The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1

Zhicheng Dong†, Meng-Hsuan Han‡, and Nina Fedoroff§

Biology Department and Huck Institutes of Life Sciences, Pennsylvania State University, University Park, PA 16802

Contributed by Nina Fedoroff, May 1, 2008 (sent for review March 30, 2008)

The results of genetic studies in Arabidopsis indicate that three proteins, the RNase III DICER-Like1 (DCL1), the dsRNA-binding protein HYPOCENTRAL LEAVES1 (HYL1), and the C2H2 Zn-finger protein SERRATE (SE), are required for the accurate processing of microRNA (miRNA) precursors in the plant cell nucleus. To elucidate the biochemical mechanism of miRNA processing, we developed an in vitro miRNA processing assay using purified recombinant proteins. We find that DCL1 alone releases 21-nt short RNAs from dsRNA as well as synthetic miR167b precursor RNAs. However, correctly processed miRNAs constitute a minority of the cleavage products. We show that recombinant HYL1 and SE proteins accelerate the rate of DCL1-mediated cleavage of pre- and pri-miR167b substrates and promote accurate processing.

miRNA | biogenesis | Dicer | Arabidopsis

Micro RNAs (miRNAs) are ~21-nt regulatory RNAs found in viruses, plants, and animals. miRNAs inhibit gene expression by translational repression and by pairing with their target mRNA to promote their cleavage (1, 2). miRNA regulation is known to play an important role in development, stress responses, and carcinogenesis (3, 4). miRNAs are transcribed by RNA polymerase II as long primary transcripts, termed pri-miRNAs, which are capped and polyadenylated (5, 6). In animals, the pri-miRNA, which contains the miRNA sequence embedded within a hairpin, is processed in the nucleus and the cytoplasm sequentially by two RNAase III-family enzymes (7). The class II RNase III Drosha, together with the dsRNA-binding protein (dsRBP) DGCR8/Pasha, cleaves the stem loop of pri-miRNA in the nucleus to a hairpin RNA (pre-miRNA) of ~70 nt (8–11). The pre-miRNAs are transported out of the nucleus by the Ran-binding protein exportin 5 (12–14). Dicer, another class III RNase III, cleaves the pre-miRNA in the cytoplasm to release the ~22-nt miRNA/miRNA* duplex (15–17). Like Drosha, the animal Dicer has dsRBP partners, including Loquacious/R3D1-L in Drosophila (18–20) and TRBP and PACT in mammals (21–23). There is also a Drosha-independent pathway for generating certain intronic miRNAs (24, 25).

The plant miRNA biogenesis mechanism is somewhat different from that of animals. The pri-miRNAs transcripts appear to be RNA polII transcripts in plants, as they are in animals, but the hairpins are substantially more variable in length (26, 27). Plants contain multiple Dicer homologs, termed the Dicer-like (DCL) enzymes. Of the four Dicer homologs in Arabidopsis, DCL1 carries out both cleavage steps in the nucleus to generate miRNAs (28–31). DCL2 produces 24-nt natural antisense siRNAs from complementary overlapping mRNA transcripts involved in the salt stress response (32). DCL3 generates the 24-nt RNA repeat sequence-associated siRNAs that direct heterochromatin formation (33), whereas DCL4 produces 21-nt trans-acting siRNAs that control leaf development (34–36). In addition, DCL4 and DCL2 generate 21- and 22-nt virus-derived small RNAs (vsRNA) redundantly functioning in antiviral defense (37).

Animal Dicers and plant DCL1 enzymes are large proteins of ~220 kDa that contain an RNA helicase domain, a PAZ domain, two RNase III domains, and either one (animal) or two (plant) dsRNA-binding domains (dsRBD) (Fig. L4). Based on the observation that pri-miRNA levels increase and pre-miRNA levels decrease in the weak dcll-9 insertion mutant, DCL1 is believed to be required to cleave pri-miRNA to pre-miRNA (31, 38). Several lines of evidence suggest that the dsRNA-binding protein (dsRBP) HYL1 and a C2H2 zinc-finger protein SE (Fig. L4) are DCL1 cofactors. The mature miRNA levels are low in both hyl1 and se mutants (39–42), and pri-miRNAs accumulate (38, 43). DCL1 and HYL1 recombinant proteins form a complex in vitro (44, 45), and DCL1 and HYL1 has been reported to interact with SE (45). DCL1 and HYL1 colocalize with DCL1 in small nuclear bodies containing pri-miRNAs, as does a fraction of the nuclear SE (38, 46). A protein complex immunoprecipitated by using anti-HYL1 antiserum has been reported to process miR169 pri-miRNAs into mature miRNAs (47). A nuclear methyltransferase, HUA ENHANCER1 (HEN1), methylates the 2’ hydroxyl group of 3’end of mature miRNAs (28, 48, 49). In addition, HASTY, the Arabidopsis homolog of exportin 5, is required for miRNA accumulation and may transport miRNA into the cytoplasm (50).

