Cell

The Loop Position of shRNAs and Pre-miRNAs Is Critical for the Accuracy of Dicer Processing In Vivo

Shuo Gu,[1] Lan Jin,[1] Yue Zhang,[1] Yong Huang,[1] Feijie Zhang,[1] Paul N. Valdmanis,[1] and Mark A. Kay[1,*]
[1]Departments of Pediatrics and Genetics, Stanford University, Stanford, CA 94305, USA
*Correspondence: markay@stanford.edu
http://dx.doi.org/10.1016/j.cell.2012.09.042

SUMMARY

Short hairpin RNA (shRNA)-induced RNAi is used for biological discovery and therapeutics. Dicer, whose normal role is to liberate endogenous miRNAs from their precursors, processes shRNAs into different biologically active siRNAs, affecting their efficacy and potential for off-targeting. We found that, in cells, Dicer induced imprecise cleavage events around the expected sites based on the previously described 5′/3′ counting rules. These promiscuous noncanonical cleavages were abrogated when the cleavage site was positioned 2 nt from a bulge or loop. Interestingly, we observed that the ~1/3 of mammalian endogenous pre-miRNAs that contained such structures were more precisely processed by Dicer. Implementing a "loop-counting rule," we designed potent anti-HCV shRNAs with substantially reduced off-target effects. Our results suggest that Dicer recognizes the loop/bulge structure in addition to the ends of shRNAs/pre-miRNAs for accurate processing. This has important implications for both miRNA processing and future design of shRNAs for RNAi-based genetic screens and therapies.

INTRODUCTION

MicroRNAs (miRNAs), 21–23 nt in length, are responsible for the regulation of at least one-half of all protein-encoding genes in mammals (Friedman et al., 2009). Primary miRNA transcripts (pri-miRNA) are initially transcribed from genome-encoded sequences and are then further processed into pre-miRNA and finally small RNA duplexes (reviewed in Bartel, 2004; Carthew and Sontheimer, 2009; Liu and Paroo, 2010). Based on the thermodynamic stabilities of the duplex ends, one strand of the resulting duplex (the miRNA strand or guide strand) is preferentially loaded into Argonaute proteins, the core component of the RNA-induced silencing complex (RISC) (Hammond et al., 2001; Khvorova et al., 2003; Schwarz et al., 2003). Gene expression is reduced by a process referred to as RNA interference (RNAi) through site-specific cleavage or noncleavage repression. Although the more efficient means of knocking down gene expression is induced when the target sequence has complete complementarity with the small RNA, the major mode of miRNA-induced gene regulation occurs when complementarity is maintained in the first third of the small RNA and target mRNA, but mismatches arise in the remainder of the aligned sequence (Gu and Kay, 2010; Huntzinger and Izaurralde, 2011).

The RNAi pathway can be induced to mediate transient sequence-specific gene silencing by directly transfecting chemically synthesized small interfering RNA (siRNA) duplexes into tissues or cells (Elbashir et al., 2001). Alternatively, DNA-based transcriptional templates expressing a small hairpin RNA (shRNA), which is processed into siRNA, can be used to achieve long-term gene silencing (Brummelkamp et al., 2002; McCaffrey et al., 2002; McManus et al., 2002; Paddison et al., 2002; Zeng et al., 2002). Both approaches have become routine in biological research and are used in therapeutic applications to treat various diseases (Kim and Rossi, 2007).

Most of the current transcriptional RNAi approaches developed for therapeutics or biological discovery (e.g., shRNA libraries) utilize polymerase III transcription cassettes because they are relatively simple to construct and provide high levels of expression (Paddison et al., 2004). More recently, regulated and/or cell-type-specific transcription of shRNA can be accomplished by using polymerase II promoters, but this requires that the shRNA sequences be embedded within a native or artificial pre-miRNA sequence that may affect the processing and creation of the desired siRNA product (Pan et al., 2012).

RNase III enzymes are crucial in the biogenesis of miRNAs (Kim et al., 2009). In particular, Dicer recognizes the hairpin-shaped pre-miRNA and cuts the terminal loop to generate a duplex miRNA/miRNA* containing 2 nt 3′ overhangs, a classical feature of the RNAi pathway (Gurtan et al., 2012; Zhang et al., 2002, 2004). The precise processing of Dicer is critical as inaccurate cleavage events generate miRNAs with different seed regions, altering the set of genes a particular miRNA regulates (Lewis et al., 2003). A shifted Dicer cleavage site changing the nucleotide composition of duplex ends can have profound effects on which miRNA strand is loaded into RISC. Dicer is also responsible for processing shRNAs into siRNAs (Siolas et al., 2005). One of the problems and limitations of shRNA-based RNAi approaches is the fact that an unpredictable number of various duplex RNAs are generated within a cell (McIntyre et al., 2011), which can limit their effectiveness. Therefore,

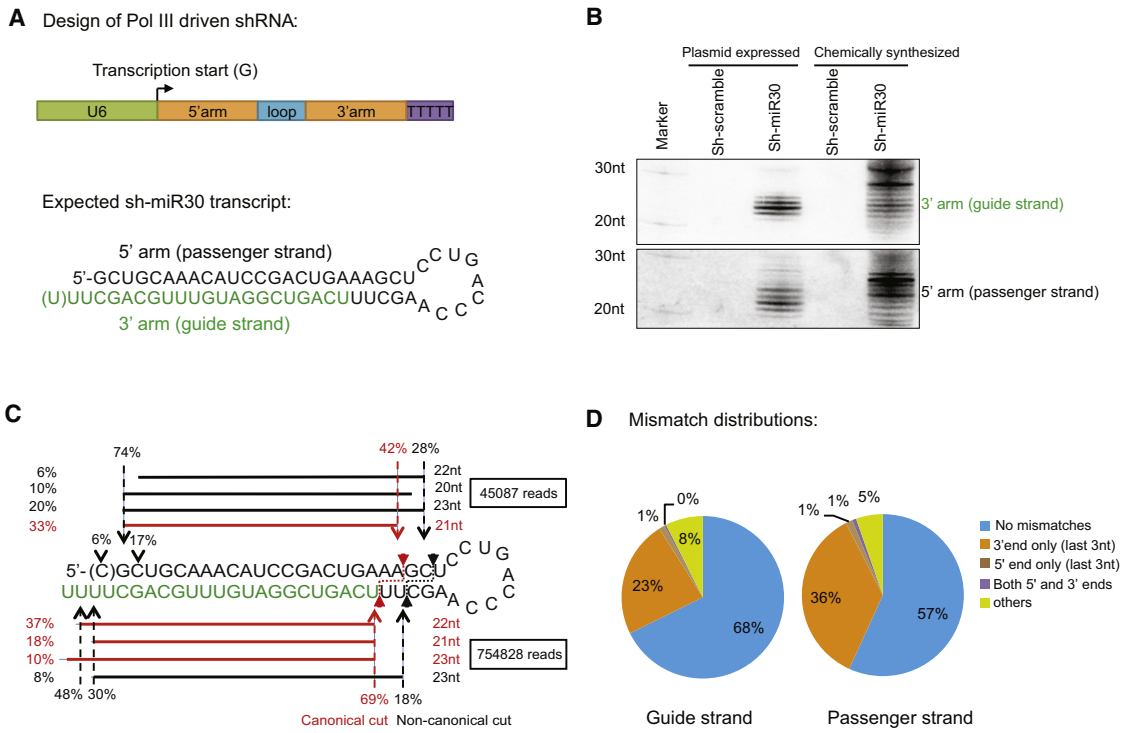

Figure 1. Dicer Performs Noncanonical and Canonical Cuts In Vivo

(A) Schematic representation of the U6-driven sh-miR30 used in this study. The expected guide-strand sequence resulting from canonical Dicer cleavage is labeled in green text.

(B) DNA plasmids expressing sh-miR30 or synthetic sh-miR30 RNAs were transfected into HEK293 cells. shRNA with a scrambled sequence served as a negative control in both cases. Extracted RNAs were run on 20% polyacrylamide 7 M urea denaturing gels 36 hr posttransfection. Guide and passenger strand were identified by sequential Northern blotting.

(C) Small RNAs from cells transfected with sh-miR30-expressing plasmids were subjected to deep sequencing. After being mapped to sh-miR30 (with up to three mismatches), the four most abundant sequences originated from the 3′ arm (guide strand), and four from the 5′ arm (passenger strand) were labeled in the figure along with their lengths and percent abundance relative to all sh-miR30-derived reads from either 5′ arm or 3′ arm accordingly. Dotted arrows and the associated numbers indicate the percentage of sequences starting or ending at those positions relative to the total number of guide-strand reads or total number of the passenger strand reads. Inferred Dicer cleavages were labeled with a small solid arrow and a dotted line. Information associated with canonical cuts is in red, whereas those with noncanonical cut are in black.

