9694–9705 *Nucleic Acids Research, 2017, Vol. 45, No. 16* *Published online 17 July 2017*
doi: 10.1093/nar/gkx613

A structural map of oncomiR-1 at single-nucleotide resolution

Saikat Chakraborty[1] and Yamuna Krishnan[2,3,*]

[1]National Centre for Biological Sciences-TIFR, Bangalore, Karnataka 560065, India, [2]Department of Chemistry, University of Chicago, Chicago, IL 60637, USA and [3]Grossman Institute for Neuroscience, Quantitative Biology and Human Behavior, University of Chicago, Chicago, IL 60637, USA

Received May 15, 2017; Revised July 03, 2017; Editorial Decision July 04, 2017; Accepted July 05, 2017

ABSTRACT

The miR-17–92a cluster, also known as 'oncomiR-1', is an RNA transcript that plays a pivotal regulatory role in cellular processes, including the cell cycle, proliferation and apoptosis. Its dysregulation underlies the development of several cancers. Oncomir-1 comprises six constituent miRNAs, each processed with different efficiencies as a function of both developmental time and tissue type. The structural mechanisms that regulate such differential processing are unknown, and this has impeded our understanding of the dysregulation of oncomiR-1 in pathophysiology. By probing the sensitivity of each nucleotide in oncomiR-1 to reactive small molecules, we present a secondary structural map of this RNA at single-nucleotide resolution. The secondary structure and solvent accessible regions of oncomiR-1 reveal that most of its primary microRNA domains are suboptimal substrates for Drosha-DGCR8, and therefore resistant to microprocessing. The structure indicates that the binding of trans-acting factors is required to remodel the tertiary organization and unmask cryptic primary microRNA domains to facilitate their processing into pre-microRNAs.

INTRODUCTION

The discovery of RNA interference and microRNAs (miRNAs) has brought to light the remarkable ability of RNA to control gene expression networks by targeting specific mRNAs for destruction or translational repression (1,2). The further realization that several miRNAs are organized in clusters and are co-transcribed, yet processed differentially, raised the possibility that these miRNAs exhibit a complex choreography of maturation. An important system in which to probe how such activation is achieved is the human oncogene oncomiR-1, also called the primary microRNA cluster 17–92a (pri-miR-17–92a), which is an ∼800 nucleotide long RNA transcript located in the third intron of the *c13orf25* gene (3–6). Oncomir-1 has the capacity both to spur and contain tissue growth by differentially expressing its constituent miRNAs in specific spatial as well as temporal patterns.

OncomiR-1 encodes a set of six tandem miRNAs that are always co-transcribed and are hence produced in equimolar ratios (Figure 1A). Yet, under different physiological conditions, the relative abundance of the mature miRNAs produced from oncomiR-1 differ in various tissues at a given developmental time, and also differ within a given tissue at different developmental times (Figure 1B) (7,8). Aberrant processing of oncomiR-1 alters the relative abundance of these constituent mature miRNAs and, due to the varied and opposing functions of these constituent miRNAs, this leads to pathological conditions such as B-cell lymphomas or Feingold syndrome, a disorder characterized by extreme physical deformity (9,10). For example, one of its constituent miRNAs, miR-17, acts as a tumor-suppressor, while another, miR-92a, promotes tissue growth and angiogenesis (11,12). In fact, the overexpression of miR-92a leads to erythroleukemia, a form of acute myeloid leukemia, which can be reversed by co-expressing miR-17 (13). Such contrasting functions of its constituent miRNAs necessitates stringent control mechanisms to selectively and differentially process oncomiR-1.

The molecular and structural mechanisms that enable context-dependent, differential processing of individual miRNAs from a single transcript remain unknown. Moreover, the elucidation of potential processing mechanisms for clinically important RNA transcripts such as oncomiR-1 is highly important, as potential modulators of these mechanisms could offer new therapeutic routes. We have previously shown that this ∼0.8 kb long RNA transcript folds into a well-defined, compact and globular tertiary structure and that, early in miRNA biogenesis this tertiary structure imposes a kinetic barrier on the rate at which it is processed into its constituent precursor miRNAs (pre-miRNAs) (3).

To understand precisely how the tertiary structure of this RNA transcript impedes processing by the micropro-

*To whom correspondence should be addressed. Tel: +1 773 834 4761; Email: yamuna@uchicago.edu
Present address: Saikat Chakraborty, Molecular Structural Biology, Max Planck Institute of Biochemistry, Martinsried 82152, Germany.

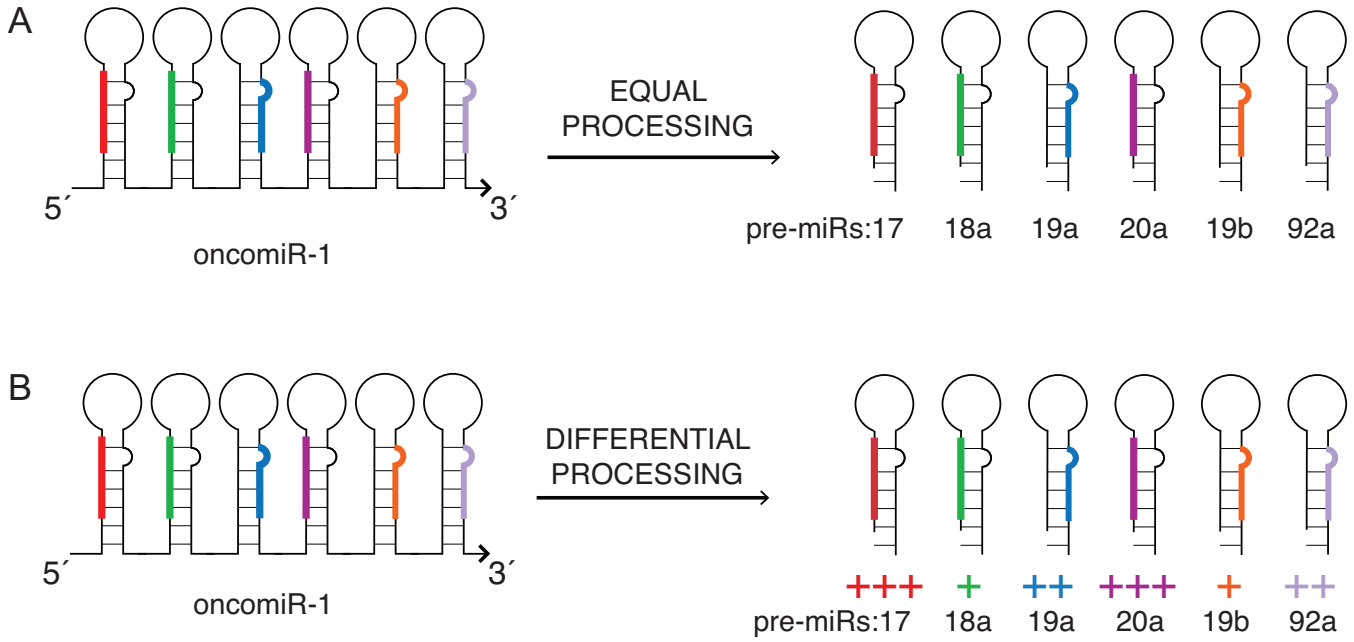

Figure 1. Differential processing of constituent miRNAs in OncomiR-1. OncomiR-1 contains a cluster of six tightly spaced miRNAs that are co-transcribed. (**A**) Uniform processing should yield equal proportions of its constituent miRNAs. (**B**) MiRNAs in oncomiR-1 are selectively processed in different contexts, yielding mature miRNAs in significantly different molar ratios.

cessor complex (MPC), we sought to elucidate the secondary structure of oncomiR-1 at single-nucleotide resolution. To do this, we determined the differential reactivity of base-paired and unpaired nucleotides to chemically reactive small molecules such as benzoyl cyanide (BzCN) and dimethyl sulphate (DMS) (14,15). We could thus pinpoint which nucleotides in oncomiR-1 are paired through their Watson–Crick faces, which nucleotides are paired through their Hoogsteen faces, and which ones are unpaired. By generating hydroxyl radicals in solution that cleave the RNA chain upon contact, we could also decipher which residues on the RNA transcript are exposed to the solvent and which ones are buried (16). Through structure-prediction programs that were guided by this information, along with phylogenetic analysis, we have constructed the first detailed secondary structure map of any primary miRNA cluster.

