In vivo mRNA structure regulates miRNA cleavage in *Arabidopsis*

Minglei Yang[1,*], Hugh C. Woolfenden[1,*], Yueying Zhang[1,*], Xiaofeng Fang[1,*], Qi Liu[1], Maria Louisa Vigh[2], Jitender Cheema[1], Xiaofei Yang[1], Matthew Norris[1], Sha Yu[1,3], Alberto Carbonell[4], Peter Brodersen[2], Jiawei Wang[3], Yiliang Ding[1,†]

[1] Department of Cell and Developmental Biology, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom

[2] Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200, Copenhagen N, Denmark

[3] National Key Laboratory of Plant Molecular Genetics (NKLPMG), CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology (SIPPE), Shanghai Institutes for Biological Sciences (SIBS), Shanghai, 200032, People's Republic of China

[4] Instituto de Biología Molecular y Celular de Plantas (Consejo Superior de Investigaciones Científicas-Universidad Politécnica de Valencia), Valencia, 46022, Spain

* These authors contributed equally to this work.
† Corresponding author. Email: yiliang.ding@jic.ac.uk

Abstract: MicroRNA(miRNAs)-mediated cleavage is involved in numerous essential cellular pathways. miRNAs recognize target RNAs via sequence complementarity. In addition to complementarity, *in vitro* and *in silico* studies have suggested that RNA structure may influence the accessibility of mRNAs to miRNA-Induced Silencing Complexes (miRISCs), thereby affecting RNA silencing. However, the regulatory mechanism of mRNA structure in miRNA cleavage remains elusive. Here, we investigated the role of *in vivo* RNA secondary structure in miRNA cleavage by developing the new CAP-STRUCTURE-seq method to capture the intact mRNA structurome in *Arabidopsis thaliana*. This approach revealed that miRNA target sites were not structurally accessible for miRISC binding prior to cleavage *in vivo*. Instead, the unfolding of the target site structure is the primary determinant for miRISC activity *in vivo*. Notably, we found that the single-strandedness of the two nucleotides immediately downstream of the target site, named Target Adjacent structure Motif (TAM), can influence miRNA cleavage but not miRNA binding, thus decoupling target site binding from cleavage. Indeed, TAM formation in the floral homeotic gene, *APETALA2*, can switch on miRNA cleavage, resulting in floral organ identity defects similar to those in loss-of-function *apetala2* mutants. Our findings demonstrate that mRNA structure *in vivo* can regulate miRNA cleavage, providing evidence of mRNA structure-dependent regulation of biological processes.

One Sentence Summary: The single-strandedness of the two nucleotides immediately downstream of the target sites on mRNAs can regulate miRNA cleavage *in vivo*.

Main Text:

MicroRNAs (MiRNAs) are ~21 nucleotide RNAs that impact various aspects of development and stress responses by post-transcriptionally regulating gene expression (*1*). MiRNAs are loaded onto ARGONAUTE proteins (AGO) to form functional post-transcriptional gene silencing effector complexes, miRNA-Induced Silencing Complexes (miRISCs) (*2*). miRISC is guided by the miRNA to target RNAs through sequence complementarity and cleave the target RNAs (*3*, *4*). However, previous studies found that sequence complementarity is not the sole factor dictating miRNA cleavage (*2*). The structure of an RNA has been suggested to influence the silencing efficiency (*5–7*). However, these studies were unable to reveal native RNA structure features adopted through evolution because they indirectly assessed RNA structure by inserting a long sequence that was predicted to form a strong structure, such as a hairpin (*5–7*). Additionally, these structure assessments examined the target site together with its long flanking regions (*5–7*). This led to difficulties in dissecting the individual contributions from the different regions and confounded the identification of a specific RNA structure motif that regulated miRNA cleavage. Furthermore, these *in vitro* and *in silico* studies could not reflect the RNA structure folding status in living cells (*8–10*). Recently, several transcriptome-wide structure probing methods for RNA *in vivo* have been established (*8–10*), which provide powerful tools to understand RNA structures under physiological conditions. However, these methods have not assessed the causal relationship between RNA structure and miRNA cleavage because the methods cannot discern whether the obtained RNA structure information is from endogenous degraded RNAs or from intact RNAs.

Deciphering the *in vivo* relationship between mRNA structure and miRNA cleavage requires *in vivo* structures of target mRNAs before cleavage and the outcome after miRNA-mediated cleavage. To obtain the RNA structure of intact mRNA, we performed *in vivo* selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) chemical probing on *Arabidopsis thaliana* (*A. thaliana*) with optimized conditions (Fig. 1A and Fig. S1A). Next, we enriched the intact mRNAs through the terminator exonuclease treatment (Fig. S1B) (*11*), and then polyA+ purification to remove the degraded mRNAs. We generated two independent biological replicates of (+)SHAPE (samples with SHAPE treatment) and (-)SHAPE (control samples without SHAPE treatment) libraries according to the protocol described previously (*8*, *12*). Between 90–97% of 340~380 million reads were mapped onto mRNAs (Fig. S1C and D) with the reproducibility of the CAP-STRUCTURE-seq library confirmed by comparing the two biological replicates (Fig. S2A and B). Nucleotide occurrence was consistent in both (-)SHAPE and (+)SHAPE libraries (Fig. S1E). To validate CAP-STRUCTURE-seq, we compared the SHAPE reactivity of the 18S rRNA with the corresponding phylogenetic covariance structure (Fig. S3A) and the 3D structure (Fig. S3B). We found that CAP-STRUCTURE-seq can accurately probe RNA structure *in vivo*, and in addition it outperforms the previous dimethyl sulphate (DMS)-based method, STRUCTURE-seq (*8*) (Fig. S1F).

To further validate CAP-STRUCTURE-seq we performed meta-property analyses with over 16,576 transcripts of sufficient mRNA structure information (Fig. S4A). We found that the region immediately upstream of the start codon showed particularly high SHAPE reactivity (Fig. S4B). This result further supports the notion that less structured regions near the start codon may facilitate translation (*13*, *14*). We also found a periodic reactivity trend in the coding sequence (CDS) but it was absent in the untranslated regions (UTRs) (Fig. S4C), which is consistent with

previous studies (*8*, *15*). Similar to a RNase-based structure study in human (*15*), we also observed an asymmetric mRNA structure signature at the exon–exon junction in *A. thaliana* (Fig. S4D). These conserved mRNA structure features suggest that CAP-STRUCTURE-seq successfully provides a global mRNA structure model in plants. Moreover, existing RNA structure methods (*9*, *10*, *16–19*) are not able to discern whether the RNA structure information belongs to the endogenous degraded mRNAs or to the intact mRNAs (Fig. S4E). Additionally, degraded mRNAs are capable of introducing false positive signals in the previous methods (Fig. S4E) (*9*, *10*, *20*, *21*). For instance, the miRNA cleavage products led to a skewed DMS reactivity profile at the miRNA cleavage sites in the previous DMS-based method (*8*) (Fig. S4F), while no such false positive signal occurs in our CAP-STRUCTURE-seq (Fig. S4G). Thus, our CAP-STRUCTURE-seq provides RNA structure information for intact mRNAs while excluding degradation signals (Fig. S4E and G), thereby overcoming the limitations of previous methods that include degradation products (Fig. S4E and F) (*8*, *10*, *18*, *22*, *23*).

Having developed a reliable method to accurately probe the intact mRNA structure, we then focused on estimating the *in vivo* miRNA cleavage efficiency (CE) at a genome-wide scale. Previously, qualitative measurement of miRNA cleavage events was achieved from degradome library analysis (*24–26*). Inspired by the definition of enzymatic activity (*27*), we quantitatively estimated CE by measuring how many degradation products were generated from one unit of substrate mRNA by one unit of miRNA (Methods). We generated degradome libraries to estimate the degradation products (Fig. S5A and Methods) and miRNA-seq libraries to estimate miRNA abundance (Fig. S5A and Methods), with library reproducibility confirmed by comparing the biological replicates (Fig. S2C, S2D and Methods). We then combined the degradome, (-)SHAPE and miRNA-seq libraries to estimate CE (Fig. S5A and Methods). We verified the consistency of our CE against previously reported targets (Table S1). For example, the CE of *AP2* targeted by miRNA172, which has been shown to act through translational repression rather than mRNA cleavage (*28–30*), was zero as expected (Table S1). *SNZ* (Table S1) is another target of miRNA172, and showed no evidence of miRNA cleavage, consistent with the previous result (*31*). In contrast, *TOE2*, which is cleaved by miRNA172, had relatively high CE (Table S1). Additionally, *TAS1a* and *TAS2*, which must be cleaved by miRNA173 to then serve as templates for trans-acting siRNA (tasi-RNA) (*32*), had high CE (Table S1). These observations were consistent with their previous reported biological functions (*28–32*). Then, we systematically examined the relationship between sequence complementarity and CE. Globally, we found that sequence complementarity and CE were uncorrelated (Spearman correlation -0.015, Fig. S5E). In addition, targets with mismatches and/or GU wobble pairs (sequence complementarity penalty score, SCPS>0) were sometimes more effectively cleaved than targets with perfect complementarity (SCPS=0) (Fig. S5E). Our results indicate that besides sequence complementarity between miRNA and mRNA there may be other factors affecting CE, for example, mRNA structure. In summary, both the RNA structure of the intact mRNAs and miRNA cleavage *in vivo* can be quantitatively measured.