(D) Mismatch distribution of all the reads that mapped to guide or passenger strands. The majority of the mismatches are located in the 3′ ends (last 3 nt).

See Figure S1 for repeat experiments performed in MEF cells.

defining the precise rules for Dicer cleavage will be a great benefit in shRNA design.

It has been demonstrated that Dicer determines its cleavage sites by measuring a fixed distance from either the 3′ end overhang (the 3′ counting rule) (Macrae et al., 2006, 2007) or 5′ end phosphate group (the 5′ counting rule) (Park et al., 2011), with the latter being the predominant process in mammalian systems. However, both models were established based on reconstituted noncellular Dicer cleavage studies, and the applicability of these rules for shRNAs or endogenous miRNA processing in vivo has not been clearly delineated.

Here, we wanted to establish additional rules governing the processing of endogenous miRNAs and shRNAs by Dicer in living cells. To do this, we focused our attention on pol-III-driven shRNAs to eliminate potential variables that could be introduced by the additional processing steps (e.g., Drosha) required in pol-II-based expression systems. Nonetheless, the relatively simplified pol III shRNA expression system provided insights into Dicer processing that were confirmed by bioinformatic analyses to be operational in endogenous miRNA processing. Most importantly, these parameters provide a means to design shRNA expression cassettes with enhanced efficacy in gene knockdown studies.

RESULTS

Dicer Performs Noncanonical and Canonical Cuts In Vivo

To investigate how Dicer processes pre-miRNA-like substrates in vivo, we designed a U6 promoter-driven shRNA (sh-miR30) consisting of a passenger strand at the 5′ arm, a 9 nt loop (from hsa-miR-22), and a 3′ arm guide strand (based on the hsa-miR-30-3p sequence). For efficient Pol III transcription start and termination, the shRNA sequence begins with a guanine (G) and ends with a track of five thymines (T). The expected transcript is a 24-bp-long-stem shRNA with two or more uracils (U) in the 3′ overhangs, which are processed by Dicer into a guide/passenger duplex (Figure 1A). Plasmids containing the shRNA

expression cassette were transfected into HEK293 cells. The guide and passenger strands processed from sh-miR30 were separately detected by Northern blot 36 hr later. Interestingly, multiple products varying in length were identified from each strand, indicating heterogeneous Dicer cleavages and/or post-cleavage modifications (Figure 1B). Of note, these small RNA products were absent when we performed the same experiments in Dicer knockout (KO) embryonic stem (ES) cells (Calabrese et al., 2007) but were rescued by complementing Dicer expression, confirming their specificity to Dicer processing (Figure S1A available online).

To analyze the Dicer-processed cleavage products in detail, we deep sequenced all 18–32 nt small RNAs from the transfected cells. More than 750,000 reads were mapped to the guide-strand sequences (with three or fewer mismatches), whereas less than 50,000 reads were identified from the passenger strand, suggesting that the passenger strand is quickly degraded during RISC loading. Although the cleavage products were highly heterogeneous, nearly 90% of the reads that mapped to the 3′ arm (designated guide strand) started at two distinct positions. One group began at the expected cleavage site of Dicer based on the 5′ counting rule, and the other (about 18% of all the 3′ arm reads) started 2 nt upstream. 5′ RNA extension is an extremely rare event in cells (Seitz et al., 2008), making it unlikely that the latter group was the result of nucleotide addition to the former group. Rather, the results strongly suggest that, in addition to the canonical cut, Dicer was also able to make a noncanonical cut. In support of this idea, the majority (~60%) of the 5′ arm reads ended at two corresponding positions predicted by the two-cleavage model. Of note, none of the frequent reads overlapped with the endogenous hsa-miR30-3p, which represented a minor fraction (<0.5%) of the reads and were therefore omitted from our analysis. A similar cleavage pattern was observed when the same experiment was performed in mouse embryonic fibroblast (MEF) cells (Figure S1B), indicating that the noncanonical cuts were not limited to human cells nor were they tissue specific.

The 3′, but not the 5′, ends of small RNA are subject to intensive modifications such as trimming and tailing in cells (Ameres et al., 2010; Burroughs et al., 2010; Seitz et al., 2008). This was consistent with our finding that the majority of mismatches in mapped reads to sh-miR30 were located in the last 3 nt of the 3′ end, whereas less than 1% of the total reads contained mismatches in the first 3 nt at the 5′ end (Figures 1D and S1C). It was unclear whether the enriched mismatches at the 3′ end were caused by sequencing errors or by 3′ nontemplated nucleotide addition. Nonetheless, this clearly indicates that the 5′ end of the 3′ arm strand, but not the 3′ end of the 5′ arm, is the appropriate manner to infer the Dicer cleavage pattern, despite the fact that both were generated by the same cleavage events. Thus, we focused our analyses on the 5′ ends of the 3′ arm (guide strand) to further investigate Dicer processing.

Dicer's Heterogeneous Cleavages Generate Downstream Off-Target Effects

Because Dicer is a core component of the RISC loading complex (Chendrimada et al., 2005), we wanted to determine whether the noncanonical Dicer cleavage products were able to associate with downstream RISC. To address this directly, we analyzed the RISC-associated small RNAs by deep sequencing after Ago2 immunoprecipitation from cells cotransfected with Flag-tagged Ago2 and sh-miR30. The relative percentage of the Dicer cleavage products was unchanged after Ago2 immunoprecipitation (Figures 2A, S2A, and S2B), indicating that the small RNAs generated from both noncanonical and canonical Dicer cleavage can associate with the downstream RNAi pathway.

shRNAs can be designed such that, after canonical Dicer cleavage, the siRNA contains the appropriate end nucleotides that make one strand more favorable for RISC loading. Our sh-miR30 constructs are designed to ensure that the 3′ arm is processed into the guide strand that is preferentially loaded into RISC. However, noncanonical Dicer cleavage generates siRNAs with different nucleotide composition at the ends, making it possible for the passenger strands to be loaded into RISC and induce off-target effects. To verify this experimentally, we prepared two synthetic siRNAs, si-miR30-can (canonical) and si-miR30-non (noncanonical), to mimic the two predominant products of Dicer cleavage (Figure 2A). Dual-luciferase reporters containing target sequences perfectly complementary to either the guide or passenger strands in their 3′ untranslated region (UTR) (Figure 2B and Table S1) were separately cotransfected with each of the siRNAs into HEK293 cells. Although both siRNAs induced robust gene repression, si-miR30-non, but not si-miR30-can, inhibited the expression from the reporter containing a target complementary to the passenger strand (Figure 2C). Similar results were obtained when we used a reporter containing mismatched target sequences (Figure S2C). As expected, when the transcriptional-based sh-miR30 was used as a template to generate siRNAs in similar reporter knockdown experiments, passenger-strand-mediated off-target effects were observed in human (Figures 2D and S2D) and mouse cells (Figure S2E). Taken together, our results demonstrated that noncanonical Dicer cleavage products generated off-target effects from unintentional loading of the passenger strands.

Heterogeneous Processing of shRNAs Is Not Unique to sh-miR30

To generalize our observation, we designed two additional shRNAs, sh-Bantam and sh-Bantam-P, based on the *Drosophila* miRNA bantam, which has no known homologous sequence in mammals. Multiple 3′ arm small RNA products were generated when these two shRNAs were expressed in human (Figures 3A and 3B) or mouse cells (Figure S3A). Similar results were obtained when we tested a third shRNA (sh-LSW) containing an artificial sequence in both human (Figures 3A and 3C) and mouse cells (Figure S3B). The detection of canonical and noncanonical Dicer products containing various nucleotides at the 5′ end suggests that cleavage site selection was not dependent on the sequence of the RNA. Interestingly, noncanonical Dicer cleavage was as efficient as (sh-Bantam and sh-LSW) or even more prevalent (sh-Bantam-P) to the canonical cutting, suggesting that the noncanonical processing is not a minor event.

