Processing of plant microRNA precursors

Nicolas G. Bologna, Arnaldo L. Schapire and Javier F. Palatnik

Advance Access publication date 11 November 2012

Abstract

MicroRNAs are endogenous small RNAs known to be key regulators of gene expression in animals and plants. They are defined by their specific biogenesis which involves the precise excision from an imperfect fold-back precursor. These precursors contain structural determinants required for their correct processing. Still, there are significant differences in the biogenesis and activities of plant and animal microRNAs. This review summarizes diverse aspects of precursor processing in plants, contrasting them to their animal counterparts.

Keywords: microRNA; precursor; structural determinants; biogenesis; processing; Arabidopsis thaliana; plants; animals

INTRODUCTION

Small RNAs (sRNAs) are major regulators of gene expression in animals and plants. They are usually ~19–24 nt long and processed from double-stranded RNAs (dsRNAs) by Rnase-III enzymes [1, 2]. Plant sRNAs can be classified into different groups such as microRNAs (miRNAs), small interfering RNAs (siRNAs), trans-acting siRNAs, natural antisense-siRNAs and the recently discovered diRNAs, which are generated after DNA lesions [3–6]. Although each class possesses its own characteristics in terms of their biogenesis and size, all of them are ultimately incorporated into RNA silencing complexes, which contain a member of the ARGONAUTE (AGO) family as a main component [3]. The sRNAs guide AGO proteins to target RNAs using base complementarity as a search tool. In turn, AGO complexes regulate gene expression at transcriptional or post-transcriptional level [2].

MiRNAs are an appealing group of sRNAs, as they originate from endogenous loci and regulate other target RNAs than those generating the sRNA [2, 7]. Plant miRNAs are usually around 21–22 nt long and mediate gene silencing at post-transcriptional level [8, 9], which entails the (i) endonucleolytic cleavage (slicing) and/or (ii) transcriptional repression of a target mRNA. The importance of miRNA-mediated regulation is evident by the severe pleiotropic phenotypes in miRNA-deficient mutants [10–13].

The miRNA-binding sites in target RNAs usually have an extensive sequence complementarity to the miRNAs and most experimentally validated targets contain five mismatches or less [14–16]. MiRNAs that are conserved during evolution generally regulate transcription factors that control plant development [17]. Plant miRNAs regulate few targets, usually duplicated members of one protein-coding gene family, yet, disruption of this regulation causes severe defects [2]. In contrast, animal miRNAs regulate many target genes through the base pairing of their 6-nt seed region [18]. However, the biological role of this regulation is unclear in most cases.

Although miRNAs mediate diverse aspects of development and physiology in both plants and animals, there are substantial differences between them,
suggesting they have appeared independently in the two lineages [19]. Still, plant and animal miRNAs are processed from precursors containing imperfect stem–loops and generally regulate longer target mRNAs in both systems. Here, we focus on the processing of plant miRNA precursors and compare them with their animal counterparts.

**MiRNA BIOGENESIS PATHWAY: AN OVERVIEW**

Most plant miRNAs are generated from their own transcriptional units. They are transcribed by RNA polymerase II and then capped, spliced and polyadenylated [20, 21]. Although many animal miRNAs are derived from introns or untranslated regions of coding messages or primary transcripts containing tandem precursors [22, 23], most plant miRNA-encoding loci comprise independent, non-protein-coding transcription units [24]. However, there are some known examples of transcripts harboring tandem precursors in plants [25, 26] and precursors located in mRNA untranslated regions [27].

The miRNA primary transcripts contain an internal stem–loop secondary structure (miRNA precursor) with the miRNA located in one of the arms. The miRNA processing machinery recognizes structural determinants in the miRNA precursors and produces staggered cleavages in the dsRNA, separated ~21 nt of each other. These cuts release the miRNA together with the opposing fragment of the precursor that is interacting with it, called miRNA* (Figure 1A).