With CAP-STRUCTURE-seq elucidating the RNA structure, we could begin to answer the elusive question about whether miRNA target sites were structurally accessible *in vivo*. Since our CAP-STRUCTURE-seq directly measures the *in vivo* structural accessibility via SHAPE reactivity (*33*), we assessed the SHAPE reactivity profiles across the miRNA target sites on the intact mRNAs. SHAPE reactivities of the target sites showed no significant difference from the upstream region (two-sided Mann-Whitney-U test, *P* value is 0.31, Fig. S4G), and were lower than the

downstream region (one-sided Mann-Whitney-U test, *P* value is 0.0038, Fig. S4G). These features indicate that under physiological conditions the target sites are not highly accessible, which may provide a protective mechanism for target sites, mitigating against other cellular ribonucleases. These relatively inaccessible target sites prompted us to ask whether the target site structure affects miRNA cleavage *in vivo*. To address this question, we examined two alternative energetic landscapes associated with the miRISC cleavage process *in vivo*: an enzyme-limiting scenario and a structure-limiting scenario (Fig. 1B). In the enzyme-limiting scenario, the energy barrier ($\Delta G^{\ddagger}_{open}$) between the inaccessible and accessible structural states (i.e., the unfolding of the target site) is lower than the barrier for catalytic cleavage (black line in Fig. 1B). Thus, the target sites equilibrate quickly between inaccessible and accessible structural states during the binding step prior to the catalytic step of miRNA cleavage. In this scenario, the CE would vary with the free energy required to surmount the AGO catalytic barrier, $\Delta G^{\ddagger}_{cutting}$ (Methods), and would be less affected by the RNA structure of the target site. In the structure-limiting scenario, the energy barrier ($\Delta G^{\ddagger}_{open}$) between the inaccessible and accessible structural states is higher than the barrier for cleavage (red line in Fig. 1B). Therefore, the target sites cannot achieve equilibrium binding with miRISC before catalytic cleavage. In this scenario, CE would vary with the free energy of opening the target site structure, $\Delta G^{\ddagger}_{open}$, rather than $\Delta G^{\ddagger}_{cutting}$. We used our *in vivo* structures to computationally approximate these two scenarios and explored a range of flanking lengths upstream and downstream of the target site (Fig. 1C, D and Fig. S6A, B). Analysis of our SHAPE reactivity-informed structures revealed that, for most flank sizes, CE anti-correlated with $\Delta G^{\ddagger}_{open}$ with a broad maximum centered around flanks of 50 nucleotides upstream and downstream (Spearman correlation of -0.23, $P = 6.3e\text{-}9$) (Fig. 1C and Fig. S6A). However, for most flank sizes, CE had no correlation with $\Delta G^{\ddagger}_{cutting}$ (Fig. 1D and Fig. S6B), contrary to the reaction kinetics where the energy barrier is anti-correlated with reaction processivity. These results indicate that the unfolding of target site structure is the rate-limiting step that determines miRISC activity *in vivo*. Furthermore, this structure-limiting scenario reveals that the ribonuclease AGO undergoes "sticky regime" activation (*34*), where substrate mRNAs associate and dissociate with AGO more slowly than they are being cleaved. This scenario further explains previous AGO RIP-seq results, where target transcripts were only captured in the slicing-defective AGO1 mutant ($AGO1^{DAH}$) but not in wild type AGO (*35*). We found that $\Delta G^{\ddagger}_{open}$ anti-correlated with the enrichment ratio of target RNAs associated with AGO1 protein from previous $AGO1^{DAH}$-RIP-seq results (*35*) (Spearman correlation -0.21, $P = 0.05$). In contrast, the free energy of binding of the miRNA-target duplex (ΔG_{duplex}) and $\Delta G^{\ddagger}_{cutting}$ show no correlations (Spearman correlation 0.06 with $P = 0.32$ and -0.11 with $P = 0.16$, respectively). These observations suggest that the target sites are not structurally accessible *in vivo*, but rather the unfolding of the target site structure is the primary determinant for target RNA processing by AGO.

Having revealed that the target site structure affects cleavage *in vivo*, we then investigated whether the structure of the target site flanking regions is involved with miRNA cleavage. We assessed the RNA secondary structure by separating the miRNA targets into non-cleaved (zero CE) and cleaved (positive CE) groups. We found higher SHAPE reactivity at the +1 and +2 nucleotides immediately downstream of target sites in the cleaved group relative to the non-cleaved group (Fig. 2A), suggesting that these two nucleotides are more single-stranded than their neighbors. To confirm this observation, we used the SHAPE reactivity with the ViennaRNA *RNAfold* utility (*36*) to calculate the base pairing probabilities (BPPs) for each nucleotide. We found that the BPPs of the +1 and +2 nucleotides were much lower than their neighboring

nucleotides (Fig. S7), indicating an increased likelihood of single-strandedness in the cleaved group compared to the non-cleaved group. Furthermore, the single-strandedness of the two nucleotides was unlikely to be due to sequence composition (Fig. 2B) or AT content (Fig. 2C) because there was no difference between the non-cleaved and cleaved groups. Our results reveal that a secondary structure feature, specifically single-strandedness of the two nucleotides adjacent to the 3' end of the miRNA target site, generally exists *in vivo* in intact mRNAs that will undergo cleavage. We named this structure feature 'Target Adjacent structure Motif' (TAM).

To explore the functional role of TAM in miRNA cleavage, we designed a structure assay (Methods) by concatenating miRNA156 target sites (20 nt) with a designed stable structure module (either a G-quadruplex structure or a stem-loop) to mimic the base-pairing state of the two nucleotides immediately downstream of the target site (Fig. 3A and Fig. S8A). To maintain the single-strandedness of the TAM we inserted two Adenines (AA), or "slippery sequence", immediately between the target site and the designed stable structure module (Fig. 3A and Fig. S8A). We confirmed the formation of TAM *in vivo* by using capillary electrophoresis (*37*) to resolve the *in vivo* RNA structure for each design (Methods, Fig. 3B and Fig. S8B). We then assessed the miRNA cleavage *in vivo* by measuring the levels of non-cleaved substrate mRNA. We found that the non-cleaved mRNA level in the substrates with TAM was significantly lower than those without TAM (Fig. 3C and Fig. S8C), which suggested increased cleavage when TAM was present. To further confirm whether the presence of TAM exclusively determines target cleavage, we performed an *in vitro* AGO cleavage assay using HA immuno-affinity-purified wildtype AGO protein and we found that the target RNA was cleaved only when TAM was present (Fig. 3D). Our results reveal that TAM is essential for miRISC nuclease activity. One might expect TAM in the target mRNA to facilitate AGO binding instead of directly triggering the nuclease activity of AGO proteins. To test the possibility that TAM affects target binding, we conducted an *in vivo* binding assay (Methods) by using the slicing-defective AGO1 mutant ($AGO1^{DAH}$). We found that AGO1 was able to bind the target RNAs with the same binding affinity irrespective of whether the TAM was present or absent (Fig. 3E and F). Therefore, our data reveal that TAM regulates miRISC cleavage activity rather than affecting target binding.

Following our investigation into the functional role of TAM, we further examined the biological impact of TAM by using a native target gene. An important flower development gene, *APETALA2* (*AP2*), is targeted by miRNA172 and the interaction between *AP2* and miRNA172 leads to translational inhibition rather than cleavage (*29*). In our CE results, *AP2* cleavage efficiency was zero, representing non-cleavage. Interestingly, the *in vivo* RNA structure of *AP2* in our CAP-STRUCTURE-seq shows that the two nucleotides downstream of the target site (+1 and +2 positions) were base-paired with the 879^{th} and 880^{th} nucleotides in *AP2*, indicating that TAM was absent (Fig. 4A). To assess whether the non-cleavage of *AP2* is due to the absence of TAM, we introduced TAM by mutating the A at the 879^{th} position with G, U, and C (labelled A879G, A879U, A879C) to disrupt the base-pairing states while preserving the codon. We then transformed the wild-type construct without TAM (non-TAM) and the synonymous mutation constructs with TAM (TAM) into protoplasts of the stable *MIR172* over-expression line. We measured the miRNA cleavage of *AP2* by primers designed to span the cleavage site to detect the non-cleaved mRNAs (Fig. 4B). We found increased cleavage of *AP2* (measured by decreased levels of non-cleaved *AP2*) when the TAM was present compared to when TAM was absent (Fig. 4B), indicating the TAM can switch on the miRNA172 cleavage of *AP2*. Therefore, our results suggest that the TAM is able to change the cleavage fate of target genes.

To further assess the genetic effects of TAM, we performed stable transformations of these non-TAM and TAM constructs (A879G, A879U, A879C) into the loss-of-function *ap2-5* mutant. As a floral homeotic gene, *AP2* determines floral organ identity and the *ap2-5* mutant results in the replacement of the perianth organs by the reproductive organs (Fig. 4C and D). In the non-TAM transgenic lines, floral organ defects were restored by complementing *ap2-5* mutants with the wild-type *AP2,* in which no TAM formed (Fig. 4E and F). In contrast to the non-TAM transgenic lines, the severe floral organ defects of *ap2-5* could not be restored when *ap2-5* mutants were transformed with the TAM constructs (A879G, A879U, A879C) (Fig. 4G-4L). We further assessed the miRNA cleavage efficacy of *AP2* where the phenotypic restorations are different between non-TAM and TAM transgenic lines. We measured the levels of non-cleaved *AP2* in these transgenic lines. We found increased cleavage of *AP2* (measured by decreased levels of non-cleaved *AP2*) in the TAM transgenic plants compared to the non-TAM transformants (Fig. 4M), while the miRNA172 levels were largely unchanged between non-TAM and TAM transgenic plants (Fig. 4N). Taken together, our results indicate that switching on TAM is sufficient to turn on the miRNA cleavage of *AP2*, and so significantly affects plant development.

In summary, by deciphering intact mRNA structures *in vivo* through CAP-STRUCTURE-seq, we found that miRNA target sites are not structurally accessible *in vivo* and we demonstrated that the unfolding of the miRNA target site structure predominantly affects miRISC activity *in vivo*. Furthermore, we discovered that the native RNA structure motif, TAM, is sufficient to regulate miRNA cleavage *in vivo*. The mechanism that we found here provides evidence of mRNA structure-dependent regulation of biological processes *in vivo*. Our study reveals that *in vivo* mRNA structure serves as an additional regulator of miRISC activity, which will facilitate the biotechnological engineering of gene silencing, and possibly provide an additional avenue towards crop improvement.