To further support our conclusion through use of a different promoter, we compared a shRNA directed against human α-antitrypsin (sh-hAAT-25) under the transcriptional control of either the U6 or H1 promoter (Grimm et al., 2006). Although

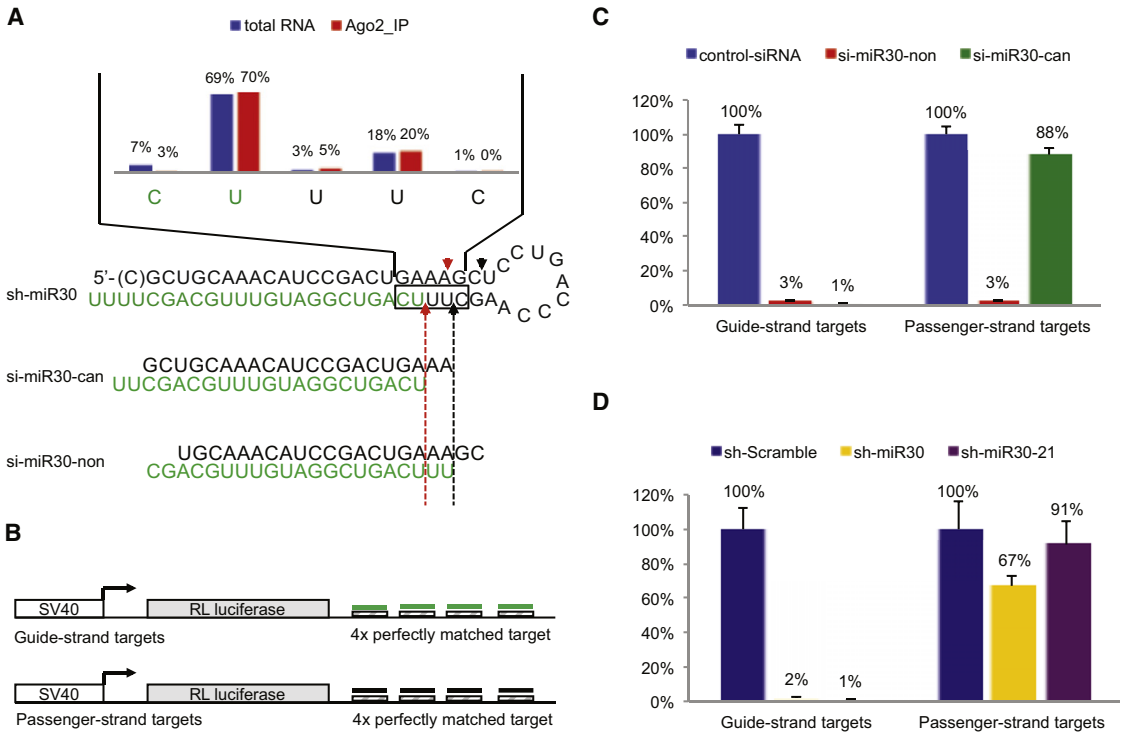

Figure 2. Heterogeneous Dicer Cleavage Generates Downstream Off-Target Effects

(A) Deep sequencing results before (Figure 1C) and after (Figure S2A) Ago2 immunoprecipitation were directly compared. In both cases, the majority of the guide-strand sequences start at two distinct positions, corresponding to two different Dicer cleavage events. The design of si-miR30-can (canonical) and si-miR30-non (noncanonical) were indicated in the figure.

(B) PsiCHECK vectors with four tandem target sites in the 3′ UTR, which were complementary to the guide or passenger strand of sh-miR30, were used to monitor the repression activity from either strand.

(C and D) (C) si-miR30-can/si-miR30-non or (D) sh-miR30 or sh-miR30-21(details in Figure 5) were cotransfected separately with psiCHECK reporters into HEK293 cells. Dual-luciferase assays were performed 36 hr posttransfection. RL-luciferase activities were normalized with FF-luciferase, and the percentage of relative enzyme activity compared to the negative control (treated with either control-siRNA or sh-scramble) was plotted. Error bars represent the SD from two independent experiments, each performed in triplicate transfections.

See also Figure S2.

the expression level of H1-sh-hAAT-25 was weaker than that of U6-sh-hAAT-25 (Figure 3A), the pattern of Dicer processing based on deep sequencing was strikingly similar (Figures 3D and S3C), suggesting that the high expression level of shRNA substrate was not likely the cause of Dicer's noncanonical processing. Of note, sh-hAAT-25 contains a different 7 nt loop sequence that has been widely used in other studies (McIntyre et al., 2011), indicating that the heterogeneous processing was also not limited to certain loop sequences/structures.

Dicer's Noncanonical Cleavage Is Independent of RNA End Heterogeneity

Pol-III-driven transcripts contain a triphosphate group at the 5′ end, which may not be efficiently recognized by Dicer and may interfere with the 5′ counting rule. In addition, the termination of Pol III polymerase leaves variable numbers of uridines at the 3′ end, which may shift the Dicer cleavage sites according to the 3′ counting rule. Therefore, it is possible that the noncanonical cleavage events observed in vivo are unique to Pol-III-driven shRNAs and can be explained by known rules of Dicer processing. To test this idea, we chemically synthesized the sh-miR30 sequences with a monophosphate group at the 5′ end and two uridines at the 3′ end. After being transfected into HEK293 cells, the small RNAs processed from synthetic sh-miR30 were analyzed by Northern blot and deep sequencing. Results from both experiments showed that the processing products of synthetic sh-miR30 were even more diversified in length and start position compared to those of expressed sh-miR30 (Figures 1B and 4A). It was possible that some of the transfected synthetic sh-miR30 were trapped in endosomes and were subject to nonspecific cleavage. Indeed, only a portion of those small RNAs originated from synthetic sh-miR30 were found to be specific to Dicer processing when we performed the same experiments in Dicer KO ES cells with or without the complementary Dicer expression (Figure S1A). Furthermore, after Ago2 immunoprecipitation, only those reads with the same start position as the predominant products of expressed sh-miR30 were enriched, indicating that in vivo Dicer processing is the same between synthetic and expressed sh-miR30 (Figure 4A). As

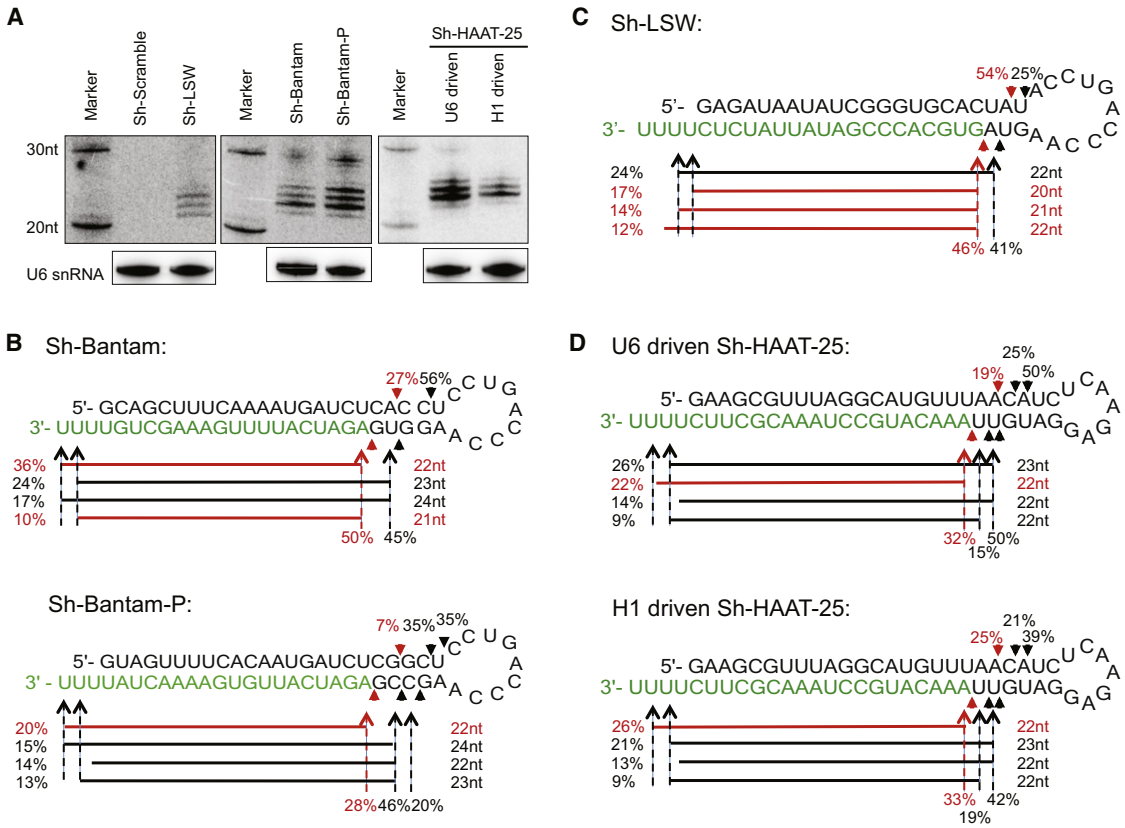

Figure 3. Heterogeneous Processing of shRNA Is a Widespread Phenomenon In Vivo

(A) Plasmids expressing various shRNAs with different stem and/or loop sequences were transfected into HEK293 cells. Extracted RNAs were run on 20% polyacrylamide 7 M urea denaturing gels 36 hr posttransfection. Small RNAs processed from the 3′ arm were identified by Northern blotting with corresponding probes. Endogenous U6 snRNA was also detected as an internal control.