The secondary structure at single-nucleotide resolution offers new molecular insights into how the tertiary structure of a primary miRNA (pri-miRNA) cluster such as oncomiR-1 can guide the differential processing of its constituent miRNAs. Drosha, the enzyme that processes pri-miRNAs, has two RNase III domain that along with a double-stranded RNA binding domain (dsRBD) recognize a specific motif called a basal helix on the pri-miRNAs, and this recognition is essential for cleavage to occur (17–19). The secondary-structure map revealed that the basal helices of five constituent pri-miRNAs were inaccessible. All the Drosha cleavage sites on oncomiR-1 are known (20). We found that Drosha cleavage sites corresponding to all the individual pri-miRNAs are buried and inaccessible to solvent when they are in the context of their 0.8 kb parent transcript. This indicates that the adoption of tertiary structure limits the access of Drosha to cleavage sites of constituent pri-miRNAs in oncomiR-1. In the absence of any factors that bind in trans, this tertiary structure can intrinsically control the relative processing efficiencies of its individual miRNAs by imposing differential accessibility to Drosha.

This reveals, for the first time, that disruption of the tertiary structure and/or structural remodeling of these pri-miRNA stems is necessary to reveal cryptic basal helices or to expose buried Drosha cleavage sites for Drosha-mediated processing. This suggests the involvement of one or more trans-acting factors to facilitate such structural remodeling triggering a conformational switch that alters the microprocessing outcome. We have identified one such potential switch from the structural map, involving miR-92a, which reconciles two apparently contradictory experimental observations related to the processing of oncomiR-1 (5,6).

MATERIALS AND METHODS

RNA preparation and annealing

Plasmid for generating the templates for transcribing RNA is described in our previous study (3). For biophysical and chemical probing experiments, samples were prepared at desired concentrations by heating the RNA in nuclease free water (Ambion) at 90°C for 3 min followed by flash cooling in ice for 1 min. Then the RNA was dissolved in 50 mM HEPES–KOH pH 7.5, 100 mM KCl and 5 mM $MgCl_2$ and folded by incubating at 37°C for 2 h.

SHAPE analysis

For SHAPE analysis, 0.5 μM folded RNA was taken in folding buffer (50 mM HEPES–KOH, pH 7.5; 100 mM KCl; 5 mM $MgCl_2$). The RNA solution was treated with

one-tenth volume of 3 mM BzCN (dissolved in anhydrous DMSO) at 37°C for 2 s (equal to five hydrolysis half-lives). A no-reagent control reaction was also performed with only DMSO in parallel. The RNA was subsequently precipitated with addition of 2 μl 5 M NaCl, 20 μg glycogen and ethanol. Next, the RNA was resuspended in water and 15 pmol of a 6-FAM labeled primer for the (+) BzCN reactions or 20 pmol of a NED labeled primer for (-) BzCN reactions was added to a total volume of 12 μl. Primers were annealed by heating at 85°C for 3 min followed by flash cooling in ice and then kept at 37°C for 15 min. Extension buffer consisting of 5 mM DTT, 1 mM each dNTP, 50 mM Tris–HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$ and 100 units of 'SuperScript' III Reverse Transcriptase (Invitrogen) was added to the annealed RNA-primer mixture on ice, and the reactions were incubated for 5 min at 37°C, 30 min at 52°C, and 5 min at 65°C. Reactions were quenched by addition of 4 μl 1:1 mixture of 50 mM EDTA, 1.5 M sodium acetate (pH 5.3) and placed on ice. Dideoxy sequencing ladders were generated by primer extension using unmodified RNA and primers labeled with VIC in the presence of 0.25 mM ddCTP keeping dNTP: ddNTP ratio 2. (+) and (–) reagent and sequencing reactions were combined and cDNAs were precipitated with ethanol, sodium acetate and 20 μg glycogen; pellets were dried and resuspended in 10 μl of deionized formamide (Ambion); fluorescently labeled cDNAs were resolved by capillary electrophoresis using a ABI 3130xl genetic analyzer.

Raw sequencing traces from the ABI 3130XL capillary electrophoresis instrument were analyzed using 'ShapeFinder' (21). In brief, 'ShapeFinder' was used to adjust the baseline, correct different dye induced mobility shifts, perform signal decay correction and scale the (+) and (–) reagent traces. SHAPE reactivities of each nucleotide in the (+) and (–) reagent traces were quantified by Gaussian integration, and were placed on a normalized scale by discarding the top 2% of the most reactive peaks and dividing by the average intensity of the next 8% of peaks. This calculation renders the data on a scale of 0 to ∼2 where 1.0 is the average intensity of highly reactive peaks. 'RNAstructure' was used to generate structure models for the RNA using SHAPE data as pseudo-free energy change terms (slope and intercept were 1.8 and -0.6 kcal/mole, respectively). Due to noise in electrophoretogram, last ∼50nts were excluded from the analysis.

DMS probing

For DMS (dimethyl sulphate) probing experiments, 25 μl 0.5 μM folded oncomiR-1 in folding buffer (50 mM HEPES–KOH, pH 7.5; 100 mM KCl; 5 mM $MgCl_2$) was reacted with 1 μl of 1.76 M DMS (final concentration, 68 mM) dissolved in 95% ethanol for 1 min at 27°C. No-reagent control was performed with addition of 1 μl of 95% ethanol only. Reaction was quenched with 4 μl of 14.6 M β-mercaptoethanol. RNA was precipitated by addition of 50 μl of 0.3 M sodium acetate, pH 5, 250 μl cold absolute ethanol and 4 μl of 5 mg/ml glycogen. After precipitation, (+) reagent and (–) reagent reaction mixture were used for extension using FAM and NED labeled primers respectively. Sequencing ladder using ddCTP was generated with VIC-labeled primer. Primer annealing and reverse transcription was carried out as outlines in the previous section. Capillary electrophoresis, data treatment was exactly similar as in SHAPE analysis.

Hydroxyl radical foot printing

For hydroxyl radical footprint, 0.5 μM folded RNA was taken in folding buffer (25 mM sodium cacodylate, pH 7.0; 100 mM KCl; 5 mM $MgCl_2$). Folded RNA was reacted with 1 μl each of 3.75 mM Fe(II)–EDTA, 150 mM sodium ascorbate and 0.3% H_2O_2 for 15 s. Concurrently, no-reagent control was performed with only ascorbate and H_2O_2 (–·OH). After precipitation, (+) reagent and (–) reagent reaction mixture were used for extension using FAM and NED labeled primers respectively. Reverse transcription and capillary electrophoresis protocol is similar to SHAPE analysis described above. Due to noise in electrophoretogram, nucleotides 390–470 and 775–826 were excluded from the analysis.

Small angle X-ray scattering

Samples for SAXS were annealed and folded as described for above. Then, they were subjected to size-exclusion chromatography (Superdex 200, GE) in RNA folding buffer. The primary miRNA cluster under study at high concentrations exhibited fractions of aggregated and monomeric RNA. Fractions containing the monomers were pooled and concentrated using Amicon Ultra (Millipore). Samples were diluted to a final concentration of ∼0.5 and 1.0 mg/ml prior to shipment of the sample to the synchrotron facility. SAXS experiments were conducted at the Advanced Photon Source ID-12 BESSRC. Data were collected in the same manner as described elsewhere (22). No concentration dependence of R_g was observed in 5 mM Mg^{2+}. For single measurements, the uncertainty in the R_g reported is taken from the fit to the data in the Guinier plot. SAXS data analysis was performed using either ATSAS package (23), Igor (Wavemetrics) or GNOM (24).

Bioinformatics analysis

For phylogenetic analysis, full length pri-miR-17–92a sequences and their corresponding alignments for 36 different species were retrieved from *Ensemble* data base. Also, for pri-miR-17, 19a and 92a stem loops, sequence alignments are retrieved from Rfam database. Alignments are further checked manually or using Clustal W in 'Jalview 2.9' program. This alignment is used to generate a starting phylogenetic model using 'RNAalifold v2.3.1' in Vienna RNA package (25,26). Output from RNAalifold was visualized using VARNA v3.92 (27). Covariations, base pairing probabilities and conservation of base pairs are determined from RNAalifold output file as well as manual checking in the alignment files. All RNA structure visualizations are done using VARNA. Positional entropy was calculated from 'RNAfold v2.3.1' program in Vienna RNA package with SHAPE constraints using standard settings.

RESULTS

SHAPE analysis of OncomiR-1

In a previous study (3), we have outlined a smaller-scale production and folding protocol, which yielded a homogeneous population of well-folded oncomiR-1, as evidenced by both size exclusion chromatography and small angle X-ray scattering (Supplementary Figure S1).

SHAPE (selective 2′-hydroxyl analyzed by primer extension) on RNA molecules, uses benzoyl cyanide (BzCN), a small molecule that selectively targets ribose 2′-hydroxyl groups of those nucleotides that are present in single stranded regions of a given structured RNA (Supplementary Figure S2) (28,29). SHAPE has been used extensively to elucidate the secondary structure of large RNA molecules such as the HIV RNA genome, and long noncoding RNAs such as HOTAIR and steroid receptor coactivator RNA (SRA) (30–32). SHAPE reactivity of well-folded 0.8 kb oncomiR-1 cluster was measured at single-nucleotide resolution using three different primers that bind to the three indicated regions in the transcript (Supplementary Table S1). The SHAPE reactivity of each nucleotide was used to provide constraints for secondary structure prediction by 'RNAstructure' (33) (Supplementary Figures S3 and S4A).

