MicroRNA Biogenesis and Cancer

Richard I. Gregory and Ramin Shiekhattar

Gene Expression and Regulation Program, and Molecular and Cellular Oncogenesis Program, Wistar Institute, Philadelphia, Pennsylvania

Abstract

MicroRNAs (miRNA) are a recently discovered family of short non–protein-coding RNAs that negatively regulate gene expression. Recent studies of miRNAs highlight a requirement for cell viability. Posttranscriptional silencing of target genes by miRNAs occurs either by targeting specific cleavage of homologous mRNAs, or by targeting specific inhibition of protein synthesis. We recently identified a multisubunit protein complex termed Microprocessor that is necessary and sufficient for processing miRNA precursor RNAs. Microprocessor contains Drosha, an RNase III endonuclease, and *DGCR8*, a gene deleted in DiGeorge syndrome. We consider recent findings that link miRNA perturbation to cancer. (Cancer Res 2005; 65(9): 3509-12)

The discovery that small (~22 nucleotides), noncoding, double-stranded RNA molecules can mediate the expression of target genes with complementary sequence led to the identification of a large family of evolutionary conserved, regulatory RNAs, dubbed microRNAs (miRNAs; refs. 1, 2). In mammals, hundreds of miRNAs have now been identified, some of which are expressed in a tissue-specific and developmental stage–specific manner. For the few miRNAs whose function has been uncovered, they are important regulators of various aspects of developmental control in both plants and animals (1, 2). In the few years since the inception of this regulatory RNA phenomenon, much progress has been made towards an understanding of the mechanisms by which this occurs and the identification of cellular machinery involved in RNA-mediated silencing (1, 2). In this review, we focus on recent discoveries related to the miRNA biogenesis pathway and discuss the implications for human diseases including cancer.

MicroRNA Transcription and Stepwise Maturation

The majority of the characterized miRNA genes are intergenic or oriented antisense to neighboring genes and are therefore suspected to be transcribed as independent units. The primary transcripts (pri-miRNA) are generated by polymerase II and recent data indicate that pri-miRNAs possess a 5′ 7-methyl guanosine cap and are polyadenylated (3, 4). Probably the best-characterized human pri-miRNA is that of miR-23a~27a~24-2, a 2.2-kb transcript containing three miRNAs. Pri-miR-23a~27a~24-2 is 5′ capped and polyadenylated ~1.8 kb downstream of the 3′ end of miR-24-2. A minimal (~600 bp) polymerase II–dependent promoter was identified at this miRNA gene (3). Interestingly, this promoter, although able to direct polymerase II–dependent transcription, lacks all typical promoter elements normally required for transcription initiation. The significance and generality to miRNA gene promoters of this observation is currently unknown. A subset of human miRNA genes are located within introns of pre-mRNAs. Because these miRNAs have the same orientation as pre-mRNAs, it is likely that they are processed from the introns and one would therefore expect the tissue-specific expression profile of the miRNA to correlate with that of the mRNA. The remaining miRNAs are clustered in the genome predicting a long transcript encompassing several coordinately expressed miRNAs. Irrespective of how different miRNAs originate, the pri-miRNA transcripts are predicted to form specific "hairpin-shaped" secondary structures and are processed to yield a mature 22-nucleotide miRNA.

Previously, it was established that processing of pri-miRNAs in the nucleus is mediated by Drosha, a RNase III endonuclease (5). Drosha asymmetrically cleaves both strands of the hairpin stem at sites near the base of the primary stem loop thus releasing a 60- to 70-nucleotide pre-miRNA that has a 5′ phosphate and a 2-nucleotide 3′ overhang. Specific RNA cleavage by Drosha predetermines the mature miRNA sequence and provides the substrate for subsequent processing events. The pre-miRNAs are transported to the cytoplasm by Exportin-5, in a Ran GTP-dependent manner. The interaction of Exportin-5 with the pre-miRNA intermediate and its subsequent transport requires correctly processed pre-miRNAs with hallmarks of Drosha-mediated cleavage and specific "hairpin" secondary structure (6, 7). Once in the cytoplasm, a second RNase III endonuclease, Dicer, acts on the pre-miRNA. It is thought that the PAZ domain of Dicer recognizes the 2-nucleotide 3′ overhang at the base of the stem loop of the pre-miRNA. Dicer cleaves double stranded RNA 22-nucleotides from the end of the substrate (8). In the case of pre-miRNA, one end of the miRNA has already been predetermined by Drosha cleavage site selection. Therefore, a subsequent cleavage by Dicer releases a 22-nucleotide mature double-stranded miRNA with 5′ phosphates and a 2-nucleotide 3′ overhangs. One strand of the miRNA duplex is subsequently incorporated into an effector complex termed RNA-induced silencing complex that mediates target gene expression (reviewed in detail in ref. 2).

