



Molecular dynamics simulations of Ago silencing complexes reveal a large repertoire of admissible 'seed-less' targets

Zhen Xia^{1,2}, Peter Clark³, Tien Huynh¹, Phillippe Loher³, Yue Zhao³, Huang-Wen Chen^{3#}, Isidore Rigoutsos^{3,4*} & Ruhong Zhou^{1,5*}

¹Computational Biology Center, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, ²Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX 78712, ³Computational Medicine Center, Thomas Jefferson University, Philadelphia, PA 19107, ⁴Department of Pathology, Anatomy and Cell Biology; Department of Cancer Biology; Department of Biochemistry and Molecular Biology; Thomas Jefferson University, Philadelphia, PA 19107, ⁵Department of Chemistry, Columbia University, New York, NY 10027.

SUBJECT AREAS:

RNAI

NON-CODING RNA'S

SMALL RNA'S

GENE EXPRESSION

Received
7 March 2012

Accepted
20 July 2012

Published
10 August 2012

Correspondence and requests for materials should be addressed to I.R. (isidore.rigoutsos@jefferson.edu) or R.Z. (ruhongz@us.ibm.com)

* These authors contributed equally.

Current address: Bloomberg L.P., New York, NY 10022.

To better understand the recognition mechanism of RISC and the repertoire of guide-target interactions we introduced G:U wobbles and mismatches at various positions of the microRNA (miRNA) 'seed' region and performed all-atom molecular dynamics simulations of the resulting Ago-miRNA:mRNA ternary complexes. Our simulations reveal that many modifications, including combinations of multiple G:U wobbles and mismatches in the seed region, are admissible and result in only minor structural fluctuations that do not affect overall complex stability. These results are further supported by analyses of HITS-CLIP data. Lastly, introduction of disruptive mutations revealed a bending motion of the PAZ domain along the L1/L2 'hinge' and a subsequent opening of the nucleic-acid-binding channel. Our findings suggest that the spectrum of a miRNA's admissible targets is different from what is currently anticipated by the canonical seed-model. Moreover, they provide a likely explanation for the previously reported sequence-dependent regulation of unintended targeting by siRNAs.

MiRNAs are short RNAs, approximately ~22 nucleotides (nts) in length, which post-transcriptionally regulate their protein-coding targets in a sequence-dependent manner¹. The typical transcribed precursor molecule (pri-miRNA) assumes a double-stranded RNA (dsRNA) conformation with a characteristic hairpin-like structure; the latter is pre-processed in the nucleus by the "microprocessor" complex into a pre-miRNA before being shuttled into the cytoplasm by XPO5². Mirtrons represent an exception to this rule: their pre-miRNAs are directly generated from *introns* during the splicing of nascent mRNAs³. In the cytoplasm, the pre-miRNA hairpin is cleaved by the DICER endonuclease to form dsRNA, 20–25 nts in length, with 3' overhangs; one of the two strands, the miRNA, is incorporated into the RISC, also known as the miRNA ribonucleoprotein complex (miRNP), where the interaction with the target mRNA takes place¹. This interaction typically results in degradation of the target mRNA or inhibition of its translation by the ribosome⁴. Originally believed to form heteroduplexes mainly with the 3' untranslated region (3'UTR) of the target, miRNAs have since been shown to also target protein-coding regions and 5'UTRs^{5–12}.

MiRNAs have been shown to be involved in many fundamental processes that include developmental timing^{13–16}, the induction of organ asymmetry¹⁷, tumor suppression and oncogenic activity^{18–20}, invasion and metastasis²¹, modulation of embryonic stem cell differentiation⁹, neurodegeneration²² etc. Moreover, extensive work has revealed tissue- and cellular-state-dependent miRNA profiles in cancers^{23–30}, cardiovascular disease^{31,32}, immunity³³, Alzheimer's^{34,35}, Tourette's syndrome³⁶, schizophrenia³⁷, and others. Given their involvement in such diverse contexts, and their ever-increasing numbers³⁸, understanding the identity and cardinality of a given miRNA's targets represents an important endeavor.

Ever since the first report on the *lin-4:lin-14* heteroduplex^{14,16}, it was clear that the 5' region of a miRNA played a central role in the recognition of its target. This region, spanning positions 2–7 from the miRNA's 5' end, was originally referred to as the 'core element' but eventually became known as the 'seed.' Its apparent importance has



been a central component of many computer-based miRNA target prediction schemes^{39–46} whereby the presence of the reverse complement of a miRNA's seed sequence is used as a filter before generating lists of candidate mRNA targets. Methods that do not rely on such filtering schemes have also been in use^{6,47}. All computational attempts to tackle the problem have met with various degrees of success and for all practical purposes the problem remains open^{48,49}.

Some of the very early work^{13,16} indicated that formation of functioning heteroduplexes did not require a strict reverse-complementarity relationship between the miRNA-seed sequence and its target. Since then, evidence has continued to accumulate steadily in support of such an 'expanded' interaction mode^{9,50–63}, in turn suggesting a potential repertoire of targets for a given miRNA that is more diverse than the 'seed'-reverse-complementarity constraint might suggest. We revisit this very question with the help of molecular dynamics (MD) simulations and the recently released crystal structure of ternary complexes of eubacterial *Thermus thermophilus* Ago (*TtAgo*)^{64,65}.

Results

TtAgo is considered an appropriate model for studying the properties of Ago complexes thanks to the high structural and functional similarities with the eukaryotic Ago families. For this study, we introduced selected modifications to the currently available Ago-DNA:mRNA co-crystal structure. The collection of heteroduplexes

to simulate was informed by previously reported non-canonical examples^{6,9,13,14,16,52,62} and includes heteroduplexes with a) G:U wobbles in the seed in conjunction with adjacent Watson-Crick pairs; b) G:U wobbles in the seed but *without* adjacent Watson-Crick pairs; c) a single bulge on the *target* side (mRNA) at each of several different seed locations; and, d) a single bulge on the *guide* side (miRNA) at each of several different seed locations. To build our Ago-miRNA:mRNA ternary complexes we replaced the bases of the guide DNA by the corresponding RNA (miRNA). Each MD simulation spanned a minimum of 100 ns, generating a trajectory that is sufficiently long for the current analysis of the mRNA recognition dynamics and also for observing any potential conformational changes (discussed further below).

The Ago-complex is stable in the presence of multiple seed-region G:U wobbles. In earlier *in vivo* studies, the impact of G:U wobbles in the seed region (positions 2–7) was examined in *D. melanogaster*⁶⁶ and in *C. elegans*⁵² and led to different conclusions. The fruit-fly study examined the impact of one, two and three G:U wobbles and concluded that “[...] a G:U wobble in the seed region is always detrimental [...]”⁶⁶. On the other hand, the worm study examined the impact of one and two wobbles in the seed region and found the mutants to still be functionally regulated by the targeting miRNA lsy-6⁵². In addition to these two studies, luciferase assays were used to

Table 1 | Sequences of the 11-nt guide miRNA and target mRNA heteroduplex used in the simulation. The nucleotides of the miRNA's seed and their bond-partners are shown in gray background. Mutated nucleotides are indicated in red. In all cases, the first row of the heteroduplex shows the target mRNA whereas the second row shows the guide miRNA. All shown numbering is with regard to the targeting miRNA. 5' and 3' are also indicated