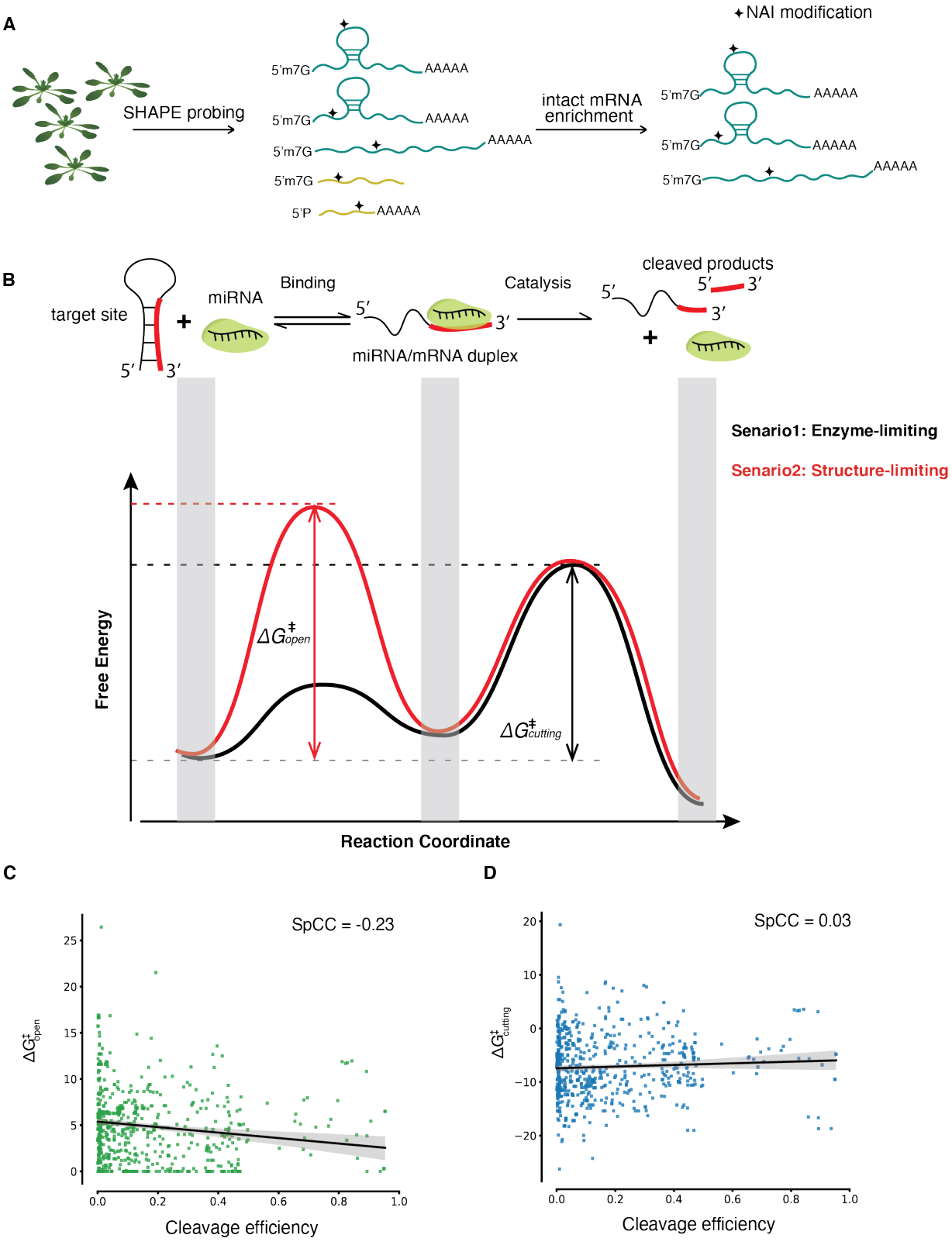

Fig. 1. Measurement of *in vivo* RNA structure of intact mRNA and *in vivo* miRNA cleavage efficiency. A, Schematic of (+)SHAPE sequencing library generation showing NAI treatment, nucleotide modification and purification of intact mRNA steps. *A. thaliana* etiolated seedlings were treated with NAI. After extraction of total RNA, degraded mRNAs (dark yellow) were removed leaving intact mRNAs characterized by 5'CAP and 3' poly(A) tails (dark blue). cDNAs were obtained and subjected to an established library construction. **B,** The miRISC cleavage reactions include target binding and cleavage catalysis. Two alternative scenarios demonstrate the energetic landscape of miRNA cleavage (**black** and **red**). In the enzyme-limiting scenario (**black**), target site structure equilibrates quickly between inaccessible (closed) and accessible (open) states in the binding step compared to the catalytic step of miRNA cleavage. In this scenario, the apparent activation energy is $\Delta G^{\ddagger}_{cutting}$, which measures the energy required to raise the initial substrate target RNA to the transition catalysis-compatible state. Alternatively, in the structure-limiting scenario (**red**), the target site cannot achieve equilibrium binding before cleavage. In this scenario, the energy barrier between the target site and the transient state is higher than the barrier for cleavage. And the apparent activation energy is equal to $\Delta G^{\ddagger}_{open}$, which measures the energy required to open the target site structure. **C,** Spearman correlation between $\Delta G^{\ddagger}_{open}$ and cleavage efficiency (647 target sites with the upstream and downstream flank lengths of 50 nucleotides, P = 6.3e-9 ***). **D,** A similar analysis to **C**, but for $\Delta G^{\ddagger}_{cutting}$ and cleavage efficiency (P = 0.46).

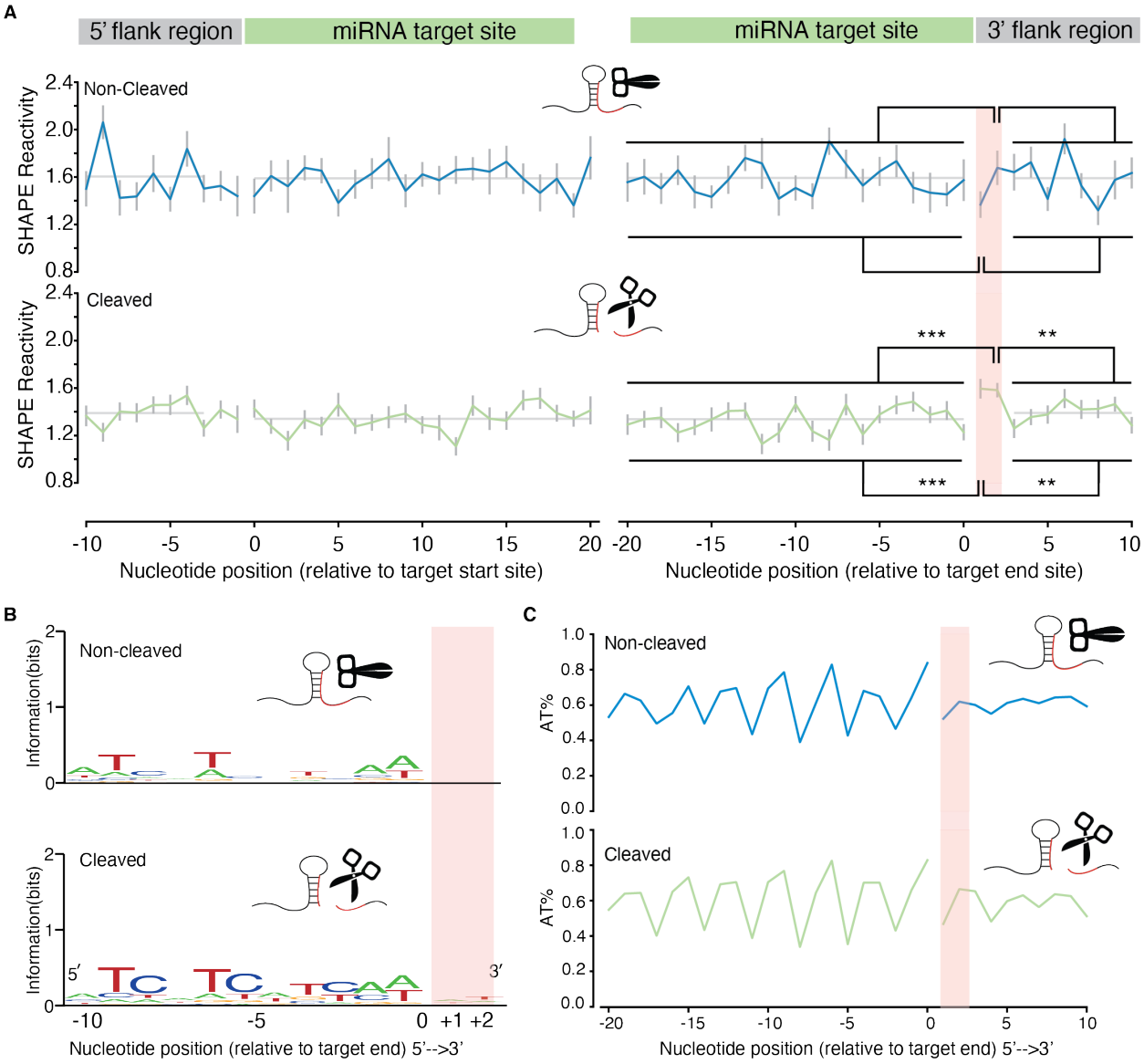

Fig. 2. Discovery of a structure motif: Target Adjacent Structure Motif (TAM), which is independent of sequence composition. A, SHAPE reactivity profiles for miRNA target sites in the non-cleaved group (387 flanked target sites with reactivity values) and cleaved group (567 flanked target sites with reactivity values). The profiles show the per nucleotide mean +/- SEM across transcripts, aligned by target site start (left panels) and end position (right panels). Two nucleotides in the cleaved group (positive CE group), immediately downstream of target sites (TAM region), show significantly higher SHAPE reactivities compared to their neighbors (by Mann-Whitney-U tests). Compared to the upstream region (target sites), P = 4.8e-7*** for both 1st and 2nd nucleotides; Compared to the downstream region, P = 3.9e-3** for both 1st and 2nd nucleotides. The two individual nucleotides of the TAM region in the non-cleaved group (zero CE group) are not significantly higher than their neighbors by Mann-Whitney-U tests. **B,** Sequence composition around the target sites for the total cleaved (675 target sites) and non-cleaved groups (571 target sites). **C,** AT content around the target sites for the total cleaved (675 target sites) and non-cleaved (571 target sites) groups.

