



OPEN

## Determining the defining lengths between mature microRNAs/small interfering RNAs and tinyRNAs

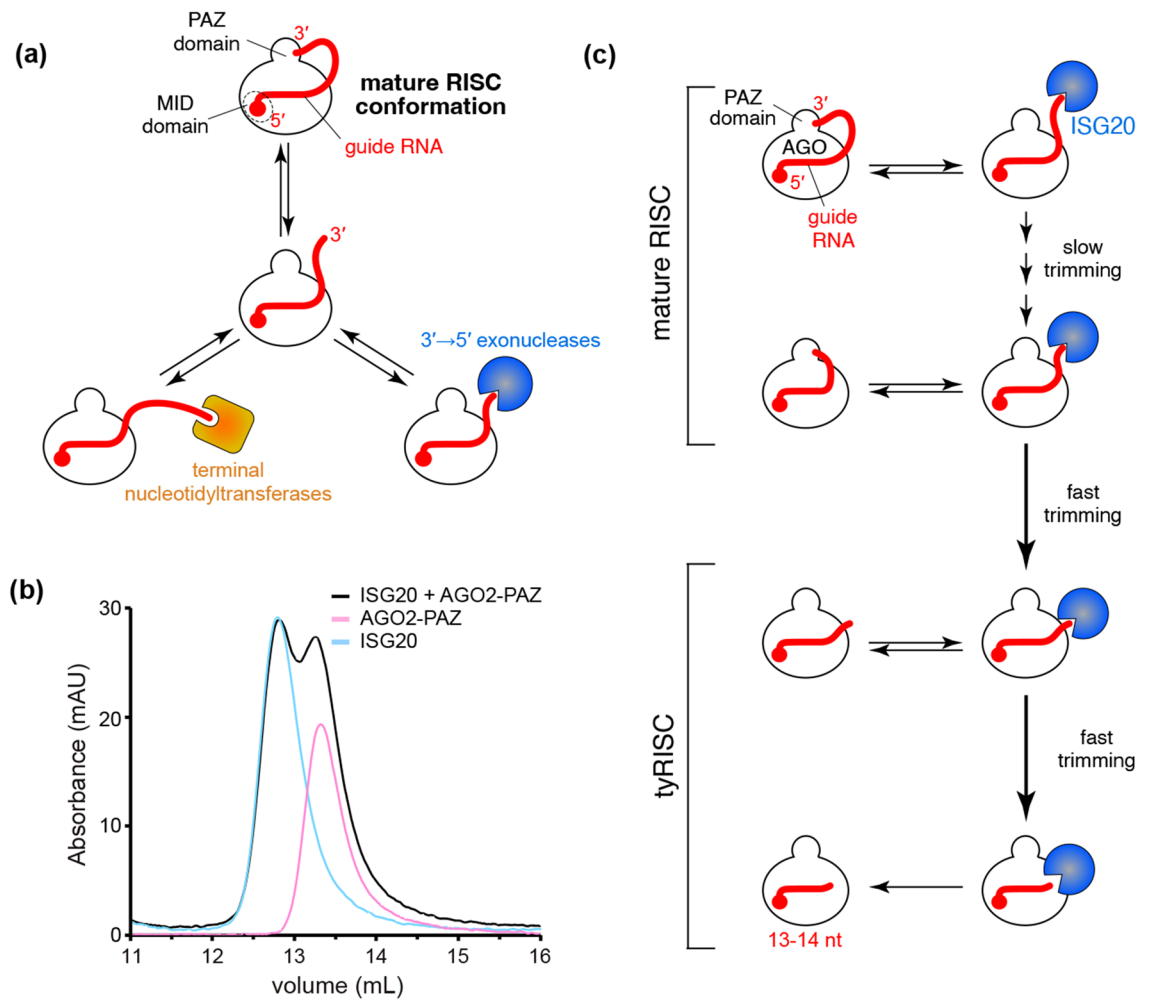
GeunYoung Sim<sup>1,2</sup>, Audrey C. Kehling<sup>3</sup>, Mi Seul Park<sup>3</sup>, Cameron Divoky<sup>4</sup>, Huaqun Zhang<sup>3</sup>, Nipun Malhotra<sup>3</sup>, Jackson Secor<sup>3</sup> & Kotaro Nakanishi<sup>1,2,3,4</sup>✉

MicroRNAs (miRNAs) and small interfering RNAs (siRNAs) are loaded into Argonaute (AGO) proteins, forming RNA-induced silencing complexes (RISCs). The assembly process establishes the seed, central, 3' supplementary, and tail regions across the loaded guide, enabling the RISC to recognize target RNAs for silencing. This guide segmentation is caused by anchoring the 3' end at the AGO PAZ domain, but the minimum guide length required for the conformation remains to be studied because the current miRNA size defined by Dicer processing is ambiguous. Using a 3' → 5' exonuclease ISG20, we determined the lengths of AGO-associated miR-20a and let-7a with 3' ends that no longer reach the PAZ domain. Unexpectedly, miR-20a and let-7a needed different lengths, 19 and 20 nt, respectively, to maintain their RISC conformation. This difference can be explained by the low affinity of the PAZ domain for the adenosine at g19 of let-7a, suggesting that the tail-region sequence slightly alters the minimum guide length. We also present that 17-nt guides are sufficiently short enough to function as tinyRNAs (tyRNAs) whose 3' ends are not anchored at the PAZ domain. Since tyRNAs do not have the prerequisite anchoring for the standardized guide segmentation, they would recognize targets differently from miRNAs and siRNAs.

