Supplemental Information

Mechanism of pri-miRNA recognition by DGCR8: an RNA-binding heme domain clamps the hairpin

Jen Quick-Cleveland[1], Jose P. Jacob[1], Sara H. Weitz[2], Grant Shoffner[1], Rachel Senturia[1] and Feng Guo[1,*]

[1] Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, CA 90095

[2] Molecular, Cell and Integrative Physiology, University of California, Los Angeles, CA 90095

* Correspondence: Email, fguo@mbi.ucla.edu; Phone, (310)206-4576; FAX, (310)206-7286.

This Supplemental Information contains the Supplemental Experimental Procedures, 6 Supplemental Figures and 1 Supplemental Table.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Plasmids

Mutagenesis was performed using the 4-primer PCR method. For N-flag-DGCR8 mammalian expression plasmids, mutant DGCR8 coding sequences were inserted between BamHI and EcoRI sites in the pCMV-Tag2A vector. For bacterial expression of NC1 R322A/R325A and R341A/K342A, DGCR8-coding sequences were inserted between NdeI and EcoRI sites in the pET-24a$^+$ vector. For expression of NC1 with an N-terminal His_6 tag (G1-G6 and corresponding wild type), DGCR8 sequences were inserted between BamHI and EcoRI sites in the pRSF-Duet1 vector. For Rhed-His_6 mutants, the coding sequences were inserted between NdeI and NotI in pET-24a$^+$. The transcription templates for pri-miR-30a, pri-miR-21 and pri-miR-380 have been reported (Faller et al., 2007; Faller et al., 2010; Barr et al., 2012). For pri-miR-9-1 and pri-miR-23a

transcription templates, pri-miRNA sequences were amplified from human genomic DNA, and were inserted between EcoRI and PstI sites in the pUC19 vector along with a T7 promoter. The coding sequences in all plasmids were confirmed using sequencing.

Expression, purification and characterization of recombinant DGCR8 proteins

Human NC1 (wild type and mutants; heme-bound and apo forms) and NC9 proteins were expressed and purified as previously described (Faller et al., 2007; Barr et al., 2011; Barr et al., 2012). Rhed-His_6 (wild type and mutants) proteins were expressed and purified using cation exchange followed by size exclusion chromatography, same as those for NC1. Immobilized metal ion affinity chromatography was not performed to avoid the use of imidazole, which might compromise the stability of heme when incubated with the proteins for an extended period of time. The purified proteins were stored in the SEC buffer containing 20 mM Tris pH 8.0, 400 mM NaCl and 1 mM DTT, except the wild-type apoNC1 which was stored in 50 mM MOPS at pH 6.0, 400 mM NaCl and 1 mM DTT. Electronic absorption spectra were recorded at room temperature on a Cary 300 spectrophotometer with bandwidth set to 1 nm.

The ability of DGCR8 mutants to bind the nitrocellulose membrane was examined by filtering them through the membrane in the presence or absence of pri-miRNAs and blotting using anti-DGCR8 antibodies.

Transcription and purification of pri-miRNAs

pri-miRNA fragments were produced using *in vitro* transcription and were purified using denaturing PAGE. Linearized plasmids were used as the transcription templates for pri-miRNAs. The transcription templates for apical junctions were amplified from pri-miRNA-coding plasmids using

PCR. The transcription templates for basal junctions were synthesized. The RNAs were purified using denaturing PAGE and their concentrations were determined using the extinction coefficients listed in Table S1.

