



# Loading and pre-loading processes generate a distinct siRNA population in *Tetrahymena*



Kazufumi Mochizuki\*, Henriette M. Kurth<sup>1</sup>

Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Dr. Bohr-Gasse 3, A-1030 Vienna, Austria

## ARTICLE INFO

### Article history:

Received 16 May 2013

Available online 11 June 2013

### Keywords:

siRNA

Argonaute

Dicer

*Tetrahymena*

## ABSTRACT

The various properties of small RNAs, such as length, terminal nucleotide, thermodynamic asymmetry and duplex mismatches, can impact their sorting into different Argonaute proteins in diverse eukaryotes. The developmentally regulated 26- to 32-nt siRNAs (scnRNAs) are loaded to the Argonaute protein Twi1p and display a strong bias for uracil at the 5' end. In this study, we used deep sequencing to analyze loaded and unloaded populations of scnRNAs. We show that the size of the scnRNA is determined during a pre-loading process, whereas their 5' uracil bias is attributed to both pre-loading and loading processes. We also demonstrate that scnRNAs have a strong bias for adenine at the third base from the 3' terminus, suggesting that most scnRNAs are direct Dicer products. Furthermore, we show that the thermodynamic asymmetry of the scnRNA duplex does not affect the guide and passenger strand decision. Finally, we show that scnRNAs frequently have templated uracil at the last base without a strong bias for adenine at the second base indicating non-sequential production of scnRNAs from substrates. These findings provide a biochemical basis for the varying attributes of scnRNAs, which should help improve our understanding of the production and turnover of scnRNAs *in vivo*.

© 2013 The Authors. Published by Elsevier Inc. Open access under CC BY-NC-ND license.

## 1. Introduction

Short interfering RNAs (siRNAs) are typically 20–25 nucleotides (nt) in length and are produced by the ribonuclease III enzyme Dicer. Dicer converts long double-stranded RNAs (dsRNAs) into siRNA duplexes that have 3' overhangs of 2 nt. Canonical Dicers, such as human Dicer, bind to the end of long dsRNA substrates through their PAZ domains to process the dsRNA substrates into siRNAs, and the distance between the PAZ domain and the ribonuclease active site determines the length of the siRNA product [1,2]. In contrast, budding yeasts possess a non-canonical Dicer that forms dimers that allow the consecutive production of both the 5' and 3' ends of a siRNA; thus, the size of the siRNA is determined by the distance between the two ribonuclease active sites [3]. siRNA length can also be modified by endoribonucleases and nucleotidyltransferases [4].

The siRNA duplexes are loaded to the Argonaute proteins, one strand of the duplex (passenger strand) is discarded, and the

remaining RNA strand (the guide strand) directs the Argonaute protein to target RNAs [5]. In many eukaryotes, the strand with a thermodynamically less stable 5' end is preferentially selected as the guide strand [6–9]. In plants, the terminal nucleotide of small RNAs has a strong impact on sorting (i.e., to which Argonaute protein the small RNA is loaded). For example, AGO1 selectively binds small RNAs that begin with a 5' uracil, while AGO5-bound RNAs display a strong bias for a 5' cytosine [10]. In contrast, in flies, the degree of base pairing in the small RNA duplex dominates sorting. Small RNA duplexes with perfect complementarity, typically siRNAs, are likely to bind to Ago2, whereas small RNAs with central mismatches, typically microRNAs, usually bind to Ago1 [11,12].

The ciliated protozoan *Tetrahymena thermophila* expresses two types of siRNAs. One is the constitutively expressed ~23- to 24-nt siRNAs [13], which are produced by the Dicer protein Dcr2p [14] and are mainly loaded to the Argonaute proteins Twi2p and Twi8p [15]. The other type of siRNAs, known as scnRNAs, is the ~29-nt siRNAs that are detected exclusively during conjugation, the sexual reproduction process in *Tetrahymena* [16,17]. scnRNAs are produced in the germline micronucleus by the Dicer protein Dcl1p [18] [19] and are exported to the cytoplasm where they are loaded to the Argonaute protein Twi1p [16]. The passenger strand of the scnRNA duplex is cleaved by the endoribonuclease activity of Twi1p and is discarded [20]. The remaining guide strand is 2'-O-methylated at the 3' end by the RNA methyltransferase Hen1p [21]. Twi1p-scnRNA complexes are eventually transferred

\* Corresponding author.

E-mail address: [kazufumi.mochizuki@imba.oeaw.ac.at](mailto:kazufumi.mochizuki@imba.oeaw.ac.at) (K. Mochizuki).

<sup>1</sup> Present address: Viollier AG, Spalenring 145/147, 4002 Basel, Switzerland.

to the somatic macronucleus [16] [20] where they play a key role in programmed DNA elimination [22].

scnRNAs are ~29 nt in length [16], and the majority (~85%) of scnRNAs that are isolated from the late conjugation stages have a 5' uracil (5' U) [13,15]. However, the processes that contribute to the length distribution and the 5' U bias of scnRNAs have not been addressed. In this study, we use deep sequencing to compare Twi1p-loaded scnRNAs to unloaded scnRNAs to understand how the production of the scnRNAs by Dicer and the loading of siRNAs to Argonaute shape the population of scnRNAs *in vivo*.

