Rethinking the Microprocessor

Hervé Seitz[1] and Phillip D. Zamore[1,*]
[1]Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01532, USA
*Contact: phillip.zamore@umassmed.edu
DOI 10.1016/j.cell.2006.05.018

MicroRNAs (miRNAs) are tiny regulators of gene expression that are processed from longer primary transcripts. In this issue, Han et al. (2006) report some of the structural features of the primary transcript that ensure that the Drosha-DGCR8 enzyme complex liberates precisely the correct precursor sequence, enabling production of a fully functional miRNA.

MicroRNAs (miRNAs) are single-stranded, 21–23 nucleotide RNAs that are able to repress specific target genes. They do this by base-pairing to target mRNAs, and then either accelerating degradation of the mRNA or inhibiting its translation. Base-pairing of miRNAs to their mRNA targets is often imperfect because miRNA nucleotides 2 through 7 have a disproportionate influence on target RNA selection. This "seed sequence" typically is exactly complementary to the target RNA, whereas the rest of the duplex tolerates imperfections in pairing. Because miRNA function requires such a remarkably small amount of complementarity to a target RNA, each miRNA species may regulate hundreds of distinct mRNA sequences. Given that the human genome may contain genes producing ~1,000 different miRNAs (Berezikov et al., 2005), it is possible that a majority of human genes are regulated posttranscriptionally by miRNAs (Lewis et al., 2005). Not surprisingly, miRNAs regulate a wide range of biological processes, including developmental timing, differentiation, apoptosis, insulin secretion, and even innate immunity against viruses.

Like mRNAs, miRNA genes are initially transcribed by RNA polymerase II as long primary transcripts (pri-miRNAs) that require subsequent processing to yield a functional mature miRNA (see Figure 1). In animals, pri-miRNAs are processed in the nucleus by the RNase III enzyme Drosha, acting with its double-stranded RNA binding partner protein DGCR8 (in vertebrates) or Pasha (in invertebrates). This RNase III enzyme/dsRNA binding protein partnership converts pri-miRNAs into small stem-loop structures called precursor miRNAs (pre-miRNAs), which are then processed further by a second RNase III enzyme/dsRNA binding protein duo into mature miRNAs. New work by Han et al. (2006) in this issue of *Cell* reveals some of the structural features of pri-miRNAs that encourage the Drosha-DGCR8 complex to liberate precisely the correct pre-miRNA, hence ensuring production of the correct miRNA sequence.

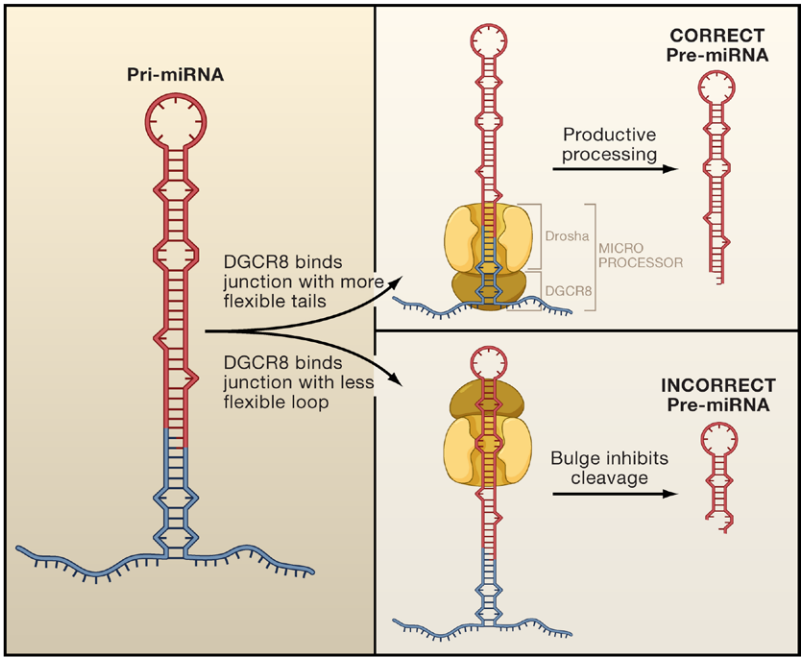

Figure 1. Pri-miRNA Processing by Drosha and DGCR8, Components of the Microprocessor
Structural features of pri-miRNAs promote their accurate processing into pre-miRNAs by the RNase III enzyme Drosha and its double-stranded RNA binding protein partner, DGCR8 (Pasha in invertebrates). DGCR8 is thought to bind more favorably to the junction between the rigid double-stranded stem and the 5′ and 3′ flexible, single-stranded segments of the pri-miRNA than to the junction between the stem and the considerably more constrained loop. Correct binding of DGCR8 to the base of the stem is proposed to position the processing center of Drosha ~11 bp up along the stem, where it makes a staggered pair of breaks in the RNA to create the ~65 nucleotide-long pre-miRNA. Binding of DGCR8 at the loop end of the stem positions Drosha inappropriately. Unpaired or weakly paired nucleotides at this site serve to discourage such unproductive cleavage, reducing the number of abortive Drosha products and favoring accurate pre-miRNA production.