(B–D) The processing of (B) sh-Bantam and sh-Bantam-P, (C) sh-LSW, and (D) sh-hAAT-25 driven by either U6 or H1 promoter in HEK293 cells was examined by deep sequencing.

Results were presented as described in Figure 1. See also Figure S3 for repeat experiments performed in MEF cells.

expected, passenger-strand-mediated off-target effects were also observed with synthetic sh-miR30 (Figure 4B).

Similar observations were made when synthetic sh-miR30 was transfected in mouse cells (Figure S4). Altogether, our results strongly suggest that noncanonical Dicer cleavage is independent of RNA end heterogeneity and is an inherent feature of the Dicer processing in vivo.

Loop-Counting Rule of Dicer Cleavage In Vivo

Because the heterogeneous Dicer cleavage is not sequence specific and is not the result of the inevitable heterogeneity of shRNA ends, we elected to investigate whether we could design shRNAs that would be homogeneously processed by Dicer. To do this, we generated a number of different sh-miR30 variants that varied in length and stem structure and then examined their processing in HEK293 cells. Whereas multiple 3′ arm processing products were detected in all designs by Northern blot (Figure 5A), the individual Dicer cleavage pattern revealed by deep sequencing was distinct (Figure 5B).

Consistent with the 5′ counting rule, Dicer made canonical cleavages at all shRNAs at a position strictly 19 bp (passenger)/21 bp (guide) away from the 5′ end. Interestingly, whereas a single mismatch in the middle of the shRNA stem did not affect the position of the canonical cut (sh-miR30-M), placement of an asymmetrical bulge at the 5′ or 3′ arm of the shRNA partially shifted the canonical cuts in opposite directions. Specifically, after a bulge was introduced at the 5′ arm of sh-miR30, the percentage of canonical cuts was reduced (69% to 51%), whereas a new cleavage site (22%) emerged one nucleotide upstream (compare sh-miR30-24B with sh-miR30 in Figure 5B), indicating that a portion of the canonical cuts were shifted toward the loop termini. Conversely, introducing a 3′ bulge shifted a portion of the canonical cuts toward the open terminus (compare sh-miR30-23B with sh-miR30). Similar observations were also made with sh-miR30-22B and sh-miR30-21B, suggesting that asymmetrical bulges should be avoided in order to achieve a distinct canonical cut.

In contrast to the canonical cut, noncanonical cutting was apparently not determined by the distance to the open ends of the shRNAs. Instead, noncanonical cleavages were observed at various positions near the site of the canonical cut without a specific pattern. Interestingly, although the mismatch/bulge

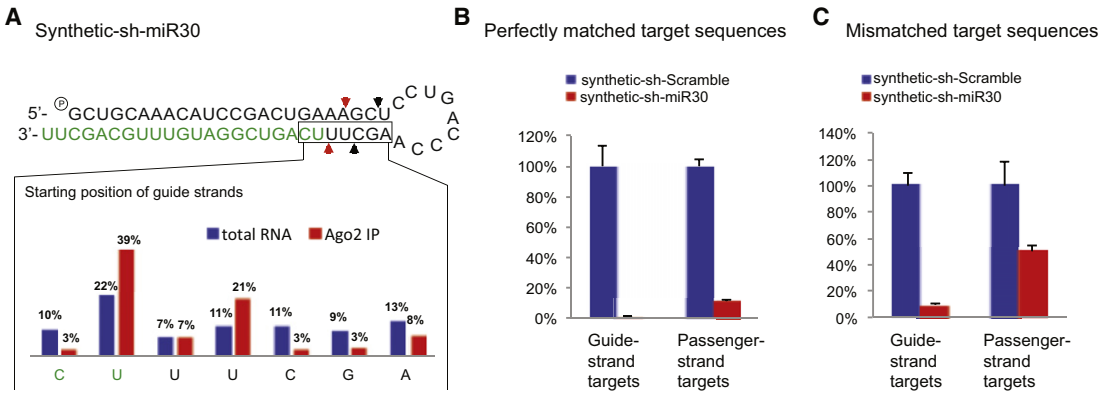

Figure 4. Noncanonical Dicer Cleavage Is Independent of RNA End Heterogeneity

(A) Synthetic sh-miR30 was coexpressed with Flag-Ago2 in HEK293 cells. Small RNAs from the 3′ arm (guide strands) were analyzed by deep sequencing. The starting positions of small RNAs were directly compared before and after Ago2 immunoprecipitation.

(B and C) PsiCHECK reporters with four tandem target sites in the 3′ UTR that were perfectly complementary (B) or mismatched (C) to the guide or passenger strand of sh-miR30 were cotransfected with synthetic sh-miR30 into HEK293 cells. Dual-luciferase assays were performed 36 hr posttransfection. RL-luciferase activities were normalized with FF-luciferase, and the percentage of relative enzyme activity compared to the negative control (synthetic shRNA with a scrambled stem sequence) was plotted. Error bars represent the SD from two independent experiments, each performed in triplicate transfections.

See also Figure S4 for the repeat experiments performed in MEF cells.

at the stem had little impact on the position and/or relative abundance of noncanonical cleavages (compare sh-miR30 with sh-miR30-M, sh-miR30-24B, and sh-miR30-23B), the pattern of noncanonical cleavage can be substantially altered when the distance between the site of the canonical cut and loop vary (such as sh-miR30 versus sh-miR30-22). Particularly, we found that the noncanonical cleavages were almost completely abrogated when the canonical cut was 2 nt from the loop (sh-miR30-21) (Figure 5B). This result was further confirmed when we examined the small RNAs associated with Ago2. More than 98% of the reads mapping to the guide strand (3′ arm) of sh-miR30-21 started at the position of Dicer's canonical cleavage (Figure S5A). As expected, the bioactivity measurements were corroborative as the homogeneous Dicer processing of sh-miR30-21 resulted in reductions of the passenger strand mediated off-target effects in both human and mouse cells (Figures 2D, S2D, and S2E; compare to sh-miR30).

By moving the loop to the location 2 nt away from the expected position of canonical cleavage, we observed near-homogenous Dicer processing when additional shRNAs, which were previously shown to be heterogeneously processed, were tested (Figure S5B). Taken together, our results indicate that the accuracy of Dicer cleavage is determined by the distance between the site of cleavage and upstream loop structure, which we refer to as a loop-counting rule.

Loop-Counting Rule Applies to Endogenous miRNAs

To evaluate the biological relevance of our findings, we sought to investigate Dicer processing of endogenous miRNAs. In contrast to the overexpressed Pol-III-driven shRNAs, the majority of the pre-miRNAs are generated by Pol II transcription. Thus, we elected to study the Dicer processing of endogenous pre-miRNAs directly by analyzing the 5′ end of the 3p miRNAs (or miRNA*) that are generated by Dicer cleavage.

In contrast to shRNAs, natural pre-miRNAs have more complex structures, making it difficult to establish the existence of a stem-bulge, internal, and/or terminal loop. For example, the RNA structure of the pre-hsa-let-7a miRNA can be drawn with either multiple internal bulges or one large terminal loop. Nonetheless, we found that the distance between the start position of the 3p miRNA/miRNA*, indicative of Dicer cleavage, and the position of the most adjacent noncomplementary region (bulge, terminal, or internal loop) upstream was not randomly distributed. Rather, 314 out of 970 (32.4%) miRNAs in human and 209 out of 624 (33.5%) miRNAs in mouse share a structure contained within a noncomplementary region at a 2 nt distance upstream the Dicer cleavage site (Figure 6A). This indicates that evolutionary selection may be operative to maintain the relative position of the loop/bulge structure of the pre-miRNA in order to achieve accurate miRNA biogenesis as predicted by the loop-counting rule.