SUPPLEMENTAL FIGURES

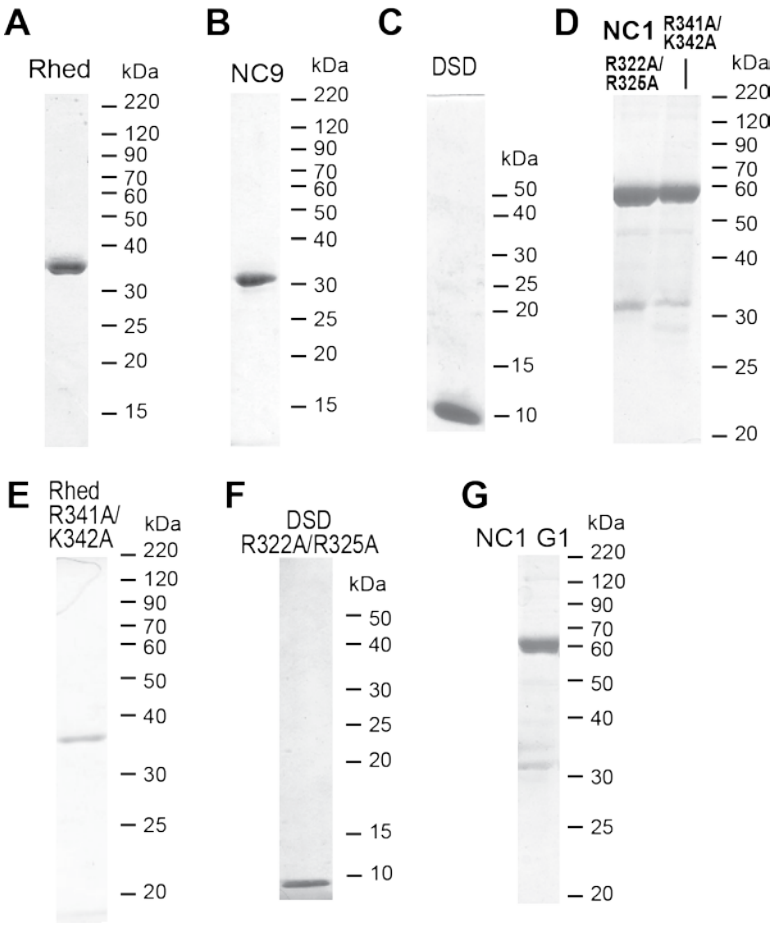

Figure S1. Coomassie-stained SDS-PAGE of purified recombinant DGCR8 proteins. Related to Figures 1-4

