Cell

# Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex

Jinju Han,[1,4] Yoontae Lee,[1,4] Kyu-Hyeon Yeom,[1,4] Jin-Wu Nam,[2] Inha Heo,[1] Je-Keun Rhee,[2] Sun Young Sohn,[3] Yunje Cho,[3] Byoung-Tak Zhang,[2] and V. Narry Kim[1,*]
[1] School of Biological Sciences and Research Center for Functional Cellulomics
[2] School of Computer Science and Engineering
Seoul National University, Seoul
[3] National Creative Research Center for Structural Biology and Department of Life Science, Pohang University of Science and Technology, Pohang, KyungBook, South Korea
[4] These authors contributed equally to this work.
*Contact: narrykim@snu.ac.kr
DOI 10.1016/j.cell.2006.03.043

## SUMMARY

The Drosha-DGCR8 complex initiates microRNA maturation by precise cleavage of the stem loops that are embedded in primary transcripts (pri-miRNAs). Here we propose a model for this process that is based upon evidence from both computational and biochemical analyses. A typical metazoan pri-miRNA consists of a stem of ~33 bp, with a terminal loop and flanking segments. The terminal loop is unessential, whereas the flanking ssRNA segments are critical for processing. The cleavage site is determined mainly by the distance (~11 bp) from the stem-ssRNA junction. Purified DGCR8, but not Drosha, interacts with pri-miRNAs both directly and specifically, and the flanking ssRNA segments are vital for this binding to occur. Thus, DGCR8 may function as the molecular anchor that measures the distance from the dsRNA-ssRNA junction. Our current study thus facilitates the prediction of novel microRNAs and will assist in the rational design of small hairpin RNAs for RNA interference.

## INTRODUCTION

MicroRNAs (miRNAs) are single-stranded RNA molecules of 19–25 nt in length that are generated from endogenous hairpin-shaped transcripts (Bartel, 2004; Kim, 2005b). MiRNAs act as posttranscriptional gene suppressors by base-pairing with their target mRNAs and inducing either translational repression or mRNA destabilization. Genetic, biochemical, and computational studies have implicated essential and diverse roles of miRNAs in multicellular organisms. A strong link between miRNA and cancer has recently been demonstrated, opening up a new area of investigation in the field of cancer biology (reviewed by Croce and Calin, 2005). The expression of miRNAs dramatically changes during development and cell differentiation. MiRNA profiling has been shown to faithfully reflect both developmental lineages and disease states (Lu et al., 2005). In order to further dissect the regulatory networks in which miRNAs function, it will be crucial to first understand how these molecules are generated and controlled (Kim, 2005a).

MiRNA biogenesis is initiated via transcription by RNA polymerase II (Cai et al., 2004; Kim, 2005a; Lee et al., 2002, 2004). The primary transcripts (pri-miRNAs) are usually species of over several kilobases long and contain both a 5′ cap and a poly(A) tail. pri-miRNAs are first cropped to release ~65 nt of hairpin-shaped precursor (pre-miRNAs) by a member of the ribonuclease III family (RNase III), Drosha (Lee et al., 2003). Drosha and its cofactor, DGCR8 (Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004) (known as Pasha in *D. melanogaster* and *C. elegans*; Denli et al., 2004), form a complex called "Microprocessor." Neither recombinant DGCR8 nor Drosha alone is active in pri-miRNA processing, whereas combining these two proteins restores this activity, indicating that both proteins play essential roles in pri-miRNA processing (Gregory et al., 2004; Han et al., 2004). pri-miRNA processing is a critical step in miRNA biogenesis because it defines the miRNA sequences embedded in long pri-miRNAs by generating one end of the molecule. Following this initial processing, the resulting pre-miRNAs are exported by the nuclear transport factor, exportin-5 (Exp5) (Bohnsack et al., 2004; Lund et al., 2004; Yi et al., 2003). In the cytoplasm, Dicer, a cytoplasmic RNase III type protein, dices the transported pre-miRNAs to generate ~22 nt miRNA duplexes (Grishok et al., 2001; Hutvagner et al., 2001; Ketting et al., 2001). One strand of the Dicer product remains as a mature miRNA and is then assembled into the effector complex called miRNP

or miRNA-induced silencing complex (miRISC) (Khvorova et al., 2003; Schwarz et al., 2003).

RNA interference (RNAi), the gene silencing mechanism that is mediated by small RNAs, is now a powerful genetic tool in mammalian systems. Effective and stable gene knockdown can be achieved by the expression of short hairpin RNAs (shRNAs) which are processed into small interfering RNAs (siRNAs). Recent breakthroughs in RNAi technology have been made by generating shRNA expression cassettes that can mimic a natural miRNA gene (Dickins et al., 2005; Silva et al., 2005; Zeng et al., 2002). MiRNA-based shRNAs, driven by RNA polymerase II promoters, can induce efficient, stable, and regulated silencing in cultured cells as well as in animal models. The expression of such shRNAs is dependent upon the presence of miRNA biogenesis factors. Therefore, a mechanistic understanding of miRNA processing is crucial for the rational design of accurate and efficient shRNAs.

RNase III type proteins have major functions in RNA silencing pathways. RNase III cleaves double-stranded RNA (dsRNA) in a staggered manner and creates a 2 nt overhang on the 3′ end of its products (Kim, 2005a; Tomari and Zamore, 2005). This enzyme family can be grouped into three classes based on their domain organization. Class I proteins are found in yeast and bacteria and have an RNase III domain (RIIID) and a dsRBD. Drosha homologs belong to the class II grouping and possess tandem RIIIDs and a dsRBD, in addition to an extended N terminus that contains a proline-rich region and a serine/arginine-rich region of unknown function. Dicer homologs, which are class III proteins within this enzyme family, contain two RIIIDs, one dsRBD, and a long N terminus. The N-terminal region of Dicer is composed of an RNA helicase/ATPase domain, a DUF283 domain, and a PAZ domain.

RNase III proteins show a high degree of conservation in their catalytic domains and also share a basic action mechanism. Two RIIIDs interact with each other to constitute a single processing center where two catalytic sites are placed closely and each of the two catalytic sites cleaves one strand of an RNA duplex (Blaszczyk et al., 2001; Zhang et al., 2004). Both the Dicer and Drosha enzymes form an intramolecular dimer of two RIIIDs (Han et al., 2004; Zhang et al., 2004). The C-terminal RIIID (RIIIDb), proximal to the dsRBD, cleaves the 5′-strand of the hairpin, whereas the other RIIID (RIIIDa) cleaves the 3′-strand.

Despite the similarities in their basic modes of action, RNase III proteins are different in many interesting ways, particularly in their substrate specificities. Dicer will act on any dsRNA with a simple preference toward the terminus of the molecule and produce ~22 nt fragments progressively from the terminus (Zhang et al., 2002). The PAZ domain of Dicer may interact with the 3′ overhang at the terminus and determines the processing site in a ruler-like fashion that measures ~22 nt segments away from the terminus (Lingel et al., 2004; Ma et al., 2004; Macrae et al., 2006; Song et al., 2003; Vermeulen et al., 2005; Yan et al., 2003).

Less is known about how Drosha recognizes its substrates. Although Drosha is the only known enzyme that can process a variety of pri-miRNAs, no common sequence motif has been found among the human pri-miRNA species. Thus, it is plausible that the Drosha-DGCR8 complex recognizes common structural feature(s) in these molecules. A typical animal pri-miRNA comprises a stem, a terminal loop, and long flanking sequences. Zeng et al. have shown previously that a large terminal loop is in fact critical for processing and that the cleavage site may be determined largely by the distance (~2 helical turns) from this structure (Zeng et al., 2005). The sequences flanking the stem loop have also been shown to be important for efficient processing in vitro (Lee et al., 2003; Zeng et al., 2005), as well as in vivo (Chen et al., 2004a). It remains to be determined which protein is responsible for specific recognition of pri-miRNAs.

With the aim of elucidating the molecular basis for pri-miRNA processing in our current study, we first analyzed the thermodynamic features of pri-miRNAs and carried out a series of systematic mutagenesis experiments with these molecules. We also demonstrate that DGCR8 is capable of recognizing the structural features of the pri-miRNAs both specifically and directly. Based on our results, we propose a new model for substrate recognition and processing by the Drosha-DGCR8 complex.

## RESULTS

### Thermodynamic Stability Profiling of pri-miRNAs

For insight into the common structural features of pri-miRNAs, we performed computational analysis to deduce the general structure of these RNA molecules. This is somewhat analogous to the procedures used to compute an "average face," which typically involve (1) compartmentalizing the face into defined parts, (2) quantifying the features of each part, (3) averaging the values obtained from individual faces, and (4) re-assembly of these averaged features into a face.

