The Core Microprocessor Component DiGeorge Syndrome Critical Region 8 (DGCR8) Is a Nonspecific RNA-binding Protein

Braden M. Roth, Daniella Ishimaru, and Mirko Hennig

From the Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, South Carolina 29425

**Background:** Double-stranded RNA-binding domain-containing proteins play key roles in microRNA biogenesis. miRNA, microRNA; pre-miRNA, precursor miRNA; pri-miRNA, primary miRNA; PPF, pulsed field gradient; dsRBD, double-stranded RNA-binding domain; ssRNA, single-stranded RNA; nt, nucleotide; BPPLED, bipolar pulse pair longitudinal eddy current delay; SEC, size exclusion chromatography; PDB, Protein Data Bank.

**Results:** The Microprocessor protein DGCR8 binds RNA targets nonspecifically. Therefore, a sequential model of DGCR8 recognition followed by Drosha recruitment is unlikely. Instead, DGCR8 recognizes the hairpin's basal ss-dsRNA junction and recruits Drosha for cleavage.

**Conclusion:** DGCR8 alone is not responsible for specific RNA target recognition by the Microprocessor.

**Significance:** Faithful RNA processing is not causatively coupled to specific substrate binding properties of DGCR8.

MicroRNA (miRNA) biogenesis follows a conserved succession of processing steps, beginning with the recognition and liberation of an miRNA-containing precursor miRNA hairpin from a large primary miRNA transcript (pri-miRNA) by the Microprocessor, which consists of the nuclear RNase III Drosha and the double-stranded RNA-binding domain protein DGCR8 (DiGeorge syndrome critical region protein 8). Current models suggest that specific recognition is driven by DGCR8 detection of single-stranded elements of the pri-miRNA stem-loop followed by Drosha recruitment and pri-miRNA cleavage. Because countless RNA transcripts feature single-stranded-dsRNA junctions and DGCR8 can bind hundreds of mRNAs, we explored correlations between RNA binding properties of DGCR8 and specific pri-miRNA substrate processing. We found that DGCR8 bound single-stranded, double-stranded, and random hairpin transcripts with similar affinity. Further investigation of DGCR8/pri-miR-16 interactions by NMR detected intermediate exchange regimes over a wide range of stoichiometric ratios. Diffusion analysis of DGCR8/pri-miR-16 interactions by pulsed field gradient NMR lent further support to dynamic complex formation involving free components in exchange with complexes of varying stoichiometry, although in vitro processing assays showed exclusive cleavage of pri-miR-16 variants bearing single-stranded flanking regions. Our results indicate that DGCR8 binds RNA nonspecifically. Therefore, a sequential model of DGCR8 recognition followed by Drosha recruitment is unlikely. Known RNA substrate requirements are broad and include 70-nucleotide hairpins with unpaired flanking regions. Thus, specific RNA processing is likely facilitated by preformed DGCR8-Drosha heterodimers that can discriminate between authentic substrates and other hairpins.

miRNAs are a conserved class of ~22-nucleotide (nt) noncoding RNAs that regulate gene expression by base pairing with target mRNA sequences. They play key roles in modulating cellular and developmental processes (1, 2) and have been implicated in a broad range of metabolic, immunological, psychiatric, and cell cycle disorders (3–6). Mature miRNAs are generated through a conserved succession of processing steps. Typically, miRNA biogenesis begins in the nucleus with the generation of a long pri-miRNA transcript by RNA polymerase II. A hairpin harboring the pri-miRNA is selected for processing by the nuclear RNase III Drosha, which liberates the miRNA-containing pre-miRNA. Exportin-5 transports the pre-miRNA to the cytoplasm where it undergoes a second round of processing (7, 8). Another RNase III, Dicer, trims the pri-miRNA to generate a 21–24-nt RNA duplex bearing 2-nt 3’ overhangs (9).

Processing of pri- and pre-miRNAs by members of the RNase III family is facilitated by accessory proteins comprising multiple dsRBDs featuring characteristic α-β-β-α motifs. In humans, DGCR8 is required for Drosha processing and contains tandem dsRBDs (10). DGCR8 interacts with ferric heme as a cofactor to increase processing activity (11–13), and although the cooperative formation of oligomeric DGCR8/pri-miRNA assemblies has been described as another positive modulator of pri-miRNA processing (14), lysine acetylation within the tandem dsRBDs decreases the affinity of DGCR8 for pri-miRNA thereby down-regulating processing (15).

Two modes of pri-miRNA recognition have been presented. In one model, the Microprocessor binds a large (>10-nt) terminal loop to position Drosha’s catalytic center roughly two helical turns (~22-nt) from the stem-loop junction, resulting in the liberation of the ~70-nt pre-miRNA (16). A more recent model suggests that the terminal loop is dispensable for Microprocessor recognition. Instead, DGCR8 recognizes the hairpin’s basal ss-dsRNA junction and recruits Drosha for cleavage.
DGCR8 Binds RNA Nonspecifically

(17). However, the generality of this model is unclear because further experiments revealed that the length and symmetry of 5′- and 3′-nonstructured ssRNAs required for efficient processing are variable and pri-miRNA-dependent (18). Most Caenorhabditis elegans pri-miRNAs apparently lack determinants required for processing in human cells, yet one-fifth of all human pri-miRNAs lack primary sequence determinants such as downstream SRp20 binding, basal UG, and the apical GUG motifs, recently described by Bartel and co-workers (19). Collectively, this leaves 70-nt hairpins with unpaired flanking regions as the only feature and common denominator across Microprocessor substrates.

Although most dsRBBDs have been reported to bind duplex RNA as the name implies, unpaired ssRNA in loops, bulges, and mismatched pairs can also be recognized (20–23). This flexibility in substrate recognition presents a challenge in identifying both protein and RNA features required for specific protein-RNA interaction in the absence of structural data. A recent genome-wide view of DGCR8 function employing RNA cross-linking immunoprecipitation sequencing studies revealed a much broader RNA target base than anticipated (24). Specifically, these results showed that DGCR8 targets hundreds of mRNAs in addition to the anticipated interactions with pri-miRNA substrates. Implications for DGCR8 function range from roles in small nucleolar RNA and long noncoding RNA stabilization to regulation of alternatively spliced isoforms.