## 2. Materials and methods

### 2.1. cDNA library preparation, sequencing and RNA analysis

Wild-type *Tetrahymena thermophila* strains B2086 and CU428 and the complete *Twil1* KO strains [17] were used for total RNA preparation. Twi1p was immunoprecipitated [20] from either wild-type cells using an anti-Twi1p antibody or from the strains expressing HA-Twi1p and HA-Twi1p-hAGO2Lmut (Supplementary Information) using an anti-HA antibody. The cDNA libraries were constructed as described [23]. For multiplex sequencing, oligonucleotides listed in Supplementary Information were used. Single-read 36- or 50-base sequences were generated using the Illumina GAII or HiSeq2000, respectively. The raw sequencing datasets have been deposited at the NCBI Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) as GSE47650. Sequence reads were filtered as described [23]. To study the base composition and thermodynamic stability, the reads were collapsed to non-redundant datasets. To model scnRNA duplexes, 29-nt RNA sequences were mapped to the micronuclear LMR region [23] and the passenger strand sequences were estimated. The thermodynamic stabilities of 4 base pairs at the ends of the modeled scnRNA duplexes were calculated using the free energy parameters at 37 °C in 1 M NaCl [24] without including the initiation value. To study potential trimming and tailing reactions, RNA sequences were mapped to the LMR region by allowing two mismatches. Terminator exonuclease treatment is described in Supplementary Information.

## 3. Results and discussion

### 3.1. Twi1p-loaded scnRNAs contain a Dicer signature

The *Tetrahymena* Argonaute family protein Twi1p interacts with the ~29-nt siRNAs, known as scnRNAs [16,17]. To analyze Twi1p-loaded scnRNAs, cell lysate was prepared from wild-type cells at 3 h post-mixing and the Twi1p-containing complex was immunoprecipitated using an anti-Twi1p antibody [20]. Co-precipitated small RNAs (~24–34 nt) were excised after gel fractionation, and analyzed with high-throughput sequencing (Table S1).

scnRNAs are reportedly ~26–32 nt in length and the majority are 28–30 nt long [13,16]. Consistent with these data, 73.4% of the sequenced Twi1p-associated RNAs were 28–30 nt in length, and 29-nt RNAs were the most abundant (30.6%) species (Fig. 1A). The base compositions of the unique set of Twi1p-associated 29-nt scnRNAs are shown in Fig. 1B. The overall adenine (A)/uracil (U) richness of the scnRNAs most likely reflects the adenine (A)/thymine (T)-rich *Tetrahymena* genome (78.0% A/T). In addition, there were three notable base biases at specific positions of the 29-nt scnRNAs: a strong bias (89.1%) for U as the first nucleotide (5' U bias), a bias (52.9%) for A as the third nucleotide from the 3' end (-3A bias), and a bias (48.7%) for U as the last nucleotide (3' U bias) (Fig. 1B). The 5' U bias was also detected in previous studies of scnRNAs cloned from total *Tetrahymena* cells [13] as well as from another ciliate, *Paramecium tetraurelia* [25]. We believe that these

base biases were not experimental artifacts because these features were not detected in the 29-nt RNAs in the same cDNA library that matched the sense strand of the rRNAs or tRNAs and thus were most likely the degradation products of these RNAs (Fig. 1C).

The presence of both a 5' U and a -3A bias in the 29-nt scnRNAs strongly indicates that these scnRNAs were formed as dsRNAs that had 3'-overhangs of 2 nt (Fig. 1D). A slightly higher incidence of G as the second and third bases and of C as the 25th and 26th bases (Fig. 1B) also supports this view. The 5' U bias was significantly more prominent than the -3A bias, likely because 5' U scnRNAs were preferentially loaded to Twi1p, which will be discussed below in more detail. Similar signatures of dsRNAs possessing 2-nt 3'-overhangs were also observed in Twi1p-associated RNAs of other lengths (5' U bias: Fig. 1E; -3A bias: Fig. 1F). Small dsRNAs with 2-nt 3'-overhangs are characteristic Dicer products and the results above are consistent with the Dcl1p-dependent production of scnRNAs from long dsRNA precursors [18,19].

All Twi1p-bound scnRNAs at a late conjugation stage have been reported to possess a 5'-monophosphate [15], and we confirmed this result by treating total RNAs from cells at early and late conjugation stages (4 h and 8 h post-mixing, respectively) with a 5'-monophosphate-dependent exonuclease (Fig. S1). These results likely indicate that none of the scnRNAs are direct RNA-dependent RNA polymerase (RdRP) products, which typically have 5'-triphosphates [26], although it is possible that a *Tetrahymena* RdRP amplifies the dsRNA substrates before Dicer processing.