We predicted the secondary structures of 321 human and 68 fly pri-miRNAs (miRBase release 7.0) using mfold program version 3.1 (Zuker, 2003) (http://www.bioinfo.rpi.edu/~zukerm/rna/mfold-3.html). We then selected 280 human and 55 fly pri-miRNAs (Table S1), excluding those that are predicted to have multiloops because it is difficult to assign position numbers to the bases when they are on the branched structure. The human data comprise 157 pri-miRNAs harboring mature miRNA sequences in their 5′-strands (5′-donors) and 123 pri-miRNAs containing miRNA sequences in their 3′-strands (3′-donors). The fly data consist of 21 5′-donors and 34 3′-donors. Based on our secondary structure prediction, we assigned position numbers to each base pair. The 5′ end of a miRNA from a 5′-donor is given the +1 position (Figures 1A and 1B), whereas the 3′ end of a miRNA from a 3′-donor is placed at the −2 position (Figures 1C and 1D). We based these designations on the following assumptions: (1) Drosha creates a 2 nt overhang for all

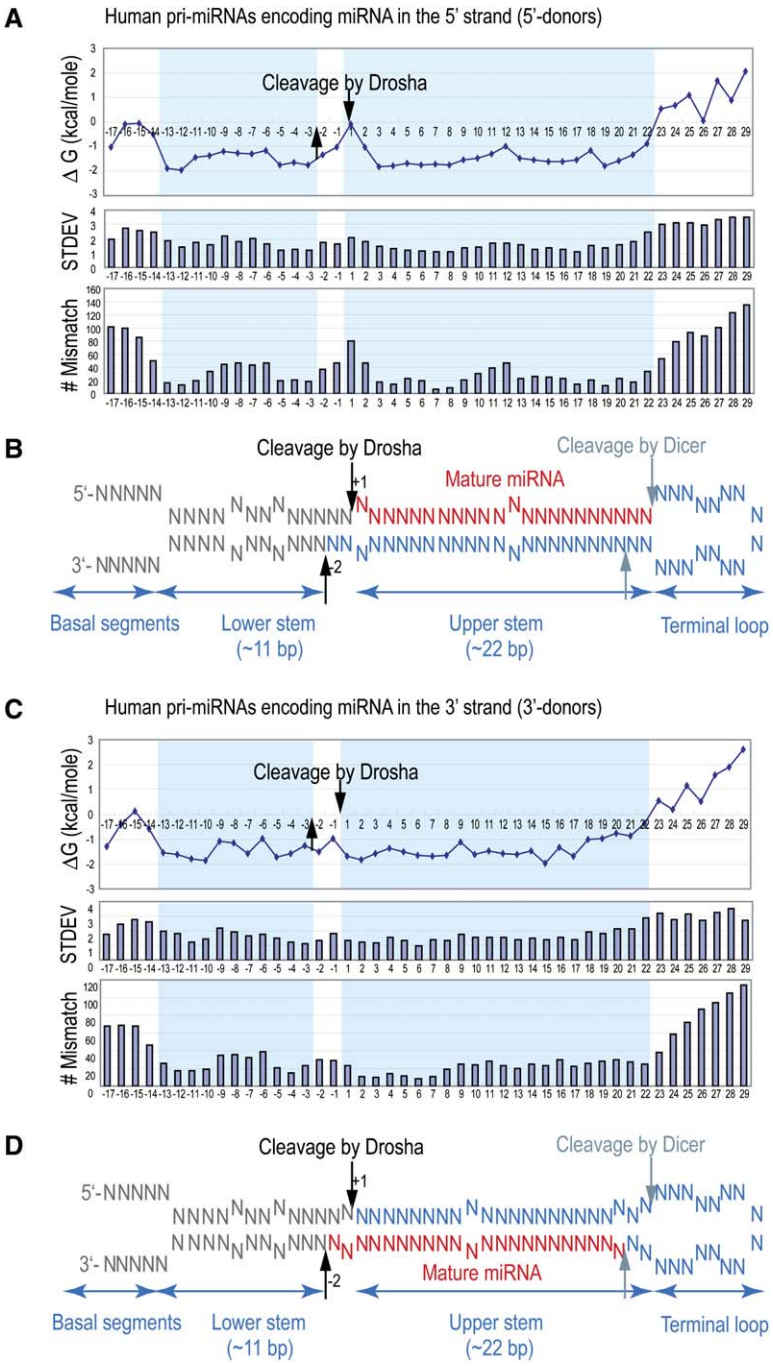


**Figure 1. Thermodynamic Stability Profiling of Human pri-miRNAs**

(A) Thermodynamic stability profiles of pri-miRNAs harboring miRNA sequences in the 5′ side of the hairpin (5′-donors). Average stabilities (ΔG in kcal/mole at each position) were calculated for 157 pri-miRNAs. The 5′-most nucleotide of mature miRNA was assigned the +1 position. The middle profile shows the standard deviation (in kcal/mole) at each position. The lowest profile presents the number of mismatches such as an internal loop or a bulge at a given position.

(B) The average face of pri-miRNA inferred from the thermodynamic profiling. A pri-miRNA can be divided into four parts: a terminal loop, the upper stem, the lower stem, and basal segments. Predicted sites for Drosha cleavage and Dicer cleavage are indicated with black arrows and gray arrows, respectively.

(C) Thermodynamic stability profiles of pri-miRNAs containing miRNA sequences in the 3′ side of the hairpin (3′-donors). Free energy of the 123 pri-miRNAs was calculated as in (A). The 3′-most nucleotide of mature miRNA is placed at the −2 position.

(D) The average face of pri-miRNA inferred from the thermodynamic profiling as in (B).

of its substrates and (2) no further modification of the ends takes place after Drosha processing (Basyuk et al., 2003; Lee et al., 2003). The thermodynamic stability at each position in the pri-miRNAs was then calculated according to the nearest neighbor method (Mathews et al., 1999). We averaged the free energy values at each position and plotted the results, shown in Figure 1 (human) and Figure S1 (fly).

Based on these plots, the overall structure can be generalized, as shown in Figures 1B and 1D. In both humans and flies, pri-miRNAs consist of an imperfect stem structure of ~3 helical turns, which is surrounded by unstable segments at both ends. pri-miRNAs can be further divided into four parts comprising the terminal loop, upper stem, lower stem, and flanking sequences ("basal segments") (Figures 1B and 1D). Importantly, Drosha cleaves the RNA molecule at ~2 helical turns away from the terminal loop and ~1 helical turn away from the basal segments. The upper stem is stable at about the +3 position, whereas the free energy is relatively high in the middle of the upper

stem (+9 ∼ +12 positions). The lower stem is ∼11 bp long and often contains small internal loops at the −6 to about the −9 positions. The basal segments are often single-stranded or contain large bulges and/or internal loops at irregular positions and therefore the energy values are variable in these segments.

In the case of pri-miRNAs encoding miRNA at the 5′ arm (5′-donors), the most unstable position inside the stem corresponds to the +1 position (Figures 1A and S1A). This may explain why the 5′-strand is selected as a mature miRNA during miRISC assembly, as the miRNA duplex derived from these pri-miRNAs is expected to be less stable at the 5′ side of the mature miRNA. Similar calculations for the 3′-donors show relatively high free energy values at the opposite side (positions +19 ∼ +20) compared to the values obtained at positions +1 ∼ +2 (Figures 1C and 1D and S1B). The p values for the differences in free energy at the +1 ∼ +2 position compared to the +19 ∼ +20 position in humans were 5.9e-12 and 4.2e-08 for the 5′-donors and the 3′-donors, respectively. In *Drosophila* pri-miRNAs, the p values were 1.9e-04 and 7.4e-03 for 5′-donors and 3′-donors, respectively. This supports the current model in that the relative instability of the termini of the miRNA duplex may be the major determinant in strand selection, as is the case for siRNA duplexes (Khvorova et al., 2003; Schwarz et al., 2003).

Interestingly, the +12 position is relatively unstable in the 5′-donors, whereas the +9 position is unstable in the 3′-donors (Figure 1) (p values; 4.1e-05 for the +12 position of 5′-donors and 4.9e-03 for the +9 position of 3′-donors). When an miRNA is 22 nt long, the +9 position corresponds to the +12 position relative to the 5′ end of the mature molecule. Thus, in both the 5′- and 3′-donors, the 12th position relative to the 5′ end of the mature miRNA is unstable. A similar profile is observed for *Drosophila* miRNAs (Figure S1). One intriguing possibility is that thermodynamic stability of this position may influence strand selection and/or other steps during RISC assembly.

## Experimental Approach: Systematic Mutagenesis and In Vitro Processing Assay

Based on the above observations, mutations were introduced into each part of the pri-miRNAs to examine their significance in pri-miRNA processing. Because miRNA maturation is a multistep process in vivo, certain mutations may affect not only the pri-miRNA processing step but also other steps such as pre-miRNA export, cytoplasmic processing, and RNA turnover. To avoid such complications, we assayed the pri-miRNA cleavage reaction in vitro using labeled transcripts and an immunopurified Drosha-DGCR8 complex.

Mutagenesis was carried out based on the "minimal pri-miRNAs" that we previously developed (Han et al., 2004) (Figure S2). Minimal pri-miRNAs contain pre-miRNA sequences plus ∼20 nt sequences outside of the Drosha cleavage sites. For efficient transcription by T7 RNA polymerase, two additional Gs were incorporated between the promoter and the pri-miRNA sequences. An in vitro processing assay was carried out by incubating RNA with immunoprecipitated FLAG-tagged Drosha (Han et al., 2004). The Drosha-DGCR8 complex (Microprocessor) cleaves pri-miRNAs, yielding three kinds of fragments which are the 5′ flanking fragment (F1, ∼25 nt), pre-miRNA (F2, ∼65 nt), and the 3′ flanking fragment (F3, ∼20 nt) (Figure S2). To identify these fragments, processing reactions were carried out using 5′ end-labeled RNA as well as internally labeled RNA (Figures 2, 4, 6, 8, and S6). When necessary, the cleavage products were gel-purified, ligated to 3′ and 5′ adapters, reverse-transcribed, PCR-amplified, inserted into pGEM-T easy vector, and confirmed by sequencing. Alternatively, some fragments were gel-purified and analyzed by primer extension.

## The Terminal Loop Is Dispensable for pri-miRNA Processing

To investigate the role of the terminal loop in pri-miRNA processing, we eliminated it by converting it into two separate ssRNA segments (16-TL1) (Figure 2A). To prepare such a substrate, two strands of RNA were transcribed separately and annealed prior to processing. To assist with the identification of each processing product, either strand A or strand B was labeled at the 5′ end using [$\gamma$-$^{32}$P] ATP. Remarkably, this variant (16-TL1) lacking a terminal loop was processed at the original site (Figure 2A). A similar mutant, 16-TL2, containing altered sequences in the sliced loop also served as a good substrate for Drosha (Figure 2A). We then determined whether a large internal loop would substitute for a terminal loop by generating a longer substrate containing extended stems at both ends (16-TL3, Figure 2B). This variant was also processed accurately. Our results clearly demonstrate that the terminal loop structure itself is unnecessary for pri-miRNA processing. The Drosha-DGCR8 complex may therefore process not only hairpin RNAs but also other substrates such as long dsRNAs with large internal loops.