In addition to the nonrandom position of the loop in mammalian miRNAs, we predicted, based on our loop-counting rule, that these miRNAs would result in precise Dicer cleavage. Therefore, we sought to measure the accuracy of Dicer cleavage of miRNAs in vivo and to ask whether it correlated with the relative position of nearby loop/bulge structures. To do this, we used a well-documented sequencing result consisting of more than 60 million mouse small RNAs from various tissues and developmental stages (Chiang et al., 2010). The variation in the 5′ end start position was calculated (see Experimental Procedures for detail) for each individual miRNA (or miRNA*). Indeed, the most precise Dicer cleavage, which was inferred by the least variation at the 5′ ends of the 3p miRNA, was observed when such cleavage was 2 nt away from a loop/bulge structure (Figure 6B). In contrast, the variation at the 5′ ends of 5p miRNA, which was created by Drosha, did not follow the same pattern (Figure 6C), indicating that the precision at the 5′ end was the result of Dicer processing. Taken together, these results demonstrate that the

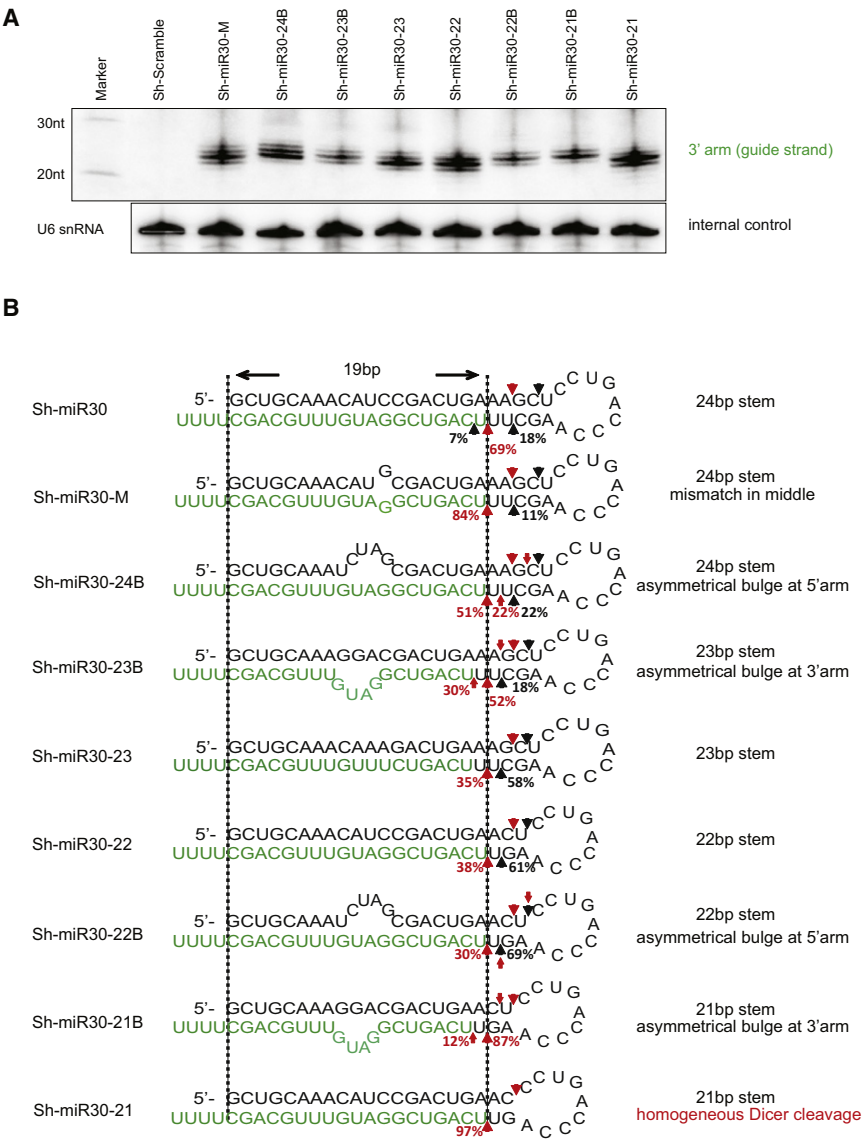

Figure 5. The Loop-Counting Rule: Noncanonical Dicer Cleavages Were Suppressed when the Canonical Cut Was 2 nt away from the Loop

(A) Plasmids expressing a set of sh-miR30 variants with different stem and/or loop sequences were transfected into HEK293 cells separately. Extracted RNAs were run on 20% polyacrylamide 7 M urea denaturing gels 36 hr posttransfection. Small RNAs processed from the 3′ arm were identified by Northern blotting. Endogenous U6 snRNA was also detected as an internal control.

(B) The processing of sh-miR30 variants was analyzed by deep sequencing. Results were presented as described in Figure 1. Although the canonical cleavage was determined by the 5′ counting rule, the noncanonical cuts were affected by the distance between the site of canonical cut and loop/bulge structure.

See also Figure S5.

loop-counting rule obtained with artificial shRNAs is part of the natural role that Dicer plays in generating endogenous miRNAs.

Generating Potent Anti-HCV shRNAs with Minimal Passenger-Strand-Mediated Off-Target Effects

On the basis of our findings, shRNA with a 21-bp-long stem loop should be precisely processed by Dicer and generate fewer RNA species capable of generating off-target effects from the passenger strand. To validate this principle in a relevant preclinical setting, we created 11 shRNAs that varied in stem length from 19 to 29 bp, which all produced the same guide strand targeting the 5NSB region of HCV (sh-HCV-19 to sh-HCV-29). Consistent with our prediction, noncanonical Dicer cuts were observed with all other shRNAs with the exception of the one with a 21 bp stem (sh-HCV-21) (Figures 7A and S6A). Consistent with what we learned from the processing of pre-miRNAs, introducing an internal loop at 2 nt away from the expected site of Dicer canonical cut into sh-HCV-29 (sh-HCV-29iB) shifted the pattern of Dicer cleavage to what was obtained with sh-HCV-21, indicating that the precision of Dicer cleavage was not specific to the stem length or terminal loop but was a result of an optimal distance to an upstream noncomplementary structure.

Despite the difference in stem length or structure, all anti-HCV shRNAs had relatively potent on-target activity (repression from guide-strand target), with the relative amount of knockdown correlating with the abundance of a mature guide strand (Figures 7B and 7C). In contrast, the passenger-strand-mediated off-target activity was not always coupled with the level of passenger strands. Rather, it paralleled with the heterogeneity of Dicer processing with the exception resulting from sh-HCV-19 and sh-HCV-20, which seemed to be poorly processed by Dicer and generated little passenger strand in the first place (Figure 7B). Similar results were also obtained in mouse cells (Figure S6). Together, our results demonstrated that potent shRNAs with improved safety could be achieved by applying the loop-counting rule and by increasing the accuracy of Dicer processing.

DISCUSSION

We performed high-throughput small RNA sequencing with 30 designed shRNAs to establish differential sites of Dicer cleavage. In contrast to Northern blot analysis, which is unable to distinguish similar but heterogeneous sequences of the same length, deep sequencing provides a more complete delineation of the cleavage products. These observations allowed us to make accurate predictions on how altering stem length and

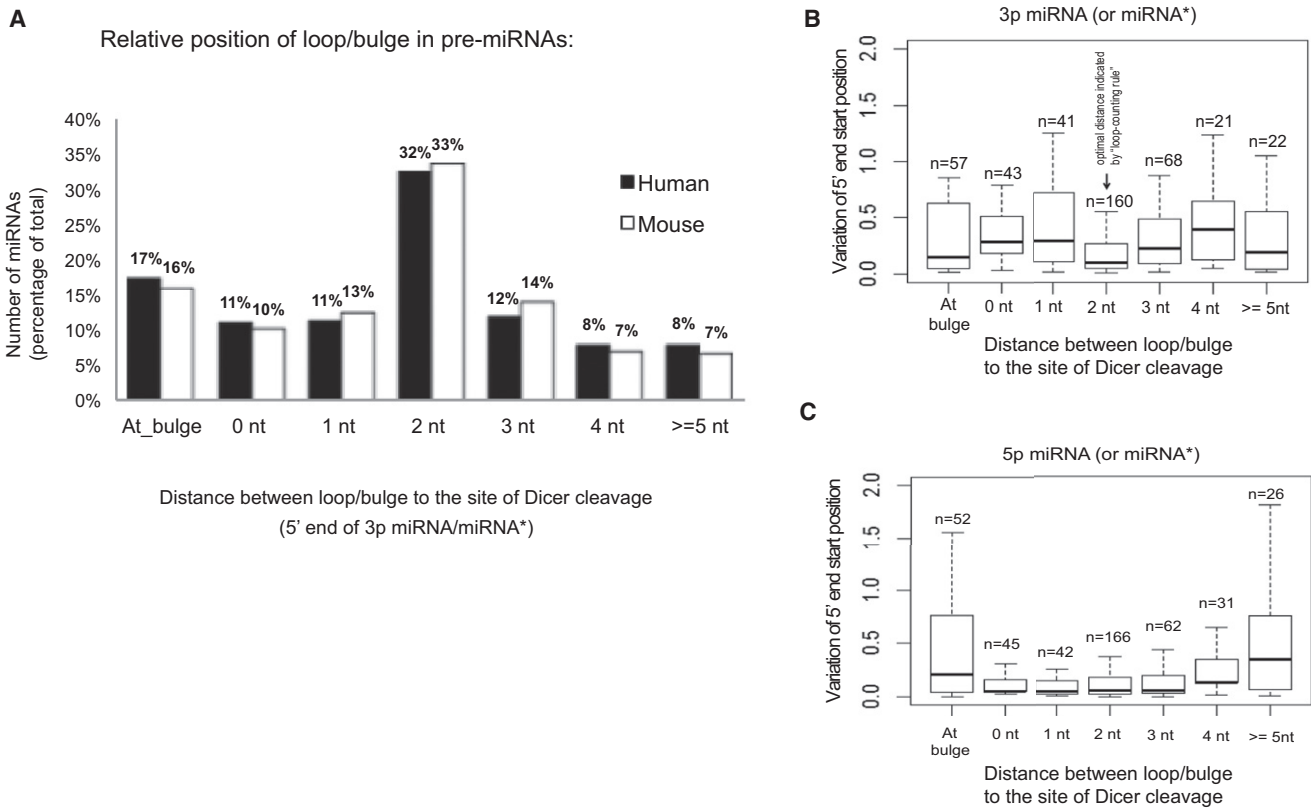

Figure 6. Loop-Counting Rule Is Applicable to the Dicer's Processing of Endogenous miRNAs