Processing of Primary MicroRNAs by the Microprocessor Complex

As discussed, Drosha plays an essential role in the genesis of miRNAs; however, it was not known whether other Drosha-associated components contribute to the processing of pri-miRNAs. Recently, four groups independently investigated the potential involvement of other proteins in pri-miRNA processing (9–12). To address the question of whether Drosha-associated components contribute miRNA biogenesis, we biochemically purified Drosha from human cells by affinity chromatography

Requests for reprints: Ramin Shiekhattar, Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104. Phone: 215-898-3896; Fax: 215-898-3868; E-mail: shiekhattar@wistar.upenn.edu.

and identified ~20 Drosha-associated proteins by mass spectrometric sequencing (9). Fractionation of the affinity eluate by a gel filtration column revealed one complex (>2 MDa) containing most of the Drosha-associated proteins, and a smaller Drosha-containing complex of ~600 kDa (9). Previously, an *in vitro* assay was established to monitor pri-miRNA processing using *in vitro* transcribed RNAs corresponding particular to pri-miRNAs (13). We cloned a miRNA gene cluster corresponding to miRNAs 17-18-19a-20-19b. After *in vitro* transcription, this fragment, containing five miRNAs, generates a pri-miRNA of ~800 nucleotides. Using this pri-miRNA as a substrate, we assayed the column fractions for processing activity and observed accumulation of a 63-nucleotide pre-miRNA that coeluted with the smaller (~600 kDa) Drosha-containing complex (9). We determined the polypeptide composition of this complex by mass spectrometric sequencing and found two proteins: Drosha and the double-stranded RNA-binding protein DGCR8 (9). We confirmed this interaction by purification of Drosha as a DGCR8-associated protein and showed pri-miRNA processing by the complex. Concomitantly, the Hannon group was investigating the association of a candidate protein (CG1800) that was previously shown to potentially interact with Drosha based on yeast two hybrid data (10). Consistent with our discovery in human cells, when extract from *Drosophila* S2 cells was fractionated by gel filtration, a peak of miRNA processing (of pri-bantam RNA) was observed in fractions corresponding to a protein complex of ~600 kDa, that comprised Drosha and the *Drosophila* orthologue of DGCR8, a protein they named Pasha (partner of Drosha; ref. 10). This complex, comprising Pasha/DGCR8 and Drosha, was named the Microprocessor complex (9, 10). A third group investigated the molecular mechanism of pri-miRNA processing by Drosha, when they fractionated nuclear extract from human cells and assayed it for pri-let-7-a-1 miRNA processing activity, a distinct peak of activity was detected in fractions corresponding to a molecular mass of >700 kDa (11). This peak of activity shifted to ~650 kDa after the nuclear extract was treated with RNase A before gel filtration. This complex most likely corresponds to Microprocessor, and indicates that the integrity of this complex is mediated by protein-protein interactions and does not depend on RNA (11). Together, these three studies provide very strong support for the existence of Microprocessor, a Drosha-containing complex of around ~600 kDa that is conserved from flies to human (see Fig. 1; refs. 9–11).

