MicroRNAs directing siRNA biogenesis

Bonnie Bartel

MicroRNAs have established roles in negatively regulating messenger RNAs. Two plant microRNAs have recently been shown to target certain non-protein-coding RNAs for cleavage, adding a new dimension to the known roles of these tiny riboregulators.

© 2005 Nature Publishing Group http://www.nature.com/nsmb

MicroRNAs (miRNAs) are endogenous ~21-nucleotide RNAs processed by Dicer-like ribonucleases from imperfectly paired hairpin precursor RNAs found in plants and metazoans[1]. Short interfering RNAs (siRNAs) are similarly sized and also require Dicer for biogenesis, but are found more widely in eukaryotes. siRNAs are derived from perfectly paired double-stranded trigger RNA molecules that can be endogenous or derived from introduced RNAs, transgenes or viruses[1]. Both miRNAs and siRNAs are incorporated with Argonaute proteins into complexes—often referred to as RNA-induced silencing complexes, or RISCs—where they provide specificity as these complexes cleave, destabilize or otherwise silence mRNA targets (**Fig. 1**). In addition, some siRNAs direct chromatin modifications that result in transcriptional or heterochromatic silencing (a role that has not been ascribed to miRNAs). Now, in a recent paper in *Cell*, Allen *et al.*[2] show that certain plant miRNAs function in the biogenesis of a special class of siRNAs, the *trans*-acting siRNAs[3,4]. This intriguing observation provides an important addition to the known roles of miRNAs and reinforces the notion that we are only beginning to understand the myriad functions of small RNAs.

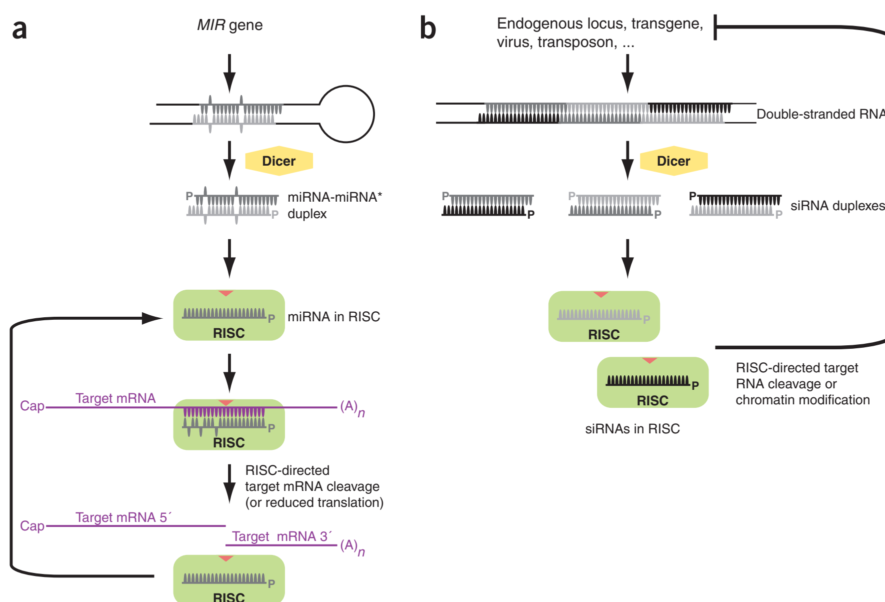

Figure 1 MicroRNAs and short interfering RNAs. (**a**) MicroRNA primary transcripts include imperfectly paired hairpins from which the miRNA-miRNA* duplex is processed by a Dicer ribonuclease. Within the RISC, the miRNA pairs with a target mRNA and an Argonaute protein cleaves the target mRNA[20]. Alternatively, the target mRNA may be destabilized by another mechanism or less efficiently translated. (**b**) siRNAs result from Dicer action on long double-stranded RNA molecules derived from endogenous loci, viruses, transposons or transgenes. siRNAs are incorporated into the RISC, where they direct target RNA cleavage or chromatin modification of the locus from which they were transcribed, or very closely related loci. For references, see ref. 1.