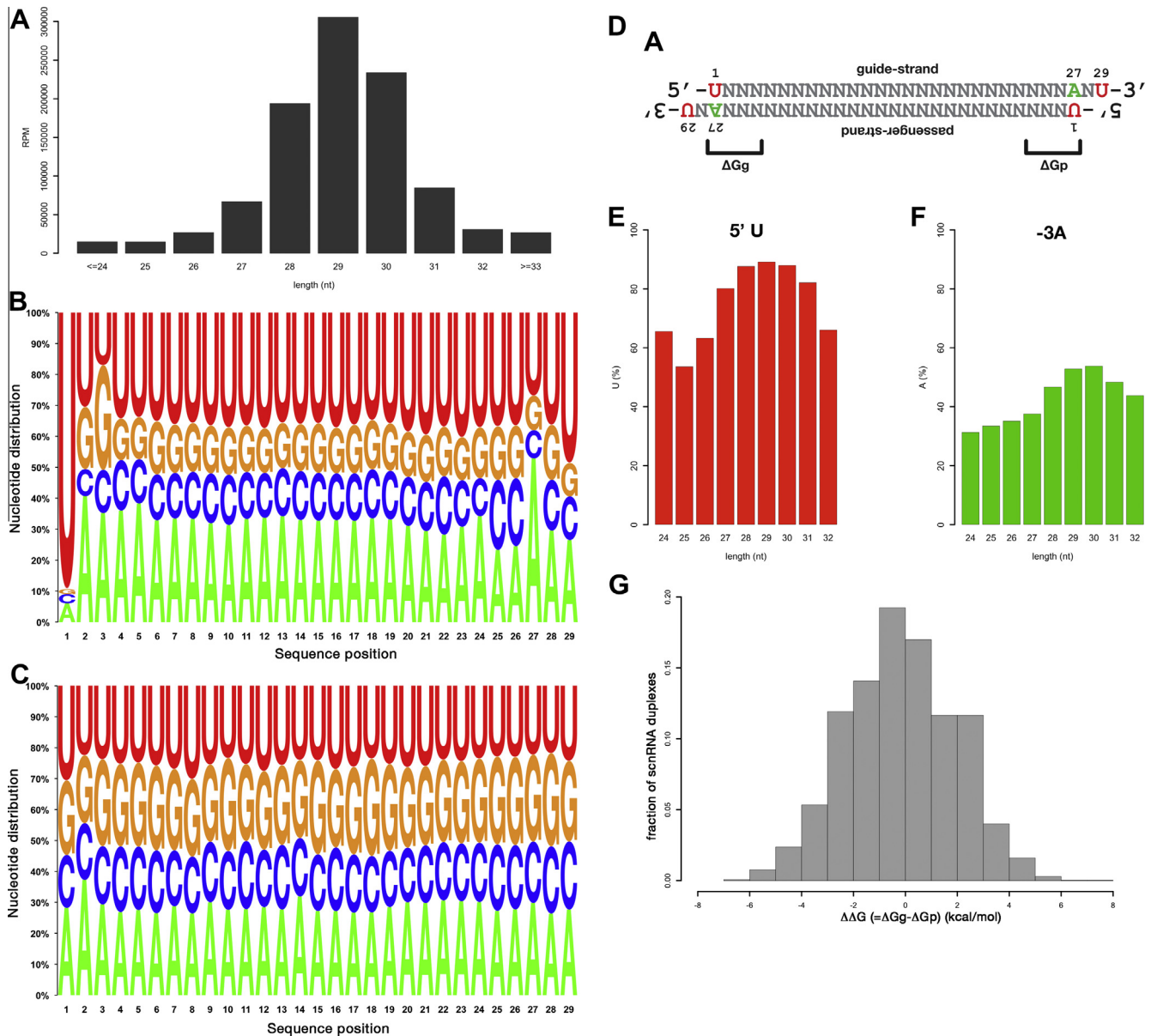
### 3.2. Thermodynamic asymmetry of the scnRNA duplex is unrelated to the guide strand decision

The thermodynamic differences in the base-pairing stability of the respective 5' ends of the two small RNA strands determine which strand of the small RNA duplex is assembled into fly Ago2 [8], plant AGO1 [9] or mammalian Argonaute proteins [6,7]. In these cases, the strand with the less stable 5' end is preferentially selected as the guide strand. We asked if the thermodynamic asymmetry of the scnRNA duplex plays a role in the choice of the guide strand of scnRNAs in *Tetrahymena*.

To calculate the thermodynamic stability of both ends of a scnRNA duplex, we first modeled the guide-passenger strand pairs of the scnRNAs. The Twi1p-associated (guide strand) 29-nt scnRNAs were mapped to a representative 100-kb micronuclear genome locus (LMR locus). A total of 4179 unique 29-nt scnRNAs were mapped to the LMR locus without mismatches. We then deduced the passenger strand sequence by examining the genomic sequences surrounding the mapped sites. Using the modeled guide and passenger pairs, we calculated the thermodynamic stability of the 4 base pairs at the termini containing the 5' ends of the guide strand scnRNAs ( $\Delta G_g$ ) and those of the passenger strand scnRNAs ( $\Delta G_p$ ) (see Fig. 1D) [24]. If the thermodynamic asymmetry of the scnRNA duplex determines the guide strand selection, as observed for the small RNAs in the other eukaryotes mentioned above, the  $\Delta G_g$  must be larger (i.e., more unstable) than the  $\Delta G_p$ . However, the  $\Delta \Delta G$  ( $=\Delta G_g - \Delta G_p$ ) values of the individual scnRNA duplexes were distributed nearly equally to both the positive and negative sides (Fig. 1G). Therefore, we conclude that the base-pairing stability of the ends of the scnRNA duplexes does not significantly influence scnRNA loading to Twi1p.