3 and 4 G:U wobbles in seed	Mutant #1	Mutant #2	Wild-Type
	5'–UCACUACUUUG–3'	5'–UCACUAAUUUG–3'	5'–UCACUACCUUG–3' (mRNA)
	3'–GAUGAUGGGGU–5'	3'–GAUGAUGGGGU–5'	3'–GAUGAUGGAGU–5' (miRNA)
3, 4 and 5 G:U wobbles in seed with no adjacent Watson-Crick pairs	Mutant #3	Mutant #4	Mutant #5
	5'–UCACUACUUUG–3'	5'–UCACUAAUUUG–3'	5'–UCACUAAUUUG–3'
	3'–GAGUAUGGGGU–5'	3'–GAGUAUGGGGU–5'	3'–GAGUCUGGGGU–5'
bulge on mRNA side at different seed positions	Mutant #6	Mutant #7	Mutant #8
	5'–UCACUA CUUUG–3'	5'–UCACUAC UUUG–3'	5'–UCACUACU UUG–3'
	3'–GAGUAU–GGGU–5'	3'–GAGUAUG–GGGU–5'	3'–GAGUAUGG–GGU–5'
bulge on miRNA side at different seed positions	Mutant #9	Mutant #10	Mutant #11
	5'–UCACUA–CUUUG–3'	5'–UCACUAC–UUUG–3'	5'–UCACUACU–UUG–3'
	3'–GAGUAU–GGGU–5'	3'–GAGUAUG–GGGU–5'	3'–GAGUAUGG–GGU–5'
	210987 54321 6	2109876 4321 5	21098765 321 4

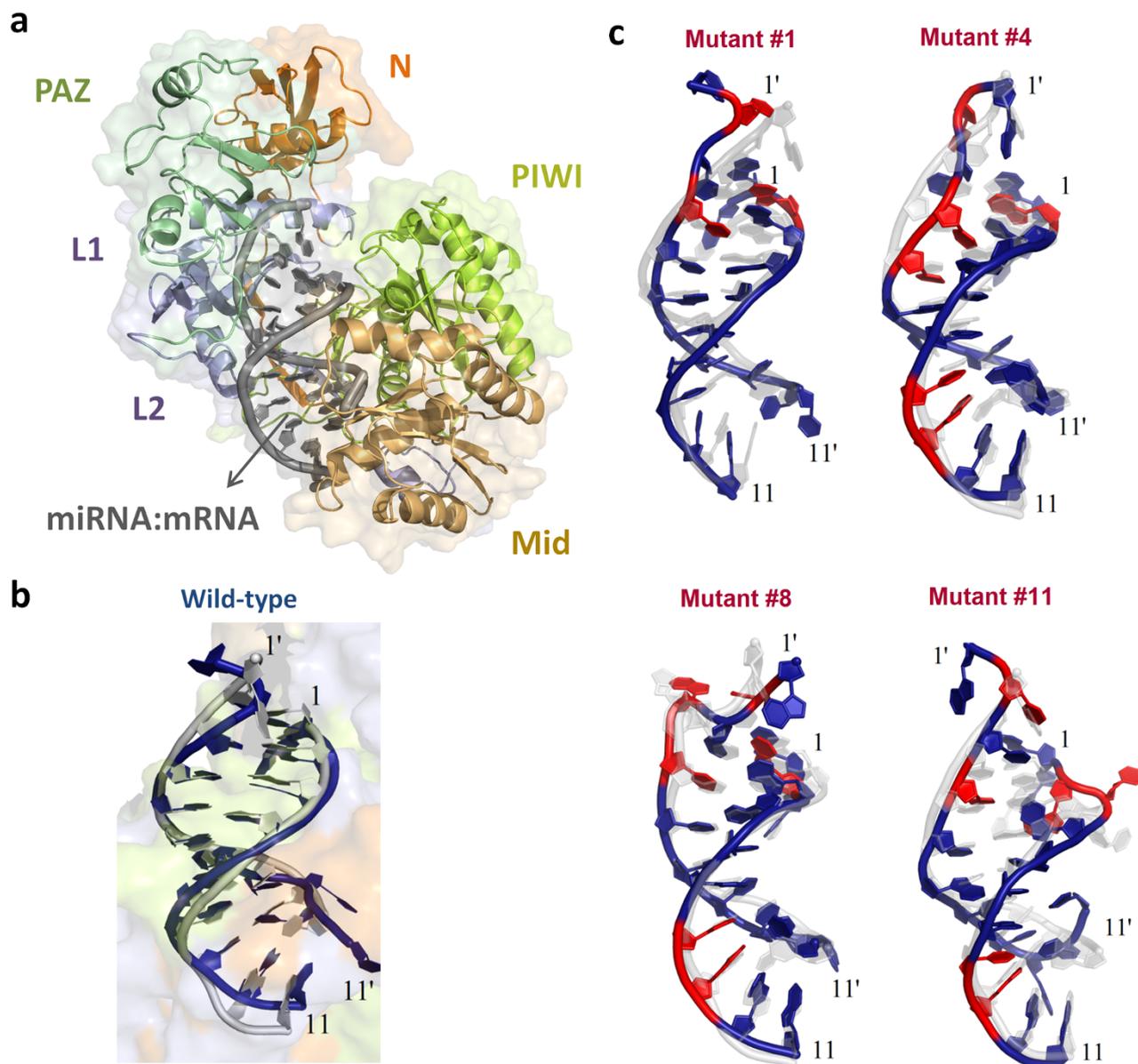


Figure 1 | Structural views of an 11-nt guide (miRNA) and target (mRNA) heteroduplex for the wild-type and mutants during the simulation. (a) The overall structure of *TtAgo*-miRNA:mRNA complexes. The Ago protein is rendered as cartoon and molecular surface, and each of its domains is colored differently. The miRNA:mRNA heteroduplex is presented as cartoon and shown in gray. (b) The structure of the guide-target heteroduplex for the wild-type during the 100-ns molecular dynamics simulation. The conformational change is shown by superimposing the final snapshot (shown in blue) to the starting native structure (shown in gray). The backbone of the heteroduplex is rendered as cartoon; the ribose and the base are represented as plates. (c) The structure of selected mutants in simulations. The conformational changes of the miRNA:mRNA heteroduplex are shown by superimposing the final snapshot (mutated sites are indicated in red) to the starting native structure (colored in light gray) with the ribose and the base shown as plates. Primed (') numbers indicates bases that belong to the target strand.

demonstrate the validity of functional heteroduplexes involving as many as five G:U wobbles in the seed⁶. For our studies, we created a first mutant with three G:U wobbles at positions 2, 3 and 4 of the seed region (Mutant #1 in Table 1). In a second experiment, we added a fourth G:U wobble at position 5 of the seed region (Mutant #2). Our analysis shows that in both configurations the resulting structures are stable and their Ago backbones compared to the backbone of the native Ago ternary complex exhibit small RMSD values ($\sim 2\text{--}3\text{\AA}$) (see Fig. 1c, Fig. 2a, and Supp. Fig. S1). In the mutants, study of position 6 or 7 of the seed region reveals that the conformations of Watson-Crick pairs remain intact between guide strand and the target strand (Supp. Fig. S2).