In this study, we investigated the RNA binding properties of the integral Microprocessor component DGCR8. Along with ssRNA and perfectly paired dsRNA substrates, we designed a comprehensive series of pri-mir-16-1 variants to elucidate the role of RNA secondary structure elements in DGCR8-RNA in general as well as the initiating event of pri- to pre-miRNA processing. These mutants were also used to investigate the link between recognition and processing of pri-miRNA substrates. Because detailed structural information on DGCR8-pri-miRNA complexes remains elusive, we utilized NMR spectroscopy to elucidate the interaction at single residue resolution and diffusion measurements to probe complex formation in solution. Although DGCR8 in isolation did not discriminate between the broad array of RNAs tested, functional processing assays confirm the specificity of Microprocessor assemblies. Thus, our results point to an intricate relationship between DGCR8 and Drosha in which both proteins are required for binding and processing, rather than a hierarchical DGCR8-mediated pri-miRNA recognition and subsequent Drosha recruitment model.

**EXPERIMENTAL PROCEDURES**

**RNA Transcription and Purification**—Pri-miRNA-containing plasmids were purified with the Omega Plasmid Giga kit, linearized with the appropriate 3′ restriction endonuclease, phenol/chloroform-extracted, and ethanol-precipitated. The minimal pri-mir-16-1 construct includes 18 nt upstream of the 5′- and 16 nt downstream of the 3′-cleavage site. Although two additional Gs were added to the 5′ overhang for efficient transcription, an additional AUUU sequence was added to the 3′ overhang after linearization employing Swal (New England Biolabs). In vitro reactions were performed in transcription buffer (40 mM Tris, pH 8.0, 1 mM spermidine, 10 mM dithiothreitol (DTT), 0.01% Triton X-100) containing 0.05 μg/ml plasmid template, 30 mM MgCl2, 1 μM T7 RNA polymerase and 6.5 mM each ATP, CTP, GTP, and UTP. Transcription reactions were incubated at 37 °C for 4 h.

Transcripts were purified as described previously (25). Briefly, the RNA was exchanged to buffer containing 10 mM NaPO4, pH 6.5, 25 mM KCl, 0.5 mM EDTA, and 25 μM Na2S. Full-length transcripts were separated from the protein, unincorporated nucleotides, and abortive transcripts by anion exchange chromatography (GE HiTrapQ HP).

**Protein Expression and Purification**—N-terminal, tobacco etch virus protease-cleavable His6-GB1/DGCR8D275, DGCR8PC, DGCR8core, and DGCR8RBBD1 (residues 509–582) sequences were transformed into *Escherichia coli* BL21-CodonPlus (DE3)-RPII competent cells, and expression was induced with 0.5 μM isopropyl β-D-thio-galactoside for 2–4 h at 37 °C (DGCR8D275 cultures were supplemented with 1 mM δ-aminolevulinic acid) (13) and purified with a HisTrapFF affinity column per the manufacturer’s instructions. Partially purified fusion proteins were cleaved with tobacco etch virus protease overnight at room temperature and reapplied to the HisTrap column. The purified protein was then exchanged into buffer appropriate for NMR analysis or biochemical assays using a PD-10 desalting column. Cultures of *E. coli* BL21-CodonPlus (DE3)-RPIII cells expressing DGCR8core and DGCR8RBBD1 (residues 509–582) were grown from a single colony in 2 ml of LB/Amp100 at 37 °C overnight. These cultures were then adapted to growth in 500 ml of 100% D2O-15N,13C-enriched M9 minimal media (26) with a succession of small volume cultures grown at 37 °C with 200 rpm agitation. Expression was induced with 1 mM isopropyl β-D-thio-galactoside for 4 h at 37 °C and purified as described previously (27). Identical batches of DGCR8 protein variants were used in successive binding and processing assays.

**Microfiltration Double-filter Binding Assays**—Filter binding was performed essentially as described in Arraiano et al. (28). Briefly, in vitro transcribed, purified RNA was end-labeled with γ-32P and then re-purified by urea-PAGE and gel extraction. The labeled RNA stocks were quantified and diluted to 30 μM. [32P]RNA was diluted to 200 pm (2× stock) with microfiltration buffer (20 mM Tris, pH 7.5, 50 mM NaCl, 1 mM DTT). The RNA was annealed by incubating at 65 °C for 5 min and then ice for 3 min. Purified protein for filter binding was prepared as 2× stocks by serial dilution in 1× microfiltration buffer. Equal volumes of 2× RNA and protein stocks were combined to yield 50-μl binding reactions. The reactions were incubated at room temperature for 1 h.

The binding reactions were filtered through nitrocellulose and nylon membranes sequentially using the BioDot microfiltration apparatus. Sample wells were washed with 2× 100 μl of microfiltration buffer before and after sample loading. Membranes were air-dried, and the intensity of the spots was quantified by a phosphorimager. The data were fit using a nonlinear least square analysis with Equation 1,

\[
FB = \frac{FB_{\text{min}} + \frac{FB_{\text{max}}[DGCR8]^{\text{P}}}{K_D + [DGCR8]^{\text{P}}}}{\text{[Equation 1]}}
\]
where FB is the fraction of RNA bound to protein; FB_{min} is the minimum fitted value of FB; FB_{max} is the maximum fitted value of FB; [DGCR8] is the molar concentration of protein, and n is the fitted value of the slope of the curve.

Electrophoretic Mobility Shift Assay (EMSA)—Purified pri-miRNA transcripts were γ-32P-end-labeled and purified as described previously (29). A 20× pri-miRNA stock was heated to 65 °C for 5 min and annealed on ice for 3 min. DGCR8 variants were diluted to the appropriate concentration as 20× stocks with EMSA reaction buffer. EMSA samples were prepared in 20 μl in reaction buffer containing 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 5% glycerol, 0.1% Triton X-100, and 0.2 mM pri-miRNA. Reactions were incubated for 1 h at 37 °C and resolved on a 6% nondenaturing polyacrylamide gel in 1× TBE.