### 3.3. The length distribution of scnRNAs is determined before loading

The length distributions and the base compositions of the scnRNAs described above can reflect the preferences of the scnRNA processing, export from the micronucleus or the loading of the scnRNAs to Twi1p. To understand which features are attributed to which of these preferences, we analyzed the small RNAs of



**Fig. 1.** Analysis of Twi1p-associated small RNAs. (A) Size distribution of the Twi1p-associated RNAs from the wild-type cells. The number of sequenced RNAs (reads per million sequences (RPM)) of each size is shown. (B) Base composition of 29-nt scnRNAs. (C) Base composition of the 29-nt RNAs derived from rRNAs or tRNAs. (D) Deduced structure of the 29-nt scnRNAs before loading. (E) Fraction of the RNAs with uracil as the first base (5' U). (F) Fraction of the RNAs with adenine as the third base from the 3' end (-3A). (G) Thermodynamic stability differences between the two ends of the modeled scnRNA duplexes. The thermodynamic stability of the 4 base pairs containing the 5' end of the guide ( $\Delta G_g$ ) and the passenger ( $\Delta G_p$ ) strand (see (D)) of the individual modeled scnRNA pairs were calculated, and the differences ( $\Delta\Delta G = \Delta G_g - \Delta G_p$ ) are shown as a histogram.

*Twil* KO cells. We believe the scnRNAs in the *Twil* KO cells remain unloaded because Twi1p is the only Argonaute protein that is known to associate with scnRNAs at 3 h post-mixing [15], which was when we extracted RNAs.

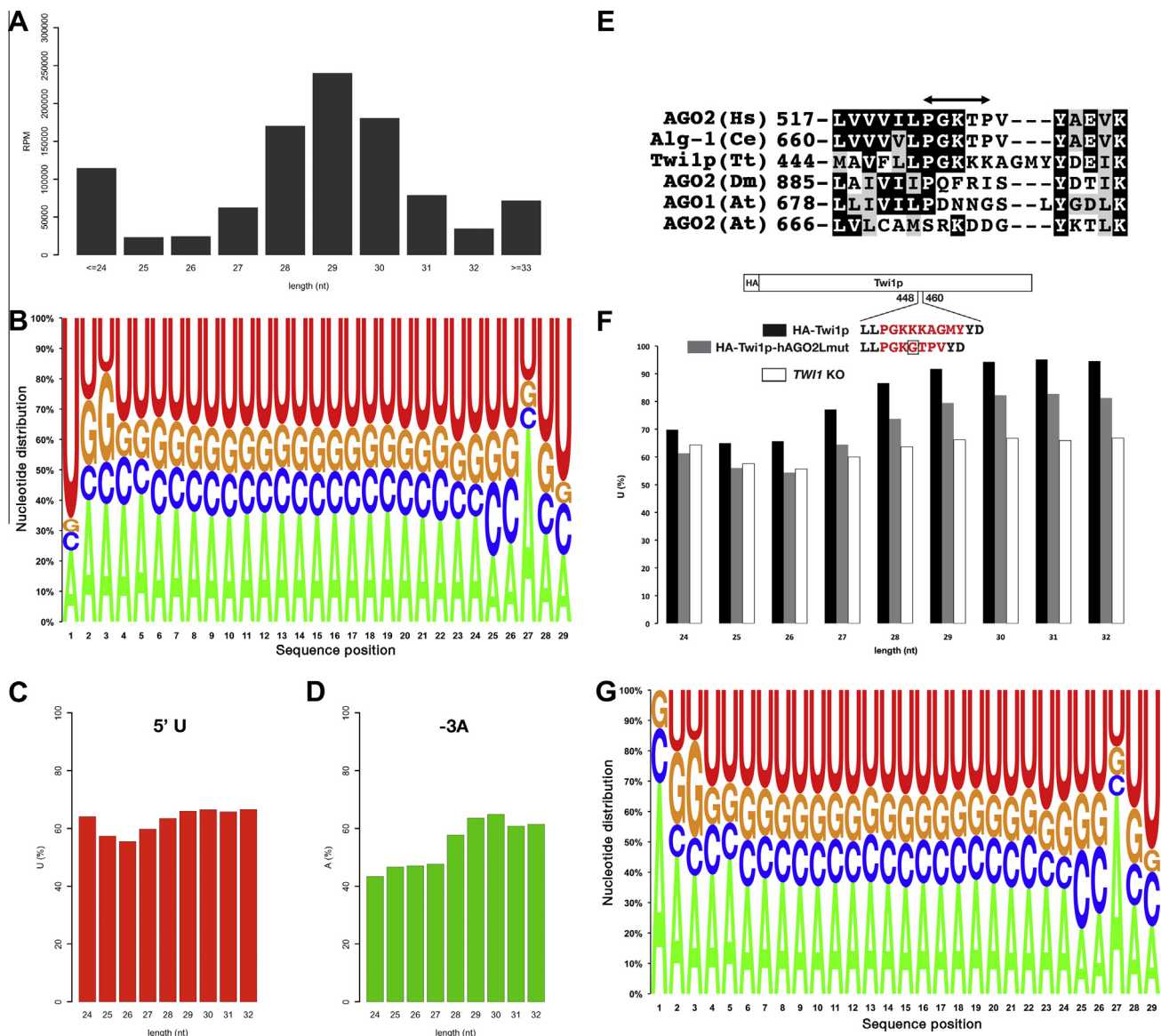
Total RNA was extracted from the *Twil* KO cells at 3 h post-mixing, and small RNAs (~24–34 nt) were analyzed. More RNAs that were detected in the *Twil* KO cells were either shorter than 25 nt or longer than 33 nt (Fig. 2A) than among the Twi1p-associated RNAs from the wild-type cells (Fig. 1A). The shorter RNAs likely represent the constitutively expressed 23- to 24-nt siRNAs, which are mainly associated with Twi2p and Twi8p [15]. The longer small RNAs were likely the degradation products of mRNAs and other non-coding RNAs. Nonetheless, the length distribution of the 26- to 32-nt RNAs from the *Twil* KO cells (Fig. 2A) was similar to that of the Twi1p-associated RNAs from the wild-type cells

(Fig. 1A). Therefore, we conclude that the length distribution for the scnRNAs in the wild-type cells is mainly determined at a pre-loading process, most during processing by Dcl1p, and that Twi1p is able to flexibly incorporate ~26- to 32-nt scnRNAs.