The Ago-complex is stable in the presence of multiple seed-region G:U wobbles and no compensating Watson-Crick pairs immediately adjacent to the seed. As can be seen from Figure 1b and Table 1, the heteroduplexes of Mutants #1 and #2 contain two Watson-Crick pairs immediately past the seed region, at positions 8 and 9. In order to determine the extent to which these two base pairs play a compensatory role that contributes to the stability of the complex we removed both and repeated the previous simulations (Mutant #1 \rightarrow Mutant #3, Mutant #2 \rightarrow Mutant #4). The two resulting heteroduplexes, were they stable, would rely primarily on coupling that spans the seed region and is rooted in the presence of three- (case of Mutant #3) and four- (case of Mutant #4) G:U wobbles

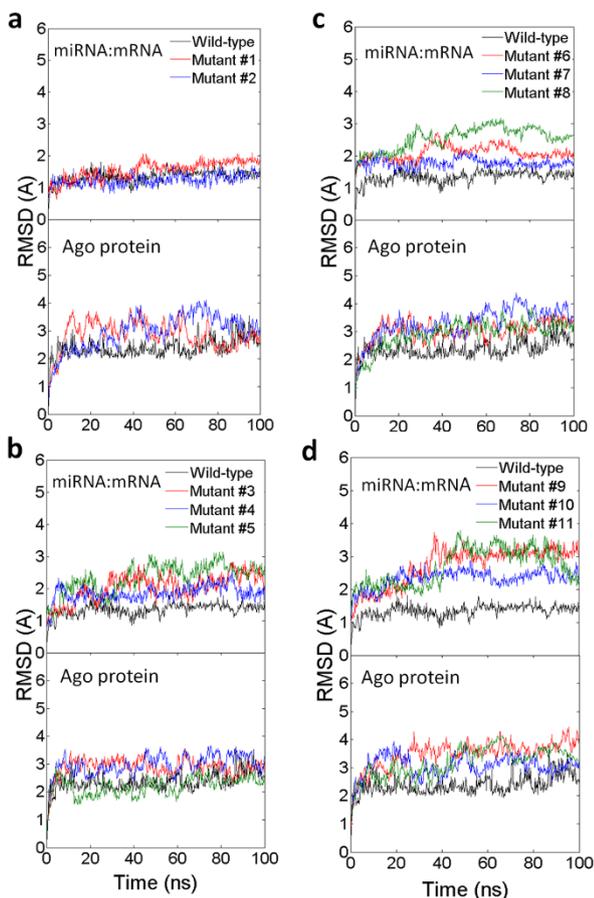


Figure 2 | Comparison of the structural variations during the simulations for the wild-type and the eleven mutants. (a) Mutants with G:U wobbles in the seed and adjacent Watson-Crick pairs; (b) mutants with G:U wobbles in the seed and with no adjacent Watson-Crick pairs; (c) mutants with one bulge on the target (mRNA) side at different seed positions; (d) mutants with one bulge on the guide (miRNA) side at different seed positions. The plot shows the RMSD values of the miRNA:mRNA heteroduplex (subplot on top) and Ago protein (subplot at bottom) in the ternary complexes. The RMSD values are calculated by comparing each snapshot to the backbone of the starting crystal structures during the simulations in the complexes (11-nt). The results are obtained from NPT ensemble simulations ($T=310$ K, $P=1$ atm) with the simulation time of 100 ns.

respectively. Our simulations show that these mutants are indeed stable: the RMSD values of the Ago backbones from the wild-type remain low and reach a plateau of ~ 2.5 Å after only ~ 40 ns (Fig. 1c, Fig. 2b, and Supp. Fig. S1). This indicates that the Watson-Crick pairs already present in the seed region (at positions 6 and 7) are sufficient for maintaining the overall stability of the heteroduplex – see also base pair distances in Supp. Fig. S2

The Ago-complex is stable in the presence of only partial seed-region coupling and no compensating Watson-Crick pairs immediately adjacent to it. The observed stability of the ternary complex in the presence of multiple G:U wobbles and without any compensating Watson-Crick pairs adjacent to the seed prompted us to also examine a somewhat extreme situation. In particular, we mutated the miRNA's adenosine at position 7 of the seed to a cytosine, thus "breaking" the base pairing at that location (Mutant #5) – shown in cyan in Table 1. The resulting heteroduplex, if realized, would be brought about by only five base pairs in the seed region, with four of them being G:U wobbles, and without any

compensating Watson-Crick pairs beyond it. Interestingly, and somewhat surprisingly, we found that this arrangement also leads to a stable structure. In fact, the resulting RMSD is only slightly larger than the wild-type arrangement, remaining well below 3 Å for the length of the simulation (green curve of Fig. 2b).

The Ago-complex is stable in the presence of a seed-region bulge on the messenger-RNA-side of the heteroduplex. For let-7, the second miRNA ever reported, it was shown that it regulates the heterochronic gene *lin-41* by binding to two locations of *lin-41*'s 3'UTR¹⁵. These two target locations, referred to as LCS1 and LCS2, were later demonstrated *in vivo* to be simultaneously required for *lin-41*'s regulation⁶². For the purpose of this discussion, the heteroduplex formed between let-7 and LCS1 contains a bulge on the *lin-41* (mRNA) side between positions 4 and 5 of the seed region^{15,62}. Subsequently, examples of functioning heteroduplexes comprising messenger-RNA-side bulges in the seed region were reported and validated for mouse *Oct4* (between seed positions 4 and 5) and mouse *Sox2* (between seed positions 5 and 6), and concomitant physiological effects were shown for these heteroduplexes⁹. We thus sought to investigate the impact on the stability of the Ago complex that a bulge located on the target-side (mRNA) might have as a function of the bulge's actual location within the seed. Notably, in these experiments we maintained the three G:U wobbles that were previously introduced in the seed region and removed the two Watson-Crick pairs that were originally adjacent to the seed at positions 8 and 9; arguably this generates a rather demanding context for the complex stability in our simulation study. We investigated three bulge placements in the seed region: between seed positions 6 and 5 (Mutant #6), between seed positions 5 and 4 (Mutant #7), and, finally, between seed positions 4 and 3 (Mutant #8). We found that all three placements of the bulge generate stable structures, with a slight dependence on the actual position of the bulge within the seed's span. The resulting RMSD from the wild-type arrangement is small and ranges between 2 and 3 Å (Fig. 2c). These findings demonstrate that more extreme and challenging scenarios than the one reported very recently⁶³ (namely, the presence of heteroduplexes containing a bulge between positions 5 and 6 of the mRNA) are also possible.

The Ago-complex stability is affected minimally by a miRNA-side bulge in the seed region. In one of the very early publications on miRNA-driven RNA interference¹³ it was shown that bulged *lin-4:lin-14* heteroduplexes, with the bulge being on the side of the targeting miRNA, at position 6 of the seed, were functional and sufficient for *lin-14* temporal gradient formation in *C. elegans*. More recently, similarly bulged interactions were shown for mouse miRNA:mRNA heteroduplexes^{6,9}. In this group of experiments we investigated the impact on the stability of the Ago-complex of a single bulge that is increasingly closer to the 5' end of the miRNA at seed positions 6, 5 and 4 (Mutants #9, #10 and #11, respectively). Just as before, we maintained in all experiments the three G:U wobbles introduced in the seed region thus creating an extreme context for our simulation study. We also preserved a single Watson-Crick base pair immediately adjacent to the seed, at position 8. In all three cases, the resulting RMSD values were similar to what we observed prior to having introduced the miRNA-side bulge (Fig. 2d). Also, there was some distortion of individual base pairs when the bulge was placed at position 4 of the seed (Mutant #11) – see Fig. 1c. Bulge placements at seed positions 6 and 4 (Mutant #9 and Mutant #11, respectively) led to slightly larger RMSD values (~ 4 Å). Placement of the bulge at position 5 (Mutant #10) exhibited smaller structural deviations from the native structure for both the Ago protein and the RNA heteroduplex compared to the other two placements (Fig. 2d).

