Base-pair conformational switch modulates miR-34a targeting of Sirt1 mRNA

https://doi.org/10.1038/s41586-020-2336-3

Received: 28 September 2018

Accepted: 12 March 2020

Published online: 27 May 2020

Check for updates

Lorenzo Baronti[1], Ileana Guzzetti[1,6], Parisa Ebrahimi[2,6], Sarah Friebe Sandoz[1,6], Emilie Steiner[1,4,6], Judith Schlagnitweit[1], Bastian Fromm[3], Luis Silva[1], Carolina Fontana[1,5], Alan A. Chen[2] & Katja Petzold[1]✉

MicroRNAs (miRNAs) regulate the levels of translation of messenger RNAs (mRNAs). At present, the major parameter that can explain the selection of the target mRNA and the efficiency of translation repression is the base pairing between the 'seed' region of the miRNA and its counterpart mRNA[1]. Here we use $R_{1\rho}$ relaxation-dispersion nuclear magnetic resonance[2] and molecular simulations[3] to reveal a dynamic switch—based on the rearrangement of a single base pair in the miRNA–mRNA duplex—that elongates a weak five-base-pair seed to a complete seven-base-pair seed. This switch also causes coaxial stacking of the seed and supplementary helix fitting into human Argonaute 2 protein (Ago2), reminiscent of an active state in prokaryotic Ago[4,5]. Stabilizing this transient state leads to enhanced repression of the target mRNA in cells, revealing the importance of this miRNA–mRNA structure. Our observations tie together previous findings regarding the stepwise miRNA targeting process from an initial 'screening' state to an 'active' state, and unveil the role of the RNA duplex beyond the seed in Ago2.

MicroRNAs—non-coding RNA molecules—regulate gene expression by targeting mRNAs. Each mature miRNA of roughly 22 nucleotides is bound to one Argonaute protein (Ago1 to Ago4 in humans), forming an RNA-induced silencing complex (RISC). In the RISC, nucleotides 2–6 of the guide miRNA (g2–g6) are prearranged to recognize mRNA targets through Watson–Crick base pairing[6–8] in the seed (Fig. 1a, b). This base-pair complementarity (involving up to g2–g8) largely determines RISC activity[9,10]: for example, complementarity involving just g2–g6 (a 5-mer) is rejected as unspecific. In human Ago2 (hereafter, Ago2 refers to human Ago2 unless specified otherwise), sites with prolonged base pairing, using at least g2–g7 base-pairing (a 6-mer or larger), can override the checkpoint imposed by Ago2's flexible helix-7 (ref. [11]) and induce a conformational transition in Ago2, allowing extended 3′-pairing of the RNA[12]. However, bioinformatics analysis of validated miRNA–mRNA pairs cannot discern sequence determinants in this region, beyond a preference for forming bulges[13]. Moreover, X-ray structures of ternary complexes are unable to resolve the central region of the duplex, supporting the idea of its flexibility[14]. In vitro biochemical studies[15] showed that mismatches in this region contribute little to target binding affinity but can impair catalytic cleavage of short interfering RNAs (siRNAs) in *Drosophila* Ago2. This implied that the 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 remained unclear.

Here we use nuclear magnetic resonance (NMR) to observe the dynamic process underlying miRNA–mRNA targeting. To elucidate the effects of the conformational transition on RISC function, we combine these measurements with molecular simulations and dual-luciferase reporter (DLR) assays in human cells.

We study hsa-miR-34a-5p ('miR-34a'), part of the evolutionarily conserved miR-34/449 family of miRNAs[16], which targets the mRNA encoding silent information regulator 1 (Sirt1)—a p53-deacetylating enzyme—in a tumour-suppressive feedback loop. Using $R_{1\rho}$ NMR relaxation dispersion, we show that the weak so-called 7-mer–A1 seed of the miR-34a–mSirt1 duplex (Fig. 1a, b) is in equilibrium with a transient and low-populated excited state that results in an 8-mer 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 the RISC, as shown by simulations. In a cell-based assay, a structural mimic of the extended seed produces a roughly two-fold increase in target downregulation. Our data suggest a model whereby RISC undergoes a structural transition mediated by RNA dynamics: the RISC first screens targets for correct seed pairing, than transitions into an active complex, releasing the miR-34a 3′-end which is allowed to fully bind the Sirt1 mRNA in the compensatory region.

Seed dynamics of miR-34a–mSirt1 binding site

Given that an RISC recognizes thousands of distinct binding sites in its target mRNA, with no apparent sequence preference beside the seed, we hypothesize that miRNA–mRNA pairs possess distinct conformational characteristics in the central bulge, facilitating their accommodation within Ago2.

[1]Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden. [2]Department of Chemistry and RNA Institute, University at Albany, State University of New York, Albany, NY, USA. [3]Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden. [4]Present address: Laboratoire SABNP, UMR INSERM 1204, Université d'Evry-Val-d'Essonne–Université Paris-Saclay, Evry, France. [5]Present address: Departamento de Química del Litoral, Facultad de Química and CENUR Litoral Norte, Universidad de la República, Paysandú, Uruguay. [6]These authors contributed equally: Ileana Guzzetti, Parisa Ebrahimi, Sarah Friebe Sandoz, Emilie Steiner. ✉e-mail: katja.petzold@ki.se

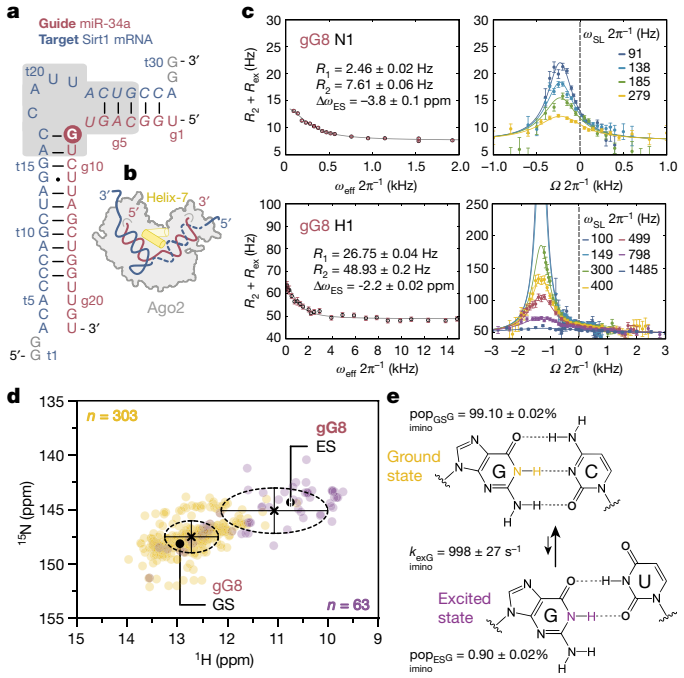


Fig. 1 | Conformational dynamics in the seed of miR-34a–Sirt1 mRNA. **a**, Secondary structure of the miR-34a–mSirt1 duplex determined by NMR. The seed (g2–g6 bound to t23–t27) comprises five base-pairs. The grey box indicates nucleotides selected to generate the reduced construct for $R_{1\rho}$ relaxation-dispersion measurements. **b**, Sketch of human Ago2 accommodating the miR-34a–mSirt1 duplex. Helix-7 is part of Ago2 and is shown with yellow cylinders. **c**, ^{15}N and ^{1}H $R_{1\rho}$ individual relaxation-dispersion profiles of gG8N1 and gG8H1, revealing the single-base-pair switching of gG8:C to gG8:U (circled in **a**). R_1 and R_2 are the longitudinal and transverse relaxation rates, respectively; R_{ex} is the exchange contribution to the relaxation rate; $\omega_{eff}\,2\pi^{-1}$ is the effective measured spinlock power; $\Omega\,2\pi^{-1}$ is the offset; $\Delta\omega_{ES}$ is the chemical shift, a structural parameter of the ground-state-to-excited-state transition; $\omega_{SL}\,2\pi^{-1}$ is the measured spinlock power. Error bars represent one standard deviation (s.d.), derived from Monte Carlo simulations of experimental uncertainty (see Supplementary Methods). **d**, Chemical-shift distribution for ^{1}H1–^{15}N1 moieties of guanosines in G:C (yellow) or G:U (purple) base pairs from the BMRB[20]. Crosses indicate average chemical shifts for G:C and G:U; dashed ellipses show 1 s.d.; black dots indicate chemical shifts for gG8 in the ground state (GS) and relaxation-dispersion-derived excited state (ES). **e**, The G:C to G:U base-pair switch, highlighting the guanosine ^{1}H1–^{15}N1 (imino-global fit, one-sided *F*-test, $n = 1$) groups in the ground state (yellow) and excited state (purple). Errors represent 1 s.d. derived from Monte Carlo simulations of experimental uncertainty (see Supplementary Methods).

First, we solved the secondary structure of miR-34a bound to the validated target site in Sirt1 mRNA (miR-34a–mSirt1 duplex)[17] by NMR (Fig. 1a, Supplementary Fig. 2 and Supplementary Discussion section 2). The overall fold confirmed the secondary structure predicted using MC-Fold (Fig. 1a and Extended Data Fig. 1): the five-nucleotide seed constitutes an A-form 5′-helix between the gG2:tC27 and gG6:tC23 base pairs; meanwhile gG8:tC17 and gG18:tC7 form a 3′-helix containing a wobble gU11:tG14 base pair ('t' refers to the target mRNA). These two helices are separated by a four-nucleotide asymmetric bulge on the mSirt1 side, comprising tC18–tU21 (Figs. 1a, 2a).

To study the structure and dynamics of the bulge, we designed a shortened hairpin construct (miR-34a–mSirt1 bulge) containing the four-nucleotide bulge and enclosing regions (Fig. 1a, grey box, and Fig. 2a). The correct fold was confirmed by a chemical-shift comparison of the shared residues (Supplementary Fig. 2e–h). 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 nuclear Overhauser effects (NOEs) data derived by NMR (Supplementary Fig. 3). We varied the temperature in the simulations in order to explore the RNA conformations that fulfil the experimental constraints, resulting in an ensemble of 153 structures. One representative structure from the ensemble is shown in Fig. 2d, with the relative stem-to-stem angle distribution shown in Fig. 2g (left).

Although classified as a 7-mer–A1 binding site by prediction tools (for example, Targetscan[18]), we found that the miR-34a–mSirt1 duplex and the reduced construct represent a less stable structure: NMR shows that the stability of the gU7:tA22 closing base pair at the 3′-end of the seed is substantially reduced (Fig. 2a, Supplementary Information Figs. S2a, S3a and Supplementary Discussion sections 2, 3). We suggest, therefore, that weak pairing at position 7 might explain previously observed sequence-specific differences in the binding affinity between RISC and target[12]. In agreement with nearest-neighbour models for A:U closing hairpins[19], we propose that 6-mer/7-mer–A1 seeds ending with closing A:U base pairs at position 7 might not suffice for stable displacement of helix-7 of Ago2, resulting in much lower binding affinities, closer to the predicted affinity of the 6-mer.

To assess the base-pair dynamics, we carried out ^{15}N, ^{13}C and ^{1}H $R_{1\rho}$ NMR relaxation-dispersion experiments[2]. ^{1}H–^{15}N NMR of gG8H1 and gG8N1, and ^{13}C NMR of gG8C8, tU21C6, tC17C1′, tU20C1′, tA19C8, tA19C2 and tA22C8, revealed a global exchange process. In this process, the base pair gG8:tC17 interconverts from the most stable structure, the ground state, to a low-populated excited state. The exchange-rate constant (based on ^{1}H1–^{15}N1) for gG8 ($k_{exGimino}$) is 998 ± 27 s^{-1}, with an excited-state population ($pop_{ESGimino}$) of 0.90 ± 0.02; the global k_{ex} (k_{exG}) is 1,008 ± 12 s^{-1}, with an excited-state population (pop_{ESG}) of 0.90 ± 0.01% (Fig. 1c, e, Supplementary Fig. S6c and Supplementary Data S1 Tab 1). Most importantly, we obtained the individual chemical-shift difference between the ground and excited states, $\Delta\omega_{ES} = \Omega_{ES} - \Omega_{GS}$—describing the structure of the excited state—by measuring ^{1}H ($\Delta\omega_{ES}$ = −2.20 ± 0.02 ppm) and ^{15}N ($\Delta\omega_{ES}$ = −3.8 ± 0.1 ppm) in $R_{1\rho}$ relaxation-dispersion data sets. This approach allows us to infer that chemical shifts in the gG8 excited state reside in a region of the ^{1}H–^{15}N heteronuclear single quantum coherence (HSQC) spectrum that is a signature for G:U wobble-base-paired guanosines. This was validated by querying the Biological Magnetic Resonance Bank (BMRB)[20] for ^{1}H1–^{15}N1 chemical shifts of G:U base-paired Gs in RNA-only entries and comparing them with the G:C distribution (Fig. 1d, e, Supplementary Fig. 6c and Supplementary Data S1 Tab 1).