More than 2000 microRNAs (miRNAs) have been reported in humans as of 2019<sup>1</sup>. MiRNAs were defined to be about 22 nt because precursor miRNAs (pre-miRNAs) are processed by Dicer, a molecular ruler which generates size-specific miRNA duplexes<sup>2,3</sup>. The ~22-nt duplexes are sufficiently long enough to be loaded into Argonaute proteins (AGOs) efficiently. While the passenger strand is ejected, the remaining guide strand and AGO form the mature RNA-induced silencing complex (RISC)<sup>4</sup>. Since both 5' and 3' ends of the guide RNA are recognized by the MID and PAZ domains, respectively (Fig. 1a)<sup>5</sup>, the guide acquires the standardized guide segmentation composed of the seed (guide nucleotide positions 2–8, g2–g8), central (g9–g12), 3'-supplementary (g13–g16), and tail (g17–3' end) regions<sup>6</sup>. The mature RISCs anchor the 2' hydroxyl and phosphate groups of the seed region thoroughly and pre-organize the segment in the A-form helical structure to enhance the affinity for target RNAs<sup>7</sup>. Guides usually do not involve their central region in target recognition, but their 3' supplementary region will enhance the affinity for targets when the seed region is mismatched with the target<sup>6</sup>. When RISCs bind to targets through the seed and 3' supplementary regions, the PAZ domain of the AGO still captures the 3' end of the loaded guide RNA<sup>8</sup>. However, once the guide-target base pairing proceeds into the tail region, the accumulated topological stress forces the 3' end of the guide to be released from the PAZ domain<sup>9,10</sup>. As a result, the release of the 3' end enables the RISC to pair the central region with the target and cleave it if both strands are fully complementary<sup>6,11</sup>. To employ this typical target recognition, the loaded guide must be long enough to maintain the mature RISC conformation. Although the 5' end of the guide RNA is not accessible to any nucleases due to its thorough recognition between the MID and PIWI domains<sup>5</sup>, the 3' end of the guide is sometimes released to the solvent and undergoes trimming by 3' → 5' exonucleases or extension by terminal nucleotidyltransferases (Fig. 1a), which generates miRNA isoforms<sup>12,13</sup>. Changing the guide length is known to manipulate the corresponding miRNAs' abundance<sup>14</sup> and RISC's half-life, target specificity, and activity<sup>15</sup>. Despite its significance, the minimum guide length required for maintaining the mature RISC conformation remains to be studied due to the lack of requisite technology and thus has been a long-standing question.

Meanwhile, AGO-associated guide RNAs shorter than regular miRNAs and small interfering RNAs (siRNAs) have been reported and referred to as tinyRNAs (tyRNAs)<sup>16</sup>. In humans, 17-nt unusually small RNAs were derived

<sup>1</sup>Center for RNA Biology, The Ohio State University, Columbus, OH 43210, USA. <sup>2</sup>Molecular, Cellular and Developmental Biology, The Ohio State University, Columbus, OH 43210, USA. <sup>3</sup>Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH 43210, USA. <sup>4</sup>Ohio State Biochemistry Program, The Ohio State University, Columbus, OH 43210, USA. ✉email: nakanishi.9@osu.edu



**Figure 1.** ISG20 is an ideal tool to determine the minimum length of miRNAs. **(a)** Schematic of the competitive interactions of the 3' end of guide RNA between the AGO PAZ domain and the processing enzymes. In the mature RISC conformation, both 5' and 3' ends of the guide RNA are recognized by the MID and PAZ domains. **(b)** Size-exclusion chromatography testing the interaction of ISG20 with the isolated AGO2 PAZ domain. **(c)** Definition of miRNAs and tyRNAs based on the guide length. Mature RISCs become tyRNA-associated RISCs (tyRISCs) upon the conversion of the miRNA to a tyRNA.

from Kaposi sarcoma-associated herpesvirus K12-1 miRNAs<sup>17</sup>, while 15–19-nt tRNA fragments bound to human Argonaute2 (AGO2) to regulate gene expression<sup>18</sup>. In plants, 10–17-nt tyRNAs were discovered inside extracellular vesicles<sup>19</sup>. These previous studies demonstrated that AGO-associated tyRNAs are biologically relevant, but little was known about how these tyRNAs were synthesized<sup>16</sup>. In this context, our recent study discovered a tyRNA-biogenesis pathway in which specific 3'→5' exonucleases, including interferon-stimulated gene 20 kDa (ISG20), trim 21–23-nt miRNAs down to 13–14-nt tyRNAs<sup>20</sup>. Neither 23-nt miR-20a nor 21-nt let-7a, but rather their 14-nt variants can catalytically activate human Argonaute3 (AGO3)<sup>21</sup>, indicating that tyRNAs play different roles from miRNAs and siRNAs. However, how the size of tyRNAs differs from that of miRNAs and siRNAs has not yet been defined. In this study, we used ISG20 as a tool to determine the defining lengths of tyRNAs and miRNAs/siRNAs.