### 3.4. 5' U scnRNAs are preferentially complexed with Twi1p

Although ~90% of the Twi1p-associated 28- to 30-nt scnRNAs from the wild-type cells had a 5' U (Fig. 1B and E), only about two-thirds of the 28- to 30-nt scnRNAs from the *Twil* KO cells has a 5' U (Fig. 2B and C). Therefore, 5' U scnRNAs are preferentially loaded to Twi1p. It is possible that the 5' U scnRNAs are preferentially exported from the micronucleus to the cytoplasm, where the scnRNAs are loaded to Twi1p. Alternatively, the 5' U scnRNAs could preferentially interact with Twi1p. To distinguish these two possi-



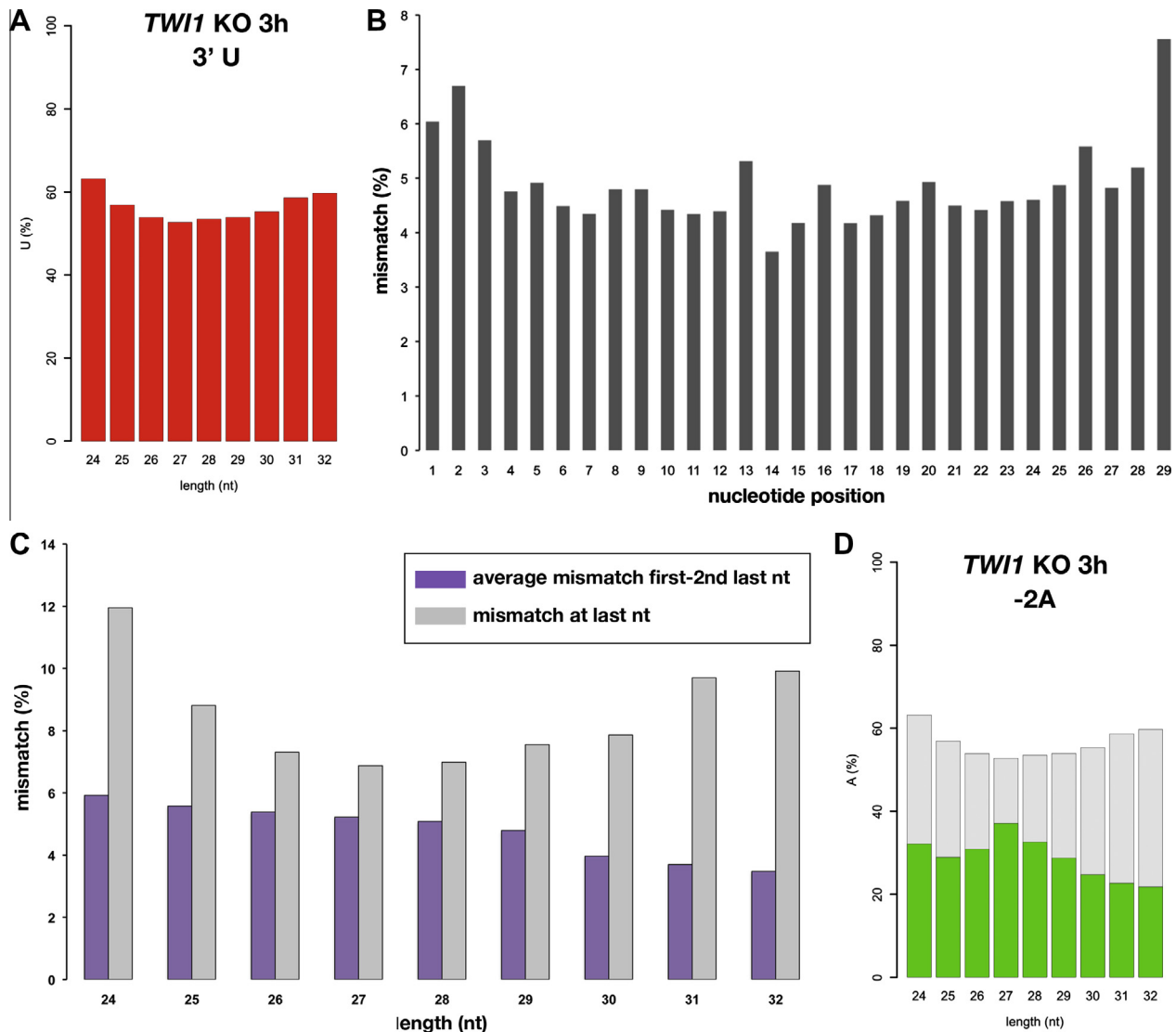


**Fig. 2.** Analyses of unloaded scnRNAs and the nucleotide specificity loop. (A) Size distribution of the small RNAs from the *TWI1* KO cells. The number of sequenced RNAs (reads per million sequences (RPM)) of each size is shown. (B) Base composition of the 29-nt RNAs. (C) Fraction of the RNAs with uracil as the first base (5' U). (D) Fraction of the RNAs with adenine as the third base from the 3' end (-3A). (E) Comparison of the nucleotide specificity loops of the different Argonaute proteins. The region corresponding to the proposed nucleotide specificity loop [27] of human AGO2 is marked with a double-headed arrow. Human (Hs) AGO2 and *C. elegans* (Ce) Alg-1, which tend to bind to 5' U/A RNAs, share a conserved sequence in the loop. Tt: *Tetrahymena thermophila*; Dm: *Drosophila melanogaster*; At: *Arabidopsis thaliana*. (F) Fraction of the 5' U RNAs in HA-Twi1p-associated (black), HA-Twi1p-hAGO2Lmut-associated (gray), or total *TWI1* KO (white, same data as (C)) small RNAs. The replaced nucleotide specificity loop region is shown in red. The extra glycine inserted into the nucleotide specificity loop of human AGO2 is marked with a square. (G) Base composition of the non-5' U 29-nt RNAs from the *TWI1* KO cells. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

bilities, we examined the importance of the nucleotide specificity loop in the 5' U bias of Twi1p-loaded scnRNAs.

Human AGO2 (hAGO2) has nucleotide-specific interactions to ensure the preference for a 5' U or A and to exclude G or C through the nucleotide specificity loop [27]. Interestingly, when one glycine residue is inserted into the nucleotide specificity loop of hAGO2, hAGO2 binds with similar affinity to any ribonucleotide [27]. Because the region including this loop is partially conserved in Twi1p (Fig. 2E), we hypothesized that this region could contribute to selectively load 5' U scnRNAs to Twi1p. We expressed, in the *TWI1* KO cells, either wild-type, N-terminally HA-tagged Twi1p (HA-Twi1p) or HA-Twi1p-hAGO2Lmut in which the nucleotide specificity loop region of Twi1p was replaced with that of hAGO2 with the glycine insertion (Fig. 2F). HA-Twi1p and HA-Twi1p-hAGO2Lmut were immunoprecipitated