Distinguishing miRNAs and siRNAs

Although similar in size, siRNAs and miRNAs have distinct characteristics. Individual siRNAs are typically less conserved than are miRNAs, likely reflecting the different origins and operational modes of the two small RNAs. siRNAs generally recognize targets with perfect complementarity, a necessary consequence of being derived from the loci that they regulate or from very closely related loci (**Fig. 1b**). MicroRNAs, in contrast, typically are derived from loci distinct from those that they regulate. In general, plant miRNAs are nearly perfectly complementary to a single site in their target mRNAs[5]. Animal miRNAs show less than perfect complementarity to target regions, which are often found in 3′ untranslated regions of mRNAs[1,6]. This difference in the extent of miRNA-target complementarity may partially account for the propensity of plant miRNAs to direct target mRNA cleavage (**Fig. 1a**), whereas animal miRNAs are more likely to silence by other mechanisms[1,7].

MicroRNAs and their targets are usually evolutionarily conserved. Most *Arabidopsis thaliana* miRNA families have obvious homologs in rice[8], and several miRNA-mRNA target pairs are conserved all the way to primitive land plants[9,10], indicating an ancient origin of miRNAs in the plant lineage. Similarly, numerous miRNAs are conserved in the metazoan lineage; the miRNA fervor was presaged not by the identification of the first miRNA[11], but rather when the second discovered roundworm miRNA[12] was found to have vertebrate and *Drosophila melanogaster* homologs[13]. No miRNAs conserved between plants and animals have been found, however, suggesting the intriguing possibility that these regulators arose twice during the independent development of multicellularity in plants and animals.

Trans-acting siRNAs mimic miRNAs

The tidy division of labor between miRNAs and siRNAs was muddied, however, with the discovery of *trans*-acting siRNAs (ta-siRNAs)[3,4]. In the initial isolations of *Arabidopsis* small RNAs, many miRNAs distinguished themselves from siRNAs simply because they were repeatedly cloned, whereas siRNAs were generally singletons[14]. However, a few siRNAs were cloned repeatedly. These exceptions seemed to be derived from loci encoding multiple siRNAs, and included a pair with two-nucleotide-offset reverse complementarity[14]. These unusual siRNAs were subsequently found to

*The author is in the Department of Biochemistry and Cell Biology, Rice University, 6100 Main Street, Houston, Texas 77005, USA.
e-mail: bartel@rice.edu*

