

RNA stem-loop to G-quadruplex equilibrium controls mature miRNA production inside the cell

Satyaprakash Pandey, Prachi Agarwala, Gopal Gunanathan Jayaraj, Raimundo Gargallo, and Souvik Maiti

Biochemistry, **Just Accepted Manuscript** • DOI: 10.1021/acs.biochem.5b00574 • Publication Date (Web): 10 Nov 2015

Downloaded from <http://pubs.acs.org> on November 14, 2015

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.



RNA stem-loop to G-quadruplex equilibrium controls mature miRNA production inside the cell

Satyaprakash Pandey^{1, 2}, Prachi Agarwala^{1, 2}, Gopal G Jayaraj^{1, 2}, Raimundo Gargallo⁴

Souvik Maiti^{1, 2, 3, *}

¹ Chemical and Systems Biology Unit, CSIR-Institute of Genomics and Integrative Biology, Mathura Road, New Delhi-110020, India

² CSIR-National Chemical Laboratory, Dr. Homi Bhabha Road, Pune-400008. India

³ Academy of Scientific and Innovative Research (AcSIR), Anusandhan Bhawan, 2 Rafi Marg, New Delhi-110001, India

⁴ Solution Equilibria and Chemometrics Group, Department of Analytical Chemistry, University of Barcelona, Diagonal 645, E-08028 Barcelona, Spain

*To whom correspondence should be addressed-

Tel: +91-11- 2766-615; Fax: (+) 91-11- 2766-7471; Email: souvik@igib.res.in

Keywords- RNA G-quadruplex, pre-miRNA, stem-loop, porphyrin

ABSTRACT

Biological role for existence of overlapping structures in RNA is possible yet remains very less explored. G-rich tracts of RNA form G-quadruplexes while GC-rich sequences prefer stem-loop structures. Equilibrium between alternate structures within RNA may occur and influence its functionality. We tested equilibrium between G-quadruplex and stem-loop structure in RNA and its effect on biological processes using pre-miRNA as a model system. Dicer enzyme recognizes canonical stem-loop structures in pre-miRNA to produce mature miRNAs. Deviation from stem-loop leads to deregulated mature miRNA levels, providing readout of existence of alternate structure per se G-quadruplex mediated structural interference in miRNA maturation. *In vitro* analysis using beacon and Dicer cleavage assays indicated that mature miRNA levels depend on relative amounts of K^+ and Mg^{2+} ions suggesting an ion-dependent structural shift. Further *in cellulo* studies with and without TmPyP₄ (RNA G-quadruplex destabilizer) demonstrated that miRNA biogenesis is modulated by G-quadruplex-stem-loop equilibrium in a subset of pre-miRNAs. Our combined analysis thus provides evidence for formation of non-canonical G-quadruplexes in competition with canonical stem-loop structure inside the cell and its effect on miRNA maturation in a comprehensive manner.

INTRODUCTION

G-quadruplexes are formed in G-rich stretches of nucleic acids. G-quadruplex is one of the most widely characterized, biologically relevant non-canonical structures in RNA. Four stretches of two or more guanines fold upon each other to form a planar arrangement termed as G-quartet. Two or more G-quartets stack upon each other to form three dimensional G-quadruplex structures. The metal ions stabilize the inwardly oriented oxygen atoms to stabilize the structure

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

in order of $K^+ > Na^+ > Li^+$.¹ G-quadruplexes are evolutionarily conserved at telomeres and act as important regulatory elements in promoters and untranslated regions (UTRs).^{2, 3} Formation and stability of these RNA structures is dependent on the presence of different ions under *in vitro* conditions.⁴⁻⁸ Functionally, RNA G-quadruplexes act as modulators of gene expression and have been shown to regulate a number of processes such as mRNA translation,⁹⁻¹² telomere biology,¹³ mRNA localisation,¹⁴ pre-mRNA splicing¹⁵ to name a few. Absence of complementary strand for RNA makes it more amenable to form secondary structures in comparison to duplex DNA. After years of debate, recent evidences by Balasubramanian and group have shown t the formation of G-quadruplexes inside the cell further highlighting its role as a biologically relevant structure.^{16,17}

Simultaneous existence of different overlapping secondary structures in RNA is a possibility and equilibrium between them upon specific stimulus might play an important regulatory role inside the cell. Till date, very few studies have investigated the equilibrium between G-quadruplex and hairpin duplex in a population of RNA and if studied have been limited to *in vitro* conditions.^{6,18} Recently, Balasubramanian and group tested the competition between a G-quadruplex and hairpin structure under cellular mimicking but *in vitro* conditions. NMR studies concluded G-quadruplex to be more stable conformer than hairpin in RNA under physiologically relevant conditions.⁶ However the occurrence of transition between alternate RNA structures and equilibrium between two structures in same RNA molecule inside the cell still remains elusive.

Most of the studies with RNA G-quadruplexes are limited to understanding its regulatory role in coding mRNAs. With emerging functions of non-coding RNAs in biology, studying the role of G-quadruplexes in non-coding repertoire of the cell is a requisite.^{19,20} MiRNAs are the most extensively studied class of non-coding RNAs that regulate gene expression and maintain

homeostasis of the cell.²¹ Deregulation of miRNAs has been linked to various physiological abnormalities and pathological conditions highlighting their importance as key players in cellular playground.²²⁻²⁵ Small functional miRNAs are produced by a highly co-ordinated series of enzymatic cleavages from its primary form (pri-miRNA) to its mature form (miRNA) via an intermediate premature form (pre-miRNA). Pre-miRNAs form a hairpin structure with a duplex stem region followed by a terminal single stranded loop. Stem-loop region of pre-miRNA is recognized by Dicer, a cytoplasmic enzyme which cleaves it to generate mature miRNAs inside the cell.^{26,27} Mature miRNAs are ~20-23 nucleotides long transcripts which bind to their cognate mRNA targets and inhibit/regulate the gene expression. Interference with/or deviation from the canonical stem-loop structure leads to reduced cleavage by Dicer effectively decreasing mature miRNA production. Small molecules effectively decrease miRNA levels by blocking Dicer cleavage and hence are considered to be attractive potential therapeutics for various pathological conditions characterized by deregulated miRNA levels.²⁸⁻³⁰

Herein, we examined the equilibrium between Hoogsteen bonded G-quadruplexes and Watson-Crick bonded stem-loop in RNA in a more biologically relevant environment (inside the cell), using pre-miRNA as a model system. Our combined analysis of *in vitro* and *in cellulo* studies report the occurrence of G-quadruplexes in pre-miRNAs and its influence in governing the biogenesis of miRNAs in the cell by a “structural interference” mechanism. Similar observations have been reported for miR-92b when our manuscript was in preparation.³¹ However, our study demonstrates both G2 (two stacked G-quadruplexes formed by four runs of two guanines separated by intervening bases) and G3 quadruplex (three stacked G-quadruplexes formed by four runs of three guanines separated by intervening bases) as metastable switches that form inside the cell and regulate pre-miRNA maturation. We further present a more global

role for G-quadruplex mediated structural interference in premature miRNAs upon destabilization of RNA G-quadruplexes using well characterized porphyrin molecule TmPyP₄.