## Results

### ISG20 competes with the PAZ domain for the 3' end of guide

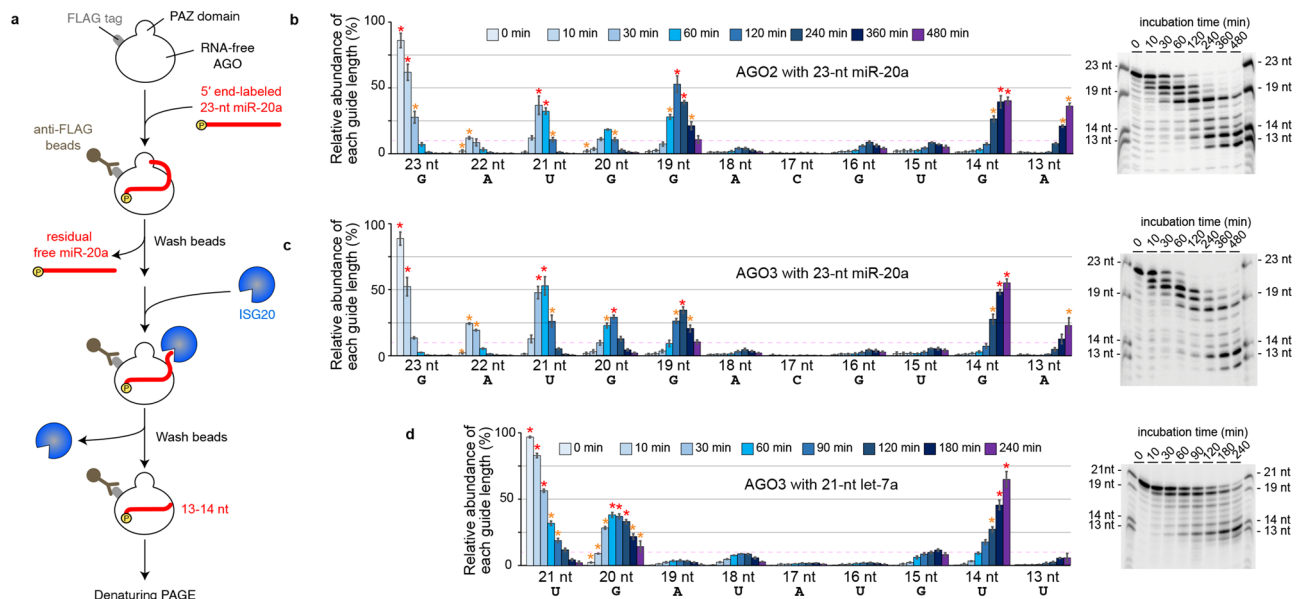
Our recent study identified several 3'→5' exonucleases, including ISG20, capable of trimming AGO-associated miRNAs down to 13–14 nt *in vitro* and *in vivo*<sup>20</sup>. In the current study, we used ISG20 as a tool to determine the minimum guide length required to be captured by the PAZ domain. A previous study reported that small RNA degrading nuclease 1 (SDN1) shortens miRNAs bound to plant AGO<sup>22</sup>. While SDN1 needs to interact with the PAZ domain to initiate the 3'→5' guide trimming<sup>22</sup>, ISG20 showed no interaction with the isolated PAZ domain of human AGO2 (Fig. 1b and Supplementary Fig. S1). This result demonstrates that ISG20 binds to the free 3' end only when the PAZ domain transiently releases the 3' end or cannot capture the 3' end of short guide RNAs (Fig. 1c). Therefore, we hypothesized that the 3' end of AGO-associated guides can go back and forth between the PAZ domain and ISG20 (Fig. 1c). The PAZ domain could intermittently protect the 3' end of the guide from

exposure to ISG20 and slow down the trimming rate when the miRNA is long enough to be captured by the PAZ domain. On the other hand, after ISG20 trims the miRNA down to the length of a tyRNA, the 3' end no longer reaches the PAZ domain and is continuously shortened by ISG20, which accelerates the trimming rate (Fig. 1c).

### miR-20a requires 19 nt to maintain the mature RISC

To validate the potential of ISG20 as a suitable tool, we performed a time-course *in vitro* trimming assay (Fig. 2a and Table 1). After incubation with a 5'-end radiolabeled 23-nt miR-20a, FLAG-tagged AGO2 (FLAG-AGO2) was immobilized on anti-FLAG beads. Any residual free guide was washed out, and ISG20 was added to trim the AGO2-associated guides. The reaction was sampled at different time points, and the trimmed guides were resolved on a denaturing gel. As expected, the trimmed guides were classified into three groups based on their lengths (Fig. 2b). The first group of 19–23-nt guides accumulated well, suggesting that their 3' end was intermittently captured by the PAZ domain and thus trimmed slowly. The second group of 15–18-nt guides was barely detected, with the 17-nt guides being the least abundant, indicating that their 3' end was always exposed to ISG20 because the lengths are too short to reach the PAZ domain. The third group of 13–14-nt guides accumulated at the end point of the time course because the nucleic acid-binding channel of RISC sequesters 13–14-nt guides from exposure to ISG20<sup>20</sup>. These results validated our hypothesis and showed that the PAZ domain can capture the 3' end of the 19-nt miR-20a, but not 18 nt or shorter, if any.