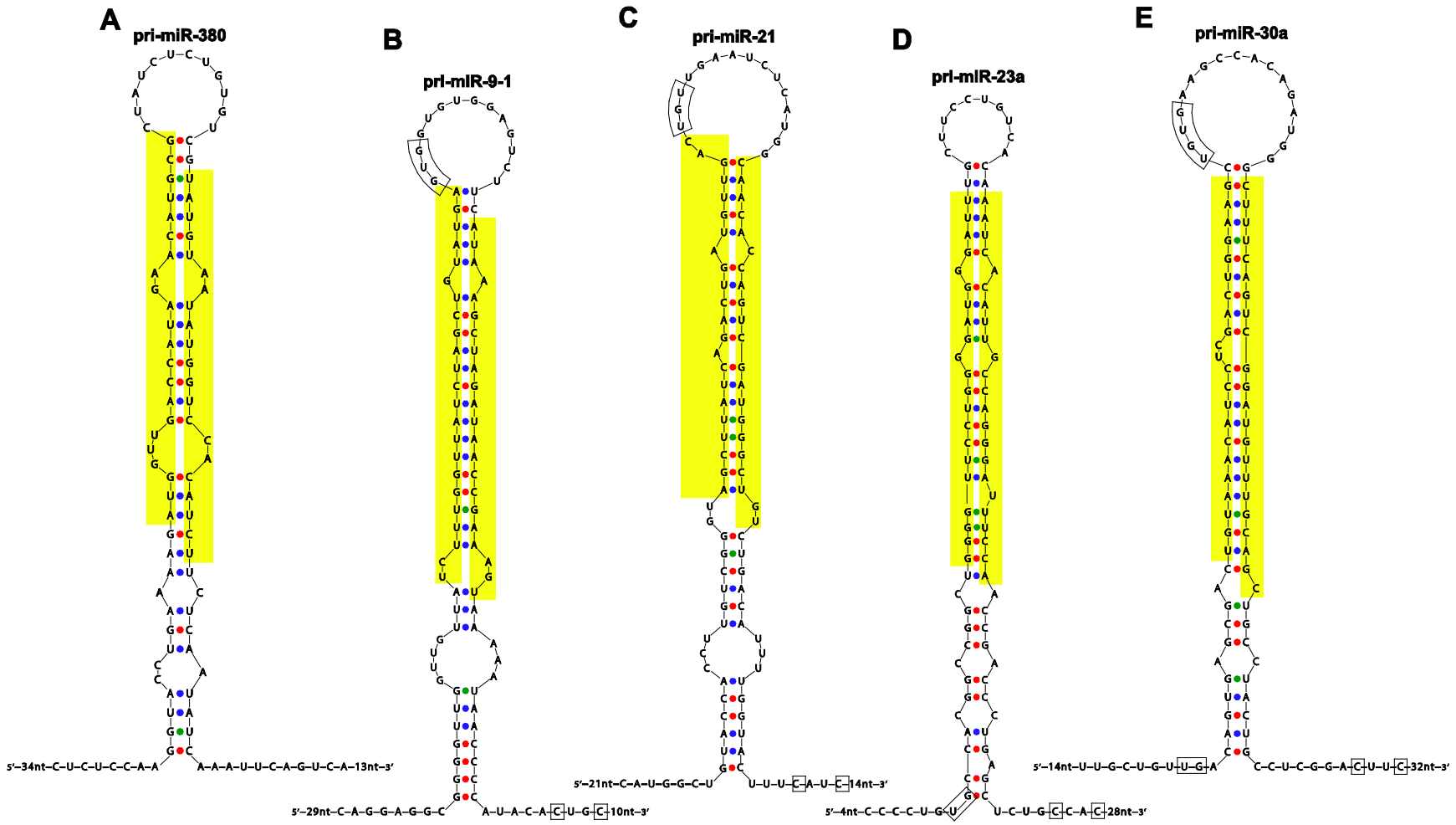

Figure S2. Secondary structures of pri-miRNA fragments. Related to Figures 1 and 2, and Table 1

(**A**) pri-miR-380. (**B**) pri-miR-9-1. (**C**) pri-miR-21. (**D**) pri-miR-23a. (**E**) pri-miR-30a. The structures were generated and graphed using MFOLD (Zuker, 2003) with the hairpin loop constrained to be single-stranded. The motifs that were recently shown to be important for human pri-miRNA processing (Auyeung *et al*, 2013) are boxed.