EXPERIMENTAL METHODS

***In vitro* transcription of pre-miRNAs.** *In vitro* transcription of pre-miRNAs was carried out using Megascript T7 IVT kit (Ambion) as per manufacturer’s instructions. Briefly, forward and reverse primers designed for each pre-miRNA were heated and annealed to form a partially complementary double stranded-DNA which was elongated to produce a fully complementary double stranded-DNA template with T7 promoter. *In vitro* transcription was carried out overnight at 37 °C in a thermally controlled PCR machine. Template DNA was digested using TURBO DNase and RNAs were purified using NucAway columns (Ambion). Pre-miRNAs were further checked on polyacrylamide gel electrophoresis for their integrity and purity using ethidium bromide staining. The primer sequences used for pre-miRNAs were purchased from Sigma Aldrich and are as follows-

Pre-miRNA	Forward Primer	Reverse Primer
Quad-pre-miR-27a	TAATACGACTCACTATAGGG GATGGGCAGGGCUUAGCUGCU UGUGAGCAGGGUCCACACCAA GUCGUGUUC	GGGTAGGGCGGAACTTAGCC ACTGTGAACACGACTTGGTGTG GACCCTGCTCACAAGCAGCTAA GCCCTGC
Stem-loop pre-miR-27a	TAATACGACTCACTATAGCU GAGGAGCAGGGCUUAGCUGCU UGUGAGCAGGGUCCACACCAA	CTGGGGGGCGGAACTTAGCC ACTGTGAACACGACTTGGTGTG GACCCTGCTCACAAGCAGCTAA

	GUCGUGUUC	GCCCTGC
Pre-let-7e	TAATACGACTCACTATAGCC CGGGCTGAGGTAGGAGGTTGT ATAGTTGAGGAGGACACCCAA	CCTGGGGAAAGCTAGGAGGC CGTATAGTGATCTCCTTGGGTG TCCTCCTCAACTATAACAACC
Pre-let-7e MQ (Mutant quadruplex)	TAATACGACTCACTATAGCC CGAGCTGAGGTAGGAGGTTGT ATAGTTGAAGAGGACACCCAA	CCTGGGGAAAGCTAGGAGGC CGTATAGTGATCTCCTTGGGTG TCCTCCTCAACTATAACAACC

Molecular beacon based Screening. The design and sequence of the molecular beacon is mentioned in Supplementary Scheme S2. Stem-loop pre-miR-27a and Quad pre-miR-27a was heated to 90 °C followed by slow cooling to 37 °C under different ionic conditions namely Magnesium [Mg^{2+}] (1 mM $MgCl_2$), Potassium [K^+] (150 mM), Magnesium and different concentrations of potassium [$Mg^{2+} + K^+$] (1 mM $MgCl_2$ + 5, 25, 50, 100 and 150 mM KCl) in 10 mM Tris Buffer (pH-7.4). Further pre-mir-27a was incubated with 0.5 U Dicer (Gel Atlantis) and molecular beacon at a concentration of 25 nM for four hours. The fluorescence spectra were collected in a Fluorolog spectrofluorometer equipped with a thermally controlled cuvette holder at 37 °C. Fluorophore FAM was excited at 492 nm and its emission spectra were recorded from 510 nm to 535 nm at a slit width of 5 nm. The emission spectrum was further plotted using Origin 7.0. Dicer cleavage efficiency was calculated and is represented in Figure 2.

For time course kinetics experiment, heat cooled pre-miRNAs were incubated with molecular beacon for 10 minutes. Dicer was added and fluorescence was measured at regular intervals till a saturation was reached using a Tecan plate reader. The data was analysed using Origin 7.0.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Dicer Cleavage Assay. 5' end P³² labelled pre-miRNA was heated to 90 °C for 2 minutes followed by slow cooling to 37 °C under different ionic conditions of magnesium (1 mM MgCl₂), potassium (150 mM KCl), magnesium and potassium (1 mM MgCl₂ and 150 mM KCl) in 10 mM Tris buffer (pH-7.4). Dicer enzyme was added at a concentration of 0.5 U and reaction was incubated at 37 °C for 90 minutes. Equal volume of gel loading buffer II (Ambion) containing 95 % formamide and 18 mM EDTA was used to stop the reaction. The products were separated onto a 15 % denaturing PAGE. The gel was exposed overnight and analysed using Typhoon FLA 7000 and ImageQuant 5.2 software.

Enzymatic Footprinting. 5' end P³² labelled pre-miRNA (75000 cpm) was denatured by heating to 90 °C followed by slow cooling to 37 °C under different ionic conditions of magnesium (1 mM MgCl₂), potassium (150 mM KCl), magnesium and potassium (1 mM and 150 mM potassium) in 10 mM Tris buffer (pH-7.4). RNase T1 (0.05 U) was added and reaction was incubated for 10 minutes at 37 °C. Equal volume of Gel loading Buffer II (Ambion) was used to stop the reaction followed by snap chill. Digested products were further separated onto a 15% denaturing PAGE in a sequencing gel apparatus (BioRad). The gel was exposed overnight and the image was analysed in a Typhoon FLA 7000 and ImageQuant 5.2 software.

Circular Dichroism (CD) spectroscopy. CD spectra of all the sequences were recorded using a Jasco 815 spectropolarimeter. The samples were prepared by slow annealing (after heating to 90 °C for 5 min, followed by programmed cooling at 0.2 °C/min) in 10 mM Tris buffer (pH-7.4) containing magnesium and potassium (1 mM MgCl₂ and 150 mM KCl) at strand concentrations of 2 µM each. CD scans were collected with increasing temperature and subjected to MCR-ALS analysis. The spectra obtained were the average of three consecutive scans for each sample with path length of 10 mm and bandwidth of 1 nm. CD spectra for Quad-pre-miR-27a and pre-let-7e

were collected from 320-200 nm both in absence and presence of 10 μ M of TmPyP₄. Further analysis was performed using Origin 7.0.

Multivariate data analysis. Data recorded along melting experiments were analyzed by means of Multivariate Curve Resolution based on Alternating Least Squares (MCR-ALS), a soft-modelling-based multivariate method.³² This method has been used extensively to analyze data recorded along melting experiments of nucleic acids,^{33,34} genomic data from DNA microarrays³⁵ or image analysis of biological material.³⁶

Briefly, the goal of the method is the resolution of the complex spectroscopic data into the contributions of the n individual species or conformations present in the experiment. The method provides the distribution profile, which reflects the variation of the concentration of each individual species or concentration with increasing temperature, as well as the corresponding pure spectra, i.e., the spectrum corresponding to each one of the individual species or concentrations. A more detailed description of the method can be found elsewhere and in the supplementary section.^{32,37}

Quantitative real-time PCR (qRT-PCR). MCF-7 cells were cultured in DMEM high glucose media supplemented with 10 % FBS at 37 °C in a 5% CO₂ incubator. MCF-7 cells were seeded at a density of 3 x 10⁶ cells per well in a six well plate. Pre-miRNAs were transfected at concentration of 100 nM and 50 nM for pre-miR-27a and pre-let-7e respectively. Cells were harvested after 48 hours of transfection and RNA was isolated using Trizol method. cDNA was prepared from 1 μ g of RNA using MMuLV reverse transcriptase and stem-loop specific primers against pre-miRNAs of interest. Quantitative real time PCR was performed using SyBr Green in a Roche LightCycler 480. The levels of mature miRNA levels were quantified using specific forward primers and universal stem-loop primers against miR-27a and let-7e (sequences given

below in the table). MiRNA levels were normalised to U6 gene as a control. Fold change analysis was performed using standard delta delta Ct method.³⁸

For TmPyP₄ treatment, 10 µM of TmPyP₄ (Sigma) was added to the cells after 24 hours of pre-miRNA transfection. Cells were grown for 24 hours after TmPyP₄ and RNA was isolated using Trizol method. Subsequently, qRT-PCR was performed with the transfected and treated MCF7 cells.

Pre-miRNA	Stemloop primer	Forward Primer
miR-27a	CTCAACTGGTGTCGTGGAGTCGGCA ATTCAGTTGAGGCGGAACT	ACACTCCAGCTGGGTTCA CAGTGGCTAAG
Let-7e	CTCAACTGGTGTCGTGGAGTCGGCA ATTCAGTTGAGTTCCTTCT	ACACTCCAGCTGGGTGAG GTAGGAGGTTG
U6 snRNA	CTCAACTGGTGTCGTGGAGTCGGCA ATTCAGTTGAGGAATTTGC	ACACTCCAGCTGGGCGCA AGGATGACACG

The universal primer sequence for all stem-loop primers is- GTGTCGTGGAGTCGGCAATTC.

