Title: Single Base-Pair Conformational Switch Modulates miR-34a Targeting of Sirt1 mRNA

Authors: Lorenzo Baronti[1], Ileana Guzzetti[1†], Parisa Ebrahimi[2†], Sarah Friebe Sandoz[1†], Emilie Steiner[1†§], Judith Schlagnitweit[1], Bastian Fromm[3], Luis Silva[1], Carolina Fontana[1‡], Alan A. Chen[2] and Katja Petzold[1*]

Affiliations:

[1]Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm Sweden.

[2]Department of Chemistry and RNA Institute, University at Albany, SUNY, Albany, NY, 12222 USA

[3]Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden

[§]Current affiliation: Laboratoire SABNP, UMR INSERM 1204, Université d'Evry-Val-d'Essonne – Université Paris-Saclay, Evry, France

[‡]Current affiliation: Departamento de Química del Litoral, Facultad de Química & CENUR Litoral Norte, Universidad de la República, Paysandú, Uruguay

[†]These authors contributed equally to this work

[*]Correspondence to: katja.petzold@ki.se

Summary:

MicroRNAs (miRNAs) regulate translation levels of messenger RNAs (mRNAs). Currently, the main parameter explaining target selection and repression efficiency is the base-pairing between the miRNA and mRNA in the seed region. Employing $R_{1\rho}$ relaxation dispersion Nuclear Magnetic Resonance and molecular simulation, we reveal a dynamic switch based on a single base-pair rearrangement in the miRNA–mRNA duplex, elongating the weak 7mer seed to a complete 8mer seed. This switch additionally causes a co-axial stacking of the seed and supplementary helix in human Argonaute 2 protein (hAgo2) such that the active state, reminiscent of prokaryotic Ago, becomes favored. Stabilizing this transient state exhibits enhanced target repression in cells, indicating the importance of the miRNA–mRNA structure. Our observations tie together the current understanding of the step-wise miRNA targeting process from the initial “screening” state to an “active” state and unveil the role of the RNA duplex in hAgo2 beyond the seed.

Main Text:

MicroRNAs (miRNAs), non-coding RNA molecules, regulate gene expression by targeting messenger RNAs (mRNAs). Each mature miRNA, ~22 nucleotides (nts), is bound to one Argonaute protein (hAgo1-4), forming the RNA-Induced Silencing Complex (RISC). In RISC, nts 2-6 of the guide miRNA (g2-g6) are pre-arranged to recognize mRNA targets through Watson-Crick (WC) base-pairing[1-3] in the seed (Fig. 1a–b). This base-pair complementarity (g2-g8) largely determines RISC activity[4,5], e.g. sites with only g2-g6 (5mer) complementarity are rejected as unspecific. In *h*Ago2, sites with prolonged base-pairing, ≥g2-g7 (≥6mer), are able to override the checkpoint imposed by helix-7 and induce a conformational transition in the protein allowing an extended 3'-pairing of the RNA[6]. The interplay between the 3'-end of the seed and helix-7 modulates this process, due to inherent flexibility[7]. Bioinformatics analysis of validated miRNA–mRNA pairs cannot discern sequence determinant in this region, beyond a preference for forming bulges[8]. X-ray structures of ternary complexes are unable to resolve the central region of the duplex, supporting its flexibility[9]. *In vitro* biochemical studies showed that mismatches in this region contribute little to target binding affinity but can impair catalytic siRNA cleavage in *Drosophila* Ago2[10]. Implying that dynamics of the central RNA bases are essential for the fate of target mRNAs, however, the precise nature of the guide-target interaction beyond the seed region remains unclear.

We employ Nuclear Magnetic Resonance (NMR) to observe this dynamic process underlying miRNA–mRNA targeting. To deduce a structural model of the RNA's conformational transition, we combine these measurements with molecular simulations and Dual-Luciferase Reporter (DLR) assay.

We study hsa-miR-34a-5p (miR-34a), part of evolutionary conserved miR-34/449 family[11-15], targeting the Silent Information Regulator 1 (Sirt1), a p53 deacetylating enzyme, in a tumor-suppressive feedback loop. Using $R_{1\rho}$ NMR Relaxation Dispersion (RD), we show that the weak 7mer-A1 seed of the miR-34a–mSirt1 duplex (Fig. 1b) is in equilibrium with a transient and low-populated excited state (ES) that results in an 8mer seed with a G:U base-pair at its 3'-end. The extended seed alters the topology of the duplex by shifting the bending angle between the seed and the 3'-helix in RISC, as shown by simulations. In a cell-based assay, the structural mimic of the extended seed exhibits ~2-fold increase in target downregulation. Our data suggest a model for RISC undergoing a structural transition mediated by RNA dynamics, where RISC first recognizes its target by screening for correct seed pairing, and then transitions into an active complex, where miR-34a 3'-end is released and fully binds mSirt1.

Seed dynamics of miR-34a–mSirt1 binding site

Since RISC recognizes thousands of distinct binding sites with no apparent sequence preference beside the seed, we hypothesize that miRNA–mRNA pairs possess distinct conformational characteristics in the central bulge, which facilitate accommodation within hAgo2.

First, we solved the secondary structure of miR-34a bound to the validated target site in Sirt1 mRNA (miR-34a–mSirt1 duplex)[16] by NMR (Fig. 1a, Fig. S2, SD2). The overall fold confirmed the secondary structure predicted using MC-Fold (Fig. 2a, Extended Data Fig. 1): the 5nts seed constitutes an A-form 5'-helix between gG2:tC27 and gG6:tC23 base-pairs, gG8:tC17 and gG18:tC7 form a 3'-helix containing a wobble gU11:tG14 base-pair. These two helices are separated by a 4nts asymmetric bulge on the mSirt1 side, (tC18-tU21) (Fig. 1a&2a).

To study the structure and dynamics of the bulge, we designed a shortened hairpin construct (miR-34a–mSirt1 bulge) containing the 4nts bulge and enclosing regions (Fig. 1a grey box, 2a). The correct fold was confirmed by chemical-shifts (CSs) comparison of the shared residues (Fig. S2). The intrinsic flexibility of the miR-34a–mSirt1 complex precluded a traditional NMR tertiary structure calculation with a single, static conformation. Therefore, we used an NMR-informed computational approach and computed the RNA's conformational ensemble using Replica-Exchange Molecular Dynamics (REMD) simulations constraining the base-pairing determined from imino ^{1}H-^{1}H NOEs derived by NMR (Fig. S3). Simulations were varied in temperature, for exploration of RNA conformations fulfilling the experimental constraints, resulting in an ensemble of 153 structures. One representative structure of the ensemble is shown in Fig. 2d and the relative stem-to-stem angle distribution in Fig. 2g.

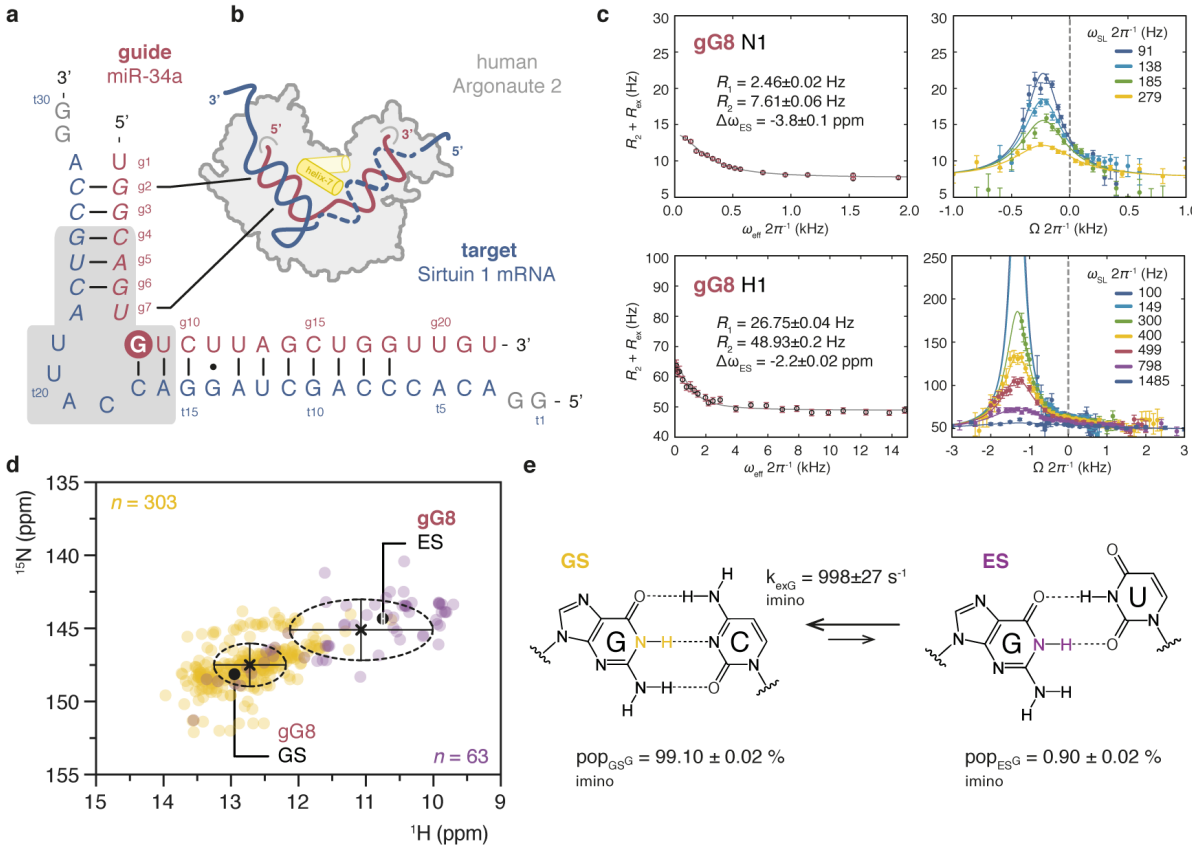

Figure 1 | Conformational dynamics in seed of miR-34a–Sirt1 mRNA. a, Secondary structure of miR-34a–mSirt1 duplex determined by NMR. Seed (*italic*), 5 base-pairs. Grey box indicates nts selected for reduced construct for $R_{1\rho}$ RD measurements. **b**, Sketch of hAgo2 accommodating miR-34a–mSirt1 duplex. **c**, ^{15}N and ^{1}H $R_{1\rho}$ RD profiles of gG8N1 and gG8H1 report on a single base-pair switching of gG8:C→gG8:U (red in **a**). **d**, CSs distribution of ^{1}H1-^{15}N1 moieties of G's in G:C (yellow) or G:U (purple) from the BMRB[17]. Crosses: average for CSs of G:C and G:U, respectively; dashed ellipses: 1s.d.; black dots: CSs for gG8 in ground state (GS) and RD-derived excited state (ES). **e**, G:C→G:U base-pair switch highlighting the G's ^{1}H1-^{15}N1 groups in yellow and purple.