We next generated an inverted hairpin, 16-TL4, in which the basal segments were ligated to create a new terminal loop, whereas the original terminal loop was cleaved into separate ssRNA segments (Figure 2C). This "inverted hairpin" variant (16-TL4) was processed at the original cleavage site albeit with less accuracy and efficiency (Figure 2C, lanes 6 and 10). When the sequences in the cleaved terminal loop were modified to create a less stable ssRNA region (16-TL5), this variant was processed more efficiently at the precise cleavage site (Figure 2C, lanes 7 and 11). An additional inverted hairpin variant (16-TL6) containing an extended stem was cleaved similarly to 16-TL4, at the original site (Figure 2C, lanes 8 and 12). This result clearly shows that a terminal loop structure per se is not important for cleavage site selection in pri-miR-16-1. To test whether this conclusion could be generalized, we generated two more inverted hairpin variants, 31-TL1 and 23-TL1, based on pri-miR-31 and pri-miR-23a, respectively (Figure S3). Both of these variants were cleaved efficiently at their natural cleavage sites. Another miR-23 variant, 23-TL2, was also cleaved efficiently

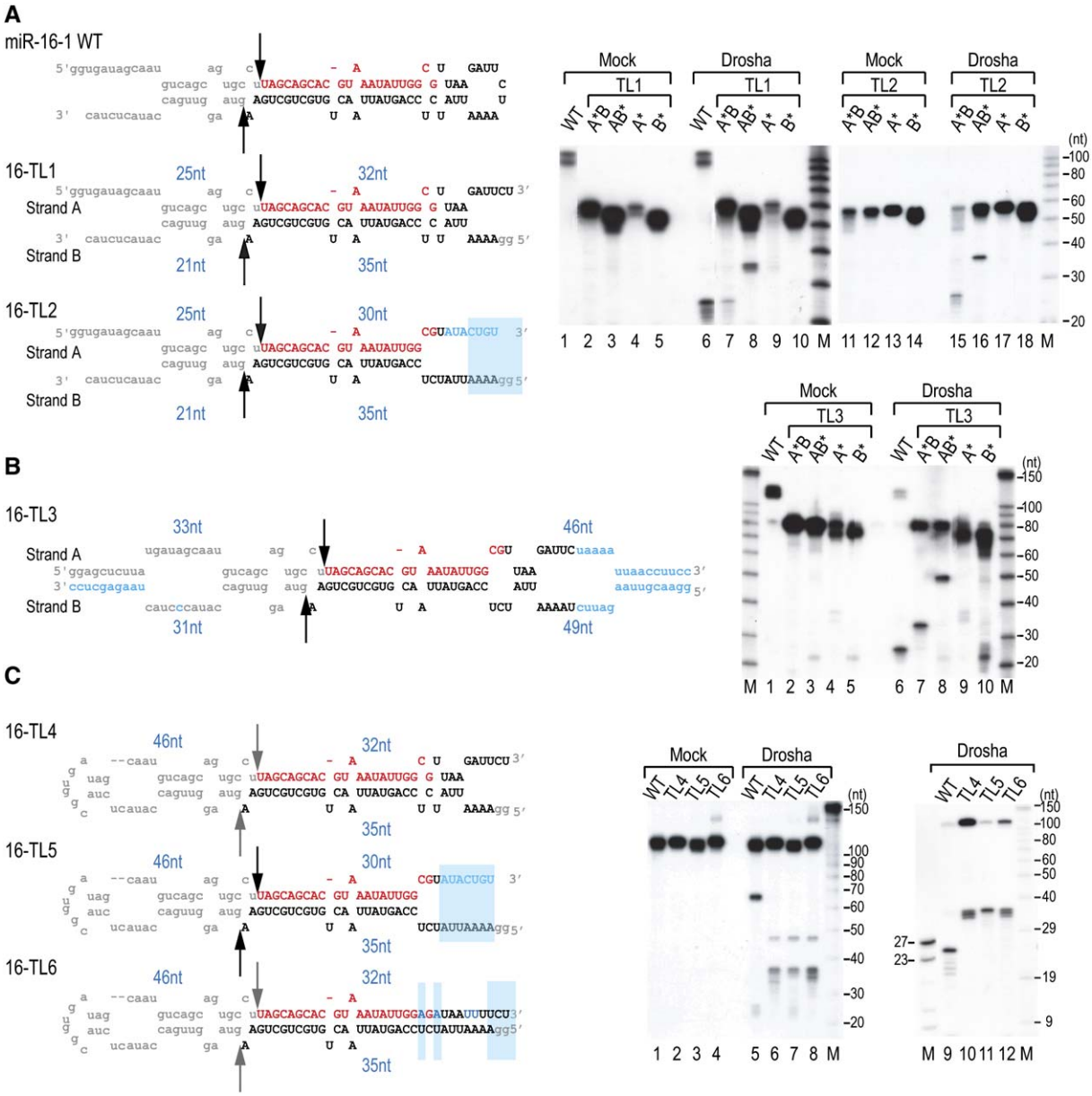


**Figure 2. A Terminal Loop Is Not Critical for pri-miRNA Processing**

(A) In vitro processing of the "sliced terminal loop" mutants (16-TL1 and 16-TL2). Each strand was labeled at the 5′ end using [$\gamma$-$^{32}$P] ATP and annealed with an unlabeled opposite strand before processing. The labeled strand is indicated with an asterisk (A* or B*). The altered sequences are indicated in light blue.

(B) In vitro processing of 5′-end-labeled dsRNA that has large internal loops at both termini.

(C) In vitro processing of inverted hairpin mutants. The RNA was labeled either internally (left panel, lanes 1–8) or at the 5′ end (right panel, lanes 9–12).

at the original site, indicating that the distance from the terminal loop is not imperative for cleavage site selection.

## The Single-Stranded Basal Segments Are Critical for pri-miRNA Processing

We and others have previously shown that the segments flanking the miRNA hairpin are important for efficient miRNA biogenesis (Lee et al., 2003; Yekta et al., 2004). To further investigate the molecular basis of this requirement, we generated a series of mutations in this region (Figures 3 and S6). When the basal segments were removed from the pri-miR-16-1 species, processing was abolished (Figure 3A, 16-ΔBS). Mutants retaining only one side of the flanking strands (16-5′BS and 16-3′BS) were processed although the efficiency was

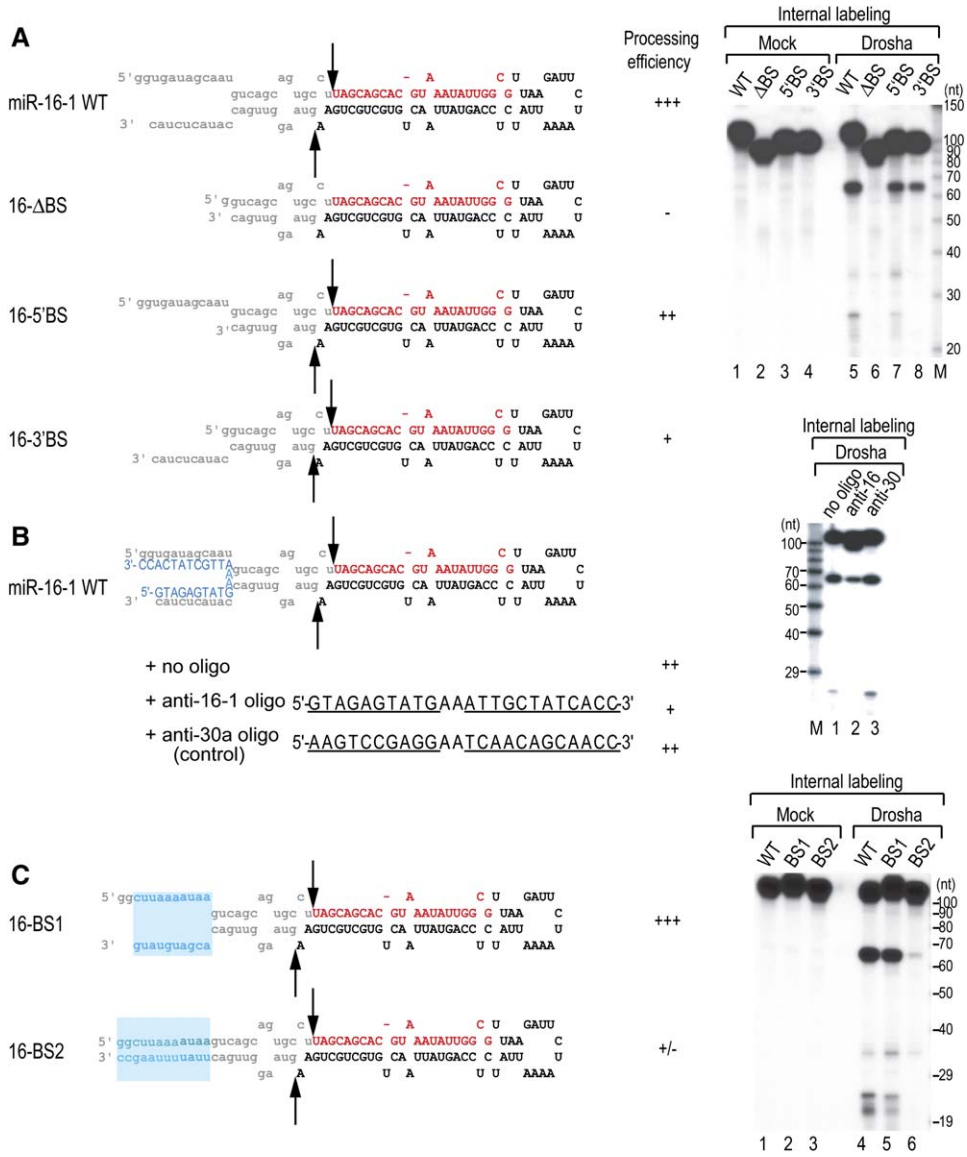


**Figure 3. Single-Stranded Basal Segments Are Critical for pri-miRNA Processing**

(A) In vitro processing of pri-miR-16-1 variants that are devoid of the basal segments (16-ΔBS, 16-5′BS, 16-3′BS).