RNA Secondary Structure Determination by NMR—All NMR spectra were recorded at 288 K or 298 K on Bruker Avance 800 or 850 MHz spectrometers equipped with a TCI cryoprobe. NMR experiments were performed on samples of 500-μl volume containing 0.1–1.6 mM of pri-mir-16 RNA variants. Data were processed using nmrPipe (30) and analyzed using Sparky (31). One-dimensional imino proton spectra were acquired using a jump-return echo sequence (32). Imino proton resonances were assigned sequence-specifically from water flip-back two-dimensional NOESY spectra (τ_{max} = 200 ms) (33). Assignments were confirmed using a jump-return (32) 15N-HMQC (34). Elucidation of base pairing and secondary structure was verified from the 2J_{NN} coupling through hydrogen bonds in the TROSY-based HNN-COSY data (35, 36).

The pri-mir-16_{lower} RNA tertiary structure was predicted using the MC-sym software (37). All information required to generate and edit the scripts are found on line on the MC-pipe-line website. Briefly, all-atom models for the 65-nucleotide pri-mir-16_{lower} satisfying the NMR-derived secondary structure were generated. 132 structures were clustered using MC-Sym, and five representative structures were selected. AMBER energy minimization was employed to refine the final five atomistic three-dimensional structures as implemented in MC-sym.

DGCR8 NMR Analysis—Backbone 1H, 15N, 13CO, 13Ca, and 13Cβ chemical shift assignments for DGCR8_{core} have been obtained using a strategy based on three-dimensional triple resonance experiments employing transverse relaxation-optimized spectroscopy (TROSY), and chemical shifts have been deposited in the BioMagResBank under accession number 177773 (27). Standard three-dimensional NMR techniques (38) were employed to obtain and confirm backbone assignments for DGCR8_{NBD1} (residues 509–582) (39). Line width and crosspeak intensities of 1H, 15N-TROSY experiments at varying protein/RNA ratios were determined using Sparky, assuming Gaussian line shapes (31).

DGCR8core/pri-mir-16_{lower} Titration—2H, 15N-DGCR8core and pri-mir-16_{lower} were buffer-exchanged to NMR buffer (20 mM NaPO4, pH 7.0, 150 mM KCl, 5 mM DTT) with the PD-10 desalting column. Titrations were performed by the addition of diluted RNA to protein in NMR buffer. Samples were concentrated with the Amicon concentrator (10K MWCO) in 500 μl of 90% H2O, 10% 2H2O. RNA was added to 250 μM DGCR8core to reach protein/RNA molar ratios of 48:1, 24:1, 12:1, 6:1, 3:1, 2:1, 1:1, and 1:2. Alternatively, diluted DGCR8core was added to 250 μM pri-mir-16_{lower} to reach RNA/protein molar ratios of 48:1, 24:1, 12:1, 6:1, 3:1, 2:1, and 1:1, respectively. Titrations were performed in NMR buffer in the absence and presence of 5 mM MgCl2.

Measurement of DGCR8-pri-miRNA Diffusion Constants by PFG NMR—All PFG-NMR experiments were recorded at 298 K on a Bruker Avance III spectrometer operating at 600 MHz and equipped with a 5-mm QXI quadruple resonance (1H,19F-13C,31P) room temperature probe with triple-axis shielded gradients. The bipolar pulse pair longitudinal eddy current delay (BPPLED) sequence (40, 41) was used, and one-dimensional spectra were recorded after two-axis gradient encoding and decoding separated by a diffusion delay of 100 ms. Dephasing of the NMR signal in the BPPLED experiment was analyzed according to Equation 2 (40, 42).

\[
A(g) = A(g_0)\exp\left(-\frac{D(g_\delta G_{\text{max}})}{2}(g^2 - g_0^2)(\Delta - \delta/3 - \tau/2)\right)
\]

where \(A(g)\) is the measured intensity of the NMR signal as a function of the fractional gradient strength \(g\); \(A(g_0)\) is the peak intensity of the reference spectrum \(g_\delta = 2\%\); \(D\) is the translational diffusion constant; \(g_\delta\) is the gyromagnetic ratio of a proton \((2.675197 \times 10^4 G^{-1} s^{-1})\); \(\delta\) is the gradient duration; \(G_{\text{max}}\) is the combined maximum power \((87.8 G \cdot cm^{-1} \cdot g^{-1})\) offered by the \(x\) and \(z\) gradients; \(\Delta\) is the time between gradients, and \(\tau\) the recovery time. By taking the natural logarithms of both sides of Equation 2, a linear equation results, and \(D\) can be calculated from least squares linear fits. To exclude complications with any exchangeable protons, the diffusion analysis of deuterated protein samples was performed by integrating over the 0.35–1.05 ppm section of the methyl region of the spectrum (primarily CH2D and CHD2 isotopomers). RNA samples were analyzed by integrating over the 7.3–8.0 ppm section.

The diffusion rate of a molecule depends on its size and shape, its concentration, the temperature, and solvent viscosity. The Stokes-Einstein equation shows that a sphere's diffusion constant \(D\) is inversely related to the hydrodynamic radius \(R_h\) and solvent viscosity \(\eta\) as shown in Equation 3,

\[
D = k_B T/(6 \pi \eta R_h)
\]

where \(k_B\) is the Boltzmann constant, and \(T\) is the temperature (298 K).