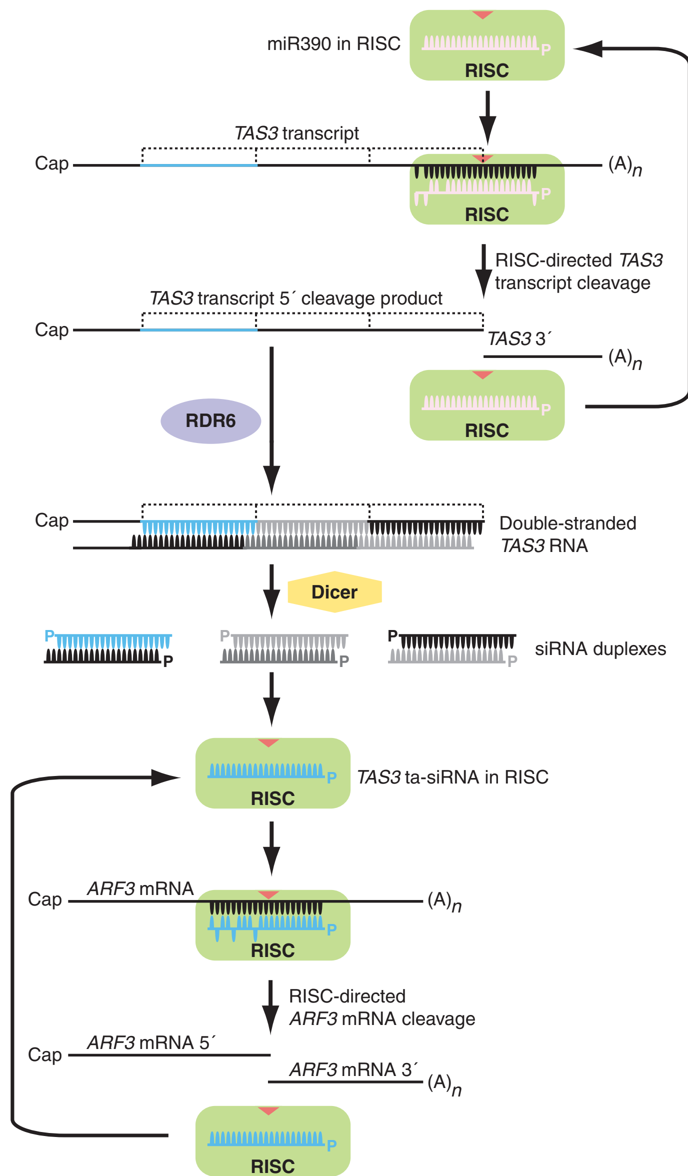

Figure 2 MicroRNAs can direct *trans*-acting siRNA phasing. miR390 (pink) is incorporated into a RISC and targets cleavage of the *TAS3* transcript, which does not encode a protein but rather gives rise to *trans*-acting siRNAs. The *TAS3* cleavage product is reverse transcribed by an RNA-dependent RNA polymerase (RDR6), resulting in a double-stranded RNA that can be processed by a Dicer-like protein into pairs of 21-nucleotide siRNAs with two nucleotide 3′ overhangs (only three pairs are shown for simplicity; more are encoded by *TAS3*). Because of the positioning of the miR390-directed cleavage, these siRNAs include one (blue) that is in the correct register (highlighted by dashed brackets) to direct cleavage of the *ARF3* mRNA[1]. Thus, this siRNA acts in *trans* on a locus distinct from the *TAS3* locus in which it is encoded.

© 2005 Nature Publishing Group http://www.nature.com/nsmb

be part of one of several ta-siRNA clusters. Members of these clusters direct cleavage of target messages encoded in distant loci[3,4], thus acting in *trans* similarly to plant miRNAs. Like miRNAs, particular ta-siRNAs can target multiple members of a gene family. However, ta-siRNAs are clearly distinct from miRNAs. Multiple siRNAs with 21-nucleotide phasing accumulate from a single *trans*-acting siRNA locus[3,4], whereas generally only a single miRNA accumulates per hairpin precursor. Moreover, an RNA-dependent RNA polymerase (RDR6) is required for biogenesis of ta-siRNAs, but not miRNAs[3,4], suggesting a transcript is reverse transcribed into a double-stranded RNA and then processed by Dicer to produce the effective siRNA moieties. Indeed, a Dicer-like enzyme is required for the accumulation of ta-siRNAs[3,4], as it is for miRNAs[14,15].

A new role for miRNAs

That the same siRNA could be repeatedly cloned[14] and that the ta-siRNAs were found in 21-nucleotide register[3,4] suggested that Dicer was acting on long double-stranded RNA with a defined end. The requirement for RDR6 (ref. 3,4) suggested how the transcript could be converted into a double-stranded molecule. But how was the end defined? The missing link was uncovered when two *Arabidopsis* miRNAs, miR173 and miR390, were shown to target primary transcripts encoding ta-siRNAs[2]. Remarkably, the sites of miRNA complementarity within the precursor ta-siRNA transcripts seem to set the register of the 21-nucleotide siRNAs generated from the locus[2]. For example, Allen *et al.*[2] show that as with typical miRNA-directed cleavage, miR390 targets a RISC to the *Trans-Acting siRNA3* (*TAS3*) transcript. RISC cleavage results in two products, one retaining the 5′ cap and the other retaining the polyadenosine tail, and the RDR6 polymerase initiates transcription at or near the 3′ end of the *TAS3 5*′ cleavage product (**Fig. 2**). Just how RDR6 is recruited to the transcript and initiates transcription, and whether the RDR6 product is cleaved directly by a Dicer-like protein or whether further processing precedes dicing, are not yet known. In the simplest model, the double-stranded *TAS3* RNA serves as the substrate for end-dependent, sequential Dicer-like cleavages of the transcript into short duplexes. Setting the correct cleavage register is important because out-of-register processing would generate siRNAs with insufficient complementarity to their target. Functional ~21-nucleotide siRNAs generated through correct ta-siRNA biogenesis would then be loaded into the RISC like other siRNAs and miRNAs and serve as guides for target mRNA cleavage (**Fig. 2**).