Although classified as 7mer-A1 binding site by prediction tools (e.g. Targetscan[18]), we found that the miR-34a–mSirt1 duplex and the reduced construct represent a less stable structure, as stability of gU7:tA22 closing base-pair at the 3'-end of the seed is significantly reduced as shown by NMR[19] (Fig. 2a, S2a and S3a, SD2–3). We suggest therefore that a weak pairing in position 7 might explain previously observed sequence-specific differences in RISC-target binding affinity[6]. In agreement with nearest-neighbor models for A:U closing hairpins[20], we propose that 6mer/7mer-A1 seeds ending with closing A:U base-pairs in position 7, might not suffice for a stable displacement of hAgo2's helix-7, resulting in significantly lower binding affinities, closer to the predicted 6mer affinity.

To assess the base-pair dynamics, we measured ^{15}N, ^{13}C and ^{1}H $R_{1\rho}$ NMR RD[21-24] experiments. ^{1}H1-^{15}N1 of gG8H1, gG8N1 and ^{13}C gG8C8, tU21C6, tC17C1', tU20C1', tA19C8, tA19C2 and tA22C8 revealed a global exchange process, where the base-pair gG8:tC17 interconverts from the most stable structure, the ground state (GS), to a minor populated excited state (ES) with an exchange rate constant for ^{1}H1-^{15}N1 of gG8 ($k_{exGimino}$) of 998±27s^{-1} with ES population ($pop_{ESGimino}$) of 0.90±0.01% and global k_{ex} (k_{exG}) of 1008±12s^{-1} with ES population (pop_{ESG}) of 0.90±0.01% (Fig. 1c, S6c and Data S1). Most importantly,

the individual CSs difference between GS and ES ($\Delta\omega_{ES} = \Omega_{ES} - \Omega_{GS}$), describing the structure of the ES, was obtained by measuring ^{1}H ($\Delta\omega_{ES}$ = -2.20±0.02 ppm) and ^{15}N ($\Delta\omega_{ES}$ = -3.8±0.1 ppm) $R_{1\rho}$ RD datasets. This approach allowed us to infer that gG8 ES CSs resides in a region of the ^{1}H-^{15}N HSQC spectrum that is a signature for G:U wobble base-paired guanosines. This was validated by querying the BMRB[17] for ^{1}H1-^{15}N1 CSs of G:U base-paired Gs of RNA-only entries and comparing them to the G:C distribution (Fig. 1d–e, S6c and Data S1 Tab 1).

Base-pair switch alters topology of complex

When analyzing MC-Fold[25] output for alternative secondary structures fulfilling the NMR-derived model (Fig. 1e), we found that a base-pairing partner switch between gG8:tC17 to gG8:tU21 occurs within the third most energetically favorable structure (Fig. Extended Data Fig. 1a–b). To characterize the nature of this process we measured additional ^{13}C $R_{1\rho}$ RD experiments on aromatic C2/6/8 and sugar C1' nuclei, known reporters of sugar pucker, stacking and base-pairing. The additional, individually fitted nuclei resulted in an exchange process with average parameters of k_{EX} = 1371Hz and pop_{ES} = 1.9%; similar to the global fit of ^{1}H, ^{13}C and ^{15}N datasets (Fig. S6–7 and Data S1 Tab 1). Based on known correlations between our measured ^{13}C $R_{1\rho}$ $\Delta\omega$ values and structural propensities[26-28], we proposed a refined secondary structure of the ES summarized in Fig. 2a.

To derive a 3D structural model of ES, high temperature REMD simulations of the GS were performed, restraining 5 experimentally-determined base-pairs (Data S1 Tab 11)[29]. A putative conformation of the ES was identified as a cluster within simulations of GS in which gG8:tU21 is base-paired to gG8:tC17. The ES conformer was sampled restraining gU9:tA16 and gG6:tC23. Mg^{2+} addition, experimentally and by simulation[30], did not show any influence (Extended Data Fig. 5 and SD6). As with GS, one representative structure of the ES ensemble (210 structures) is presented in Fig. 2e. The topology of the ES is altered compared

to GS, indicated by a stem-to-stem co-axial stack resulting in an angle distribution peaking around 90° (Fig. 2g, middle).

To experimentally validate the candidate ES structure, we used the NMR Mutate-and-Chemical-Shift-Fingerprint (MCSF) approach[26], where a substitution or chemical modification is used to trap the proposed ES. CSs are then compared between the trapped ES and $R_{1\rho}$ RD-derived data. We introduced a two-point isosteric substitution in the bulge construct (swapping tC17→tU17 and tU21→tC21) promoting the repositioning of gG8 to the seed 5'-helix, base-paired with tC21 (miR-34a–mSirt1 trapped ES), without impacting the overall binding affinity. We determined the secondary structure of trapped ES by NMR (Fig. 2b and S4) and used imino ^{1}H-^{1}H NOE's as sparse constraints to calculate structural ensembles via REMD (Fig. 2f). As expected, the trapped ES forms an additional gG8:tC21 base-pair elongating the seed 5'-helix, resulting in identical base-pairing patterns and inter-helical bending angles as observed in the ES (Fig. 2g, bottom).

The MCSF showed remarkable agreement for C1's, tA22C2 and gG8C8 (Fig. 2c, green) confirming that the trapped ES well represents the overall topology of the ES modeled from $R_{1\rho}$ RD data. The sugar puckers measured by $^{3}J_{H1'\text{-}H2'}$ for tU20, tU21 (dominant C2'-*endo*) and tC18 (dominant C3'-*endo*) that were expected to interconvert to their opposite configuration in ES from $R_{1\rho}$ RD (Fig. 2a), were successfully recapitulated in the trapped ES (Fig. 2b). Furthermore, co-axial stacking between the two helices is validated by CSs of tA22H8/C8, tA16H8/C8 and gG8H8/C8 shifting to a region characteristic of nts embedded in uninterrupted A-form helix[27] (Extended Data Fig. 4 and Fig. S5a).

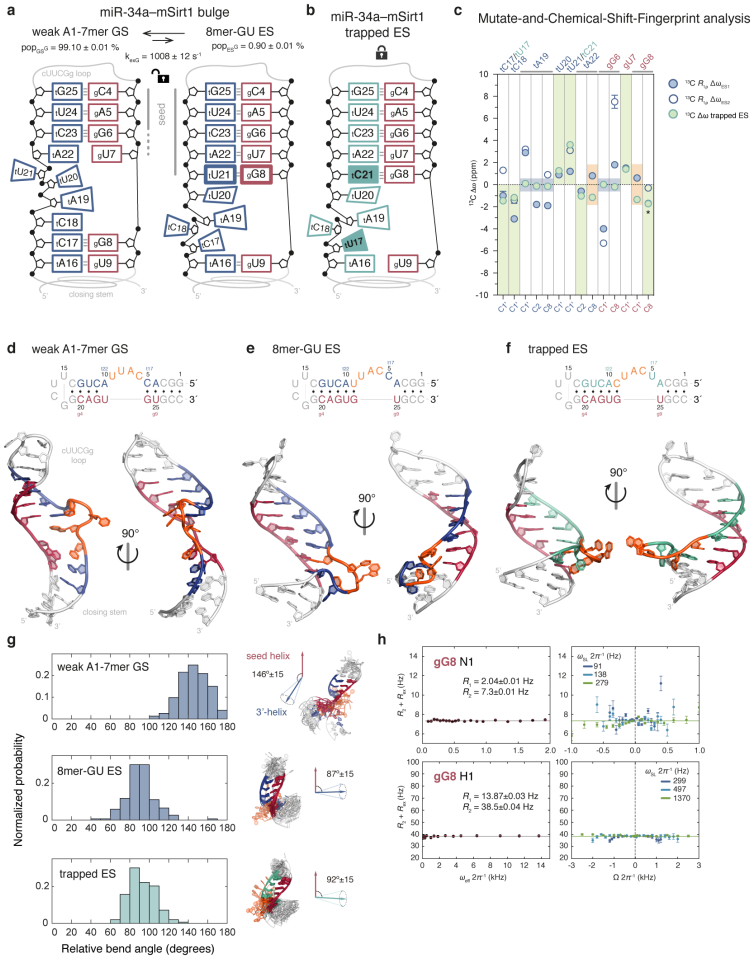

Figure 2 | Structure and conformation of miR-34a-mSirt1 excited state. a, Secondary structures of bulge region (Fig. 1a, grey). Left: GS solved by NMR. Right: ES resulting from $R_{1\rho}$ RD-derived CSs (gG8:tC17→gG8:tU21, bold box). **b**, Stabilization of ES conformation by isosteric two-point substitution tC17→tU17/tU21→tC21 (trapped ES). Secondary structure solved by NMR. **c**, MCSF analysis of trapped ES validates the ES model (green). Expected perturbations are observed at site of modification (orange), tA19 and gG6 (blue) explained in SD5, Extended Data Fig. 2. Individually fitted ^{13}C $R_{1\rho}$ RD-derived Δω are in blue, full dots for ES1 and hollow dots ES2, for 3-state fitting datasets. **d–f**, Representative conformations from NMR-informed REMD of GS **d**, ES **e**, trapped ES **f**. **g**, Inter-helical bend-angle distributions for GS (top), ES (middle) trapped ES (bottom). Mean and s.d. of angles distribution derived from REMD. **h**, $R_{1\rho}$ RD profiles of trapped gG8N1&H1 show loss of dynamics, in the timescale probed.