Obtaining the hydrodynamic radius \(R_h\) for a spherical particle using Equation 3 requires an accurate measure of the solvent viscosity \(\eta\). All samples were dissolved in identical aqueous (90:10 H2O:D2O) buffer solutions containing 150 mM KCl. Using viscosities of 1.097 and 0.8929 kg cm\(^{-1}\) s\(^{-1}\) for 100% D2O and a 100% H2O solution (43), the viscosity of a mixture of light and heavy water can be represented by a linear function of concentration (44) shown in Equation 4,

\[
\eta_0 = x(H_2O)\eta(H_2O) + (1-x)(D_2O)\eta(D_2O)
\]

which yields 0.91331 kg cm\(^{-1}\) s\(^{-1}\) at 298 K. A salt effect correction on viscosity was performed according to the Jones-Dole Equation 5 (45),
**RESULTS**

**pri-miRNA Recognition by DGCR8**—To investigate the secondary structure elements required for pri-miRNA recognition by DGCR8, we performed quantitative binding studies using a series of pri-mir-16-1 variants. Along with the 104-nt hairpin that includes single-stranded regions at both the base and terminal loop, we generated a terminal loop mutant (pri-mir-16Δloop) by replacing 14 nucleotides of the loop with a G:C-G C clamp and UUCG tetraloop. We also created a pri-mir-16 mutant lacking single-stranded 5’ and 3’ overhangs (pri-mir-16Δss) and a pri-mir-16ΔΔ mutant that incorporates both modifications (Fig. 1A). Each pri-miRNA secondary structure was confirmed by NMR spectroscopy monitoring imino proton resonances (Fig. 2).

To complement the pri-mir-16 constructs, we produced three recombinant DGCR8 variants to be tested in our binding studies, DGCR8Δ275, DGCR8PC, and DGCR8core (Fig. 1B). DGCR8Δ275 lacks the first 275 amino acids but retains WW and heme-binding domains, two dsRBDs and the Drosha interaction/trimerization domain and is competent for both dimerization and pri-miRNA processing (13, 14, 48). DGCR8PC includes residues 484–750 and represents the processing-competent (PC) core (48). DGCR8core (residues 493–706) lacks the Drosha interaction domain but retains the tandem dsRBDs and is therefore unable to participate in pri-miRNA processing but is RNA binding-competent (14, 49).

DGCR8Δ275/pri-miRNA interactions were monitored by filter binding assays. As shown in Fig. 3B, DGCR8Δ275 bound pri-mir-16-1 with low nanomolar affinity (3.7 ± 0.9 nM), in qualitative agreement with previously published filter binding experiments (14). Surprisingly, DGCR8Δ275 bound pri-mir-16Δloop, Δss, and ΔΔ constructs with similar affinities (Fig. 3, C–E, summarized in Table 1) and Hill coefficients, suggesting that neither the ss-flanking regions nor the terminal loop are determinants for specific recognition by DGCR8.

To test whether pri-mir-16 mutants contain enough ssRNA in internal loops or bulges to attract DGCR8, we have modified the pri-mir-16ΔΔ construct to produce a fully paired hairpin (pri-mir-16Δfp) by altering unpaired and bulged nucleotides to effectively eliminate ssRNA and maximize duplex stability (Fig. 1A). DGCR8Δ275 was still able to bind pri-mir-16Δfp with affinity and cooperativity equivalent to that of pri-mir-16-1, Δloop, Δss, and ΔΔ constructs with similar affinities (Fig. 3F). To rule out the possibility that DGCR8Δ275 recognition of pri-mir-16-1 is driven by a sequence determinant common to all of our RNA substrates, we tested pri-mir-16x (Fig. 1A), a Quikfold (50)-predicted 82-nt hairpin co-transcribed with mir-16-1 and mir-15a on the DLEU2 lincRNA (long intergenic noncoding RNA) (51, 52). Located 397 nt upstream of mir-16-1, pri-mir-16x is overlooked by the Microprocessor. As shown in Fig. 3G, filter binding analysis revealed that DGCR8Δ275 binds pri-mir-16x with no discernable loss of affinity or cooperativity relative to the other RNAs tested. To test whether the lack of significant variation in our results could be explained by DGCR8Δ275 recognition of predominantly dsRNA elements, we produced an 87-nt ssRNA principally composed of nine tandem CCCCUAAA repeats (Fig. 1A). NMR analysis of exchangeable imino protons confirmed that no sta-
ble secondary structure was formed in this transcript (data not shown). Finally, we also generated a pri-let-7b transcript to serve as a second bona fide pri-miRNA substrate. Again, DGCR8 bound both constructs with \( K_d \) and Hill values consistent with other RNA substrates (Fig. 3, H and I).

Similarly, binding affinities and cooperativity measurements remained consistent when we analyzed DGCR8\(^{\text{core}}\) and a subset of the pri-miRNA variants, even though the affinities were reduced \( \sim 10 \)-fold for DGCR8\(^{\text{core}}\), compared with DGCR8\(^{\Delta 275}\) (Table 1). Collectively, our filter binding experiments demonstrate that DGCR8 recognizes the tested RNA variants with high affinity and Hill coefficients ranging from 0.9 to 1.8. None of the DGCR8 constructs tested in filter binding demonstrated a pronounced tendency to bind cooperatively beyond a dimer model.

To directly detect oligomeric DGCR8-RNA complex formation, we performed electrophoretic mobility shift assays (EMSA) with DGCR8\(^{\Delta 275}\), DGCR8\(^{\text{PC}}\), or DGCR8\(^{\text{core}}\) and the
pri-mir-16 transcript. As seen in Fig. 4A, progressively more truncated variants of DGCR8 all populate similar heterogeneous complexes with pri-mir-16-1. Additionally, DGCR8/pri-miRNA interactions reveal a “smearing” effect rather than discrete band formation, which could indicate that DGCR8 binds pri-miRNA transiently and with variable stoichiometry, sampling different regions of the substrate until excess protein effectively saturates all potential binding sites on the RNA at any given time. EMSA titrations of DGCR8core into radiolabeled pri-mir-16-1 in the presence of varying amounts of non-specific competitor yeast tRNA\(^{\text{Phe}}\) revealed that increasing amounts of tRNA increase the apparent \(K_d\) values for the DGCR8core/pri-mir-16-1 complex (Fig. 4B). Thus, on the basis of our EMSA results, DGCR8’s mode of binding resembles other dsRBDs (such as protein kinase R (PKR) (53)) that interact with RNA through transient and/or noncooperative interactions featuring multiple, potentially overlapping RNA-binding sites with comparable affinities. Taken together, our binding data imply that DGCR8 alone is not able to distinguish miRNA-containing hairpins from other RNA species but rather retains high, indiscriminate affinity for RNA within its dsRBDs.