Next, dimethyl sulphate (DMS) was used to probe oncomiR-1 to verify the secondary structure map derived from the SHAPE data (15). DMS methylates adenosines and cytidines that are not base-paired. We analyzed DMS reactivity using the same analysis methods described for SHAPE reactivity (see methods) to reveal the base pairing status of adenosines and cytosines (15,32). The extent of DMS reactivity of most reactive nucleotides on oncomiR-1 directly correlated with their SHAPE reactivity. Only 9% of the nucleotides on oncomiR-1 showed high SHAPE reactivity and low DMS reactivity, or vice versa, possibly corresponding to regions of tertiary structure.

The secondary structure map corresponding to the lowest empirical free energy computed by the program 'RNAstructure' with SHAPE constraints, is represented in Figure 2A (also Supplementary Figures S3 and S5). In Figure 2A, the SHAPE and DMS reactivities of a given nucleotide position are categorized into either low, intermediate or high reactivity ranges respectively corresponding to values of 0–0.3 (duplex), 0.3–0.7 (possibly single stranded) and ≥0.7 (single stranded) (Figure 2A, Supplementary Figure S5). Noticeably, many regions show low reactivity, consistent with them being in a duplexed state. Moderate and highly reactive nucleotides are found mainly in loops and internal bulges, and are consistent with them being single stranded. This agrees with similar observations in other structured RNAs also analyzed by SHAPE (30,31). The accuracy of this SHAPE-based structure was compared with the structure predicted from phylogenetic analysis as well as the minimum free energy model generated by the 'RNAstructure' algorithm (33) (Supplementary Figures S3 and S4B). Importantly, base-paired regions that are predicted to be evolutionarily conserved by phylogenetic analysis are also observed experimentally by SHAPE analysis as discussed later.

OncomiR-1 has a highly helical architecture

To describe its structure effectively, we divide oncomiR-1 into six empirical pri-miRNA domains with each domain containing a pre-miRNA helix with almost always an additional one helical turn beyond the Drosha-cleavage site. These domains are referred to P-I to P-VI as shown in Figure 3 (also Figure 2A), which also identifies the miRNA in each domain, account for much of the helical structure in oncomiR-1. Additionally, two smaller stem loop structures found between P-I and P-II and between P-V and P-VI, are referred to as NPSL1 and NPSL2 respectively (non-pre-miR stem loop) since they do not host a pre-miRNA sequence.

The secondary structure of oncomiR-1 determined by SHAPE confirmed that it is highly structured, with more than 70% of the nucleotides being base paired (Figures 2A and 3). Most unpaired nucleotides that show high SHAPE reactivity occur either in terminal loops, junctions of helices or internal bulges. The secondary structure is dominated by long helical domains that are capped by terminal loops. OncomiR-1 contains 61 helical segments, 31 internal loops, 10 terminal loops and 5 junction regions and is comparable to other highly structured RNAs, such as 16S rRNA and SRA (Supplementary Table S2). The sequences flanking the 5′ and 3′ regions of five of the six pre-miRNA regions are helical. These flanking regions extend at least one helical turn beyond the Drosha-cleavage site of a constituent pri-miRNA and correspond to the basal helix of each pre-miRNA domain. Basal helices are a critical structural requirement for effective primary miRNA processing by MPC to yield the resultant pre-miRNA (19,34).

Features of the secondary structure map indicate that oncomiR-1 adopts a tertiary structure

Our secondary structural map of oncomiR-1 reveals key features that point toward the adoption of tertiary structure by this RNA (Figures 2A and 3). The structural map shows that oncomiR-1 possesses highly stable helices that are nearly ∼10–11 nm long. Yet, electron microscopic and dynamic light scattering indicate that oncomiR-1 adopts a globular state of ∼6–8 nm radius (5,6). Thus, to achieve these dimensions, these helices must undergo dramatic compaction through numerous tertiary contacts. We know that oncomiR-1 undergoes a folding reaction to achieve this globular form (3). Such folding should also result the burial and protection of substantial regions of this RNA from the surrounding solvent, as seen in other structured RNAs (35,36).

A notable feature of oncomiR-1 are the long purine-rich terminal loops that cap every helical domain. Each domain P-I to P-VI, NPSL1 and NPSL2 possesses a loop of length between 5 and 11 nt (Figures 2A and 3). These loops are larger in size compared to the thermodynamically predicted pri-miRNA structures. SHAPE reactivity of the terminal loop nucleotides in the oncomiR-1 cluster revealed that a few terminal loops in oncomiR-1 are single stranded, yet are highly conformationally constrained, e.g. the loops of pre-miR-92a, pre-miR-19b, NPSL1 and NPSL2. Some of these loops show a preponderance of adenosines, which together indicate their involvement in mediating tertiary con-

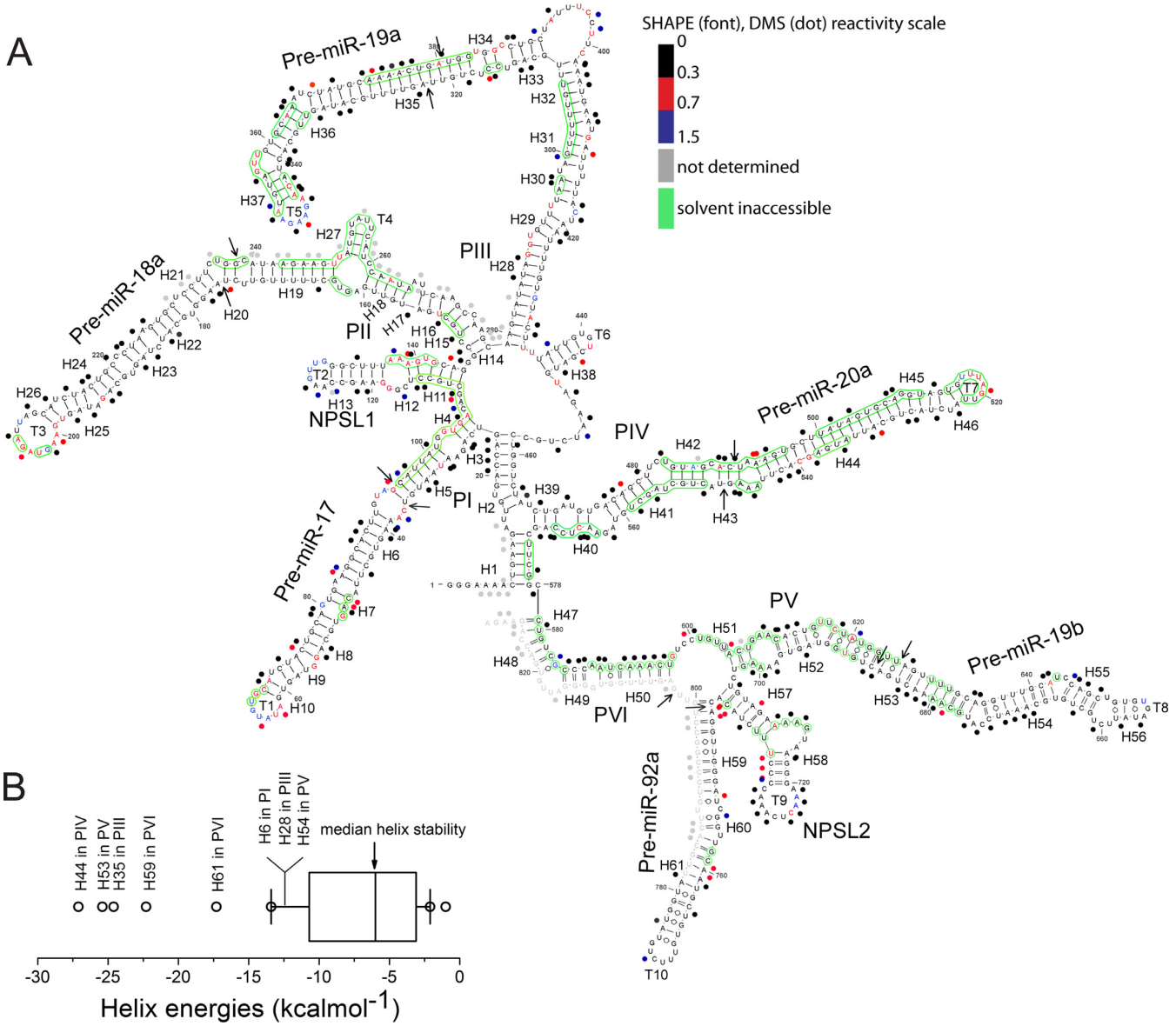

Figure 2. SHAPE analysis of oncomiR-1. (**A**) SHAPE derived secondary structure model of oncomiR-1 cluster. The nucleotides are color coded according to SHAPE (font) and DMS (dot) reactivity. Solvent inaccessible nucleotides are shown in green circles. Pri-miRNA domains (P) and structural elements (H, helix; T, terminal loops) are indicated in the figure. Black arrows indicate Drosha cleavage sites. (**B**) Thermodynamic stability of all the helices constituting oncomiR-1 cluster. Free energy calculated using nearest neighbor analysis from the SHAPE derived model structure.