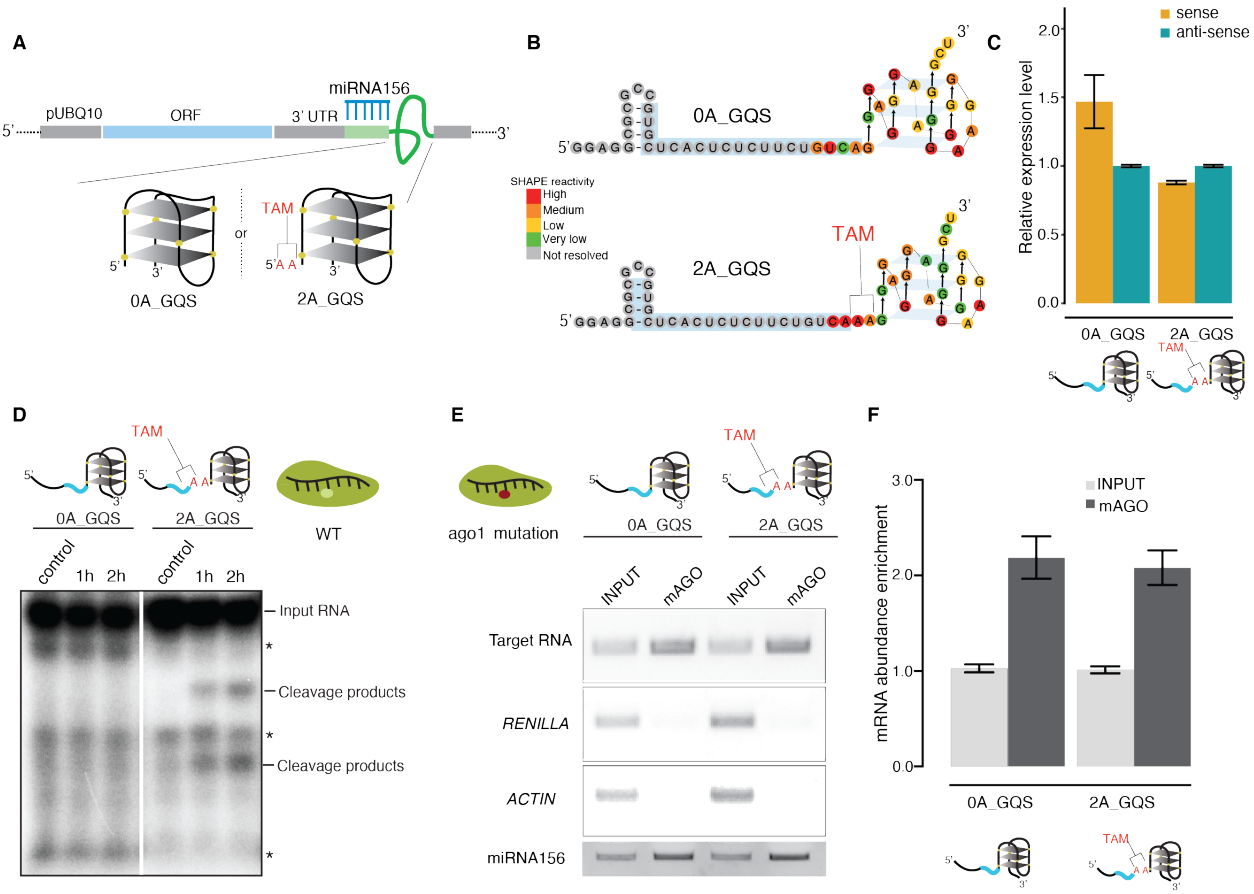

Fig. 3. Validation of TAM functionality by a designed structure assay. A, Cartoon representation of the protoplast transformation assay to validate the TAM functionality using a designed structure assay. GQS refers to a G-quadruplex. The miRNA156 target sites (blue comb) followed by 0 or 2 Adenines (As) and ending with a GQS. The prefixes, "0A" and "2A", indicate the number of Adenines. **B,** *In vivo* RNA structures of 0A_GQS and 2A_GQS. **C**, The non-cleaved mRNA abundance for the substrates in **B** was measured by qRT-PCR (dark yellow bars). The antisense target sites were used as controls (teal bars). Data are mean +/- SEM from three independent biological replicates. **D,** *In vitro* AGO1 cleavage assay shows that target RNA is cleaved when the TAM is present. The target RNAs were incubated for 1h or 2h, where two cleavage products were present in the target with TAM on the X-ray film (as indicated). The asterisk indicates the background bands present in both control and experiment groups. **E,** *In vivo* AGO1 binding assay shows no difference between target RNAs with TAM and without TAM. *RENILLA* and *ACTIN* were used as the control. miRNA156 levels were measured in all the samples. **F,** The RNA abundance enrichment in **E** was quantified by amplicon intensities and normalized by input. Data are mean +/- SEM from three independent biological replicates.

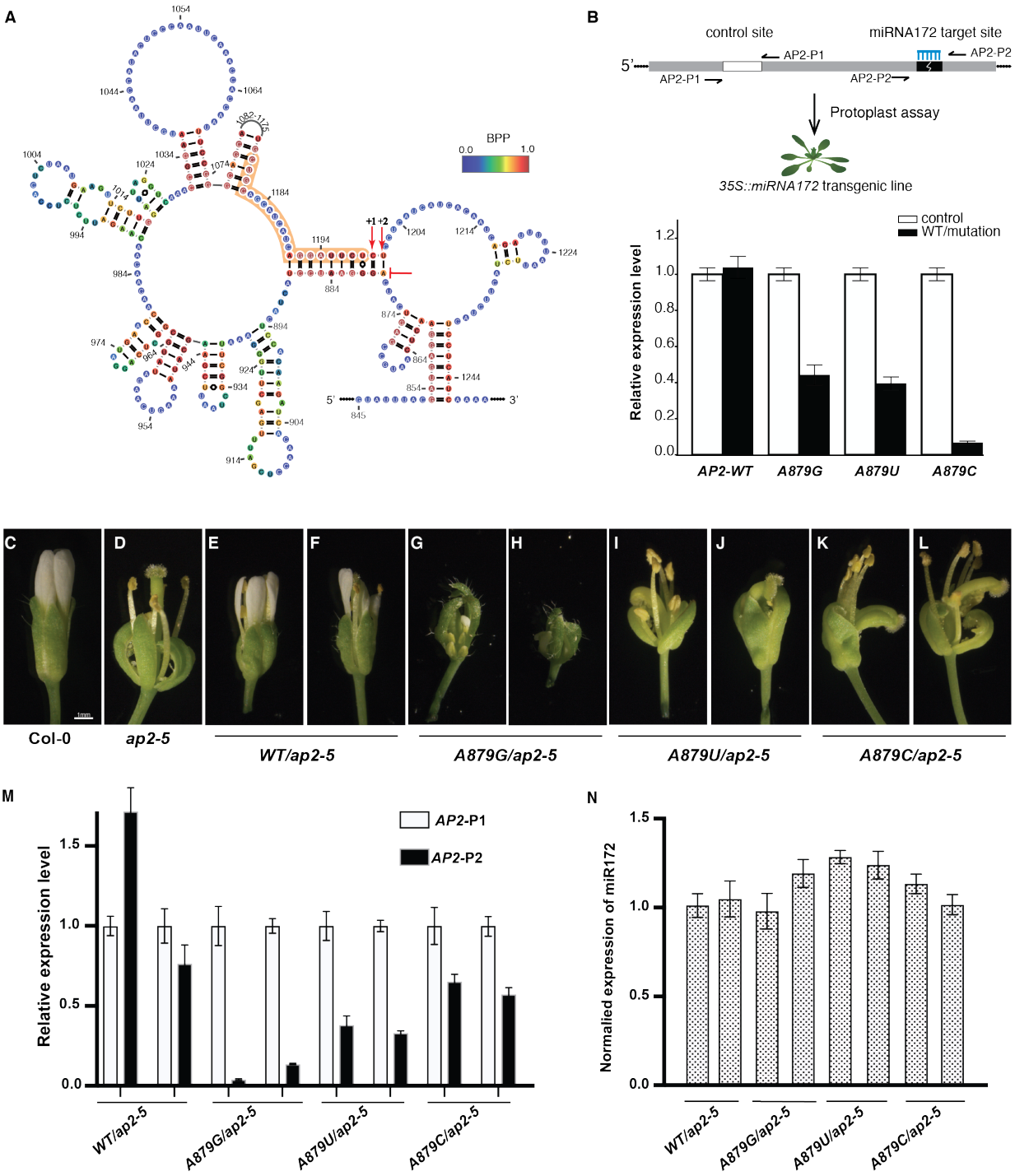

Fig. 4. Introduction of TAM can switch on the miRNA172 cleavage of *AP2*. A, *In vivo* RNA structure model of *AP2* showing the miRNA172 target site (highlighted in orange) and the double-strandedness in the TAM region as indicated by the red arrows. The nucleotides are colored according to base-paring probability (BPP). **B,** The primers, AP2-P1, measured the level of both non-cleaved and degraded *AP2*, and were used as control (white bars). The primers, AP2-P2, that spanned the cleavage site enabled the level of non-cleaved *AP2* to be measured (black bars). Adenine at position 879 (A879) was mutated to Guanine (A879G), Uracil (A879U) or Cytosine (A879C) without changing the codon. This synonymous mutation (labelled as a red flat arrow in **A**) resulted in increased cleavage (a decrease of RNA level, black bars). The experiment was conducted in a stable transgenic line with constant overexpression of the miRNA172 gene. **C,** The wild type with complete flower. **D,** The *AP2* loss-of-function mutant, *ap2-5*, with severe floral defects. **E-F,** Genetic complementation analysis. The severe floral defects in *ap2-5* were restored by complementing *ap2-5* with the wild-type *AP2,* in which no TAM formed (two independent representative lines **E** and **F**). **G-L,** Genetic complementation analysis. The severe floral defects in *ap2-5* could not be restored by complementing *ap2-5* with the mutated *AP2,* in which TAM formed (two independent representative lines per each construct, **G-L**). **M**, The corresponding levels of non-cleaved *AP2* in **G-L** were distinctly reduced (black bars), different from those in **E-F**. **N,** The levels of miRNA172 remained largely unchanged throughout all the transgenic lines (hatched bars). All the bar plots are mean +/- SEM from three independent biological replicates.