Disruptive mutations lead to a large bending motion of PAZ domain along the L1/L2 'hinge' and a subsequent opening of the

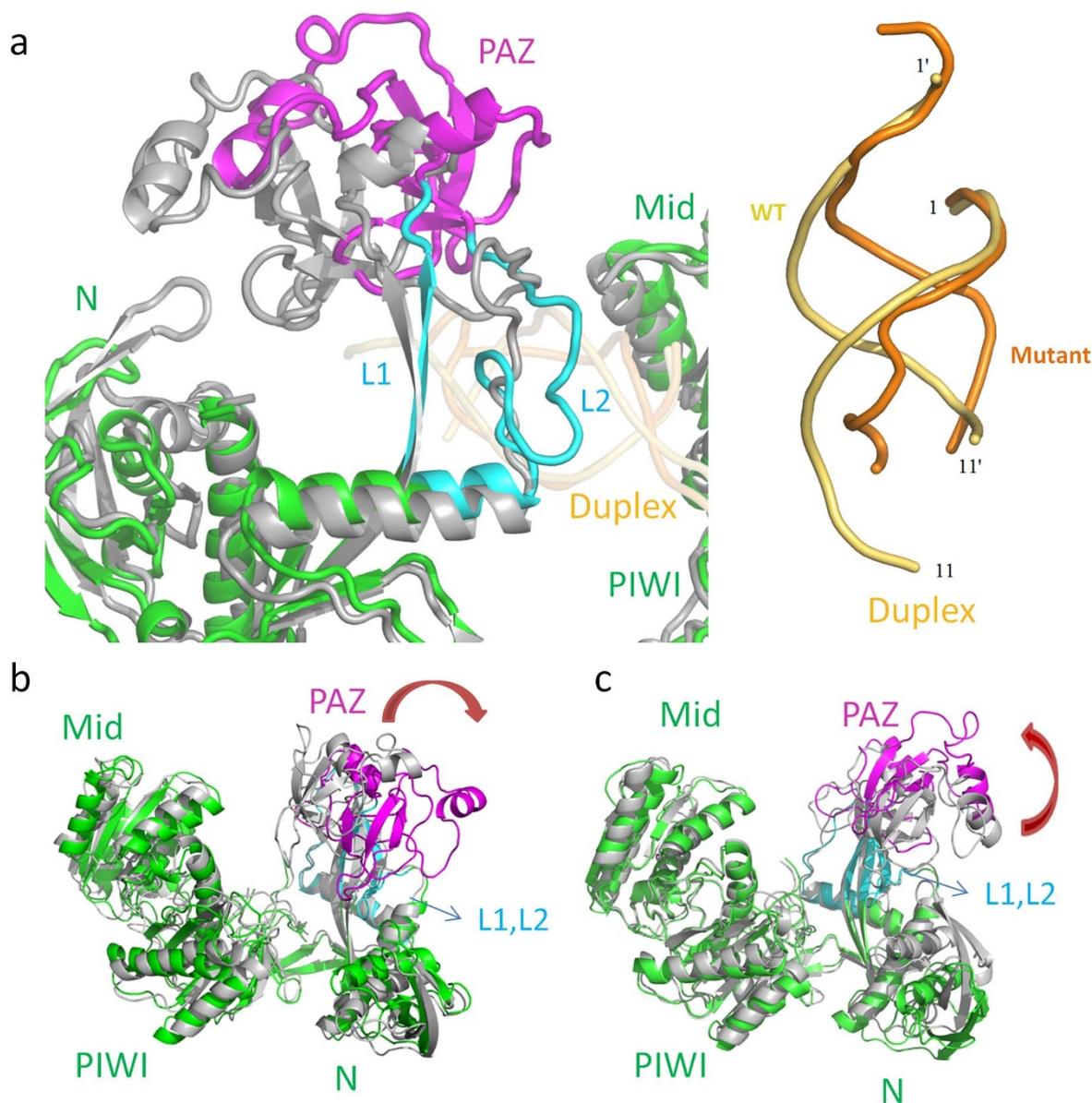


Figure 3 | Structural views of the guide-target heteroduplex distortion and the domain motions of Ago protein with extreme disruptive mutations. (a) The disassociation of the “hinge-like” L1/L2 segment and the nucleic acid heteroduplex in Mutant #15 (four G-C disruptions). The final conformation and the starting structure are superimposed. The nucleic acid duplex is colored in orange (mutant) and yellow (wild-type), and the Ago protein is colored in green (mutant) and gray (wild-type), respectively. The PAZ domain is shown in magenta and the L1/L2 segment is shown in cyan for the mutant. The Ago protein is represented as cartoon with the domain name labeled, and the backbone of the nucleic acids is shown as tube. (b) and (c) Structural view of the domain motions in the four-G-C-disruptions mutant. Two structures (one colored light gray and the other colored green) are picked from a 100-ns trajectory for each by the principal component analysis (PCA) and the domain motion analysis. The 1st principal component (b) and the 2nd principal component (c) are shown. The PAZ domain and the L1/L2 segment are shown using different colors. The red arrows indicate the motions of the PAZ domain.

nucleic-acid-binding channel. We also examined whether our 100 ns simulations are long enough to capture large Ago-complex motions, as would be the case when attempting to simulate unsuitable, disruptive mutations. In order to address this question, we introduced several G→C mutations in the seed region aimed at “disrupting” the structure of the complex. Each G→C mutation broke a triple bond and led to non-bonded bases between the guide (miRNA) strand and the target (mRNA) strand. The first mutation we introduced broke the G-C bond at position 8 immediately adjacent to the seed (Mutant #12). Three more mutations gradually increased the number of non-bonded bases *inside* the seed region from one (Mutant #13) to two (Mutant #14) to three (Mutant #15), while maintaining the mutation at position 8 – see Methods and Supp. Table S1 for details. Not surprisingly, as the number of non-bonded bases *increased*, the stability of the

miRNA-mRNA heteroduplex *decreased* (Supp. Fig. S3). The average RMSD of each nucleotide in the mRNA strand increased significantly and in proportion to the number of mismatches (Supp. Fig. S4). The comparable structural stability of the wild-type and mutants with a single non-bonded base supports earlier experimental work showing that a single nucleotide mismatch at the seed region only slightly reduces the cleavage activity of Ago complexes⁶⁵. On the contrary, for Mutant #15 (which contains four G-C disruptions) a mere ~10 ns of simulation sufficed to disrupt most of the base pairing and base stacking, even for the canonical Watson-Crick base pairs (Supp. Fig. S5). The severe distortion of the backbone in the guide-target duplex caused the overall “decoupling” of the miRNA-mRNA heteroduplex, which in turn indicates that nucleation at the seed region cannot be achieved in the mutant with the four G-C-disruptions. The final snapshot of the wild type



Table 2 | Probability estimates indicating the support by HITS-CLIP data of the canonical model and of an expanded model where one or more G:U wobbles are permitted in the seed region. Data are shown for “Brain A” (Ago antibody 2A8) and “Brain D” (Ago antibody 7G1-1*), two of five previously reported mouse brain datasets⁵⁵. Results for the remaining three brain datasets are shown in the Supplement. As can be seen, HITS-CLIP data indeed support miRNA:mRNA interactions where the nucleation in the seed region is provided by one or more G:U wobbles. In some instances, HITS-CLIP provides stronger support for the expanded model than for the canonical one. The miRNAs in each case are listed in order of decreasing abundance in the respective ~110 kDa set