tacts (37). For example, NPSL2 is a ∼37 nt long domain that folds into a stem loop structure. Six out of the ten nucleotides in its loop are adenosines, all of which show low SHAPE reactivity. Overall, the NPSL2 domain has 15 adenosines of which 10 are unpaired (A_U) giving an unpaired:paired adenosine ratio (A_U/A_P) ratio of 1.5, which is characteristic of adenosine mediated tertiary contacts. However, most terminal loops show relatively high flexibility, suggesting that those SHAPE unreactive loops may not be involved in specific tertiary contacts. Flexible terminal loops may be involved in functions distinct from mediating tertiary structure such as trans-acting protein binding site. For example, the flexible 11 nt terminal loop of pri-miR-18a recruits hnRNP-A1 which facilitates the processing of pri-miR-18a (38).

The presence of internal loops and bulges along the long helical pri-miRNA containing domains is another characteristic feature in oncomiR-1(Supplementary Table S3). Bulges are critical structural motifs because their incorporation within a helix changes the helix axis, produces a kink in the helical RNA segment, enabling RNA compaction (39–42). Many of the bulges in oncomiR-1 show moderate to highly reactive residues. However, some residues located in bulges show low SHAPE reactivity, indicating that they are conformationally constrained, suggesting their possible involvement in stacking or tertiary contacts. A few examples include U213 and C235-U236 in the P-II helix, A365-U366

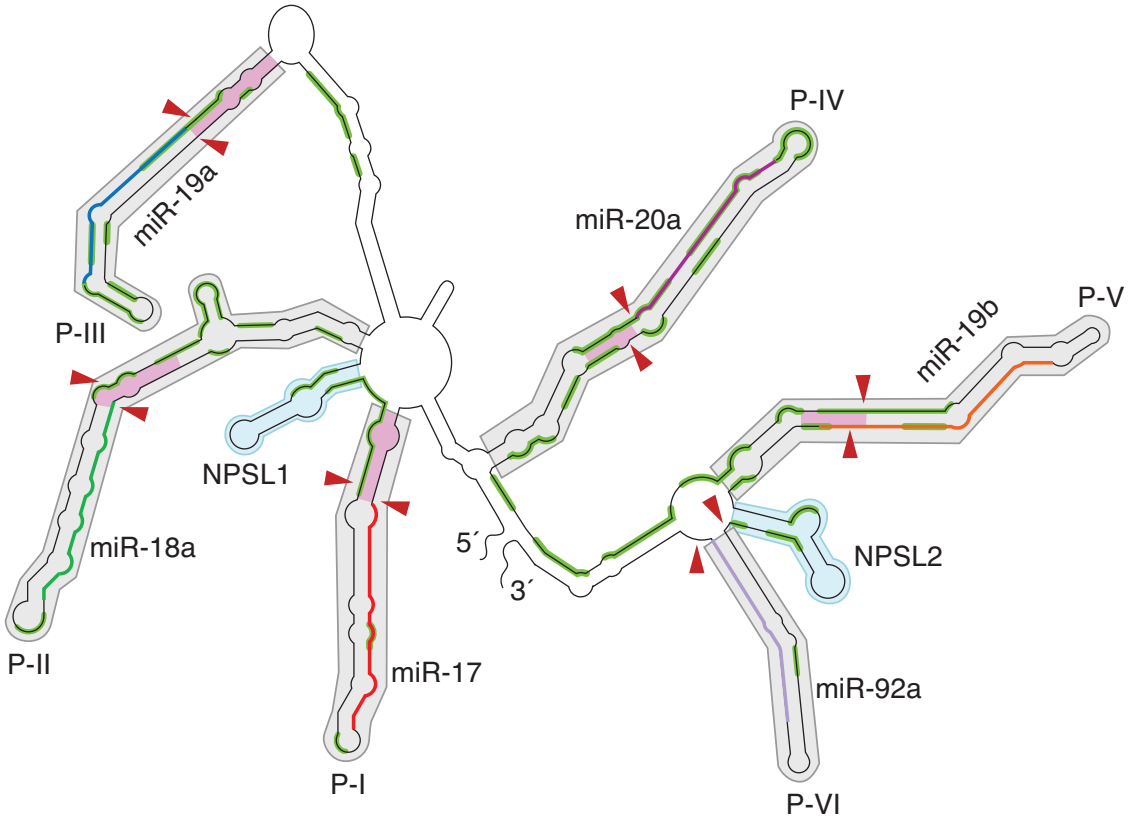

Figure 3. Schematic of the structural map showing domain architecture in oncomiR-1. OncomiR-1 is empirically divided into helical domains, P-I to P-VI (grey shaded regions), containing individual pri-miRNA helices and two non-pri-miRNA stem loop domains NPSL1 and NPSL2 (cyan shaded regions). The regions corresponding to mature miRNA sequences in P-I to P-VI are indicated as colored segments (e.g. miR-17 is in red, miR-19a in blue, miR-19b in orange etc.). Red arrowheads indicate Drosha cleavage sites on oncomiR-1. The pink shaded regions indicate the basal helix of each pri-miRNA domain. Solvent inaccessible segments on oncomiR-1 are highlighted in green.

in the P-III helix and U659-U661 in the P-V helices (Figure 2A).

Thermodynamic stabilities of helical domains hint at structural rearrangement

Thermodynamic stability profiling of all helical segments suggests a potential role for its secondary structure and its associated energetics in regulating pri-miRNA processing. The basal helix of a pri-miRNA is a critical target for the processing machinery since any structural perturbation in the basal helix architecture is known to impede pri-miRNA processing (18,19,34). All the basal helices of oncomiR-1 contain a number of mismatches and are marginally stable compared to helical pre-miRNA segments. They are probably further stabilized by coaxial stacking against the pre-miRNA helices (Figures 2B and 3). In the cluster, consecutive pri-miRNA domain is linked by conserved sequences that can also form hairpin structures such as NPSL1, as well as by unstructured/flexible regions. Interestingly, the latter regions sometimes sequester basal helix forming sequences to form hairpin loops such as NPSL2 and can possibly modulate the basal helix structure of pri-miRNAs leading to impaired miRNA processing. Notably, NPSL1 and NPSL2 have stabilities in the range of 9–12 kcal/mol and are much less stable than their flanking pri-miRNA helices that have stabilities in the 25–35 kcal/mol range (Figure 2B). In fact, in other higher energy predicted conformations, NPSLs lose their identity and merge with the neighboring pri-miRNA stems only with a few kcal/mole energy loss relative to the most stable conformation. As discussed later, these kind of naturally engineered bi-stability can significantly affect the overall pri-miRNA helix architecture as well as its recognition and cleavage by the MPC.

Hydroxyl radical foot-printing reveals that Drosha cleavage sites are solvent inaccessible

Since all the Drosha cleavage sites within oncomiR-1 are known, we sought to gauge the accessibility of these sites in the context of the whole transcript. Therefore, we subjected folded oncomiR-1 to hydroxyl radical footprinting, which yields a measure of the solvent exposure of each nucleotide position. Here, hydroxyl radicals generated *in situ* cleave nucleotides on the RNA backbone in proportion to their solvent exposure (43). Hydroxyl radicals were generated *in situ* using Fe(II)-EDTA, H_2O_2 in a solution of *in vitro* transcribed, folded oncomiR-1 in the presence of Mg^{2+} and analysed by using primer extension using fluorescently labelled primers that were then resolved by capillary electrophoresis. This yielded cleavage information at single-nucleotide resolution that reported on the solvent exposure

of ∼80% of the nucleotides present in the oncomiR-1 (Figure 2A, green circles). Cleavage intensities were normalized on a scale of 0 to 1.5, where 1.0 is defined as the average intensity of highly reactive, or highly exposed nucleotides. On this scale, nucleotides with reactivity ≤0.5 of the mean reactivity are considered solvent inaccessible based on the analysis of previously investigated RNA structures (44,45).