When compared to the other three AGO paralogs, the AGO2 PAZ domain was reported to have a lower affinity for the 3' end of the guide because the 3'-end binding pocket possesses Arg315 and His316, instead of lysine and tyrosine, respectively, in the other AGOs<sup>23</sup>. To see whether the trimming pattern varies in paralogous RISCs,



**Figure 2.** Determining the defining length between miRNAs and tyRNAs. **(a)** Schematic of the *in vitro* trimming assay of FLAG-AGO programmed with a 5'-end radiolabeled 23-nt miR-20a. **(b, c)** Time-course *in vitro* trimming assay of a 5'-end radiolabeled 23-nt miR-20a. *Left:* The relative abundance of 13–23-nt miR-20a bound to FLAG-AGO2 **(b)** and -AGO3 **(c)** over 8 h was quantified from three independent experiments and shown with mean  $\pm$  SD. The guide lengths with the highest and second highest abundances at each time point are highlighted with red and orange asterisks, respectively. 10% relative abundance is depicted as dotted pink lines. *Right:* Representative gel images of the guide trimming on FLAG-AGO2 **(b)** and -AGO3 **(c)**. **(d)** Time-course *in vitro* trimming assay of 5'-end radiolabeled 21-nt let-7a bound to FLAG-AGO3. *Left:* The relative abundance of 13–21-nt let-7a. *Right:* A representative gel image of the guide trimming.

Guide RNAs	Sequence				
	1	5	10	15	20
23-nt miR-20a	5'	-pUAAAGUGCCUUAUAGUGCAGUAG-3'			
21-nt let-7a	5'	-pUGAGGUAGUAGGUUGUAUAGU-3'			
23-nt chimeric poly(U)	5'	-pUAAAGUGCCUUAUUUUUUUUUUUU-3'			
23-nt chimeric poly(A)	5'	-pUAAAGUGCCUUAUUUUUUUUUUUU-3'			
23-nt chimeric poly(C)	5'	-pUAAAGUGCCUUAUUUUUUUUUUUU-3'			

**Table 1.** Guide RNAs used in this study.

the same assay was repeated with FLAG-tagged AGO3 (FLAG-AGO3). While 19–23-nt guides were trimmed slowly, 15–18 nt lengths were barely detected (Fig. 2c). The trimming stopped when the guides became 13–14 nt. The observed trimming profile of FLAG-AGO3-associated 23-nt miR-20a was similar to that of FLAG-AGO2 (Fig. 2b,c). These results demonstrate that ISG20 is a good tool for determining the minimum guide length required to retain the mature RISC conformation and that miR-20a functions as a miRNA as long as the size retains 19 nt.

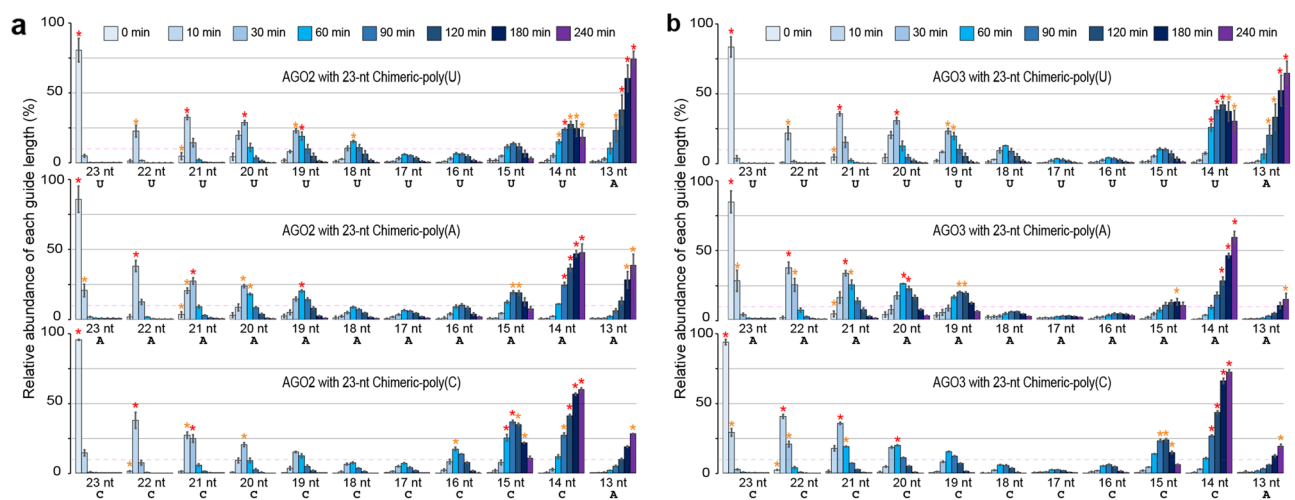
### let-7a needs 20 nt to stably maintain the mature RISC