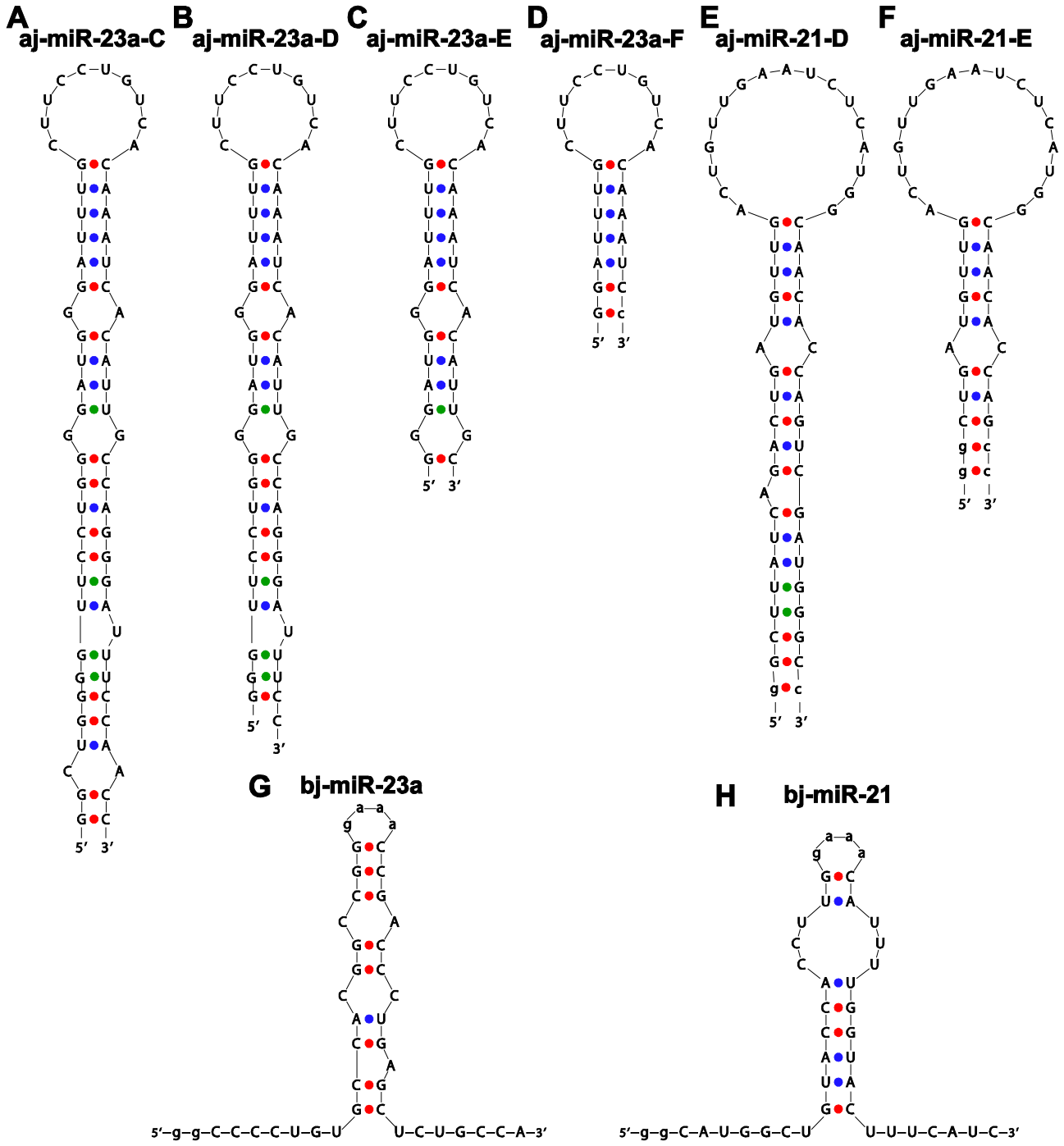

Figure S3. Secondary structures of pri-miRNA apical hairpins and basal junctions. Related to Figure 2 and Table 1

(**A**) aj-miR-23a-C. (**B**) aj-miR-23a-D. (**C**) aj-miR-23a-E. (**D**) aj-miR-23a-F. (**E**) aj-miR-21-D. (**F**) aj-miR-21-E. (**G**) bj-miR-23a. (**H**) bj-miR-21. Non-native residues that were introduced to stabilize the structures or facilitate *in vitro* transcriptions are represented by lowercase letters.