A significant portion, corresponding to nearly 30% of the residues on oncomiR-1, proved to be solvent inaccessible (Figures 2A and 3). This finding is consistent with the adoption of tertiary structure. The longest tract of solvent inaccessible nucleotides is A243-A265 that encompasses the basal helix of pri-miR-18a in the P-II domain. Many solvent-protected sites are present on inter pri-miRNA stem loop regions. A case in point is the inter pri-miRNA domain lying between pri-miR-17 and pri-miR-18a comprising the residues C-110 to C-114 and A-137 to G-142 that are part of NPSL1. This reveals that the NPSL1 stem and internal loop get buried due to tertiary structure formation by oncomiR-1. The internal bulge of NPSL2 is also completely buried (Figures 2A and 3) and the adenosine residues present here are unreactive to both BzCN and DMS. This indicates that despite being single stranded, they are also conformationally constrained and hydrogen bonded, strongly implicating their involvement in tertiary structure contacts.

The hydroxyl radical footprint shows that regions comprising the Drosha cleavage sites of at least three constituent pri-miRNAs - miR-18a, miR-19a and miR-20a - are solvent inaccessible in folded oncomiR-1. In fact, in P-IV domain nearly all of the pri-miR-20a helix is buried, indicating that a trans-acting factor is likely to be required to provide access to Drosha for pri-miR-20a to be processed. In fact, KH-type splicing regulatory protein (KSRP) is known to bind pri-miR-20a and facilitate its processing, although how KSRP accesses pri-miR-20a in the context of the fully folded transcript is as yet unknown (Supplementary Figure S6) (46). In other examples, the regions U236-C239, A243-U248, A373-G383 and G615-U625 found in the P-II, P-III and P-V helices respectively show that one strand of these helices bearing Drosha cleavage sites is solvent inaccessible while the opposite strand is solvent accessible (Figure 2A), indicating the requirement of trans-acting factors to remodel the structure to allow accessibility to Drosha. This is supported by the fact that the MPC itself comprises many RNA helicases (47).

Another striking feature of the hydroxyl radical footprint is that the basal helix regions of nearly every pri-miRNA is solvent inaccessible. In pri-miR-17 and pri-miR-18a one of the strands of the basal helix is solvent inaccessible, while in pri-miR-20, pri-miR-19b, the basal helices are almost completely buried. Pri-miR-92a is particularly dramatic - our structural map shows that its basal helix is absent, and the sequences that should have comprised its basal helix form part of a four-way junction. This strongly supports the requirement of trans-acting factors to expose Drosha recognition motifs critical for processing.

The structural map also reveals additional evidence that trans-acting factors are necessary to process key constituent miRNAs. For example, the binding sites for hnRNP A1 on the terminal loop of pri-miR-18a comprising G202-U203 and G205-A206 (Supplementary Figure S6) were found to be solvent inaccessible in folded oncomiR-1. This reaffirms that disruption of the tertiary structure of oncomiR-1 must occur in order to recruit hnRNP A1 for pri-miR-18a processing.

OncomiR-1 presents a number of potential sites for A-minor interactions - a motif that mediates tertiary structure in RNAs (37,48,49). We can identify several stretches of solvent inaccessible U-A base pairs that can function as A-minor receptors, e.g. positions 325–328:373–376 and 628–631:679–682, on pri-miRs 19a and 19b respectively (Figure 2A). In addition, there are several proximal, unpaired, solvent inaccessible adenosines that are conformationally constrained. Yet they are also DMS reactive indicating their involvement in interactions via their Hoogsteen faces. The profusion of such adenosines and A-minor receptors indicates significant potential for A-minor interactions in oncomiR-1 (50).

Phylogenetic model supports SHAPE derived solution structure

Phylogenetic analysis of oncomiR-1 is in good agreement with the experimentally derived secondary structure map as shown in Figure 4A. This conformity strongly indicates that Figure 2A represents the most predominant structure of oncomiR-1 cluster in solution. Up to 36 sequences from different phyla were used to derive a phylogenetic model using covariations in nucleotides involved in base pairing (Figure 4A, Supplementary Figure S7). We mapped covariations that are Watson–Crick base paired and mutated to other Watson–Crick base pairs, (green, Figure 4A). We also looked for mutations that convert Watson–Crick base pairs to a GU wobble pair and vice versa, referred to as half flips (cyan, Figure 4A). We also pinpointed highly conserved base pairs (red, Figure 4A).

At the sequence level, oncomiR-1 is highly conserved, with primate sequences being >90% conserved. The highly conserved base pairs are mostly limited to pre-miRNA domains found in regions P-I to P-VI (red, Figure 4A) that are invariant in order to retain the capacity to recognize their mRNA targets. Thus, there are fewer mutations in the mature miRNA sequences and more in the miRNA* sequences. For half flips in the miRNA/miRNA* regions, the half-flipped nucleotide is nearly always in the miRNA* region. For example, in P-V, containing pri-miRNA 19b-1, the miRNA/miRNA* region contains three half flips, with all of them occurring in miRNA*. All the pre-miR-containing domains, i.e. P-I to P-VI, possess at least three covariant base pairs confirming their helical structure.

Covariations and base pairing probability analyses reveal the importance of the basal helix regions in oncomiR-1. Covariations are abundant at nucleotide positions corresponding to the basal helix of every pri-miRNA. For instance, in P-I, ∼25% of all covariations are located in the basal helix of pri-miR-17, indicating the importance of helicity in this region (18,19). This is supported by an analysis of base pairing probabilities (BPP) as well as sequence conservation of all the basal helices in oncomiR-1, which are plotted as a function of nucleotide position, where zero represents the Drosha cut site (Figure 4B). The high BPP of the basal helical regions despite their low sequence conservation reveals

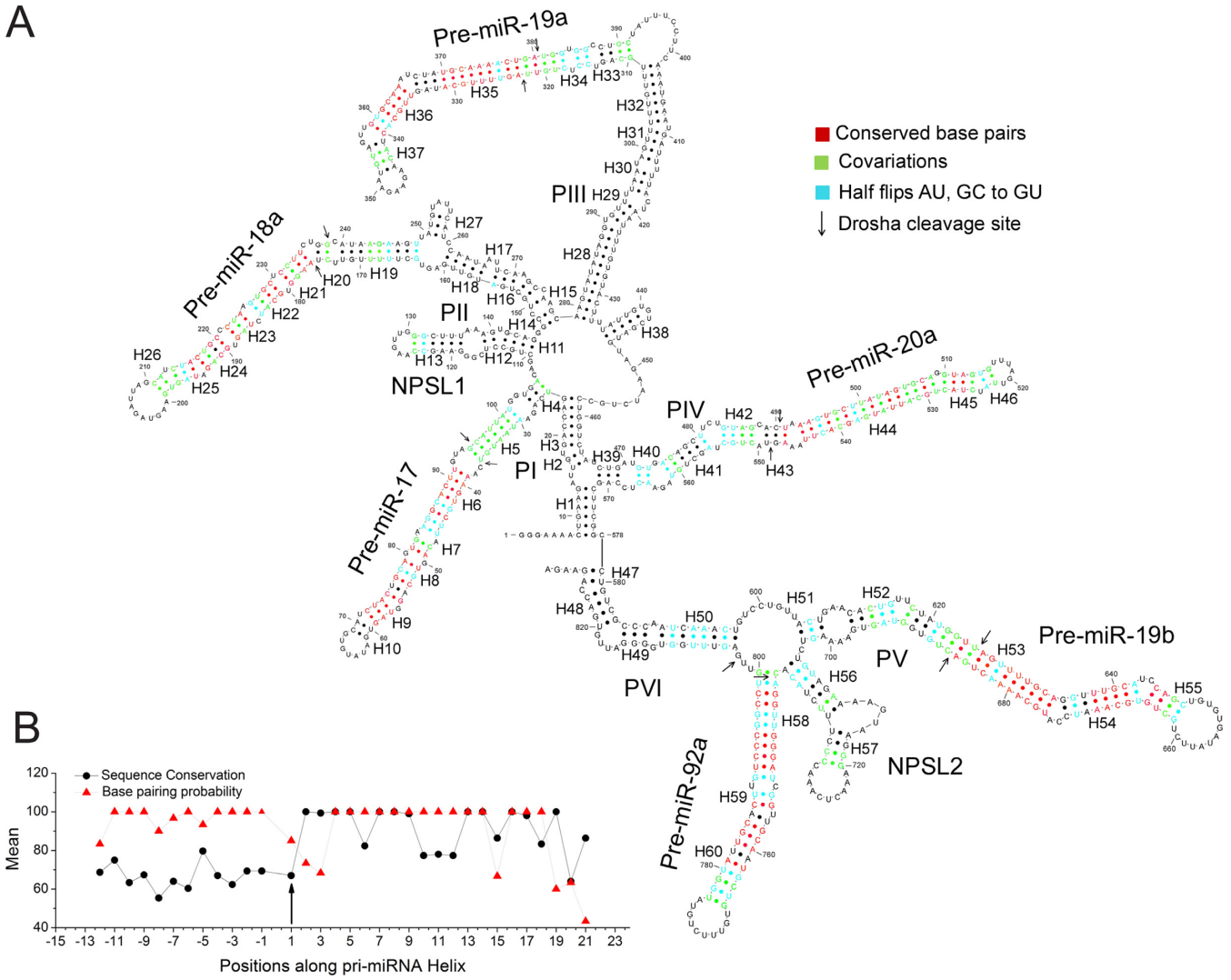

Figure 4. Phylogenetically predicted secondary structure of oncomiR-1. (**A**) Phylogenetically derived secondary structure model of oncomiR-1 cluster. Conserved and co varying base pairs are color coded. (**B**) Average base pairing probability and conservation of the basal helix nucleotides for all the pri-miRNA helices are shown as a function of their position along the pri-miRNA helix. Black arrows indicate Drosha cleavage sites.

the importance of helicity in these regions for the purpose of processing.