Inconsistencies for tA22C8 and gG8C1' are a consequence of the substitution (Fig. 2c, orange, Extended Data Fig. 2). tA19 and gG6 datasets report on presence of a second, thermodynamically similar excited state (ES2) (Fig. 2c, blue). However, this conformation could not be trapped experimentally and is discussed in SD5 and Extended Data Fig. 2. Interestingly, when probing dynamics of the trapped ES with RD experiments, no exchange with alternative conformations was detected in the probed timescale (Fig. 2h, Fig. S8–9, Data S1 Tab 3).

In summary, our results show that the miR-34a–mSirt1 binding site is in equilibrium between a high-populated weak 7mer-A1 GS and a low-populated 8mer-GU seed-elongating ES ('8mer-GU ES'), where position 8 is occupied by a G:U base-pair, a motif previously observed for miR-48[31,32]. Upon GS→ES switch, both $R_{1\rho}$ RD data and REMD indicate re-arrangement of the bulge and a stacking of the two helices.

Functional relevance of 8mer-GU excited state

We compared wild-type (WT) miR-34a–mSirt1 and miR-34a–mSirt1 trapped ES (Fig. 3a) by measuring thermal stability followed by UV absorption, RNA-RNA binding affinity by Electrophoretic Mobility Shift Assay (EMSA) and RISC target affinity by Filter-Binding Assay (FBA) of miR-34a-loaded hAgo2 (Fig. 3b-d). We found that melting temperature (T_m) and dissociation constant (K_d) were unchanged (Fig. 3b and c, Extended Data Fig. 3), showing that the substitution does not affect the duplex stability *in vitro*. Similarly, binding affinity of hAgo2 towards the target RNAs in FBA is the same within error (Fig. 3d). Next, we asked if the two binding sites, despite similar stability, have different target downregulation levels in cells. DLR assay in HEK293T cells of miR-34a co-transfected for the WT weak 7mer-A1 results in 52.3±3.5%, as previously reported[16] (Fig. 3e blue), while trapped ES 8mer 31.0±5.7% (Fig. 3e turquoise), showing that the two-point substitution trapping the ES results in an ~2-fold significant increase in downregulation of the target.

Taken together, the FBA and DLR assays suggest that, for stably-bound 3'-paired targets, the binding affinity cannot fully explain the observed biological data.

This difference prompted us to compute the RNA structure in the context of RISC[6]. We used slow-growth simulation to test whether the calculated ensembles of miR-34a–mSirt1 bulge GS, ES and trapped ES (Fig. 2d–f) could be accommodated in hAgo2 binding site. Starting from the crystal structures[6] the visible crystallographic A-form seed helix was replaced with conformations from the miR-34a–mSirt1 bulge GS, ES and trapped ES ensembles and than aligned with the seed of the co-crystal (Fig. S10).

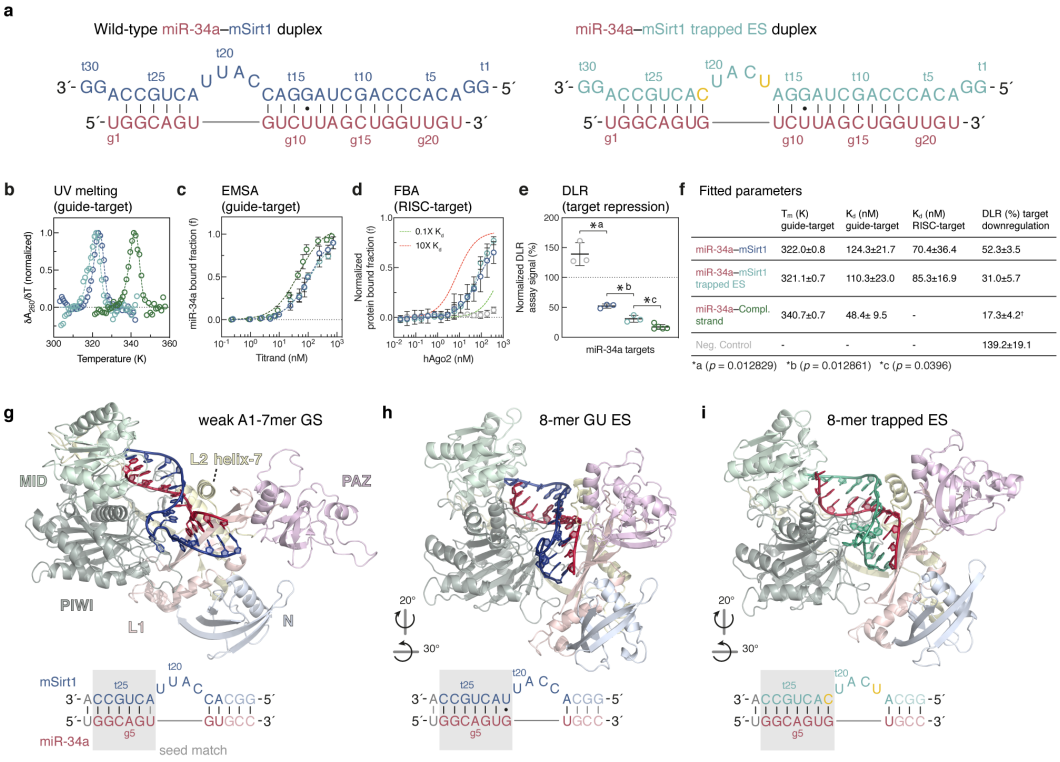

Figure 3 | Biophysical and functional characterization of WT and trapped ES miR-34a–mSirt1 duplexes. a, Secondary structures of the WT and trapped ES (tC17→tU17/tU21→tC21 in yellow). **b–d**, WT (blue) and trapped ES (turquoise) show comparable stability as indicated by equivalent T_m (**b**, UV melting), K_d (**c**, EMSA) and binding to miR-34a loaded RISC (**d,** FBA) within error. (Extended Data Fig. 3, Extended Data Table 1). **e**, DLR assay reveals an ~2-fold increased miR-34a

downregulation of trapped ES (8mer - turquoise) with respect to WT (weak 7mer-A1 - blue). Grey: Scrambled negative control. Green: highest level of downregulation (siRNA-type). **f,** Table with fitted parameters for a–e and *p*-values from DLR (paired, two-tailed t-test), †performed independently. **g–i**, Slow-growth simulated RNA structures bound to hAgo2 (PDB: 4W5O)[6]. **g**, GS (Fig. 2a&d) PAZ-oriented conformation. **h**, Oppositely, ES (Fig. 2a&e), co-axial, N-PIWI-oriented conformation, recapitulated the 8mer trapped ES **i**.

The resulting ternary complexes are shown in Fig. 3g–i. GS ensemble samples the 3'-helix of miRNA–mRNA within the PAZ domain (Fig. 3g), where the miRNA is bound prior to target binding[1,2,6,33]. In contrast, the 8mer–GU ES conformation adopts a global bend angle stacking the 3'-helix co-axially with the seed and favoring binding along the PIWI-N domains (Fig. 3h), also recapitulated in the trapped ES (Fig. 3i).

While only small conformational changes in the crystal structure of hAgo2 are needed to bind the microRNA–mRNA complex in the GS state, accommodating the ES state requires pivoting of the PAZ domain (Fig. 4e, Video S1), consistently with prior studies, where simulations identified these PAZ domain movements leading to more "open" hAgo2 conformations[34-36]. Intriguingly, the slow-growth induced-fit conformation of hAgo2–ES is reminiscent of binding modes observed for DNA-bound prokaryotic Ago ternary complexes[37-40] (Extended Data Fig. 6), suggesting that hAgo2 undergoes structural changes during target recognition and down-regulation activity.

We therefore performed a sequence search for other GS→ES instances in the 28,653 isoforms of 19,432 genes of 3'-UTRs of human protein coding genes. A minimal A1-6mer seed resulted in 3,269 predicted target sites for miR-34a (Fig. 4a), where a sequence search and secondary structure prediction for the GS→ES motif was then performed using MC-Fold[25], resulting in bulge sizes from 1nt (139 and 74 representatives), 2nt (109 & 45), 3nt

(123 & 33), 4nt (105 & 26) and 5nt (117 & 15), respectively. In a more stringent cluster, with 3 WC base-pairs following the bulge, 22 targets could be identified (Fig. 4c). We selected 5 different targets for further investigation using DLR assay in HEK293T (HEBP1, ADAM22, ATG9A[41], ANKS1A and CCND1[42]). All 5 mRNA candidates we tested were downregulated in the trapped ES form compared to WT, ranging from 50-80% downregulation efficiency (Fig. 4d and Materials and Methods section), suggesting that conformational switching of bulged miRNA–mRNA complexes is a general mechanism for modulating downregulation efficiency.