The PAZ domain is known to have a preference for UMP, GMP, CMP, and AMP, in this order<sup>24</sup>. Accordingly, we thought that when the 3' end of the guide has adenine (A), it must be readily released from the PAZ domain and thus trimmed by ISG20 quickly. Such guides would be less prevalent at lengths where A is at their 3' end, compared with guides of similar lengths that have uridine (U), cytidine (C), or guanine (G) at their 3' end. To examine this idea, the *in vitro* guide trimming assay was performed with FLAG-AGO3-associated 21-nt let-7a (Fig. 2d and Table 1). miR-20a and let-7a have G and A, respectively, at g19, albeit both have G at g20 and U at g21 (Fig. 2c,d). Since the PAZ domain has a higher affinity for GMP than AMP<sup>24</sup>, we surmised that let-7a must not accumulate a length of 19 nt stably, but rather 20 nt. As expected, the 19-nt guide of let-7a remained less than 10% abundance at any time point, unlike miR-20a (Fig. 2d). In contrast, the 18-nt guide of let-7a accumulated slightly more than that of miR-20a (Fig. 2d), which could be explained by the higher affinity of the PAZ domain for UMP over AMP<sup>24</sup>. However, 20-nt let-7a was always accumulated more than 18-nt at each time point (Fig. 2d), indicating that 20 nt is the minimum length for let-7a to stably maintain the mature RISC conformation. The trimming profile of let-7a also showed the least accumulation at 17 nt (Fig. 2d), which is consistent with the data for miR-20a (Fig. 2b,c). These results suggest that let-7a needs at least 20 nt to employ the standard guide segmentation for target recognition.

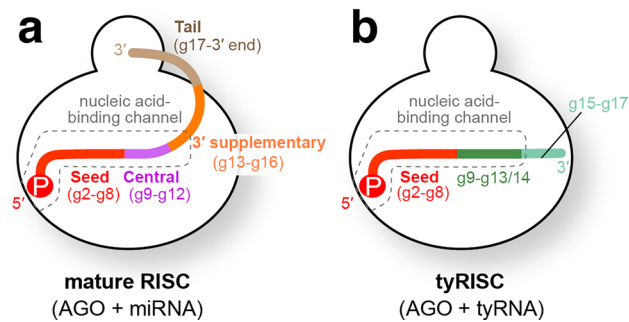
### Unbiased guide trimming

The affinity of the PAZ domain for the 3'-end nucleotide impacts the guide trimming rate of ISG20. In addition, the preference of ISG20 for the 3'-end nucleotide of guide RNA would affect the trimming rate as well. We investigated the minimum guide length required for mature RISC in a condition independent of these two biases. To this end, we made three 23-nt chimeric guides whose g1–g13 is derived from that of miR-20a while the g14–g23 is poly(U), poly(A), or poly(C) and named them Chimeric-poly(U), -poly(A), and -poly(C), respectively (Table 1). Chimeric-poly(G) was highly structured, as reported previously<sup>25</sup>, and could not be used for this assay. Since the g14–g23 of a chimeric guide is composed of identical nucleotides, the 14–23-nt derivatives expose the same 3'-end nucleotide to ISG20 during trimming. Therefore, ISG20 must trim those derivatives at the same rate unless the PAZ domain sequesters their 3' end. Likewise, the PAZ domain has the same affinity for the 3'-end nucleotides across the 14–23-nt chimeric guide derivatives, so long as it can interact with the 3' end of the guide. As a result, any observed difference in the accumulation between the 14–23-nt derivatives of a chimeric guide is expected to be attributed to the accessibility of their 3' end to the PAZ domain.

We programmed FLAG-AGO2 with either of three chimeric guides and tested them for trimming by ISG20. Although the trimming rates were different among the three chimeric guides, which necessitated that a different amount of ISG20 had to be used for each chimeric guide, all of them showed similar trimming patterns in that their 19–23 nt guides accumulated well (Fig. 3a and Supplementary Fig. S2a–c). The 18-nt Chimeric-poly(U) accumulated more than that of Chimeric-poly(A) and -poly(C), as seen in let-7a whose g18 has U. This result further supports that 18-nt guides could stabilize the mature RISC conformation when the g18 is U. Notably,



**Figure 3.** 17-nt or shorter guides behave as tyRNAs. **(a, b)** Time-course *in vitro* trimming assay of 5'-end radiolabeled 23-nt chimeric guides on FLAG-AGO2 **(a)** and -AGO3 **(b)**. 500, 200, and 80 pmol of ISG20 were used for Chimeric-poly(U) (top), -poly(A) (middle), and -poly(C) (bottom), respectively.



**Figure 4.** Schematics of mature RISC and tyRISC. **(a)** Mature RISC is composed of an AGO and a miRNA. The standardized guide segmentation, the seed (red), central (magenta), 3' supplementary (orange), and tail (wheat) regions, is shown. The nucleic acid-binding channel is depicted as dotted boxes. **(b)** tyRISC is composed of an AGO and a tyRNA. The seed (red) and the g9-g13/14 (dark green) are protected by the nucleic acid-binding channel from any nucleases. The solvent-exposed g15-g17 (faint green) is easily trimmed by 3'→5' exonucleases.

all chimeric guides were the least abundant at lengths of 17 nt, which is consistent with the data for miR-20a (Fig. 2b,c) and let-7a (Fig. 2d). The same trend in guide trimming was observed when FLAG-AGO3 was programmed with these chimeric guides (Fig. 3b and Supplementary Fig. S2d-f). Altogether, our data suggest that the mature RISC conformation remains stable so long as the loaded guide maintains 19 nt or longer.