**DGCR8/pri-miRNA Binding Monitored by NMR**—In the interest of reducing overall molecular weight for NMR studies, we used a 65-nt version of pri-mir-16-1 (pri-mir-16\(^{\text{lower}}\), Fig. 1A) that is correctly processed \textit{in vitro} (data not shown) (17). Using NMR spectroscopy, we have determined that the secondary structure of pri-mir-16\(^{\text{lower}}\) includes 5\(^{-}\)/H9004- and 3\(^{-}\)/H9004-ssRNA overhangs, a 21-nt stem featuring tandem G-A mismatches and a C-A mismatch, and a stable 8-nt loop (Fig. 5).

Our recent NMR analysis confirmed that our purified DGCR8\(^{\text{core}}\) contains all of the secondary structure features of the reported crystal structure (Fig. 7A) (27, 49). Furthermore, we observed pronounced chemical shift differences between the backbone amide \(^{1}\text{H},^{15}\text{N}\)-correlations of an isolated RBD1 domain and those of the RBD1-RBD2-containing DGCR8\(^{\text{core}}\) construct (Fig. 6). This demonstrates that the two tandem RBD domains interact with each other in solution confirming earlier reports based on the crystal structure (49) and MD simulations (54).

To characterize the interactions between DGCR8 and pri-miRNA at single residue resolution, an NMR titration study was performed in which \(^{1}\text{H},^{15}\text{N}\)-TROSY spectra of \(^{2}\text{H},^{15}\text{N}\)-laabeled DGCR8\(^{\text{core}}\) were monitored upon addition of increasing...
amounts of RNA. pri-mir-16-1lower was titrated into 250 μM
$^2$H,15N-DGCR8core to produce protein/RNA solutions ranging
from 48:1 to 1:2 molar ratios. We observed differential broadening
of backbone amide $^1$H,15N correlations evident even at low concentrations of pri-mir-16-1lower (Fig. 7). Because binding and
dissociation between DGCR8core and pri-mir-16-1lower occurred on the intermediate time scale, $^1$H,15N correlation experiments do not reveal discrete resonances for RNA-bound and free states commonly observed for systems in fast or slow exchange. Thus, our protein-RNA interface analysis focused on differential broadening at a molar ratio where most protein resonances were still observed (24:1 ratio of DGCR8core/pri-mir-16-1lower).

Differences in amide proton LW (∆LW) upon RNA binding were calculated by subtracting the LWbound of a given residue in the presence of sub-stoichiometric RNA from the LWfree observed in the absence of RNA (Fig. 8A). A value of zero indicates no change in LW. The ∆LW values for infinitely broadened resonances (Fig. 8A, gray bars) were calculated assuming a conservative LWbound of 50 Hz, which corresponds to the largest $^1$H-TROSY LWbound observable at 800 MHz. The residues with the largest change in line width (most negative ∆LW) values map to the first and second helices of the each dsRBD motif (RBD1:α1 and α2, RBD2:α5 and α6) (Figs. 7 and 8B). This pattern of RNA-dependent line broadening correlates with previously characterized canonical RBD interactions in which approximately one helical turn of a successive minor and major groove dsRNA is occupied by the α-helical regions of the α-β-β-β-α dsRNA-binding domain (20, 55). However, no significant broadening was detected for either the β2-β3 loop located in RBD1 or the β5-β6 loop in RBD2, which one would predict to contact an adjacent minor groove in a canonical dsRBD-dsRNA interaction (Figs. 7 and 8B). Instead, noticeable broadening occurs within the β4, β7, and α7 regions. For its part, α7 has been shown to stabilize the DGCR8core through extensive intramolecular contacts with both dsRNA (49, 54). Therefore, line broadening in this region is not likely attributed to direct association with pri-mir-16-1lower but is probably reporting on conformational exchange on an intermediate time scale between bound and free DGCR8core forms. Ultimately, line broadening analysis of our DGCR8core/pri-mir-16-1lower titration suggests some variation in archetypal dsRNA binding for both RBD subunits of the DGCR8core where three structural elements of each dsRBD, the N-terminal α-helix, the loop between β1 and β2, and the C-terminal second helix, engage adjacent minor grooves and the intervening groove of dsRNA (56).

Finally, we tested whether protein resonances, which are broadened beyond detection at substoichiometric RNA ratios,
would sharpen as a result of changes in exchange regime brought about by the formation of specific, stable interactions in the presence of excess pri-mirR16lower. Over the course of the titration and at the end point ratio 1:2 DGCR8core/pri-mir-16lower, the vast majority of assigned backbone amide resonances was broadened beyond detection. The remaining 30 residues experienced negligible chemical shift perturbations and map primarily to the flexible linker between RBD1 and RBD2 (residues Asp-579 to Glu-594), the N terminus, including strand β1 (residues Gln-495 to Ile-505), and the C terminus of the protein (residues Ser-702 to Val-706) (Figs. 7 and 8C). Overall, our findings indicate that DGCR8core exchanges between and among pri-mir-16lower-binding sites on the intermediate chemical shift time scale leading to differential line broadening. Such a model was further corroborated by a titration revealing global imino proton line broadening of pri-mir-16lower in the presence of increasing amounts of DGCR8core (Fig. 9A).

Divalent cations often stabilize important RNA tertiary structures. Therefore, titrations were repeated in the presence of 5 mM MgCl₂. Imino proton chemical shift changes observed for pri-mir-16lower and larger RNA variants (Fig. 1A) in the presence of Mg²⁺-cations were invariably small and did not exceed 0.04 ppm (data not shown). Moreover, intermediate exchange kinetics prevailed in the presence of Mg²⁺-cations (Fig. 9B). The NMR titration data also suggest that the structure of DGCR8core may not be significantly perturbed by the pri-mir-16lower interaction because chemical shifts of resonances observable during the titration are mostly unchanged.