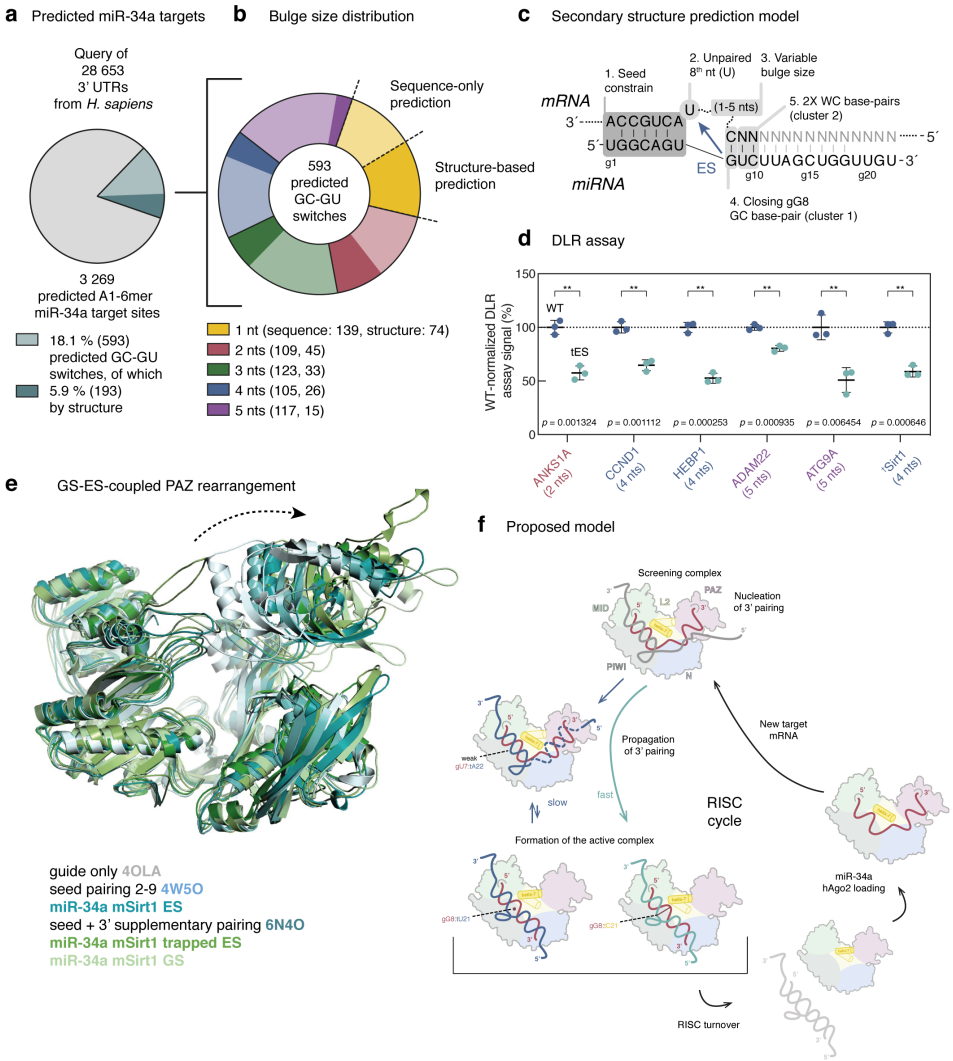

Figure 4 | Proposed mechanism of down-regulation for GC-GU switches in miR-34a loaded RISC. a, Predicted miR-34a targets in human 3'-UTRs (grey), that can experience GC-GU switch by sequence (18.1%) and by structure (5.9%). **b,** Distribution of bulge sizes (by sequence, lighter, or structure, darker color). **c,** Model of criteria used to search for GC-GU switches. **d,** DLR assay of 5 target mRNAs repression with respective trapped ES. All datasets were internally and WT-normalized for comparison, †performed independently. Shown are *p*-values (paired, two-tailed t-test) **e,** Overlay of X-ray structures and slow-growth simulations. **f**, Proposed mechanism for the miRNA-mRNA switch in RISC from a screening into an active state.

Discussion

While seed matching is important, it is only the first step of the RISC cycle. Subsequently, it is believed that nucleation from the 3'-helix can propagate towards the central region. This, in concert with disengagement of the miRNA 3'-end from the PAZ domain leads to an active complex, or rather the final step in the RISC activity cycle[43].

We propose that in the case of miR-34a–mSirt1, this process is mediated by a conformational transition triggered by gG8 switching its base-pairing partner. In its GS, miR-34a–mSirt1 adopts a 7mer-A1 seed, closed by a weak base-pair in position g7, better described as a 6mer-A1 seed, which is unable to fully displace hAgo2's helix-7 (Fig. 2a). The GS accesses a distribution of inter-helical bend angles placing the miR-34a 3'-end towards the PAZ domain, favoring initial target engagement and nucleation of the 3'-helix[43] (Fig. 4f). During the GS→ES transition, gG8 repositions to the seed helix and pairs with tU21, resulting in an extended 8mer–GU seed. The re-arrangement of gG8 causes co-axial stacking of the two helices and therefore release of the 3'-end of miR-34a from PAZ, re-orienting the RNA duplex towards PIWI domain (Fig. 3g–h). This process is accommodated by the concerted widening of the N-PAZ channel[6] that facilitates binding of the new stem-to-stem orientation to the cleft and repositioning along the PIWI-N channel in a second binding mode. This ES conformation is similar to the catalytically competent state reported for prokaryotic Ago[37-40] (Fig. 3I&J, 4e, Extended Data Fig. 6a–b) and a recent hAgo2 structure confirms that the 3'-helix is mobile[9].

We thus propose that the GS→ES transition described here presents a mechanism to achieve an active, "catalytically competent" RISC, promoting mRNA down-regulation[43,44]. While hAgo2-bound miRNA is not known to cleave targets with partial complementarity, it is

possible that these conformational changes enable RISC to achieve multiple turnovers, that will increase downregulation of the target mRNA[45-47] (Fig. 4f).

Our biophysical and in-cell functional results support this hypothesis by showing a ~2-fold increase of downregulation upon GS→ES stabilization while maintaining RNA–RNA stability. We show that 5 selected mRNA targets of miR-34a have similar increases in downregulation efficiency upon being trapped in their ES (Fig. 4d). This finding indicates that the mechanism proposed in our work could be a widespread feature of bulged binding sites containing partial or extended 3'-pairing.

We have demonstrated that the structural transitions of the guide–target RNA in RISC provides a mechanistic explanation for bulged complexes, enabling a more accurate prediction of microRNA target downregulation. With ever-increasing interests in adapting RNA-guided nuclease machineries for therapeutic, diagnostic, and technology applications, we believe that leveraging the power of RNA conformational dynamics will lead to the design of better guide RNAs as well as a deeper understanding of these macromolecular complexes.

Methods

RNA sample preparation

RNA samples were produced *in-house* by T7 *in vitro* transcription[48,49], unless otherwise stated. Modified DNA templates (Integrated DNA Technologies (IDT)) with oxy-methylated C2' groups in the first two 5′ nucleotides were used to reduce the 3′-OH heterogeneity of the product[50]. *In vitro* transcription reactions were supplemented with 20% DMSO to improve reaction yield and reduce side products[51]. ^{13}C, ^{15}N labeled NMR samples were produced by supplementing the transcription reaction with ^{13}C, ^{15}N fully labeled NTPs (Merck Sigma-Aldrich). High-Performance Liquid Chromatography Ultimate3000 uHPLC system (Thermo Scientific) was used to purify the product of interest from abortive transcripts in two chromatographic steps (Ion-Pair Reverse Phase and Ion-Exchange under denaturing conditions) further described the Supplementary Information Methods section. hsa-miR-34a-5p 3'-Cy3 labeled and single stranded mSirt1 trapped ES were purchased from Integrated DNA Technologies (IDT) as chemically synthesized RNA oligonucleotides purified by RNase free HPLC purification. A complete list of RNA and DNA sequences used in this work is available in Data S1 Tab 10.

hAgo2 protein sample preparation and RISC reconstitution

Human Argonaute 2 (hAgo2) cloned into pFastBac HT plasmid was obtained as described in the original publication[52]. hAgo2 was expressed in Sf9 insect cells and purified from the clarified cell lysate through a nickel affinity chromatography and a gel-filtration chromatography step. The fractions containing hAgo2 were pooled together, concentrated and stored at -80 °C. Further details of hAgo2 sample preparation are described in details in the Supplementary Information Methods section.

Purified hAgo2 was incubated with a ~2-fold excess of in vitro transcribed miR-34a in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM Imidazole, and 0.5 mM TCEP supplemented with 10 μg/ml of Bovine Serum Albumin (BSA) (Sigma-Aldrich) for 6 h at 37 °C. The assembled RISC (hAgo2-miR-34a complex) was then separated from unbound excess RNA using a gel filtration chromatography step. Loading of the guide miR-34a into RISC was assessed by an improved Northern Blot for the detection of small RNA[53,54]. Further experimental details of RISC reconstitution are described in the Supplementary Information Methods section.

Thermal denaturation monitored by UV absorption

Thermal denaturation monitored by UV absorption at 260 nm (A_{260}) was performed using an Evolution 260 Bio UV-Vis Spectrophotometer (Thermo Scientific) equipped with PCCU1 Peltier Control and Cooling Unit (Thermo Scientific). All samples were dissolved in NMR buffer (15 mM Sodium Phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 6.5). Fitting of the normalized differential melting curves (DMCs), as described in Supplementary Information Methods section, allowed for estimation of the melting temperature T_m and thermodynamic parameters presented in Extended Data Table 1 and Data S1 Tab 7.

Electrophoretic Mobility Shift Assay (EMSA)

hsa-miR-34a-5p 3'-Cy3 was incubated at a final concentration of 24 nM with increasing amounts of unlabeled single stranded partner (mSirt1, mSirt1 trapped ES or Complementary strand) in NMR buffer (15 mM Sodium Phosphate, 25 mM NaCl, 0.1 mM EDTA, pH 6.5) to a final volume of 10 μl. The total reaction volumes were mixed with 10 μl of 100% glycerol (Sigma-Aldrich) and subsequently loaded into a 10% non-denaturing Tris-Borate-EDTA (TBE) poly-acrylamide gel. Fluorescence signals relative to the free and bound form of hsa-miR-34a-5p 3'-Cy3 were quantified using ImageJ software[55]. Fitting of the binding curves to

a standard binding isotherm, as described in Supplementary Information Methods section, allowed for estimation of the equilibrium dissociation constants K_d presented in Extended Data Table 1 and Data S1 Tab 8.

Equilibrium filter binding assay

3'-Cy3 labeled target RNAs (mSirt1, mSirt1 trapped ES or scrambled control) were incubated at a constant concentration of 0.5 nM with increasing amounts of hAgo2-miR-34a complex in target binding buffer[6] (30 mM Tris-HCl pH 8.0, 100 mM Potassium Acetate, 2 mM Magnesium Acetate, 2.5 mM TCEP, 0.005% v/v NP-40 supplemented with 10 µg/ml yeast tRNA (Sigma-Aldrich) and 10 µg/ml BSA (Sigma-Aldrich)) to a final volume of 100 µl and incubated for 1h at 37 °C. After incubation, samples were readily applied to a DHM-48 dot-blot apparatus (Scie-Plas) and filtered through a nitrocellulose membrane Amersham Protran (GE Healthcare Life Sciences) and a positively charged nylon membrane Amersham Hybond-N+ (GE Healthcare Life Sciences). Fluorescence signals relative to the free (nylon) and protein bound (nitrocellulose) form of 3'-Cy3 target RNAs were quantified using ImageJ software[55]. Fitting of the binding curves to a standard binding isotherm, as described in Supplementary Information Methods section, allowed for estimation of the equilibrium dissociation constants K_d presented in Extended Data Table 1 and Data S1 Tab 9.