The sequence of the miRNA is embedded in one of the two arms of the stem of a stem-loop structure within the pri-miRNA. The Drosha-DGCR8 heterodimer, a component of a large protein complex dubbed the "Micropocessor," cuts the stem loop containing the future miRNA out of the pri-miRNA (Denli et al., 2004; Gregory et al., 2004; Han et al., 2004; Landthaler et al., 2004). The resulting stem-loop structure, the pre-miRNA, is about ~65 nucleotides long and is exported to the cytoplasm. There, another RNase III enzyme, Dicer, aided by its own double-stranded RNA binding partner, cleaves the pre-miRNA approximately two helical turns, that is, ~21 base pairs, from the site of Drosha cleavage. Dicer cleavage liberates an imperfect duplex comprising the mature miRNA paired to the miRNA* strand, which derives from the other arm of the stem of the pre-miRNA. The arms of both the pri-miRNA and pre-miRNA stem-loop structures are imperfectly base-paired, containing G:U wobble pairs, single-nucleotide insertions, and outright mismatches. Consequently, the miRNA:miRNA* duplex contains imperfections in its pairing. These imperfections cause one strand of the duplex to be less stably paired at its 5′ end (Khvorova et al., 2003; Schwarz et al., 2003)—this is the miRNA strand. In the process of loading the miRNA into the RNA binding pocket of a member of the Argonaute family of proteins, the miRNA is separated from the miRNA* strand and the miRNA* is degraded. At last, a functional miRNA-Argonaute protein complex is born.

The precision of Drosha-DGCR8 cleavage is crucial for the fidelity of miRNA maturation: if the position of the Drosha cut is shifted by a single nucleotide on the pri-miRNA, then Dicer cleavage, too, will be shifted, and the final miRNA will have different 5′ and 3′ ends. Now imagine that the single-nucleotide shift inverts the relative stabilities of the 5′ end of the miRNA and of the miRNA* strand. In this event, the wrong strand may be chosen as the mature miRNA. Even if the relative stabilities are preserved and the correct strand is loaded into a functional protein-RNA complex, the shift in the 5′ end of the miRNA will reposition the seed sequence, redefining the set of mRNAs it regulates.

How then do Drosha and DGCR8-Pasha collaborate to excise pre-miRNA so precisely from pri-miRNA transcripts? As there is no strong sequence bias in pri-miRNAs, some structural feature of the RNA must determine the site of Drosha cleavage. The two most obvious distinguishing features are the extremities of the miRNA stem, that is, the flanking single-stranded RNA segments at the base of the hairpin and the terminal loop at its top.

Initially, Drosha was thought to cut the stem by measuring two helical turns from the loop (Zeng et al., 2005). Enter Han et al. (2006) in this issue with their report demonstrating that the terminal loop is unlikely to be the reference point for the molecular ruler that positions the site of cleavage. The terminal loop can be replaced by single-stranded RNA with no major effect on pri-miRNA processing, but the single-stranded RNA segments flanking the base of the stem are indispensable for Drosha cleavage. Han and colleagues show that deleting these single-stranded regions or converting them to double-stranded RNA by annealing a synthetic oligonucleotide to them greatly impairs the conversion of pri-miRNA to pre-miRNA (see also Zeng and Cullen, 2005). Modifying the length of the base of the stem also shifts the cleavage site. So it seems that the molecular ruler is anchored by the junction between the 5′ and 3′ single-stranded segments and the base of the double-stranded stem. Drosha or DGCR8-Pasha must then recognize this junction of single-stranded and double-stranded RNA and count up ~11 bp, one helical RNA turn, to the scissile phosphodiester bond. As DGCR8, but not Drosha, can be crosslinked to pri-miRNA, DGCR8-Pasha is the better candidate for the molecular ruler.

Yet, when Han et al. fed the Microprocessor an "inverted hairpin" in which the single-stranded RNA segments at the base of the stem were replaced by a loop and the loop was replaced by two single-stranded RNA segments, they observed that the correct cleavage site was nonetheless selected. For this pri-miRNA, the Microprocessor appears to measure from the loop. What precisely does DGCR8-Pasha see if both single-stranded tails and a loop suffice to anchor the complex? One explanation is that the terminal loop and the basal single-stranded RNA segments are both unpaired, predicting that a highly structured loop—such as a GNRA tetraloop—would not anchor accurate processing. For a more open loop, DGCR8 may bind to the loop-to-stem junction positioning the Drosha processing center ~11 bp away.