Base-pair switch alters the complex topology

When analysing the MC-Fold[21] output for alternative secondary structures that could fulfil the NMR-derived model (Fig. 1e), we found that a switch in base-pairing partner from gG8:tC17 to gG8:tU21 occurs within the third most energetically favourable structure (Extended Data Fig. 1a, b). To characterize the nature of this process, we carried out additional ^{13}C $R_{1\rho}$ relaxation-dispersion 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} = 1,371 Hz and pop_{ES} = 1.9%, similar to the global fit obtained with ^{1}H, ^{13}C and ^{15}N data sets (Supplementary Figs. 6, 7 and Supplementary Data S1 Tabs 1–3). On the basis of known correlations between our measured ^{13}C $R_{1\rho}$ $\Delta\omega$ values and structural propensities[22–24], we propose a refined secondary structure of the excited state (Fig. 2a).

To derive a three-dimensional structural model of the excited state, we carried out high-temperature REMD simulations of the ground state, restraining five experimentally determined base pairs (Supplementary Data S1 Tab 11)[3]. We identified a putative conformation of the

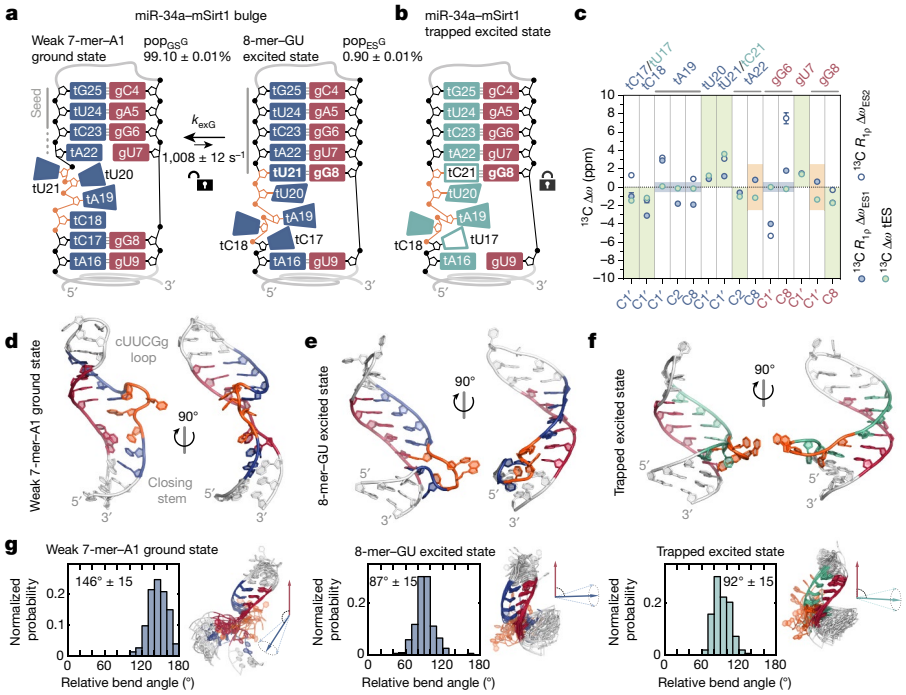


Fig. 2 | Structure and conformation of the excited state of miR-34a–mSirt1. **a**, Secondary structures of the bulge region (the grey area in Fig. 1a). Left, the ground state as solved by NMR; right, the excited state, resulting from $R_{1\rho}$ relaxation-dispersion-derived chemical shifts (gG8:tC17 to gG8:tU21). **b**, Stabilization of the excited-state conformation by isosteric two-point substitution of tC17 with tU17 and tU21 with C21 (trapped excited state). This secondary structure was solved by NMR. **c**, MCSF analysis of the trapped excited state (tES) validates the excited-state model (green shading). Expected perturbations are observed at the sites of modification (orange shading); tA19 and gG6 (blue shading) are explained in Supplementary Discussion 5 and Extended Data Fig. 5. Individually fitted ^{13}C $R_{1\rho}$ relaxation-dispersion-derived $\Delta\omega$ values are in blue, with filled dots for excited state 1 and hollow dots for excited state 2, for three-state-fitting data sets. Error bars for $R_{1\rho}$ relaxation-dispersion-derived $\Delta\omega$ values represent 1 s.d. from fitting (see Supplementary Methods). **d**–**f**, Representative conformations from NMR-informed REMD of the 7-mer–A1 ground state (**d**), the 8-mer–GU excited state (**e**) and the trapped excited state (**f**). **g**, Interhelical bend-angle distributions for the ground state (cluster size $n = 153$), excited state (cluster size $n = 210$) and trapped excited state (cluster size $n = 222$). Means ± s.d. of angle distribution are derived from REMD.

excited state as a cluster within simulations of the ground state, with gG8:tU21 being base paired to gG8:tC17. We sampled the excited-state conformer, restraining gU9:tA16 and gG6:tC23. Addition of magnesium ions, experimentally and in simulations, had no effect (Extended Data Fig. 2 and Supplementary Fig. 6). As for the ground state, we show one representative structure from the excited-state ensemble (210 structures) in Fig. 2e. The topology of the excited state is altered compared with the ground state, indicated by a stem-to-stem coaxial stack that results in an angle distribution peaking around 90° (Fig. 2g, middle).

To experimentally validate the candidate excited-state structure, we used the NMR mutate-and-chemical-shift-fingerprint (MCSF) approach[22], where a substitution or chemical modification is used to trap the proposed excited state. Chemical shifts are then compared between the trapped excited-state and the $R_{1\rho}$ relaxation-dispersion-derived data. We introduced a two-point isosteric substitution in the bulge construct, swapping tC17 with tU17 and tU21 with tC21. This promotes the repositioning of gG8 to the seed 5′-helix, base paired with tC21 (we name this the 'miR-34a–mSirt1 trapped excited state'), without affecting the overall binding affinity. We determined the secondary structure of the trapped excited state by NMR (Fig. 2b and Supplementary Fig. 4) and used imino ^{1}H-^{1}H NOEs as sparse constraints to calculate structural ensembles via REMD (Fig. 2f). As expected, the trapped excited state forms an additional gG8:tC21 base pair that elongates the seed 5′-helix, resulting in identical base-pairing patterns and interhelical bending angles to those in the excited state (Fig. 2g, right).

The MCSF showed remarkable agreement for C1's, tA22C2 and gG8C8 (Fig. 2c, green) confirming that the trapped excited state well represents the overall topology of the excited state modelled from $R_{1\rho}$ relaxation-dispersion data. The sugar puckers measured by $^{3}J_{H1'-H2'}$ for tU20, tU21 (dominant C2′-endo) and tC18 (dominant C3′-endo)—which were expected on the basis of $R_{1\rho}$ relaxation dispersion to interconvert to their opposite configuration in the excited state (Fig. 2a)—were successfully recapitulated in the trapped excited state (Fig. 2b). Furthermore, coaxial stacking between the two helices is validated by tA22H8/C8, tA16H8/C8 and gG8H8/C8 chemically shifting to a region that is characteristic of nucleotides embedded in the uninterrupted A-form helix[23] (Extended Data Fig. 3 and Supplementary Fig. 5a).

Inconsistencies observed for tA22C8 and gG8C1′ are a consequence of the substitution (Fig. 2c, orange, and Extended Data Fig. 3). Data sets for tA19 and gG6 reveal the presence of a second, thermodynamically similar excited state (ES2 in Fig. 2c, blue). However, this conformation could not be trapped experimentally and is discussed in Supplementary Discussion 5 and Extended Data Fig. 3. Interestingly, when probing the dynamics of the trapped excited state in relaxation-dispersion experiments, we detected no exchange with alternative conformations during the timescale probed (Supplementary Figs. 8, 9 and Supplementary Data S1 Tab 3).

In summary, our results show that the miR-34a–mSirt1 binding site is in equilibrium between a high-populated but weak 7-mer–A1 ground state and a low-populated 8-mer–GU seed-elongating excited state, where position 8 is occupied by a G:U base pair, a motif seen previously for miR-48[25,26]. During the ground-state to excited-state switch, both $R_{1\rho}$ relaxation-dispersion data and REMD indicate rearrangement of the bulge and stacking of the two helices.

Functional relevance of 8-mer–GU excited state

We compared wild-type miR-34a–mSirt1 and the miR-34a–mSirt1 trapped excited state by measuring thermal stability followed by

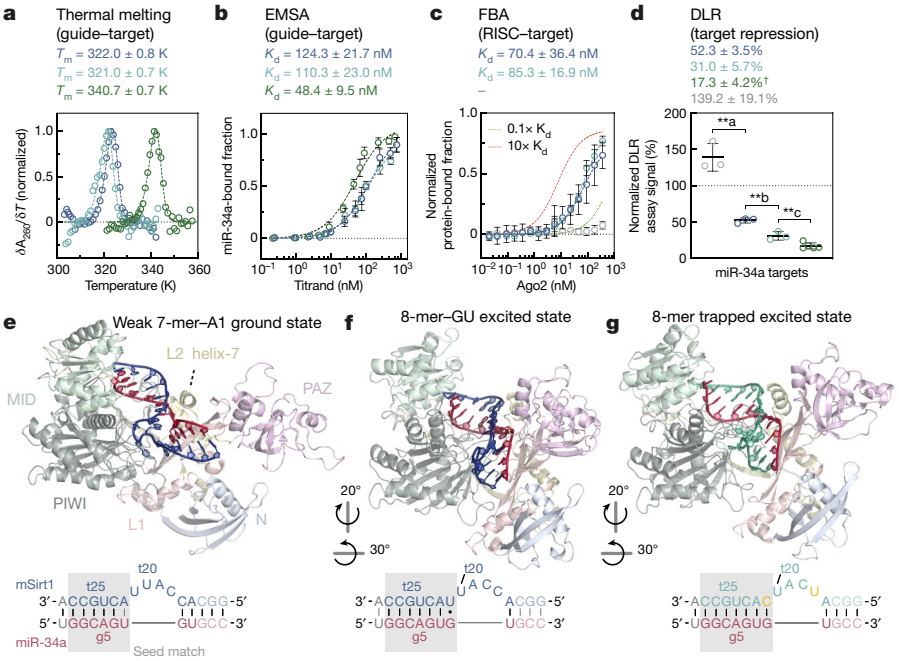


Fig. 3 | Biophysical and functional characterization of wild-type and trapped excited state miR-34a–mSirt1 duplexes. a–c, The wild-type miR-34a–mSirt1 (blue) and the complex in its trapped excited state (turquoise) show comparable stability as indicated by their equivalent melting temperature (T_m) (**a**), binding affinity (K_d) (**b**) and filter-binding assay (FBA; **c**) values. The green curves show the perfect complement for purposes of comparison of the RNAi pathway downregulation efficiency and stability. The melting temperatures (**b**) were obtained by thermal denaturation monitored by ultraviolet absorption at 260 nm (A_{260}); shown are plots from single technical replicates. T_m values are presented as means ± of fitted T_m values from three individual technical replicates. K_d values (**b**) were obtained by EMSA; means are at the plot centres, and error bars represent 1 s.d. from three independent replicates. Fitted K_d values are presented with a confidence interval of 95% as an estimation of the experimental error. For FBA analysis of binding to the miR-34a-loaded RISC (**c**), means are at the plot centres for each data point, and error bars represent 1 s.d. from three independent replicates. See Extended Data Fig. 4, Extended Data Table 1a–c and Supplementary Methods for further details. **d**, DLR assays reveal a roughly two-fold increase in miR-34a-mediated downregulation for the trapped excited state (turquoise, $n = 3$) with respect to the wild-type (blue, $n = 3$). Grey, scrambled negative control, $n = 3$. Green, the highest level of downregulation (siRNA-type), $n = 5$ (performed independently). *P*-values: **a = 0.0015, **b = 0.0054, **c = 0.0076 (** indicates $P < 0.01$, unpaired, two-tailed *t*-test), The centre line shows the mean and error bars represent 1 s.d. from independent replicates (see Supplementary Table 12). **e–g**, Slow-growth simulated RNA structures bound to Ago2 (Protein Data Bank (PDB, https://www.rcsb.org) code 4W5O)[12]. **e**, The ground-state conformation (Fig. 2a, d), orients the compensatory region towards the PAZ domain. **f**, The excited-state conformation (Fig. 2a, e), with coaxial stacking of the helices, orients it towards the N-PIWI domain. **g**, The trapped excited state recapitulates the excited state in **f**.