DGCR8/pri-miRNA Binding Monitored by PFG-NMR—A distinct advantage of PFG-NMR methods is that translational self-diffusion coefficients can be determined under experimental conditions (protein and RNA concentration, temperature, and viscosity) identical to those used for collecting the NMR titration data described above. Because diffusion measurements are sensitive to changes in molecular size and shape, these experiments provide a supplementary characterization of DGCR8/pri-mir-16-1 association to complement our line broadening analysis. The results of PFG-NMR experiments are reported in Table 2. A single (average) diffusion coefficient was observed in all experimental measurements of free DGCR8 protein variants or pri-mir-16lower, as evident from linear correlations with R values ≥0.996 (Fig. 10A).

Possible explanations for a slower diffusion rate (8.7 ± 0.01 × 10⁻¹¹ m²·s⁻¹) of the 214-residue DGCR8core (larger apparent R₅₀) in comparison with calculated values (9.8 ± 10⁻¹¹ m²·s⁻¹) include differences between the solution structure and the crystal structure used for the prediction, intermolecular oligomerization, and/or the presence of unfolded protein regions. In the case of intermolecular oligomerization, one would expect the slower diffusion to be caused by partial dimerization. Assuming exchange averaging between a monomer and a dimer on the diffusion time scale, we calculated a 47% dimer population of the DGCR8core. In good agreement with results previously published (49), our SEC analysis confirmed an estimated molecular mass of 31 kDa for the DGCR8core, significantly larger than the theoretical molecular mass of 24.2 kDa (data not shown). The overall agreement of experimental and calculated translational diffusion coefficients and hydrodynamic radii for the DGCR8 variants and pri-mir-16lower in isolation confirms that physically meaningful data could be obtained for macromolecular DGCR8/pri-mir-16 complexes.
For the DGCR8core-pri-mir-16lower complex, diffusion rates of the protein in the presence of excess RNA (1:2 molar ratio) revealed a deviation from linearity ($R^2 = 0.990$) (Fig. 10B), suggesting a polydisperse system described by Equation 2. Linear correlations obtained after subtracting the contributions from the slowly diffusing complex (i.e. elimination of last five points) or the faster diffusing component (i.e. elimination of first five points) yielded upper and lower boundaries of $D = 10.1 \pm 1.02 \times 10^{-11}$ m$^2$s$^{-1}$ with ($R = 0.999$) and $D = 6.4 \pm 0.15 \times 10^{-11}$ m$^2$s$^{-1}$ ($R = 0.999$). Extrapolated to infinite dilution, Stokes radii of $R_H = 24.1 \pm 2.6$ Å and $38.1 \pm 0.9$ Å can thus be estimated for the smallest and largest diffusing spherical particles for the DGCR8core-pri-mir-16lower complex exchanging on the diffusion time scale. The smallest diffusing molecule is likely unbound DGCR8core given the close agreement with the predicted $R_H$ calc value of 24.7 Å.

Recombinant DGCR8$^{\Delta 275}$ Is an Active Microprocessor Component—We performed in vitro Microprocessor assays that mimic the first step in miRNA biogenesis to examine the relationship between observed DGCR8/pri-miRNA binding patterns and processing. When incubated with immunoprecipitated Microprocessor, the pri-mir-16-1 and 16loops were efficiently processed to release respective 65- and 59-nt pre-miRNAs as well as the 20-nt 5' and 3' reaction by-products (Fig. 11A). However, pri-mir-16 variants lacking ss-flanking RNA (16ss and 16Δss) were not cleaved by the Microprocessor, in agreement with previously published work (14, 17, 18). In addition, no cleavage products were observed when pri-mir-16x was used in the processing assay. These data demonstrate that only a distinct subgroup, including pri-mir-16-1 and 16loop, is able to productively engage the Microprocessor, even though all RNA constructs display similar binding by DGCR8.

To address whether DGCR8's RNA specificity determinant could reside in the truncated portion of the DGCR8$^{\Delta 275}$ constructs, we have performed in vitro processing with a DGCR8$^{\Delta 275}$-reconstituted Microprocessor. c-Myc-Drosha alone did not efficiently process pri-mir-16-1, 16Δss, or 16Δloop (Fig. 11B). However, when incubated with purified recombinant DGCR8$^{\Delta 275}$, 5' cleavage products were observed for pri-mir-16-1 and 16Δloop but not 16Δss, confirming that recombinant DGCR8$^{\Delta 275}$ is a properly folded, active component of a highly specific in vitro pri-miRNA processing complex.
Taken together, our binding and processing data strongly suggest that DGCR8 alone is not the specificity determinant for recognition of bona fide miRNA-containing hairpins. Instead, DGCR8 works in concert with Drosha to promote specific recognition and processing of pri-miRNA transcripts.

DGCR8-Drosha Co-immunoprecipitation—In the assumption that the DGCR8-Drosha complex provides the specificity for pri-miRNA recognition, it is necessary that DGCR8 and Drosha can form heterodimers prior to pri-miRNA binding. To probe for the prerequisite heterodimer, a co-immunoprecipitation assay was performed in the absence or presence of RNases. Western blot analysis of Drosha-immunoprecipitated complexes revealed the presence of DGCR8 in extracts treated with either RNase inhibitor or RNase A/T1 mix (data not shown). This observation confirms earlier reports by Han et al. (10) and suggests that association of Drosha and DGCR8 does not depend on the availability of RNA substrate molecules.

DISCUSSION

Accessory dsRNA-binding proteins involved in small RNA pathways have been shown to be required for efficient and accurate modulation of specific RNA species, when part of multiprotein complexes (10, 57, 58). Despite the importance and characterized functions of endonuclease-associated dsRBD-containing proteins in miRNA biogenesis, few reports have described the detailed mechanistic and structural nature of their interactions with RNase IIIIs or specific RNA substrates (59).

Previous results from filter binding experiments form the basis for a recognition model in which highly cooperative binding by DGCR8 is the distinguishing characteristic of specific pri-miRNA recognition (14). Although our $K_d$ values are in agreement with those previously reported for filter binding (13, 49), we have been unable to demonstrate conclusive deviations in binding cooperativity among various RNA substrates. Our direct, EMSA-based observation of higher order complex formation involving DGCR8core, which lacks a previously identified trimerization domain, also questions a generalized model of pri-miRNA recognition involving stepwise, cooperative binding mediated by the DGCR8 trimerization domain (14).