NMR spectroscopy

All NMR assignment and $R_{1\rho}$ relaxation dispersion (RD) experiments were acquired on a Bruker AVANCE III 600 NMR spectrometer operating at 600 MHz for ^{1}H, equipped with a cryogenically cooled QCI probe.

Sequence-specific resonance assignment – Sequence-specific resonance assignment experiments were performed on ^{13}C, ^{15}N fully labeled RNA samples dissolved in 15 mM

Na_2HPO_4/NaH_2PO_4, 25 mM NaCl, 0.1 mM EDTA, pH 6.5. Unless otherwise stated, assignment of aromatic ^{13}C2/5/6/8-^{1}H/2/5/6/8, sugar ^{13}C1'-^{1}H1' and imino ^{15}N1/3-^{1}H1/3 resonances was achieved using a standard set of ^{1}H-^{13}C, ^{1}H-^{15}N 2D HSQCs, 3D HCNs, HNN COSY and ^{1}H-^{1}H NOESY (all acquired using 175 ms mixing time) NMR experiments as described in Fürtig *et al.*,[56] recorded at different temperatures (9.0 °C, 22.4 °C and 35.9 °C) (Fig. S1, S3 and S4). For miR-34a–mSirt1 duplex, only a reduced set of imino ^{15}N1/3-^{1}H1/3 resonances were assigned using ^{1}H-^{15}N 2D HSQCs, HNN COSY and ^{1}H-^{1}H NOESY NMR experiments (Fig. S2). Assigned CSs were deposited to the Biological Magnetic Resonance Bank (BMRB)[17] for hsa-miR-34a-5p (entry 27225), miR-34a–mSirt1 bulge (entry 27226) and miR-34a–mSirt1 trapped ES (entry 27229).

^{1}H, ^{13}C and ^{15}N $R_{1\rho}$ RD – ^{1}H, ^{13}C and ^{15}N $R_{1\rho}$ RD NMR experiments were acquired as previously described in the respective original publications[21-23], using ^{13}C, ^{15}N fully labelled (^{13}C and ^{15}N $R_{1\rho}$) or natural abundance unlabeled (^{1}H $R_{1\rho}$) RNA samples dissolved in 15 mM Na_2HPO_4/NaH_2PO_4, 25 mM NaCl, 0.1 mM EDTA, pH 6.5. In brief, for each spin-lock power (ω_{SL}), data points were recorded as a function of different relaxation delays (T_{EX}). For each residue, variable delay (VD) lists were optimized in order to achieve a maximum decay of 1/3 of the starting peak intensity (T_{EX} = 0 s). To account for a reduced loss in peak intensity for large offsets (Ω $2\pi^{-1}$), a subset of off-resonance datasets was recorded with an extended VD list comprising longer maximal T_{EX} values, care was taken, that no additional heating occurred. In all datasets, data points with signal-to-noise ratio (S/N) < 20 for ^{1}H, ^{13}C and S/N < 10 for ^{15}N were discarded. VD lists for each dataset are reported in Supplementary Data S1.

Peak intensities were extracted from deconvoluted 1D datasets and plotted as a function of T_{EX}. $R_{1\rho}$ values were obtained from fitting of the data to a mono-exponential decay and error estimates were computed as one standard deviation (s.d.) using a Monte Carlo simulation

method[57] with 500 iterations. Potential artefacts (e.g. arising from Hartmann-Hahn matching conditions or strong ^{1}H-^{1}H and ^{13}C-^{13}C homo-nuclear coupling that results in deviation from mono-exponential behavior, were excluded from subsequent analysis by discarding exponential fits with $R^2 < 0.985$. $R_{1\rho}$ values as a function of ω_{SL} were subsequently fitted, using the Laguerre approximation[58] (Supplementary Methods Eq. (4)) assuming absence of exchange ($R_{ex} = 0$), fast exchange regime (Reduced Laguerre form where $k_{ex} >> \Delta\omega$) (Supplementary Methods Eq. (5)), 2-state exchange (Supplementary Methods Eq. (6)) or a 3-state exchange (Supplementary Methods Eq. (7–8)) using the models and fitting methods further described in the Supplementary Information Methods section. Selection of the best fitting model for each dataset was performed using a statistical F-test[27]. Degrees of freedoms were calculated as the number of data points minus the number of fitted parameters for each model. Fitted parameters, reduced χ^2 values resulting from the fit and exact *p*-values from the statistical F-test for each dataset are reported in Data S1.

Global fitting was performed assuming the presence of one collective exchange process to a minor populated state (ES^G) characterized by the global parameters k_{ex}^G (global exchange rate) and p_{ES}^G (population of ES^G) shared across the datasets. Each dataset was fitted using the best fitting model resulting from the individual fits and the fitted parameters as initial guesses for the global fit using a 2-state (Supplementary Methods Eq. (6)) and a 3-state (Supplementary Methods Eq. (7–8)) exchange model. For those dataset, globally fitted using 3-state exchange model (Supplementary Methods Eq. (7–8)), we assigned one excited state to the global fit (ES^G) while leaving the fitting of the parameters relative to the second state (k_{EX2}, pop_{ES2} $\Delta\omega_{ES2}$) unconstrained during the fit, fundamentally equivalent to fit them individually. Error estimates of the fitted global parameters were computed as one s.d. using a Monte Carlo simulation method[57] with 500 iterations. Selection of the best fitting model was performed using a statistical F-test[27]. Degrees of freedoms were calculated as the number

of data points minus the number of fitted parameters for each model. Fitted global parameters, reduced χ^2 values resulting from the fit and exact *p*-values from the statistical F-test for the global fittings are reported in Data S1.

Exponential fittings, individual and global fittings and model selection were performed using an *in-house* written Python (2.7) code (https://www.python.org/) available upon request.

Secondary structure prediction

All secondary structure predictions were performed using MC-Fold 1.6.0[25], unless otherwise stated, providing as input the nucleotide sequence of each construct. Structures consisting of two strands where simulated by using a UUCG connection loop.

Chemical-shift distribution of G:C and G:U base-pairs

PDB IDs and nucleotide numbers of Guanosines involved in either G:C or G:U base-pairs were obtained using RNA FRABASE 2.0[59]. PDB ID matching BMRB entries and were selected using the "Matched submitted BMRB-PDB entries" list. Subsequently, chemical-shifts from ^{1}H1-^{15}N1 assigned couples only, were extracted from the BMRB entries, duplicates and mis-referenced couples were removed. A total of 303 G:C and 63 G:U unique ^{1}H1-^{15}N1 couples were obtained (Fig. 1).

All atom, explicit solvent Molecular Dynamics simulations

Atomistic simulations of miR34a-mSirt1 bulge, were initialized using starting structures generated by MC-Fold and MC-Sym[25]. All-atom, explicit solvent molecular dynamics simulations were performed using GROMACS 5.0.7[60] and the modified Chen-Garcia force field for RNA[61] including backbone phosphate modifications of Steinbrecher *et al.,*[62]. The structure was solvated with 6,664 TIP4P-Ew[63] waters in a 6.1 nm cubic box, and salt

conditions of 1 M excess, KCl were represented by 161 K+ and 134 Cl- ions using the activity-coefficient calibrated parameters of Cheatham and Joung[64]. In order to enhance exploration of diverse bulge conformations using temperature replica-exchange without inadvertently inducing RNA melting, five harmonic restraints were assigned with a force constant of 500 kJ mol^{-1} nm^{-2} on the middle H-bond of the three initial G:C base-pairs and C14:G19 and G13:C20 (tG25:gC4) in the seed region, which are all observed to be well formed under NMR experimental conditions of 9 – 35.9 °C. The initial structures were energy minimized and equilibrated at 1 atm constant pressure with random initial velocities drawn from a Boltzmann distribution. Using Replica Exchange Molecular Dynamics (REMD), 24 individual replicas spanning a temperature range of 77 – 147 °C were simulated to evaluate the conformational flexibility of miR34a–mSirt1 bulge. The exchange rate was 25% with attempted temperature swaps every 1000 steps (2 ps), which is also how often coordinates were saved. Once equilibrated, production simulations were propagated for ~670 ns per replica, a total of 16.08 μs of cumulative simulation time. Structural clustering based on all-heavy-atom RMSD was accomplished utilizing the Daura *et al*., algorithm[65] on 30,000 evenly spaced snapshots taken from the lowest temperature replica (27 °C) using a cut-off of 5.0 Ångströms. The most highly populated cluster, which contains >60% of all structures in the 27 °C replica, is reported as the GS ensemble (Fig. 2g). A separate set of REMD simulations consisting of 25 replicas spanning 25 – 77 °C were also simulated using identical settings as above. Each replica sampled for 478 ns for a cumulative total of 11.95 μs of sampling, and identical cluster analysis was performed on the 25 °C replica. Details on REMD simulations of miR-34a bulge ES and trapped ES, as well as inter-helical bending angle distributions are further described in Supplementary Information Methods section.

Alignment of GS/ES ensembles into AGO2 crystal structure – 250 randomly picked snapshots from each REMD ensemble (GS, ES and trapped ES) were initially aligned into the 4W5T[6]

PDB structure**.** Each simulation structure was aligned such that the backbone phosphates positions of bases g2-g8 to matched those of the crystal structure. For visual clarity, only 20 of the 250 conformations are graphically depicted in Fig. S10.

Slow-growth simulations of GS/ES insertion into hAgo2 complex – In order to ascertain the ability of the hAgo2 protein to physically accommodate the miR-34a–mSirt1 RNA complexes in the GS and ES states, representative snapshots from each ensemble were inserted into the hAgo2 protein using slow-growth binding simulations[66]. Starting with the 4W5O[6] PDB structure, the existing partial miRNA–mRNA complex was deleted and missing hAgo2 amino acids were modelled in. The UUCG tetraloop used to anchor the NMR construct was mutated in-place to match the native miR-34a–mSirt1 seed sequence, and the initial RNA conformation was determined by aligning the backbone positions of bases g2-g8 to match the crystallographic RNA seed helix. The RNA was then inserted using a slow-growth process in which RNA-protein VdW and electrostatic interactions were completely decoupled at t = 0 s, and then linearly increased to 100% interaction in a 100 ps stochastic dynamics simulation at 47 °C with 1 fs time-steps. This method only succeeds if the RNA could be accommodated by flexing of the protein to resolve minor steric overlaps. Successful slow-growth attempts were then solvated in explicit solvent and ions, minimized, and simulated for a ~10 ns N,P,T simulation at 25 °C and 1 atm. The conformations shown in Fig. 3e-g are from the final frames of these simulations.