(B) In vitro processing of pri-miR-16-1 in the presence of antisense oligonucleotide. The antisense oligo of 24 nt is complementary to the 5′ basal segment (12 nt) as well as to the 3′ basal segment (10 nt). The oligo also contains the 2 nt linker sequence, AA. Complementary sequences to the basal segments are underlined. Anti-30a oligo is complementary to pri-miR-30a and was used as a control.

(C) In vitro processing of pri-miR-16-1 variants that have alterations in the basal segments (16-BS1 and 16-BS2).

compromised (Figure 3A), suggesting that only one side of the flanking strands could support processing. We then blocked the basal segments by treating the pri-miRNA with an oligonucleotide that is complementary to these segments (Figure 3B). The oligonucleotide was in fact designed to bind to both the 5′ and 3′ basal segments simultaneously and was found to specifically repress the cleavage of pri-miR-16-1.

To examine how the basal segments contribute to processing, we altered their sequences in a further series of experiments. The mutant molecule 16-BS1, which retains its single-stranded structure in the basal segment region, was processed efficiently (Figure 3C). However, when the single-strands of the basal segments were converted into a double strand (16-BS2), the cleavage reaction was blocked. Thus, it is the single-stranded nature of the basal segments, rather than the nucleotide sequences, that may be critical for Drosha processing. It is noted that comparable observations have recently also been reported (Zeng and Cullen, 2005).

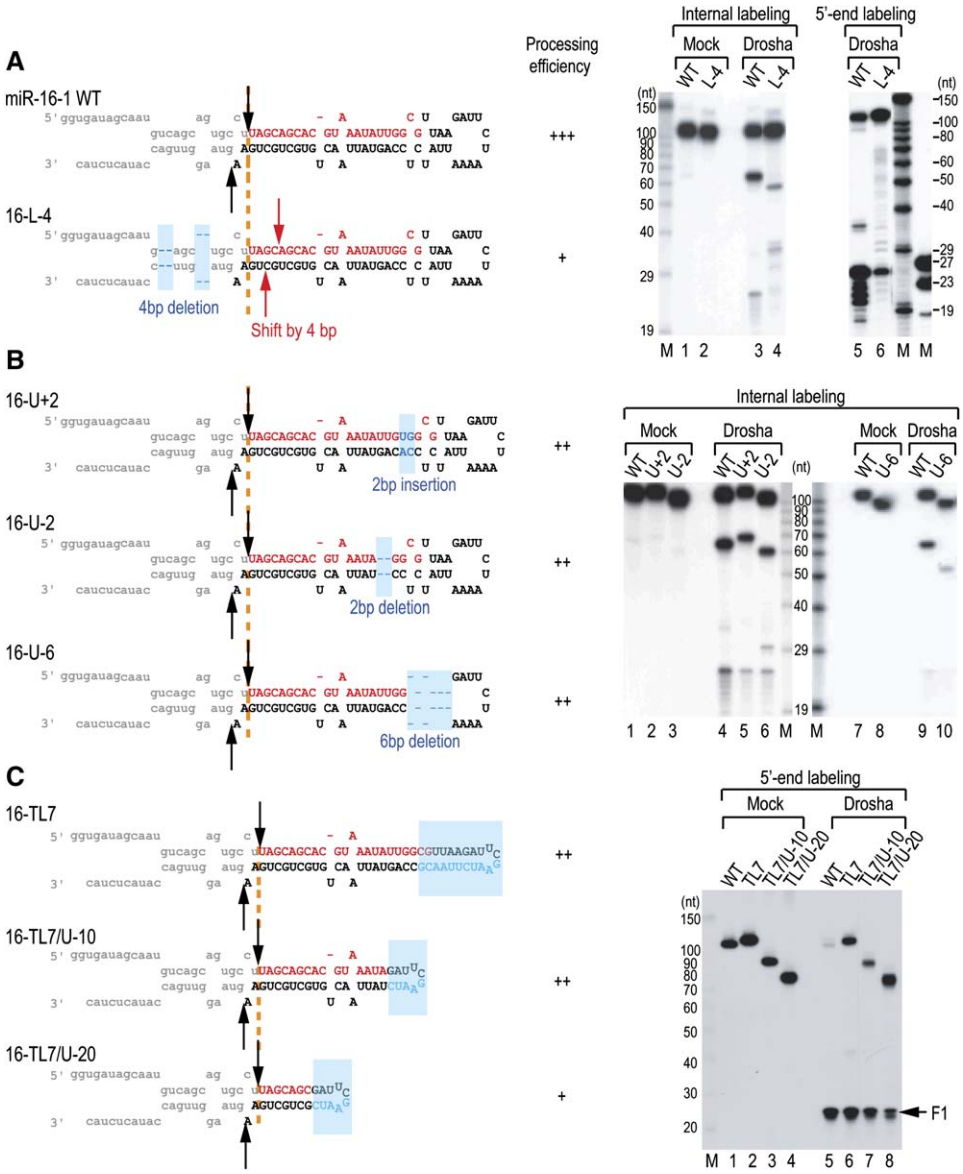


**Figure 4. Cleavage Site Is Determined Largely by the Distance from the Basal Segments**

(A) In vitro processing of the mutant-containing deletion in the lower stem. The altered cleavage site is indicated with red arrows. The original cleavage site is represented with orange dotted lines.

(B) In vitro processing of the mutants containing either insertion or deletion in the upper stem.

(C) In vitro processing of small terminal loop mutants.

## The Distance from the ssRNA Basal Segments Is Significant for Cleavage Site Selection

Animal pri-miRNAs typically contain a stem of ~3 helical turns. Zeng et al. previously suggested that the sites of Drosha cleavage may be determined largely by the distance (~22 nt) from the terminal loop (Zeng et al., 2005). In their study, the authors were able to show that when a pri-miR-30a variant was modified such that the loop-stem junction moves either 1 bp up the stem or 1 bp down the stem, the cleavage site shifted either 1 bp up the stem or 1 bp down the stem, respectively (Zeng et al., 2005). We have independently performed extensive experiments to address this same issue, but our data are in fact inconsistent with this "loop-stem junction anchoring" model.

We first introduced a deletion to the lower stem of the pri-miR-16-1 species to reduce the distance from the basal segments (Figure 4A). In a mutant where the distance from the basal segments decreases by 4 bp (16-L-4), the cleavage site was shifted by 4 bp away from these

segments. We then altered the distance between the cleavage site and the terminal loop by either replacement, deletion, or insertion in the upper stem of pri-miR-16-1 (Figures 4B and S6). Three upper-stem mutants were found to have been cleaved at the original site, in spite of such deletions or insertions (16-U+2, 16-U-2, and 16-U-6). A mutant containing a smaller terminal loop was also cleaved at the same site (Figure 4C, 16-TL7). We then introduced further deletions in the upper stem to generate variants with a small terminal loop and a shorter stem (16-TL7/U-10 and 16-TL7/U-20) and observed that these substrates were also cleaved at the original sites, albeit at a lower efficiency (Figure 4C). Together with our observations that mutants lacking the terminal loop are cleaved at their original sites (Figure 2), our results demonstrate that the distance from the terminal loop is unlikely to be the major determinant of cleavage site selection.

However, we did observe that the processing efficiency of the "small terminal loop" mutant (16-TL7) was slightly affected, which suggests that a flexible terminal loop may be beneficial in the reaction. Also, the processing became less efficient when there were reduced lengths of the stem (16-TL7/U-10 and 16-TL7/U-20), suggesting that the Microprocessor may need to contact the whole length (~33 bp) of the hairpin for full activity.

## Specific Interaction of DGCR8 with pri-miRNAs

Our data indicate that the Microprocessor may recognize the ssRNA basal segments and thereby measure the distance (~11 bp) from the junction between the basal segments and the stem. This model hypothesizes the existence of a molecular anchor that can distinguish the ssRNA-dsRNA junction. The Drosha-DGCR8 complex contains at least three dsRBDs, one on Drosha and two on DGCR8, but no known ssRNA binding domain has been identified in either of these proteins. Although partial fragments of Drosha and DGCR8 proteins have previously been shown to bind to ssRNA and dsRNA in simple GST-pull down experiments, the relative affinities of these proteins to various RNA species have not been determined (Zeng and Cullen, 2005).

To investigate which component(s) of the Microprocessor directly interact(s) with pri-miRNAs, a UV-crosslinking experiment was carried out by incubating Drosha and DGCR8 with an internally radiolabeled pri-miR-16-1. Both proteins were fused to the FLAG epitope (FLAG-DGCR8 and Drosha-FLAG), and then coexpressed and immunopurified from HEK293T cells (Figure 5A, left panel). Subsequent silver staining indicated that full-length DGCR8, full-length Drosha, and two truncated forms of Drosha were purified to near homogeneity. After UV crosslinking and treatment with an RNase A/T1 mixture, radioactivity was detected on DGCR8, but not on Drosha (Figure 5A, right panel). We have been unable to detect any significant RNA binding activity of Drosha in any other experiments (gel mobility shift assays and pull-down experiments) (data not shown). It is possible that Drosha may interact with its substrate only transiently during the catalytic reaction, whereas DGCR8 associates directly with the substrate in a more stable manner.

To further examine the mode of interaction between DGCR8 and RNA, the FLAG-DGCR8 protein product was immunopurified by intensive washing with high salt buffer (Figure 5B). Dissociation of endogenous Drosha from the immunoprecipitates was then verified by Western blotting using anti-Drosha antibodies (Figure S4A) and by an in vitro processing assay (Figure S4B). UV crosslinking experiments were then carried out using purified FLAG-DGCR8 protein and various RNA molecules (Figure S4C). FLAG-DGCR8 was found to have crosslinked strongly to pri-miRNA but less efficiently to an siRNA duplex or 23 nt ssRNA (Figure S4C).