## Discussion

Our unbiased trimming assays indicate that 19-nt or longer guide RNAs maintain the mature RISCs stably (Fig. 4a). This minimum guide length slightly differs depending on the sequence of the tail region due to the preference of the PAZ domain for different 3'-end nucleotides (Fig. 2c,d). In the current study, we defined that miRNAs and siRNAs are long enough to anchor their 3' end at the PAZ domain. Now we also define guide RNAs whose length is too short to be captured by the PAZ domain as tyRNAs. All of our trimming results showed that 17-nt guides accumulated the least, regardless of guide sequence, demonstrating that 17 nt is sufficiently short enough to keep the 3' end from being recognized by the PAZ domain. Thus, 17-nt or shorter guide RNAs are defined as tyRNAs, and their ribonucleoprotein complex is termed “tyRNA-associated RISC (tyRISC)” (Fig. 4b).

Since the nucleic acid-binding channel sequesters the g1-g13/14 from the solvent<sup>20</sup>, the seed of tyRNA is never accessible to any 3'→5' exonucleases. tyRISCs would use the intact seed region similarly to mature RISCs when recognizing target RNAs. Unlike miRNAs, tyRNAs do not anchor their 3' end at the PAZ domain of AGO, allowing the g9-g17 to move vigorously. The g9-g17 of tyRNAs is unlikely to follow the more restrictive standardized guide segmentation of miRNAs. Therefore, after recognizing target RNAs through their seed regions, the tyRISC and mature RISC would treat the target differently. It is known that base-pairing with a fully complementary target occurs sequentially in the order of the seed, 3' supplementary, tail, and central regions, after which the mature AGO2 RISC cleaves the scissile phosphate group between the t10 and t11<sup>6</sup>. Although AGO3 loaded with specific 14-nt tyRNAs similarly cleaves complementary RNAs between the t10 and t11<sup>21,26</sup>, the actual molecular mechanism must differ from that of mature AGO2 RISCs. Further research is required to understand the mechanisms by which tyRISCs recognize and cleave target RNAs. Future studies should consider this defining size between tyRNAs and miRNAs/siRNAs.

## Methods

### Purification of recombinant proteins

The genes of FLAG-AGO2 and -AGO3 were cloned into pFastBac-HTB (Invitrogen). Their recombinant proteins were overproduced and purified from insect cells as previously reported and stored in Buffer E (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 0.5 mM TCEP)<sup>26,27</sup>. ISG20 and AGO2 PAZ were cloned into a SUMO-fused pRSFDuet-1 vector (Novagen) and expressed in BL21 Rosetta 2 (DE3) pLysS *Escherichia coli* (*E. coli*) cells (Novagen). All plasmids were verified by Sanger sequencing (OSU Shared Resources). A recombinant protein of ISG20 was purified as previously reported<sup>20</sup>.

*E. coli* cells expressing the isolated AGO2 PAZ domain were lysed with an Emulsiflex C3 homogenizer (Avestin) in Lysis buffer [Buffer A (1× PBS, 500 mM NaCl, 40 mM imidazole, 5% glycerol, 10 mM β-mercaptoethanol) with 1 mM PMSF], followed by centrifugation for 1 h. The supernatant was loaded onto a 5 mL HisTrap HP column (Cytiva) and washed with Buffer A, followed by elution over a linear gradient to 50% Buffer B (1× PBS, 500 mM NaCl, 1.5 M imidazole, 5% glycerol, 10 mM β-mercaptoethanol). The sample was dialyzed overnight against Dialysis Buffer 1 (1× PBS, 500 mM NaCl, 10 mM β-mercaptoethanol) with Ulp1 protease to cleave the His-SUMO-tag. To remove the tag, the sample was applied to a 5 mL HisTrap HP column (Cytiva) and the flow-through collected for dialysis against Dialysis Buffer 2 (1× PBS, 10 mM β-mercaptoethanol). The sample was loaded to a 5 mL HiTrap Heparin HP Column (Cytiva) and washed with Buffer C (1× PBS, 5% glycerol, 10 mM β-mercaptoethanol) to remove the RNA-bound population. The RNA-free PAZ domain was eluted over a linear gradient to 100% Buffer D (1× PBS, 1 M NaCl, 5% glycerol, 10 mM β-mercaptoethanol) and concentrated by ultrafiltration. The sample was loaded onto a HiLoad 16/600 Superdex 75 pg column (Cytiva) equilibrated in

Buffer E (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM DTT) and peak fractions were concentrated by ultrafiltration. The final concentration was determined by BioSpectrometer (Eppendorf), and samples were flash-frozen in liquid nitrogen to store at  $-80^{\circ}\text{C}$ .