Dual Luciferase Reporter (DLR) assay

Plasmids – All mRNA targets dual luciferase reporter (DLR)[67] plasmids were generated by cloning a synthetic dsDNA (Data S1 Tab 10a) into the *XhoI* and *NotI* restriction sites of psiCHECK2-miR-34 WT[68]. Fully complementary binding site is the unmodified psiCHECK2-miR-34 MT plasmid[68]. As a negative control, the mutated hsa-miR-34a-5p

binding site of psiCHECK2-miR-34 MT was used[68]. These plasmids were a gift from Joanne Weidhaas (Addgene plasmid #78258 and #78259). The newly generated plasmids were verified by sequencing.

Cell lines and culture – HEK293T cells (CRL-11268) were obtained from ATCC and cultured in Dulbecco's Modified Essential Medium (DMEM, Gibco) supplemented with 10% Fetal Calf Serum (FCS, Gibco).

DLR Assay – HEK293T cells were seeded 24 h prior to transfection in 12 well plates. Cells were transfected at 70 - 90% confluency with 1.6 µg of plasmid DNA and with or without 40 pmol of hsa-miR–34a-5p/hsa-miR–34a-3p (guide/passenger) duplex using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. After 24 h, cells were washed with Phosphate Buffer Saline (PBS) once and luciferase activity was measured with the DLR assay system (Promega) according to the manufacturer's protocol on a Promega GloMax 96 microplate luminometer using a 1s delay and 10s integration time. For each sample the signal corresponding to the Renilla luciferase activity was acquired and normalized relative to the Firefly luciferase signal. Samples without co-transfected miR-34a were set to 100% and downregulation of samples co-transfected with miR-34a was calculated based on this. Results show the average and standard deviation of at least three independent biological replicates. For statistical analysis, a paired, two-sided t-test was performed. Error bars represent one s.d. ** $p \leq 0.01$, *$p \leq 0.05$.

Predicted target screening of GC to GU switches

In total 28'653 3'-UTR sequences, including all isoforms, of all 19'432 protein-coding genes for human were downloaded from TargetScan[18]. The sequences were bioinformatically screened for putative mir-34a targets using regular expression. Specifically, sequences were selected that included the reverse complement sequence of a canonical 6mer-A1, followed by

a U or C as first nucleotide of the bulge. Thereafter, to allow for a bulge of up to 6 nucleotides, the sequence was unspecified for position 1-5 and the bulge was closed with a C base-pairing with the G from the miR-34a leading to this conformational switch model (5'-'C[A,G,U,C]{1,5}[U,C]ACUGCCA' – 3').

Each of the 532 mRNA targets (593 with all isoforms) were screened according to the possibility to form different bulges sizes (from 1 to 5 nts) with a G-C or a G-U as closing base pair. Thereafter, the secondary structure of each mRNA-UUCG-miR34a complex was simulated using MC-Fold 2.32[25]; different mRNA lengths were tested, until maximum 8 nts were added to a mRNA sequence of 22nts. Each length was screened to identify examples of ground and excited states similar to those observed for Sirt1 and defined according to the following structural features. A Ground State (GS) was defined having: 1) a non-base-paired U (position t21 in mSirt1) after the seed, followed by a number of unpaired bases equaling the length of the bulge; 2) a GC Watson-Crick base pair closing the bulge, followed by 2 base pairs, in the 3'-helix (Fig 4c sketch), cluster 1. 3) A second more stringent cluster (cluster 2) is described by 2 additional base pairs, Watson-Crick in addition, instead after the GC closing base pair. An Excited State (ES) was defined having the U (position t21in mSirt1) after the seed pairs with the G in position gG8 (in miR-34a). For obvious structural reason, in all clusters, we excluded structures where the miR-34a sequence was folding onto itself or where shortening of the seed was occurring. Only sequences were considered if the GS and the ES were present for at least 3 different lengths and all the lengths have at least a GS or an ES. Of the 5 target tested, only CCND1 and ATG9A were previously confirmed a miR-34a targets[41,42].

Data availability

NMR sequence specific resonance assignment of hsa-miR-34a-5p (entry 27225), miR34a–mSirt1 bulge (entry 27226) and miR34a–mSirt1 trapped ES (entry 27229) constructs were deposited in the BMRB. The plasmids used in this work for the DLR assay were a gift from Joanne Weidhaas (Addgene plasmid #78258 and #78259). All data and code used for data analysis are available upon request. The ensembles of REMD simulations have been deposited to the PDB-DEV #264 (https://pdb-dev.wwpdb.org).

Supplementary Information Supplementary Information including Supplementary Discussion, Supplementary Methods, Figures, Tables and Data can be found in the attached documents.

Acknowledgments The authors acknowledge the members of the Petzold lab and the Emma R. Andersson lab for the insightful discussions. Furthermore, the authors acknowledge the time and effort and valuable comments of the anonymous reviewers. K.P. acknowledges funding from the Swedish Research Council (2014-04303), the Swedish Foundation for Strategic Research (Project No. ICA14-0023), Harald och Greta Jeansson Stiftelse (JS2014-0009), Eva och Oscar Ahréns Stiftelse, Åke Wiberg Stiftelse (#467080968 & M14-0109), Cancerfonden (CAN 2015/388), Karolinska Institute and the department of Medical Biochemistry and Biophysics (KID 2-3707/2013 & the support for the purchase of a 600MHz Bruker NMR spectrometer) and Ragnar Söderberg Stiftelse (M91/14). A.A.C. acknowledges support from National Science Foundation grant MCB1651877. B.F. acknowledges funding from the Strategic Research Area (SFO) program of the Swedish Research Council (VR) through Stockholm University. This work used the computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE) [allocation TG-MCB140273 to AAC], which is supported by National Science Foundation grant number ACI-1548562. J.S. acknowledges funding through a Marie Sklodowska-Curie IF (EU H2020 / Project No. 747446). We kindly acknowledge Prof. Ian MacRae, Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA for providing us with plasmid expressing hAgo2 as gift. We acknowledge the Lund University Protein Production Platform (LP3) (Lund, Sweden) and the Protein Science Facility at the Karolinska Institute department of Medical Biochemistry and Biophysics (Stockholm, Sweden) for the help with hAgo2.

Author contributions: L.B. and K.P. conceived the project. L.B. carried out the majority of the experiments and data analysis with assistance from I.G, E.S., S.F.S., K.P. and L.S.. P.E. and A.A.C. carried out the REMD simulation and slow-growth docking. I.G. and B.F. performed the target screening of GC-to-GU switches. J.S. and C.F. provided advice in design and analysis. L.B. K.P. and A.A.C. wrote the manuscript and all authors contributed to the final version.

Author Information The authors declare no competing interests. Correspondence and requests for materials should be addressed to K.P. (katja.petzold@ki.se).

Extended Data and Supplementary Figure Captions

Extended Data Table 1 | T_m and K_d fitted parameters. a, Thermal denaturation monitored by UV absorption. Parameters derived from fitting of Supplementary Equation (1a or 1b) as described in the Supplementary Methods section. T_m and h obtained from the fit are presented with their 95% CI as estimate of the experimental error. Mean T_m is presented with the associated s.d. Complete fitting details and statistics are presented in Data S1 Tab 7. **b**&**c**, EMSA and FBA. Parameters derived from fitting of Supplementary Equation (2) as described in the Supplementary Methods section. K_d values obtained from the fit are presented with their 95% CI as estimate of the experimental error. Complete fitting details and statistics are presented in Data S1 Tab 8 and 9. **d**, Northern blot. The fraction of hAgo2 loaded with the guide RNA of interest was estimated using Northern blot as described in the Supplementary Methods section.

Extended Data Figure 1 | Secondary structure prediction using MC-Fold. Secondary rearrangements among the 10 lowest energy structures are calculated using MC-Fold[25]. Ranking according to the predicted energy difference from the MFE is indicated in each label ($\Delta\Delta G(n)$ in unreferenced kcal mol^{-1} as described in the MC-Fold original paper[25]). Secondary structures with single base pair opening in the cUUCGg region were omitted from the representation. **a**, miR-34a (red) and mSirt1 (blue) duplex connected by a cUUCGg loop (black). The MFE correspond to an A1-7mer binding site. Suboptimal structures (3) and (5) suggest possible modulation of the binding site to a 8mer GU and an A1-6mer configuration, respectively. **b**, miR-34a (red) mSirt1 (blue) bulge construct comprising an cUUCGg loop and a closing stem (black). The secondary structure distribution of miR-34a–mSirt1 bulge follow the same trend as the full-length duplex, where dashed lines connect identical bulge structures. Suboptimal structures were used to validate or reject models of excited states (ES)

secondary structures based on $R_{1\rho}$ NMR RD data: (1) MFE correspond the assigned ground state structure (GS). (3) Satisfies the ^{1}H1 and ^{15}N1 $R_{1\rho}$ NMR RD data on gG6(G24) being G:U base paired with tU20(U9). (5) is mutually exclusive with (3) in structural terms and satisfies the ^{13}C $R_{1\rho}$ NMR RD data measured on tA19 that indicate this residue adopting a base paired conformation. Therefore (3) was proposed as ES1 and (5) as a second ES2. Conformations (6) and (7) are not agreeing and partially clashing with our $R_{1\rho}$ NMR RD data and can therefore be excluded to be an excited state. **c**, miR-34a (red) and mSirt1 (turquoise) trapped ES duplex connected by a cUUCGg loop (black). Substituted nts used for trapping the ES are highlighted in yellow. The MFE correspond to a 8mer binding site. **d**, miR-34a (red) mSirt1 (turquoise) trapped ES construct comprising an cUUCGg loop and a closing stem (black). Substituted nts used for trapping the ES are highlighted in yellow. The secondary structure distribution of miR-34a–mSirt1 trapped ES follow a similar trend as the full-length duplex, where dashed lines connect identical bulge structures.