The relative affinity of pri-miRNA to DGCR8 was next determined by competition experiments (Figure 5C). Internally labeled pri-miR-30a was crosslinked to FLAG-DGCR8 in the presence of different amounts of cold competitors such as siRNA duplex, 23 nt ssRNA, 80 bp dsRNA, or pre-miR-30a hairpin. Whereas cold pri-miR-30a successfully competed with hot pri-miR-30a, the other RNA molecules did not compete efficiently in this reaction (Figure 5C). Importantly, pre-miR-30a was found to have barely competed with pri-miR-30a under these conditions, suggesting that the main binding site for DGCR8 resides outside the upper stem and the terminal loop. This result also indicates that DGCR8 may dissociate from pre-miRNA upon processing. It is noteworthy also that long dsRNA has a relatively high affinity to DGCR8 and that 23 nt ssRNA is also capable of competing with pri-miR-30a weakly but reproducibly (Figure 5C). This suggests that DGCR8 may interact with pri-miRNAs by recognizing both ssRNA and dsRNA structures.

To test this possibility, we also carried out crosslinking experiments using mutated pri-miRNAs as cold competitors (Figure 5D). These mutants were disrupted either in their basal segments (16-ΔBS) or in their terminal loop structures (16-TL7). The basal segment mutant (16-ΔBS) could not compete for the binding to DGCR8 as efficiently as wild-type RNA. The pri-miRNA species with a smaller terminal loop (16-TL7) was slightly impaired in this binding but was still able to compete with wild-type RNA (Figure 5D). This result demonstrates that the ssRNA segments of the pri-miRNAs are critical for DGCR8 binding.

In addition, the "shorter-stem" mutants (16-TL7/U-10 and 16-TL7/U-20) were tested to examine the requirements for the minimal stem length for DGCR8 binding (Figure 5D). The stem lengths of mutants 16-TL7/U-10 and 16-TL7/U-20 are predicted to be 31–33 bp and 21–23 bp, respectively. As the stem becomes shorter, the mutants competed gradually less efficiently than the longer mutant 16-TL7. In fact, the DGCR8 binding affinities of these mutants correlated well with their processing efficiencies (Figures 3 and 4).

## Artificial Substrates

In order to confirm our present findings, we generated an artificial substrate bearing no sequence homology to any

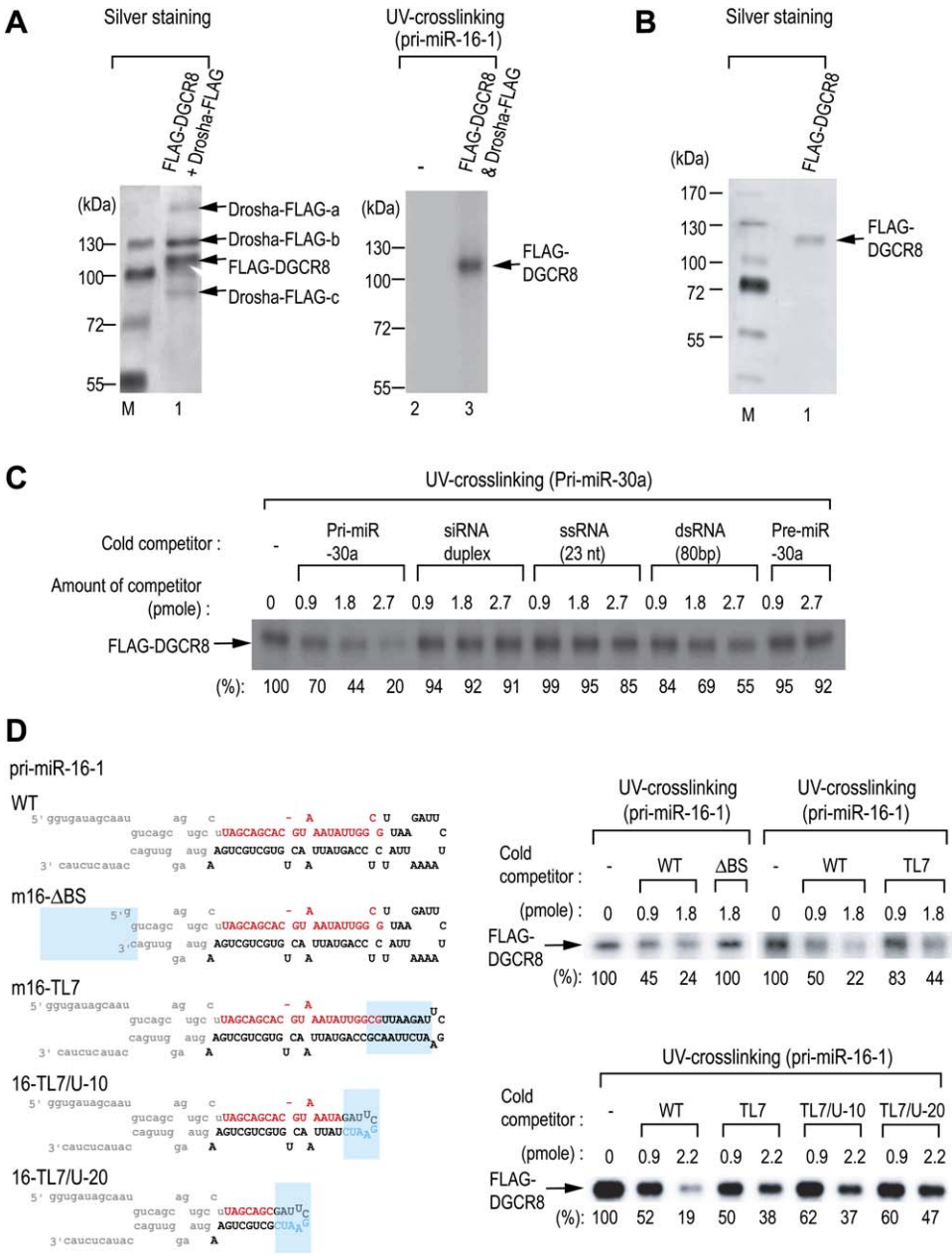


**Figure 5. DGCR8 Interacts Preferentially and Directly with pri-miRNAs**

(A) UV crosslinking. The left panel shows silver staining of the purified Drosha-FLAG and FLAG-DGCR8 proteins. These proteins were then incubated with internally labeled pri-miR-16-1, irradiated with UV and resolved on SDS-PAGE gel (right panel).

(B) Silver staining of immunopurified DGCR8. The FLAG-DGCR8 protein was immunopurified under high-salt condition to remove endogeneous Drosha protein.

(C) UV-crosslinking experiment. Various cold RNAs were used as competitors against internally labeled pri-miR-30a. Band intensity was quantified by using the UviDoc program and normalized against the band intensity obtained in the absence of a competitor (–).

(D) UV-crosslinking experiment. Internally labeled pri-miR-16-1 was incubated with purified DGCR8 protein and cold competitors.

known pri-miRNAs (Figure 6A). When annealed, the two RNA strands of this molecule are expected to form a simple structure of "ssRNA tails-3 helical turns-ssRNA tails." Either one of the two strands was then labeled at the 5′ end in a given reaction to allow for easy identification of the cleavage products. This artificial substrate was found to have been cleaved either at ~11 bp from the left junction (cleavage I) or at ~11 bp from the right junction (cleavage II), at a comparable efficiency (Figure 6A). Next, strand A was replaced by strand B in order to convert the ssRNA tails in one side into

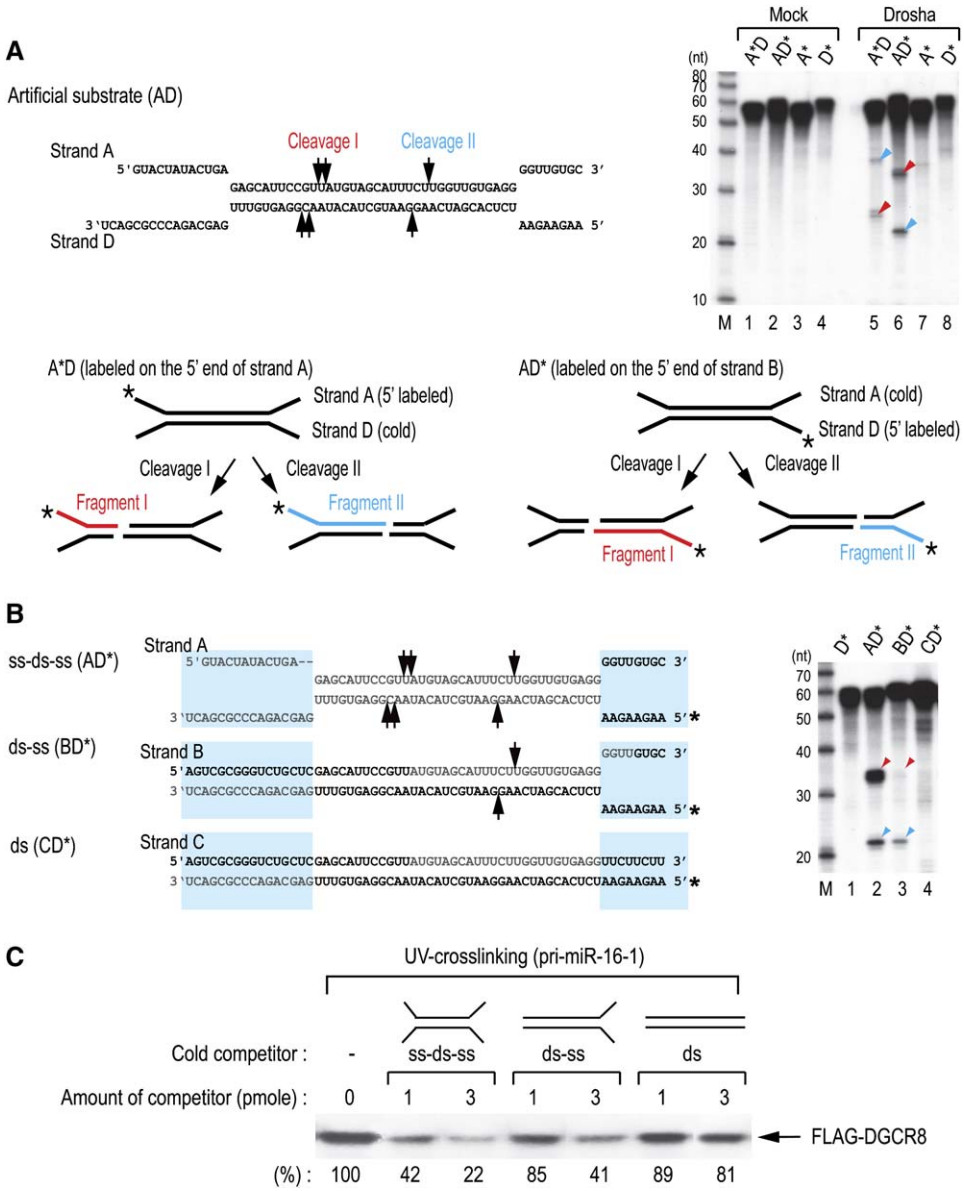


**Figure 6. Artificial Substrates for the Microprocessor**

(A) In vitro processing of an artificial substrate. Each strand was labeled at the 5′ end using [$\gamma$-$^{32}$P] ATP and annealed with the cold opposite strand before processing. The labeled strand is indicated with an asterisk (A* or D*). Two different processing events occurred (cleavage I or cleavage II). The products from cleavage I are indicated with red arrowheads while the products from cleavage II are indicated with blue arrowheads.