DISCUSSION

Differential accessibility of pri-miRNAs to the microprocessor complex

The most notable feature emerging from the structural map of oncomiR-1 is that several constituent pri-miRNAs of the oncomiR-1 cluster are not readily accessible for processing. We had shown previously that oncomiR-1 folds into a higher order structure that impeded the processing of individual pri-miRNAs and that disruption of this higher order structure lifted this impediment (3). However, the reason for this was unclear. Our findings reveal that the Drosha cleavage sites and/or basal helices of many constituent pri-miRNAs are solvent inaccessible due to tertiary structure formation by oncomiR-1. Pri-miR-20a and pri-miR-19b are the most solvent inaccessible miRNAs. These two domains along with pri-miR-92a constitute a buried '3′-core' as hypothesized by Chaulk et al (5). Thus, the tertiary structure of the pri-miRNA cluster provides a first layer of regulation by making selected Drosha cleavage sites inaccessible for processing.

Individual pri-miRNA domains exist as proto-substrates in the transcript.

We address with specific examples how a trans-acting factor might regulate pri-miRNA processing, based on the experimentally derived structural map of oncomiR-1 at single-nucleotide resolution. Selected individual pri-miRNA domains, such as pri-miR-18a, are present as suboptimal ''proto-substrates'' in oncomiR-1, in conformations that are expected to be refractory to processing. The stem-loop structure of such 'proto-substrates' show subtle structural differences that results in the overall architecture of the pri-

miRNA domain being a sub-optimal substrate for Drosha (38,51).

For example, pri-miR-18a, which is a constituent pri-miRNA of oncomiR-1, has a paralog called pri-miR-18b (38). The sequence of pri-miR-18a makes it intrinsically a proto-substrate, in contrast to pri-miR-18b, as slight differences in their sequences structure their stems differently (Supplementary Figure S8). The experimentally derived structure of just the pri-miR-18a domain by Michlewski et al differs slightly from its structure when it is part of oncomiR-1 (38). The isolated pri-miR-18a domain cannot be processed, and requires the binding of hnRNP A1 to remodel it into the stem structure found in pri-miR-18b, for its processing (38). However, pri-miR-18b can be processed as is, without hnRNP A1 and is not a proto-substrate (38). When it is part of oncomiR-1, pri-miR-18a also exists in a non-processable, proto-substrate form, due to the CU bulge and GU wobble near the Drosha cleavage site. Thus, a relatively minor rearrangement must occur to remodel this CU bulge into a pri-miR-18b-like stem (Supplementary Figure S8) which is a good substrate for Drosha that is likely to be facilitated by hnRNP A1. Interestingly, proto-substrates such as the pri-miR-18a, is the least thermodynamically stable among pri-miR helices indicating capacity of being remodeled probably scales with the helical stability.

A conformational switch that regulates processing

We found a novel example of another 'proto-substrate' in the case of pri-miR-92a that constitutes a conformational switch that can turn pri-miR-92a processing off or on. In fully folded oncomiR-1, pri-miR-92a adopts a structure that does not have a basal helix since its Drosha cleavage sites are sequestered in a four-way junction (Figure 2A, Supplementary Figure S9). SHAPE analysis reveals that NPSL2 must exist as a hairpin to invoke a tertiary contact with NMSL-R observed by Chaulk et al. (A679-A682 on pri-miR-19b) (50). In order to do so, NPSL2 hairpin sequesters a crucial part of the sequence that is supposed to form the basal helix of pri-miR-92a. Thus, in fully folded oncomiR-1 with tertiary structure, NPSL2 exists as a hairpin and pri-miR-92a has no basal helix. The latter is therefore refractory to processing by Drosha. This explains why pri-miR-92a is the least processed of all constituent miRs *in vitro*. Interestingly, mutated oncomiR-1 transcripts where the pri-miR-19b or the NPSL2 hairpin is either deleted or, where the unpaired adenosines of NPSL2 are mutated, showed significant de-compaction and enhanced pre-miR-92a expression (5,50). Thus, our structural map explains why pri-miR-92 cannot be processed in fully folded oncomiR-1. We therefore explored predicted, alternate conformations where pri-miR-92a might become a substrate for Drosha.

We found a higher energy conformation predicted by Mfold (52) where pri-miR-92a has a basal helix, in which nearly 55% of base pairs were conserved compared to the lowest energy SHAPE derived structure. This higher energy structure showed base-pairing between the NPSL-2 domain (green segment, Supplementary Figures S9, S10) and the 3′ end of the transcript (magenta segment, Supplementary Figures S9, S10) that together generate a basal helix for pri-miR-92a (6). This could potentially turn the processing of pri-miR-92a off or on, and could explain apparently contradictory observations of oncomiR-1 processing across different contexts (Supplementary Figure S11). This conformer was higher by only 9 kcal/mol compared to the lowest energy structure in Figure 2A. This energy difference is comparable to the free energies of ligand binding to non-coding RNAs that undergo conformational switching, e.g. riboswitches (Supplementary Table S4). It is thus energetically possible for a trans-acting factor to bind and switch the thermodynamically favored structure shown in Supplementary Figure S10A into the conformation shown in Supplementary Figure S10B.

The NPSL-2 domain forms the heart of this conformational switch, serving as a toggle to switch on miR-92a processing. It either exists in a stem loop structure invoking tertiary interactions to compact oncomiR-1 thereby, make miR-92a refractory to processing (50), or it can disassemble and generate the basal helix of pri-miR-92a. NPSL2 is conserved across phyla with abundant covariation to support its secondary structure (Supplementary Figure S12). Importantly, despite such high covariation, the free energy of the NPSL-2 domain predicted by Mfold remains the same across phyla (∼–9.3 kcal/mol), indicating that the energy difference between both conformations in Supplementary Figure S9 is critical to preserve.

The existence of a structural switch reconciles two different previous observations where oncomiR-1 is processed differently (6,50). Chaulk et al. have shown that pri-miR-92a is processed least of all *in vitro* (5). The lowest energy structure of oncomiR-1 where pri-mir-92a is without a basal helix and highly solvent protected, is also consistent with the findings of Chaulk *et al.* who have described an inaccessible '3′ core" for this RNA (5,50).

Yet pri-miR-92a is constitutively expressed in mouse embryonic stem cells, suggesting the structural barriers that impede its processing are somehow lifted endogenously (6). Thus, Gregory et al. hypothesized that, the 5′-segment of murine oncomiR-1 (5′-RD, red segment, Supplementary Figure S10B) base pairs with the region between pri-miR-19b and pri-miR-92a defined as 3′-RD*, (green segment, Supplementary Figure S10B), which corresponds to the NPSL2 domain in our SHAPE derived structure (green segment, Figure 5A) (6). Based on an endonucleolytic cleavage event that removes the 5′-RD domain, they postulated that association between 5′-RD and 3′-RD* results in an overall conformation that resists miRNA processing (6). Cleavage of 5′-RD by ISY1 and pri-miR-92a by the MPC results in a new species called the 'progenitor miRNA', that permits the processing of its constituent pri-miRNAs (6). They predicted a model for the oncomiR-1 in mouse ESCs where pri-miR-92a has a basal helix comprising NPSL2. This is consistent with the higher energy conformation we have identified, where a region corresponding to the NPSL-2 (dark blue segment, Supplementary Figure S10B) base pairs with the 5′-RD.