BRAIN A				
miRNA ID	Canonical Model	Expanded Model	Canonical Model P-value	Expanded Model P-value
mmu-miR-30d-5p	GTTTAC	GTTRY	1.81E-10	1.24E-07
mmu-miR-30a-5p	GTTTAC	GTTRY	1.78E-10	1.29E-07
mmu-miR-30e-5p	GTTTAC	GTTRY	1.77E-10	1.20E-07
mmu-miR-27a-3p	CTGTGA	YTGTR	4.64E-04	5.60E-06
mmu-miR-708-5p	GCTCCT	GYYYT	9.33E-01	5.06E-01
mmu-let-7b-5p	TACCTC	TRYTY	0.00E+00	0.00E+00
mmu-miR-26a-5p	ACTTGA	RYTGR	2.59E-04	7.64E-06
mmu-miR-21-5p	TACGCT	TRRGT	2.38E-03	3.03E-01
mmu-miR-374-5p	ATTATA	RTTRR	0.00E+00	1.07E-13
mmu-miR-153-3p	TATGCA	TRGYR	5.99E-08	6.59E-09
mmu-let-7i-5p	TACCTC	TRYTY	0.00E+00	0.00E+00
mmu-miR-101a-3p	TACTGT	TRYTG	0.00E+00	0.00E+00
mmu-miR-27b-3p	CTGTGA	YTGTR	5.80E-04	1.03E-05
mmu-miR-222-3p	TGTAGC	TGRGY	4.18E-04	4.17E-07
mmu-miR-204-5p	AAGGGA	RRGGR	9.58E-01	1.00E+00
mmu-miR-31-5p	CTTGCC	YTYGY	3.21E-03	1.02E-14
mmu-miR-9-5p	CCAAAG	YRRRG	2.21E-11	1.00E+00
mmu-miR-101b-3p	TACTGT	TRYTG	0.00E+00	0.00E+00
mmu-miR-449a-5p	ACTGCC	RYTYGY	1.20E-03	2.43E-13
mmu-miR-217-5p	TGCAGT	TGRGT	1.05E-02	3.06E-07
mmu-miR-30b-5p	GTTTAC	GTTRY	7.15E-10	2.59E-07
mmu-miR-23b-3p	ATGTGA	RTGTR	5.03E-02	3.32E-01
mmu-miR-17-5p	CACTTT	YRYTT	6.57E-01	3.59E-03
mmu-miR-136-5p	ATGGAG	RTGGR	9.28E-01	1.00E+00
mmu-miR-125a-5p	TCAGGG	TYRGG	4.97E-01	9.79E-01
mmu-miR-106b-5p	CACTTT	YRYTT	6.93E-01	3.69E-03
mmu-miR-181c-5p	GAATGT	GRRTG	8.12E-01	8.22E-01
mmu-miR-221-3p	TGTAGC	TGRGY	1.28E-04	3.49E-07
mmu-miR-20a-5p	CACTTT	YRYTT	6.81E-01	3.23E-03
mmu-miR-340-5p	TTTATA	TTTRR	8.70E-01	1.23E-10
BRAIN D				
miRNA ID	Canonical Model	Expanded Model	Canonical Model P-value	Expanded Model P-value
mmu-miR-30e-5p	GTTTAC	GTTRY	0.00E+00	0.00E+00
mmu-miR-30d-5p	GTTTAC	GTTRY	0.00E+00	0.00E+00
mmu-miR-30a-5p	GTTTAC	GTTRY	0.00E+00	0.00E+00
mmu-miR-27a-3p	CTGTGA	YTGTR	0.00E+00	0.00E+00
mmu-miR-193b-3p	TGTGGG	TGTGG	8.69E-01	1.00E+00
mmu-miR-153-3p	GACTAT	GRYRT	0.00E+00	0.00E+00
mmu-miR-22-3p	CTGGCA	YTGGR	1.00E+00	1.00E+00
mmu-miR-449a-5p	TACACT	TRYRY	3.82E-04	0.00E+00
mmu-miR-136-5p	ATGGAG	RTGGR	1.00E+00	1.00E+00
mmu-miR-340-5p	TTTATA	TTTRR	3.25E-04	0.00E+00
mmu-miR-27b-3p	CTGTGA	YTGTR	0.00E+00	0.00E+00
mmu-let-7i-5p	ACTACT	RYTRY	0.00E+00	0.00E+00
mmu-let-7b-5p	CCTACT	YYTRY	0.00E+00	0.00E+00
mmu-miR-221-3p	TGTAGC	TGRGY	1.78E-05	0.00E+00
mmu-miR-21-5p	TGATAA	TGRTR	2.44E-06	9.99E-01
mmu-miR-374-5p	ATTATA	RTTRR	1.65E-05	0.00E+00
mmu-miR-31-5p	CATCTT	YRYTT	0.00E+00	0.00E+00
mmu-miR-217-5p	TGCAGT	TGRGT	9.66E-01	6.48E-01
mmu-miR-324-3p	GCAGTG	YRGTG	1.00E+00	1.00E+00
mmu-miR-34a-5p	GACACT	GRYRY	1.60E-04	0.00E+00
mmu-miR-17-5p	CACTTT	YRYTT	2.31E-06	6.35E-11
mmu-miR-186-5p	TTCTTT	TYTTT	4.44E-02	5.31E-01
mmu-miR-20a-5p	ATAAGC	RTRRY	1.85E-07	3.06E-13
mmu-miR-23b-3p	ATGTGA	RTGTR	2.89E-15	3.93E-06
mmu-miR-34c-5p	TCACTT	TRYRY	6.55E-04	0.00E+00
mmu-miR-451	ACGGTT	RYGGT	9.99E-01	1.00E+00
mmu-miR-126-5p	AATAAT	RRTRT	2.01E-09	7.54E-01



Table 2 | Cont

BRAIN D				
miRNA ID	Canonical Model	Expanded Model	Canonical Model P-value	Expanded Model P-value
mmu-miR-671-5p	AGGGCT	RGGGYT	8.66E-01	7.54E-01
mmu-miR-137-3p	GCAATA	GYRRTT	0.00E+00	1.00E+00
mmu-miR-204-5p	AAGGGA	RRGGGR	1.00E+00	1.00E+00

and of the four-G-C-disruption mutant are shown in Fig. 3a (see also Supplement). Mutant #15 also helped us observe significantly large motions of the PAZ domain. Strikingly, the PAZ domain bent away from both the N-domain (rotated by 55.7° and translated by 1.7Å – Fig. 3b) and the PIWI-domain (rotated by 42.5° and translated by 5.3Å – Fig. 3c), via two “hinges” close to PAZ in the L1 and L2 regions. The rotation and translation of the PAZ domain caused the nucleic-acid-binding channel to open between the PAZ and PIWI lobes, indicating that the bending motion of the PAZ domain plays a pivot role in the miRNA recognition process (see Supplement notes for more details).

The findings persist when simulating the 15-nt ternary complex.