We employed concentrations of 100 pm of end-labeled RNA and a low salt buffer (50 mM NaCl) in filter binding assays and observed tighter affinity and lower Hill coefficients for the heme-bound DGCR8$^{32775}$/pri-mir-16-1 interaction when com-
pared with previously published studies by Faller et al. (13, 14) using unspecified concentrations of body-labeled pri-mir21 and pri-mir30a in complex with DGCR8core/275 in a reaction buffer containing 85 mM NaCl. We observed reduced affinity for the DGCR8core/pri-mir-16-1 interaction using EMSA in a high salt buffer (150 mM KCl, data not shown). Thus, the apparent $K_d$ values for DGCR8 binding to pri-mir-16-1 generally decrease with increasing salt concentration.

Hill coefficients often do not provide accurate estimates of the number of binding sites (60), particularly when multisite complex formations are involved. In general, nitrocellulose filter binding describes only macroscopic association parameters that are composite averages of individual binding sites and cooperativity constants. Here, one must assume a relationship between the number of ligands bound and complex retention by the nitrocellulose membrane to accurately describe the interaction. Senear et al. (61) have concluded that it is impossible to correctly resolve microscopic binding and cooperativity constants of a multisite complex by filter binding. The apparent steepness of dose-response curves and the corresponding Hill coefficients ($n$) depend on invariant $^32P$-labeled RNA concentration below the $K_d$. When RNA concentrations approach the $K_d$ value, the resulting best fits to the Hill equation no longer provide accurate information on either $K_d$ or $n$ (60, 62). Dimerization at higher DGCR8 concentration can potentially alter binding characteristics that would further complicate reliable curve fitting. Using SEC and NMR diffusion measurements, we and others (49) have detected a dimerization tendency for DGCR8 variants. Finally, because we investigated DGCR8-pri-mir-16-1 and pri-let-7b complex formation, we cannot rule out the possibility that the highly cooperative association previously observed could be a specific feature of DGCR8 binding to pri-mir30a and pri-mir21 studied by Faller et al. (13, 14). Although processed by the Microprocessor, alter-

**FIGURE 8.** $^1H,^15N$-TROSY resonance broadening map for binding of DGCR8core to pri-mir-16lower.

**A**, changes in $^1H$ LW ($\Delta LW$) of 250 $\mu$M $^1H,^15N$-labeled DGCR8core in the presence of 10.4 $\mu$M pri-mir-16lower versus the amino acid sequence of DGCR8core. $\Delta LW$ values were calculated by subtracting the LW bound of a given residue from the LW free. $\Delta LW$ for infinitely broadened resonances (gray bars) were calculated assuming a conservative LW bound of 50 Hz corresponding to the largest $^1H$-TROSY LW bound observable at 800 MHz. Secondary structure elements ($\alpha$1–7 and $\beta$1–7) as determined by chemical shift index using TALOS+ (75) are identified below the residue number. $B$, color-coded worm representation of DGCR8core/pri-mir-16lower titration at a 24:1 protein/RNA ratio. Secondary structure elements are numbered and indicated. Worm thickness of the DGCR8core crystal structure (PDB code 2YT4) varies between 2.00 ($\Delta LW = -38.3$ Hz, red) and 0.25 ($\Delta LW = 9.2$ Hz, blue). Missing crystal residues were modeled and energy-minimized in Chimera (76). $C$, DGCR8core/pri-mir-16lower titration at a 1:2 protein/RNA ratio. Individual RBDs are numbered; N and C termini are shown and indicated. The 39 visible TROSY peaks remaining in the presence of excess RNA predominantly map to the N and C termini, and flexible linker regions of the DGCR8core crystal structure (PDB code 2YT4) and are mapped and highlighted in blue on the protein sequence ($A$).
native biogenesis pathways for pri-mir21 independent of DGCR8 have been suggested (63).

UV cross-linking experiments in the presence of varying amounts of competitors have previously been employed to establish relative affinities of DGCR8 to pri-miRNA (17). Although DGCR8 could be cross-linked with decreasing efficiency to diverse targets such as pri-mir-16-1, a small interfering (si)RNA duplex and ssRNA, the ability to compete for DGCR8 binding revealed some advantage for ss-dsRNA junctions. Furthermore, a length dependence where shorter RNA constructs competed gradually less efficiently with the longer pri-miRNA was observed. Our DGCR8core/pri-mir-16-1 EMSA and increasing apparent Kd values in the presence of varying amounts of nonspecific competitor yeast tRNA suggest that DGCR8core/RNA interactions are nonspecific. Consistent with a nonspecific binding mode, the results obtained from competition experiments may have uncovered a correlation between efficient competition and the number of nonspecific binding sites provided by the various competitors.

To overcome the binding assay limitations described above and to obtain independent information about the DGCR8core/pri-mir-16 complex equilibrium in solution, we conducted NMR titration and diffusion measurements. Results from NMR titrations strongly suggest that the DGCR8core/pri-mir-16lower complex represents a nonspecific interaction. The observed line broadening in the absence of chemical shift perturbations indicates the formation of transient, nonspecific complexes (64). Results obtained from diffusion measurements further corroborated nonspecific DGCR8core/pri-mir-16lower binding models as we observed interconverting free DGCR8core and complexes with hydrodynamic radii of ~38.1 Å.

These findings are consistent with reports of other dsRBD-containing proteins involved in small RNA pathways. For example, dsRBDs of both HYL1 (HYTONASTIC LEAVES 1) and PKR that exhibit noncanonical binding and extensive line broadening have been shown to bind RNA without apparent specificity (55, 65, 66). Similarly, we have been unable to detect by NMR nucleotide-specific changes to RNA targets complexed with DGCR8core. Analogous to the protein spectra, pri-mir-16 imino-proton peaks were broadened globally and severely. This result suggests that DGCR8 dsRBDs simultaneouly bind overlapping interfaces of the RNA in a manner resembling nonspecific PKR-dsRNA complex formation (66). Finally, the Dicer accessory dsRBD-containing protein-trans-activating response RNA-binding protein and two homologs, including PKR, were most recently shown to diffuse along dsRNA substrates, which is clearly an activity conceivably adopted by DGCR8 to anchor Drosha-DGCR8 heterodimers (67).