using an anti-HA antibody at 3 h post-mixing and co-precipitated small RNAs were analyzed. While the HA-Twi1p-bound RNAs exhibited a strong 5' U bias (Fig. 2G), which was similar to those co-precipitated with the endogenous Twi1p (Fig. 1B and E), the Twi1p-hAGO2Lmut-associated RNAs had a weaker 5' U bias (Fig. 2G). There was no other obvious difference between the HA-Twi1p-associated and HA-Twi1p-hAGO2Lmut-associated RNAs (Fig. S2). We conclude that the 5' U bias of the Twi1p-loaded scnRNAs in the wild-type cells is attributed, at least in part, to the preferential interaction of the nucleotide specificity loop of Twi1p to 5'U RNAs. However, because the Twi1p-hAGO2Lmut-bound scnRNAs had a higher 5' U incidence than the scnRNAs from *TWI1* KO cells (Fig. 2G), the nucleotide specificity loop might not be the only determinant of the 5' U-biased loading of scnRNAs to Twi1p.



**Fig. 3.** Analysis of the 3' U bias of scnRNAs. (A) Fraction of the RNAs from the *TWI1* KO cells with uracil as the last base (3' U). (B) Analysis of the potential trimming and tailing reactions on the 29-nt RNAs from the *TWI1* KO cells. The 29-nt RNAs from the *TWI1* KO cells were mapped to the 100-kb LMR genomic locus by allowing for two base mismatches and were analyzed to determine the positions at which the mismatches were located. (C) Fraction of the 24- to 32-nt RNAs from the *TWI1* KO cells with mismatches in the LMR locus. The 24- to 32-nt RNAs from the *TWI1* KO cells were analyzed as in (B), and the average frequency of mismatches at the first base to the second base from the end (purple) and at the last base (gray) are shown. (D) Fraction of the RNAs from the *TWI1* KO cells with adenine as the second base from the end (-2A) is shown in green. Expected fraction of -2A RNAs, if scnRNA processing occurs sequentially (this is equal to the fraction of 3'U RNAs in (A)), is shown in gray. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.5. The 5' U bias of scnRNAs is also attributable to a pre-loading process

Although the fraction of scnRNAs with a 5' U was lower in the scnRNAs from the *TWI1* KO cells than in the *Twi1p*-associated scnRNAs from the wild-type cells, the scnRNAs from the *TWI1* KO cells were still highly biased toward 5' U (Fig. 2B and C). This strongly suggests that the 5' U bias of the *Twi1p*-associated scnRNAs is partly due to one of the steps that occurs prior to their loading to *Twi1p*. *Dcl1p* likely preferentially digests phosphodiester bonds that are followed by U in substrate dsRNAs. Another possibility is that an unknown mechanism in the pre-loading steps stabilizes the 5' U scnRNAs. If the bias is due to preferential digestion by *Dcl1p*, this would be the first example where a siRNA terminal base bias can be significantly attributed not only to Argonaute proteins but also to the siRNA production machinery, although minor contributions of Dicer processing to the 5' U bias of fission yeast siRNAs have been reported [28].