Although DCL1, HYL1, SE, HEN1, and HASTY have all been identified as playing a role in miRNA biogenesis in Arabidopsis, the detailed mechanism remains unknown. In the present study, we analyzed the ability of recombinant DCL1, HYL1, and SE proteins to process miRNA precursors in vitro. We show that recombinant DCL1 cleaves dsRNA, pre-miRNA, and pri-miRNA to release short RNAs of predominantly 21 nt, requiring only divalent cations and ATP. Both HYL1 and SE stimulate DCL1 activity on both pri- and pre-miRNA substrates and markedly increase the fidelity of cleavage. These results suggest that DCL1, HYL1, and SE form a heteropolymeric complex that catalyzes efficient and accurate processing of pri-miRNA to miRNA.

Results

Preparation of Recombinant DCL1, HYL1, and SE Proteins. Recombinant His-HA-DCL1-FLAG, His-HYL1, and His-SE were expressed in insect cells. To overcome the protease cleavage routinely observed in insect cell extracts expressing recombinant DCL1, we epoite-tagged the DCL1 with a 6× His tag at the N terminus and a FLAG tag at the C terminus (Fig. L4). We obtained intact DCL protein by a two-step affinity purification procedure. Recombinant His-HYL1 and His-SE proteins were purified by using Talon (cobalt) resin, followed by NTA/Ni resin.

Author contributions: Z.D., M.-H.H., and N.F. designed research; Z.D. and M.-H.H. performed research; Z.D. and N.F. analyzed data; and Z.D. and N.F. wrote the paper.

The authors declare no conflict of interest.

See Commentary on page 9851.

1Present address: Department of Molecular Biology, Cell Biology, and Biochemistry, Brown Medical School, Providence, RI 02912.
2Present address: Department of Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, IL 60208-3500.
3Present address: Department of State, 2101 C Street NW, Washington, DC 20520.
4To whom correspondence should be addressed. E-mail: nvf1@psu.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0803356105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA
As judged by SDS/PAGE analysis, >90% of each purified protein preparation had the expected mobility of the intact protein (Fig. 1B). In addition, we confirmed the identity and integrity of the purified recombinant His-HA-DCL1-FLAG protein by Western blot analysis with both anti-FLAG and anti-His sera (data not shown).

**Recombinant DCL1 Processes dsRNA to 21-nt RNA.** We first tested the activity of recombinant DCL1 protein using a dsRNA substrate with 2-nt 3′ overhangs. Consistent with a previous report that used immunoprecipitated DCL1 (51), we observed the release of predominantly 21-nt small RNAs in a reaction that was time- and DCL1 protein concentration-dependent (Fig. 1C and D). The optimal salt concentration for the DCL1-mediated cleavage reaction was found to be 25–50 mM NaCl (Fig. 1E). No activity was observed in the absence of ATP, and the activity was substantially reduced when ATP was replaced by GTP (Fig. 1F). Moreover, DCL1 recombinant protein is able to hydrolyze ATP in vitro (data not shown). We conclude that DCL1 cleavage is ATP-dependent (52, 53).

**DCL1 protein concentration-dependent (Fig. 1 C and D).** The optimal salt concentration for the DCL1-mediated cleavage reaction was found to be 25–50 mM NaCl (Fig. 1E), similar to that observed for human Dicer (52). At 50 mM NaCl, DCL1 requires both a divalent cation and ATP to cleave dsRNA; addition of EDTA inhibited the reaction completely (Fig. 1F). No activity was observed in the absence of ATP, and the activity was substantially reduced when ATP was replaced by GTP (Fig. 1F). Moreover, DCL1 recombinant protein is able to hydrolyze ATP in vitro (data not shown). We conclude that DCL1 cleavage is ATP-dependent (52, 53).