The core component of the miRNA processing machinery in *Arabidopsis* is DICER–LIKE1 (DCL1), which is the RNAse type III that produces all cuts in the miRNA precursors [10, 11, 28]. DCL1 is assisted by the dsRNA-binding protein HYponastic LEAVES1 (HYL1) [29, 30] and the C2H2 zinc–finger protein SERRate (SE) [31, 32] (Figure 1). Both HYL1 and SE, improve the efficiency and precision of cleavage by DCL1 [33].

Although not lethal, *hyl1* null mutations severely impair miRNA maturation, which in turn causes developmental defects [30, 34, 35]. Mutations in *SE* accumulate high levels of miRNA primary transcripts and less mature miRNAs [31, 32]. These mutants exhibit also general mRNA splicing defects [36]. *ABH1/CBP80* and *CBP20*, which encode subunits of the nuclear CBC also display general mRNA splicing defects and accumulate more mature primary transcripts [36, 37]. Thus, dual roles in splicing and miRNA processing distinguish SE and CBC from HYL1, which is more specialized in miRNA biogenesis.

DCL1 and HYL1 co-localize in subnuclear regions that have been termed dicing bodies or D-bodies [38–40]. Analysis of miRNA primary transcripts revealed that they are also recruited to these bodies, suggesting that these regions function as miRNA-processing centers in plants [38, 39]. The DCL1-/HYL1-containing regions have similar features as the Cajal bodies, which are known to participate in RNA metabolism (reviewed in [41]). However, while the DCL1-/HYL1-containing bodies contain SmD3 and SmB [39], they do not express coilin [38, 39, 42], a main marker of the Cajal bodies [43].

DAWDLE, a DCL1 interacting protein, is thought to stabilize miRNA primary transcripts until they are processed by DCL1 [44]. HASTY, a homolog of the animal Exportin5, also contributes to the levels of certain miRNAs [45]. However, whereas Exportin5 transports animal pre-miRNAs to the cytoplasm (see below), the molecular role of HASTY is unclear as all processing steps occur in the plant nucleus. Still, HASTY might be associated with other cargo, such as the miRNA/miRNA*.

Recently, it has been reported that the lack of TOUGH, an RNA-binding protein, reduces the accumulation of miRNAs and siRNAs in vivo [46]. TOUGH binds to miRNA precursors in vivo and might aid in the recruitment of primary transcripts or contribute to the efficiency of the DCL1 processing complex. Finally, after the processing of the precursors, the miRNA/miRNA* duplex is released. The 3’-ends of both miRNA and miRNA* are 2’-O-methylated by the nuclear protein HEN1, preventing its degradation [47, 48] (Figure 1A).

In contrast to plants, miRNA biogenesis is compartmentalized in animals. First, primary transcripts are trimmed in the nucleus to separate the stem–loop precursor from the rest of the transcript. This process is achieved by a Microprocessor complex formed by an RNase III-like enzyme termed Drosha and the dsRNA-binding protein DiGeorge syndrome critical region gene 8 (DGCR8; Pasha in *Drosophila melanogaster* and *Caenorhabditis elegans*) [49–51]. The fold-back precursors are then translocated into the cytoplasm by Exportin 5 [52, 53]. Once there, these stem–loops are cleaved by a Dicer, releasing ~21-nt miRNA duplexes [54–56] (Figure 1B). Dicer acts in
concert with other dsRNA-binding proteins: Loquacious, that contains three dsRNA-binding domains (dsRBDs) [57–59] in flies, and HIV trans-activator RNA-binding protein (TRBP) [60, 61] and protein activator of the interferon induced protein kinase (PACT) [62] in humans. In contrast to plants, animal miRNAs are usually not methylated. Still, methylation by HEN1 orthologs has been observed in other classes of sRNAs, such as the single-stranded piRNAs and certain siRNAs [48, 63–65].
Structural determinants of miRNA precursors