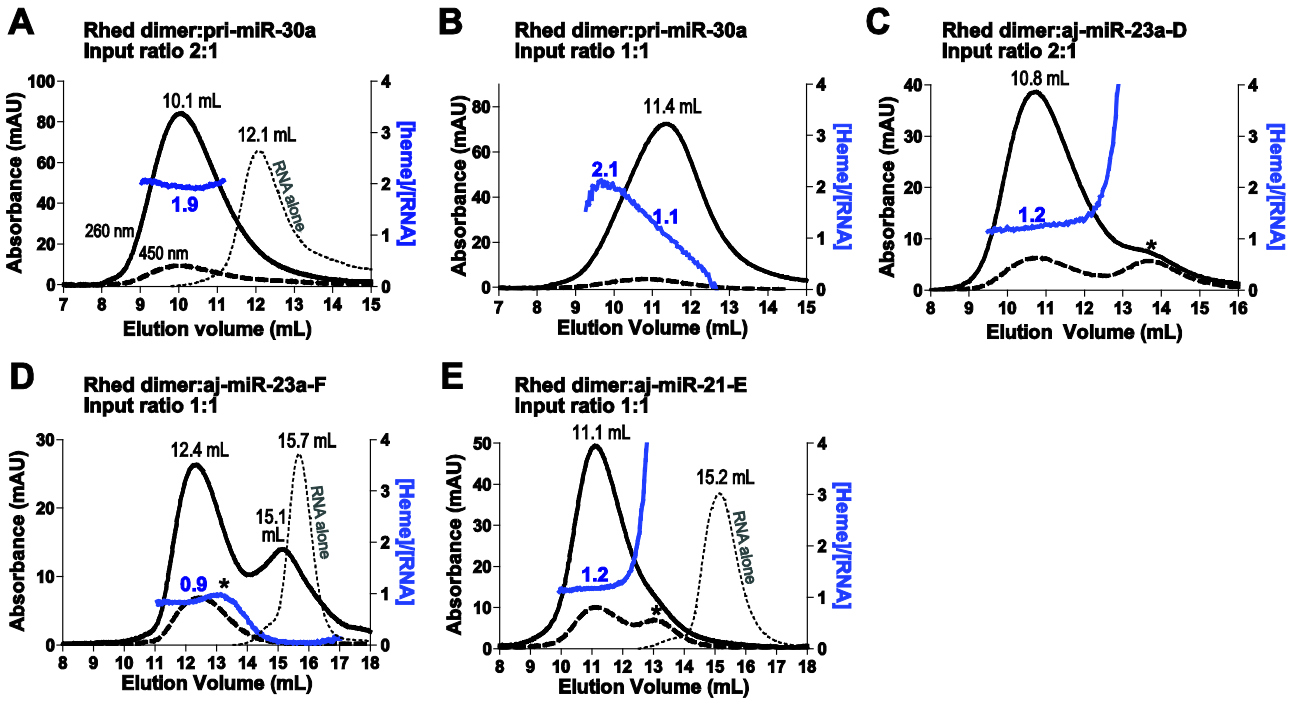

Figure S4. Size exclusion chromatograms of the Rhed in complex with pri-miRNAs. Related to Figure 2

The procedure and condition are similar to those used in Figure 2. The inputs contained 2 μM of pri-miR-30a (**A**,**B**), 4 μM of aj-miR-23a-D (**C**), 2 μM of aj-miR-23a-F (**D**), or 4 μM of aj-miR-21-E (**E**). Solid black lines indicate A_{260}, dashed lines show A_{450} and dotted lines are A_{260} of the RNA-only injections. Solid blue lines represent the heme-RNA ratios calculated from A_{450} and A_{260}, following the scale on the right *y* axis. The asterisk indicates a free-Rhed peak. The chromatogram of the Rhed and aj-miR-23a-F displayed a minor A_{260} peak at 15.1 mL (corresponding to free RNA) and an increase of the heme:RNA ratio at 13.2 mL (hinting the presence of free Rhed), indicating that the Rhed has a low affinity for aj-miR-23a-F.