**Recombinant DCL1 Processes pre-miR167b into Mature miRNA.** Because pre-miRNAs have an imperfect dsRNA hairpin structure with a 2-nt 3′ overhang at one end, we reasoned that DCL1 might itself be able to cleave the pre-miRNA to the mature 21-nt miRNA. Because miR167 level were substantially reduced in both hyll and se mutant plants (40, 42, 45), we chose miR167b...
Recombinant DCL1 Processes pri-miR167b into Mature miRNA. To ask whether DCL1 can process pri-miRNA, we incubated DCL1 protein with a synthetic pri-miR167b substrate containing 93-nt pre-miRNA 5′-flanked by an additional 115 and 90 nt on the 5′- and 3′-sides of the pre-miRNAs, respectively (Fig. 3A). We observed 21-nt cleavage products as well as a weaker 19-nt band (Fig. 3B, lane 4). We confirmed that the small RNA products contain both miR167b and miR167b*, the short digestion products were cloned and sequenced (see Methods). After removal of the sequences that do not match the pri-miR167b substrate sequence or that are shorter than 19 nt, we obtained 83 sequences. Among them, nine were identical to mature miR167b, one was a close variant of miR167b, and one was derived from the vicinity of the miR167b* sequence (Fig. 4A). We note that the small RNA databases of Arabidopsis contain, in addition to the 21-nt miR167b sequence, 20-, 22-, 23-, 24-, and 1-nt offset reads (54, 55), suggesting that the shorter and offset products also result from DCL1 cleavage. The remaining 72 sequences were derived from the other part of the pri-miR167b substrate, comprising ~80% of the total. These sequences mostly matched the 3′ end of the pri-miR167b substrate, indicating that DCL1 alone predominantly processes the substrate incorrectly.

HYL1 and SE Increase the Accuracy of pri-miR167b Processing. In view of the reports that the hyl1 and se mutations cause substantial reductions in the accumulation of mature miRNAs (39–41) and that miscleavage of miR163 pri-miRNA is observed in hyl1 mutant plants (43), we next asked whether recombinant HYL1 and SE proteins affect either the rate or the accuracy of DCL1 processing of pri-miR167b. We performed reactions using a fixed DCL1 concentration and increasing amounts of HYL1, SE, or both proteins. When judged by the amount of ~21-nt RNA produced, we found that at the highest protein concentration used, HYL1 increased the amount of the cleavage product by ~1.5-fold (Fig. 3B and E, lanes 5–7), SE by 3.8-fold (Fig. 3B and E, lanes 8–10), and both together by almost 5-fold (Fig. 3B and E, lanes 11–13). When the miR167b present in the digest was estimated by using an antisense oligonucleotide probe on a Northern blot, the added HYL1 and SE proteins appeared to described in E. Arrowheads, pri-miR167b substrates; asterisk, ~21-nt RNA products; arrows, ~21-nt small RNAs detected by miR167 antisense probe; diamond, ~21-nt miR167b* small RNAs detected by the miR167b* antisense probe; D, DCL1; H, HYL1; S, SE.
have an even more pronounced stimulatory effect. Quantification of the results indicates that HYL stimulated DCL-mediated release of miR167b by 2-fold (Fig. 3 C and F, lanes 5–7), SE almost 6-fold (Fig. 3 C and F, lanes 8–10), and both together by >7-fold (Fig. 3 C and F, lanes 11–13). Quantification of miR167b* by image analysis of Northern blots shows an even more pronounced stimulatory effect of HYL1 and SE proteins (HYL1, 3-fold; SE, 11-fold; and HYL1+SE, 17-fold; Fig. 3 D and G). We also tested the influence of HYL1 and SE on the cleavage of a pri-miR171a substrate and observed a similar stimulatory effect of both proteins [supporting information (SI) Fig. S1]. These results suggested that both proteins improve the accuracy with which DCL1 processes pri-miRNA to mature miRNA.