References and Notes

1. W. Fang, D. P. Bartel, The Menu of Features that Define Primary MicroRNAs and Enable De Novo Design of MicroRNA Genes. *Mol. Cell*. **60**, 131–145 (2015).

2. Y. Yu, T. Jia, X. Chen, The 'how' and 'where' of plant microRNAs. *New Phytol.* **216**, 1002–1017 (2017).

3. C. Zhang, D. W. K. Ng, J. Lu, Z. J. Chen, Roles of target site location and sequence complementarity in trans-acting siRNA formation in Arabidopsis. *Plant J.* (2012).

4. Q. Liu, F. Wang, M. J. Axtell, Analysis of complementarity requirements for plant MicroRNA targeting using a Nicotiana benthamiana quantitative transient assay. *Plant Cell*. **26**, 741–53 (2014).

5. S. L. Ameres, J. Martinez, R. Schroeder, Molecular Basis for Target RNA Recognition and Cleavage by Human RISC. *Cell*. **130**, 101–112 (2007).

6. M. Kertesz, N. Iovino, U. Unnerstall, U. Gaul, E. Segal, The role of site accessibility in microRNA target recognition. *Nat. Genet.* **39 VN-r**, 1278–1284 (2007).

7. D. Long, R. Lee, P. Williams, C. Y. Chan, V. Ambros, Y. Ding, Potent effect of target structure on microRNA function. *Nat. Struct. Mol. Biol.* **14**, 287–294 (2007).

8. Y. Ding, Y. Tang, C. K. Kwok, Y. Zhang, P. C. Bevilacqua, S. M. Assmann, In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*. **505**, 696–700 (2014).

9. S. Rouskin, M. Zubradt, S. Washietl, M. Kellis, J. S. Weissman, Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature*. **505 VN-**, 701–705 (2014).

10. R. C. Spitale, R. A. Flynn, Q. C. Zhang, P. Crisalli, B. Lee, J. W. Jung, H. Y. Kuchelmeister, P. J. Batista, E. A. Torre, E. T. Kool, H. Y. Chang, Structural imprints in vivo decode RNA regulatory mechanisms. *Nature*. **519**, 486–490 (2015).

11. V. Pelechano, W. Wei, L. M. Steinmetz, Genome-wide quantification of 5'-phosphorylated mRNA degradation intermediates for analysis of ribosome dynamics. *Nat Protoc.* **11**, 359–376 (2016).

12. Y. Ding, C. K. Kwok, Y. Tang, P. C. Bevilacqua, S. M. Assmann, Genome-wide profiling of in vivo RNA structure at single-nucleotide resolution using structure-seq. *Nat. Protoc.* **10**, 1050–1066 (2015).

13. S. M. Studer, S. Joseph, Unfolding of mRNA Secondary Structure by the Bacterial Translation Initiation Complex. *Mol. Cell*. **22**, 105–115 (2006).

14. D. H. Burkhardt, S. Rouskin, Y. Zhang, G. W. Li, J. S. Weissman, C. A. Gross, Operon mRNAs

are organized into ORF-centric structures that predict translation efficiency. *Elife*. **6** (2017).

15. Y. Wan, K. Qu, Q. C. Zhang, R. a Flynn, O. Manor, Z. Ouyang, J. Zhang, R. C. Spitale, M. P. Snyder, E. Segal, H. Y. Chang, Landscape and variation of RNA secondary structure across the human transcriptome. *Nature*. **505**, 706–9 (2014).

16. R. A. Flynn, Q. C. Zhang, R. C. Spitale, B. Lee, M. R. Mumbach, H. Y. Chang, Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat. Protoc.* **11**, 273–290 (2016).

17. J. Talkish, G. May, Y. Lin, J. L. Woolford, C. J. McManus, Mod-seq: High-throughput sequencing for chemical probing of RNA structure. *RNA*. **20**, 713–720 (2014).

18. M. Zubradt, P. Gupta, S. Persad, A. M. Lambowitz, J. S. Weissman, S. Rouskin, DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods*. **14**, 75–82 (2016).

19. N. A. Siegfried, S. Busan, G. M. Rice, J. A. E. Nelson, K. M. Weeks, RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods*. **11**, 959–965 (2014).

20. J. Talkish, G. May, Y. Lin, J. L. Woolford, C. J. McManus, Mod-seq: high-throughput sequencing for chemical probing of RNA structure. *RNA*. **20**, 713–20 (2014).

21. R. A. Flynn, Q. C. Zhang, R. C. Spitale, B. Lee, M. R. Mumbach, H. Y. Chang, Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat. Protoc.* **11**, 273–290 (2016).

22. J. Talkish, G. May, Y. Lin, J. L. Woolford, C. J. McManus, Mod-seq: High-throughput sequencing for chemical probing of RNA structure. *Rna*. **20**, 713–720 (2014).

23. M. J. Smola, K. M. Weeks, In-cell RNA structure probing with SHAPE-MaP. *Nat. Publ. Gr.* **13** (2018), doi:10.1038/nprot.2018.010.

24. D.-H. Jeong, S. R. Thatcher, R. S. H. Brown, J. Zhai, S. Park, L. a Rymarquis, B. C. Meyers, P. J. Green, Comprehensive investigation of microRNAs enhanced by analysis of sequence variants, expression patterns, ARGONAUTE loading, and target cleavage. *Plant Physiol.* **162**, 1225–45 (2013).

25. J. Zhai, S. Arikit, S. A. Simon, B. F. Kingham, Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. *Methods*. **67**, 84–90 (2014).

26. K. R. Franke, S. A. Schmidt, S. Park, D. H. Jeong, M. Accerbi, P. J. Green, Analysis of Brachypodium miRNA targets: Evidence for diverse control during stress and conservation in bioenergy crops. *BMC Genomics*. **19** (2018).

27. NC-IUB, Units of enzyme activity. *Eur. J. Biochem.* **97**, 319–320 (1979).

28. M. J. Aukerman, H. Sakai, Regulation of Flowering Time and Floral Organ Identity by a MicroRNA and Its APETALA2 -Like Target Genes. *Plant Cell*. **15**, 2730–2741 (2003).

29. S. Li, L. Liu, X. Zhuang, Y. Yu, X. Liu, X. Cui, L. Ji, Z. Pan, X. Cao, B. Mo, F. Zhang, N. Raikhel, L. Jiang, X. Chen, MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in arabidopsis. *Cell*. **153**, 562–574 (2013).

30. X. Chen, A MicroRNA as a Translational Repressor of APETALA2 in Arabidopsis Flower Development. *Science (80-.).* **303**, 2022–2025 (2004).

31. R. Schwab, J. F. Palatnik, M. Riester, C. Schommer, M. Schmid, D. Weigel, Specific effects of microRNAs on the plant transcriptome. *Dev. Cell*. **8**, 517–527 (2005).

32. M. Yoshikawa, A. Peragine, M. Y. Park, R. S. Poethig, A pathway for the biogenesis of trans - acting siRNAs in Arabidopsis. *Genes Dev.*, 2164–2175 (2005).

33. J. L. McGinnis, J. A. Dunkle, J. H. D. Cate, K. M. Weeks, The mechanisms of RNA SHAPE chemistry. *J. Am. Chem. Soc.* **134**, 6617–6624 (2012).

34. N. Bisaria, I. Jarmoskaite, D. Herschlag, Lessons from Enzyme Kinetics Reveal Specificity Principles for RNA-Guided Nucleases in RNA Interference and CRISPR-Based Genome Editing. *Cell Syst.* **4** (2017), pp. 21–29.

35. A. Carbonell, N. Fahlgren, H. Garcia-Ruiz, K. B. Gilbert, T. A. Montgomery, T. Nguyen, J. T. Cuperus, J. C. Carrington, Functional Analysis of Three Arabidopsis ARGONAUTES Using Slicer-Defective Mutants. *Plant Cell*. **24**, 3613–3629 (2012).

36. R. Lorenz, I. L. Hofacker, P. F. Stadler, RNA folding with hard and soft constraints. *Algorithms Mol. Biol.* (2016).

37. C. K. Kwok, Y. Ding, Y. Tang, S. M. Assmann, P. C. Bevilacqua, Determination of in vivo RNA structure in low-abundance transcripts. *Nat. Commun.* **4**, 1–12 (2013).

Acknowledgments: We thank Dame Prof. Caroline Dean (John Innes Centre), Prof. Giles Oldroyd (SLCU, Cambridge), Dr. Desmond Bradley (John Innes Centre) and the group members in Ding lab for discussions with this work. We thank Mr. Mirko Ledda (UC. Davis) for discussions on data analysis. We thank Jim Carrington's lab for providing us with the AGO construct for performing the binding assay.