Global profiling for quadruplex harbouring pre-miRNAs was performed upon 10 µM TmPyP₄ (a porphyrin compound that destabilises RNA G-quadruplex structure) a treatment of MCF7 cells for 48 hours. RNA was isolated using Trizol method and cDNA synthesis was performed using Human miRnome profiler kit (SBI System Biosciences) and as per manufacturer's

instructions. Primers from human miRnome profiler plates were suspended in a new 384 plate and further used to quantify expression changes in miRNAs harbouring PQs in their premature forms. Ten miRNAs (including U6 snRNA) with no PQS in their premature forms were used as input controls to normalise the data.

Luciferase Assay. MCF-7 cells were cultured in DMEM high glucose media supplemented with 10 % FBS at 37 °C in a 5% CO₂ incubator. Cells were seeded (50000 cells/well) in a 24 well plate and transfection was performed at 70 % confluency. 200 ng of target plasmid was transfected per well using Cellfectin 2 (Invitrogen) along with pre-miRNAs. For pre-miR-27a, psicheck-2 vector containing 3' UTR of Prohibitin gene under Renilla luciferase was used as target for miR-27a and Firefly luciferase acted as normalising control. For pre-let-7e, vector containing 3' UTR of IL-13 gene cloned under Firefly luciferase was used as a let-7e target and Renilla plasmid served as a normalising control. Pre-miRNAs were transfected along with target luciferase plasmids. The concentrations employed for Stem-loop pre-miR-27a, Quad-pre-miR-27a, pre-let-7e and mutant pre-let-7e were 100 nM, 100 nM, 50 nM and 50 nM respectively. Cells were harvested after 48 hours of transfection and luciferase levels were quantified using a dual luciferase assay (Promega).

For TmPyP₄ treatment, 10 µM of TmPyP₄ was added to the cells after 24 hours of pre-miRNA transfection. Cells were grown for 24 hours after TmPyP₄ and lysate was prepared using manufacturers instructions. Subsequently, luciferase assays were performed with the transfected and treated MCF7 cells.

RESULTS

Experimental design and *in vitro* evidence of equilibrium shift between alternate structures in a synthetic pre-miR-27a. Pre-miRNAs are cleaved by Dicer enzyme to produce

1
2
3 mature miRNAs. Cleavage reaction is structure dependent wherein stemloop structure of pre-
4
5 miRNAs is recognised by Dicer for efficient miRNA production. To understand the effect of
6
7 equilibrium and/or a transition between different structures namely G-quadruplex and stem-loop
8
9 in RNA, we chose well characterized premature miR-27a as model RNA molecule. Pre-miR-27a
10
11 does not contain a putative quadruplex sequence (PQS) in its precursor form. Systematic point
12
13 mutations in pre-miR-27a (Stem-loop pre-miR-27a) were designed to introduce a PQS in its
14
15 sequence such that new pre-miR-27a (referred as Quad-pre-miR-27a) can potentially form a G-
16
17 quadruplex and alter the canonical stem-loop structure (Scheme 1). Care was taken that both pre-
18
19 miRNA sequences-Quad- pre-miR-27a and Stem-loop pre-miR-27a potentially gave rise to same
20
21 mature miRNA sequence and their GC content remains similar.
22
23
24
25
26

27 *In vitro* quantification of mature miR-27a was performed using a molecular beacon
28
29 approach.³⁹ A beacon containing a fluorophore (FAM) at one end and a quencher at the other
30
31 was designed with its terminal loop complementary to mature miR-27a sequence. Binding of
32
33 beacon to mature miR-27a separates the fluorophore from quencher resulting in an enhanced
34
35 fluorescence signal (Supplementary figure Scheme S1). Different ions stabilize different
36
37 structures in RNA. Magnesium ions aid in hairpin formation while potassium ions stabilize G-
38
39 quadruplexes.⁶ Levels of mature miRNA produced in different ionic conditions were used to
40
41 examine the role of G-quadruplex mediated structural interference on miRNA maturation. A
42
43 time dependent kinetics experiment was performed to check for any differences in Dicer
44
45 cleavage efficiency of Stem-loop-pre-miR-27a and Quad-pre-miR-27a. Under conditions of only
46
47 magnesium (Mg^{2+}), kinetics was found to be similar for both the pre-miRNAs but under
48
49 conditions of magnesium and potassium (Mg^{2+} and K^{+} in conditions of magnesium (Mg^{2+}),
50
51 kinetics was found to be different (Figure 1A). Quad-pre-miR-27a showed a lag phase in
52
53
54
55
56
57
58
59
60

fluorescence levels as opposed to Stem-loop-pre-miR-27a suggesting that an alternate structure may hinder the Dicer processing of Quad-pre-miR-27a. A saturation in mature miRNA level was achieved after 3 hours of incubation with Dicer for both pre-miRNAs under different conditions (Figure 1). To normalize for the effect of cations, Dicer cleavage of Stem-loop pre-miR-27a and Quad-pre-miR-27a was carried out in presence of both magnesium and potassium ions after four hours (under saturating conditions). Stem-loop-pre-miR-27a and Quad-pre-miR-27a were incubated with Dicer under different salt conditions (1 mM Mg^{2+} + varying concentration of K^+) and subjected to Dicer cleavage. In each condition, cleavage activity of Dicer for Stem-loop pre-miR-27a was normalized to one and compared with Quad-pre-miR-27a. Fluorescence analysis revealed that increasing concentrations of K^+ led to lesser cleavage of Quad-pre-miR-27a than Stem-loop pre-miR-27a (Figure 1B). It is known that potassium ions drastically reduce the Dicer activity by locking the Dicer-dsRNA complex in a catalytically inactive form thus reducing the production of mature miRNAs.⁴⁰ Significantly less levels of mature miR-27a in presence of only potassium was attributed to the inhibitory role of potassium ions on Dicer activity (Supplementary figure S1A). Normalisation of data to unity for Stem-loop pre-miR-27a did not show any significant change between Quad-pre-miR-27a and Stem-loop pre-miR-27a in potassium only condition (150 mM KCl) (Supplementary figure S1B). However at physiologically relevant concentrations of 1mM Mg^{2+} and 150 mM KCl, less cleavage occurred in Quad-pre-miR-27a than Stem-loop-pre-miR-27a (Supplementary figure S1B). Production of varying amounts of mature miRNA could be due to reduction in Dicer activity alone or due to formation of G-quadruplexes that in turn interfere with cleavage of pre-miRNA. Since the reduction in miRNA production was significantly more for Quad-pre-miR-27a as compared to

Stem-loop pre-miR-27a, it pointed towards a role of G-quadruplex mediated interference on Dicer processing.

To validate the results obtained from molecular beacon approach, Dicer cleavage assays were performed using radiolabelled Stem-loop pre-miR-27a and Quad-pre-miR-27a. Labelled pre-miRNAs were folded and subjected to Dicer cleavage under different ionic conditions (Mg [1 mM Mg²⁺], K [150 mM K⁺] and Mg + K [1 mM Mg²⁺ and 150 mM K⁺]. Cleaved products were analyzed using denaturing gel electrophoresis and corresponding band intensities were calculated. Autoradiograph obtained further corroborated the beacon based results showing lower levels of precleavage product in Quad-pre-miR-27a than pre-miR-27a under physiologically relevant salt conditions of 1 mM Mg²⁺ and 150 mM K⁺ (Supplementary Figure S2). Since the production of cleavage product is different for Stem-loop pre-miR-27a and Quad-pre-miR-27a; reduced Dicer activity due to cation effect alone seemed insufficient to explain this behaviour. Varying amounts of cleaved product could be attributed to an equilibrium shift in Quad-pre-miR-27a between non-canonical G-quadruplex and canonical stem-loop that interferes with Dicer cleavage leading to lesser cleaved product for Quad-pre-miR-27a as compared to Stem-loop pre-miR-27a.