Lastly, we carried out simulations with the longer 15-nt complex⁶⁷ – see Methods, Supp. Table 1, and Supp. Fig. S7–S9. In all cases, we were able to recapitulate the observations we made with the 11-nt complex. The extra 3'-compensatory pairing of the longer 15-nt complex has a minor contribution on maintaining the stability of the heteroduplex when the seed region bonds are broken (Supp. Fig. S7–S9). Our results are consistent with a recent finding that the complementary base-pairing beyond the seed region is not relevant for the repression of the *cog-1* 3'UTR and other *C. elegans* 3'UTRs by the *lcy-6* miRNA^{68,69}. It is noteworthy that, just like the case of Mutant #15 above, the same four G-C disruptions in the longer 15-nt complex (Mutant #19) result in large distortions in the seed region (Supp. Fig. S8, S9) and a disruption of the complex. This indicates that lack of nucleating base pairs in the seed region cannot be rescued even in the presence of a significant number of compensatory bonds beyond the seed.

Analyses of HITS-CLIP data corroborate the findings of the molecular dynamics simulations. HITS-CLIP (i.e. high-throughput sequencing of RNAs isolated by cross-linking immunoprecipitation) is a method that was introduced recently⁵⁵ for the analysis of miRNA targets from mouse brain: following ultraviolet irradiation, Ago was immunoprecipitated under stringent conditions; as expected, the Ago protein was cross-linked with miRNAs, resulting in complexes of ~110 kDa, and further with mRNAs, resulting in larger size complexes of ~130 kDa. In the original report⁵⁵, the immunoprecipitation was carried out using two distinct monoclonal antibodies and with biological replicates, and gave rise to a total of 10 datasets, five with enriched miRNAs and five with enriched mRNAs. We created genomic maps for the sequenced reads, paired up the maps of matching ~110 kDa and ~130 kDa sets and examined the data for support of the canonical miRNA targeting model and of an ‘expanded’ model that allows one or more G:U wobbles in the seed region (see Methods). Table 2 shows the results for these two models for the Brain A (Ago antibody 2A8) and Brain D (Ago antibody 7G1-1*) datasets. For each of the shown miRNA seeds, the Table lists the P-value that the matching HITS-CLIP ~130 kDa dataset could support the corresponding targeting model accidentally. As can be seen, HITS-CLIP data indeed support the ‘expanded model.’ For some of the miRNAs, e.g. miR-449a-5p, miR-222-3p, etc., the expanded model is a better fit for the HITS-CLIP data.

Discussion

We have presented a series of molecular dynamics simulations on Ago ternary complexes that focused on investigating the influence of seed-located wobbles, bulges and combinations thereof on the structural stability of the Ago-miRNA:mRNA complex and the motion of its domains, and, by extension its ability to cleave its target. We found that introduction of multiple G:U wobbles in the seed region only minimally affects the miRNA-mRNA heteroduplex and does not compromise the stability of the complex. With regard to bulge insertions in the seed region, and for a variety of possible arrangements, we find that they are tolerated on both the miRNA and the mRNA sides. Seed-region bulges that occur on the miRNA side of the heteroduplex give rise to slight distortions in the nucleic acid duplex and induce somewhat larger conformational changes but do not disrupt the complex. Seed-region bulges that occur on the mRNA side appear to be better tolerated by comparison. We also find that arrangements involving simultaneously multiple G:U wobbles and a single bulge lead to stable structures as well. Moreover, we examined the impact of artificially introduced disruptive mutations to the seed region and found a novel recognition mechanism that involves an important bending motion of the PAZ domain along the L1/L2 ‘hinge’ link followed by the opening of the nucleic-acid-binding channel. Lastly, we made use of several distinct publicly available HITS-CLIP datasets and found that they corroborate the conclusions of our current molecular simulations.

Our analyses provide additional evidence in support of and are consistent with earlier work that emphasized the importance of strong base-pairing interactions spanning positions 2 through 7 of a miRNA, or a subset of those positions (e.g. *in vivo* examples involving *lin-4* and *let-7* comprising seed region bulges). However, it is important to also realize that as our molecular dynamics analyses show such strong interactions can be realized in a multitude of ways that obviate the requirement that the exact reverse complement of the miRNA’s seed sequence be present in the target. In turn, this suggests that a given miRNA can give rise to non-canonical functioning heteroduplexes with targets that do not contain the miRNA seed. Taken together, these findings indicate that the spectrum of potential targets for a miRNA can admit a wide-spectrum of seed-less targets and thus substantially differs from what is anticipated by the canonical seed model. Consequently, our findings indicate that similar conclusions can be drawn about the potential spectrum of a given siRNA’s targets, considering that user-designed siRNAs and miRNAs share the same pathway downstream of the DICER cleavage. In other words, it follows that those mRNAs harboring sequences that are proximal to the seed of the transfected siRNA, either because they would induce G:U wobbles or the introduction of a bulge in the seed region, could also be down-regulated by the siRNA.

Methods

Molecular dynamics simulations. Following similar protocols as in our previous studies^{70–73}, the X-ray crystal structure of wild-type *TtAgo* bound to a 21-nt guide DNA and a 20-nt target RNA complex (PDB entry: 3F73, released in 2008.12) was used as the starting structure for the MD simulation⁶⁵. The DNA and the RNA strands can only be partly traced from position 1 to 11, and the base coordinates at position 10 and 11 are not available in the reported crystal structure⁶⁵. Therefore, the missing coordinates at position 10 and 11 were built from the known backbone structures. We repeated the simulations using the more recently released Ago complex with longer traceable guide-target duplex (length of the duplex is 15-bp, positions 2–16, PDB



entry: 3HK2, released in 2009.10)⁶⁷. The Ago-miRNA:mRNA complexes were generated by replacing the bases of the guide DNA by corresponding RNA (deoxyribose was replaced by ribose in A, C, and G whereas T was replaced by U). All the Ago complexes were solvated in $\sim 110 \times 100 \times 90 \text{ \AA}^3$ water boxes. A total of 32 Na^+ ions and 29 Cl^- ions were added to neutralize and mimic the biological environment (100 mM NaCl concentration). The solvated systems contain approximately 100,000 atoms. We utilized the NAMD2⁷⁴ package for the MD simulations with the NPT ensemble. The CHARMM (parameter set c32b1) force field was used for the protein and nucleic acid^{75,76}, and the TIP3P water model was used as the explicit solvent⁷⁷. The Particle Mesh Ewald (PME) method⁷⁸ as applied to treat the long-range electrostatic interactions and a 12 Å cutoff was employed for the van der Waals interactions. All the Ago complexes systems were equilibrated via a 20,000-step energy minimization to remove bad contacts. The minimized configurations were used as the starting point for 1-ns NPT MD equilibrations with 0.5 fs time-step at 1 atm and 310 K. The equilibrated configurations were then subjected to production runs for a minimum of 100 ns. The time step for all production runs was 1.5 fs with SHAKE/RATTLE algorithm⁷⁹.