Interestingly, miR173 and miR390 differ in the positioning of their cleavage sites of the primary transcripts; miR390-directed cleavage is 3′ of the siRNA cluster (**Fig. 2**), whereas miR173-directed cleavage is 5′ of the siRNA clusters in the corresponding *TAS1* and *TAS2* transcripts[2]. Thus, it may be that either uncapped or non-polyadenylated RNAs can serve as signals for RDR6 processing.

Although miR173 has been found only in *Arabidopsis*[15], miR390 has been cloned from *Arabidopsis*[16] and rice[17] and is present in fern and moss species as well[10]. Moreover, the *TAS3* transcript upon which miR390 acts is conserved in rice, along with the corresponding mRNA targets of the miR390-dependent *TAS3* siRNAs[2]. This conservation suggests that setting the phasing of ta-siRNAs is an ancient function of miRNA in plants. Even before the work of Allen *et al.*[2], miR173 and miR390 were distinguished by being among the few

cloned *Arabidopsis* miRNAs for which mRNA targets had not been suggested (miR173) or experimentally validated (miR390). However, a conventional mRNA target (encoding a receptor-like kinase) of rice miR390 was recently validated[17], so miR390 may have dual roles regulating both mRNAs and loci encoding ta-siRNAs, and perhaps additional, yet-to-be-discovered ta-siRNA loci also will be targeted by miRNAs with mRNA targets. It is not yet clear whether there is a biological relationship between the receptor-like kinase mRNAs directly targeted by miR390 (ref. 17) and the auxin response factor mRNAs (*ARF3*/*ETTIN* and *ARF4*) targeted by the ta-siRNAs generated after miR390-directed cleavage of the siRNA precursor transcript[2], or whether one or both of these targets is conserved along with miR390 in moss[10].

With the work of Allen *et al.*[2], another function is added to the miRNA repertoire—promoting the phasing of ta-siRNAs (**Fig. 2**), which are themselves repressors of target messages. It remains to be seen how many additional functions will be uncovered for small silencing RNAs, but the pace of progress suggests that more surprises await us. Several years after plant miRNAs were shown to direct target cleavage[18], an animal miRNA (widely known for their repressive effects on translation independent of cleavage) was shown to direct mRNA cleavage[19]. It will be interesting to learn whether any animal miRNAs have learned the newest trick of their green cousins. In the case of the miRNAs directing siRNA phasing, sequence specificity is used not only to recognize a target but also to produce a molecule with a specific terminus that provides functionality to the target. Even in insects and mammals, which seem to lack an RNA-dependent RNA polymerase such as would be needed for ta-siRNA biogenesis, miRNAs could be enlisted more broadly to solve other problems requiring a defined, non-capped RNA terminus.

ACKNOWLEDGMENTS
I thank D. Bartel, D. Dugas and A. Woodward for comments on the manuscript. MicroRNA research is supported in the author's laboratory by the US National Institutes of Health (R24-GM069512), the G. Harold and Leila Y. Mathers Charitable Foundation and the Robert A. Welch Foundation (C-1309).

1. Bartel, D.P. *Cell* **116**, 281–297 (2004).
2. Allen, E., Xie, Z., Gustafson, A.M. & Carrington, J.C. *Cell* **121**, 207–221 (2005).
3. Peragine, A., Yoshikawa, M., Wu, G., Albrecht, H.L. & Poethig, R.S. *Genes Dev.* **18**, 2368–2379 (2004).
4. Vazquez, F. *et al. Mol. Cell* **16**, 69–79 (2004).
5. Rhoades, M. *et al. Cell* **110**, 513–520 (2002).
6. Lewis, B.P., Burge, C.B. & Bartel, D.P. *Cell* **120**, 15–20 (2005).
7. Lim, L.P. *et al. Nature* **433**, 769–773 (2005).
8. Bartel, B. & Bartel, D.P. *Plant Physiol.* **132**, 709–717 (2003).
9. Floyd, S.K. & Bowman, J.L. *Nature* **428**, 485–486 (2004).
10. Axtell, M.J. & Bartel, D.P. *Plant Cell* **17**, 1658–1673 (2005).
11. Lee, R.C., Feinbaum, R.L. & Ambros, V. *Cell* **75**, 843–854 (1993).
12. Reinhart, B.J. *et al. Nature* **403**, 901–906 (2000).
13. Pasquinelli, A.E. *et al. Nature* **408**, 86–89 (2000).
14. Reinhart, B.J., Weinstein, E.G., Rhoades, M.W., Bartel, B. & Bartel, D.P. *Genes Dev.* **16**, 1616–1626 (2002).
15. Park, W., Li, J., Song, R., Messing, J. & Chen, X. *Curr. Biol.* **12**, 1484–1495 (2002).
16. Gustafson, A.M. *et al. Nucleic Acids Res.* **33**, D637–D640 (2005).
17. Sunkar, R., Girke, T., Jain, P.K. & Zhu, J.K. *Plant Cell* **17**, 1397–1411 (2005).
18. Llave, C., Xie, Z., Kasschau, K.D. & Carrington, J.C. *Science* **297**, 2053–2056 (2002).
19. Yekta, S., Shih, I.H. & Bartel, D.P. *Science* **304**, 594–596 (2004).
20. Liu, J. *et al. Science* **305**, 1437–1441 (2004).