(B) In vitro processing of various artificial substrates. Strand D was labeled at the 5′ end.

(C) UV-crosslinking experiment. Cold artificial substrates were used as competitors against internally labeled pri-miR-30a transcript. Band intensity was quantified by using the UviDoc program and normalized against the band intensity that was obtained in the absence of a competitor (–).

an extended stem (Figure 6B, lane 3). Cleavage on the left side was abolished whereas cleavage on the right side was only slightly affected. When the ssRNA tails on the right side were also converted into a dsRNA stem, this simple dsRNA was not cleaved by Drosha (Figure 6B, lane 4). This result clearly shows that cleavage takes place at ~11 bp from the dsRNA-ssRNA junction.

To examine whether DGCR8 binds to the ssRNA tails, the relative affinity of these RNAs for DGCR8 was determined by further UV crosslinking-competition experiments (Figure 6C). The duplexes containing ssRNA tails in both sides displayed the highest affinity to DGCR8, whereas a simple duplex could not compete for this interaction effectively.

## DISCUSSION

### A ssRNA-dsRNA Junction Anchoring Model

Our analyses illustrate the importance of ssRNA tails as well as a ~33 bp stem for pri-miRNA processing. The average face of pri-miRNA consists of a stem of 3 helical turns surrounded by ssRNA segments at both ends (the basal segments on one side and the terminal loop on the other side). Drosha cleaves both natural and artificial substrates at a site ~11 bp away from the ssRNA-dsRNA junction (SD junction) (Figure 7). Manipulating the length of the outer stem affects cleavage site selection, implicating the existence of a molecular device that measures the distance from the SD junction. We contend that DGCR8 is likely to function as the "molecular anchor."

pri-miRNA processing may consist of two sequential steps; substrate recognition and catalytic reaction (Figure 7). First, DGCR8 may recognize the substrate by tight anchoring at the SD junction and, at the same time, interacting with the ~33 bp stem. Drosha, on the other hand, is not in direct contact with RNA at this stage. After this "precleavage" complex is formed, the dsRBD of Drosha may interact transiently with the stem to locate the processing center of the enzyme at ~11 bp from the SD junction.

Our data also show that the terminal loop is rather irrelevant to pri-miRNA processing. We note, however, that both processing and DGCR8 binding were slightly impaired in a mutant with a small loop (16-TL7) (Figures 4C and 5D), suggesting that the presence of a large loop may be beneficial to some extent. It is possible that a terminal loop that is too small in size may impose structural constraints upon the stem and affect processing. An additional and nonmutually exclusive possibility is that a large terminal loop may act as a flexible ssRNA and loosely interact with DGCR8. Supporting this notion, terminal loop mutants were cleaved more efficiently when the loop was converted into more flexible ssRNA (Figure 2, compare 16-TL1 with 16-TL2; 16-TL5 with 16-TL6). Moreover, an artificial substrate with ssRNA tails at only one end was cleaved less efficiently than a substrate with ssRNA tails at both ends (Figure 6B). This may explain why a large loop appears to be required for the processing of pri-miR-30a (Zeng et al., 2005).

We generated a small terminal loop mutant of pri-miR-30a in our current study (30-TL1) and found that this mutant was still cleaved, although both the efficiency and accuracy of the reaction was greatly compromised (Figure S5A, lane 7). Because the basal segments of pri-miR-30a can form a short stem (Figure S5A), this region may not interact strongly with DGCR8. Thus, the interaction of pri-miR-30a with DGCR8 may be more dependent on the terminal loop. In fact, the processing efficiency of pri-miR-30a is considerably lower than other pri-miRNAs (data not shown) and at least twice the amount of protein is required for pri-miR-30a processing, compared with pri-miR-16-1, to achieve comparable efficiency. When the basal segments of pri-miR-30a were mutated into more flexible ssRNA segments (30-DBS and 30-TL1/DBS), both the efficiency and accuracy of the processing was improved (Figure S5A, lanes 6 and 8).

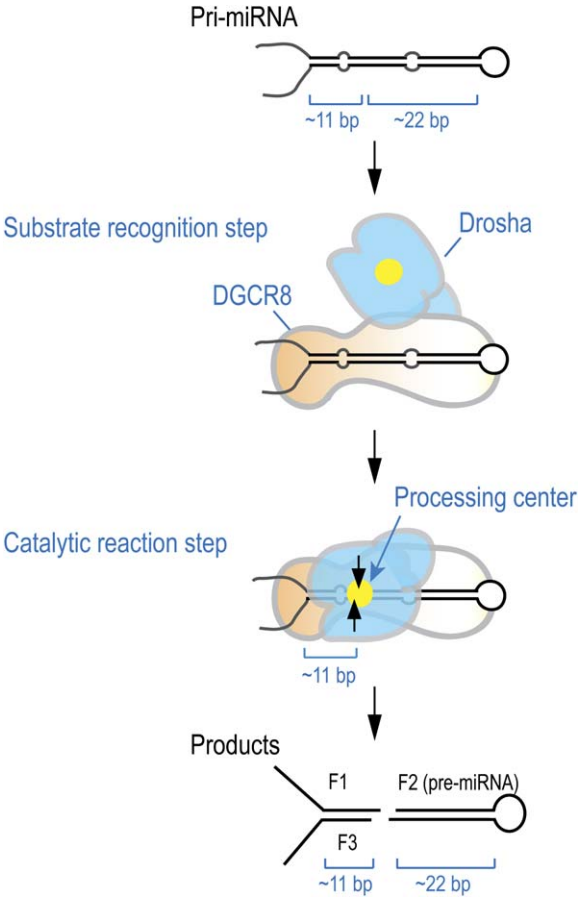


**Figure 7. A "ssRNA-dsRNA Junction Anchoring" Model for the Processing of pri-miRNA**

DGCR8 may play a major role in substrate recognition by directly anchoring at the ssRNA-dsRNA junction. DGCR8 also interacts with the stem of ~33 bp and the terminal loop for a full activity although the terminal loop structure is not critical for DGCR8 binding and cleavage reaction. After the initial recognition step, Drosha may transiently interact with the substrate for catalysis. The processing center (yellow circle) of Drosha is placed at ~11 bp from the basal segments.

### Productive Processing versus Abortive Processing

Because some large terminal loops can be seen as unstructured ssRNA segments, pri-miRNAs may be considered to be a "ssRNA-dsRNA (~3 helical turns)-ssRNA" structure. Yet, they are known to be cleaved at a position ~11 bp from the basal segments but not at ~11 bp from the terminal loop. In actual fact, we were able to detect, after extended exposure of the films, fragments generated by cleavage at the +16 position in pri-miR-30a (Figure 8A) and at the +12 position in pri-miR-16-1 (Figure 8B). This indicates that the Drosha-DGCR8 complex may bind to pri-miRNA in one of two alternative orientations, such that the processing center is located either at ~11 bp from the

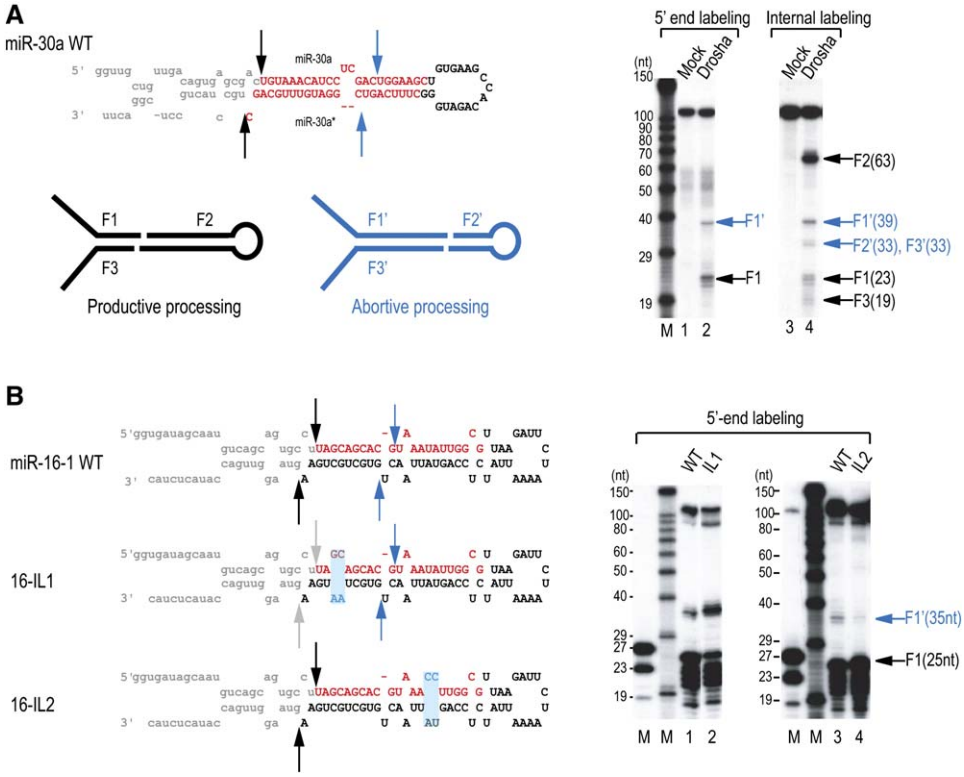


**Figure 8. Productive Processing versus Abortive Processing**

(A) The abortive processing sites in pri-miR-30a were determined by in vitro processing assay. The blue arrows indicate the abortive cleavage sites. The fragments from abortive processing are designated as F1′, F2′, and F3′.