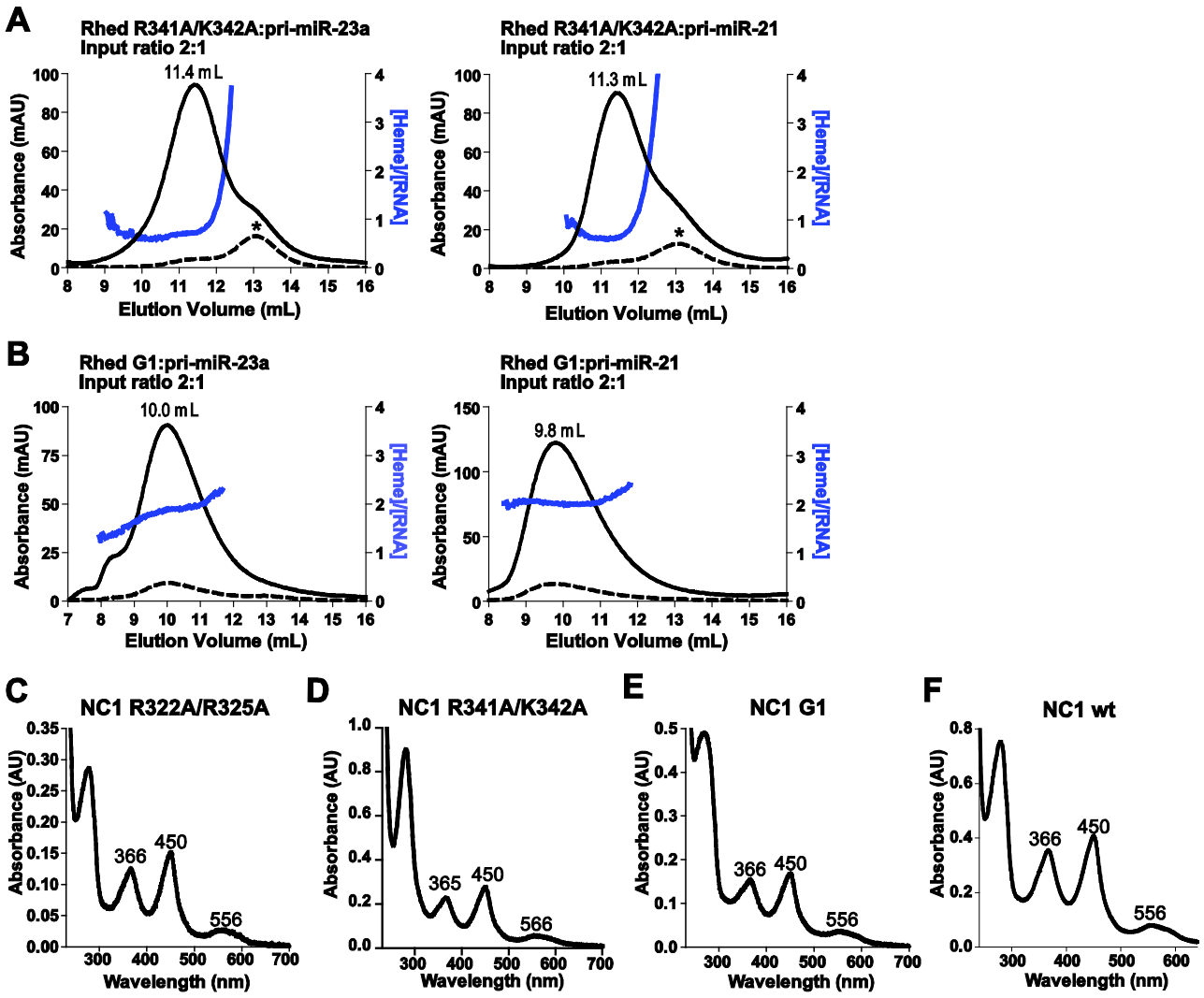

Figure S5. Characterization of RNA-binding and heme-binding properties of DGCR8 mutants. Related to Figure 4

(A, B) Size exclusion chromatograms of the Rhed R341A/K342A (**A**) and G1 (**B**) mutants in complex with pri-miRNAs. The procedure and condition are similar to those used in Figure 2. The inputs contained 4 μM Rhed dimer and 2 μM RNA. The asterisk indicates a free-Rhed peak. Solid black lines indicate A_{260}, dashed lines show A_{450} and dotted lines are A_{260} of the RNA-only injections. Solid blue lines represent the heme-RNA ratios calculated from A_{450} and A_{260}, following the scale on the right *y* axis.

(C-F) The NC1 R322A/R325A, R341/K342A and G1 mutants bind heme similarly to the wild type. Electronic absorption spectra of R322A/R325A (**C**), R341A/K342A (**D**), G1 (**E**) and wild type (**F**).