Funding: The experiments here were supported by the Biotechnology and Biological Sciences Research Council [BB/L025000/1] and a European Commission Horizon 2020 European Research Council (ERC) Starting Grant [680324]. **Author contributions:** M.L.Y. and Y.L.D. conceived the study. M.L.Y., Y.Y.Z. and Y.L.D. designed the experiments with assistance from Y.S., A.C., P.B. and J.W.W. M.L.Y. designed the analyses. M.L.Y, Y.Y.Z., Q.L., M.L.V., X.F.Y. and X.F.F. performed the experiments with assistance from P.B., J.W.W. and Y.L.D. H.C.W., J.C. and M.N. performed the analyses. M.L.Y. wrote the paper with input from all authors. M.L.Y., H.C.W., Y.Y.Z., and X.F.F. contributed equally to this work. Correspondence should be addressed to Y.L.D. (yiliang.ding@jic.ac.uk). **Competing interests:** The authors have no competing interests. **Data and materials availability:** Sequencing data are deposited in the Sequence Read Archive (SRA) on the NCBI website under the accession number SRR8444115.

Materials and Methods

Plant materials and growth conditions

A. thaliana seeds of the Columbia (Col-0) and the *xrn4* mutant accession (*38*, *39*) were sterilized with 70% (v/v) ethanol and plated on half-strength Murashige and Skoog medium (1/2 MS). The plates were wrapped in foil and stratified at 4℃ for 3–4 days and then grown in a 22–24℃ growth chamber for 5 days.

Degradome library construction

Five-day-old *A. thaliana* etiolated seedlings were grown as described above. They were collected and immediately frozen in liquid nitrogen and stored at −80°C. The seedlings were ground into powder. Total RNA was extracted using RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. On-column DNAaseI treatment was carried out according to RNase-Free DNase Set (Qiagen). To construct the Illumina library for degradome analysis, polyA+ selection was carried out using the Poly(A)Purist Kit (Ambion™). Selectively captured polyadenylated RNAs (1μg) were ligated directly to an DNA/RNA hybrid adapter (5'-CTACAC GACGCTrCrUrUrCrCrGrArUrCrUrNrNrN-3') using T4 RNA ligase (NEB) at 37°C for 30 minutes. The ligated RNAs were subjected to RT by SuperScript III First-Strand Synthesis System (Invitrogen) with random hexamers fused with Illumina TruSeq adapters (5'-CAGACGTGTGCTCTTCCGATCTNNNNNN-3'). PCR amplification was performed on the ligated cDNA using Illumina TruSeq Primers. Two different barcode indices were used for two degradome biological replicates. The final dsDNA degradome libraries were subjected to next-generation sequencing on Illumina HiSeq 4000.

miRNA library construction

The same seedling samples stored at −80°C, as described above, were ground into powder using liquid nitrogen. Total RNA was extracted using mirVana miRNA Isolation Kit (Ambion™, Austin, TX, USA) following the manufacturer's instructions. The integrity analysis was performed on a Bioanalyzer by the Beijing Genomics Institute (BGI), Shenzhen, China, which also performed the library construction according to standard protocols.

Gel-based 18S rRNA structure probing

The gel-based method of structure probing used the same *in vivo* total RNA pools as for CAP-STRUCTURE-seq. To accomplish gel-based structure probing, reverse transcription was performed using 18S gene-specific DNA primers with 5' end labelled Cy5 (TAGAATTACTACGGTTATCCGAGTA). The whole procedure was performed according to Ding, et.al(*8*). Each gel was detected by Typhoon FLA 9500 (GE Healthcare).

(+)SHAPE and (-)SHAPE CAP-STRUCTURE-seq library construction

We modified the *in vivo* chemical probing protocol (*8*) by changing the reagent from dimethyl sulphate (DMS) to the SHAPE reagent, 2-methylnicotinic acid (NAI). NAI was prepared

as reported previously (*40*). Briefly, five-day-old *A. thaliana* etiolated seedlings were suspended and completely covered in 20 ml 1X SHAPE reaction buffer (100mM KCl, 40mM HEPES (pH7.5) and 0.5mM MgCl2) in a 50 ml Falcon tube. NAI was added to a final concentration of 144mM and the tube swirled on a shaker (1,000rpm) for 15min at room temperature (22°C). This NAI concentration and reaction time had been optimized to allow NAI to penetrate plant cells and modify the RNA *in vivo* under single-hit kinetics conditions (Fig. S1A). After quenching the reaction with freshly prepared dithiothreitol (DTT), the seedlings were washed with deionized water and immediately frozen with liquid nitrogen and ground into powder. Total RNA was extracted using RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions, followed by on-column DNaseI treatment in accordance with the manufacturer's protocol. The control group was prepared using DMSO (labelled as (-)SHAPE), following the same procedure as described above.

To capture the structure information around the cleavage site of miRNA target genes, we adopted the feature of 5PSeq (*11*). The whole CAP-STRUCTURE-seq procedure is illustrated in Fig. 1A. In our method, the (+)SHAPE and (-)SHAPE RNA samples were treated with Terminator™ 5′-Phosphate-Dependent Exonuclease (TER51020, EPICENTRE co.), which processively digests RNA with 5′-monophosphate ends, thereby leaving mRNAs with 5'cap structures. Following the 5'cap enrichment, polyA+ selection was carried out using the PolyA purist Kit (AmbionTM) leaving intact (pre-cleaved) mRNAs with enriched 5'cap and 3'poly(A) tails. The resultant mRNAs were subjected to library construction following the STRUCTURE-seq procedure. The name of CAP-STRUCTURE-seq refers to 5' **cap**-enriched and 3' poly(A)-enriched RNA **structure** **seq**uencing.

Degradome analysis

Raw reads were processed to remove 5'and 3' adapter sequences. Degradome reads were mapped to the TAIR10 transcript reference and a degradome density file was generated. The degradation level of target genes was normalized by reads per kilobase per million mapped reads (RPKM).

miRNA-seq analysis

The small RNA sequences were processed by BGI to filter out the 5′ adapter sequences, 3′ adapter sequences and low-quality reads. We mapped two biological replicates against 253 miRNA sequences confidently annotated as *A. thaliana* mature miRNAs (*41*). We used Bowtie (*42*) for the mapping using the command 'bowtie -f -a -S --best --strata -v 1'. pysam (*42*) was used to count the mapped reads.

CAP-STRUCTURE-seq analysis

We merged the biological replicates of the transcript-level reverse transcription (RT) stop counts to obtain a single (-)SHAPE library and a single (+)SHAPE library. We calculated the SHAPE reactivity using a slightly modified version of the formula in Ding et al.(*8*),

$$\text{SHAPE reactivity}_i = \frac{\log(1+P_i)}{\sum_i \log(1+P_i)} - \alpha \frac{\log(1+M_i)}{\sum_i \log(1+M_i)},$$

where *Pi* is the (+)SHAPE RT count and *Mi* is the (-)SHAPE RT count at nucleotide *i*. The factor, α ($= \min(1, \sum_i \log(1+P_i) / \sum_i \log(1+M_i))$ is a simple library size correction factor. Setting α=1 recovers the reactivity formula in Ding *et al*.(*8*). The reactivities were then normalised using the box-plot method(*43*). For the SHAPE reactivity profiles, we extracted values in the 50 nts upstream and downstream of target sites and calculated a per nucleotide mean and SEM.

The cleavage efficiency (CE) calculation

miRNAs regulate mRNAs through translational repression, mRNA de-stabilization and mRNA cleavage. In animals, it has been suggested that translational repression is prevalent, which can then be followed by de-stabilization, such as shortening of the poly(A) tails at the 3' end and removal of the cap at the 5' end (*44*). In plants, cleavage is the dominant pathway for miRNAs to regulate their target mRNAs (*45*). In order to compare the ability of different miRNAs to cleave their targets on a global scale, we need to quantify the cleavage efficiency (CE) of miRNAs at their target sites.

Our CE calculation is based on two underlying facts (*41*, *46*): miRNA-mediated cleavage is the major mRNA turnover pathway for target genes; the 5' cleaved products are located within binding sites, which are transiently stable. Therefore, the degradation signal within target sites reflects the cleavage products from miRISC cleavage. These two facts were also confirmed by our analysis below.

Firstly, to confirm that the degradation signal within target sites is mainly from miRNA-mediated cleavage, we mapped the 5'end of our wildtype (WT) *A. thaliana* degradome reads to previously validated cleavage sites (*41*). We found that most of the read ends were mapped at the tenth nucleotide of the miRNA complementary sites (Fig. S5C), which provides strong evidence that miRNA-mediated cleavage is the major mRNA turnover pathway for target genes. Secondly, to confirm that the 5' cleaved products are transiently stable, we mapped the read ends of the cleavage products in the *xrn4* mutant to previously validated cleavage sites (*41*). The cleavage site distribution in the *xrn4* mutant exhibited the same pattern as WT (Fig. S5D), which is consistent with the notion that miRNA cleavage products are temporally stable intermediates, resistant to cellular XRN4 exonuclease in *A. thaliana*, although the precise mechanism is currently unknown (*46*, *47*). Thus, we counted the degradation reads within target sites as the outcomes of cleavage products from miRISC cleavage.