The existence of a structural switch also supports findings from other groups where polyadenylated oncomiR-1 clusters are often stored in subnuclear foci containing sc35 splicing factors and processed on cue (53). It is possible that polyadenylated oncomiR-1 adopts a structure similar

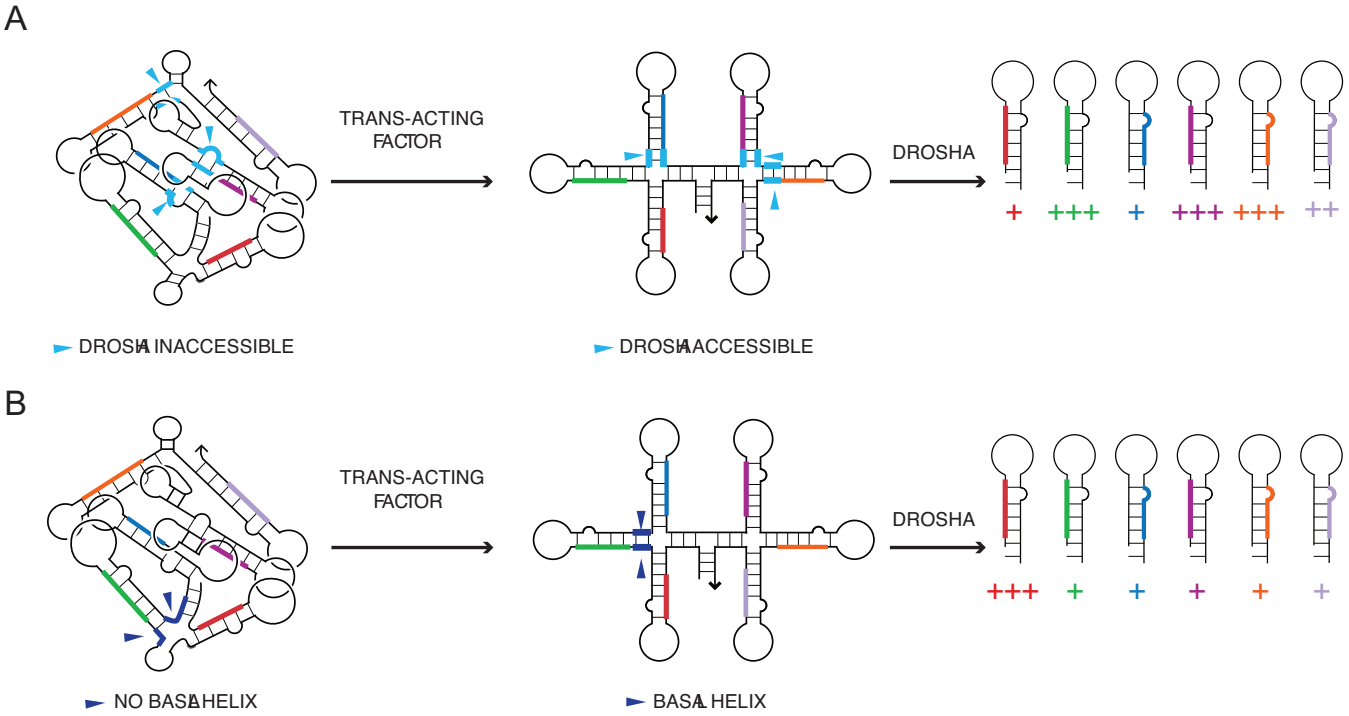

Figure 5. Modes by which tertiary structure of pri-miRNA influences differential processing of constituent miRNAs. (A) Pri-miRNA domains can have their basal helices (cyan segments) and/or Drosha cut sites buried, making them inaccessible to Drosha/DGCR8, impeding their microprocessing. Binding of a trans-acting factor is needed to unfold the tertiary structure to provide accessibility to Drosha. (B) In the fully folded structure, pri-miR-92a does not have a basal helix (dark blue segments). The binding of a trans-acting factor is necessary to generate a basal helix and make the primary microRNA domain a substrate for Drosha.

to Figure 3A where nearly every pri-microRNA is expected to be refractory to processing and represents a storable form of pri-miRNA that can be switched into a process-able form by trans-acting factors when required.

Thus, the ability of oncomiR-1 to form folded 3D structures enables differential presentation of the miRNA stems for processing. Distinct ways to unfold the globular pri-miRNA cluster could reveal different pri-miRNA stems that can be processed for transcriptional control. One of these could include a trans-acting factor simply disrupting the tertiary structure to reveal solvent inaccessible basal helices and Drosha cleavage sites as seen in pri-miR-20a (Figure 5A). Alternatively, binding of the trans-acting factor can cause more extensive structural remodeling to generate a basal helix as discussed for pri-miR-92a (Figure 5B). The remodeling can depend on the presence and/or relative abundance of protein factors in specific cellular environments that would impose differential processing.

Diverse strategies are now being employed to synthetically regulate the biogenesis of miRNAs from pri-miRNA scaffolds through the use of aptamers or peptoid ligands that target specific parts of the pri-miRNA, the binding of which can modulate subsequent processing steps (54–57). Not surprisingly, the success of such strategies hinge on *a priori* knowledge of the structure of the target primary miRNAs. The structural elucidation of a clinically significant oncomiR-1 cluster has broad implications in terms of developing synthetic strategies to modulate its structure and thereby activity.

Compared to our understanding of the RNAi pathway, where structural insights have been instrumental, our understanding of miRNA biogenesis has been so far been impeded by the unavailability of detailed structures of pri-miRNAs and processing enzymes either individually or in complex. The present solution structure of oncomiR-1 should serve as a roadmap for the future studies that seek to address the very first steps of miRNA biogenesis from clustered pri-miRNAs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Professors Joseph Piccirilli and John Kuriyan as well as Dr Shabana Shaik for valuable comments and suggestions, the Sequencing Facility at NCBS for capillary electrophoresis, Nathan Baird for SAXS data acquisition. S.C. thanks CSIR, Govt. of India for research fellowships. Y.K. thanks the Wellcome Trust DBT India Alliance, the Brain Research Foundation, the National Center for Biological Sciences and the University of Chicago for start-up funds.

FUNDING

Funding for open access charge: University of Chicago for start-up funds.
Conflict of interest statement. None declared.