The ~65% 5' U bias of the scnRNAs from the *TWI1* KO cells can be explained if one strand of the scnRNA duplex always has a 5' U while the other strand has a random base at the 5' end. If this is the case, all scnRNAs not beginning with U should have A as the third base from the 3' end (-3A). However, only 66.2% of non-5' U 29-nt scnRNAs from the *TWI1* KO cells had an A at position 27 (Fig. 2G), indicating that both strands of the scnRNA duplexes have a 5' U at a frequency of ~65% before loading. This is in sharp contrast to the *Paramecium* scnRNAs in which only non-5'U RNAs have a high incidence of -3A [25]. Therefore, the *Dicers* of *Paramecium* and *Tetrahymena* most likely have different processing preferences.

### 3.6. The unloaded scnRNAs frequently have templated 3'U

More than half of the scnRNAs from the *TWI1* KO cells ended with U (Figs. 2B and 3A). We hypothesized that this 3' U bias could be due to the replacement or addition of the 3' end nucleotides by an exonuclease and a nucleotidyltransferase, which could prefer-

entially add a U at the end. To test this possibility, we first computationally extracted the 29-nt scnRNA sequences from the *TWI1* KO cells that were complementary to the micronuclear LMR locus when two base mismatches were allowed. We then asked at which positions of the scnRNAs the mismatches were located. Using this method, we could estimate how frequently base replacements or additions post-transcriptionally occurred at each position of the scnRNAs.

Among the 29,147 unique 29-nt scnRNAs from the *TWI1* KO cells that mapped the LMR locus within 2 mismatches, 7.6% had mismatches at the last base (Fig. 3B). Although this frequency was slightly higher than the fraction of scnRNAs with mismatches at each of the other positions (4–7%; Fig. 3C), it was not enough to explain the strong 3' U bias (Figs. 2B and 3A). Moreover, because many of these mismatches could have been caused by sequencing errors and false mapping of scnRNAs that were produced from loci outside of the LMR locus, the frequency of base replacements was likely overestimated. Similar results were obtained by analyzing scnRNAs of other lengths, although slightly higher mismatches at the last base were detected in shorter (24–25 nt) and longer (31–32 nt) small RNAs (Fig. 3C). Some fraction of the 23- to 24-nt siRNAs is modified by a nucleotidyltransferase activity [15]. Similarly, some longer scnRNAs may be modified by a nucleotidyltransferase. Nonetheless, these results indicate that most of the terminal nucleotides of scnRNAs are templated by the genome.

An important implication from this result is that Dcl1p likely preferentially digests phosphodiester bonds preceded by U. If scnRNAs are sequentially produced from dsRNA substrates, the last base of a scnRNA should be base paired with the second base from the 3' terminus of the adjacent scnRNA in the substrate dsRNA. However, although more than half of the scnRNAs had a U as their last base, the fraction of A at the second position from the 3' terminus did not exceed the fraction of A in the total genome (39.0%) in the scnRNAs from the *TWI1* KO cells (Fig. 3D). Therefore, if Dcl1p preferentially produces scnRNAs ending with U, scnRNAs must be produced non-sequentially from dsRNA substrates (i.e., a digestion that produces the 3' end of one scnRNA does not produce the 5' end of another scnRNA), which is an atypical siRNA production mode that has not been reported. Because the Dcl1p of *Tetrahymena* lacks a PAZ domain [13,18,19] by which the canonical Dicers can bind to the end of dsRNA substrates [1,2], Dcl1p may produce scnRNAs by a non-canonical mechanism similar to the budding yeast, although the budding yeast Dicer produces siRNAs sequentially *in vitro* [3]. In this context, it would be interesting to analyze the production of scnRNAs by *Tetrahymena* Dcl1p in an *in vitro* system.

## Acknowledgments

We thank all members of the Mochizuki group for helpful discussions. The work was supported by an ERC Starting Grant (204986) under the European Community's Seventh Framework Programme, the Special Research Program (SFB) "RNA regulation of the transcriptome" (F4307-B09) from the Austrian Science Fund (FWF) and core funding from the Austrian Academy of Sciences.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bbrc.2013.05.133>.