Analyses of HITS-CLIP data. We used BWA⁸⁰ to quality-trim and map on the mouse genome all 10 sets of reads (Brain A through E at ~ 110 kDa, and Brain A through E at ~ 130 kDa) that resulted from the deep-sequencing of the two immunoprecipitations of the biological replicates⁵⁵. The mapping process allowed up to 2 mismatches and excluded all the reads that could not be mapped uniquely on the genome. More than 48 million reads in total were processed, of which almost 32 million were mapped uniquely. From each of the mRNA-enriched sets (~ 130 kDa) we only kept locations that had at least twenty reads mapped to them. We treated the products from the two arms of a miRNA separately and used the coordinates of each miRNA's 5p and 3p products (Release 18 of miRBase³⁸) to identify the top-30 most abundant miRNAs in the five miRNA-enriched sets (~ 110 kDa). Not unexpectedly, and given that our searches were carried out using a miRBase release that was significantly more enriched than the one used in the original work⁵⁵, we found some of the top-spots to be occupied by miRNAs that were added to miRBase only recently. For the rest of the analyses, the read sets were paired up: Brain A ~ 110 kDa \leftrightarrow Brain A ~ 130 kDa, Brain B ~ 110 kDa \leftrightarrow Brain B ~ 130 kDa, etc. Targets for the top miRNAs of a given ~ 110 kDa read set were sought in the genomic maps of the matching ~ 130 kDa set among locations to which 20 or more reads mapped. We focused only on those HITS-CLIP reads that mapped on the exons of known mouse protein-coding genes. We then carried out two types of searches. In one, we sought the exact reverse complement of the seed of a top-ranking miRNA (canonical model) in the mRNA maps. In the second, we examined whether the mRNA maps support an 'expanded' model of miRNA:mRNA interactions where one or more G:U wobbles are allowed in the seed region. To this end, we replaced every G in the seed by a *pyrimidine* (C or T; represented as Y) in the reverse complement of the seed, and every T in the seed by a *purine* (A or G; represented as R) and computed the P-value of HITS-CLIP *accidentally* supporting targeting by the corresponding miRNA under each of the two models (see Supplement for details).

- Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
- Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
- Berezikov, E., Chung, W.-J., Willis, J., Cuppen, E. & Lai, E. C. Mammalian mirtron genes. *Molecular Cell* **28**, 328–336 (2007).
- Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet* **9**, 102–114 (2008).
- Orum, U. A., Nielsen, F. C. & Lund, A. H. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular Cell* **30**, 460–471 (2008).
- Miranda, K. C. *et al.* A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* **126**, 1203–1217 (2006).
- Duursma, A. M., Kedde, M., Schrier, M., le Sage, C. & Agami, R. miR-148 targets human DNMT3b protein coding region. *Rna* **14**, 872–877 (2008).
- Lal, A. *et al.* p16(INK4a) translation suppressed by miR-24. *PLoS ONE* **3**, e1864 (2008).
- Tay, Y., Zhang, J., Thomson, A. M., Lim, B. & Rigoutsos, I. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature* **455**, 1124–1128 (2008).
- Forman, J. J., Legesse-Miller, A. & Coller, H. A. A search for conserved sequences in coding regions reveals that the let-7 microRNA targets Dicer within its coding sequence. *Proc Natl Acad Sci U S A* **105**, 14879–14884 (2008).
- Shen, W. F., Hu, Y. L., Uttarwar, L., Passegue, E. & Largman, C. MicroRNA-126 regulates HOXA9 by binding to the homeobox. *Mol Cell Biol* **28**, 4609–4619 (2008).
- Rigoutsos, I. New tricks for animal microRNAs: targeting of amino acid coding regions at conserved and nonconserved sites. *Cancer Res* **69**, 3245–3248 (2009).
- Ha, I., Wightman, B. & Ruvkun, G. A bulged lin-4/lin-14 RNA duplex is sufficient for *Caenorhabditis elegans* lin-14 temporal gradient formation. *Genes Dev* **10**, 3041–3050 (1996).
- Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843–854 (1993).
- Reinhart, B. J. *et al.* The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* **403**, 901–906 (2000).
- Wightman, B., Ha, I. & Ruvkun, G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in *C. elegans*. *Cell* **75**, 855–862 (1993).
- Didiano, D. & Hobert, O. Molecular architecture of a miRNA-regulated 3' UTR. *RNA* **14**, 1297–1317 (2008).
- Esquela-Kerscher, A. & Slack, F. J. Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* **6**, 259–269 (2006).
- Hammond, S. M. MicroRNAs as tumor suppressors. *Nat Genet* **39**, 582–583 (2007).
- Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038 (2010).
- Ma, L., Teruya-Feldstein, J. & Weinberg, R. A. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449**, 682–688 (2007).
- Nelson, P. T., Wang, W. X. & Rajeev, B. W. MicroRNAs (miRNAs) in neurodegenerative diseases. *Brain Pathol* **18**, 130–138 (2008).
- Calin, G. A. & Croce, C. M. MicroRNA signatures in human cancers. *Nat Rev Cancer* **6**, 857–866 (2006).
- Deng, S., Calin, G. A., Croce, C. M., Coukos, G. & Zhang, L. Mechanisms of microRNA deregulation in human cancer. *Cell Cycle* **7**, 2643–2646 (2008).
- Godshalk, S. E., Melnik-Martinez, K. V., Pasquinelli, A. E. & Slack, F. J. MicroRNAs and cancer: a meeting summary of the eponymous Keystone Conference. *Epigenetics* **5**, 164–168 (2010).
- Ryan, B. M., Robles, A. I. & Harris, C. C. Genetic variation in microRNA networks: the implications for cancer research. *Nat Rev Cancer* **10**, 389–402 (2010).
- Saeey, T. Cancer's little helpers: Tiny pieces of RNA may turn cells to the dark side. *Science News* (2010).
- Spizzo, R., Nicoloso, M. S., Croce, C. M. & Calin, G. A. SnapShot: MicroRNAs in Cancer. *Cell* **137**, 586–586.e581 (2009).
- Ventura, A. & Jacks, T. MicroRNAs and cancer: short RNAs go a long way. *Cell* **136**, 586–591 (2009).
- Voorhoeve, P. & Agami, R. Classifying microRNAs in cancer: The good, the bad and the ugly. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **1775**, 274–282 (2007).
- Small, E. M., Frost, R. J. A. & Olson, E. N. MicroRNAs add a new dimension to cardiovascular disease. *Circulation* **121**, 1022–1032 (2010).
- Small, E. & Olson, E. Pervasive roles of microRNAs in cardiovascular biology. *Nature* **469**, 336–342 (2011).
- Taganov, K. D., Boldin, M. P. & Baltimore, D. MicroRNAs and immunity: tiny players in a big field. *Immunity* **26**, 133–137 (2007).
- Wang, W.-X. *et al.* miR-107 regulates granulin/progranulin with implications for traumatic brain injury and neurodegenerative disease. *Am J Pathol* **177**, 334–345 (2010).
- Wang, W. X. *et al.* The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *J Neurosci* **28**, 1213–1223 (2008).
- Abelson, J. F. *et al.* Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* **310**, 317–320 (2005).
- Perkins, D. O. *et al.* microRNA expression in the prefrontal cortex of individuals with schizophrenia and schizoaffective disorder. *Genome Biol* **8**, R27 (2007).
- Griffiths-Jones, S. miRBase: the microRNA sequence database. *Methods Mol Biol* **342**, 129–138 (2006).
- Burgler, C. & Macdonald, P. M. Prediction and verification of microRNA targets by MovingTargets, a highly adaptable prediction method. *BMC Genomics* **6**, 88 (2005).
- Enright, A. J. *et al.* MicroRNA targets in *Drosophila*. *Genome Biol* **5**, R1 (2003).
- John, B. *et al.* Human MicroRNA targets. *PLoS Biol* **2**, e363 (2004).
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U. & Segal, E. The role of site accessibility in microRNA target recognition. *Nat Genet* **39**, 1278–1284 (2007).
- Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
- Moxon, S., Moulton, V. & Kim, J. A scoring matrix approach to detecting miRNA target sites. *Algorithms for Molecular Biology* **3**, 3 (2008).
- Rajewsky, N. & Socci, N. D. Computational identification of microRNA targets. *Dev Biol* **267**, 529–535 (2004).
- Stark, A., Brennecke, J., Russell, R. B. & Cohen, S. M. Identification of *Drosophila* MicroRNA targets. *PLoS Biol* **1**, E60 (2003).
- Rehmsmeier, M., Steffen, P., Hochsmann, M. & Giegerich, R. Fast and effective prediction of microRNA/target duplexes. *RNA* **10**, 1507–1517 (2004).
- Ritchie, W., Flamant, S. & Rasko, J. E. J. Predicting microRNA targets and functions: traps for the unwary. *Nat Methods* **6**, 397–398 (2009).
- Rigoutsos, I. & Tsigonis, A. in *MicroRNAs in Development and Cancer* Vol. 1 *Molecular Medicine and Medicinal Chemistry* (ed Frank, J. Slack) Ch. 10, (Imperial College Press, 2010).
- Selbach, M. *et al.* Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58–63 (2008).