The processing of the fold-back precursors by the RNase III complexes causes the release of miRNA/miRNA* duplexes (Figure 1). Although the miRNA is incorporated into an AGO complex, the miRNA* is generally degraded. The precision in the position of the cuts along the precursor is of key importance, as they determine the sequence of the miRNA and therefore its target specificity. Of special importance is the selection of the position for the first cleavage reaction because the second cut is usually performed by measuring a fixed distance from the end of the precursor. The evidence accumulated so far, indicates that the secondary structure of the fold-back precursor directs the activity of the processing machinery represented by Drosha and DCL1 complexes. In some, the biogenesis of the miRNAs could proceed by other pathways. For example, a few recently evolved miRNAs have been shown to depend on DCL4 rather than DCL1 in *Arabidopsis thaliana* [27] and the activity of Drosha can be bypassed by the splicing machinery, generating mirtrons [66, 67], which have been also described in plants [68, 69]. The biogenesis of most animal and plant miRNAs, however, is canalized through Drosha and DCL1, respectively.

**Animal precursors**

A typical miRNA precursor in animals comprises a stem of approximately three helical turns (~65 nt), a terminal loop and long single-stranded RNA (ssRNA) flanking sequence below the fold-back structure (Figure 2A). The stem corresponds to the miRNA/miRNA* plus an additional ~11 nt of a lower stem located below the miRNA. Experiments performed *in vitro* have shown that the terminal loop is dispensable for processing the miRNA precursors, whereas the region below the miRNA is essential [70, 71]. Furthermore, it was demonstrated that the distance from the ssRNA basal segments to miRNA/miRNA* duplex is critical for cleavage site selection [70].

The model for the biogenesis of animal miRNAs suggests that DGCR8/Pasha recognizes the precursor by anchoring at the ssRNA–dsRNA junction while interacting with the precursor stem. After this initial recognition, the processing center of Drosha might be located, performing the first cleavage 11 bp away from the ssRNA–dsRNA junction and releasing the fold-back from the rest of the transcript [23, 70]. The second cleavage can then be performed by Dicer in the cytoplasm by recognizing the free 3′-end of the precursor stem and cutting 22 nt (approximately two helical turns) from this end [57].

**Plant precursors**

In contrast to the defined *in vitro* systems used to analyze miRNA biogenesis in animals (e.g. [70, 71]), most of the studies carried out in plants used transgenic plants expressing miRNA precursors [72–74]. In many plant precursors, a single change in the lower stem of 15 nt below the miRNA is sufficient
to completely abolish its processing [72–75]. A random mutagenesis approach performed on miR172a precursor revealed that point mutations located in the lower stem significantly affected its processing, whereas mutations located in the terminal loop were largely neutral [72]. Interestingly, these unbiased mutations that affected miRNA processing identified by this approach also destabilized the precursor structure [72].

A systematic analysis of plant precursors revealed that many of them have a lower stem of ~15 nt below the miRNA/miRNA* that follows the presence of a large bulge [72–74] (Figure 2B). The current model implicates the recognition of this lower stem to position the initial DCL1 cleavage event. The second cut is produced in the stem at 21 nt of this first cut. These structural determinants are similar but not identical to those identified in animal miRNA precursors (Figure 2C).

**Loop-to-base processing of plant precursors**

The animal stem–loop structures of miRNA precursors have highly uniform sizes of ~65 nt, whereas in plants there is a wide range of precursors from 50 to >500 nt [76] (Figure 3, inset). This heterogeneity in the precursor structures might also reflect differences in the processing pathways. The precursors of miR159 and miR319 are among those with long fold-backs [77–79].

Detailed mutagenesis studies revealed that the processing of these precursors begin with a cleavage next to the loop [80]. DCL1 then continues to cut the precursor three more times at 20- to 22-nt intervals in a ‘loop-to-base’ direction, until the miRNA is finally released [80, 81] (Figure 3). Removing the lower stem below the miRNA/miRNA* did not affect miR319 processing. On the contrary, modifying the upper part of the precursor severely impaired the biogenesis of miR319 [80].