(B) The abortive processing sites in pri-miR-16-1 were determined by 5′-end labeling followed by in vitro processing assay. To generate the mutants 16-IL1 and 16-IL2, an artificial internal loop of 2 bp was introduced to the upper stem region by changing the sequences in the stem.

basal segments or less efficiently at ~11 bp from the terminal loop. In this case, cleavage at the +1 position would be "productive" in the sense that it produces functional pre-miRNA. The processing at ~11 bp from the terminal loop would thus be "abortive" because the cleavage product does not contain miRNA sequences in full.

pri-miRNAs may have evolved to allow the Microprocessor to bind more favorably in the productive orientation. This biased binding and processing may be attributable to multiple elements that influence both the substrate binding and the subsequent catalytic reaction. The flexible ssRNA in the basal segment region appeared to suppress abortive processing in pri-miR-30a (Figure S5A, 30-DBS), suggesting that terminal structure influences the orientation of DGCR8 binding to pri-miRNA. The internal loop structures in the stem may also influence the ratio between productive processing and abortive processing (Figure 8B). When a small internal loop was introduced near the productive site, productive processing was suppressed while abortive processing was facilitated (Figure 8B, 16-IL1). But when the internal loop was placed close to the abortive processing site, abortive processing was reduced while productive processing was more efficient (Figure 8B, 16-IL2).

## Additional Substrates for Drosha?

The substrate specificity of Drosha revealed in this study may facilitate the computational prediction of novel miRNA genes. In addition, our findings suggest that the natural substrates of Drosha may not be restricted to pri-miRNAs as this enzyme can cleave not only hairpin-like structures but also dsRNA composed of two separate strands. Long dsRNAs with large internal loops separated by ~3 helical turns may also be recognized by the Drosha-DGCR8 complex. For example, Drosha may cleave the transcripts derived from slightly different repeat sequences that may form long dsRNAs interrupted by large internal loops. It would also be interesting to determine whether Drosha is involved in the metabolism of the abundant sense-antisense RNA pairs identified by recent genomic studies (Chen et al., 2004b; Yelin et al., 2003).

## Implications for RNA Interference

The regulated expression of shRNA using inducible pol II promoters would be desirable, particularly in clinical applications. However, a major technical hurdle when using pol II promoters is that the primary transcript must be processed efficiently and accurately by the Drosha-DGCR8 complex. The rational design of shRNAs may therefore

benefit from our current thermodynamic stability profiling and mutagenesis studies. For instance, the stem needs to be of 33 to 35 bp in length. The cleavage site should be placed at a position that is 11 bp from the basal segments. If the guide sequences are placed in the 5′ arm of the hairpin, it is also desirable that the +1 position is mismatched, whereas the +19 ∼ 20 positions are made relatively stable. In addition, abortive processing has to be avoided to increase the yield of functional siRNAs. Having flexible ssRNA segments in the basal segment region is therefore important, not only to maximize the efficiency of productive processing but also to minimize the possibility of abortive processing. The inclusion of small internal loops/bulges at the +9 ∼ +12 positions may be beneficial to suppress abortive processing and possibly to assist strand selection/RISC assembly. We contend that the fundamental principles of pri-miRNA recognition by Microprocessor that we present in this study will provide the basis for the future design of an efficient shRNA expression system.

## EXPERIMENTAL PROCEDURES

### Thermodynamic Stability Profiling of pri-miRNAs

We collected pri-miRNA sequences of 110 nt in length from genome sequences using genome annotation information of miRBase release 7.0 which at the time of study enlisted 321 human and 68 fly miRNAs. We initially prepared pri-miRNA sequences from human genome assembly NCBI35 (http://www.ncbi.nlm.nih.gov/Ftp/) and pri-miRNAs from *D. melanogaster* genome assembly BDGP3 (http://www.fruitfly.org/sequence/download.html). The secondary structure of RNA was predicted using the mfold program version 3.1 (Zuker, 2003) (http://www.bioinfo.rpi.edu/∼zukerm/rna/mfold-3.html). We then selected 280 human and 55 fly pri-miRNAs, excluding pri-miRNAs that are predicted to form multiloops.

Thermodynamic stability at each position in pri-miRNAs was calculated according to the nearest neighbor method using thermodynamic parameters determined at 37°C for all stacking energy values, taking into account all the different destabilizing elements such as internal loops and bulges (Mathews et al., 1999) (http://www.bioinfo.rpi.edu/∼zukerm/cgi-bin/efiles-3.0.cgi). We averaged the free energy at each position and plotted our findings in Figures 1 and S1.

The method employed in this study was previously devised by Krol and colleagues to calculate the thermodynamic features of pre-miRNAs (Krol et al., 2004). Our approach is different from that of Krol and colleagues in the following aspects: (1) our calculation included ∼20 nt outside the miRNA hairpin whereas Krol et al. considered only the hairpin region; (2) we plotted the thermodynamic profile by calculating the ΔG values at each individual position while Krol et al. calculated the ΔG values from the windows comprising three nucleotides; (3) we included most known human and fly miRNAs in our calculation while only 13 miRNAs were calculated in Krol et al.'s.

It is noted that bulges and internal loops were considered as one position regardless of their size because the energy values at the bulges or internal loops cannot be divided into individual nucleotides when using the nearest neighbor model. The energy values are assigned at the base pair before the bulge and internal loop. Therefore, when an internal loop or bulge exits, the accurate comparison of each position becomes inevitably more difficult with increasing distance from the standard position. To avoid the positioning problem, we introduced blanks at the position by the size of the bugles or internal loops and assigned the free energy value at the 3′ position of the bulges or internal loops. Finally, we constructed the free energy profile including blanks, which indicate unmatched pairs, insertions, and deletions (indels). The blanks are not considered when calculating the means and the standard deviations at each position.

The significance of the differences in free energy values at a given position compared to other positions were determined by two-sample t test performed in the R language. For the +12 position of the 5′-donors, we collected the first sample data at the +12 position and the second sample data from the +4 ∼ +11 and +13 ∼ +17 positions. For the +9 position of the 3′-donors, the first sample data was obtained from the +9 position while the second sample data was collected from the +4 ∼ +8 and +10 ∼ +17 positions. The resulting p values for the +12 position of human 5′-donors and the +9 position of human 3′-donors were 4.1e-05 and 4.9e-03, respectively. In *Drosophila* pri-miRNAs, the p values were 0.011 and 0.033 for the +12 and +9 positions, respectively. The p values for the differences in free energy at the +1 ∼ +2 positions compared to the +19 ∼ +20 positions are 5.9e-12 and 4.2e-08 for the 5′-donors and the 3′-donors in humans, respectively, and 1.9e-04 and 7.4e-03 for the 5′-donors and the 3′-donors in flies, respectively.

### Generation of Mutant miRNAs that Contain Extended or Shortened Upper Stem

To generate the mutant 16-U+2, a pair of oligonucleotides that contain mutated sequences (Table S2) was annealed and filled in using the expanded high-fidelity PCR system (Roche). The PCR product was then subcloned into pGEM-T easy (Promega). The same method was used to modify the upper stem length. The sequences of the oligonucleotides are described in Table S2.

### Generation of Double-Stranded RNA

To prepare 16-TL1, 16-TL2, and 16-TL3, each strand of the RNA duplex was synthesized by in vitro transcription. The template for transcription was amplified by PCR using the following primers. For strand A of 16-TL1, primers 16-WT-F and 16-TL1A-R were used for PCR. For strand B of 16-TL1, primers 16-TL1A-F and 16-WT-R were used for PCR. For strand A of 16-TL2, primers 16-WT-F and 16-TL2A-R were used for PCR. For strand A of 16-TL3, primers 16-TL3A-F and 16-TL3A-R were used for PCR. For strand B of 16-TL3, primer 16-TL3B-F and 16-TL3B-R were used for PCR. Following in vitro transcription using the PCR products, the sense-antisense pairs were heated in annealing buffer (Invitrogen) at 95°C for 15 s and cooled down slowly to allow annealing.

### Generation of the Inverted Hairpin Mutant

Two strands were transcribed separately as described above and ligated at the basal segments using T4 DNA ligase and a DNA "bridge" (a DNA oligonucleotide complimentary to the joint area). The DNA bridge contains the sequences complementary to the basal segment of miR-16-1 (5′-CATTGCTATCACCGTAGAGTATGG-3′). First, strand A of 16-TL1 was dephosphorylated by calf intestine phosphatase (TAKARA). Strand A, strand B, and the DNA bridge were then mixed and ligated using T4 DNA ligase (TAKARA) at 30°C for 4 hr. RNA was extracted from the reaction mixture by phenol extraction and the shifted band was gel-purified. This ligated RNA was subsequently used for reverse transcription using SUPERSCRIPT II (Invitrogen). The primer employed for reverse transcription was the 16-WT-R. The sequences of the primers used for PCR amplification of 16-TL4, 16-TL5, and 16-TL6 are described in Table S3.

### Hybridization of the Basal Segments of pri-miR-16-1 with Antisense DNA Oligonucleotide

One hundred picomole of antisense DNA oligonucleotide was mixed in the processing reaction. The sequence of anti-16-1, which is complementary to the basal segments of pri-miR-16-1, is 5′-GTAGAGTATG AAATTGCTATCACC-3′. A control oligonucleotide, anti-30a, is 5′-AA GTCCGAGGAATCAACAGCAACC-3′.

### Preparation of Artificial Substrates

Strands A, B, C, and D of artificial substrates for Microprocessor were synthesized (Samchully) and then phosphorylated at the 5′end using T4 polynucleotide kinase (TAKARA) and [$\gamma$-$^{32}$P] ATP. The RNAs were heated in 1 × TE at 95°C for 15 s and cooled down slowly to allow annealing.