ultraviolet absorption; RNA–RNA binding affinity by electrophoretic mobility shift assay (EMSA); and RISC–target affinity by filter-binding assay (FBA) of miR-34a-loaded Ago2 (Fig. 3a–c). We found that the melting temperature (T_m) and dissociation constant (K_d) were unchanged (Fig. 3a, b and Extended Data Fig. 4), showing that the substitution does not affect the duplex stability in vitro. Similarly, the binding affinity of miR-34a-loaded Ago2 for the target RNAs in FBA is the same within error (Fig. 3c). Next, we asked whether the two binding sites, despite their similar stabilities, produce different degrees of target downregulation in cells. DLR assays in HEK 293T cells of miR-34a co-transfected with the wild-type weak 7-mer–A1 results in 52.3 ± 3.5% downregulation, as previously reported[17] (Fig. 3d, blue), while the trapped 8-mer excited state leads to 31.0 ± 5.7% downregulation (Fig. 3d, turquoise), showing that the two-point substitution that traps the excited state causes a roughly two-fold increase in target downregulation. 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[12]. We used slow-growth simulations to test whether the calculated ensembles of the miR-34a–mSirt1 bulge ground state, excited state and trapped excited state (Fig. 2d–f) could be accommodated in the Ago2 binding site. Starting from the crystal structures[12], we replaced the visible crystallographic A-form seed helix with conformations from the miR-34a–mSirt1 bulge ground-state, excited-state and trapped excited-state ensembles and aligned them with the seed of the cocrystal (Extended Data Fig. 5).

The resulting simulated ternary complexes are shown in Fig. 3e–g. The ground-state ensemble samples the 3′-helix of the miRNA–mRNA complex within the PAZ domain (Fig. 3e), where the miRNA is bound before target binding[6,7,12,27]. By contrast, the 8-mer–GU excited-state conformation adopts a global bend angle that stacks the 3′-helix coaxially with the seed and favours binding along the PIWI-N domains (Fig. 3f), also recapitulated in the trapped excited state (Fig. 3g).

Although only small conformational changes in the crystal structure of Ago2 are needed to bind the miRNA–mRNA complex in the ground state conformation, accommodating the excited state conformation requires pivoting of the PAZ domain (Fig. 4e and Supplementary Video 1), consistent with prior studies, in which simulations identified these PAZ-domain movements as leading to more ‘open’ Ago2 conformations[28]. Intriguingly, the slow-growth induced-fit conformation of Ago2 bound to the excited state is reminiscent of the binding modes observed for DNA-bound prokaryotic Ago ternary complexes[4,5] (Extended Data Fig. 6), suggesting that Ago2 undergoes structural changes during target recognition and downregulation activity.

We therefore performed a sequence search for other instances of ground-state to excited-state transitions in the 28,653 isoforms of 19,432 human protein-coding genes (specifically, in their 3′-untranslated regions). Requiring a minimal 6-mer–A1 seed resulted in 3,269 predicted target sites for miR-34a (Fig. 4a). Using MC-Fold[21], we then

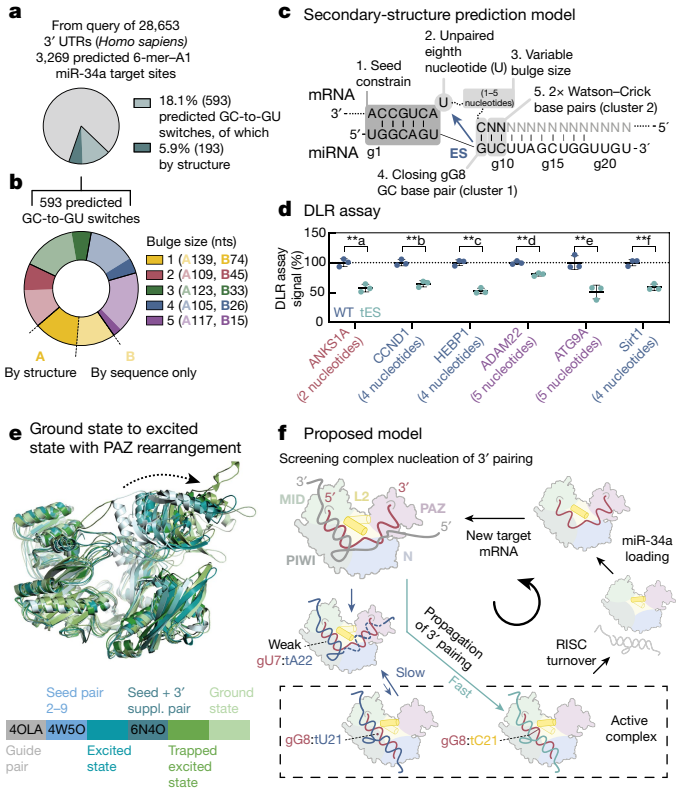


Fig. 4 | Proposed mechanism of downregulation for GC–GU switches in miR-34a-loaded RISC. a, Predicted miR-34a targets in human 3′-UTRs (grey), showing those predicted by sequence (18.1%) and by structure (5.9%) to experience the GC-to-GU switch. **b**, Distribution of bulge sizes (predicted by sequence, lighter colours, or structure, darker colours). **c**, Model of criteria used to search for GC-to-GU switches. **d**, DLR assay for repression of five target mRNAs with respective trapped excited states. All data sets were normalized internally and to the wild type for comparison. The experiment for Sirt1 was performed independently. *P*-values: **a = 0.001324, **b = 0.001112, **c = 0.000253, **d = 0.000935, **e = 0.006454, **f = 0.000646 (**$P < 0.01$, unpaired, two-tailed multiple *t*-test, $n = 3$). The centre line shows the mean; error bars represent 1 s.d from independent replicates (see Supplementary Table 12). **e**, Overlay of X-ray structures and slow-growth simulations. The code at bottom explains the colour of each structure. **f**, Proposed mechanism for the miRNA–mRNA switch in the RISC from a 'screening' into an 'active' state. Domains of Ago2 are indicated in the top left structure. See Discussion for more details.

carried out a sequence search and secondary-structure prediction for the ground-state-to-excited-state switch motif, resulting in bulge sizes from one nucleotide (139 and 74 representatives from sequence search and secondary-structure prediction, respectively), two nucleotides (109 and 45), three nucleotides (123 and 33), four nucleotides (105 and 26) and five nucleotides (117 and 15), respectively. In a more stringent cluster, with three Watson–Crick base pairs following the bulge, we identified 22 targets (Fig. 4c). We selected five different mRNA targets for further investigation in DLR assays in HEK 293T cells (HEBP1, ADAM22, ATG9A, ANKS1A and CCND1 mRNAs). All five candidates were more downregulated in the trapped excited-state form compared with the wild type, with a 50–80% increase in downregulation efficiency (Fig. 4d, Extended Data Fig. 7 and Methods), suggesting that conformational switching of bulged miRNA–mRNA complexes is a general mechanism for modulating downregulation efficiency.

Discussion

Although seed matching is important, it is only the first step of the RISC cycle. Subsequently, it is thought that nucleation from the 3′-helix can propagate towards the central region. This, together 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[1].

We propose that, in the case of miR-34a–mSirt1, this process is mediated by a conformational transition that is triggered by gG8 switching its base-pairing partner. In its ground state, miR-34a–mSirt1 adopts a 7-mer–A1 seed, closed by a weak base pair (g7)—better described as a 6-mer–A1 seed—that is unable to fully displace helix-7 of Ago2 (Fig. 2a). The ground state accesses a distribution of interhelical bend angles that place the miR-34a 3′-end towards the PAZ domain, favouring initial target engagement and nucleation of the 3′-helix[1] (Fig. 4f). During the ground-state to excited-state transition, gG8 repositions to the seed helix and pairs with tU21, resulting in an extended 8-mer–GU seed. The rearrangement of gG8 causes coaxial stacking of the two helices and therefore release of the 3′-end of miR-34a from the PAZ domain, reorienting the RNA duplex towards PIWI domain (Fig. 3e–g) in the simulated structures. This process is accommodated by concerted widening of the N-PAZ channel[12], which 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 excited-state conformation is similar to the catalytically competent state reported for prokaryotic Ago[4,5] (Figs. 3e–g, 4e and Extended Data Fig. 6a, b); moreover, a recent human Ago2 structure confirms that the 3′-helix is mobile[14].

We thus propose that the ground-state to excited-state transition described here provides a mechanism to achieve an active, 'catalytically competent' RISC, promoting mRNA downregulation[1,29]. Although Ago2-bound miRNA is not known to cleave centrally bulged targets, it is possible that these conformational changes enable the RISC to achieve multiple turnovers, which will increase downregulation of the target mRNA[30,31] (Fig. 4f).

Our biophysical and in-cell functional results support this hypothesis, showing a roughly two-fold increase in downregulation upon excited-state stabilization while maintaining RNA–RNA stability. We find that five selected mRNA targets of miR-34a show similar increases in downregulation efficiency when trapped in their excited state (Fig. 4d). Thus the mechanism proposed here could be a widespread feature of bulged binding sites containing partial or extended 3′-pairing.

We have shown that the structural transitions of the guide–target RNA modelled in the RISC provide a mechanistic explanation for bulged complexes, enabling a more accurate prediction of target downregulation by miRNAs. With ever-increasing interest in adapting RNA-guided nuclease machineries for therapeutic, diagnostic and technology applications, we suggest 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.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2336-3.

1. Bartel, D. P. Metazoan microRNAs. *Cell* **173**, 20–51 (2018).
2. Marušič, M., Schlagnitweit, J. & Petzold, K. RNA dynamics by NMR spectroscopy. *ChemBioChem* **20**, 2685–2710 (2019).
3. 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* **162-163**, 96–107 (2019).
4. Wang, Y. et al. Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. Nature 461, 754–761 (2009).
5. Sheng, G. et al. Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl Acad. Sci. USA* **111**, 652–657 (2014).
6. Elkayam, E. et al. The structure of human argonaute-2 in complex with miR-20a. *Cell* **150**, 100–110 (2012).

7. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336**, 1037–1040 (2012) e
8. Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. Structure of yeast Argonaute with guide RNA. *Nature* **486**, 368–374 (2012).
9. Grimson, A. et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* **27**, 91–105 (2007).
10. Nielsen, C. B. et al. Determinants of targeting by endogenous and exogenous microRNAs and siRNAs. *RNA* **13**, 1894–1910 (2007).
11. 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).
12. Schirle, N. T., Sheu-Gruttadauria, J. & MacRae, I. J. Structural basis for microRNA targeting. *Science* **346**, 608–613 (2014).
13. 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).
14. Sheu-Gruttadauria, J., Xiao, Y., Gebert, L. F. & MacRae, I. J. Beyond the seed: structural basis for supplementary microRNA targeting by human Argonaute2. *EMBO J.* **38**, e101153 (2019).
15. 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. *Cell* **151**, 1055–1067 (2012).
16. He, L. et al. A microRNA component of the p53 tumour suppressor network. *Nature* **447**, 1130–1134 (2007).
17. Yamakuchi, M., Ferlito, M. & Lowenstein, C. J. miR-34a repression of SIRT1 regulates apoptosis. *Proc. Natl Acad. Sci. USA* **105**, 13421–13426 (2008).
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. Vecenie, C. J. & Serra, M. J. Stability of RNA hairpin loops closed by AU base pairs. *Biochemistry* **43**, 11813–11817 (2004).
20. Ulrich, E. L. et al. BioMagResBank. *Nucleic Acids Res.* **36**, D402–D408 (2008).
21. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452**, 51–55 (2008).
22. Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A. & Al-Hashimi, H. M. Visualizing transient low-populated structures of RNA. *Nature* **491**, 724–728 (2012).
23. Xue, Y. et al. in *Laboratory Methods in Enzymology: RNA* **vol. 558** (eds Woodson, S. A. & Allain, F. H. T.) 39–73 (Academic Press, 2015).
24. 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 Res.* **45**, e134 (2017).
25. Broughton, J. P., Lovci, M. T., Huang, J. L., Yeo, G. W. & Pasquinelli, A. E. Pairing beyond the seed supports microRNA targeting specificity. *Mol. Cell* **64**, 320–333 (2016).
26. Brancati, G. & Großhans, H. An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity. *Nucleic Acids Res.* **46**, 3259–3269 (2018).
27. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. *Nat. Struct. Mol. Biol.* **11**, 576–577 (2004).
28. 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).
29. Gebert, L. & MacRae, I. J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **20**, 21–37 (2019).
30. Park, J. H., Shin, S.-Y. & Shin, C. Non-canonical targets destabilize microRNAs in human Argonautes. *Nucleic Acids Res.* **45**, 1569–1583 (2017).
31. De, N. et al. Highly complementary target RNAs promote release of guide RNAs from human Argonaute2. *Mol. Cell* **50**, 344–355 (2013).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

RNA sample preparation

RNA samples were produced in-house by T7 in vitro transcription[32,33], unless otherwise stated. Modified DNA templates (Integrated DNA Technologies) with oxy-methylated C2′ groups in the first two 5′ nucleotides were used to reduce the 3′-OH heterogeneity of the product[34]. In vitro transcription reactions were supplemented with 20% dimethylsulfoxide (DMSO) to improve reaction yield and to reduce side products[35]. ^{13}C- and ^{15}N-labelled NMR samples were produced by supplementing the transcription reaction with ^{13}C and ^{15}N fully labelled nucleotide triphosphates (Merck Sigma Aldrich). A 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) (see Supplementary Methods). hsa-miR-34a-5p 3′-Cy3 labelled and single-stranded mSirt1 in the trapped excited state were purchased from Integrated DNA Technologies as chemically synthesized RNA oligonucleotides purified by RNase-free HPLC purification. A complete list of RNA and DNA sequences used here is given in Supplementary Data S1 Tab 10.

Ago2 preparation and RISC reconstitution

Human Argonaute 2 cloned into the pFastBac HT plasmid was obtained as described[36]. Ago2 was expressed in Sf9 insect cells and purified from the clarified cell lysate through nickel affinity chromatography and gel-filtration chromatography. Sf9 cells were obtained from Invitrogen (catalogue number 11496-015, lot 1296885) and, to our knowledge, were not authenticated. All cell lines were visually inspected throughout the experiments and can be easily identified through their morphology and growth. No misidentified cells were used. The fractions containing Ago2 were pooled together, concentrated and stored at −80 °C. Further details of Ago2 sample preparation are described in the Supplementary Methods.

Purified Ago2 was incubated with a roughly two-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 tris(2-carboxyethyl)phosphine (TCEP) supplemented with 10 μg ml^{-1} bovine serum albumin (BSA; Sigma Aldrich) for 6 h at 37 °C. The assembled RISC (Ago2–miR-34a complex) was then separated from unbound excess RNA by gel filtration chromatography. Loading of the guide miR-34a into the RISC was assessed by an improved northern blot for the detection of small RNA[37,38]. Further details of RISC reconstitution are given in the Supplementary Methods.

Thermal denaturation monitored by UV absorption

Thermal denaturation monitored by UV absorption at 260 nm (A_{260}) was carried out using an Evolution 260 Bio UV-vis spectrophotometer (Thermo Scientific) equipped with a 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 ; see Supplementary Methods) allowed for estimation of the melting temperature (T_m) and thermodynamic parameters presented in Extended Data Table 1a and Supplementary Data S1 Tab 7.

EMSA

hsa-miR-34a-5p 3′-Cy3 was incubated at a final concentration of 24 nM with increasing amounts of unlabelled single-stranded partner (mSirt1, trapped excited-state mSirt1 or the 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-bBorate-EDTA (TBE) polyacrylamide gel. Fluorescence signals relative to the free and bound forms of hsa-miR-34a-5p 3′-Cy3 were quantified using ImageJ software[39]. Fitting of the binding curves to a standard binding isotherm (see Supplementary Methods) allowed for estimation of the equilibrium dissociation constants (K_d) presented in Extended Data Table 1b and Supplementary Data S1Tab 8.

Equilibrium filter binding assay

3′-Cy3-labelled target RNAs (mSirt1, trapped excited-state mSirt1 or scrambled control) were incubated at a constant concentration of 0.5 nM with increasing amounts of Ago2–miR-34a complex in target binding buffer[12] (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^{-1} yeast transfer RNA (Sigma Aldrich) and 10 μg ml^{-1} BSA (Sigma Aldrich)) to a final volume of 100 μl and incubated for 1 h 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) forms of 3′-Cy3 target RNAs were quantified using ImageJ software[39]. Fitting of the binding curves to a standard binding isotherm (see Supplementary Methods) allowed for estimation of the K_d values presented in Extended Data Table 1c and Supplementary Data S1 Tab 9.

NMR spectroscopy

All NMR assignment and $R_{1\rho}$ relaxation-dispersion 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. These experiments were performed on ^{13}C and ^{15}N fully labelled 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/C5/C6/C8–^{1}H1′/H2/H5/H6/H8, sugar ^{13}C1′–^{1}H1′ and imino ^{15}N1/N3–^{1}H1/H3 resonances was achieved using a standard set of ^{1}H–^{13}C, ^{1}H–^{15}N two-dimensional HSQCs, three-dimensional ^{1}H-^{13}C-^{15}Ns, ^{1}H-^{15}N-^{15}N correlation spectroscopy (COSY) and ^{1}H–^{1}H nuclear Overhauser effect spectroscopy (NOESY) NMR experiments (all acquired using a mixing time of 175 ms) as described[40], recorded at different temperatures (9.0 °C, 22.4 °C and 35.9 °C; Supplementary Figs. 1, 3 and 4). For the miR-34a–mSirt1 duplex, only a reduced set of imino ^{15}N1/N3–^{1}H1/H3 resonances were assigned using ^{1}H–^{15}N two-dimensional HSQCs, HNN COSY and ^{1}H-^{1}H NOESY NMR experiments (Supplementary Fig. 2). Assigned chemical shifts were deposited to the BMRB[20] for hsa-miR-34a-5p (entry 27225), the miR-34a–mSirt1 bulge (entry 27226) and the miR-34a–mSirt1 trapped excited state (entry 27229).

^{1}H, ^{13}C and ^{15}N $R_{1\rho}$ relaxation-dispersion NMR. These experiments were carried out as described[41–44], using ^{13}C and ^{15}N fully labelled (^{13}C and ^{15}N $R_{1\rho}$) or natural-abundance unlabelled (^{1}H $R_{1\rho}$) RNA samples dissolved in 15 mM Na_2HPO_4/NaH_2PO_4, 25 mM NaCl and 0.1 mM EDTA, pH 6.5. In brief, for each spinlock power (ω_{SL}), data points were recorded as a function of different relaxation delays (T_{EX}). For each residue, variable-delay 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 ($\Omega 2\pi^{-1}$), we recorded a subset of off-resonance data sets with an extended variable-delay list comprising longer maximal T_{EX} values; we took care that no additional heating occurred. In all data sets, we discarded data points with signal-to-noise ratios (S/N) of less than 20 for ^{1}H and ^{13}C, and of less than 10 for ^{15}N.

Peak intensities were extracted from deconvoluted one-dimensional data sets 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[45] with 500 iterations. Potential artefacts (for example, arising from Hartmann–Hahn matching conditions or strong ^{1}H–^{1}H and ^{13}C–^{13}C homonuclear coupling that results in deviation from mono-exponential behaviour) were excluded from subsequent analysis by discarding exponential fits with R^2 values of less than 0.985. $R_{1\rho}$ values as a function of ω_{SL} were subsequently fitted, using the Laguerre approximation[46] (see Supplementary Methods equation (5)), and assuming an absence of exchange ($R_{ex} = 0$), a fast exchange regime (reduced Laguerre form where k_{ex} is much greater than $\Delta\omega$; Supplementary Methods equation (6)), two-state exchange (Supplementary Methods equation (7)) or three-state exchange (Supplementary Methods equation (8) and (9)) using the models and fitting methods further described in the Supplementary Methods. Selection of the best-fitting model for each data set was performed using a statistical *F*-test[23]. Degrees of freedom were calculated as the number of data points (as represented by values in Supplementary Tables 1, 5) minus the number of fitted parameters for each model (two for no exchange, three for reduced Laguerre, five for Laguerre approximation two-state, and eight for Laguerre approximation three-state, for the global fits indicated in Supplementary Table 1). Fitted parameters, reduced χ^2 values resulting from the fit and exact *P*-values from the F-tests (one-sided) for each data set are reported in Supplementary Data S1 Tabs 1, 2, 4.

We carried out global fitting by assuming the presence of one collective exchange process to a minor populated state (ES^G), characterized by the global parameters $k_{ex}{}^G$ (the global exchange rate) and $pop_{ES}{}^G$ (population of ES^G) shared across the data sets. Each data set 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 two-state (Supplementary Methods equation (7)) and a three-state (Supplementary Methods equation (8) and (9)) exchange model. For those data sets that were globally fitted using three-state exchange model, 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} and $\Delta\omega_{ES2}$) unconstrained during the fit, fundamentally equivalent to fit them individually. Error estimates of the fitted global parameters were computed as one standard deviation using a Monte Carlo simulation method[45] with 500 iterations. Selection of the best-fitting model was performed using a statistical *F*-test[23], where the simpler fit (a global fit, with a smaller number of parameters) was selected if *P* was greater than 0.05. 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 *F*-test (one-sided) for the global fittings are in Supplementary Data S1 Tab 1.

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 carried out using MC-Fold 1.6.0[21], 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 identification codes and nucleotide numbers of guanosines involved in either G:C or G:U base pairs were obtained using RNA FRABASE 2.0[47]. PDB identification codes that have matching BMRB entries 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, and duplicates and misreferenced couples were removed. A total of 303 G:C and 63 G:U unique ^{1}H1–^{15}N1 couples were obtained (Fig. 1d).

All-atom, explicit solvent molecular dynamics simulations

Atomistic simulations of the miR34a–mSirt1 bulge were initialized using starting structures generated by MC-Fold and MC-Sym[21]. All-atom, explicit solvent molecular dynamics simulations were performed using GROMACS 5.0.7[48] and the modified Chen–Garcia force field for RNA[49], including backbone phosphate modifications[50]. The structure was solvated with 6,664 TIP4P-Ew[51] 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 activity-coefficient calibrated parameters[52]. In order to enhance exploration of diverse bulge conformations using temperature replica-exchange without inadvertently inducing RNA melting, we assigned five harmonic restraints 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 a constant pressure of 1 atm, with random initial velocities drawn from a Boltzmann distribution.

Using REMD, we simulated 24 individual replicas spanning a temperature range of 77–147 °C to evaluate the conformational flexibility of the miR34a–mSirt1 bulge. The exchange rate was 25% with attempted temperature swaps every 1,000 steps (2 ps), which is also how often coordinates were saved. Once equilibrated, production simulations were propagated for roughly 670 ns per replica, a total of 16.08 μs of cumulative simulation time. Structural clustering based on all-heavy-atom r.m.s.d. was accomplished using the algorithm of ref. [53] with 30,000 evenly spaced snapshots taken from the lowest temperature replica (27 °C), using a cut-off of 5.0 Å. The most highly populated cluster, which contains more than 60% of all structures in the 27 °C replica, is the ground-state ensemble (Fig. 2g). We also carried out a separate set of REMD simulations consisting of 25 replicas spanning 25–77 °C, using the same settings as above. Each replica was sampled for 478 ns for a cumulative total of 11.95 μs, and identical cluster analysis was carried out on the 25 °C replica. Details of REMD simulations of the miR-34a bulge excited state and trapped excited state, as well as interhelical bending-angle distributions, are further described in the Supplementary Methods.