Extended Data Figure 2 | Mutate-and-Chemical-Shift-Fingerprint (MCSF) analysis of miR-34a–mSIRT1 bulge and trapped ES, and analysis of 13C tA22C8 outliers. We used the MCSF approach[26] to cross-validate the candidate ES (ES1) modelled using $R_{1\rho}$ derived GS→ES chemical shift differences (**a** ^{13}C $R_{1\rho}$ Δω data, blue dots and **b**, left). We generated an ES1 mimic (trapped ES) using a two-point substitution predicted to stabilize the proposed conformation (**b**, bottom). For each reporter, we compared ^{13}C $R_{1\rho}$ Δω data with the chemical shift differences derived from the assignment of the bulge and the trapped ES constructs (**a**, ^{13}C Δω trapped ES data, turquoise dots). **a**, The MCSF analysis validates our ES1 model (green boxes) with exceptions arising from the limitation of the mimic (orange boxes) and from the presence of a second ES (ES2 b, right) (violet boxes). **b**, The proposed model for ES2 satisfies the ^{13}C $R_{1\rho}$ Δω data measured for tA19 and gG6 and **c** allows us to draw a free

energy landscape for the entire star-like 3-state exchange process. Additional discussion of the MCSF analysis and ES2 can be found in Supplementary Discussion 5. Since the transition coeffient (κ), was assumed to be 1[27], the transition states energy (TS1 and TS2), calculated using Supplementary Eq. (11), has to be considered an upper limit of this exchange process. **e&f**, The substitution site (tU21→ tC21) perturbs the chemical environment of tA22C8 that is directly neighboring (3' top) the substituted nucleobase (orange sphere in **b**). Conversely tA22C2 (green sphere), pointing towards the miR-34a strand (red) experience an equivalent chemical environment in bulge (blue) and trapped ES (turquoise) constructs. Thus, explaining the inconsistency in the MCSF profile for tA22C8 (Fig. S12a, orange box). **e**, Secondary structure of tA22 in the miR-34a–mSirt1 bulge excited state (left) and trapped ES (right) constructs. The substitution site (tU21→ tC21) is highlighted (grey wavy line). **f**, Overlay of average structures of the bulge ES (blue) and trapped ES (turquoise) from REMD ensembles, aligned according to residues gU7 and tA22. Residues gU7, gG8, tU21 and tA22 are shown. tA22C8 (orange) and tA22C2 (green) ^{13}C atoms are shown.

Extended Data Figure 3 | Biophysical and biochemical characterization of the constructs. a, Individual A260 UV melting profiles for the constructs used in this study. miR-34a, miR-34a–mSirt1 duplex, miR-34a–mSirt1 trapped ES duplex and miR-34a–Complementary strand duplex were measured as three independent replicates. Normalized differential melting curves (δA260/δT) plotted as a function of Temperature (K) (circles) and fitted to either Supplementary Equation (1a) or (1b) (line), depending on the molecularity of the system. For each construct, each replicate is presented in a different shade of the corresponding colour. **b**, Electrophoretic Mobility Shift Assay (EMSA) titration profiles for miR-34a–mSirt1 duplex, miR-34a–mSirt1 trapped ES duplex and miR-34a–Complementary strand duplex measured as three independent replicates. Gel images were acquired by Cy3

fluorescence detection. During the titration miR-34a 3'-Cy3 was kept at a constant concentration of 24 nM, which set the sensitivity limit to estimate the dissociation constant (K_d) (Supplementary Figure 1 a–c). mSirt1 and mSirt1 trapped ES were found to be equivalent in their ability to form a stable RNA-RNA duplex with miR-34a. Tighter binding was observed for the complementary strand (48.4±9.5 nM), compared to the mSirt1 (124.3±21.7 nM) and mSirt1 trapped ES (110.3±23.0 nM) titrations provided a control for the dynamic range of K_d estimation with this technique. Bound/total miR-34a 3'-Cy3 ratio plotted as a function of titrant concentration (circles) and fitted a standard binding isotherm (line) (Supplementary Equation (2). **c**, Equilibrium Filter Binding Assay profiles for mSirt1, mSirt1 trapped ES and a scrambled control binding miR-34a-loaded hAgo2. The three targets were measured as three independent replicates and fitted to a standard binding isotherm (line) (Supplementary Equation (2). The 95% Confidence Interval is presented as estimate of the experimental error. Similarly to the EMSA, mSirt1 and mSirt1 trapped ES were found to be equivalent in their ability to form a stable ternary complex within RISC. The simulated dataset (dotted line) indicate curves corresponding to K_d values ten times lower (red) or ten times higher (green) than the average value for mSirt1, mSirt1 trapped ES, providing a frame for the amplitude of our experimental error. **d**, Northern blot for the detection of miR-34a loaded in hAgo2. A standard calibration curve (naked miR-34a) was used to obtain an estimate of miR-34a in RISC. R^2 from the calibration curved is shown, whereas the two outer curves indicate the 95% Confidence interval of the calibration line fit. The average ratio of hAgo2 (pmole) and miR-34a loaded (pmole), was used to obtain a fraction of hAgo2 loaded with our guide (~1.5%). The complete list of fitted parameters for UV melting, EMSA titration, FBA titration and Northern blot can be found in Table S1a, b, c and d, respectively. The complete fitting analysis of UV melting, EMSA titration, FBA titration can be found in Data S1 Tab 7, 8 and 9, respectively.

Extended Data Figure 4 | DLR assay of additional miR-34a targets. 5 targets of different bulge sizes were studied as described in the Materials and Method section of the main manuscript. **a**, Standard DLR normalization (relative to control condition where no miR-34a duplex is transfected) is show. Despite the large variability between replicates, consistent increase in downregulation is observed among replicates for WT and tES constructs (connecting lines). **b**, When the dataset are internally normalized and the WT condition is set to 100% (mean value), the variation due to experimental replicas is attenuated and the trend observed in a is maintained.

Extended Data Figure 5 | miR-34a–mSirt1 bulge Mg2+ titration followed by NMR. HSQCs overlay of different Mg2+ titration steps. **a**, ^{1}H-^{13}C aromatic 2/6/8 HSQC. **b**, ^{1}H-^{13}C sugar 1' HSQC. **c**, ^{1}H-^{15}N imino 1/3 HSQC. The titration steps are color-coded as indicated by the inset (**a**, top left). Additional overlay of the miR-34a–mSirt1 trapped ES is presented in grey in **a** and **b**. The arrows indicate the chemical shift trajectory during the titration. The dashed lines connect equivalent peaks in the miR-34a–mSirt1 bulge and trapped ES constructs.

Extended Data Figure 6 | Slow-growth insertion of ES-RNA into hAgo2 predicts ability of bulged miR/mRNA complexes to access an alternate dsRNA binding mode of hAgo2. Comparison of slow-growth induced-fit hAgo2 structures with existing X-ray structures via structural alignment. **a**, hAgo2 after induced-fit with ES, RNA binds in the PIWI-adjacent groove rather than PAZ domain. **b**, *T. thermophilus* Ago crystal structure reported by Patel and coworkers[69] similarly shows an DNA/RNA duplex binding in the analogous PIWI-

adjacent groove (PDB 3hm9). **c–j**. RMSD for each indicated pair of Ago structures were measured after structural alignment of either all protein atoms (all) or excluding the PAZ, PIWI loops and helix-7 atoms (subset aligned) - these atoms still count towards RMSD. The subset-aligned structures are shown to illustrate that most of the RMSD difference arises from pivoting motions of the PAZ domain coupled with small shifts in helix 7 and PIWI loops to accommodate the inserted ES RNA structures. **c**, GS (yellow) and 4ola (pink) RMSD: 2.065Å (all) and 2.62Å (subset aligned). **d**, ES (green) and 4ola (pink) RMSD: 1.4Å (all) and 1.65Å (subset align). **e**, Trapped ES (bright pink) and 4ola (dark pink) RMSD 1.9Å (all) 2.18 (subset align). **f**, ES (green) and GS (yellow) RMSD: 2.1Å (all) and 2.2 (subset align). **g**, Trapped ES (pink) and ES (green) RMSD: 1.6Å (all) 1.33Å (subset align). **h**, ES (green) and 6n4o (gray) RMSD: 2.05Å (all) and 2.065 Å (subset align). **i**, ES (green) and 3hm9 (gray) RMSD 4.52Å (all). **j**, GS(yellow) and 3hm9 (gray) RMSD: 3.85Å (all).

Supplementary Figure 1 | Uncropped images of EMSA gels and Northern Blot membrane. a–c, Uncropped gel images acquired as described in Supplementary Information Methods section. For each target RNA strand three, independently performed, replicates are presented. Values obtained from the gel presented are plotted in Fig. 3c and Extended Data Fig. 3b. Red arrows indicate free miR-34a 3'-Cy3 labelled (22 nts), blue arrows indicate target bound miR-34a 3'-Cy3 labelled (hetero-duplexes). **d**, Uncropped image of Northern Blot positively-charged nylon membrane aquired as described in Supplementary Information Methods section. Values obtained from the membrane presented here were used to estimate hAgo2 loading with miR-34a and presented in Extended Data Table 1.