### Cell Culture and Transfection

HEK293T cells were cultured in DMEM (WelGENE) supplemented with 10% FBS (WelGENE). Eight micrograms of pCK-Drosha-FLAG and/or 5 μg of pCK-FLAG-DGCR8 were transfected into HEK293T cells grown in a 100 mm dish using the calcium-phosphate method.

### Immunoprecipitation and In Vitro Processing of pri-miRNAs

In vitro processing of pri-miRNAs was carried out as previously described (Lee et al., 2002, 2003). Briefly, in 30 μl reaction, 6.4 mM $MgCl_2$, 1 unit/μl of Ribonuclease Inhibitor (TAKARA), the labeled transcripts of 1 × $10^4$ to 1 × $10^5$ cpm, and 15 μl of the beads in buffer D′ from immunoprecipitation were included in the mixture. The reaction mixture was incubated at 37°C for 60 min. RNA was phenol-extracted from the reaction mixture and analyzed on 12.5% denaturing urea-polyacrylamide gel. RNA size markers (Decade marker, Ambion) labeled at the 5′ end were used. When necessary, two synthetic RNAs of 23 nt and 27 nt were labeled at the 5′ ends and used as additional size markers.

### Purification of Recombinant DGCR8 Proteins

HEK293T cells were transfected with FLAG-DGCR8 expression vector. Two days after transfection, cells were harvested and sonicated in ice-cold buffer D-K'200 (20 mM Tris, pH 8.0, 200 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF). After centrifugation at 13,200 rpm at 4°C for 15 min, the supernatant was treated with 50 μg/ml of RNase A at 4°C for 30 min. This extract was then incubated with anti-FLAG antibody conjugated to agarose beads with constant rotation for 120 min at 4°C. The beads were washed four times in buffer D-Na'2500 (20 mM Tris, pH 8.0, 2.5 M NaCl, 0.2 mM EDTA, 0.2 mM PMSF, 1% Triton X-100) and then three times with FLAG-elution buffer (50 mM Tris, pH 7.4, 150 mM NaCl). The protein was eluted with FLAG-elution buffer containing 400 μg/ml of 3X FLAG peptide (Sigma) at 4°C for 60 min and then concentrated into 20 ng/μl by using Centricon YM-30 (Millipore).

### Preparation of 80 bp dsRNA, ssRNA, and siRNA Duplex for UV Crosslinking Experiments

To prepare template DNA for sense or antisense transcript of dsRNA, 80 bp regions of firefly luciferase cDNA were amplified by PCR using 100 ng of pGL3 vector (Promega) as the template. The forward primers for PCR contain the T7 promoter sequences at their 5′ ends. The sequences of the primers are 5′-TTAATACGACTCACTATAGGGCATTTCGCAGCCTACCGTGG-3′ (forward primer for sense strand), 5′-TTGGGAGCTTTTTTTGCACGTTC-3′ (reverse primer for sense strand), 5′-TTAATACGACTCACTATAGGGAGCTTTTTTTGCACGTTCAA-3′ (forward primer for antisense strand), and 5′-ATGGGCATTTCGCAGCCTACCG-3′ (reverse primer for antisense strand). The PCR products were then used as the templates for in vitro transcription to prepare sense and antisense transcripts of dsRNA. The sense-antisense pair was annealed into duplex in 1× universal buffer (6 mM HEPES-KOH, pH 7.5, 20 mM KCl, and 0.2 mM $MgCl_2$) by boiling at 90°C for 2 min followed by incubation at 30°C for 1 hr. The sequences of ssRNA of 23 nt and siRNA duplex are 5′-UCUUUGGUUAUCUAGCUGUAUGA-3′ (23 nt ssRNA) and 5′-UUAAGGCACGCGGUGAAUGCCA-3′(sense strand for siRNA duplex) and 5′-GCAUUCACCGCGUGCCUUAAUU-3′ (antisense strand for siRNA duplex), respectively.

### UV Crosslinking

Twenty to fifty ng of FLAG-DGCR8 and radiolabeled RNAs of 1 × $10^6$ c.p.m. (50 ∼ 100 fmole) were mixed in 15 μl of binding buffer (10 mM Tris, pH 7.5, 50 mM KCl, 0.5 mM DTT, 1U of RNasin) in 96-well plates and then incubated at 4°C for 30 min. For competition assay, cold competitors were added to the reaction mixture. The RNAs were either prepared by in vitro transcription (pri-miR-16-1, m16-ΔBS, m16-TL1, pri-miR-30a, 80 bp dsRNA) or purchased from Samchully Phamaceuticals (siRNA duplex, 23 nt ssRNA, pre-miR-30a, artificial substrates). The 96-well plate containing the reaction mixture was brought into contact with a UV lamp in a UV crosslinker (CL-1000 UV-crosslinker, UVP) for 5 min. The mixture was then treated with the RNase A/T1 mixture and subsequently loaded on 7.5% SDS-polyacrylamide gel.

### Supplemental Data

Supplemental Data include Experimental Procedures, six figures, and three tables and can be found with this article online at http://www.cell.com/cgi/content/full/125/5/887/DC1/.

## ACKNOWLEDGMENTS

We thank members of our laboratory and Dr. Ji-Sook Hahn for critical reading of the manuscript. This work was supported by grants from the Basic Research Program (R02-2004-000-10173-0), the National Research Laboratory Programs (M1050000010905J000010910 and M10412000095-04J0000-03610), the SRC program (R11-2005-009-01003-0), the Molecular and Cellular BioDiscovery Research Program (2005-00518), and the BK21 program. J.H. was supported by the Seoul Science Fellowship.

Received: December 21, 2005
Revised: February 22, 2006
Accepted: March 23, 2006
Published: June 1, 2006

## REFERENCES

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

Basyuk, E., Suavet, F., Doglio, A., Bordonne, R., and Bertrand, E. (2003). Human let-7 stem-loop precursors harbor features of RNase III cleavage products. Nucleic Acids Res. *31*, 6593–6597.

Blaszczyk, J., Tropea, J.E., Bubunenko, M., Routzahn, K.M., Waugh, D.S., Court, D.L., and Ji, X. (2001). Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. Structure (Camb) *9*, 1225–1236.

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA *10*, 185–191.

Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA *10*, 1957–1966.

Chen, C.Z., Li, L., Lodish, H.F., and Bartel, D.P. (2004a). MicroRNAs modulate hematopoietic lineage differentiation. Science *303*, 83–86.

Chen, J., Sun, M., Kent, W.J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R.Z., and Rowley, J.D. (2004b). Over 20% of human transcripts might form sense-antisense pairs. Nucleic Acids Res. *32*, 4812–4820.

Croce, C.M., and Calin, G.A. (2005). miRNAs, cancer, and stem cell division. Cell *122*, 6–7.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231–235.

Dickins, R.A., Hemann, M.T., Zilfou, J.T., Simpson, D.R., Ibarra, I., Hannon, G.J., and Lowe, S.W. (2005). Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. Nat. Genet. *37*, 1289–1295.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235–240.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. Cell *106*, 23–34.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev. *18*, 3016–3027.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834–838.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev. *15*, 2654–2659.

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

Kim, V.N. (2005a). MicroRNA biogenesis: coordinated cropping and dicing. Nat. Rev. Mol. Cell Biol. *6*, 376–385.

Kim, V.N. (2005b). Small RNAs: Classification, biogenesis, and function. Mol. Cells *19*, 1–15.

Krol, J., Sobczak, K., Wilczynska, U., Drath, M., Jasinska, A., Kaczynska, D., and Krzyzosiak, W.J. (2004). Structural features of microRNA (miRNA) precursors and their relevance to miRNA biogenesis and small interfering RNA/short hairpin RNA design. J. Biol. Chem. *279*, 42230–42239.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). The human DiGeorge syndrome critical region gene 8 and Its D. melanogaster homolog are required for miRNA biogenesis. Curr. Biol. *14*, 2162–2167.

Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. EMBO J. *21*, 4663–4670.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415–419.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. *23*, 4051–4060.

Lingel, A., Simon, B., Izaurralde, E., and Sattler, M. (2004). Nucleic acid 3′-end recognition by the Argonaute2 PAZ domain. Nat. Struct. Mol. Biol. *11*, 576–577.

Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., et al. (2005). MicroRNA expression profiles classify human cancers. Nature *435*, 834–838.

Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95–98.

Ma, J.B., Ye, K., and Patel, D.J. (2004). Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature *429*, 318–322.

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.

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. *288*, 911–940.

Schwarz, D.S., Hutvagner, 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.

Silva, J.M., Li, M.Z., Chang, K., Ge, W., Golding, M.C., Rickles, R.J., Siolas, D., Hu, G., Paddison, P.J., Schlabach, M.R., et al. (2005). Second-generation shRNA libraries covering the mouse and human genomes. Nat. Genet. *37*, 1281–1288.

Song, J.J., Liu, J., Tolia, N.H., Schneiderman, J., Smith, S.K., Martienssen, R.A., Hannon, G.J., and Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. Nat. Struct. Biol. *10*, 1026–1032.

Tomari, Y., and Zamore, P.D. (2005). MicroRNA biogenesis: Drosha can't cut it without a partner. Curr. Biol. *15*, R61–R64.

Vermeulen, A., Behlen, L., Reynolds, A., Wolfson, A., Marshall, W.S., Karpilow, J., and Khvorova, A. (2005). The contributions of dsRNA structure to Dicer specificity and efficiency. RNA *11*, 674–682.

Yan, K.S., Yan, S., Farooq, A., Han, A., Zeng, L., and Zhou, M.M. (2003). Structure and conserved RNA binding of the PAZ domain. Nature *426*, 468–474.

Yekta, S., Shih, I.H., and Bartel, D.P. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. Science *304*, 594–596.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., et al. (2003). Widespread occurrence of antisense transcription in the human genome. Nat. Biotechnol. *21*, 379–386.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. *17*, 3011–3016.

Zeng, Y., and Cullen, B.R. (2005). Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. J. Biol. Chem. *280*, 27595–27603.

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.

Zeng, Y., Yi, R., and Cullen, B.R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. EMBO J. *24*, 138–148.

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.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.