Alignment of ground-state/excited-state ensembles into the Ago2 crystal structure. We initially aligned 250 randomly picked snapshots from each REMD ensemble (ground state, excited state and trapped excited state) into the 4W5T PDB structure[12]. Each simulation structure was aligned such that the backbone phosphate positions of bases g2–g8 matched those of the crystal structure. For visual clarity, only 20 of the 250 conformations are graphically depicted in Extended Data Fig. 5.

Slow-growth simulations of insertion of the ground state/excited state into the Ago2 complex. In order to ascertain the ability of the Ago2 protein to physically accommodate the miR-34a–mSirt1 RNA complexes in the ground and excited states, we inserted representative snapshots from each ensemble into the Ago2 protein using slow-growth binding simulations[54]. Starting with the 4W5O PDB structure[12], we deleted the existing partial miRNA–mRNA complex and modelled in missing Ago2 amino acids. 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 van der Waals 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 succeeds only if the RNA can 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 roughly 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. The structural models resulting from slow-growth insertion of the ground-state, excited-state and trapped excited-state RNA into the Ago2 protein have been deposited in Model Archive (www.modelarchive.org) under accession codes ma-bc9uo, ma-z54y4 and ma-g8e5z.

Plasmids. All mRNA-targeting DLR[55] plasmids were generated by cloning a synthetic double-stranded DNA (Supplementary Table 10a) into the XhoI and NotI restriction sites of wild-type psiCHECK2-miR-34 (ref. [56]). The fully complementary binding site is the unmodified psiCHECK2-miR-34 WT plasmid[56]. As a negative control, we used the mutated hsa-miR-34a-5p binding site of psiCHECK2-miR-34 MT[56]. These plasmids were a gift from J. Weidhaas (Addgene plasmids 78258 and 78259). The newly generated plasmids were verified by sequencing.

Cell lines and culture. HEK 293T cells were obtained from ATCC (catalogue number CRL-11268) and authenticated by short tandem repeat (STR) analysis by the manufacturer. These cells were used soon after purchase and therefore were not tested for mycoplasma contamination. All cell lines were visually inspected throughout the experiments and can be easily identified by their morphology and growth. No misidentified cells were used. For DLR, HEK 293T cells were cultured in Dulbecco's modified essential medium (DMEM, Gibco) supplemented with 10% fetal calf serum (FCS, Gibco).

DLR assay. HEK 293T cells were seeded 24 h before 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-buffered saline (PBS) once, and luciferase activity was measured with a DLR assay system (Promega) according to the manufacturer's protocol, using a Promega GloMax 96 microplate luminometer, with a 1-s delay and 10-s 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 on this basis. Results show the average and standard deviation of at least three independent biological replicates. For statistical analysis, we performed unpaired, two-sided single (Fig. 3d) and multiple (Fig. 4d) *t*-tests. Error bars represent one s.d. $^{**}P \leq 0.01$. Details from the fit are presented in Supplementary Data Tab 12.

Predicted target screening of GC-to-GU switches

In total we downloaded 28,653 3′-UTR sequences, including all isoforms, of all 19,432 human protein-coding genes from TargetScan[18]. The sequences were bioinformatically screened for putative mir-34a targets using regular expression. Specifically, sequences were selected that included the reverse complementary sequence of a canonical 6-mer–A1, followed by a U or C as the first nucleotide of the bulge. Thereafter, to allow for a bulge of up to six nucleotides, the sequence was unspecified for positions one to five, 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) was screened according to its potential for forming different bulge sizes (from one to five nucleotides) with a G:C or a G:U as the closing base pair. Thereafter, the secondary structure of each mRNA-UUCG–miR34a complex was simulated using MC-Fold 2.32[21]; different mRNA lengths were tested, until a maximum of eight nucleotides was added to an mRNA sequence of 22 nucleotides. 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 was defined as having: first, a non-base-paired U (position t21 in mSirt1) after the seed, followed by a number of unpaired bases equalling the length of the bulge; second, a GC Watson–Crick base pair closing the bulge, followed by two base pairs, in the 3′-helix (Fig. 4c, cluster 1); and third, a second more stringent cluster (cluster 2) described by two additional Watson–Crick base pairs after the GC closing base pair. An excited state is defined as having a U (position t21 in mSirt1) after the seed pairs with the G in position gG8 (in miR-34a). For obvious structural reasons, in all clusters, we excluded structures in which the miR-34a sequence was folding onto itself or where shortening of the seed was occurring. Sequences were considered only if the ground state and excited state are present for at least three different lengths, and if all the lengths have at least a ground state and an excited state. Of the five targets tested, only CCND1 and ATG9A were previously confirmed as miR-34a targets[57,58].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

NMR sequence-specific resonance assignments have been deposited in the Biological Magnetic Resonance Data Bank under accesssion numbers 27225 (hsa-miR-34a-5p), 27226 (the miR34a–mSirt1 bulge) and 27229 (the miR34a–mSirt1 trapped excited state). The plasmids used for the DLR assay were a gift from J. Weidhaas (Addgene plasmids 78258 and 78259). All data and code used for data analysis are available upon request. The ensembles of REMD simulations have been deposited in Model Archive (www.modelarchive.org) under accession numbers ma-bc9uo, ma-z54y4 and ma-g8e5z.

32. Beckert, B. & Masquida, B. in *Rna* 29–41 (Springer, 2011).
33. Baronti, L., Karlsson, H., Marušič, M. & Petzold, K. A guide to large-scale RNA sample preparation. *Anal. Bioanal. Chem*. **410**, 3239–3252 (2018).
34. 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).
35. Helmling, C. et al. Rapid NMR screening of RNA secondary structure and binding. *J. Biomol. NMR* **63**, 67–76 (2015).
36. De, N. & MacRae, I. J. in *Argonaute Proteins* 107–119 (Springer, 2011).
37. Pall, G. S. & Hamilton, A. J. Improved northern blot method for enhanced detection of small RNA. *Nat. Protocols* **3**, 1077–1084 (2008).
38. Rio, D. C. Northern blots for small RNAs and microRNAs. *Cold Spring Harbor Protocols*, https://doi.org/10.1101/pdb.prot080838 (2014).
39. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
40. Fürtig, B., Richter, C., Wöhnert, J. & Schwalbe, H. NMR spectroscopy of RNA. *ChemBioChem* **4**, 936–962 (2003).
41. Schlagnitweit, J., Steiner, E., Karlsson, H. & Petzold, K. Efficient detection of structure and dynamics in unlabeled RNAs: the SELOPE approach. *Chemistry* **24**, 6067–6070 (2018).
42. 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 $R_{1\rho}$ NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 3818–3819 (2009).
43. 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).
44. Steiner, E., Schlagnitweit, J., Lundström, P. & Petzold, K. Capturing excited states in the fast-intermediate exchange limit in biological systems using ^{1}H NMR spectroscopy. *Angew. Chem. Int. Edn* **55**, 15869–15872 (2016).
45. Metropolis, N. & Ulam, S. The Monte Carlo method. *J. Am. Stat. Assoc.* **44**, 335–341 (1949).
46. Palmer, A. G., III & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev*. **106**, 1700–1719 (2006).
47. 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).
48. Abraham, M. J. et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25 (2015).
49. Chen, A. A. & García, A. E. High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Proc. Natl Acad. Sci. USA* **110**, 16820–16825 (2013).
50. Steinbrecher, T., Latzer, J. & Case, D. A. Revised AMBER parameters for bioorganic phosphates. *J. Chem. Theory Comput*. **8**, 4405–4412 (2012).
51. Horn, H. W. et al. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys*. **120**, 9665–9678 (2004).

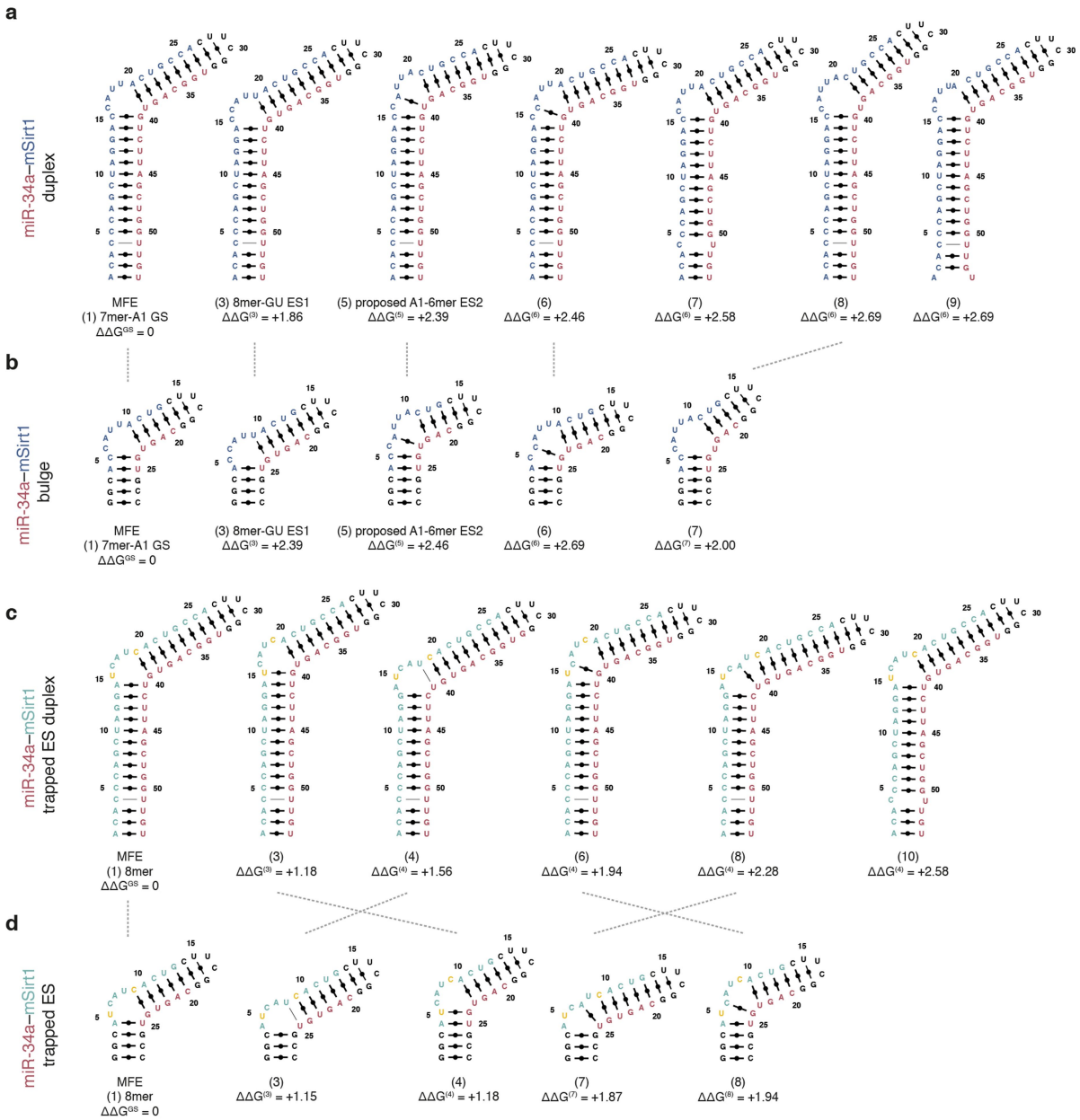
52. Joung, I. S. & Cheatham, T. E., III. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).
53. Daura, X. et al. Peptide folding: when simulation meets experiment. *Angew. Chem. Int. Edn* **38**, 236–240 (1999).
54. 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). *Mol. Simul.* **28**, 67–80 (2002).
55. Grentzmann, G., Ingram, J. A., Kelly, P. J., Gesteland, R. F. & Atkins, J. F. A dual-luciferase reporter system for studying recoding signals. *RNA* **4**, 479–486 (1998).
56. Salzman, D. W. et al. miR-34 activity is modulated through 5′-end phosphorylation in response to DNA damage. *Nat. Commun.* **7**, 10954 (2016).
57. Sun, F. et al. Downregulation of CCND1 and CDK6 by miR-34a induces cell cycle arrest. *FEBS Lett.* **582**, 1564–1568 (2008).
58. 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).

