhibited if SHAPE is performed in the presence of a perfect DNA target (Supplementary Data S2, lane 12). When a target DNA sequence that is complementary to the first nucleotide and the seed region of mir-58 is annealed to miRNPs, NMIA still modifies all positions outside of the seed sequence (Figure 4B, lane 8). Likewise, a target that is complementary to all bases outside of the seed region (bases 8–22) does not protect any microRNA bases in the microRNP from NMIA modification (Figure 4B, lane 9). This demonstrates that if the seed-region is not complementary to a target, the microRNA in a miRNP complex cannot base pair efficiently to a target sequence even if all of the other nucleotides (positions 8–22) are complementary. A bulged DNA target that is complementary to mir-58 except at base positions 9–12, protects the full length of mir-58 miRNPs from modification by NMIA (Figure 4B lane 11). SHAPE and kethoxal modification patterns of synthetic mir-58 RNA do not change in the presence of a DNA target complementary only to the seed region (Supplementary Data S2, lanes 4 and 13), indicating that mir-58 RNA on its own does not anneal strongly to the seed region match target. However, the 3′-region of mir-58 RNA is protected from both kethoxal and NMIA modification in the presence of a target that is complementary to all bases outside of the seed region (bases 8–22) (Supplementary Data, lanes 5 and 14). In the absence of the RNP complex, mir-58 RNA can base pair to a non-seed matched target DNA. These data are consistent with the model that in the microRNP, argonaute proteins confer to the miRNA the ability to target mRNAs based on the sequence of the seed region.

DISCUSSION

Our findings directly show that the seed regions of miRNAs are conformationally pre-arranged by argonaute proteins. The complete protection of seed regions from NMIA modification may arise from extensive hydrogen bond interactions between ALG-1/ALG-2 and each of the 2′-OH positions in the seed region. Alternatively, since NMIA only reacts with the 2′-OH on ribose when it adopts a C2′-endo conformation, the lack of NMIA reactivity in the seed region could indicate that miRNPs stabilize a C3′-endo conformation of the ribose groups. This can occur without direct contact with the ribose, as seen in A-form double-stranded RNA helices in which the helical conformation stabilizes the C3′-endo conformation and provides protection from NMIA modification (42). The high level of NMIA reactivity outside of the seed region indicates that although argonaute specifically interacts with the seed region, it also allows flexibility in the rest of the miRNA, consistent with the tolerance for target mismatches and bulges outside of the seed sequence. The NMIA protection of the seed region that we observe in miRNPs provides direct evidence that eukaryotic argonaute proteins promote structural restraint and potential pre-organization of the seed region to promote target recognition. Consistent with this model, we show by chemical protection assays that complementarity between

the seed region and a target sequence is required for the propagation of microRNA binding to its target.

DMS and kethoxal probing experiments demonstrate that most of the bases of miRNAs in miRNPs are solvent-exposed and therefore accessible for base pairing with substrates. Surprisingly, protection from chemical modifications is observed within the first four bases in native *C. elegans* miRNPs. Nucleotides outside of the seed region are solvent accessible and strongly reactive with both DMS and kethoxal. Reduced chemical reactivity at positions 2–4, most dramatically seen for G2 and G4 of mir-58 (Figure 3B) which are predicted to be accessible for base pairing with substrates, may indicate protein protection or tight nucleotide stacking within the seed region, making the chemical modification of the bases less efficient.

The association of miRNP complexes with miRNAs results in an altered conformation of specific seed region bases. A structural property of the seed region that has been revealed in our experiments is an unusual position-dependent DMS reactivity with uracils at position 6. Several DMS-reactive uracils were observed in the prokaryotic ribosome probed at pH 8.5 (48), and this induced DMS reactivity was attributed to a perturbed pKa resulting in a keto-to-enol tautomeric shift. DMS typically modifies N1 of adenosine and N3 of cytosine. The N3 of uracil is protonated and therefore excluded from DMS reactivity. Thus, it is logical to suggest that the unusual DMS reactivity of U6 in miRNAs from *C. elegans* may be due to enolate tautomerization induced by contacts with the argonaute protein. The function of an enol tautomer at miRNA position 6 is presently unclear, but an intriguing possibility is that altered base pairing potential, and an expanded breadth of functional targets of miRNAs with uracils at position 6, may be consequences of enolate tautomer stabilization.