RNase T1 footprinting reveals formation of G-quadruplex in Quad-pre-miR-27a. To reveal the identity of alternate structure formed in Quad-pre-miR-27a, RNase T1 footprinting was performed. RNase T1 cleaves after guanine residues present in RNA. Guanine residues involved in formation of structure are constrained and less amenable to T1 cleavage thus leading to their protection and less intense bands on the gel.¹² Quad-pre-miR-27a was heated and allowed to fold in different conditions (Mg²⁺, K⁺, Mg²⁺ + K⁺ and Li⁺) before enzymatic probing. Lithium ions do not promote or stabilize G-quadruplexes and hence served as control

1
2
3 monovalent ions. RNase T1 digestion revealed that guanine residues predicted to be involved in
4
5 G-quadruplex formation in Quad-pre-miR-27a were protected and displayed less intense bands
6
7 on gel in presence of ($\text{Mg}^{2+} + \text{K}^+$) and K^+ ions. Moreover, protected guanine bands showed a
8
9 decreased intensity in order of $\text{K}^+ > \text{Mg}^{2+} + \text{K}^+$ ions. No protection was observed in presence of
10
11 lithium ions suggesting the existence of a potassium dependent structure per se G-quadruplex
12
13 (Figure 2).
14
15
16

17
18 **MCR-ALS analysis resolves and estimates distribution of G-quadruplex and stem-loop in**
19
20 **pre-miRNAs.** UV and CD spectroscopy have been used to assess the formation and stability of
21
22 structures in nucleic acids. However, presence of more than two structural species that are ionic
23
24 dependent may result in a mixture of absorption spectra from both species making it difficult to
25
26 delineate the stability and formation of these structures. CD signature for duplex RNA and G-
27
28 quadruplex RNA display positive maxima at 263 nm and negative minima at 240 nm.⁴¹ Thus
29
30 determining the contribution of G-quadruplex and stem-loop to the CD spectrum for G-
31
32 quadruplex containing pre-miRNA is relatively difficult. The structural equilibria and associated
33
34 conformational changes between different nucleic acid structures can be resolved using
35
36 multivariate curve resolution analysis method based on alternating least squares (MCR-ALS).⁴²
37
38 MCR-ALS assesses and resolves different possible conformations and estimates their
39
40 distribution profile and pure spectra based on the data collected for multiple wavelengths under
41
42 varying temperature and salt concentrations.⁴² CD melting experiments were performed for
43
44 Quad-pre-miR-27a, Stem-loop pre-miR-27a, pre-let-7e and mutant pre-let-7e in different ionic
45
46 conditions (magnesium [Mg^{2+}], magnesium + potassium [$\text{Mg}^{2+} + \text{K}^+$] and potassium [K^+]) to
47
48 understand the conformational equilibrium between G-quadruplex and stem-loop and data was
49
50 subjected to MCR-ALS (refer to supplementary for a detailed description of the analysis).
51
52
53
54
55
56
57
58
59
60

MCR-ALS analysis indicated presence of three components- G-quadruplex [blue], stem loop [green] and single stranded structures [pink] for Quad-pre-miR-27a (Figure 3) Distribution profiles for Quad-pre-miR-27a showed presence of G-quadruplex species at physiological temperature that unfolds to form stem-loop and single stranded species with increasing temperature (Figure 3 panel a). Third component was found to be absent for Stem-loop pre-miR-27a further supporting a G-quadruplex-stemloop equilibrium shift in pre-miRNA (Figure 3 panel c). Pure spectra obtained from MCR-ALS for Quad-pre-miR-27a demonstrated G-quadruplex to be the major conformer in comparison to stem-loop structure (Figure 3 panel b). CD spectra for duplex and G-quadruplex RNA display a maxima at 263 nm and a minima at 240 nm making it difficult to determine the identity and contribution of both structures towards the spectrum.⁴¹ Careful analysis of CD melting data at 263 nm showed a broad melting domain for the G-quadruplex harbouring pre-miRNAs and a sharper melting domain for non-G-quadruplex harbouring pre-miRNAs (Supplementary figure S3A). The broad melting domain could be attributed to presence of more than two structural species (G-quadruplex and stem-loop) with similar CD signature.

Thus, MCR-ALS resolved and determined the presence of G-quadruplex in pre-miRNAs- Quad-pre-miR-27a and pre-let-7e and suggested G-quadruplex to be major conformer between stem-loop and G-quadruplex at physiological temperature.

Quadruplex harbouring pre-miR-27a produces less mature miRNA inside the cell. Although *in vitro* studies point towards existence of G-quadruplex in Quad-pre-miR-27a, cellular conditions differ considerably from favourable *in vitro* conditions. Molecular crowding conditions may affect structural transitions and have been shown to favour G-quadruplexes over canonical Watson-Crick base pairing structures in DNA.⁴³ To determine G-quadruplex formation

and its regulation of mature miRNA levels inside the cell, quantitative real time PCR (qRT-PCR) was performed for miR-27a. Full length pre-miR-27a and Quad-pre-miR-27a were transfected into MCF-7 cells and mature miRNA levels were measured using specific stem-loop primers against mature miR-27a. U6 RNA was used as a normalisation control. qRT-PCR analysis showed less mature miR-27a levels for Quad-pre-miR-27a than Stem-loop pre-miR-27a. A difference of more than three-fold was observed between Quad-pre-miR-27a and Stem-loop pre-miR-27a transfected samples (Figure 4A). The difference in mature miRNA levels points towards a possible interference in Dicer processing owing to formation of G-quadruplex in Quad-pre-miR-27a as opposed to Stem-loop-pre-miR-27a.

To further ascertain the biological significance of G-quadruplex in Quad-pre-miR-27a, luciferase assays were performed for miR-27a target- prohibitin. Plasmid containing miR-27a target was used to determine the change in luciferase levels upon transfection of Quad-pre-miR-27a and Stem-loop pre-miR-27a. Luciferase assays showed more reporter luciferase signals in Quad-pre-miR-27a than Stem-loop pre-miR-27a (Figure 4B). The resultant increase in reporter signal for Quadpre-miR-27a is an effect of deregulated mature miRNA levels owing to equilibrium shift thus favoring G- quadruplex population. Luciferase studies hence provided a proof for biological role of G-quadruplex in pre-miRNA processing which in turn would regulate levels of target mRNA of these pre-miRNAs. Thus presence of putative G-quadruplex sequence along with stem-loop prone regions in RNA may form G-quadruplex inside cellular milieu under specific stimuli and influence miRNA processing machinery. Similar results have been reported recently for miR-92b using luciferase reporter systems.³¹

***In silico* analyses of occurrence of G-quadruplexes in pre-miRNAs.** A comprehensive *in silico* analysis was performed across all pre-miRNA sequences using an in house tool –

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Quadfinder to examine the occurrence of G-quadruplexes in premature microRNAs.⁴⁴ The consensus motif used for searching putative quadruplex sequences was based on the total number of bases in the loop denoted as N_{sum} as opposed to the conventional pattern search for G-quadruplexes. G2 with $N_{sum} \leq 7$ and G3 quadruplexes with $N_{sum} \leq 22$ was used to predict the occurrence of stable and metastable G-quadruplexes in pre-miRNA. The cut off values for the sum of total number of bases in loop were chosen based upon the experimental results obtained Quad-pre-miR-27a ($N_{sum}=22$). Biophysical characterizations of G2 and G3 RNA quadruplexes from our lab have shown G2U3 ($N_{sum}=9$) and G3U15 ($N_{sum}=45$) to be stable structures with T_m values of 30 °C and hence a conservative cut-off of $N_{sum}=7$ with a predicted T_m of 37 °C for G2 quadruplexes.⁷ Bioinformatics' prediction was also performed with cut off of G2 ($N_{sum}=9$) and G3 ($N_{sum}=45$) (Supplementary table 1). We observed occurrence of overlapping putative G-quadruplexes in premature miRNAs of mirbase v20.⁴⁵ Data was further filtered to obtain unique pre-miRNAs which house at least one putative G-quadruplex in its sequence. Bioinformatics analyses predicted presence of 180 G2 ($N_{sum} \leq 7$) and 76 G3 ($N_{sum} \leq 22$) suggesting that 13 % of pre-miRNAs harbour atleast one G-quadruplex motif in their sequence (Figure 5A). Prediction with G2 ($N_{sum} \leq 9$) and G3 ($N_{sum} \leq 45$) further increased the prediction to 23 % of pre-miRNAs housing a G-quadruplex motif. We further analyzed the G-quartets, loop length and composition of positive hits to examine the stability of G-quadruplexes. Majority of PQS in pre-miRNAs were found to contain shorter loops with a major enrichment of mononucleotide loops suggesting a potential existence of stable G-quadruplexes under favourable conditions (Figure 5C and 5D).^{5,7} More than half of all the putative PQS contain at least two mononucleotide loops further strengthening the existence of potential G-quadruplex structures in pre-miRNAs. A more comprehensive analysis of positive hits and their loop distribution profiles is provided in