Acknowledgements We thank the members of the Petzold and the E. R. Andersson laboratories for insightful discussions; N. H. Hopkins for providing the curated chemical shifts and for contributing to the MC-Fold secondary-structure-prediction script; I. MacRae for the gift of plasmid expressing human Ago2; and 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 help with human Ago2. K.P. acknowledges funding from the Swedish Research Council (grant number 2014-04303), the Swedish Foundation for Strategic Research (project number ICA14-0023), Harald och Greta Jeansson Stiftelse (JS20140009), Carl Tryggers stiftelse (CTS14-383 and 15-383), Eva och Oscar Ahréns Stiftelse, Åke Wiberg Stiftelse (467080968 and M14-0109), Cancerfonden (CAN 2015/388), the Karolinska Institute Department of Medical Biochemistry and Biophysics (grant number KID 2-3707/2013 and support for the purchase of a 600-MHz 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 of the Extreme Science and Engineering Discovery Environment (XSEDE) (allocation TG-MCB140273 to A.A.C.), which is supported by National Science Foundation grant number ACI-1548562. J.S. acknowledges funding through a Marie Sklodowska-Curie Individual Fellowship (EU H2020/project number 747446).

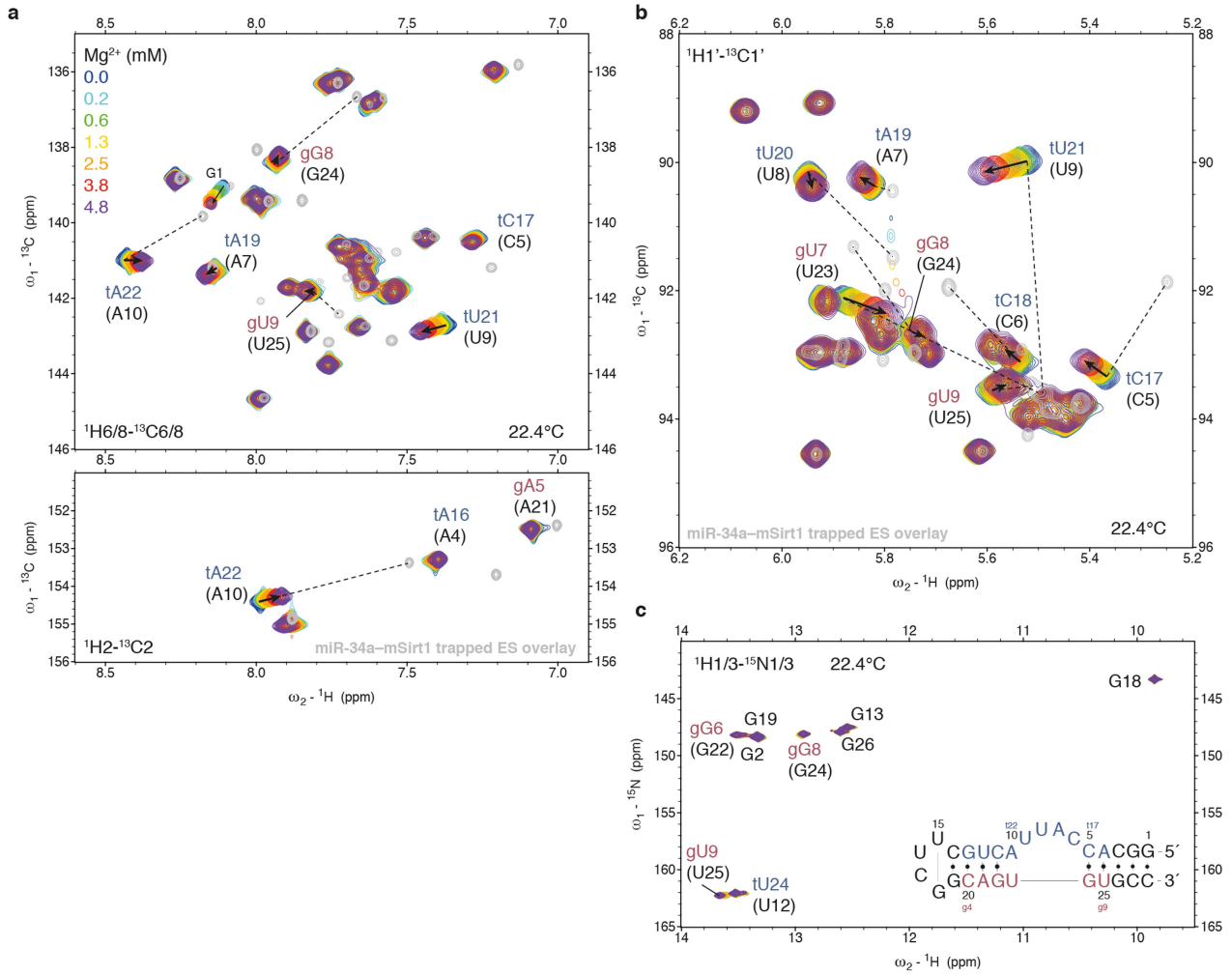
Author contributions L.B. and K.P. conceived the project. L.B. carried out most of the experiments and data analysis, with assistance from I.G., E.S., S.F.S., K.P. and L.S. P.E. and A.A.C. carried out the REMD simulations and slow-growth docking. I.G. and B.F. performed the target screening of GC-to-GU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

Competing interests The authors declare no competing interests.

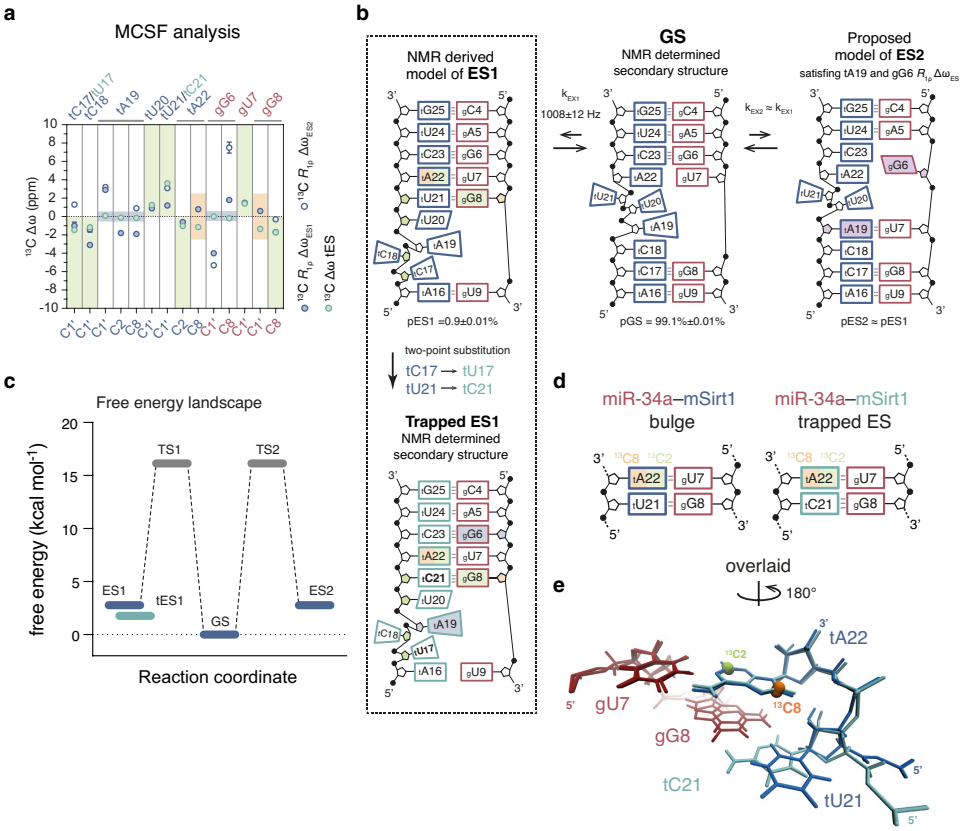
Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2336-3.
Correspondence and requests for materials should be addressed to K.P.



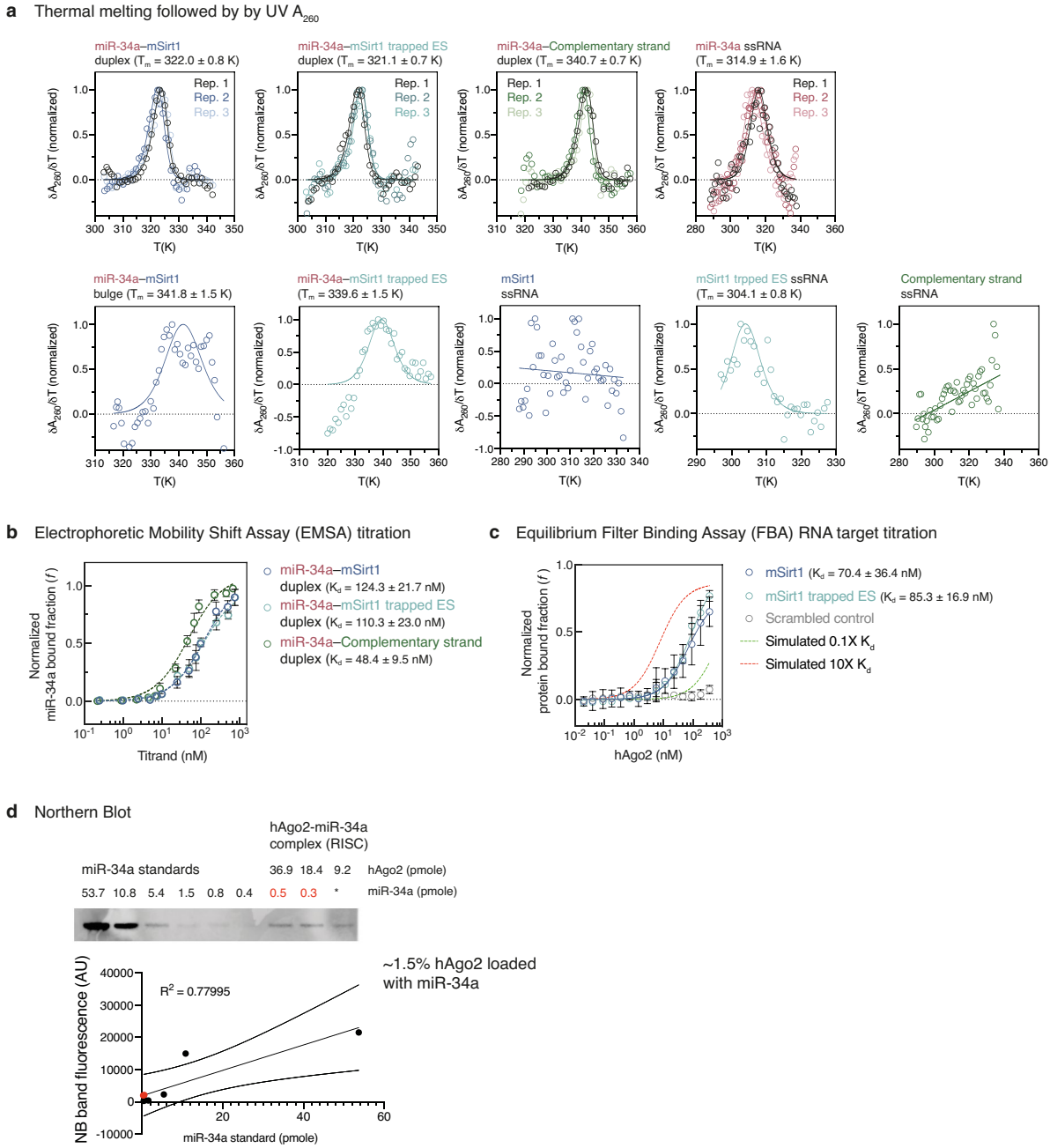
Extended Data Fig. 1 | Secondary-structure prediction using MC-Fold. Secondary-structure rearrangements among the ten lowest-energy structures were calculated using MC-Fold[21]. Ranking (numbers in parentheses) according to the predicted energy difference, based on the minimum free energy (MFE), is indicated in each label ($\Delta\Delta G(n)$ in units of unreferenced kcal mol^{-1}, as described[21]). Secondary structures with a single base-pair opening in the cUUCGg region are omitted. **a**, The miR-34a–mSirt1 duplex connected by a cUUCGg loop (black). The MFE corresponds to a 7-mer–A1 binding site. Suboptimal structures (3) and (5) suggest possible modulation of the binding site to a 8-mer–GU and an 6-mer–A1 configuration, respectively. **b**, miR-34a–mSirt1 bulge construct, comprising a cUUCGg loop and a closing stem (black). The secondary-structure distribution of the miR-34a–mSirt1 bulge follows the same trend as the full-length duplex; dashed lines connect identical bulge structures. Suboptimal structures were used to validate or reject models of excited-state (ES) secondary structures on the basis of $R_{1\rho}$ NMR relaxation-dispersion data. Structure (1), with the MFE, corresponds to the assigned ground-state structure (GS). Structure (3) satisfies the $^{1}H_1$ and $^{15}N_1$ $R_{1\rho}$ NMR relaxation-dispersion data on gG6(G24), being G:U base paired with tU20(U9). Structure (5) is mutually exclusive with (3) in structural terms and satisfies the ^{13}C $R_{1\rho}$ NMR relaxation-dispersion data measured on tA19 that indicate this residue adopting a base-paired conformation. Therefore we propose structure (3) as ES1 and structure (5) as ES2. Conformations (6) and (7) do not agree, and partially clash, with our $R_{1\rho}$ NMR relaxation-dispersion data and can therefore be excluded as excited states. **c**, miR-34a–mSirt1 trapped excited-state duplex connected by a cUUCGg loop (black). Substituted nucleotides used to trap the excited state are in yellow. The MFE corresponds to a 8-mer binding site. **d**, miR-34a–mSirt1 (turquoise) trapped excited-state construct comprising an cUUCGg loop and a closing stem (black). Substituted nucleotides used to trap the excited state are in yellow. The secondary-structure distribution of the miR-34a–mSirt1 trapped excited state follows a similar trend as the full-length duplex; dashed lines connect identical bulge structures.