### Binding assay with AGO2 PAZ domain and ISG20

A Superdex 75 increase 10/300 GL column (GE Healthcare) was equilibrated with running buffer (150 mM NaCl, 20 mM Tris-HCl pH 7.5, 2 mM DTT). 100  $\mu\text{mol}$  of the purified AGO2 PAZ domain, ISG20, or a mixture of both was injected to the column. The mixture was incubated on ice for 60 min prior to injection.

### Time-course in vitro trimming assay

For the RISC assembly, 20  $\mu\text{M}$  recombinant FLAG-AGO2 or -AGO3 was incubated for 1 h at  $37^{\circ}\text{C}$  with 2  $\mu\text{M}$  5'-end radiolabeled guide RNA (23-nt miR-20a, 21-nt let-7a, 23-nt Chimeric-poly(U), 23-nt Chimeric-poly(A), or 23-nt Chimeric-poly(C)) in 1 $\times$  Trimming Buffer (25 mM HEPES-KOH pH 7.5, 50 mM KCl, 5 mM DTT, 0.2 mM EDTA, and 0.05 mg/mL BSA). 23-nt Chimeric-poly(G), which can form a stable secondary structure<sup>25</sup>, was excluded from the assay. The assembled RISC reaction was pulled down by anti-FLAG M2 affinity gel (Sigma Aldrich), which was pre-washed with 1 $\times$  PBS twice and 1 $\times$  Trimming buffer twice, at room temperature for 2 h with gentle rotation. 50  $\mu\text{L}$  of the 50% slurry of agarose beads were used per 10  $\mu\text{L}$  of reaction volume. Residual free guide RNA was removed by centrifugation, and the coupled beads were washed with immunoprecipitation (IP) wash buffer (300 mM NaCl, 50 mM Tris-HCl pH 7.5, and 0.05% NP-40) 8 times. After removing IP wash buffer by centrifugation, the beads were washed twice with 1 $\times$  Trimming buffer including 5 mM  $\text{MnCl}_2$ . The washed beads were resuspended in 1 $\times$  Trimming buffer including 5 mM  $\text{MnCl}_2$  (25  $\mu\text{L}$  of buffer per 25  $\mu\text{L}$  of packed beads). For 0 min sample, 50  $\mu\text{L}$  of slurry was transferred to a new tube and washed with IP wash buffer 4 times, mixed with 2 $\times$  urea quenching dye (8 M urea, 1 mM EDTA, 0.05% (w/v) xylene cyanol, 0.05% (w/v) bromophenol blue, 10% (v/v) phenol), and incubated for 1 min at  $90^{\circ}\text{C}$ . ISG20 was added to the rest of the beads and incubated at  $37^{\circ}\text{C}$  for the guide RNA trimming. 500, 200, and 80 pmol of ISG20 per 50  $\mu\text{L}$  of slurry was added to 23-nt Chimeric-poly(U), -poly(A), and -poly(C), respectively. Because ISG20 trims them at different rates, we used different amounts of ISG20 for each chimeric guide to resolve the trimmed lengths. For each time point, 50  $\mu\text{L}$  of slurry was transferred to a new tube, washed with IP wash buffer 4 times, mixed with 2 $\times$  urea quenching dye, incubated for 1 min at  $90^{\circ}\text{C}$ , and resolved on an 8 M urea 16% (w/v) polyacrylamide gel. Images were analyzed by the Typhoon Imaging System (GE Healthcare) and quantified by Image Lab (Bio-Rad). All data are presented as the mean  $\pm$  SD obtained from three replicates.

### Data availability

The sequence information of the plasmids used in the study are available at National Center for Biotechnology Information (NCBI) under the accession ID OR606569, OR606570, OR606571, and OR606572. The data underlying this article will be shared on reasonable request to the corresponding author.