The X-ray structure of a eubacterial DNA-guided endoribonuclease from *Thermus thermophilus*, co-crystallized with a DNA guide strand, provides important insights into how a protein can carry a DNA guide strand such that it can base pair to a target (49–52). However, the relevance of that structure to miRNP structure is potentially limited; the crystal structure is loaded with DNA (*T. thermophilus* has no microRNAs), and the primary amino acid sequence alignment to *C. elegans* ALG-1 in BLAST searches is limited to 26% identity in a 135 amino acid stretch only in the PIWI domain. The eubacterial complex, predicts that residues 2–6 are exposed to solvent while our data indicate that in the eukaryotic complex nucleotides 5–8 are readily exposed to solvent. The eubacterial complex shows continuous stacking of nucleotides 2–11, while our data show structural organization of the backbone limited to nucleotides 1–7. These distinct differences point to the need for detailed study of an authentic eukaryotic argonaute complex in order to understand the structural basis of the miRNA seed region function.

The structural constraint of the seed region, demonstrated in our experiments, supports a previously proposed model in which the seed region is held in an optimal conformation to nucleate miRNA binding to mRNA targets (5,21,22). This model suggests that the seed region is held by argonaute in a way that is preformed for binding to its target partner. Given the importance of the seed region in target discrimination, the constraints put on this region by the argonaute protein in the miRNP must confer improved target recognition over the flexible regions of the rest of the miRNA. The absence of structural constraint placed on bases outside of the seed region could then allow for further interaction of the miRNA with the target RNA after initial binding of the seed region. Since the chemical accessibility of only the seed region is protected in miRNPs compared to naked miRNAs, it is likely that miRNA residues outside of the seed region do not make contact with argonaute proteins. These unrestrained sequences themselves, even though they are 14-nt long, are not capable of initiating base pairing to target sequences.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Many thanks to Harry Noller for many helpful discussions and critical reading of the article. Thanks to Dmitri Ermolenko, Ian MacRae, Manuel Ares, Jr., Martin Laurberg, Andrei Korostelev and Tim Nilsen for helpful discussions. Thanks to Craig Mello for strain WM84 and Ronald Plasterk for anti-TSN-1 antibody.

FUNDING

Funding for open access charge: National Institutes of Health (grant number R01 GM-0161646 to A.M.Z.).

Conflict of interest statement. None declared.

REFERENCES

1. Wightman,B., Ha,I. and Ruvkun,G. (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*, **75**, 855–862.
2. Reinhart,B.J., Slack,F.J., Basson,M., Pasquinelli,A.E., Bettinger,J.C., Rougvie,A.E., Horvitz,H.R. and Ruvkun,G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. *Nature*, **403**, 901–906.
3. Lee,R.C., Feinbaum,R.L. and Ambros,V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, **75**, 843–854.
4. Griffiths-Jones,S., Grocock,R.J., van Dongen,S., Bateman,A. and Enright,A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.*, **34**, D140–D144.
5. Bartel,D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.
6. Selbach,M., Schwanhausser,B., Thierfelder,N., Fang,Z., Khanin,R. and Rajewsky,N. (2008) Widespread changes in protein synthesis induced by microRNAs. *Nature*, **455**, 58–63.
7. Baek,D., Villen,J., Shin,C., Camargo,F.D., Gygi,S.P. and Bartel,D.P. (2008) The impact of microRNAs on protein output. *Nature*, **455**, 64–71.
8. Guo,H., Ingolia,N.T., Weissman,J.S. and Bartel,D.P. (2010) Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature*, **466**, 835–840.