Supplementary Table 1 sheet 2-4. The composition of loop nucleotides in all the hits were dominated by adenine in order of $A > C > U$ (Supplementary Figure S4). Higher percentage of adenines and less cytosines in predicted PQS further suggest that non-canonical structures such as G-quadruplexes can form in pre-miRNAs and regulate miRNA production. Bioinformatics data was further scrutinized to determine GC content and thermodynamic stability of quadruplex harbouring pre-miRNAs and non-quadruplex pre-miRNAs. It was found that although mean thermodynamic stability (ΔG) remained similar for G-quadruplex pre-miRNAs and non-Quad-pre-miRNAs, their GC contents differed considerably. Mean GC content of non-Quad-pre-miRNAs was found to be 48% while G-quadruplex harbouring pre-miRNAs possess a GC content of 60% (Figure 5B). Since G-quadruplexes are formed in G rich tracts, it is intuitive that pre-miRNAs with higher GC content have evolved structural interference mechanism using G-quadruplex motifs to regulate the production of mature miRNAs upon specific stimulus and requisite of the cell.

G2 quadruplex in pre-let-7e regulates mature let-7e production. G2 quadruplexes with shorter loop size are metastable and possess a thermal stability in range of physiological temperature. Let-7e-a G2 quadruplex with $N_{\text{sum}}=7$ was chosen as a representative to investigate the effect of metastable G-quadruplexes on pre-miRNA processing (Supplementary Scheme S2). Let-7 family of miRNA is well studied and has been shown to be deregulated in different cancers and important biological processes.⁴⁶ Pre-let-7e (quadruplex harbouring) and pre-let-7e MQ (G to A mutant for disruption of G-quadruplex) were *in vitro* transcribed, radiolabelled and subjected to Dicer cleavage assay. Dicer assay showed that under conditions of $\text{Mg}^{2+} + \text{K}^+$, mature let-7e production was less in pre-let-7e as compared to pre-let-7e MQ (Supplementary Figure S5). Dicer analysis pointed towards the presence and involvement of a potassium

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

dependent structure such as G-quadruplex in pre-let-7e that interferes with Dicer processing leading to lesser let-7e production.

G2 quadruplexes are less stable as compared to G3 quadruplexes. CD melting was performed under different ionic conditions and data was analysed using MCR-ALS method. MCR-ALS analysis revealed the distribution profiles and quantified the relative concentration of different structural species in pre-let-7e and pre-let-7e MQ. MCR-ALS of pre-let-7e showed presence of three components- A, B and C (namely G-quadruplex, stem-loop and single stranded RNA) (Supplementary figure S6 panel a). Pure spectra analysis further suggested G-quadruplex to be the major conformer at physiologically relevant temperatures (Supplementary figure S6 panel b). Distribution profiles were similar to Quad-pre-miR-27a suggesting the existence of a G-quadruplex at lower temperature which unfolds leading to an increase in duplex RNA population that in turn melts at higher temperature to form single stranded RNA (Supplementary figure S6 panel a). Pre-let-7e MQ analysis suggested existence of only two components with a pure spectra analysis showed an absence of G-quadruplex signature (Supplementary figure S6 panel c and d). CD melting at 270 nm displayed a broad melting domain for pre-let-7e as opposed to pre-let-7e MQ with a sharp melting domain (Figure S3B). The broad domain further suggested the existence of multiple structural species in pre-let-7e in turn pointing towards G-quadruplex formation in pre-let-7e.

Let-7e being a G2 quadruplex further supports the hypothesis that metastable G-quadruplexes form in presence of potassium ions and interfere with Dicer cleavage thus producing less mature let-7e under in vitro conditions. However whether G2 quadruplex forms and affects miRNA maturation inside the cellular conditions needed to be examined.. To this end, we carried out

transfection experiments in MCF-7 cells using full length pre-let-7e (G2 quadruplex) and pre-let-7e MQ (mutant quadruplex).

Pre-let-7e and pre-let-7e MQ were transfected and mature let-7e levels were quantified using real time PCR primer specific for mature let-7e. qRT-PCR analysis revealed a difference of more than five-fold of mature let-7e levels between pre-let-7e and pre-let-7e MQ transfected samples (Figure 6A). This significant change in mature let-7e levels suggests a structural interference in Dicer processing due to formation of alternate potassium dependent structure per se G-quadruplex in wild type pre-let-7e. We further validated the influence of deregulated let-7e levels on its target mRNAs using a reporter plasmid containing IL-13 gene (target of let-7e) cloned under Firefly luciferase.⁴⁷ Luciferase assays showed more reporter signals in pre-let-7e than pre-let-7e MQ corroborating the effect of G-quadruplex on pre-let-7e and its consequence on luciferase levels (Figure 6B). Thus, both *in vitro* and *in cellulo* results for let-7e revealed that a metastable G2-quadruplex could form in its precursor form and regulate mature let-7e levels inside the cell.

TmPyP4 destabilises RNA G-quadruplexes in pre-miRNAs and modulates production of mature miRNA. Selective ligands have been used to stabilize/destabilize G-quadruplexes due to importance of these structures as potential anti-cancer therapeutic modules in gene networks.⁴⁸ Cationic porphyrin TmPyP₄ (tetra-(N-methyl-4-pyridyl)porphyrin) is a potent stabilizing ligand for DNA G-quadruplexes and has recently been shown to destabilize RNA G-quadruplexes.^{48,49} A decrease in CD signal strength was observed for both Quad-pre-miR-27a and pre-let-7e RNA in presence of 10 μ M TmPyP₄ treatment confirming the destabilization effect of the ligand on RNA G-quadruplex (Supplementary Figure S7). To test the efficacy of TmPyP₄ on G-quadruplexes inside the cells, MCF-7 cells transfected with G-quadruplex and non-quadruplex

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

pre-miRNAs (Quad-pre-miR-27a and Stem-loop pre-miR27a and let-7e and let-7e MQ) were treated with 10 μ M TmPyP₄ and respective mature miRNA levels were quantified by qRT-PCR. Pre-miRNAs containing putative G-quadruplexes showed an increase in mature miRNA levels upon TmPyP₄ treatment suggesting destabilization of G-quadruplexes by porphyrin TmPyP₄ inside the cell. An increase of more than three-fold was observed for Quad-pre-miR-27a and a difference of more than four-fold was observed for let-7e (Figure 7A). A moderate increase was also seen in pre-miRNAs which do not contain putative quadruplex motifs which could be due to non-specific effects of TmPyP₄ treatment.

Further luciferase assays were performed to assess the functional role of increased mature miRNAs upon TmPyP₄ treatment. Destabilization of G-quadruplexes in pre-miRNAs by TmPyP₄ leads to more production of mature miRNA and hence decreased luciferase levels for the corresponding target genes. Decrease in luciferase levels was found to be more pronounced for Quad-pre-miR-27a and pre-let-7e as compared to their non-quadruplex harbouring counterparts- pre-miR-27a and let-7e MQ (Figure 7B). TmPyP₄ mediated deregulation in mature miRNA levels and its consequent effect on luciferase levels further provides evidence of G-quadruplex occurrence and formation in pre-miRNA and role of equilibrium and transition between alternate structures in regulating biological processes.