## References

- [1] H. Zhang, F.A. Kolb, L. Jaskiewicz, E. Westhof, W. Filipowicz, Single processing center models for human Dicer and bacterial RNase III, *Cell* 118 (2004) 57–68.
- [2] I.J. Macrae, K. Zhou, F. Li, A. Repic, A.N. Brooks, W.Z. Cande, P.D. Adams, J.A. Doudna, Structural basis for double-stranded RNA processing by Dicer, *Science* 311 (2006) 195–198.
- [3] D.E. Weinberg, K. Nakanishi, D.J. Patel, D.P. Bartel, The inside-out mechanism of Dicers from budding yeasts, *Cell* 146 (2011) 262–276.
- [4] B.W. Han, J.H. Hung, Z. Weng, P.D. Zamore, S.L. Ameres, The 3'-to-5' exonuclease Nibbler shapes the 3' ends of microRNAs bound to *Drosophila* Argonaute1, *Curr. Biol.* 21 (2011) 1878–1887.
- [5] T. Kawamata, Y. Tomari, Making RISC, *Trends Biochem. Sci.* 35 (2010) 368–376.
- [6] A. Khvorova, A. Reynolds, S.D. Jayasena, Functional siRNAs and miRNAs exhibit strand bias, *Cell* 115 (2003) 209–216.
- [7] D.S. Schwarz, G. Hutvagner, T. Du, Z. Xu, N. Aronin, P.D. Zamore, Asymmetry in the assembly of the RNAi enzyme complex, *Cell* 115 (2003) 199–208.
- [8] Y. Tomari, C. Matranga, B. Haley, N. Martinez, P.D. Zamore, A protein sensor for siRNA asymmetry, *Science* 306 (2004) 1377–1380.
- [9] A.L. Eamens, N.A. Smith, S.J. Curtin, M.B. Wang, P.M. Waterhouse, The *Arabidopsis thaliana* double-stranded RNA binding protein DRB1 directs guide strand selection from microRNA duplexes, *RNA* 15 (2009) 2219–2235.
- [10] S. Mi, T. Cai, Y. Hu, Y. Chen, E. Hodges, F. Ni, L. Wu, S. Li, H. Zhou, C. Long, S. Chen, G.J. Hannon, Y. Qi, Sorting of small RNAs into *Arabidopsis* Argonaute complexes is directed by the 5' terminal nucleotide, *Cell* 133 (2008) 116–127.
- [11] B. Czech, R. Zhou, Y. Erlich, J. Brennecke, R. Binari, C. Villalta, A. Gordon, N. Perrimon, G.J. Hannon, Hierarchical rules for Argonaute loading in *Drosophila*, *Mol. Cell* 36 (2009) 445–456.
- [12] K. Okamura, N. Liu, E.C. Lai, Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes, *Mol. Cell* 36 (2009) 431–444.
- [13] S.R. Lee, K. Collins, Two classes of endogenous small RNAs in *Tetrahymena thermophila*, *Genes Dev.* 20 (2006) 28–33.
- [14] S.R. Lee, K. Collins, Physical and functional coupling of RNA-dependent RNA polymerase and Dicer in the biogenesis of endogenous siRNAs, *Nat. Struct. Mol. Biol.* 14 (2007) 604–610.
- [15] M.T. Couvillion, S.R. Lee, B. Hogstad, C.D. Malone, L.A. Tonkin, R. Sachidanandam, G.J. Hannon, K. Collins, Sequence, biogenesis, and function of diverse small RNA classes bound to the Piwi family proteins of *Tetrahymena thermophila*, *Genes Dev.* 23 (2009) 2016–2032.
- [16] K. Mochizuki, N.A. Fine, T. Fujisawa, M.A. Gorovsky, Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in *Tetrahymena*, *Cell* 110 (2002) 689–699.
- [17] K. Mochizuki, M.A. Gorovsky, Conjugation-specific small RNAs in *Tetrahymena* have predicted properties of scan (scn) RNAs involved in genome rearrangement, *Genes Dev.* 18 (2004) 2068–2073.
- [18] C.D. Malone, A.M. Anderson, J.A. Motl, C.H. Rexer, D.L. Chalker, Germ line transcripts are processed by a Dicer-like protein that is essential for developmentally programmed genome rearrangements of *Tetrahymena thermophila*, *Mol. Cell. Biol.* 25 (2005) 9151–9164.
- [19] K. Mochizuki, M.A. Gorovsky, A Dicer-like protein in *Tetrahymena* has distinct functions in genome rearrangement, chromosome segregation, and meiotic prophase, *Genes Dev.* 19 (2005) 77–89.
- [20] T. Noto, H.M. Kurth, K. Kataoka, L. Aronica, L.V. Desouza, K.W. Siu, R.E. Pearlman, M.A. Gorovsky, K. Mochizuki, The *Tetrahymena* Argonaute-binding protein Giw1p directs a mature Argonaute-siRNA complex to the nucleus, *Cell* 140 (2010) 692–703.
- [21] H.M. Kurth, K. Mochizuki, 2'-O-Methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in *Tetrahymena*, *RNA* 15 (2009) 675–685.
- [22] M.C. Yao, J.L. Chao, RNA-guided DNA deletion in *Tetrahymena*: an RNAi-based mechanism for programmed genome rearrangements, *Annu. Rev. Genet.* 39 (2005) 537–559.
- [23] U.E. Schoeberl, H.M. Kurth, T. Noto, K. Mochizuki, Biased transcription and selective degradation of small RNAs shape the pattern of DNA elimination in *Tetrahymena*, *Genes Dev.* 26 (2012) 1729–1742.
- [24] S.M. Freier, R. Kierzek, J.A. Jaeger, N. Sugimoto, M.H. Caruthers, T. Neilson, D.H. Turner, Improved free-energy parameters for predictions of RNA duplex stability, *Proc. Natl. Acad. Sci. USA* 83 (1986) 9373–9377.
- [25] G. Lepere, M. Nowacki, V. Serrano, J.F. Gout, G. Guglielmi, S. Duhaucourt, E. Meyer, Silencing-associated and meiosis-specific small RNA pathways in *Paramecium tetraurelia*, *Nucleic Acids Res.* 37 (2009) 903–915.
- [26] T. Sijen, J. Fleenor, F. Simmer, K.L. Thijssen, S. Parrish, L. Timmons, R.H. Plasterk, A. Fire, On the role of RNA amplification in dsRNA-triggered gene silencing, *Cell* 107 (2001) 465–476.
- [27] F. Frank, N. Sonenberg, B. Nagar, Structural basis for 5'-nucleotide base-specific recognition of guide RNA by human AGO2, *Nature* 465 (2010) 818–822.
- [28] M. Buhler, N. Spies, D.P. Bartel, D. Moazed, TRAMP-mediated RNA surveillance prevents spurious entry of RNAs into the *Schizosaccharomyces pombe* siRNA pathway, *Nat. Struct. Mol. Biol.* 15 (2008) 1015–1023.