References

1. Elkayam, E. *et al.* The Structure of Human Argonaute-2 in Complex with miR-20a. *Cell* **150,** 100–110 (2012).
2. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336,** 1037–1040 (2012).
3. Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. Structure of yeast Argonaute with guide RNA. *Nature 2012 491:7426* **486,** 368 (2012).
4. Grimson, A. *et al.* MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Molecular Cell* **27,** 91–105 (2007).
5. Nielsen, C. B. *et al.* Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* **13,** 1894–1910 (2007).
6. Schirle, N. T., Sheu-Gruttadauria, J. & MacRae, I. J. Structural basis for microRNA targeting. *Science* **346,** 608–613 (2014).
7. Klum, S. M., Chandradoss, S. D., Schirle, N. T., Joo, C. & MacRae, I. J. Helix-7 in Argonaute2 shapes the microRNA seed region for rapid target recognition. *EMBO J.* **37,** 75–88 (2018).
8. Filipowicz, W., Bhattacharyya, S. N. & Sonenberg, N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature reviews genetics* **9,** 102 (2008).
9. Sheu-Gruttadauria, J., Xiao, Y., Gebert, L. F. & MacRae, I. J. Beyond the seed: structural basis for supplementary microRNA targeting by human Argonaute2. *The EMBO journal* e101153 (2019).
10. Wee, L. M., Flores-Jasso, C. F., Salomon, W. E. & Zamore, P. D. Argonaute Divides Its RNA Guide into Domains with Distinct Functions and RNA-Binding Properties.(Report). *Cell* **151,** 1055 (2012).
11. Bommer, G. T. *et al.* p53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Current biology* **17,** 1298–1307 (2007).
12. Tarasov, V. *et al.* Differential Regulation of microRNAs by p53 Revealed by Massively Parallel Sequencing: miR-34a is a p53 Target That Induces Apoptosis and G1-arrest. *Cell Cycle* **6,** 1586–1593 (2007).
13. Chang, T.-C. *et al.* Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Molecular Cell* **26,** 745–752 (2007).
14. He, L. *et al.* A microRNA component of the p53 tumour suppressor network. *Nature 2012 491:7426* **447,** 1130 (2007).
15. Raver-Shapira, N. *et al.* Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. *Molecular Cell* **26,** 731–743 (2007).
16. Yamakuchi, M., Ferlito, M. & Lowenstein, C. J. miR-34a repression of SIRT1 regulates apoptosis. *Proceedings of the National Academy of Sciences* **105,** 13421–13426 (2008).
17. Ulrich, E. L. *et al.* BioMagResBank. *Nucleic Acids Research* **36,** D402–D408 (2007).
18. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *elife* **4,** e05005 (2015).
19. Dallmann, A. *et al.* Efficient detection of hydrogen bonds in dynamic regions of RNA by sensitivity-optimized NMR pulse sequences. *Angew. Chem. Int. Ed.* **52,** 10487–10490 (2013).
20. Vecenie, C. J. & Serra, M. J. Stability of RNA Hairpin Loops Closed by AU Base Pairs †. **43,** 11813–11817 (2004).
21. Nikolova, E. N., Gottardo, F. L. & Al-Hashimi, H. M. Probing Transient Hoogsteen

Hydrogen Bonds in Canonical Duplex DNA Using NMR Relaxation Dispersion and Single-Atom Substitution. *J. Am. Chem. Soc.* **134,** 3667–3670 (2012).
22. Hansen, A. L., Nikolova, E. N., Casiano-Negroni, A. & Al-Hashimi, H. M. Extending the Range of Microsecond-to-Millisecond Chemical Exchange Detected in Labeled and Unlabeled Nucleic Acids by Selective Carbon R 1ρNMR Spectroscopy. *J. Am. Chem. Soc.* **131,** 3818–3819 (2009).
23. Schlagnitweit, J., Steiner, E., Karlsson, H. & Petzold, K. Efficient Detection of Structure and Dynamics in Unlabeled RNAs: The SELOPE Approach. *Chem. Eur. J.* **24,** 6067–6070 (2018).
24. Steiner, E., Schlagnitweit, J., Lundström, P. & Petzold, K. Capturing Excited States in the Fast-Intermediate Exchange Limit in Biological Systems Using 1H NMR Spectroscopy. *Angew. Chem. Int. Ed.* **55,** 15869–15872 (2016).
25. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature 2012 491:7426* **452,** 51 (2008).
26. Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A. & Al-Hashimi, H. M. Visualizing transient low-populated structures of RNA. *Nature 2012 491:7426* **491,** 724–728 (2012).
27. Xue, Y. *et al.* in *Laboratory Methods in Enzymology: RNA* (eds. Woodson, S. A. & Allain, F. H. T.) **558,** 39–73 (Academic Press, 2015).
28. Clay, M. C., Ganser, L. R., Merriman, D. K. & Al-Hashimi, H. M. Resolving sugar puckers in RNA excited states exposes slow modes of repuckering dynamics. *Nucleic Acids Research* **45,** e134–e134 (2017).
29. Ebrahimi, P., Kaur, S., Baronti, L., Petzold, K. & Chen, A. A. A two-dimensional Replica-Exchange Molecular Dynamics method for simulating RNA folding using sparse experimental restraints. *Methods* (2019).
30. Li, P. & Merz, K. M., Jr. Metal ion modeling using classical mechanics. *Chemical reviews* **117,** 1564–1686 (2017).
31. Broughton, J. P., Lovci, M. T., Huang, J. L., Yeo, G. W. & Pasquinelli, A. E. Pairing beyond the Seed Supports MicroRNA Targeting Specificity. *Molecular Cell* **64,** 320–333 (2016).
32. Brancati, G. & Großhans, H. An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity. *Nucleic Acids Research* **46,** 3259–3269 (2018).
33. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. *Nat Struct Mol Biol* **11,** 576 (2004).
34. Wang, Y., Li, Y., Ma, Z., Yang, W. & Ai, C. Mechanism of microRNA-target interaction: molecular dynamics simulations and thermodynamics analysis. *PLoS Comput Biol* **6,** e1000866 (2010).
35. Kong, R. *et al.* Exploring the RNA-bound and RNA-free human Argonaute-2 by molecular dynamics simulation method. *Chemical biology & drug design* **90,** 753–763 (2017).
36. Bhandare, V. & Ramaswamy, A. Structural dynamics of human argonaute2 and its interaction with siRNAs designed to target mutant tdp43. *Advances in bioinformatics* **2016,** (2016).
37. Wang, Y. *et al.* Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature 2012 491:7426* **456,** 921 EP –
38. Sheng, G. *et al.* Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proceedings of the National Academy of Sciences* **111,** 652 (2014).
39. Doxzen, K. W. & Doudna, J. A. DNA recognition by an RNA-guided bacterial

Argonaute. *PLOS ONE* **12,** e0177097 (2017).
40. Liu, Y. *et al.* Accommodation of Helical Imperfections in Rhodobacter sphaeroides Argonaute Ternary Complexes with Guide RNA and Target DNA. *Cell Reports* **24,** 453–462 (2018).
41. Huang, J. *et al.* miR-34a Modulates Angiotensin II-Induced Myocardial Hypertrophy by Direct Inhibition of ATG9A Expression and Autophagic Activity. *PLOS ONE* **9,** e94382 (2014).
42. Sun, F. *et al.* Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest. *FEBS letters* **582,** 1564–1568 (2008).
43. Bartel, D. P. Metazoan MicroRNAs. *Cell* **173,** 20–51 (2018).
44. Gebert, L. & MacRae, I. J. Regulation of microRNA function in animals. *Nature Reviews Molecular Cell Biology* (2018).
45. Park, J. H., Shin, S.-Y. & Shin, C. Non-canonical targets destabilize microRNAs in human Argonautes. *Nucleic Acids Research* **45,** 1569–1583 (2017).
46. Jo, M. H. *et al.* Human Argonaute 2 Has Diverse Reaction Pathways on Target RNAs. *Molecular Cell* **59,** 117–124 (2015).
47. De, N. *et al.* Highly Complementary Target RNAs Promote Release of Guide RNAs from Human Argonaute2. *Molecular Cell* **50,** 344–355 (2013).

Methods References

48. Beckert, B. & Masquida, B. in *Rna* 29–41 (Springer, 2011).
49. Baronti, L., Karlsson, H., Marušič, M. & Petzold, K. A guide to large-scale RNA sample preparation. *Anal Bioanal Chem* **410,** 3239–3252 (2018).
50. Kao, C., Zheng, M. & Rüdisser, S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3′ terminus of RNAs transcribed by T7 RNA polymerase. *RNA* **5,** 1268–1272 (1999).
51. Helmling, C. *et al.* Rapid NMR screening of RNA secondary structure and binding. *Journal of biomolecular NMR* **63,** 67–76 (2015).
52. De, N. & MacRae, I. J. in *Argonaute Proteins* 107–119 (Springer, 2011).
53. Pall, G. S. & Hamilton, A. J. Improved northern blot method for enhanced detection of small RNA. *Nature Protocols* **3,** 1077 (2008).
54. Rio, D. C. Northern blots for small RNAs and microRNAs. *Cold Spring Harbor Protocols* **2014,** pdb–prot080838 (2014).
55. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature methods* **9,** 671 (2012).
56. Fürtig, B., Richter, C., Wöhnert, J. & Schwalbe, H. NMR spectroscopy of RNA. *Chembiochem* **4,** 936–962 (2003).
57. Metropolis, N. & Ulam, S. The monte carlo method. *Journal of the American statistical association* **44,** 335–341 (1949).
58. Palmer, A. G. & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chemical reviews* **106,** 1700–1719 (2006).
59. Popenda, M. *et al.* RNA FRABASE 2.0: an advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC bioinformatics* **11,** 231 (2010).
60. Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2,** 19–25 (2015).
61. Chen, A. A. & Garcia, A. E. High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Proceedings of the National*

Academy of Sciences **110,** 16820–16825 (2013).
62. Steinbrecher, T., Latzer, J. & Case, D. A. Revised AMBER parameters for bioorganic phosphates. *Journal of chemical theory and computation* **8,** 4405–4412 (2012).
63. Horn, H. W. *et al.* Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of chemical physics* **120,** 9665–9678 (2004).
64. Joung, I. S. & Cheatham, T. E., III. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *The Journal of Physical Chemistry B* **112,** 9020–9041 (2008).
65. Daura, X. *et al.* Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Ed.* **38,** 236–240 (1999).
66. Hu, H., Yun, R. H. & Hermans, J. Reversibility of Free Energy Simulations : Slow Growth May Have a Unique Advantage.(With a Note on Use of Ewald Summation). *Molecular Simulation* **28,** 67–80 (2002).
67. Grentzmann, G., Ingram, J. A., Kelly, P. J., GESTALAND, R. F. & Atkins, J. F. A dual-luciferase reporter system for studying recoding signals. *RNA* **4,** 479–486 (1998).
68. Salzman, D. W. *et al.* miR-34 activity is modulated through 5′-end phosphorylation in response to DNA damage. *Nature Communications* **7,** 10954 (2016).
69. Wang, Y. *et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature 2012 491:7426* **461,** 754 (2009).