To assess the role of Drosha-DGCR8 complex in initiation of miRNA processing *in vivo*, we used RNA interference to deplete DGCR8 and Drosha in human cells. Small interfering RNA against Drosha and DGCR8 invoked a pronounced decrease in mature miRNA levels (9). Consistent with Microprocessor's role in processing pri-miRNAs, we also detected an accumulation of pri-miRNA following knock down of Drosha or DGCR8 (9). In accordance with our findings, genetic studies in *Caenorhabditis elegans* using mutant Drosha and mutant Pasha worms (10), together with RNA interference–mediated knockdown of these proteins in *Drosophila* cells (10, 12), *C. elegans* (10), and human cells (10–12), confirmed an obligatory role for these proteins in processing pri-miRNA to pre-miRNA. Although the aforementioned data show that both Drosha and DGCR8 are necessary for pri-miRNA processing, the most compelling evidence for the sufficiency of these proteins in mediating miRNA biogenesis was provided by our ability to reconstitute Microprocessor activity using recombinant proteins (9). Recombinant Drosha and DGCR8 proteins were purified and assayed for pri-miRNA processing. Neither recombinant Drosha nor DGCR8 showed any significant miRNA processing activity. However, addition of the both recombinant proteins together reconstituted the miRNA processing activity to similar levels seen with native complex (9). Moreover, Drosha protein itself displayed a nonspecific RNase activity toward the substrate, which may provide insight into the mechanism of Microprocessor action by demonstrating the requirement of DGCR8 in directing the specific cleavage of pri-miRNA by Drosha. As the DGCR8 protein contains two double-stranded RNA binding domains at its COOH terminus, it is tempting to speculate that the role of DGCR8 in the Microprocessor may be to recognize pri-miRNAs and/or to orient the catalytic RNase III core of Drosha to ensure correct cleavage site selection at the stem of the hairpin RNA structure. Furthermore, DGCR8 has an NH_2-terminal WW domain known to interact with proline-rich peptides. The WW domain of DGCR8 is most likely the interacting surface with the proline-rich NH_2-terminal domain of Drosha. This possibility requires further exploration based on the observation of Han et al., where it seemed that the proline-rich domain of Drosha is dispensable for interaction with DGCR8 (11); however, it should be noted that interpretation of this result may be complicated by the indication that a single Microprocessor complex is formed by multiple molecules of Drosha and/or DGCR8 (11).

Of particular interest is the fact that *DGCR8* was originally identified as a gene that maps to the chromosomal region 22q11.2, a region whose monoallelic deletion accounts for >90% of patients with DiGeorge syndrome, the most common human genetic deletion syndrome that effects around 1 in 3,000 live births (14). The clinical manifestations of the disease are highly variable, with 75% of patients displaying congenital heart defects. Other common features include characteristic facial appearance, immunodeficiency from thymic hypoplasia, and developmental and behavioral problems, including schizophrenia and obsessive-compulsive disorder in adulthood (15). Despite its relatively high incidence, the fact that most cases of DiGeorge syndrome are caused by haploinsufficiency of a typical deleted region (~1.5 Mb) that encompasses around 30 genes has made the task of identifying the particular gene(s) that underlie the pathogenesis a challenge. Given the emerging role of miRNAs in regulating diverse aspects of development, it is tantalizing to hypothesize that haploinsufficiency of DGCR8; thus, defects in miRNA biogenesis may contribute to the widespread developmental abnormalities affecting DiGeorge syndrome patients. In this respect, we suggest that DGCR8 is strong candidate gene for DiGeorge syndrome and are currently investigating the potential role of miRNA perturbation in the genesis and development of DiGeorge syndrome.

Cancer Connections

Insights leading to the understanding of miRNA biogenesis may affect cancer in at least two ways. First, emerging data indicate that dysregulation of miRNAs is associated with certain types of cancer (16–20). Second, exploitation of the therapeutic potential of RNA interference may be achieved through the investigation of how endogenous miRNA and produced and exert their gene regulatory function (21, 22). Chronic lymphocytic leukemia (CLL), the most common form of adult leukemia, can be attributed to a deletion at 13q14 in >50% of cases. Additionally, various other cancers