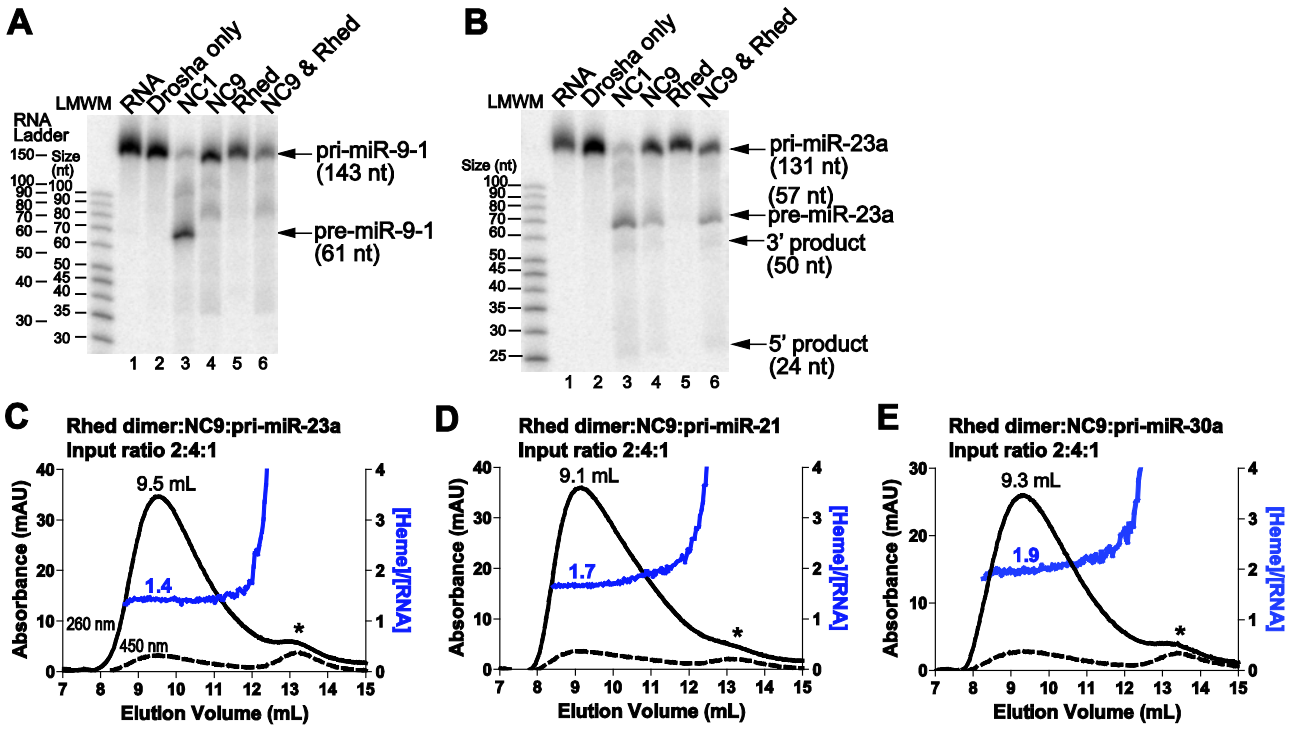

Figure S6. The covalent linkage between the Rhed and the dsRBDs is required for pri-miRNA processing activity.

(**A**, **B**) Reconstituted pri-miRNA processing assays. Uniformly ^{32}P-labeled pri-miRNAs were incubated with His$_6$-Drosha$^{390\text{-}1374}$ and indicated DGCR8 fragments. The concentrations of the DGCR8 proteins are 25 nM for NC1, 150 nM for NC9 and Rhed. The reactions were analyzed using denaturing 15% polyacrylamide gel electrophoresis (PAGE) and autoradiography. Low molecular weight marker, LMWM. Relationship between LMWM and a true RNA ladder in 15% gels is shown in panel (**A**). Purity of Rhed and NC9 is shown in Figure S1.

(C-E). Size exclusion chromatograms of the Rhed and NC9 in complex with pri-miRNAs. The procedure and condition are similar to those used in Figure 2. The inputs contained 4 μM Rhed dimer, 8 μM NC9 and 2 μM pri-miRNAs. The asterisk indicates a free-Rhed peak. Solid black lines indicate A_{260} and dashed lines show A_{450}. Solid blue lines represent the heme-RNA ratios calculated from A_{450} and A_{260}, following the scale on the right *y* axis.

Table S1. Sequences and extinction coefficients of the pri-miRNA fragments used in the study.

Related to Figures 1 and 2, and Table 1

A. pri-miRNA fragments containing all essential elements for processing

pri-miR-380 (ε_{260} = 1174 mM^{-1} cm^{-1})
GGAGAGGAAAGAGACACCGGCUCUGACCUCAGCCCUCUCCAAGGUACCUGAAAAGA
UGGUUGACCAUAGAACAUGCGCUAUCUCUGUGUCGUAUGUAAUAUGGUCCACAUCU
UCUCAAUAUCAAAUUCAGUCAUAGAGGGCUUCCC

pri-miR-9-1 (ε_{260} = 1165 mM^{-1} cm^{-1})
GGCUGCGUGGAAGAGGCGGCGACAGCAGCCAGGAGGCGGGGUUGGUUGUUAUCUUU
GGUUAUCUAGCUGUAUGAGUGGUGUGGAGUCUUCAUAAAGCUAGAUAACCGAAAGU
AAAAAUAACCCCAUACACUGCGCAGAGGGGC

pri-miR-21 (ε_{260} = 1120 mM^{-1} cm^{-1})
GGCCUACCAUCGUGACAUCUCCAUGGCUGUACCACCUUGUCGGGUAGCUUAUCAGAC
UGAUGUUGACUGUUGAAUCUCAUGGCAACACCAGUCGAUGGGCUGUCUGACAUUUU
GGUAUCUUUCAUCUGACCAUCCAUAUC

pri-miR-23a (ε_{260} = 1058 mM^{-1} cm^{-1})
GGCACCCCUGUGCCACGGCCGGCUGGGGUUCCUGGGGAUGGGAUUUGCUUCCUGUC
ACAAAUCACAUUGCCAGGGAUUUCCAACCGACCCUGAGCUCUGCCACCGAGGAUGCU
GCCCGGGGACGGGGUGGC

pri-miR-30a (ε_{260} = 1200 mM^{-1} cm^{-1})
GGAAAGAAGGUAUAUUGCUGUUGACAGUGAGCGACUGUAAACAUCCUCGACUGGAA
GCUGUGAAGCCACAGAUGGGCUUUCAGUCGGAUGUUUGCAGCUGCCUACUGCCUCG
GACUUCAAGGGGCUACUUUAGGAGCAAUUAUCUUGUUU