Received: 15 September 2023; Accepted: 2 November 2023

Published online: 13 November 2023

### References

- Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: From microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162. <https://doi.org/10.1093/nar/gky1141> (2019).
- Ambros, V. *et al.* A uniform system for microRNA annotation. *RNA* **9**, 277–279. <https://doi.org/10.1261/rna.2183803> (2003).
- Zhang, H., Kolb, F. A., Jaskiewicz, L., Westhof, E. & Filipowicz, W. Single processing center models for human Dicer and bacterial RNase III. *Cell* **118**, 57–68. <https://doi.org/10.1016/j.cell.2004.06.017> (2004).
- Shang, R., Lee, S., Senavirathne, G. & Lai, E. C. microRNAs in action: Biogenesis, function and regulation. *Nat. Rev. Genet.* <https://doi.org/10.1038/s41576-023-00611-y> (2023).
- Nakanishi, K. Anatomy of RISC: How do small RNAs and chaperones activate Argonaute proteins?. *Wiley Interdiscip. Rev. RNA* **7**, 637–660. <https://doi.org/10.1002/wrna.1356> (2016).
- Bartel, D. P. Metazoan MicroRNAs. *Cell* **173**, 20–51. <https://doi.org/10.1016/j.cell.2018.03.006> (2018).
- Parker, J. S., Parizotto, E. A., Wang, M., Roe, S. M. & Barford, D. Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. *Mol. Cell* **33**, 204–214. <https://doi.org/10.1016/j.molcel.2008.12.012> (2009).
- 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. <https://doi.org/10.15252/embj.2018101153> (2019).
- Sheu-Gruttadauria, J. *et al.* Structural basis for target-directed MicroRNA degradation. *Mol. Cell* **75**, 1243–1255. <https://doi.org/10.1016/j.molcel.2019.06.019> (2019).
- Wang, Y. *et al.* Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature* **461**, 754–761. <https://doi.org/10.1038/nature08434> (2009).
- Nakanishi, K. Anatomy of four human Argonaute proteins. *Nucleic Acids Res.* **50**, 6618–6638. <https://doi.org/10.1093/nar/gkac519> (2022).
- Bofill-De Ros, X., Yang, A. & Gu, S. IsomiRs: Expanding the miRNA repression toolbox beyond the seed. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863**, 194373. <https://doi.org/10.1016/j.bbagr.2019.03.005> (2020).
- Yashiro, Y. & Tomita, K. Function and regulation of human terminal uridylyltransferases. *Front. Genet.* **9**, 538. <https://doi.org/10.3389/fgene.2018.00538> (2018).
- Yang, A. *et al.* TENT2, TUT4, and TUT7 selectively regulate miRNA sequence and abundance. *Nat. Commun.* **13**, 5260. <https://doi.org/10.1038/s41467-022-32969-8> (2022).
- Yates, L. A., Norbury, C. J. & Gilbert, R. J. The long and short of microRNA. *Cell* **153**, 516–519. <https://doi.org/10.1016/j.cell.2013.04.003> (2013).
- Nakanishi, K. Are argonaute-associated tiny RNAs junk, inferior miRNAs, or a new type of functional RNAs?. *Front. Mol. Biosci.* **8**, 795356. <https://doi.org/10.3389/fmolb.2021.795356> (2021).
- Li, Z. *et al.* Characterization of viral and human RNAs smaller than canonical MicroRNAs. *J. Virol.* **83**, 12751–12758. <https://doi.org/10.1128/JVI.01325-09> (2009).

18. Li, Z. *et al.* Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res.* **40**, 6787–6799. <https://doi.org/10.1093/nar/gks307> (2012).
19. Baldrich, P. *et al.* Plant extracellular vesicles contain diverse small RNA species and are enriched in 10–17-nucleotide “Tiny” RNAs. *Plant Cell* **31**, 315–324. <https://doi.org/10.1105/tpc.18.00872> (2019).
20. Sim, G. *et al.* Manganese-dependent microRNA trimming by 3'→5' exonucleases generates 14-nucleotide or shorter tiny RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2214335119. <https://doi.org/10.1073/pnas.2214335119> (2022).
21. Park, M. S., Sim, G., Kehling, A. C. & Nakanishi, K. Human Argonaute2 and Argonaute3 are catalytically activated by different lengths of guide RNA. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 28576–28578. <https://doi.org/10.1073/pnas.2015026117> (2020).
22. Chen, J. *et al.* Structural and biochemical insights into small RNA 3' end trimming by Arabidopsis SDN1. *Nat. Commun.* **9**, 3585. <https://doi.org/10.1038/s41467-018-05942-7> (2018).
23. Juvvuna, P. K., Khandelia, P., Lee, L. M. & Makeyev, E. V. Argonaute identity defines the length of mature mammalian microRNAs. *Nucleic Acids Res.* **40**, 6808–6820. <https://doi.org/10.1093/nar/gks293> (2012).
24. Kandeel, M. & Kitade, Y. Molecular dynamics and binding selectivity of nucleotides and polynucleotide substrates with EIF2C2/AgO<sub>2</sub> PAZ domain. *Int. J. Biol. Macromol.* **107**, 2566–2573. <https://doi.org/10.1016/j.ijbiomac.2017.10.145> (2018).
25. Lombardi, E. P. & Londono-Vallejo, A. A guide to computational methods for G-quadruplex prediction. *Nucleic Acids Res.* **48**, 1603. <https://doi.org/10.1093/nar/gkaa033> (2020).
26. Park, M. S. *et al.* Human Argonaute3 has slicer activity. *Nucleic Acids Res.* **45**, 11867–11877. <https://doi.org/10.1093/nar/gkx916> (2017).
27. Park, M. S. *et al.* Multidomain convergence of argonaute during RISC assembly correlates with the formation of internal water clusters. *Mol. Cell* **75**, 725–740. <https://doi.org/10.1016/j.molcel.2019.06.011> (2019).

## Acknowledgements

This work was supported by Pelotonia Fellowships (to G.Y.S. and M.S.P.), a Center for RNA Fellowship (to G.Y.S.), and the NIH (R01GM138997 to K.N.).

## Author contributions

K.N. designed the research; G.Y.S., A.C.K., M.S.P., C.D., H.Z., N.M., and J.S. performed experiments and analyzed data. K.N. wrote the paper with input from the other authors.

## Competing interests

Kotaro Nakanishi is an inventor of the patent of the relevant technology. None of the rest of the authors has a conflict of interest.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-46562-6>.

**Correspondence** and requests for materials should be addressed to K.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023