(A) Distribution of the loop position in pre-miRNA structure. Pre-miRNAs were categorized based on the distance between the Dicer cleavage site (determined by the 5′ end of 3p miRNA or miRNA* sequence) and the nearest upstream terminal loop or internal loop (bulge). If a 3p miRNA/miRNA* starts inside a loop/bulge, it was counted as "At_bulge." The structure information of human and mouse miRNAs was obtained from miRBase release 17.

(B and C) Box plots show the distribution of start position variation for the 5′ ends of miRNA or miRNA* sequence based on the sequencing results from a previous publication (Chiang et al., 2010). Results were categorized based on the pre-miRNA structure as described in (A). MiRNAs originating from the 5′ arm (5p miRNA) or 3′ arm (3p miRNA) were analyzed separately. For each miRNA, the variation was calculated as the weighted mean of the absolute values of the distance between the expected end and actual end of individual reads (see Experimental Procedures for details). Box represents the first and third quartiles, and the internal bar indicates the median. Whiskers denote the lowest and highest values within 1.5× interquartile range of the first and third quartiles, respectively. n represents the number of pre-miRNAs in the particular category (pre-miRNAs generating at least 50 reads from 3′ arm in sequencing results were counted in [B], whereas those generating at least 50 reads from 5′ arm were counted in [C]).

bulge position affect Dicer cleavage patterns. The consistency of our experimental results between the myriad of shRNAs tested in both mouse and human cells confirmed our ability to design shRNAs that resulted in the creation of a homogenous population of guide-strand RNAs.

Furthermore, our finding that these artificial shRNA cleavage guidelines paralleled endogenous mammalian miRNA processing led us to propose the loop-counting rule as a critical mechanism for directing bona fide small RNA processing in mammalian cells. The loop-counting rule is as follows: Dicer cleaves precisely when it is able to recognize a single-stranded RNA sequence either from the loop region or internal bulge at a fixed distance (two nucleotides) relative to the site of cleavage. Otherwise, Dicer cleavage is not precise, leading to a range of Dicer cleavage products with variable 5′ start positions.

Our observations raised intriguing questions regarding the mechanism of how Dicer determines its cleavage site. Interestingly, the helicase domain of *Drosophila* Dicer-1 was shown to be responsible for recognizing the single-stranded loop region of pre-miRNA (Tsutsumi et al., 2011). Furthermore, the recent elucidation of the human Dicer 3D structure indicated that the helicase domain was physically adjacent to the RNase III domain, the catalytic center for Dicer cleavage (Lau et al., 2012; Sawh and Duchaine, 2012). In light of these findings and the data presented here, we propose that the cleavage site is defined in a stepwise process. First, Dicer docks the open ends of hairpin RNA and feeds the region to be cleaved into the catalytic core as previously elucidated (Macrae et al., 2006; Park et al., 2011). Then, the cleavage site is secured by a second contact between Dicer and its substrate RNA. The precise cleavage is achieved when an ssRNA region (loop/bulge) is available at the optimal position where the helicase domain can grab on to and in turn stabilize the catalytic center (Figure 7D). To test this idea, we deep sequenced the miRNAs in a Dicer mutant HCT cell line in which the helicase domain was disrupted with a 43 amino acid in-frame insertion (Cummins et al., 2006). There

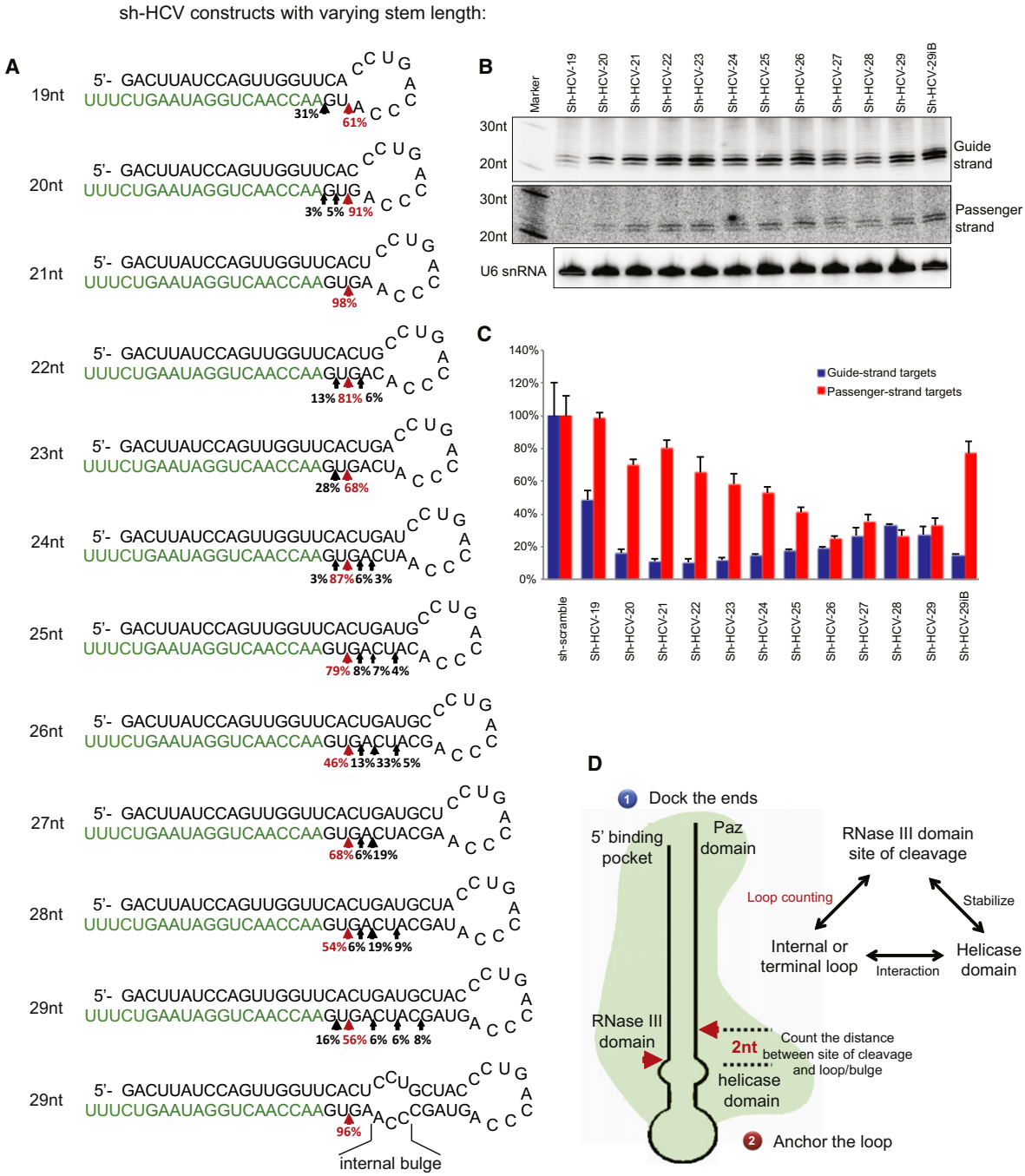

Figure 7. Generating Potent Anti-HCV shRNAs with Minimal Passenger-Strand-Mediated Off-Target Effects

(A) A set of anti-HCV shRNAs generating the same guide-strand sequences but that were varied in length were expressed in HEK293 cells. The Dicer's processing of these shRNAs was analyzed by deep sequencing. Results were presented as described in Figure 1. As predicted by the loop-counting rule, precise Dicer cleavages were observed only with sh-HCV-21 and sh-HCV-29iB.