Spliceosome assembly in yeast: one ChIP at a time?

Timothy W Nilsen

When and how the spliceosome assembles on its substrates in living yeast cells has been unclear. It is now evident that the 'when' is during transcription; the 'how' remains debatable.

Intron removal is catalyzed by a massive, remarkably complex ribonucleoprotein assemblage known as the spliceosome[1,2]. Although they have been extensively studied *in vitro*, the dynamics of spliceosome assembly *in vivo* have remained largely unexplored. Two recent reports[3,4] have used chromatin immunoprecipitation (ChIP) to address this issue in budding yeast. Collectively, the data clearly show that the splicing machinery is recruited to intron-containing transcripts cotranscriptionally and that the cap-binding complex (CBC) is important for this recruitment. Less definitive is the argument made in both reports that spliceosome assembly *in vivo* proceeds in a stepwise manner.

Current models of spliceosome assembly come in two flavors: the first model holds that assembly involves strictly ordered, piece-by-piece construction and the second model posits that the spliceosome exists as a preassembled entity. The stepwise model is supported by an extensive array of biochemical experiments wherein discrete intermediates in assembly have been isolated and characterized. The preassembled model is also supported by a substantial amount of experimental data, the most important of which is the isolation and characterization of a nearly complete, catalytically competent preassembled spliceosome from yeast, the 'penta-snRNP'[5]. The two models of spliceosome assembly may not really be in conflict; as a consequence of spliceosomal dynamics, the apparent stepwise assembly could reflect stepwise stabilization of intermolecular contacts rather than stepwise recruitment[6].

So, how does assembly occur *in vivo*? To approach this question, both groups chose to use ChIP, a widely used technique for dissecting chromatin structure that has gained increasing acceptance for the analysis of association of proteins with nascent RNA transcripts. The basic idea is to treat living cells with formaldehyde, a general crosslinking reagent. After lysis, chromatin is sheared and immunoprecipitated with the antibody of choice. At this point, the crosslinks are reversed and recovered DNA is amplified by specific PCR primers. Through judicious choice of primers, it is possible to walk through a region of DNA (gene) to determine the disposition of proteins along its length.

In principle, interpretation of protein-DNA crosslinks is straightforward, although the approach is obviously better at determining the presence rather than the absence of a specific protein. Demonstrating absence is problematic because it is not possible to predict *a priori* the efficiency of crosslinking—the 'crosslinkability'—of any specific intermolecular interaction.

When ChIP is used to analyze the binding of proteins to nascent RNA transcripts, additional complications arise. The methodology is identical to DNA ChIP except that the protein of interest binds to the nascent RNA chain and then directly or indirectly crosslinks to the DNA template (**Fig. 1**). Direct crosslinking can occur if the nascent RNA chain is close to the template, presumably when the binding site

*The author is at the Center for RNA Molecular Biology, Case Western Reserve University School of Medicine, 10900 Euclid Avenue, Cleveland, Ohio 44106-4973, USA.
e-mail: twn@cwru.edu*

© 2005 Nature Publishing Group http://www.nature.com/nsmb