9. Wu,L., Fan,J. and Belasco,J.G. (2006) MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl Acad. Sci. USA*, **103**, 4034–4039.
10. Giraldez,A.J., Mishima,Y., Rihel,J., Grocock,R.J., Van Dongen,S., Inoue,K., Enright,A.J. and Schier,A.F. (2006) Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science*, **312**, 75–79.
11. Liu,J., Rivas,F.V., Wohlschlegel,J., Yates,J.R. 3rd, Parker,R. and Hannon,G.J. (2005) A role for the P-body component GW182 in microRNA function. *Nat. Cell Biol.*, **7**, 1261–1266.
12. Lim,L.P., Lau,N.C., Garrett-Engele,P., Grimson,A., Schelter,J.M., Castle,J., Bartel,D.P., Linsley,P.S. and Johnson,J.M. (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature*, **433**, 769–773.
13. Bagga,S., Bracht,J., Hunter,S., Massirer,K., Holtz,J., Eachus,R. and Pasquinelli,A.E. (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell*, **122**, 553–563.
14. Petersen,C.P., Bordeleau,M.E., Pelletier,J. and Sharp,P.A. (2006) Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell*, **21**, 533–542.
15. Pillai,R.S., Bhattacharyya,S.N., Artus,C.G., Zoller,T., Cougot,N., Basyuk,E., Bertrand,E. and Filipowicz,W. (2005) Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science*, **309**, 1573–1576.
16. Humphreys,D.T., Westman,B.J., Martin,D.I. and Preiss,T. (2005) MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc. Natl Acad. Sci. USA*, **102**, 16961–16966.
17. Olsen,P.H. and Ambros,V. (1999) The lin-4 regulatory RNA controls developmental timing in Caenorhabditis elegans by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.*, **216**, 671–680.
18. Enright,A.J., John,B., Gaul,U., Tuschl,T., Sander,C. and Marks,D.S. (2003) MicroRNA targets in Drosophila. *Genome Biol.*, **5**, R1.
19. Stark,A., Brennecke,J., Russell,R.B. and Cohen,S.M. (2003) Identification of drosophila MicroRNA targets. *PLoS Biol.*, **1**, E60.
20. Lewis,B.P., Shih,I.H., Jones-Rhoades,M.W., Bartel,D.P. and Burge,C.B. (2003) Prediction of mammalian microRNA targets. *Cell*, **115**, 787–798.
21. Parker,J.S., Parizotto,E.A., Wang,M., Roe,S.M. and Barford,D. (2009) Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Mol. Cell*, **33**, 204–214.
22. Haley,B. and Zamore,P.D. (2004) Kinetic analysis of the RNAi enzyme complex. *Nat. Struct. Mol. Biol.*, **11**, 599–606.
23. Brennecke,J., Stark,A., Russell,R.B. and Cohen,S.M. (2005) Principles of microRNA-target recognition. *PLoS Biol.*, **3**, e85.
24. Kloosterman,W.P., Wienholds,E., Ketting,R.F. and Plasterk,R.H. (2004) Substrate requirements for let-7 function in the developing zebrafish embryo. *Nucleic Acids Res.*, **32**, 6284–6291.
25. Doench,J.G. and Sharp,P.A. (2004) Specificity of microRNA target selection in translational repression. *Genes Dev.*, **18**, 504–511.
26. Krek,A., Grun,D., Poy,M.N., Wolf,R., Rosenberg,L., Epstein,E.J., MacMenamin,P., da Piedade,I., Gunsalus,K.C., Stoffel,M. *et al.* (2005) Combinatorial microRNA target predictions. *Nat. Genet.*, **37**, 495–500.
27. Lai,E.C. (2004) Predicting and validating microRNA targets. *Genome Biol.*, **5**, 115.
28. Lai,E.C. (2002) Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. *Nat. Genet.*, **30**, 363–364.
29. Jinek,M. and Doudna,J.A. (2009) A three-dimensional view of the molecular machinery of RNA interference. *Nature*, **457**, 405–412.
30. Farazi,T.A., Juranek,S.A. and Tuschl,T. (2008) The growing catalog of small RNAs and their association with distinct Argonaute/Piwi family members. *Development*, **135**, 1201–1214.