(B) Expression of these shRNAs was also analyzed by Northern blotting 36 hr posttransfection. Extracted RNAs were run on 20% polyacrylamide 7 M urea denaturing gels. Guide strand (processed from the 3′ arm) or passenger strand (processed from the 5′ arm) were identified with corresponding probes. Endogenous U6 snRNA was also detected as an internal control.

(C) PsiCHECK vectors with a target sequence in the 3′ UTR, which was complementary to the expected guide or passenger strand of sh-HCV sets, were used to monitor the repression activity from either strands. Each shRNA was cotransfected with the psiCHECK reporters into HEK293 cells. Dual-luciferase assays were performed 36 hr posttransfection. RL-luciferase activities were normalized with FF-luciferase, and the percentage of relative enzyme activity compared to the negative control (treated with sh-scramble) was plotted. Error bars represent the SD from two independent experiments, each performed in triplicate transfections. See also Figure S6 for repeating experiments performed in MEF cells.

was a trend for less precise pre-miRNA processing in mutant versus wild-type cells (Figure S7A). Interestingly, only pre-miRNAs containing the optimal 2 nt distance between Dicer cleavage site and an upstream single-stranded region, but not the rest of pre-miRNAs, showed a statistically significant ($p = 0.005$) reduction in the precision of Dicer cleavage in mutant versus wild-type cells (Figures S7B and S7C). This result supports a role for the helicase domain in sensing the region of noncomplementarity within the pre-miRNA. Alternatively, Dicer cofactors, such as TRBP, may be responsible for recognizing the loop/bulge structure. However, such regulation is likely to be indirect, as the optimal loop position indicated by the loop-counting rule is too close to the catalytic center to be accessible by a protein other than Dicer.

This feed and clamp model not only renders a molecular explanation for the loop-counting rule but also provides insights into the requirement of the helicase domain of *Drosophila* Dicer-2 in processing blunt end double-stranded RNA (dsRNA) substrates (Welker et al., 2011). Given that blunt ends were poorly recognized by the Paz domain (MacRae et al., 2007), Dicer-2 may have to rely on its helicase domain for more efficient binding and hence more proficient cleavage. Another line of evidence supporting this model comes from the regulatory role of the helicase domain in Dicer processing. Because there is a tradeoff between precision and speed, our model is consistent with the observation that the absence of a helicase domain in human Dicer increased its catalytic activity (Ma et al., 2008). Interestingly, the Dicer helicase insertion mutation was found to increase the processivity of Dicer for long, stable hairpins but decreased its processivity for bulged hairpins, suggesting that the terminal loop and internal bulge/loop may be sensed differently by the helicase domain (Soifer et al., 2008).

Consistent with a previous report (Starega-Roslan et al., 2011), we found that asymmetrical stem bulges may twist the 3D folding of substrate shRNA and indirectly affect the position of Dicer cleavage. Therefore, the overall structure of substrate RNA, including the relative position of the loop/bulge to the cleavage site, has a much more profound impact on Dicer processing than previously believed. Further investigation into the structural conservation of pre-miRNAs based on the parameters provided in this study may finally provide an evolutionary meaning for the unique bulge-enriched structure of miRNA precursors.

The studies presented here also have great implications for RNAi technology, particularly the design of shRNAs. Despite the successes, the application of shRNAs is hampered due to unwanted off-target effects (Jackson and Linsley, 2010; Kaelin, 2012). A major source of off-targeting is unintentional loading of the passenger strand into RISC. To increase the odds of RISC loading guide strands over passenger strands, we avoided placing the guide strand in the 5p arm for two reasons: (1) the 5p small RNA generated from Pol-III-driven shRNA carries a triphosphate group at the 5′ end, which may interfere with its incorporation into Ago2/RISC; and (2) 5p transcripts starting with a guanine (G) or adenine (A) are required for efficient Pol III transcription, making them structurally unfavorable for Ago2/RISC loading. In addition, the 3p guide strand was designed to start with a uracil (U), which not only is the most preferred nucleotide for Ago2 association but also lowers the 5′ end thermodynamic stability, further enhancing its preferential loading into RISC (Frank et al., 2010; Seitz et al., 2011). Indeed, we found this design to be effective when the shRNAs were processed as expected (sh-miR30-21, sh-HCV-21, and shHCV-29iB). However, we found that the additional products generated from Dicer's noncanonical cleavage, even when present in relatively small amounts, could induce robust passenger-strand-mediated off-target effects, highlighting the importance of reducing heterogeneous processing in an shRNA design. Furthermore, the generation of shRNAs with differing seed region sequences can result in additional guide-strand-mediated off-target effects through seed region base paring. Therefore, longer hairpins, especially those expressed shRNAs where chemical modification is unavailable, would suffer from inaccurate processing and should be used with caution. Consistent with this idea, shRNAs with longer stems were generally more toxic when overexpressed in mouse liver (Grimm et al., 2006). More importantly, we have experimentally demonstrated the implementation of the loop-counting rule in designing shRNAs free of heterogeneous processing. Overall, our results provide additional guidelines of how to design potent si/shRNAs with minimal off-target effects for biological knockdown of important genes and/or treatment of diseases.

EXPERIMENTAL PROCEDURES

Plasmid Vectors and Construction

For cloning of all psiCHECK reporter systems with various targets, both strands of insert were chemically synthesized, annealed, purified, and inserted between the XhoI and SpeI sites in the psiCHECK2 vector (Promega). All of the oligo sequences to generate these reporter plasmids can be found in Table S1. A similar approach was used to generate all of the sh-RNA-expressing plasmids. shRNA sequences (as detailed in the figures) were directly cloned downstream of U6 Pol III promoter between BglII and KpnI. Plasmids expressing Flag-tagged human Argonaute2 (Ago2) and human Dicer were obtained from Addgene (http://www.addgene.org).

Cell Culture and Transfection

HEK293 and MEF cells were grown in Dulbecco's modified Eagle's medium (DMEM; GIBCO-BRL) with L-glutamine, nonessential amino acids, sodium pyruvate, and 10% heat-inactivated fetal bovine serum with antibiotics. All transfection assays were done by using Lipofectamine 2000 (Invitrogen) following the manufacturer's protocol.

Dual-Luciferase Reporter Assay

One hundred ng of psiCHECK reporter plasmids were cotransfected with either 10 ng shRNA plasmids or certain amount of synthetic siRNA/shRNA (to a final concentration of 30 nM) into HEK293 or MEF cells in 24 well plates. FF-luciferase and RL-luciferase activities were measured 36 hr posttransfection by using Promega's dual-luciferase kit (cat E1980) protocol and detected by a Modulus Microplate Luminometer (Turner BioSystems).

(D) A proposed model of Dicer cleavage and the loop-counting rule. First, Dicer docks the ends of substrate shRNA/per-miRNA and measures a fixed distance to determine where to cleave. Then, the accuracy of a cleavage is determined by its distance relative to the nearby bulge/loop structure. When such distance is optimal (2 nt), Dicer secures the catalytic center by a contact between the loop/bulge and the helicase domain to achieve precise processing.
See also Figures S6 and S7.

Northern Blots

HEK293 cells in 10 cm dishes were transfected with 5 μg of sh-RNA-expressing plasmids. Total RNA was isolated 36 hr posttransfection by using Trizol (Invitrogen) and was then electrophoresed on 20% (w/v) acrylamide/7 M urea gel. After transfer onto a Hybond-N1 membrane (Amersham Pharmacia Biotech), small RNAs were detected by using P32-labeled probes (see Table S1 for sequences).

Ago Coimmunoprecipitation of Small RNAs

IP experiments were performed in a slightly modified protocol as described previously (Gu et al., 2011). In brief, HEK293 cells in 10 cm dishes were cotransfected with 0.5 μg of plasmids expressing Flag-tagged Ago2 and 4.5 μg of plasmids expressing various shRNAs. Cells were lysed 36 hr posttransfection and incubated with anti-Flag M2 agarose beads (Sigma A2220) overnight at 4°C. After three washes with cold IP buffer, Flag-Ago-RNA complexes were eluted with 100 μg/ml 3× Flag peptide (Sigma F4799) in Tris-buffered saline (TBS). RNAs associated with Agos were extracted by Trizol and subject to small RNA deep sequencing.