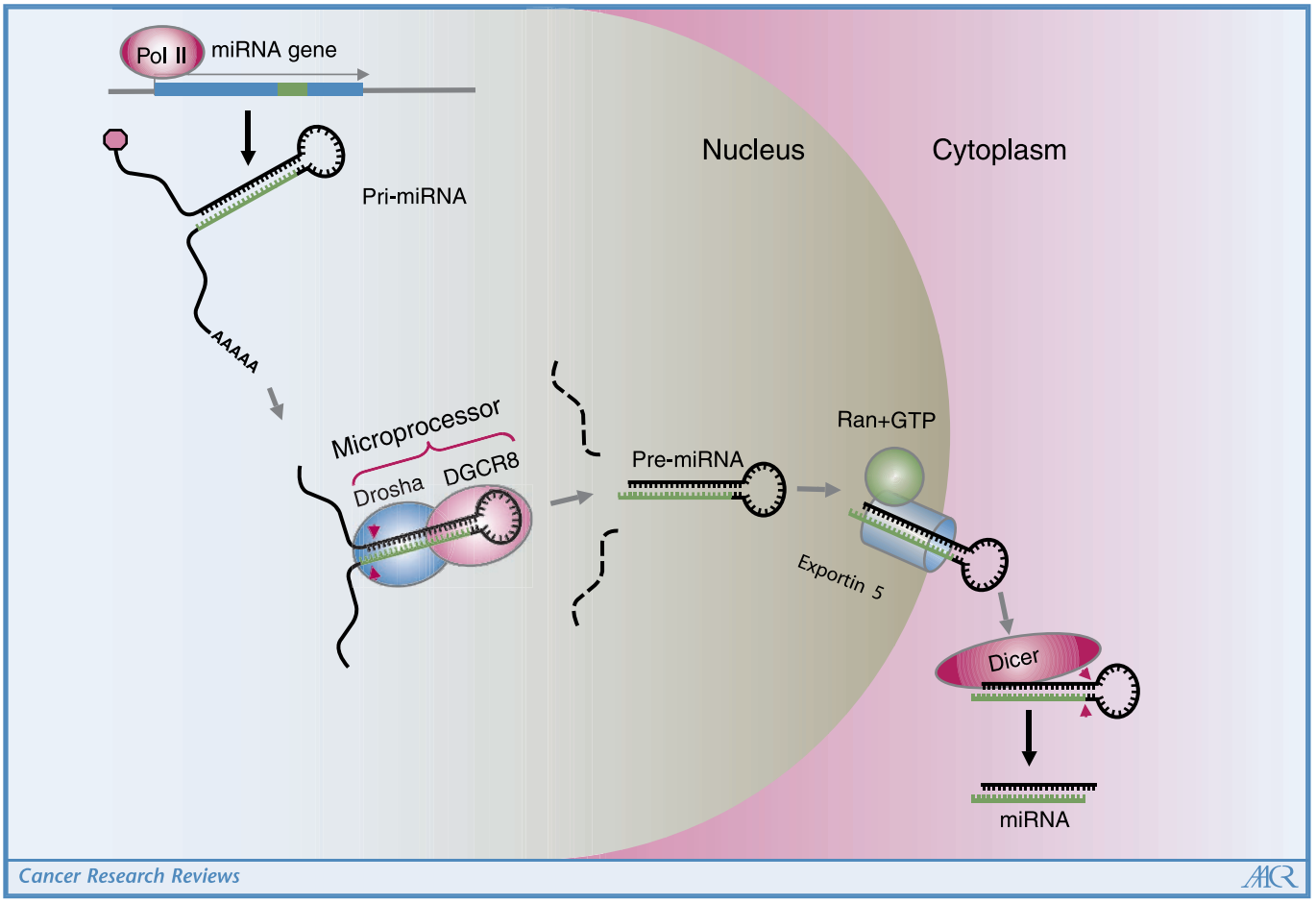

Figure 1. Model of miRNA biogenesis. Long primary transcripts (pri-miRNAs) containing one to several miRNAs are generated by polymerase II and are 5′ capped and polyadenylated. The recently identified Microprocessor complex, comprising Drosha (RNase III endonuclease) and DGCR8 (a double-stranded RNA binding protein) recognize the distinct hairpin secondary structure of the pri-miRNA and specifically cleave at the base of the stem loop releasing a 60- to 70-nucleotide pre-miRNA intermediate enabling Exportin 5–mediated cytoplasmic export where Dicer, a second RNase III endonuclease, cleaves 22-nucleotide from the Drosha cleavage site to yield the mature miRNA. Dysregulation of this processing leading to perturbation in miRNA genesis may have oncogenic consequences.