Extended Data Fig. 2 | Mg^{2+} titration of the miR-34a–mSirt1 bulge followed by NMR. Shown are HSQC overlays of different Mg^{2+} 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 colour-coded (**a**, top left). Additional overlay of the miR-34a–mSirt1 trapped ES is shown in grey in **a**, **b**. Arrows indicate the chemical-shift trajectory during titration. Dashed lines connect equivalent peaks in the miR-34a–mSirt1 bulge and trapped ES constructs.

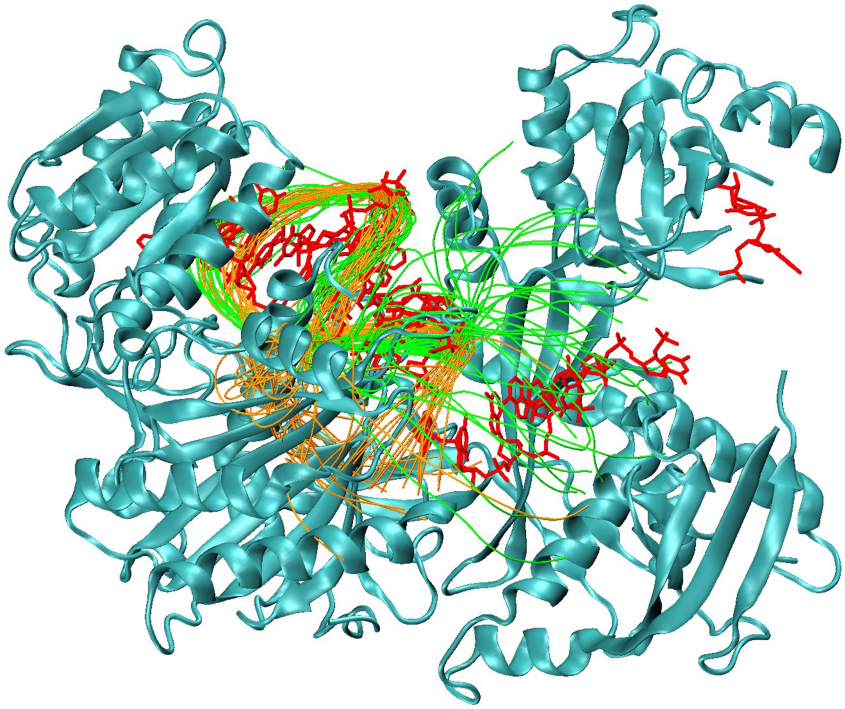


Extended Data Fig. 3 | MCSF analysis of the miR-34a–mSIRT1 bulge and trapped excited state, and analysis of ^{13}C tA22C8 outliers. a, **b**, We used the MCSF approach[22] to cross-validate our candidate excited state (ES1), modelled using $R_{1\rho}$-derived ground-state-to-excited-state chemical-shift differences (**a**, ^{13}C $R_{1\rho}$ $\Delta\omega$ data, blue dots; **b**, left). p_{ES} refers to the excited-state population (pop_{ES} in the main text). We also generated an ES1 mimic (trapped ES1) using a two-point substitution, predicted to stabilize the proposed conformation (**b**, bottom). For each reporter atom, we compared ^{13}C $R_{1\rho}$ $\Delta\omega$ data with the chemical-shift differences derived from the assignment of the bulge and the trapped ES constructs (**a**, ^{13}C $\Delta\omega$ trapped ES (tES) data, turquoise dots). In **a**, The MCSF analysis validates our ES1 model (green shading), with exceptions arising from the limitations of the mimic (orange shading) and from the presence (violet shading) of a second ES (ES2, **b**, right). Errors for $R_{1\rho}$ relaxation-dispersion-derived $\Delta\omega$ represent 1 s.d. from fitting (see also Supplementary Methods). In **b**, the proposed model for ES2 satisfies the ^{13}C $R_{1\rho}$ $\Delta\omega$ data measured for tA19 and gG6. GS, ground state. **c**, The free-energy landscape for the entire star-like three-state exchange process. (The MCSF analysis and ES2 are discussed further in the Supplementary Information, Discussion 5.) The transition coeffient (κ), is assumed to be 1 (ref. [23]), so the transition-state energies (TS1 and TS2), calculated using Supplementary equation (11), must be considered an upper limit of this exchange process. **d**, **e**, The substitution site (tU21 to tC21) perturbs the chemical environment of tA22C8 that is directly neighbouring the substituted nucleobase (orange sphere in **e**). Conversely, tA22C2 (green sphere), pointing towards the miR-34a strand (red), experiences an equivalent chemical environment in the bulge (blue) and trapped ES (turquoise) constructs. This explains the inconsistency in the MCSF profile for tA22C8 (Supplementary Fig. 12a, orange box). **d**, Secondary structure environment of tA22 in the miR-34a–mSirt1 bulge excited state (left) and trapped ES (right) constructs. The substitution site (tU21 to tC21) is highlighted. **e**, 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 and tA22C2 ^{13}C atoms are in orange and green respectively.

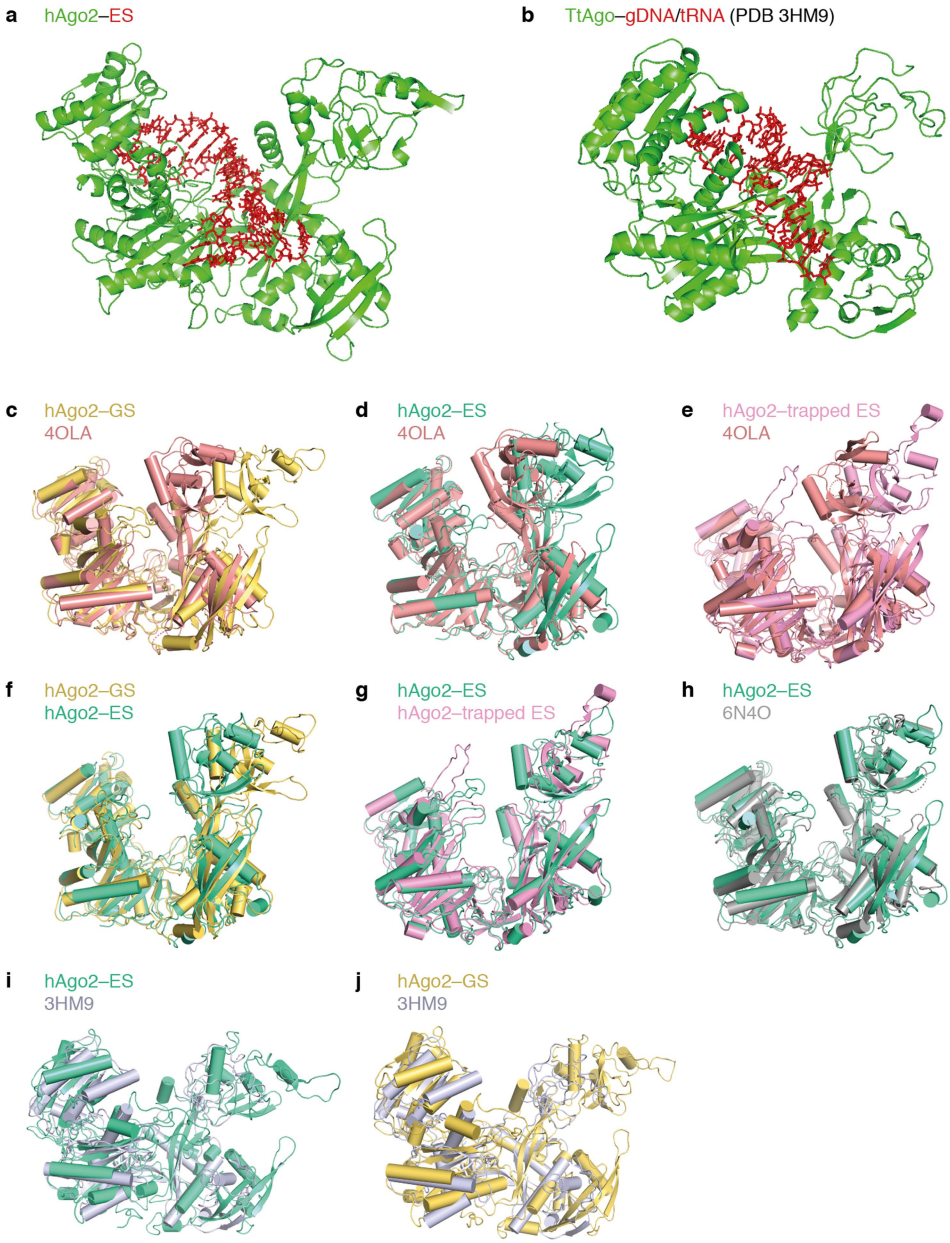


Extended Data Fig. 4 | See next page for caption.