51. Baek, D. *et al.* The impact of microRNAs on protein output. *Nature* **455**, 64–71 (2008).
52. Didiano, D. & Hobert, O. Perfect seed pairing is not a generally reliable predictor for miRNA-target interactions. *Nat Struct Mol Biol* **13**, 849–851 (2006).
53. Easow, G., Teleman, A. A. & Cohen, S. M. Isolation of microRNA targets by miRNP immunopurification. *RNA* **13**, 1198–1204 (2007).
54. Lal, A. *et al.* miR-24 Inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to “seedless” 3’ UTR microRNA recognition elements. *Molecular Cell* **35**, 610–625 (2009).
55. Chi, S. W., Zang, J. B., Mele, A. & Darnell, R. B. Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* **460**, 479–486 (2009).
56. Zisoulis, D. G. *et al.* Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nature Structural & Molecular Biology* **17**, 173–179 (2010).
57. Thomas, M., Lieberman, J. & Lal, A. Desperately seeking microRNA targets. *Nature Structural & Molecular Biology* **17**, 1169–1174 (2010).
58. Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
59. Stark, A., Brennecke, J., Bushati, N., Russell, R. B. & Cohen, S. M. Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3’ UTR evolution. *Cell* **123**, 1133–1146 (2005).
60. Farh, K. K. *et al.* The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* **310**, 1817–1821 (2005).
61. Fabian, M. R., Sonenberg, N. & Filipowicz, W. Regulation of mRNA translation and stability by microRNAs. *Annual Review of Biochemistry* **79**, 351–379 (2010).
62. Vella, M. C., Choi, E. Y., Lin, S. Y., Reinert, K. & Slack, F. J. The *C. elegans* microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3’ UTR. *Genes & Development* **18**, 132–137 (2004).
63. Chi, S. W., Hannon, G. J. & Darnell, R. B. An alternative mode of microRNA target recognition. *Nature Structural & Molecular Biology* **19**, 321–327 (2012).
64. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209–213 (2008).
65. Wang, Y. *et al.* Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature* **456**, 921–926 (2008).
66. Brennecke, J., Stark, A., Russell, R. B. & Cohen, S. M. Principles of microRNA-target recognition. *PLoS Biol* **3**, e85 (2005).
67. Wang, Y. *et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754–761 (2009).
68. Garcia, D. M. *et al.* Weak seed-pairing stability and high target-site abundance decrease the proficiency of *lisy-6* and other microRNAs. *Nat Struct Mol Biol* **18**, 1139–1146 (2011).
69. Shin, C. *et al.* Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell* **38**, 789–802 (2010).
70. Das, P., King, J. A. & Zhou, R. Aggregation of gamma-crystallins associated with human cataracts via domain swapping at the C-terminal beta-strands. *Proc Natl Acad Sci U S A* **108**, 10514–10519 (2011).
71. Liu, P., Huang, X., Zhou, R. & Berne, B. J. Observation of a dewetting transition in the collapse of the melittin tetramer. *Nature* **437**, 159–162 (2005).
72. Zhou, R., Eleftheriou, M., Royyuru, A. K. & Berne, B. J. Destruction of long-range interactions by a single mutation in lysozyme. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5824–5829 (2007).
73. Zhou, R., Huang, X., Margulis, C. J. & Berne, B. J. Hydrophobic collapse in multidomain protein folding. *Science* **305**, 1605–1609 (2004).
74. Kumar, S. *et al.* Scalable Molecular Dynamics with NAMD on Blue Gene/L. *IBM Journal of Research and Development* **52**, (2007).
75. Brooks, B. R. *et al.* CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
76. MacKerell, A. D., Jr., Banavali, N. & Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56**, 257–265 (2000).
77. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **79**, 926–935 (1983).
78. Darden, T. A., York, D. M. & Pedersen, L. G. Particle mesh Ewald: An NlogN method for Ewald sums in large systems. *J Chem Phys* **98**, 10089–10092 (1993).
79. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **23**, 327–341 (1977).
80. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

Acknowledgements

We thank Payel Das, Jingyuan Li, Bruce Berne, and Chunli Yan for helpful discussions. This work is supported by the IBM Blue Gene Science program. The work of Y.Z., H.-W.C., P.L., P.C. and I.R. was supported by Thomas Jefferson University startup funds and by a William M. Keck Foundation grant to I.R.

Author contributions

R.Z. and I.R. designed and supervised the research. I.R. and R.Z. designed the collection of mutants. I.R., Z.X., R.Z., P.R. and P.C. prepared the manuscript. Z.X., R.Z. and T.H. carried out the molecular dynamics simulations. R.Z. and Z.X. analyzed the dynamic structures of Ago complexes and mutants. Y.Z., H.-W. C. and P.L. generated several of the next-generation-sequencing mapping and analysis tools or carried out the mapping of the HITS-CLIP data to the mouse genome. P.C. carried out the HITS-CLIP studies and P.C. and I.R. analyzed the results.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

How to cite this article: Xia, Z. *et al.* Molecular dynamics simulations of Ago silencing complexes reveal a large repertoire of admissible ‘seed-less’ targets. *Sci. Rep.* **2**, 569; DOI:10.1038/srep00569 (2012).



SUBJECT AREAS:

RNAI

NON-CONDING RNAS

SMALL RNAS

GENE EXPRESSION

ERRATUM: Molecular dynamics simulations of Ago silencing complexes reveal a large repertoire of admissible 'seed-less' targets

Zhen Xia^{1,2}, Peter Clark³, Tien Huynh¹, Phillippe Loher³, Yue Zhao³, Huang-Wen Chen^{3#}, Pengyu Ren², Isidore Rigoutsos^{3,4*} & Ruhong Zhou^{1,5*}

SCIENTIFIC REPORTS:

2:569

DOI: 10.1038/srep00569

(2012)

¹Computational Biology Center, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, ²Department of Biomedical Engineering, University of Texas at Austin, Austin, TX 78712, ³Computational Medicine Center, Thomas Jefferson University, Philadelphia, PA 19107, ⁴Department of Pathology, Anatomy and Cell Biology; Department of Cancer Biology; Department of Biochemistry and Molecular Biology; Thomas Jefferson University, Philadelphia, PA 19107, ⁵Department of Chemistry, Columbia University, New York, NY 10027.

Published:

10 August 2012

Updated:

28 November 2012

Due to a technical error, Pengyu Ren was omitted from the author list of this article. The correct author details appear above.

*These authors contributed equally.

#Current address:

Bloomberg L.P., New York, NY 10022.