(including mantle cell lymphoma, multiple myeloma, and prostate cancers) have been linked to varying degrees with 13q14 deletions. Despite considerable effort, none of the known genes located in the deleted region have been shown to lead to CLL. Two miRNA genes, *miR-15* and *miR-16*, map to this deleted region and it was found that in 68% of CLL patients and the majority of prostate cancer cell lines tested, both miRNA genes are deleted or down-regulated. Perhaps even more provocative is the observation that in a subset of CLL samples, accumulation of the (~70 nucleotides) pre-miR-15 intermediate was detected by Northern blotting, pointing to a potential deficiency in miR-15 processing (16). Down-regulation of miR-143 and miR-145 was observed in colorectal neoplasia (17), and expression of the miRNA *let-7* is frequently reduced in lung cancers, a feature that is associated with poor prognosis (18). Increased expression of the precursor of miR-155 was detected in pediatric Burkitt lymphoma (19). In fact, it has been speculated based on cancer-associated alterations in miRNA expression, the observation that miRNAs are frequently located at genomic regions involved in cancers, and their gene regulatory function, that miRNAs may act as both tumor suppressors and oncogenes (20). In addition, although the function of the larger (>2 MDa) Drosha-containing complex we purified is currently unknown, attention should be drawn to our identification of Ewing's sarcoma gene product (EWS) as one component (9). Chromosomal translocations of the *EWS* gene that encodes a putative RNA binding protein, are known to cause Ewing sarcoma as well as neuroectodermal and various other tumors (9). Future investigations will undoubtedly reveal additional links between mechanisms of miRNA-mediated gene silencing and human diseases including cancer.

Acknowledgments

Received 1/28/2005; revised 2/9/2005; accepted 2/11/2005.
Grant support: Jane Coffin Childs Foundation Fund for Medical Research postdoctoral fellowship (R.I. Gregory).

References

1. Novina CD, Sharp PA. The RNAi revolution. Nature 2004;430:161–4.
2. Meister G, Tuschl T. Mechanisms of gene silencing by double-stranded RNA. Nature 2004;431:343–9.
3. Lee Y, Kim M, Han J, et al. MicroRNA genes are transcribed by RNA polymerase II. EMBO J 2004;23:4051–60.
4. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated

transcripts that can also function as mRNAs. RNA 2004;10:1957–66.
5. Lee Y, Ahn C, Han J, et al. The nuclear RNase III Drosha initiates microRNA processing. Nature 2003;425: 415–9.
6. Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U. Nuclear export of microRNA precursors. Science 2004; 303:95–8.
7. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev 2003;17:3011–6.
8. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. Cell 2004; 118:57–68.
9. Gregory RI, Yan KP, Amuthan G, et al. The Microprocessor complex mediates the genesis of microRNAs. Nature 2004;432:235–40.
10. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. Nature 2004;432: 231–5.
11. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev 2004;18:3016–27.
12. Landthaler M, Yalcin A, Tuschl T. The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. Curr Biol 2004;14:2162–7.
13. Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. EMBO J 2002;21:4663–70.
14. Shiohama A, Sasaki T, Noda S, Minoshima S, Shimizu N. Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region. Biochem Biophys Res Commun 2003; 304:184–90.
15. Lindsay EA. Chromosomal microdeletions: dissecting del22q11 syndrome. Nat Rev Genet 2001;2:858–68.
16. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002; 99:15524–9.
17. Michael MZ, O' Connor SM, van Holst Pellekaan NG, Young GP. James Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol Cancer Res 2003;1:882–91.
18. Takamizawa J, Konishi H, Yanagisawa K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 2004;64: 3753–6.
19. Metzler M, Wilda M, Busch K, Viehmann S, Borkhardt A. High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. Genes Chromosomes Cancer 2004;39:167–9.
20. Calin GA, Sevignani C, Dumitru CD, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 2004;101:2999–3004.
21. Ryther RC, Flynt AS, Phillips JA, Patton JG. siRNA therapeutics: big potential from small RNAs. Gene Ther 2005;12:5–11.
22. Stevenson M. Therapeutic potential of RNA interference. N Engl J Med 2004;351:1772–7.

MicroRNA Biogenesis and Cancer

Richard I. Gregory and Ramin Shiekhattar

Cancer Res 2005;65:3509-3512.