Although our nonspecific DGCR8/pri-miRNA interaction model differs from the proposed cooperative recognition model (14), several common observations have been described. Two studies have identified that helix 2 in each α-β-β-β-α dsRNA-binding domain of DGCR8 is important for pri-miRNA

![Summary of least squares regressions and experimental and theoretical translational self-diffusion constants and hydrodynamic radii](image)

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary of least squares regressions and experimental and theoretical translational self-diffusion constants and hydrodynamic radii</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>DGCR8RRD1</th>
<th>DGCR8core</th>
<th>DGCR8PC</th>
<th>pri-mir-16lower</th>
<th>DGCR8core/pri-mir-16lower complex</th>
<th>Complex lower limit</th>
<th>Complex upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(Kd) versus (g^2·a0)^2^a</td>
<td>R ≤ 0.997</td>
<td>R ≤ 0.999</td>
<td>R ≤ 0.997</td>
<td>R ≤ 0.996</td>
<td>R ≤ 0.999</td>
<td>R ≤ 0.999</td>
<td></td>
</tr>
<tr>
<td>D^calc</td>
<td>13.83</td>
<td>9.815</td>
<td>8.716</td>
<td>0.998</td>
<td>± 0.223</td>
<td>± 1.024</td>
<td>± 0.153</td>
</tr>
<tr>
<td>10^{-11} μm^2·s^b</td>
<td>± 0.044</td>
<td>± 0.100</td>
<td>± 0.04</td>
<td>± 0.121</td>
<td>± 0.098</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10^{-11} μm^2·s^c</td>
<td>± 0.153</td>
<td>± 0.153</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>± 1.0</td>
<td>± 0.3</td>
<td>± 0.1</td>
<td>34.5</td>
<td>± 1.1</td>
<td>24.1</td>
<td>38.1</td>
</tr>
<tr>
<td>A</td>
<td>17.5</td>
<td>24.7</td>
<td>29.4</td>
<td>23.0</td>
<td>± 0.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a Slopes were derived from linear least squares fits using three replicate measurements (30 data points total for the 1st 5 columns; 15 data points were used to determine upper and lower boundaries, 6th and 7th columns). Data analysis was performed using Prism 4.0 software.

^b Translational diffusion constants derived from Equation 2 at 298 K in aqueous (90:10% H2O/D2O) buffer solutions containing 135 mM KCl. Three replicate experiments were performed, and the resulting average values for D with corresponding standard deviations are reported.

^c Theoretical translational diffusion constants were calculated from atomic level structures available for DGCR8RRD1 and DGCR8core using HYDROPRO (PDB accession codes 1X47 and 2YT4, respectively).

^d Theoretical translational diffusion constants were calculated from atomic level structures available for DGCR8RRD1 and DGCR8core using HYDROPRO (PDB accession codes 1X47 and 2YT4, respectively).
recognition (15, 49). In good agreement, our NMR titration analysis recognizes the first and second helices of each dsRBD motif (RBD1:a1 and a2, RBD2:a5 and a6) as primary pri-miR16-1 interaction interfaces. The binding affinity of DGCR8 for the 160-nt P4–P6 domain of the Tetrahymena ribozyme is similar to the ones observed for bona fide pri-miRNA substrates (14). Likewise, we cannot distinguish ssRNA, fully paired (pri-mir-16fp) or random hairpin (pri-mir-16x) interactions involving DGCR8 from pri-miRNA-derived substrates based on $K_d$ values. Previous studies have highlighted the importance of oligomeric (DGCR8)-pri-miRNA complexes using filter binding assays, SEC, and low resolution electron tomography (14). Our NMR titration and diffusion analysis directly probe solution equilibria and offer further support for oligomeric, albeit nonspecific, binding of DGCR8 to RNA substrates. Furthermore, we were able to visualize DGCR8-pri-miRNA complexes with variable stoichiometry using EMSA.

DGCR8 alone shows no preference for a particular RNA substrate, yet we and others have demonstrated through in vitro processing assays that the Microprocessor is highly selective when cleaving its RNA targets (17, 18). Thus, we conclude that DGCR8-RNA substrate binding is not the specificity determinant of pri- to pre-miRNA processing, and a model of stepwise pri-miRNA recognition and Drosha recruitment by DGCR8 seems unlikely. Because known requirements common to all Microprocessor substrates remain broad and include 70-nt hairpins with unpaired flanking regions (19), the functional implication of our finding is that the concerted action of a DGCR8-Drosha complex is required for effective recognition of miRNA-containing transcripts. Consistently, recent studies...
revealed that association with the HIV-1 LTR requires the presence of both Microprocessor components, Drosha and DGCR8 (68). An alternative scenario, where Drosha can distinguish its target among nonspecific DGCR8-pri-miRNA complexes cannot be strictly ruled out based on our results.

Importantly, stable association of DGCR8 and Drosha occurs independently of RNA (10). Presumably, this protein/protein interaction leads to enhanced RNA binding specificity by way of conformational modulation of DGCR8’s dsRBDs, by steric limitation of binding opportunities to constrained regions of the RNA hairpin, or a combination of both scenarios. The DGCR8/RNA interaction and other examples (55, 67, 69, 70) suggest that dsRBDs associated with miRNA biogenesis are generally nonspecific binders. Although this form of association might seem less than satisfying, nonspecific and/or loose binding can be an important characteristic for an accessory protein. Thus, the biological implication is that nonspecific associations can greatly enhance the RNA binding flexibility of a protein complex, which is in good agreement with reports demonstrating DGCR8’s involvement in controlling the fate of diverse classes of RNA such as mRNAs, small nucleolar RNAs, long noncoding RNAs, and pri-miRNAs (24). For the Microprocessor in particular, the ability to bind a broad range of loosely related hairpin structures could be critical to carry out its universal function.

Acknowledgments—We thank members of the Hennig laboratory for stimulating discussions and comments on the manuscript. We acknowledge the support of the Hollings Marine Laboratory NMR Facility for this work.

REFERENCES