Small RNA Deep Sequencing and Data Analysis

Small RNA libraries were created by using a protocol similar to previous small RNA capture procedures (Lau et al., 2001; Maniar and Fire, 2011). Sequencing reads (36 nt) for all libraries were generated by using the Illumina Genome Analyzer II (Stanford Functional Genomics Facility). After removing low quality reads, all sequences were sorted based on the 5′ bar codes (four nucleotides). Further, reads without 3′ adaptor sequences or shorter than 18 nt were dropped. After removing the 3′ linker and 5′ barcode sequences, the resulting reads were aligned to either the 5′ arm or 3′ arm of shRNA sequences with up to three mismatches by bowtie (version 0.12.7) (Langmead et al., 2009) without allowing mapping to the reverse-complement reference strand (command "norc").

Calculating Variation on the 5′ End of miRNA or miRNA* Sequences

All reads were first aligned to human miRNA library sequences (miRBase; Kozomara and Griffiths-Jones, 2011) by bowtie (Langmead et al., 2009). For each particular miRNA or for particular miRNA* sequences, reads with a 5′ end within 4 n distance to the expected position were considered as small RNA generated from such loci and were taken into calculation in the next step. With very few exceptions, the expected 5′ end as indicated in miRBase was also the most abundant 5′ end for that miRNA/miRNA* measured in deep sequencing results. The variation of each particular read was calculated as the absolute value of distance between its 5′ end and the expected end. The variation for a particular miRNA/miRNA* was then calculated by averaging the individual variation and by using the relative abundance as weight.

ACCESSION NUMBERS

The NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) accession number for the sequence reported in this paper is GSE41292.

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and one table and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2012.09.042.

ACKNOWLEDGMENTS

This work was supported by NIH DK 078424 (M.A.K.) and NIH AI071068 (M.A.K.). We thank Dr. Grace Zheng for the Dicer KO ES cells and Dr. George Mias for helpful discussions on statistical methods.

Received: June 29, 2012
Revised: August 20, 2012
Accepted: September 27, 2012
Published: November 8, 2012

REFERENCES

Ameres, S.L., Horwich, M.D., Hung, J.H., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P.D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science *328*, 1534–1539.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281–297.

Brummelkamp, T.R., Bernards, R., and Agami, R. (2002). A system for stable expression of short interfering RNAs in mammalian cells. Science *296*, 550–553.

Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y., and Daub, C.O. (2010). A comprehensive survey of 3′ animal miRNA modification events and a possible role for 3′ adenylation in modulating miRNA targeting effectiveness. Genome Res. *20*, 1398–1410.

Calabrese, J.M., Seila, A.C., Yeo, G.W., and Sharp, P.A. (2007). RNA sequence analysis defines Dicer's role in mouse embryonic stem cells. Proc. Natl. Acad. Sci. USA *104*, 18097–18102.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and mechanisms of miRNAs and siRNAs. Cell *136*, 642–655.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740–744.

Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E., et al. (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev. *24*, 992–1009.

Cummins, J.M., He, Y., Leary, R.J., Pagliarini, R., Diaz, L.A., Jr., Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A.E., Labourier, E., et al. (2006). The colorectal microRNAome. Proc. Natl. Acad. Sci. USA *103*, 3687–3692.

Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature *411*, 494–498.

Frank, F., Sonenberg, N., and Nagar, B. (2010). Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. Nature *465*, 818–822.

Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. Genome Res. *19*, 92–105.

Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F., and Kay, M.A. (2006). Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. Nature *441*, 537–541.

Gu, S., and Kay, M.A. (2010). How do miRNAs mediate translational repression? Silence *1*, 11.

Gu, S., Jin, L., Zhang, F., Huang, Y., Grimm, D., Rossi, J.J., and Kay, M.A. (2011). Thermodynamic stability of small hairpin RNAs highly influences the loading process of different mammalian Argonautes. Proc. Natl. Acad. Sci. USA *108*, 9208–9213.

Gurtan, A.M., Lu, V., Bhutkar, A., and Sharp, P.A. (2012). In vivo structure-function analysis of human Dicer reveals directional processing of precursor miRNAs. RNA *18*, 1116–1122.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. Science *293*, 1146–1150.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat. Rev. Genet. *12*, 99–110.

Jackson, A.L., and Linsley, P.S. (2010). Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. Nat. Rev. Drug Discov. *9*, 57–67.

Kaelin, W.G., Jr. (2012). Molecular biology. Use and abuse of RNAi to study mammalian gene function. Science *337*, 421–422.

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. Cell *115*, 209–216.

Kim, D.H., and Rossi, J.J. (2007). Strategies for silencing human disease using RNA interference. Nat. Rev. Genet. *8*, 173–184.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell Biol. *10*, 126–139.

Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. *39*(Database issue), D152–D157.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858–862.

Lau, P.W., Guiley, K.Z., De, N., Potter, C.S., Carragher, B., and MacRae, I.J. (2012). The molecular architecture of human Dicer. Nat. Struct. Mol. Biol. *19*, 436–440.

Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787–798.

Liu, Q., and Paroo, Z. (2010). Biochemical principles of small RNA pathways. Annu. Rev. Biochem. *79*, 295–319.

Ma, E., MacRae, I.J., Kirsch, J.F., and Doudna, J.A. (2008). Autoinhibition of human dicer by its internal helicase domain. J. Mol. Biol. *380*, 237–243.

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. Science *311*, 195–198.

MacRae, I.J., Zhou, K., and Doudna, J.A. (2007). Structural determinants of RNA recognition and cleavage by Dicer. Nat. Struct. Mol. Biol. *14*, 934–940.

Maniar, J.M., and Fire, A.Z. (2011). EGO-1, a C. elegans RdRP, modulates gene expression via production of mRNA-templated short antisense RNAs. Curr. Biol. *21*, 449–459.

McCaffrey, A.P., Meuse, L., Pham, T.T., Conklin, D.S., Hannon, G.J., and Kay, M.A. (2002). RNA interference in adult mice. Nature *418*, 38–39.

McIntyre, G.J., Yu, Y.H., Lomas, M., and Fanning, G.C. (2011). The effects of stem length and core placement on shRNA activity. BMC Mol. Biol. *12*, 34.

McManus, M.T., Petersen, C.P., Haines, B.B., Chen, J., and Sharp, P.A. (2002). Gene silencing using micro-RNA designed hairpins. RNA *8*, 842–850.

Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J., and Conklin, D.S. (2002). Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. Genes Dev. *16*, 948–958.

Paddison, P.J., Silva, J.M., Conklin, D.S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., et al. (2004). A resource for large-scale RNA-interference-based screens in mammals. Nature *428*, 427–431.

Pan, Q., van der Laan, L.J., Janssen, H.L., and Peppelenbosch, M.P. (2012). A dynamic perspective of RNAi library development. Trends Biotechnol. *30*, 206–215.

Park, J.-E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5′ end of RNA for efficient and accurate processing. Nature *475*, 201–205.

Sawh, A.N., and Duchaine, T.F. (2012). Turning Dicer on its head. Nat. Struct. Mol. Biol. *19*, 365–366.

Schwarz, D.S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell *115*, 199–208.

Seitz, H., Ghildiyal, M., and Zamore, P.D. (2008). Argonaute loading improves the 5′ precision of both MicroRNAs and their miRNA* strands in flies. Curr. Biol. *18*, 147–151.

Seitz, H., Tushir, J.S., and Zamore, P.D. (2011). A 5′-uridine amplifies miRNA/miRNA* asymmetry in Drosophila by promoting RNA-induced silencing complex formation. Silence *2*, 4.

Siolas, D., Lerner, C., Burchard, J., Ge, W., Linsley, P.S., Paddison, P.J., Hannon, G.J., and Cleary, M.A. (2005). Synthetic shRNAs as potent RNAi triggers. Nat. Biotechnol. *23*, 227–231.

Soifer, H.S., Sano, M., Sakurai, K., Chomchan, P., Saetrom, P., Sherman, M.A., Collingwood, M.A., Behlke, M.A., and Rossi, J.J. (2008). A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. Nucleic Acids Res. *36*, 6511–6522.

Starega-Roslan, J., Krol, J., Koscianska, E., Kozlowski, P., Szlachcic, W.J., Sobczak, K., and Krzyzosiak, W.J. (2011). Structural basis of microRNA length variety. Nucleic Acids Res. *39*, 257–268.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., and Tomari, Y. (2011). Recognition of the pre-miRNA structure by Drosophila Dicer-1. Nat. Struct. Mol. Biol. *18*, 1153–1158.

Welker, N.C., Maity, T.S., Ye, X., Aruscavage, P.J., Krauchuk, A.A., Liu, Q., and Bass, B.L. (2011). Dicer's helicase domain discriminates dsRNA termini to promote an altered reaction mode. Mol. Cell *41*, 589–599.

Zeng, Y., Wagner, E.J., and Cullen, B.R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol. Cell *9*, 1327–1333.

Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J. *21*, 5875–5885.

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004). Single processing center models for human Dicer and bacterial RNase III. Cell *118*, 57–68.