31. Frank,F., Sonenberg,N. and Nagar,B. (2010) Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature*, **465**, 818–822.
32. Boland,A., Tritschler,F., Heimstadt,S., Izaurralde,E. and Weichenrieder,O. (2010) Crystal structure and ligand binding of the MID domain of a eukaryotic Argonaute protein. *EMBO Rep.*, **11**, 522–527.
33. Gu,S.G., Pak,J., Barberan-Soler,S., Ali,M., Fire,A. and Zahler,A.M. (2007) Distinct ribonucleoprotein reservoirs for microRNA and siRNA populations in C. elegans. *RNA*, **13**, 1492–1504.
34. Caputi,M., Mayeda,A., Krainer,A.R. and Zahler,A.M. (1999) hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing. *EMBO J.*, **18**, 4060–4067.
35. Zhang,L., Ding,L., Cheung,T.H., Dong,M.Q., Chen,J., Sewell,A.K., Liu,X., Yates,J.R. 3rd and Han,M. (2007) Systematic identification of C. elegans miRISC proteins, miRNAs, and mRNA targets by their interactions with GW182 proteins AIN-1 and AIN-2. *Mol. Cell*, **28**, 598–613.
36. Moazed,D., Stern,S. and Noller,H.F. (1986) Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension. *J. Mol. Biol.*, **187**, 399–416.
37. Maroney,P.A., Chamnongpol,S., Souret,F. and Nilsen,T.W. (2007) A rapid, quantitative assay for direct detection of microRNAs and other small RNAs using splinted ligation. *RNA*, **13**, 930–936.
38. Das,R., Laederach,A., Pearlman,S.M., Herschlag,D. and Altman,R.B. (2005) SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA*, **11**, 344–354.
39. Rasband,W.S. (2009) *US National Institutes of Health, Bethesda, Maryland, USA*. http://rsb.info.nih.gov/ij/ (December 2010, date last accessed).
40. Grishok,A., Pasquinelli,A.E., Conte,D., Li,N., Parrish,S., Ha,I., Baillie,D.L., Fire,A., Ruvkun,G. and Mello,C.C. (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing. *Cell*, **106**, 23–34.
41. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
42. Wilkinson,K.A., Merino,E.J. and Weeks,K.M. (2006) Selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.*, **1**, 1610–1616.
43. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.
44. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
45. Peattie,D.A. and Gilbert,W. (1980) Chemical probes for higher order structure in RNA. *Proc. Natl Acad. Sci. USA*, **77**, 4679–4682.
46. Parker,J.S., Roe,S.M. and Barford,D. (2005) Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature*, **434**, 663–666.
47. Ma,J.B., Yuan,Y.R., Meister,G., Pei,Y., Tuschl,T. and Patel,D.J. (2005) Structural basis for 5′-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature*, **434**, 666–670.
48. Bayfield,M.A., Dahlberg,A.E., Schulmeister,U., Dorner,S. and Barta,A. (2001) A conformational change in the ribosomal peptidyl transferase center upon active/inactive transition. *Proc. Natl Acad. Sci. USA*, **98**, 10096–10101.
49. Hall,T.M. (2005) Structure and function of argonaute proteins. *Structure*, **13**, 1403–1408.
50. Wang,Y., Juranek,S., Li,H., Sheng,G., Wardle,G.S., Tuschl,T. and Patel,D.J. (2009) Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature*, **461**, 754–761.
51. Wang,Y., Sheng,G., Juranek,S., Tuschl,T. and Patel,D.J. (2008) Structure of the guide-strand-containing argonaute silencing complex. *Nature*, **456**, 209–213.
52. Wang,Y., Juranek,S., Li,H., Sheng,G., Tuschl,T. and Patel,D.J. (2008) Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*, **456**, 921–926.