Comprehensive analysis of mature miRNA levels for G-quadruplex harbouring pre-miRNAs upon TmPyP₄ treatment. We further went on to characterize the effect of G-quadruplex in pre-miRNA processing in a more general manner. A high-throughput qRT-PCR approach was used to assess the miRNA levels of quadruplex harbouring pre-miRNAs upon TmPyP₄ treatment. A total of 100 pre-miRNAs harbouring quadruplexes were chosen along with ten non-quadruplex harbouring pre-miRNAs (as control). Corresponding mature levels were

quantified upon treatment of MCF7 cells with 10 μ M TmPyP₄. Expression analyses revealed that G-quadruplex harbouring pre-miRNAs displayed significantly higher fold changes in their mature miRNA production (Figure 8). In contrast, non-quadruplex harbouring pre-miRNAs displayed no change or significantly less up regulation in their mature miRNA levels (Supplementary Table sheet 5). Thus, TmPyP₄, a RNA G-quadruplex destabilizing ligand shifts the structural shift towards stem-loop structures and increases the production of mature miRNAs in G-quadruplex harbouring pre-miRNAs. Increased mature miRNA levels upon ligand treatment suggested a more global role of G-quadruplex mediated structural interference in pre-miRNA biology and establishes the role of structural shift between overlapping structures in RNA to play a key role inside the cell upon specific stimulus. Thus alternate structures such as G-quadruplex in pre-miRNAs add another layer of regulation to miRNA biogenesis and in turn the overall response of a cell towards its environment.

DISCUSSION

RNA folding *in vivo* differs considerably from *in vitro* conditions. Different structures can coexist and transit within same RNA molecule in response to specific cues as per the cellular requirement.^{6,18,19} Multiple overlapping structures can exist within the same RNA molecule and these structures may undergo equilibrium shifts favoring one over the other in response to varying ionic conditions, molecular crowding, other specific interacting partners and a range of stimuli.¹⁹ MicroRNAs (miRNAs) are small non-coding RNAs that act as silencing molecules inside the cell by binding to their cognate mRNAs and regulate gene expression in various physiological and pathological conditions.²¹⁻²⁵ Primary miRNAs (pri-miRNAs) are transcribed by RNA Pol II and processed by Drosha to form premature miRNAs (pre-miRNAs). Pre-miRNAs are exported to cytoplasm and processed by Dicer to produce functional miRNA

1
2
3 molecules. Various components of miRNA processing pathways have been targeted to regulate
4
5 miRNA levels and bring about homeostasis inside the cell.^{29,30} In this study, we unravel a G-
6
7 quadruplex mediated “structural interference” mechanism present in pre-miRNAs that regulates
8
9 miRNA levels. Canonical stem-loop structure of pre-miRNAs is recognised by Dicer for
10
11 efficient cleavage and production of mature miRNA levels. We observed the occurrence of
12
13 potential G-quadruplexes in pre-miRNA sequences and sought to determine whether equilibrium
14
15 shift between non-canonical G-quadruplexes and canonical stem-loop structures in premiRNAs
16
17 could act as a new regulatory mechanism for governing miRNA levels.
18
19

20
21
22 Synthetic (Quad-pre-miR-27a) and naturally occurring (pre-let-7e) pre-miRNAs containing
23
24 putative G-quadruplexes were examined for their effective role on miRNA levels. Dicer cleavage
25
26 assay using a molecular beacon specific for mature miR-27a demonstrated less production of
27
28 mature miRNA levels in Quad-pre-miR-27a than canonical Stem-loop pre-miR-27a. Presence of
29
30 monovalent and divalent cations such as potassium (K^+) and magnesium (Mg^{2+}) affect various
31
32 structures in RNA such as duplex and G-quadruplexes differently. In this study, we determined
33
34 the influence of physiologically relevant concentrations of potassium and magnesium on G-
35
36 quadruplex-Stem-loop equilibrium in Quad-pre-miR-27a. *In vitro* Dicer cleavage assays
37
38 demonstrated the existence and inhibitory role of a potassium dependent G-quadruplex structure
39
40 in Quad-pre-miR-27a thus lowering levels of mature miR-27a. The total loop length for
41
42 Synthetic Quad-pre-miR-27a was greater than 21 nucleotides (conventional cut-off used in
43
44 potential quadruplex prediction tools) suggesting metastable G-quadruplexes in pre-miRNAs to
45
46 be favoured at physiological conditions and in turn have a functional role in miRNA biology.
47
48 Balasubramanian and group recently attribute G-quadruplex to be the dominant conformer at
49
50 physiologically relevant conditions using NMR studies under varying magnesium and potassium
51
52
53
54
55
56
57
58
59
60

concentrations.⁶ Similar results have also been reported by Basu and group for miR-92b which contains a highly stable G-quadruplex in its premature form.³¹

G2 quadruplexes with shorter loops are more metastable in nature with their reported thermal stabilities in range of physiological temperatures.⁷ These metastable modules could be more dynamic inside the cell and hence need to be investigated for their functional roles. We tested the feasibility of G-quadruplex stem-loop equilibrium using a G2 quadruplex harbouring pre-miRNA let-7e. Pre-miRNA let-7e contains a G2 quadruplex with total loop length of seven nucleotides. Previous studies from our lab have successfully demonstrated G2 quadruplexes to be less stable than G3 quadruplexes in RNA under *in vitro* conditions.⁷ In this study, Dicer cleavage assay using radiolabelled pre-let-7e suggested a role of G-quadruplex in pre-let-7e under physiologically relevant concentrations of magnesium and potassium. Although favourable *in vitro* conditions implicate the role of potassium dependent G-quadruplex in miRNA biology, we further validated its effect on pre-miRNAs inside the cell.

Molecular crowding conditions that mimic the cellular environment favour the formation of G-quadruplexes over DNA duplexes.⁴³ Since canonical stem-loop structure is majorly a result of Watson-Crick base pairing, role of G-quadruplex inside the cell and its effect on miRNA maturation was further examined. Synthetic and natural pre-miRNAs harbouring potential G-quadruplexes were transfected along with mutant G-quadruplexes and mature miRNA levels were quantified using qRT-PCR. *In vivo* Dicer produced significantly less mature miRNA levels for G-quadruplex harbouring pre-miRNAs suggesting the role of structural interference by G-quadruplex in governing mature miRNA levels inside the cellular milieu. To further test the functionality of this G-quadruplex mediated structural interference in pre-miRNA processing, we quantified the target levels of G-quadruplex harbouring pre-miRNAs. Mature miRNA molecules

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

target mRNA UTR regions and regulate gene expression levels. Luciferase reporter assays for pre-let-7e and synthetic Quad-pre-miR-27a displayed more signals for quadruplex harbouring pre-miRNAs in comparison to mutant or no G-quadruplex containing pre-miRNAs. Similar observations have been reported recently for pre-miR-92b further supporting the hypothesis of structural interference role of G-quadruplex in pre-miRNAs.³¹ To confirm the identity of this potassium dependent structure to be a G-quadruplex, transfected cells were treated with TmPyP₄. Selective ligand TmPyP₄ is a RNA G-quadruplex destabiliser that distorts RNA G-quadruplexes.^{48,49} Treatment of cells with TmPyP₄ resulted in increased miRNA levels possibly by shifting the equilibrium between G-quadruplex and Stem-loop towards canonical Stem-loop structures which were more efficiently cleaved by *in vivo* Dicer. Bioinformatic analyses predicted the occurrence of G-quadruplexes in 13 % of all premature miRNAs with G2 ($N_{\text{sum}} \leq 7$) and G3 ($N_{\text{sum}} \leq 22$), a cut off based on the obtained experimental results.. The total percentage increased to 23 % with G2 ($N_{\text{sum}} \leq 9$) and G3 ($N_{\text{sum}} \leq 45$), a cut off based on the thermal stability of RNA G-quadruplexes. To understand and establish the extent of G-quadruplex mediated interference in pre-miRNA processing in a more general manner, cells were treated with TmPyP₄ and qRT-PCR was performed to quantify the production of mature miRNAs for 100 putative G-quadruplex harbouring pre-miRNAs. A significant up-regulation was observed in levels of mature miRNAs upon TmPyP₄ treatment cementing the role of structural interference mechanism of G-quadruplex in pre-miRNAs. Careful bioinformatics analyses of mean folding energies and GC content suggests that mean folding energies do not vary much among pre-miRNAs with and without putative G-quadruplex motifs. However GC content vary significantly. Average GC content of pre-miRNAs is 51.3%. Average GC content for pre-miRNAs harbouring G-quadruplexes is 63.9 % for predicted G2 and 69.12 % for predicted G3