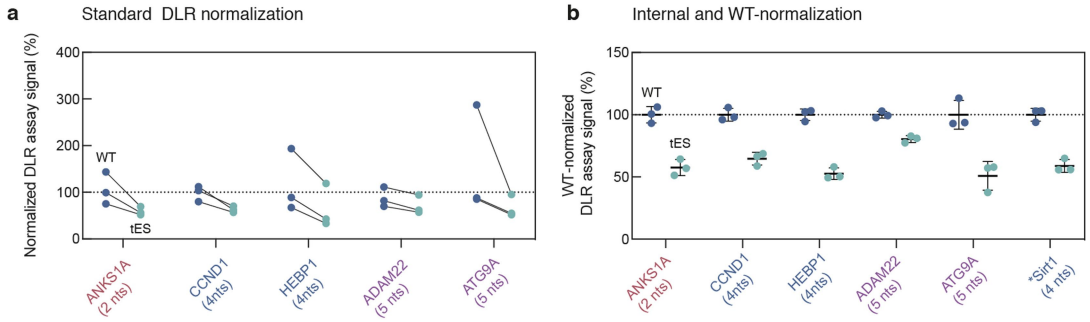
Extended Data Fig. 4 | Biophysical and biochemical characterization of the constructs. a, Individual A_{260} UV melting profiles for the constructs used here. The miR-34a–mSirt1 duplex, miR-34a–mSirt1 trapped ES duplex, miR-34a–complementary-strand duplex and miR-34a single-stranded RNA (ssRNA) were each measured as three technical independent replicates (shown in different colours; $n = 1$). Individual technical replicates are plotted. T_m values are shown as means ± s.d. of fitted T_m values in individual technical replicates ($n = 3$). The other ssRNAs (bottom row) were measured and plotted as individual technical replicates; fitted T_m values are shown with associated confidence intervals of 95% ($n = 1$) as an estimate of the experimental error. Normalized differential melting curves ($\delta A_{260}/\delta T$) are plotted as a function of temperature (in K) (circles) and fitted to Supplementary equation (1a) or (1b) (curves), depending on the molecularity of the system. **b**, EMSA titration profiles for the miR-34a–mSirt1 duplex, miR-34a–mSirt1 trapped ES duplex and miR-34a–complementary-strand duplex, measured as three independent technical replicates. The ratio of bound to total miR-34a 3′-Cy3 is plotted as a function of titrand concentration (circles) and fits a standard binding isotherm (line) (Supplementary equation (2)). The plot centre is the mean; error bars represent 1 s.d. from the three independent replicates. Fitted K_d values along with confidence intervals of 95% are shown as an estimate of the experimental error ($n = 3$). Gel images were acquired by detection of Cy3 fluorescence. During the titration, miR-34a 3′-Cy3 was kept at a constant concentration of 24 nM, setting the sensitivity limit for estimating K_d (Supplementary Fig. 1a–c). mSirt1 and its trapped-ES counterpart are equivalent in their ability to form a stable RNA–RNA duplex with miR-34a. Tighter binding is observed for the complementary strand (48.4 ± 9.5 nM) than for the mSirt1 (124.3 ± 21.7 nM) and trapped-ES mSirt1 (110.3 ± 23.0 nM), providing a control for the dynamic range of K_d estimation. **c**, Equilibrium FBA profiles for mSirt1, mSirt1 trapped ES and a scrambled control, binding to miR-34a-loaded Ago2. The three targets were each measured as three independent replicates and fitted to a standard binding isotherm (line) (Supplementary equation (2)). The plot centre is the mean; error bars represent 1 s.d. from three independent replicates. Fitted K_d values are shown with confidence intervals of 95% (an estimate of the experimental error). As in **c**, mSirt1 and mSirt1 trapped ES are equivalent in their ability to form a stable ternary complex within RISC. The simulated data set (dotted lines) indicate curves corresponding to K_d values ten times lower (red) or ten times higher (green) than the average value for mSirt1 and mSirt1 trapped ES, providing a frame for the amplitude of our experimental error. **d**, Top, northern blot showing the detection of miR-34a loaded in Ago2. Bottom, a standard calibration curve (using naked miR-34a), used to obtain an estimate of miR-34a in RISC. The centre calibration curve was used to calculate R^2. The two outer curves indicate the 95% confidence interval of the calibration-line fit (from a single repeated experiment). The average ratio of Ago2 and miR-34a-loaded Ago2 (both in pmole) was used to obtain the fraction of Ago2 loaded with our guide (roughly 1.5%). The complete lists of fitted parameters for UV melting, EMSA titration, FBA titration and northern blot are in Supplementary Table 1a–d. The complete fitting analyses of UV melting, EMSA titration and FBA titration are in Supplementary Tables 7–9.



Extended Data Fig. 5 | Crystal structure of Ago2 overlaid with REMD ensembles. Superposition of ground state (green) and excited state (orange) conformational ensembles on the Ago2 crystal structure (PDB code 4W5T), with seed sequences aligned to crystallographic miRNA–mRNA positions (red). Although the seed orientations are comparable, the ground-state and excited-state conformations sample different space within Ago2.



Extended Data Fig. 6 | Slow-growth insertion of excited-state RNA into Ago2 predicts the ability of bulged miRNA–mRNA complexes to access an alternative dsRNA-binding mode of Ago2. Slow-growth induced-fit Ago2 structures are compared with existing X-ray structures (whose PDB identification codes are shown in the figure) via structural alignment. **a**, Ago2 after induced fit with ES RNA binds in the PIWI-adjacent groove rather than the PAZ domain. **b**, The *Thermus thermophilus* (Tt) Ago crystal structure similarly shows DNA/RNA-duplex binding in the analogous PIWI-adjacent groove. **c**–**j**, The root mean square deviation (r.m.s.d.) for each indicated pair of Ago structures was measured after structural alignment either of all protein atoms, or excluding the PAZ domain, PIWI loops and helix-7 atoms ('subset aligned'; these excluded atoms still count towards the r.m.s.d.). The subset-aligned structures show that most of the r.m.s.d. 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**, Comparison of slow-growth human Ago2–GS and the existing Ago2 structure (PDB code 4OLA; r.m.s.d. = 2.065 Å (all) and 2.62 Å (subset aligned)). **d**, Comparison of slow-growth Ago2–ES and the 4OLA structure (r.m.s.d. = 1.4 Å (all) and 1.65 Å (subset aligned)). **e**, Comparison of the slow-growth Ago2–trapped ES and the 4OLA structure (r.m.s.d. = 1.9 Å (all) and 2.18 Å (subset aligned)). **f**, Comparison of the slow-growth Ago2–GS with Ago2–ES (r.m.s.d. = 2.1 Å (all) and 2.2 Å (subset aligned)). **g**, Comparison of the slow-growth Ago2–ES with Ago2–trapped ES (r.m.s.d. = 1.6 Å (all) and 1.33 Å (subset aligned)). **h**, Comparison of the slow-growth Ago2–GS and the 6N4O structure (r.m.s.d. = 2.05 Å (all) and 2.065 Å (subset aligned)). **i**, Comparison of the slow-growth Ago2–ES (green) with the 3HM9 structure (r.m.s.d. = 4.52 Å (all)). **j**, Comparison of the slow-growth Ago2–GS with the 3HM9 structure (r.m.s.d. = 3.85Å (all)).



Extended Data Fig. 7 | DLR assay of additional miR-34a targets. We studied five targets of different bulge sizes (see Methods). Individual replicates are plotted as circles; the centre line is the mean; error bars represent 1 s.d from three independent replicates; nts, nucleotides. **a**, Standard DLR normalization (relative to the control condition with no miR-34a duplex transfected). Despite the large variability between replicates, a consistent increase in downregulation (connecting lines) is observed for wild-type (WT) and trapped excited-state (tES) constructs. **b**, When the data sets 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 Table 1 | T_m and K_d fitted parameters

a	UV A_{260} thermal melting	mean T_m (K)	T_m (K)	h	ΔH (kJ mol^{-1})	ΔS (kJ K^{-1} mol^{-1})	R^2
	miR-34a–mSirt1 duplex	322.0 ± 0.8	322.5 ± 0.2	-280.0 ± 17.9	-750.8 ± 398.0	-2.3 ± 1.2	0.9657
			321.1 ± 0.3	-257.4 ± 24.2	-687.2 ± 537.2	-2.1 ± 1.7	0.9339
			322.5 ± 0.3	-254.3 ± 19.3	-681.9 ± 430.3	-2.1 ± 1.3	0.9513
	miR-34a–mSirt1 trapped ES duplex	321.1 ± 0.7	320.3 ± 0.3	-259.2 ± 24.9	-690.3 ± 550.3	-2.2 ± 1.7	0.9311
			321.3 ± 0.7	-249.9 ± 43.2	-667.6 ± 958.5	-2.1 ± 3.0	0.7994
			321.7 ± 0.4	-277.4 ± 29.7	-742.0 ± 660.6	-2.3 ± 2.1	0.9026
	miR-34a-Complentary strand duplex	340.7 ± 0.7	340.9 ± 0.3	-285.6 ± 19.4	-809.5 ± 457.2	-2.4 ± 1.3	0.9657
			339.9 ± 0.5	-321.1 ± 53.9	-907.5 ± 1265.4	-2.7 ± 3.7	0.8243
			341.3 ± 0.3	-349.8 ± 44.8	-992.6 ± 1057.0	-2.9 ± 3.1	0.9023
	miR-34a ssRNA	314.9 ± 1.6	316.3 ± 0.8	-97.5 ± 0.8	-256.4 ± 17.3	-0.8 ± 0.1	0.8369
			315.3 ± 0.8	-89.5 ± 0.8	-234.7 ± 19.2	-0.7 ± 0.1	0.8246
			313.1 ± 0.8	-93.5 ± 0.8	-243.5 ± 18.1	-0.8 ± 0.1	0.8388
	miR-34a-mSIRT1 bulge	-	341.8 ± 1.5	-77.5 ± 1.6	-220.3 ± 38.7	-0.6 ± 0.1	0.6198
	miR-34a-mSIRT1 trapped ES	-	339.6 ± 1.5	-110.1 ± 1.5	-310.9 ± 37.0	-0.9 ± 0.1	0.6477
	mSirt1 ssRNA	-	-				-
	mSirt1 trapped ES ssRNA		304.1 ± 0.8	-107.7 ± 1.5	-272.3 ± 363.7	-0.9 ± 1.2	0.7834
	Complementary strand ssRNA	-	-				-

b	EMSA	K_d (nM)	R^2
	miR-34a–mSirt1 duplex	124.3 ± 21.7	0.9821
	miR-34a–mSirt1 trapped ES duplex	110.3 ± 23.0	0.9732
	miR-34a-Complentary strand duplex	48.4 ± 9.5	0.9755

c	FBA	K_d (nM)	R^2
	miR-34a–mSirt1 duplex	70.4 ± 36.4	0.8791
	miR-34a–mSirt1 trapped ES duplex	85.3 ± 16.9	0.9822
	Scrambled control	-	-

d	Northern blot	%
	Estimated hAgo2 loaded with miR-34a	~1.5

a, Thermal denaturation was monitored by UV absorption. Mean ± s.d. T_m values were obtained from three independent replicates. Also shown are parameters derived from fitting of Supplementary equation (1a) or (1b) (Supplementary Methods). Fitting parameters T_m and h are presented with confidence intervals of 95% (as estimates of the experimental error) (T_m is the melting temperature and $h = \Delta H/RT_m$ (R = 8.31447 J K^{-1} mol^{-1})). Complete fitting details and statistics are in Supplementary Table 7. **b**, **c**, EMSA and FBA. Parameters derived from fitting of Supplementary equation (2) (see Supplementary Methods). K_d values obtained from the fit are presented with confidence intervals of 95% as estimates of the experimental error (n = 3). Complete fitting details and statistics are in Supplementary Tables 8, 9. **d**, Northern blot. The fraction of Ago2 loaded with the guide RNA of interest was estimated by northern blotting (see Supplementary Methods).

nature research

Corresponding author(s): Katja Petzold

Last updated by author(s): Feb 10, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All software and code details are described in the Method and Supplementary Methods sections.
Data analysis	All software and code details are described in the Method and Supplementary Methods sections.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on 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 were a gift from Joanne Weidhaas (Addgene plasmid #78258 and #78259). hAgo2 expressing plasmid was a gift from (Prof. Ian MacRae, Department of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA, USA). The structural models resulting from slow-growth insertion of the GS, ES, and trapped ES RNA into the hAGO2 protein have been deposited in Model Archive (www.modelarchive.org) as ma-bc9uo, ma-z54y4, and ma-g8e5z.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. Cell-based assays were performed in at least 3 independent, biological replicas.
Data exclusions	No data was excluded from analysis.
Replication	Experiments were successfully replicated to ensure that they stably support our findings. The numbers are given for relevant experiments. Cell-based assays were performed at least in 3 independent, biological replicates. Technical replicates were performed where necessary and are stated.
Randomization	Samples were not randomized for analysis.
Blinding	No blinding was performed in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T cells were obtained from ATCC (CRL-11268), Sf9 cells were obtained from Invitrogen (Cat no 11496-015, Lot 1296885).
Authentication	All human cells lines from ATCC are authenticated by STR analysis. Sf9 cells were to our knowledge not authenticated. All cell lines were visually inspected throughout the experiments and can be easily identified through their morphology and growth.
Mycoplasma contamination	HEK293T cells (CRL-11268) were readily used after purchase and therefore were not tested for Mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No misidentified cell lines were used