How can DGCR8-Pasha differentiate between binding to the junction of the single-stranded RNA segments and the base of the stem junction, thereby accurately defining one end of the miRNA, and counterproductive binding to the junction of the loop and stem, which will promote abortive processing? Two determinants seem to favor productive binding. First, DGCR8 prefers the junction between flexible single-stranded RNA and a double-stranded stem; a small loop will always be more constrained than single-stranded RNA. Second, the authors' large-scale computational analysis of human and *Drosophila* pri-miRNAs suggests that most pri-miRNAs contain internal bulges or weakly paired bases ~11 bp from the loop-to-stem junction, that is, at abortive Drosha cleavage sites. Such sites may act to deter inappropriate cleavage by Drosha.

These results have several implications. The seemingly dual abilities of DGCR8 to bind to both single-stranded RNA and double-stranded RNA may allow it to bind cooperatively to the pri-miRNA. Cleavage by Drosha, which separates the two single-stranded segments from the double-stranded stem of the pre-miRNA, might then decrease the affinity of DGCR8 for all three reaction products, facilitating their release.

Computationally, our new understanding of how DGCR8 positions Drosha to promote productive pri-miRNA cleavage should help in the search for new miRNA genes and in the design of artificial miRNA genes. To date, successful algorithms for finding miRNAs have relied on phylogenetic conservation because simply searching the genome of a plant or animal for 65 nucleotide-long hairpins yields mainly false-positive results. The phylogenetic conservation approach is powerful but cannot find species-specific miRNAs and perhaps may not even find primate-specific miRNAs. The search for miRNAs that have not been well conserved through evolution may be facilitated by seeking sequences capable of folding into a structure predicted to be bound by DGCR8-Pasha and to promote Drosha cleavage ~11 bp from the junction of a stem with single-stranded RNA tails (while discouraging Drosha cleavage ~11 bp from a terminal loop). That kind of algorithm may finally allow us to ask of the genomes of the Earth's animals: Are any of miRNAs different from yours?

REFERENCES

Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R.H., and Cuppen, E. (2005). Cell *120*, 21–24.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Nature *432*, 231–235.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). Nature *432*, 235–240.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., and Kim, V.N. (2004). Genes Dev. *18*, 3016–3027.

Han, J., Lee, Y., Yeom, K.H., Kim, Y.K., Jin, H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., et al. (2006). Cell *125*, this issue.

Khvorova, A., Reynolds, A., and Jayasena, S.D. (2003). Cell *115*, 209–216.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). Curr. Biol. *14*, 2162–2167.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Cell *120*, 15–20.

Schwarz, D.S., Hutvágner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Cell *115*, 199–208.

Zeng, Y., and Cullen, B.R. (2005). J. Biol. Chem. *280*, 27595–27603.

Zeng, Y., Yi, R., and Cullen, B.R. (2005). EMBO J. *24*, 38–48.

TFIIIC Boxes in the Genome

Lori L. Wallrath[1] and Pamela K. Geyer[1,*]
[1]Department of Biochemistry, University of Iowa, Iowa City, IA 52242, USA
*Contact: pamela-geyer@uiowa.edu

DOI 10.1016/j.cell.2006.05.016

In this issue of *Cell*, Noma et al. (2006) show that *B-boxes* and TFIIIC limit the spread of heterochromatin at the silent *mat* region in the fission yeast genome. Global analysis of TFIIIC distribution revealed dispersed sites of association that coalesce at the nuclear periphery, suggesting that TFIIIC may act as a barrier throughout the genome.

Eukaryotic chromosomes are divided into domains with distinct structural features (Wallrath et al., 2004). Heterochromatic domains are required for chromosome segregation and telomere maintenance as well as for suppressing recombination between repetitive elements. These domains encompass chromosomal regions that have few genes and are assembled into hypoacetylated, regularly spaced nucleosomal arrays containing the epigenetic mark of methylated lysine 9 of histone H3 (H3K9me). Heterochromatic histone modifications generate a condensed chromatin structure that limits the access of transcription factors to target sequences. Interspersed with these transcriptionally repressive domains are euchromatic regions that are gene rich and organized into hyperacetylated irregular arrays of nucleosomes enriched in histone H3 methylated at lysine 4 (H3K4me). Such euchromatic modifications establish chromatin packaging that is accessible to transcription factors. The ability of heterochromatin to propagate in *cis* implies that mechanisms are needed for limiting the spread of silent chromatin. A class of DNA elements, known as insulators, has been implicated in defining the junctions between structural domains (Kuhn and Geyer, 2003). Insulators are found in most eukaryotes, suggesting that these elements have a conserved role in organizing transcriptional domains. Two classes of insulators have been identified that differentially affect transcriptional processes (Kuhn and Geyer, 2003). Enhancer blockers are insulators that prevent enhancer-dependent transcription when placed between an enhancer and promoter. In contrast, barriers are insulators that impede the spread of heterochromatin emanating from an initiation site. In this issue of *Cell*, Noma et al. (2006)