REFERENCES

1. Flynt,A.S. and Lai,E.C. (2008) Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat. Rev. Genet.*, **9**, 831–842.
2. Hannon,G.J. (2002) RNA interference. *Nature*, **418**, 244–251.
3. Chakraborty,S., Mehtab,S., Patwardhan,A. and Krishnan,Y. (2012) Pri-miR-17-92a transcript folds into a tertiary structure and autoregulates its processing. *RNA*, **18**, 1014–1028.
4. Mendell,J.T. (2008) miRiad roles for the miR-17-92 cluster in development and disease. *Cell*, **133**, 217–222.
5. Chaulk,S.G., Thede,G.L., Kent,O.a., Xu,Z., Gesner,E., Veldhoen,R.a., Khanna,S.K., Goping,I.S., MacMillan,A.M., Mendell,J.T. *et al.* (2011) Role of pri-miRNA tertiary structure in miR-17∼92 miRNA biogenesis. *RNA Biol.*, **8**, 1105–1114.
6. Du,P., Wang,L., Sliz,P. and Gregory,R.I. (2015) A biogenesis step upstream of microprocessor controls miR-17∼92 expression. *Cell*, **162**, 885–899.
7. Olive,V., Li,Q. and He,L. (2013) mir-17-92: A polycistronic oncomir with pleiotropic functions. *Immunol. Rev.*, **253**, 158–166.
8. O'Donnell,K.A, Wentzel,E.A., Zeller,K.I., Dang,C.V. and Mendell,J.T. (2005) c-Myc-regulated microRNAs modulate E2F1 expression. *Nature*, **435**, 839–843.
9. Ventura,A., Young,A.G., Winslow,M.M., Lintault,L., Meissner,A., Erkeland,S.J., Newman,J., Bronson,R.T., Crowley,D., Stone,J.R. *et al.* (2008) Targeted deletion reveals essential and overlapping functions of the miR-17∼92 family of miRNA clusters. *Cell*, **132**, 875–886.
10. de Pontual,L., Yao,E., Callier,P., Faivre,L., Drouin,V., Cariou,S., Van Haeringen,A., Geneviève,D., Goldenberg,A., Oufadem,M. *et al.* (2011) Germline deletion of the miR-17∼92 cluster causes skeletal and growth defects in humans. *Nat. Genet.*, **43**, 1026–1030.
11. Hossain,A., Kuo,M.T. and Saunders,G.F. (2006) Mir-17-5p regulates breast cancer cell proliferation by inhibiting translation of AIB1 mRNA. *Mol. Cell. Biol.*, **26**, 8191–8201.
12. Bonauer,A., Carmona,G., Iwasaki,M., Mione,M., Koyanagi,M., Fischer,A., Burchfield,J., Fox,H., Doebele,C., Ohtani,K. *et al.* (2009) MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice. *Science*, **324**, 1710–1713.
13. Li,Y., Vecchiarelli-Federico,L.M., Li,Y.J., Egan,S.E., Spaner,D., Hough,M.R. and Ben-David,Y. (2012) The miR-17-92 cluster expands multipotent hematopoietic progenitors whereas imbalanced expression of its individual oncogenic miRNAs promotes leukemia in mice. *Blood*, **119**, 4486–4498.
14. Mortimer,S.A. and Weeks,K.M. (2008) Time-resolved RNA SHAPE chemistry. *J. Am. Chem. Soc.*, **130**, 16178–16180.
15. Tijerina,P., Mohr,S. and Russell,R. (2007) DMS footprinting of structured RNAs and RNA-protein complexes. *Nat. Protoc.*, **2**, 2608–2623.
16. Shcherbakova,I. and Mitra,S. (2009) Hydroxyl-radical footprinting to probe equilibrium changes in RNA tertiary structure. *Methods Enzymol.*, **468**, 31–46.
17. Burke,J.M., Kelenis,D.P., Kincaid,R.P. and Sullivan,C.S. (2014) A central role for the primary microRNA stem in guiding the position and efficiency of Drosha processing of a viral pri-miRNA. *RNA*, **20**, 1068–1077.
18. Kwon,S.C., Nguyen,T.A., Choi,Y.G., Jo,M.H., Hohng,S., Kim,V.N. and Woo,J.S. (2015) Structure of human DROSHA. *Cell*, **164**, 81–90.
19. Nguyen,T.A., Jo,M.H., Choi,Y.-G., Park,J., Kwon,S.C., Hohng,S., Kim,V.N. and Woo,J.-S. (2015) Functional anatomy of the human microprocessor. *Cell*, **161**, 1374–1387.
20. Lagos-quintana,M., Rauhut,R., Lendeckel,W. and Tuschl,T. (2001) Identification of novel genes coding for RNAs of Small expressed RNAs. *Science*, **294**, 853–858.
21. Vasa,S.M., Guex,N., Wilkinson,K.A., Weeks,K.M. and Giddings,M.C. (2008) ShapeFinder: a software system for high-throughput quantitative analysis of nucleic acid reactivity information resolved by capillary electrophoresis. *RNA*, **14**, 1979–1990.
22. Kulshina,N., Baird,N.J. and Ferré-D'amaré,A.R. (2009) Recognition of the bacterial second messenger cyclic diguanylate by its cognate riboswitch. *Nat. Struct. Mol. Biol.*, **16**, 1212–1217.
23. Konarev,P.V., Petoukhov,M.V., Volkov,V.V. and Svergun,D.I. (2006) ATSAS 2.1, a program package for small-anglescattering data analysis. *J. Appl. Crystallogr.*, **39**, 277–286.
24. Svergun,D.I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.*, **25**, 495–503.
25. Lorenz,R., Bernhart,S.S.H., Höner zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F.P. and Hofacker,I.I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
26. Hofacker,I.L., Fontana,W., Stadler,P.F., Bonhoeffer,L.S., Tacker,M. and Schuster,P. (1989) Fast folding and comparison of RNA secondary structures. *Monatshefte f. Chemie*, **125**, 167–188.
27. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
28. Merino,E.J., Wilkinson,K.A., Coughlan,J.L. and Weeks,K.M. (2005) RNA structure analysis at single nucleotide resolution by Selective 2′-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.*, **127**, 4223–4231.
29. Weeks,K.M. (2010) Advances in RNA structure analysis by chemical probing. *Curr. Opin. Struct. Biol.*, **20**, 295–304.
30. Watts,J.M., Dang,K.K., Gorelick,R.J., Leonard,C.W., Bess,J.W., Swanstrom,R., Burch,C.L. and Weeks,K.M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, **460**, 711–716.
31. Somarowthu,S., Legiewicz,M., Chillón,I., Marcia,M., Liu,F. and Pyle,A.M. (2015) HOTAIR forms an intricate and modular secondary structure. *Mol. Cell*, **58**, 353–361.
32. Novikova,I. V., Hennelly,S.P. and Sanbonmatsu,K.Y. (2012) Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic Acids Res.*, **40**, 5034–5051.
33. Mathews,D.H. (2006) RNA secondary structure analysis using RNAstructure. *Curr. Protoc. Bioinformatics*, doi:10.1002/0471250953.bi1206s13.
34. Han,J., Lee,Y., Yeom,K.H., Nam,J.W., Heo,I., Rhee,J.K., Sohn,S.Y., Cho,Y., Zhang,B.T. and Kim,V.N. (2006) Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, **125**, 887–901.
35. Adilakshmi,T., Bellur,D.L. and Woodson,S.A. (2008) Concurrent nucleation of 16S folding and induced fit in 30S ribosome assembly. *Nature*, **455**, 18–22.
36. Hampel,K.J. and Burke,J.M. (2003) Solvent protection of the hammerhead ribozyme in the ground state: Evidence for a cation-assisted conformational change leading to catalysis. *Biochemistry*, **42**, 4421–4429.
37. Gutell,R.R., Cannone,J.J., Shang,Z., Du,Y. and Serra,M.J. (2000) A Story: unpaired adenosine bases in ribosomal RNAs. *J. Mol. Biol.*, **304**, 335–354.
38. Michlewski,G., Guil,S., Semple,C.A. and Cáceres,J.F. (2008) Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol. Cell*, **32**, 383–393.
39. Ritchie,W., Legendre,M. and Gautheret,D. (2007) RNA stem-loops: to be or not to be cleaved by RNAse III. *RNA*, **13**, 457–462.
40. Hermann,T. and Patel,D.J. (2000) RNA bulges as architectural and recognition motifs. *Structure*, **8**, 47–54.
41. Meyers,B.C., Simon,S.A. and Zhai,J. (2010) MicroRNA processing: battle of the bulge. *Curr. Biol.*, **20**, R68–R70.
42. Zacharias,M. and Hagerman,P.J. (1995) Bulge-induced bends in RNA: quantification by transient electric birefringence. *J. Mol. Biol.*, **247**, 486–500.
43. Tullius,T.D. and Greenbaum,J.A. (2005) Mapping nucleic acid structure by hydroxyl radical cleavage. *Curr. Opin. Chem. Biol*, **9**, 127–134.
44. Mcginnis,J.L., Duncan,C.D.S. and Weeks,K.M. (2009) High-throughput SHAPE and hydroxyl radical analysis of RNA structure and ribonucleoprotein assembly. *Methods Enzymol.*, **468**, 67–89.
45. Duncan,C.D.S. and Weeks,K.M. (2010) The Mrs1 splicing factor binds the bI3 Group I intron at each of two tetraloop-receptor motifs. *PLoS One*, **5**, e8983.
46. Trabucchi,M., Briata,P., Garcia-Mayoral,M., Haase,A.D., Filipowicz,W., Ramos,A., Gherzi,R. and Rosenfeld,M.G. (2009) The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature*, **459**, 1010–1014.
47. Gregory,R.I., Yan,K.P., Amuthan,G., Chendrimada,T., Doratotaj,B., Cooch,N. and Shiekhattar,R. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature*, **432**, 235–240.

48. Batey,R.T., Rambo,R.P. and Doudna,J.A. (1999) Tertiary motifs in RNA structure and folding. *Angew. Chem. - Int. Ed.*, **38**, 2326–2343.
49. Hendrix,D.K., Brenner,S.E. and Holbrook,S.R. (2005) RNA structural motifs: building blocks of a modular biomolecule. *Q. Rev. Biophys.*, **38**, 221–243.
50. Chaulk,S.G., Xu,Z., Glover,M.J.N. and Fahlman,R.P. (2014) MicroRNA miR-92a-1 biogenesis and mRNA targeting is modulated by a tertiary contact within the miR-17∼92 microRNA cluster. *Nucleic Acids Res.*, **42**, 5234–5244.
51. Guil,S. and Caceres,J.F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.*, **14**, 591–596.
52. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
53. Pawlicki,J.M. and Steitz,J.A. (2008) Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J. Cell Biol.*, **182**, 61–76.
54. Diaz,J.P., Chirayil,R., Chirayil,S., Tom,M., Head,K.J. and Luebke,K.J. (2014) Association of a peptoid ligand with the apical loop of pri-miR-21 inhibits cleavage by Drosha. *RNA*, **20**, 528–539.
55. Beisel,C.L., Chen,Y.Y., Culler,S.J., Hoff,K.G. and Smolke,C.D. (2011) Design of small molecule-responsive microRNAs based on structural requirements for Drosha processing. *Nucleic Acids Res.*, **39**, 2981–2994.
56. Lünse,C.E., Michlewski,G., Hopp,C.S., Rentmeister,A., Cáceres,J.F., Famulok,M. and Mayer,G. (2010) An aptamer targeting the apical-loop domain modulates pri-miRNA processing. *Angew. Chem. - Int. Ed.*, **49**, 4674–4677.
57. Guan,L. and Disney,M.D. (2012) Recent advances in developing small molecules targeting RNA. *ACS Chem. Biol.*, **7**, 73–86.