The processing of plant precursors by multiple DCL1 cuts has been also observed in other cases [28, 82] and so far, is a distinctive feature not present in animal precursors. In principle, the processing of miR319 and miR159 precursors might lead to the accumulation of several miRNAs. The experimental evidence so far indicates that miR319 and miR159 have clear biological roles in plant development [78, 80, 83–85]. In turn, miR319 and miR159 are highly conserved across the plant kingdom and copies of miR319 can even be found in mosses, also being processed in a loop-to-base direction, indicating their ancient origin [80, 81, 86, 87].

The other sRNAs can be detected in vivo associated with AGO proteins [88]; however, their biological function remains to be elucidated. The secondary structure of these long precursors contains information regarding the final level of the different potential sRNAs [80]. It seems that the presence of large number of bulges in certain regions of the stem prevents the accumulation of those sRNAs [80].

**Young miRNAs**

It was found that some newly evolved miRNA fold-backs in plants show complementarity with their target genes that extended beyond the miRNA/miRNA* region [24, 27, 89–91]. This observation strongly suggests that many young MIRNA genes derive from inverted duplication events of their targets, giving rise to ‘proto-miRNAs’ [24, 27, 89–91].

However, these newly formed RNA structures do not necessarily have the requirements for miRNA biogenesis as described above and some of them might be processed by different DCLs (reviewed in [2]). In this model, the dsRNA with perfect base pairing formed after a gene duplication event is initially processed by DCL3 and DCL4 to release several siRNAs (or proto-miRNAs). During evolution, the accumulation of mutations in the fold-back arms would lead to changes in the secondary structure and shortening of the hairpin stem, and finally, the processing will be channeled through the miRNA-specific DCL1.

In accordance with this hypothesis, mir822, miR839 and miR869 fold-backs are processed by DCL4 rather than DCL1 [92, 93]. DCL3 has also shown to process conserved miRNAs of *Arabidopsis* [94], as well as certain precursors in rice to yield long-miRNAs (24 nt) [95]. For example, in *Arabidopsis* the activity of DCL3 causes the accumulation of a 24-nt species for certain miRNAs in addition to the main 21-nt class. This is most obvious in inflorescences, which is also the tissue where DCL3 transcripts accumulate [94]. The processing of miR1850 in rice leads to 21- and 24-nt miRNA species by the sequential action of DCL1 and DCL3, respectively. Furthermore, DCL1 and DCL3 can act in parallel on several other precursors as well [95]. It is certain that these recently evolved miRNAs contribute to the heterogeneity of plant precursors (Figure 3, inset).
Plant miRNA precursors are much more heterogeneous than their animal counterparts. In part, this heterogeneity resides in the base-to-loop and loop-to-base processing mechanisms that generate them, as well in the multiple cuts that can be produced by DCL1 to release the miRNAs. Young miRNAs, formed by inverted duplication of protein-coding sequences, contribute to the large variation of shapes and sizes of the plant precursors. Still, a more systematic approach is needed to understand the extension of the complexity of miRNA processing in plants. Interestingly, it has been recently shown that precursors containing a structural asymmetry due to the presence of bulges in either the miRNA or the miRNA* will then generate sRNAs capable of transitivity [96]. It has been suggested that AGO1 complexes with these sRNAs recruit RDR6 and SGS3, which in turn generate dsRNA from their targets and secondary siRNAs [96], therefore, linking the processing of the miRNAs with their activity. It would be interesting to determine the molecular mechanisms and the causal relationships between the processing of the precursor and the activity of the sRNAs.

**Key Points**

- Processing of plant microRNA precursors.
- Comparison of animal and plant microRNA biogenesis.
- Structural determinants of animal and plant microRNA precursors.

**FUNDING**

Grants from the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) (to J.P.).
References


40. Song L, Han M-H, Lesicka J, et al. Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a
nuclear body distinct from the Cajal body. *Proc Natl Acad Sci USA* 2007;104:5437–42.