B. Apical junction models

aj-miR-23a-C (24 bp in the stem) (ε_{260} = 520 mM^{-1} cm^{-1})
GGCUGGGGUUCCUGGGGAUGGGAUUUGCUUCCUGUCACAAAUCACAUUGCCAGGGA
UUUCCAACC

aj-miR-23a-C-GAAA (24 bp in the stem)
GGCUGGGGUUCCUGGGGAUGGGAUUUGgaaaCAAAUCACAUUGCCAGGGAUUUCCAA
CC

aj-miR-23a-D (20 bp in the stem) (ε_{260} = 450 mM^{-1} cm^{-1})
GGGUUCCUGGGGAUGGGAUUUGCUUCCUGUCACAAAUCACAUUGCCAGGGAUUUCC

aj-miR-23a-E (11 bp in the stem) (ε_{260} = 289 mM^{-1} cm^{-1})
GGGAUGGGAUUUGCUUCCUGUCACAAAUCACAUUGC

aj-miR-23a-F (7 bp in the stem) (ε_{260} = 191 mM^{-1} cm^{-1})
GGAUUUGCUUCCUGUCACAAAUCC

aj-miR-21-D (18 bp in the stem) (ε_{260} = 451 mM^{-1} cm^{-1})
gGCUUAUCAGACUGAUGUUGACUGUUGAAUCUCAUGGCAACACCAGUCGAUGGGCc

aj-miR-21-E (10 bp in the stem) (ε_{260} = 314 mM^{-1} cm^{-1})
ggCUGAUGUUGACUGUUGAAUCUCAUGGCAACACCAGcc

C. Basal junction models

bj-miR-21 (ε_{260} = 223 mM^{-1} cm^{-1})*
ggCAUGGCUGUACCACCUUGgaaaCAUUUUGGUAUCUUUCAUC

bj-miR-23a (ε_{260} = 220 mM^{-1} cm^{-1})*
ggCCCCUGUGCCACGGCCGGgaaaCCGACCCUGAGCUCUGCCA

D. ssRNA and siRNA duplex

siRNA duplex (siDGCR8-1)

```
sense strand        5′-CAUCGGACAAGAGUGUGAUUU-3′
anti-sense strand 3′-UUGUAGCCUGUUCUCACACUA-5′
```

ssRNA same as the sense strand of siDGCR8-1

All extinction coefficients were calculated using ε_{260} = M.M./(40 μg/mL), except the ones for the short RNAs marked by "*", which were determined using alkaline hydrolysis (procedure described at http://www.scripps.edu/california/research/dna-protein-research/forms/biopolymercalc2.html)

REFERENCES

Barr, I., Smith, A.T., Chen, Y., Senturia, R., Burstyn, J.N., and Guo, F. (2012). Ferric, not ferrous, heme activates RNA-binding protein DGCR8 for primary microRNA processing. Proc Natl Acad Sci USA *109*, 1919-1924.

Barr, I., Smith, A.T., Senturia, R., Chen, Y., Scheidemantle, B.D., Burstyn, J.N., and Guo, F. (2011). DiGeorge Critical Region 8 (DGCR8) is a double-cysteine-ligated heme protein. J Biol Chem *286*, 16716-16725.

Faller, M., Matsunaga, M., Yin, S., Loo, J.A., and Guo, F. (2007). Heme is involved in microRNA processing. Nat Struct Mol Biol *14*, 23-29.

Faller, M., Toso, D., Matsunaga, M., Atanasov, I., Senturia, R., Chen, Y., Zhou, Z.H., and Guo, F. (2010). DGCR8 recognizes primary transcripts of microRNAs through highly cooperative binding and formation of higher-order structures. RNA *16*, 1570-1583.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res *31*, 3406-3415.