Based on these two observations, we can derive the miRNA CE *in vivo*. Since AGO and miRNA form an enzyme complex, we defined the (CE) in a similar way to enzyme activity. In details, the catalytic ability of an enzyme can be defined as the amount of product generated by one unit of enzyme from one unit of substrate, which led us to define:

$$\boldsymbol{CE = \frac{Number\ of\ miRNA\ mediated\ cleavage\ products}{Total\ number\ of\ miRNA\ target\ RNAs \times\ Number\ of\ miRNAs}}.$$

Cleaved transcripts can be characterized by the 5' phosphate featured on 3' cleavage products. We constructed a degradome library to capture cleavage products. RNA abundance of degradation products can be measured by the Reads Per Kilobase of transcript per Million mapped reads (RPKM). RPKM is a measure of relative RNA concentration in the whole transcriptome. In the degradome, the RPKM of each mRNA means relative degraded mRNA fragment number compared to all degradation products, i.e.,

$$\boldsymbol{Degradome[RPKM]} \approx \frac{\boldsymbol{Number\ of\ cleavage\ products}}{\boldsymbol{Total\ degradation\ products}}.$$

To quantify the miRNA-mediated degradation products, we used TargetFinder (*48*) to predict the miRNA target sites on the expressed transcripts in our samples and removed any duplicated target sites from the same miRNA family. TargetFinder predicts target sites with high specificity in *A. thaliana* by assigning a sequence complementarity penalty score (SCPS) (*49*) (Fig. S5B). We then counted and designated the reads mapped within each target site as the products of miRNA-mediated cleavage. Therefore, we can label the miRNA-mediated Degradome RPKM as mirDegradome[RPKM], whereby:

$$\boldsymbol{mirDegradome[RPKM]} \approx \frac{\boldsymbol{Number\ of\ miRNA\ mediated\ cleavage\ product}}{\boldsymbol{Total\ degradation\ products}}.$$

Similar to the degradome, the RPKM of each target mRNA can be described as,

$$\boldsymbol{RNAseq[RPKM]} \approx \frac{\boldsymbol{Total\ number\ of\ miRNA\ target\ RNAs}}{\boldsymbol{Total\ number\ of\ all\ RNAs}}.$$

The miRNA RPKM estimate can be derived from the miRNA library. Similarly,

$$\boldsymbol{miRNAseq[RPKM]} \approx \frac{\boldsymbol{Number\ of\ miRNAs}}{\boldsymbol{Total\ miRNA\ products}}$$

so

$$\boldsymbol{CE} \approx \boldsymbol{\alpha} * \frac{\boldsymbol{mirDegradome[RPKM]}}{\boldsymbol{RNAseq[RPKM]} \times \boldsymbol{miRNAseq[RPKM]}}$$

where

$$\boldsymbol{\alpha} = \frac{\boldsymbol{Total\ degradation\ products}}{\boldsymbol{Total\ number\ of\ all\ RNAs} \times \boldsymbol{Total\ miRNA\ products}}$$

The total degradation products, total RNAs and total miRNAs should be constant and be reflected by the library sequencing depth. Therefore, α is a constant.

The population of each target mRNA is constant over time due to the dynamic equilibrium of an intact mRNA and its degraded products (*50*), i.e.,

$$\boldsymbol{Total\ number\ of\ mRNA = Number\ of\ intact\ mRNA + Number\ of\ degraded\ mRNA}$$

In our study, (-)SHAPE library can estimate the intact mRNA abundance. Indeed, we found that RNAseq[RPKM] was tightly correlated with (-)SHAPE[RPKM] +Degradome[RPKM] (Fig. S2E), i.e.,

$$\boldsymbol{RNAseq[RPKM]} \approx \beta(\boldsymbol{(-)SHAPE\ [RPKM] + Degradome[RPKM]}),$$

in which β is a constant. The advantage of combining the (-)SHAPE and the degradome libraries to calculate the CE lies in its focus on miRNA-mediated cleavage events. Then, for miRNA target genes,

$$\boldsymbol{RNAseq[RPKM]} \approx \beta(\boldsymbol{(-)SHAPE\ [RPKM] + mirDegradome[RPKM]}),$$

thus

$$\boldsymbol{CE \approx \frac{\alpha}{\beta} * \frac{mirDegradome[RPKM]}{((-)SHAPE\ [RPKM] + mirDegradome[RPKM]) \times miRNAseq[RPKM]}}$$

or

$$\boldsymbol{CE \propto \frac{mirDegradome[RPKM]}{((-)SHAPE\ [RPKM] + mirDegradome[RPKM]) \times miRNAseq[RPKM]}}.$$

In short, we can combine miRNA-seq, (-)SHAPE and degradome libraries to quantify miRNA CE *in vivo*. The CE pipeline is illustrated in Fig. S5A.

Multiple sequence alignment and Visualization of RNA structure

Multiple sequence alignment was conducted with Jalview (*51*) and visualized by WebLogo (*52*). The secondary structures and the corresponding SHAPE values were visualized using the VARNA Java applet (*53*).

Calculation of $\Delta G^{\ddagger}_{open}$ and $\Delta G^{\ddagger}_{cutting}$

$\Delta G^{\ddagger}_{open}$ measures the energy required to open the target sites during miRISC binding. $\Delta G^{\ddagger}_{open}$ was computed as the difference between the minimum free energy of the *in vivo* secondary structure and the minimum free energy of the ‘‘hard constrained’’ transcript, in which the target nucleotides were required to be unpaired (*6*, *54*). By exploring a range of flanking region lengths upstream and downstream of the target site, we chose the upstream and downstream flank lengths to be 50 nucleotides for the majority of analyses. We used *RNAfold* from the Vienna RNA package (*55*) together with our SHAPE reactivity data to calculate the energy terms in $\Delta G^{\ddagger}_{open}$, the RNA structures and the base pairing probabilities (BPP).

$\Delta G^{\ddagger}_{cutting}$ measures the energy required to raise the initial substrate target RNA to the transition catalysis-compatible state. $\Delta G^{\ddagger}_{cutting}$ is given by:

$$\Delta G^{\ddagger}_{cutting} = \Delta G^{\ddagger}_{open} - |\Delta G_{duplex}| + \Delta G_{catalysis},$$

where ΔG_{duplex} is the binding free energy of the miRNA-target duplex and $\Delta G_{catalysis}$ refers to the miRISC transition catalytic state energy. The ΔG_{duplex} energy was calculated for the miRNA sequence and the target region sequence using *RNAduplex* from the Vienna RNA package (*55*).

The crystal structures of *Thermus thermophilus* Argonaute (TtAgo) (*56*, *57*), human AGO2(*58*) and yeast *Kluyveromyces polysporus* Argonaute (KpAGO) (*59*) suggest that AGO proteins have a conserved catalytic mechanism. Furthermore, the transition cleavage model does not engage in any nucleotide interactions (*56*, *57*). Therefore, we assumed that the activation energy, $\Delta G_{catalysis}$, is a constant for the same type of AGO protein. Quantum mechanics simulations estimate the value to be approximately 15 kcal mol^{-1}(*60*). Therefore, $\Delta G^{\ddagger}_{cutting}$ is given by:

$$\Delta G^{\ddagger}_{cutting} = \Delta G^{\ddagger}_{open} - |\Delta G_{duplex}| + 15.$$

Plasmid Construction

For cleavage efficiency validation, the miRNA156 target sites, followed by 0 or 2 Adenines (As) and ending with a G-quadruplex (GQS) or a stem-loop (SL) were synthesized and inserted into *Afl*II and *Pac*I of Firefly 3'UTR in vector inter2. We labelled the GQS constructs as 0A_GQS and 2A_GQS, and the stem-loop constructs as 0A_SL and 2A_SL, with the prefix indicating the number of Adenines. Antisense of miRNA156 target site constructs with the same flanking sequence were also synthesized as the control for each construct.

To carry out the *AP2* mutagenesis assay, the native *AP2* sequence was inserted into the *pEAQ* vector by Invitrogen LR reaction. Mutations were introduced using Q5 site-directed mutagenesis kits (NEB, E0554S).

For the miRNA156 overexpression vector construction in AGO1 *in vivo* binding assay, the *MIR156B* genomic sequence was inserted into *Asc*I and *Sac*I of vector *pMDC32*. Primers are listed in Table S2.

Arabidopsis protoplast transformation

Protoplasts from the stable *MIR156* over-expression line and the stable *MIR172* over-expression line were prepared and transformed according to the Tape-Arabidopsis Sandwich method (*61*). 16 h after transformation, protoplasts were centrifuged at 100 g for 2 min. RNA was extracted with Qiagen RNeasy kit and qRT-PCR quantification was performed with Bio-Rad CFX. Primers are listed in Table S2.

In vivo structure validation experiments

Four-week-old tobacco leaves were co-infiltrated with agrobacterium strains harboring plasmids of 0A_SL, 2A_SL, 0A_GQS or 2A_GQS. Two days after infiltration, the leaves were treated with 144 mM of the SHAPE reagent (NAI). The control group was treated with DMSO. Total RNA was extracted using RNeasy Plant Mini Kit (Qiagen) and then by on-column DNaseI

treatment (following the respective manufacturer's protocol). Gene-specific reverse transcription was performed as previously described by Kwok et al. (*37*), with a few modifications. Of note, 2 µg of *in vivo* total RNA was resuspended in 10 µl RNase-free water. Primer extension was performed with 2 pmol of DNA gene-specific primers (5'CATGCTTAACGTAATTCAACAGAAATTATATG) by Invitrogen SuperScript III reverse transcriptase. The resulting cDNA pellet was dissolved in RNase-free water and mixed with 1 µl 50 µm Mly1-HBLPCR-5'ssDNA linker modified by a 5'-phosphate and a 3'-3-Carbon spacer group (5'P-AGATCGACTCAGCGTCGTGTAGCTGAGTCGATCTNNNNNN-C3-3'), 10 µl Quick Ligase Reaction Buffer (2X), 1U Quick Ligase (New England Biolab) in a 20 µl system. The ligation was performed at 25°C for 1 h, followed by Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v, sigma) and Chloroform:Isoamyl (24:1, v/v, sigma) purification.

The ligated cDNA samples were dissolved in 10 µl of water and used for the PCR reaction. The PCR reaction contained final concentrations of 0.5 mM VIC-labelled DNA gene-specific primers (the same as that used in the reverse transcription primers except the 5' end was labelled with Vic), 0.5 mM of linker reverse primer (AGATCGACTCAGCTACACGACGC), 200 mM dNTPs, 1X ThermoPol reaction buffer and 1.25U of NEB Taq DNA polymerase in 25 µl. The solution was then extracted with Phenol:Chloroform:Isoamyl (25:24:1, v/v, sigma) and incubated with Mly1 restriction enzyme, according to the manufacturer's protocol. Finally, the reaction pellets were dried and resuspended in Hi-Di formamide (Applied Biosystems/Life Technologies).