To determine directly whether HYL1 and SE improve the accuracy of cleavage, we cloned and sequenced the small RNAs from reactions containing DCL1 and each protein alone as well as both proteins together. We determined the accuracy of cleavage by measuring the frequency with which the short RNAs correspond precisely or almost precisely to the miRNA sequence or its complement. The location of the sequenced small RNAs cleavage products in the sequence of the pri-miR167b substrate is shown in Fig. 4 A–D. Precise processing was defined as the percentage of the small RNA sequences identical to miR167b or its complement. HYL1 increased the accuracy from 11% to 22%, whereas SE increased the accuracy to 41%. When both HYL1 and SE were added to the reaction, 68% products corresponded to the miR167b sequence (Fig. 4E). If shorter 19-nt sequences and sequences offset from the miR167b sequence (or its complement) by <2 nt were included (termed “almost precise” in Fig. 4E), the accuracy increased from 13% for DCL1 alone to 26% for DCL1 + HYL1, 58% for DCL1 + SE, and 81% for DCL1 + HYL1 + SE (Fig. 4E). Conversely, the fraction of short RNA sequences derived from the end of the pri-miRNA substrate (termed “incorrect” in Fig. 4E) decreased when either HYL1 or SE or both were included in the reaction (Fig. 4E).

The in vivo processing accuracy is ~100%, as judged by the representation of small RNAs corresponding to the mature miRNA or its complement as a fraction of all small RNAs in the small RNA deep-sequencing databases derived from the MIR167b locus. The accuracy of miR167b* sequence is substantially underrepresented among the cloned short RNA sequences, which may be a cloning artifact, because the complementary sequence can be detected by Northern blot analysis. Correction of the error introduced by this underrepresentation would further increase the fidelity of in vitro processing of the pri-miRNA to mature miRNA by DCL1 in the presence of HYL1 and SE. Combining the increasing bias toward correct cleavage with the extent of stimulation of the cleavage rate, it appears likely that the HYL1 and SE proteins function by stimulating correct cleavage rather than by suppressing incorrect cleavage of the precursor by DCL1. We conclude that both HYL1 and SE are important in defining the cleavage sites of DCL1 on miR167b pri-miRNA, and the two proteins appear to function synergistically.

Discussion

The results of the present studies on DCL1, one of the four DCL enzymes in Arabidopsis, reveal that the plant enzyme shows significant mechanistic differences from its animal counterparts DCL and Drosha in miRNA biogenesis. It was reported that DCL1 immunoprecipitated from plant tissue extracts was able to cleave dsRNA to 21-nt small RNAs (51). We characterized the nuclease activity of purified recombinant DCL1 and found that it is able to cleave dsRNA with 2-nt 3’ overhangs to 21-nt products, requiring a divalent cation and ATP for activity.

DCL1 alone is capable of cleaving both the hairpin RNA structure of the pre-miR167b and a longer pri-miR167b containing flanking single-stranded regions to release 21-nt small RNAs. However, when we analyzed the sequence content of the small RNA digestion products by cloning and sequencing those released upon digestion of the pri-miR167b substrate with DCL1 alone, we found those corresponding to the miR167b miRNA to
be a minority (13%). Based on both genetic and cytological observations that implicate the involvement of the HYL1 and SE proteins in miRNA biogenesis (38–42, 45, 46), we studied the effects on the rate and accuracy of the DCL1-mediated processing of pri-miRNA, using pri-miR167b substrates. We found that each protein increased the rate of cleavage, as well as its accuracy, as judged by the fraction of short RNA sequences that correspond nearly or exactly to miR167b and its complement miR167b*. Addition of both proteins further enhanced the rate and accuracy of processing, suggesting that the proteins act synergistically. Judging from the increase in both the rate and the accuracy of cleavage, it appears that the HYL1 and SE proteins act by stimulating correct cleavage, rather than by suppressing incorrect cleavage of the precursor.

By contrast to Drosha, which contains a single dsRBD and cannot form a stable complex with pri-miRNA (56), DCL1 has two dsRBDs and does cleave pri-miR167b in vitro. Moreover, Arabidopsis plants homozygous for the dcl1–9 mutation that disrupts the second dsRBD are defective in miRNA biogenesis (28, 29). It therefore appears likely that the presence of two dsRBDs in DCL1 is required for its ability to carry out at least the first cleavage step from pri-miRNA to pre-miRNA. The Dicer-like activity of DCL1 in the release of mature miRNA from pre-miRNA may not require both dsRBDs because human Dicer, with only one dsRBD at the C terminus, is able to process the dsRNA substrates without other cofactors (52, 53). Furthermore, the Dicer from Giardia intestinalis, which lacks a C-terminal dsRBD, is active on dsRNA substrates (57).