The Ned-labelled gene-specific primer (the same as that used in reverse transcription primer except the 5' end was labelled with NED) was used to make sequencing ladders using linear DNA and 1 µl 5 mM ddTTP by Klenow DNA Polymerase I (New England Biolab) (*62*). Then, the reaction pellets were dried, resuspended in Hi-Di formamide (Applied Biosystems/Life Technologies) and run on an Applied Biosystems 3730xl Genetic Analyzer. The resulting data were analyzed using QuSHAPE (*63*).

AGO1 *in vitro* cleavage assay

HA-tagged AGO1^{WT} was immuno-purified from *Arabidopsis* seedlings (*64*). The 0A_GQS and 2A_GQS designed RNAs were transcribed *in vitro* with T7 polymerase (NEB, 2040S) as substrates. To perform the slice assay, cleavage buffer (100mM ATP, 10mM GTP, 60mM MgCl2, 0.5M CPO4, 1mg/ml CPK) was added to 20µl beads in extraction buffer (1:1) bearing freshly purified HA-AGO1 from 3g seedling on the beads' surface. 50 cps of labelled substrate was added to the reaction and incubated at 25℃. 10µl of the resultant liquid was added to 10µl 2x RNA loading buffer (95% Formamide, 0.02% SDS, 1mM EDTA, 0.02%, Bromothymol Blue, 0.01% Xylene Cyanol), denatured for 5min at 95℃ and loaded into a 1mm PAGE gel (10% acrylamides:bis 19:1, 7M Urea, 1xTBS). Then the gel was dried and exposed to a phosphor screen for image analysis.

AGO1 *in vivo* binding assay

Four-week-old tobacco leaves were co-infiltrated with agrobacterium strains harbouring plasmids of *35S:MIR156B* and *35S:HA-AGO1*^{DAH} and 0A_GQS or 2A_GQS. Two days after infiltration, the leaves were collected and ground in liquid nitrogen. The protein/RNA complexes were extracted using two volumes of IP buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5% β-mercaptoethanol, 1 mM EDTA, 10% glycerol, 0.1% NP-40, 1 mM PMSF, and 1X complete protease inhibitor cocktail). After removing insoluble debris by centrifugation, cell extracts were incubated with anti-HA antibody (Abcam ab9110) for 1h at 4℃ with gentle mixing. The anti-HA-decorated extracts were then incubated with pre-washed protein G magnetic beads for 1h. After incubation, the beads were washed 6 times with the IP buffer. The RNA produced after co-immunoprecipitation was precipitated with ethanol and glycogen, and analysed by RT-PCR. The miRNA156 expression levels were analysed by miRNA RT-PCR (*65*).

References

38. F. F. Souret, J. P. Kastenmayer, P. J. Green, AtXRN4 degrades mRNA in Arabidopsis and its substrates include selected miRNA targets. *Mol. Cell*. **15**, 173–183 (2004).

39. M. A. German, M. Pillay, D. H. Jeong, A. Hetawal, S. Luo, P. Janardhanan, V. Kannan, L. A. Rymarquis, K. Nobuta, R. German, E. De Paoli, C. Lu, G. Schroth, B. C. Meyers, P. J. Green, Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat. Biotechnol.* **26**, 941–946 (2008).

40. R. C. Spitale, P. Crisalli, R. a Flynn, E. A. Torre, E. T. Kool, H. Y. Chang, RNA SHAPE analysis in living cells. *Nat. Chem. Biol.* **9**, 18–20 (2013).

41. C. Addo-Quaye, T. W. Eshoo, D. P. Bartel, M. J. Axtell, Endogenous siRNA and miRNA Targets Identified by Sequencing of the Arabidopsis Degradome. *Curr. Biol.* **18**, 758–762 (2008).

42. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, 25 (2009).

43. K. E. Deigan, T. W. Li, D. H. Mathews, K. M. Weeks, Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci.* **106**, 97–102 (2009).

44. S. Djuranovic, A. Nahvi, R. Green, miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science*. **336**, 237–40 (2012).

45. X. Chen, Small RNAs and Their Roles in Plant Development. *Annu. Rev. Cell Dev. Biol.* **25**, 21–44 (2009).

46. P. Jackowiak, M. Nowacka, P. M. Strozycki, M. Figlerowicz, RNA degradome-its biogenesis and functions. *Nucleic Acids Res.* **39**, 7361–7370 (2011).

47. C. Llave, Z. Xie, K. D. Kasschau, J. C. Carrington, Cleavage of Scarecrow-like mRNA Targets Directed by a Class of Arabidopsis miRNA. *Science (80-.).* **297**, 2053–2056 (2002).

48. N. Fahlgren, M. D. Howell, K. D. Kasschau, E. J. Chapman, C. M. Sullivan, J. S. Cumbie, S. A. Givan, T. F. Law, S. R. Grant, J. L. Dangl, J. C. Carrington, High-throughput sequencing of Arabidopsis microRNAs: Evidence for frequent birth and death of MIRNA genes. *PLoS One*. **2** (2007), doi:10.1371/journal.pone.0000219.

49. P. K. Srivastava, T. R. Moturu, P. Pandey, I. T. Baldwin, S. P. Pandey, A comparison of performance of plant miRNA target prediction tools and the characterization of features for genome-wide target prediction. *BMC Genomics*. **15**, 348 (2014).

50. C. Miller, B. Schwalb, K. Maier, D. Schulz, S. Dümcke, B. Zacher, A. Mayer, J. Sydow, L. Marcinowski, L. Dölken, D. E. Martin, A. Tresch, P. Cramer, Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Mol. Syst. Biol.* **7**, 458 (2011).

51. A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics*. **25**, 1189–1191

(2009).

52. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

53. K. Darty, A. Denise, Y. Ponty, VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*. **25**, 1974–1975 (2009).

54. R. Dimitrov, microRNA Gene Finding and Target Prediction - Basic Principles and Challenges. *MOJ Proteomics Bioinforma.* **1**, 105–110 (2014).

55. H. Tafer, C. Höner zu Siederdissen, P. F. Stadler, S. H. Bernhart, I. L. Hofacker, R. Lorenz, C. Flamm, ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).

56. Y. Wang, S. Juranek, H. Li, G. Sheng, T. Tuschl, D. J. Patel, Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. *Nature*. **456** (2008), pp. 921–926.

57. G. Sheng, H. Zhao, J. Wang, Y. Rao, W. Tian, D. C. Swarts, J. van der Oost, D. J. Patel, Y. Wang, Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. *Proc. Natl. Acad. Sci.* **111**, 652–657 (2014).

58. N. T. Schirle, I. J. MacRae, J. Liu, G. Meister, I. Behm-Ansmant, J. E. Braun, E. Huntzinger, M. Fauser, E. Izaurralde, M. R. Fabian, J. J. Song, S. K. Smith, G. J. Hannon, L. Joshua-Tor, Y. Wang, Y. Wang, Y. Wang, G. Sheng, S. Juranek, T. Tuschl, D. J. Patel, Y. R. Yuan, J. B. Ma, J. B. Ma, K. Ye, D. J. Patel, A. Lingel, B. Simon, E. Izaurralde, M. Sattler, J. J. Song, K. S. Yan, A. Boland, F. Tritschler, S. Heimstädt, E. Izaurralde, O. Weichenrieder, A. Boland, E. Huntzinger, S. Schmidt, E. Izaurralde, O. Weichenrieder, F. Frank, N. Sonenberg, B. Nagar, Y. L. Chiu, T. M. Rana, G. F. Deleavey, D. P. Bartel, B. P. Lewis, C. B. Burge, D. P. Bartel, D. P. Bartel, S. L. Lian, A. Eulalio, S. Helms, C. Fritzsch, M. Fauser, E. Izaurralde, M. El-Shami, N. Bies-Etheve, D. Baillat, R. Shiekhattar, S. Till, A. Boland, E. Huntzinger, S. Schmidt, E. Izaurralde, O. Weichenrieder, A. Eulalio, E. Huntzinger, E. Izaurralde, N. De, I. J. Macrae, W. Kabsch, P. D. Adams, T. C. Terwilliger, P. Emsley, K. Cowtan, G. N. Murshudov, The crystal structure of human Argonaute2. *Science*. **336**, 1037–40 (2012).

59. K. Nakanishi, D. E. Weinberg, D. P. Bartel, D. J. Patel, Structure of yeast Argonaute with guide RNA. *Nature*. **486**, 368–374 (2012).

60. E. Rosta, M. Nowotny, W. Yang, G. Hummer, Catalytic mechanism of RNA backbone cleavage by ribonuclease H from quantum mechanics/molecular mechanics simulations. *J. Am. Chem. Soc.* **133**, 8934–8941 (2011).

61. F. H. Wu, S. C. Shen, L. Y. Lee, S. H. Lee, M. T. Chan, C. S. Lin, Tape-arabidopsis sandwich - A simpler arabidopsis protoplast isolation method. *Plant Methods*. **5** (2009).

62. R. A. McGraw, Dideoxy DNA sequencing with end-labeled oligonucleotide primers. *Anal. Biochem.* **143**, 298–303 (1984).

63. F. Karabiber, J. L. McGinnis, O. V Favorov, K. M. Weeks, QuShape: rapid, accurate, and best-

practices quantification of nucleic acid probing information, resolved by capillary electrophoresis. *RNA*. **19**, 63–73 (2013).

64. Qi, Y. & Mi, S. Purification of Arabidopsis Argonaute Complexes and Associated Small RNAs. in Methods in Molecular Biology 243–254 (2010).

65. E. Varkonyi-Gasic, R. Wu, M. Wood, E. F. Walton, R. P. Hellens, Protocol: A highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods*. **3** (2007).

66. K. M. Weeks, D. M. Crothers, RNA recognition by Tat-derived peptides: Interaction in the major groove? *Cell*. **66**, 577–588 (1991).

67. J. Zuber, B. Joseph Cabral, I. McFadyen, D. M. Mauger, D. H. Mathews, Analysis of RNA nearest neighbor parameters reveals interdependencies and quantifies the uncertainty in RNA secondary structure prediction. *Rna*. **24**, 1568–1582 (2018).