The dsRNA-binding proteins TRBP and PACT are required for miRNA-mediated gene silencing and are believed to help Dicer load mature miRNA into the RISC complex. However, their precise functions remain uncertain because two reports have failed to find differences in in vitro pre-miRNA processing between Dicer and the Dicer-dsRBD complex (21, 23), whereas a third showed that depletion of TRBP reduced the Dicer activity on pre-miRNAs in vitro (22). Another study demonstrated that TRBP and PACT facilitate Dicer in siRNA production in vitro (58). The results of the present study, together with those of earlier genetic studies, establish HYL1 and SE as DCL1 cofactors in pri-miRNA processing to mature miRNA. There is both in vivo and in vitro evidence that DCL1, HYL1, and SE colocalize and interact (43–46). DCL1 has been detected in association with a very large complex (>660 kDa) that may contain DCL1, HYL1 and SE (51). As judged by gel mobility-shift assays, DCL1, HYL1, and SE are each capable of binding pri-miRNAs (data not shown). Hence, HYL1 and SE may function to present the substrate to DCL1 for correct cleavage. Among animal pri-miRNAs, the distance from the bottom of miRNA-containing stem–loop structure within the pri-miRNA to the Drosha cleavage site is consistently ~11 nt, and the dsRNA-binding protein DGCR8 is believed to measure this distance to determine the Drosha cleavage position on pri-miRNAs in generating pre-miRNA from pri-miRNA (59). The mechanism by which HYL1 and SE define the correct initial cleavage site on the pri-miRNAs for DCL1 is likely to be more complex, because the corresponding distance between the end of the stem and the initial site of cleavage is much more variable in plants than in animals.

Methods
Preparation of Baculovirus DCL1, HYL1, and SE Expression Cassettes for Insect Cells. To obtain C-terminal fusions of HA-DCL1-FLAG, oligonucleotide pairs FLAG-Asc-I, FLAG-Asci-R, and HA-Not-I, HA-Not-R were annealed and inserted into the AscI site of plasmid pENTR-D-DCL1 and the NotI site of plasmid pENTR-d-DCL1-FLAG. Although the HA tag was not used in the purification, there is an HA tag at the N terminus of the DCL1 ORF. HA-DCL1-FLAG, HYL1, and SE fragments were inserted into pDEST10 vector by LR recombination. The Invitrogen Bac-to-Bac system was used to produce the bacmid to transform the insect cells.

Purification of Recombinant DCL1, HYL1, and SE Proteins from Insect Cells. The His-HA-DCL1-FLAG recombinant protein was purified by two affinity columns by using Ni/NTA beads, followed by anti-FLAG beads. Recombinant His-HYL1 and His-SE proteins were purified by using Talon (cobalt) resin, followed by NTAs/ni resin.

Activity Assay. Briefly, the RNA cleavage assays (10 µL) contained 20 mM Tris-HCl (pH 7.0), 50 mM NaCl, 4 mM MgCl₂, 5 mM ATP, 1 mM GTP, 2 units of RNase inhibitor (RNaseOUT; Invitrogen), RNA substrate, and recombinant DCL1, HYL1, and SE proteins. After incubation at 37°C, the products were extracted with phenol/chloroform and precipitated. The processing products were fractionated by PAGE in a 12% acrylamide-8 M urea gel.

Cloning of RNA Cleavage Products. The bands corresponding to oligonucleotides in the vicinity of 21 nt were eluted from gel slices and coprecipitated with glycerol. A 5’ adapter and a 3’ adapter were ligated to the RNA cleavage products sequentially. RT-PCR was carried out by using the adapter-ligated RNA cleavage products. PCR fragments were cloned into the pGEM-T-easy vector (Promega) and sequenced with M13f primer. Sequences were analyzed with Vector NTI software (Invitrogen).

Supplemental Data. See SI Text and Table S1 for detailed methods, oligonucleotide sequences, in vitro processing of pri-miR171a, and small RNA sequences cloned from in vitro processing of pri-miR167b.

Acknowledgments. We thank M. Axtell, S. Narendra, S. Wang, and C. Vandenberg for the suggestions and discussion. This work was supported by National Science Foundation Grants IBN-0091650 and MCB-0344151 (to N.F.).