quadruplex harbouring pre-miRNAs whereas average GC content for non-quadruplex pre-miRNAs is 49.44%. (Supplementary table sheet6) These values clearly suggest that G-quadruplex in pre-miRNAs may have evolved to act as regulatory switches that form in response to appropriate stimulus and influence mature miRNA levels and in turn target mRNAs. Similar results have also been reported recently when this manuscript was in preparation by Basu and group for miR-92b. Reporter systems validate the role of G-quadruplex as a potential switch that regulates levels of miR-92b inside the cell.³¹ We further extend those observations to metastable G-quadruplexes both G2 and G3 and also provide a more comprehensive scenario of G-quadruplex mediated structural interference on Dicer processing and its effect on miRNA maturation.

Interference with miRNA processing pathways is an attractive therapeutic target and various small molecules have been used to target oncogenic miRNAs and decrease their levels inside the cell.^{29,30} Herein we unravel another layer of regulation for miRNAs inside the cell mediated by alternate structure formation in premature miRNAs. Various G-quadruplex stabilising/destabilising proteins and other interacting partners could shift the equilibrium between G-quadruplex and stem-loop in response to specific stimuli and govern mature miRNA levels. Also since these structures are sensitive to ionic change inside the cells, role of alternate structures and equilibrium between them for neuronal pre-miRNAs which experience potassium flux changes would be worth exploring in future. Role of G-quadruplexes as zipcodes in neurites have already been reported¹⁴ hinting that G-quadruplex effect on neuronal miRNAs is worth exploring in future. Thus, this study provides insights into a global role of putative G-quadruplexes in pre-miRNAs and their influence on mature miRNA levels paving way to

understanding the impact of structural interference on miRNA processing pathways and production of miRNAs inside the cell.

ASSOCIATED CONTENT

Supporting information.

Scheme S1: Design of molecular beacon approach. **Scheme S2:** Representative stem-loop structures of pre-let-7e and pre-let-7e MQ (mutant quadruplex). **Figure S1:** Molecular beacon approach to measure mature miRNA levels. **Figure S2:** Dicer cleavage assay for pre-miR-27a and Quad pre-miR-27a. **Figure S3:** CD melting profile of A] Quad-pre-miR-27a and Stem-loop pre-miR-27a. B] Pre-let-7e and pre-let-7e MQ. **Figure S4:** Loop composition profile of pre-miRNAs harbouring G-quadruplexes in their sequence. **Figure S5:** Dicer cleavage assay for pre-let-7e and pre-let-7e MQ. **Figure S6:** MCR-ALS analysis of pre-let-7e and pre-let-7e MQ (mutant quadruplex). **Figure S7:** CD spectra of Quad pre-miR-27a and pre-let7e in presence and absence of 10 μ M TmPyP₄. **Supplementary Table 1:** Bioinformatic prediction of pre-miRNAs harbouring G-quadruplexes in their sequences detailed in separate excel sheets, alongwith their GC content and folding energy parameters and log10 fold change values of miRNA levels upon TmPyP₄ treatment. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*souvik@igib.res.in

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

This work was supported by Project BSC0123 (Project: Genome Dynamics in Cellular Organization, Differentiation and Enantiostasis) from the Council of Scientific and Industrial Research (CSIR), Government of India; and by Swarnajayanti project to Dr. Souvik Maiti from the Department of Science and Technology (DST), Government of India.

Notes

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Balaram Ghosh for IL-13 UTR plasmid used in luciferase studies.

REFERENCES

1. Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K. and Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure, *Nucleic Acids Res.* 34, 5402-5415.
2. Huppert, J.L. and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome, *Nucleic Acids Res.* 35, 406-413.
3. Huppert, J.L., Bugaut, A., Kumari, S. and Balasubramanian, S. (2008) G-quadruplexes: the beginning and end of UTRs, *Nucleic Acids Res.* 36, 6260-6268.

4. Kumari, S., Bugaut, A. and Balasubramanian, S. (2008) Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5' UTR of the NRAS proto-oncogene, *Biochemistry* 47, 12664-12669.
5. Zhang, A.Y., Bugaut, A. and Balasubramanian, S. (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology, *Biochemistry* 50, 7251-7258.
6. Bugaut, A., Murat, P. and Balasubramanian, S. (2012) An RNA hairpin to G-quadruplex conformational transition, *J. Am. Chem. Soc.* 134, 19953-19956.
7. Pandey, S., Agarwala, P. and Maiti, S. (2013) Effect of loops and G-quartets on the stability of RNA G-quadruplexes, *J. Phys. Chem. B* 117, 6896-6905.
8. Jodoin, R., Bauer, L., Garant, J.M., Mahdi Laaref, A., Phaneuf, F. and Perreault, J.P. (2014) The folding of 5'-UTR human G-quadruplexes possessing a long central loop, *RNA* 20, 1129-1141.
9. Arora, A., Dutkiewicz, M., Scaria, V., Hariharan, M., Maiti, S. and Kurreck, J. (2008) Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif, *RNA* 14, 1290-1296.
10. Kumari, S., Bugaut, A., Huppert, J.L. and Balasubramanian, S. (2007) An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation, *Nat. Chem. Biol.* 3, 218-221.

11. Morris, M.J. and Basu, S. (2009) An unusually stable G-quadruplex within the 5'-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells, *Biochemistry* 48, 5313-5319.
12. Morris, M.J., Negishi, Y., Pazsint, C., Schonhoft, J.D. and Basu, S. (2010) An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES, *J. Am. Chem. Soc.* 132, 17831-17839.
13. Collie, G.W., Haider, S.M., Neidle, S. and Parkinson, G.N. (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex, *Nucleic Acids Res.* 38, 5569-5580.
14. Subramanian, M., Rage, F., Tabet, R., Flatter, E., Mandel, J.L. and Moine, H. (2011) G-quadruplex RNA structure as a signal for neurite mRNA targeting, *EMBO Rep.* 12, 697-704.
15. Fisette, J.F., Montagna, D.R., Mihailescu, M.R. and Wolfe, M.S. (2012) A G-rich element forms a G-quadruplex and regulates BACE1 mRNA alternative splicing, *Journal of Neurochemistry* 121, 763-773.
16. Biffi, G., Tannahill, D., McCafferty, J. and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells, *Nat. Chem.* 5, 182-186.
17. Biffi, G., Di Antonio, M., Tannahill, D. and Balasubramanian, S. (2014) Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells, *Nat. Chem.* 6, 75-80.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

18. Gros, J., Guedin, A., Mergny, J.L. and Lacroix, L. (2008) G-Quadruplex formation interferes with P1 helix formation in the RNA component of telomerase hTERC, *ChemBioChem* 9, 2075-2079.

19. Wan, Y., Kertesz, M., Spitale, R.C., Segal, E. and Chang, H.Y. (2011) Understanding the transcriptome through RNA structure, *Nat. Rev. Genet.* 12, 641-655.

20. Jayaraj, G.G., Pandey, S., Scaria, V. and Maiti, S. (2012) Potential G-quadruplexes in the human long non-coding transcriptome, *RNA Biol.* 9, 81-86.

21. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions, *Cell* 136, 215-233.

22. Soifer, H.S., Rossi, J.J. and Saetrom, P. (2007) MicroRNAs in disease and potential therapeutic applications, *Mol. Ther.* 15, 2070-2079.

23. Lujambio, A. and Lowe, S.W. (2012) The microcosmos of cancer, *Nature* 482, 347-355.

24. Mendell, J.T. and Olson, E.N. (2012) MicroRNAs in stress signaling and human disease, *Cell* 148, 1172-1187.

25. Zampetaki, A. and Mayr, M. (2012) MicroRNAs in vascular and metabolic disease, *Circ. Res.* 110, 508-522.

26. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. et al. (2003) The nuclear RNase III Drosha initiates microRNA processing, *Nature* 425, 415-419.

27. Kim, V.N., Han, J. and Siomi, M.C. (2009) Biogenesis of small RNAs in animals. Nature reviews, *Mol. Cell. Biol.* 10, 126-139.
28. Young, D.D., Connelly, C.M., Grohmann, C. and Deiters, A. (2010) Small molecule modifiers of microRNA miR-122 function for the treatment of hepatitis C virus infection and hepatocellular carcinoma, *J. Amer. Chem. Soc.* 132, 7976-7981.
29. Deiters, A. (2010) Small molecule modifiers of the microRNA and RNA interference pathway, *AAPS J.* 12, 51-60.
30. Bose, D., Jayaraj, G., Suryawanshi, H., Agarwala, P., Pore, S.K., Banerjee, R. and Maiti, S. (2012) The tuberculosis drug streptomycin as a potential cancer therapeutic: inhibition of miR-21 function by directly targeting its precursor, *Angew. Chem. Int. Ed. Engl.* 51, 1019-1023.
31. Mirihana Arachchilage, G., Dassanayake, A.C. and Basu, S. (2015) A Potassium Ion-Dependent RNA Structural Switch Regulates Human Pre-miRNA 92b Maturation, *Chem. Biol* 22, 262-272.
32. Jaumot J, Gargallo R, de Juan A, Tauler R. (2005) A graphical user-friendly interface of mCR-ALS; a new tool for multivariate curve resolution in MATLAB, *CHEMOMETRICS AND INTELLIGENT LABORATORY SYSTEMS* 76, 101-110.
33. Kumar, P., Verma, A., Maiti, S., Gargallo, R. and Chowdhury, S. (2005) Tetraplex DNA transitions within the human c-myc promoter detected by multivariate curve resolution of fluorescence resonance energy transfer, *Biochemistry* 44, 16426-16434.

34. Jaumot, J., Eritja, R., Tauler, R. and Gargallo, R. (2006) Resolution of a structural competition involving dimeric G-quadruplex and its C-rich complementary strand, *Nucleic Acids Res.* **34**, 206-216.
35. Jaumot J, Tauler R, Gargallo R. (2006) Exploratory data analysis of DNA microarrays by multivariate curve resolution, *Anal. Biochem.* **358**, 76-89.
36. Felten, J., Hall, H., Jaumot, J., Tauler, R., de Juan, A. and Gorzsas, A. (2015) Vibrational spectroscopic image analysis of biological material using multivariate curve resolution-alternating least squares (MCR-ALS), *Nat. Prot.* **10**, 217-240.
37. de Oliveira, R.R., de Lima, K.M., Tauler, R. and de Juan, A. (2014) Application of correlation constrained multivariate curve resolution alternating least-squares methods for determination of compounds of interest in biodiesel blends using NIR and UV-visible spectroscopic data, *Talanta* **125**, 233-241.
38. Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR, , *Nucleic Acids Res.* **29**, e45.
39. Bose, D., Jayaraj, G.G., Kumar, S. and Maiti, S. (2013) A molecular-beacon-based screen for small molecule inhibitors of miRNA maturation, *ACS. Chem. Biol.* **8**, 930-938.
40. Provost, P., Dishart, D., Doucet, J., Frendewey, D., Samuelsson, B. and Radmark, O. (2002) Ribonuclease activity and RNA binding of recombinant human Dicer, *Embo. J.* **21**, 5864-5874.
41. Kypr, J., Kejnovska, I., Renciuk, D. and Vorlickova, M. (2009) Circular dichroism and conformational polymorphism of DNA, , *Nucleic Acids Res.* **37**, 1713-1725.

42. Jaumot, J., Escaja, N., Gargallo, R., Gonzalez, C., Pedroso, E. and Tauler, R. (2002) Multivariate curve resolution: a powerful tool for the analysis of conformational transitions in nucleic acids, *Nucleic Acids Res.* 30, e92.
43. Kumar, N. and Maiti, S. (2008) Role of molecular crowding in perturbing quadruplex-Watson Crick duplex equilibrium, *Nucleic Acids Symp. Ser. (Oxf)* 52, 157-158.
44. Scaria, V., Hariharan, M., Arora, A. and Maiti, S. (2006) Quadfinder: server for identification and analysis of quadruplex-forming motifs in nucleotide sequences, *Nucleic Acids Res.* 34, W683-685.
45. Kozomara, A. and Griffiths-Jones, S. (2014) miRBase: annotating high confidence microRNAs using deep sequencing data, *Nucleic Acids Res.* 42, D68-73.
46. Jerome, T., Laurie, P., Louis, B. and Pierre, C. (2007) Enjoy the Silence: The Story of let-7 MicroRNA and Cancer, *Curr. Genomics* 8, 229-233.
47. Kumar, M., Ahmad, T., Sharma, A., Mabalirajan, U., Kulshreshtha, A., Agrawal, A. and Ghosh, B. (2011) Let-7 microRNA-mediated regulation of IL-13 and allergic airway inflammation, *J. Allergy Clin. Immunol.* 128, 1077-1085.
48. Morris, M.J., Wingate, K.L., Silwal, J., Leeper, T.C. and Basu, S. (2012) The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells, *Nucleic Acids Res.* 40, 4137-4145.
49. Zamiri, B., Reddy, K., Macgregor, R.B., Jr. and Pearson, C.E. (2014) TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)_n repeat of

the C9orf72 gene and blocks interaction of RNA-binding proteins, *J. Biol. Chem.* 289, 4653-4659.

FIGURE LEGENDS

Scheme 1. A] Stem-loop pre-miR-27a and B] Quad-pre-miR-27a are represented in form of stem-loop structures. Nucleotides in capital letters denote the mutations in natural pre-miR-27a to introduce a putative quadruplex sequence G₃N_xG₃N_xG₃N_xG₃ (G-quartets represented by red letters). Mature miR-27a sequence is presented in the blue box region.

Figure 1. Molecular beacon approach to measure mature miRNA levels. Stem-loop pre-miR-27a and Quad pre-miR-27a were folded in presence of 1 mM MgCl₂ (Mg²⁺), 1 mM MgCl₂ + 150 mM KCl (Mg²⁺ + K⁺) and subjected to Dicer cleavage. Mature miRNA levels were measured using fluorescence from molecular beacon from time zero till a saturation was reached in the fluorescence values, B] Less mature miRNA levels were observed in Quad pre-miR-27a (red) as compared to Stem loop pre-miR-27a (blue) after 90 minutes of reaction in 1 mM MgCl₂ +150 mM KCl (Mg²⁺+K⁺) while potassium decreases the Dicer activity significantly for both Quad pre-miR-27a and Stem-loop pre-miR-27a. Error bars represent the standard error calculated with a p-value < 0.05 for n=3 obtained from Students t-test.

Figure 2. Quad-pre-miR-27a was folded in presence of different salt conditions and subjected to RNase T1 digestion. Samples are- Alkaline hydrolysis ladder (lane AH), T1 ladder (lane T1), T1 digestion in presence of 1 mM MgCl₂ (lane Mg), 150 mM KCl (lane K), 1 mM MgCl₂ and 150 mM KCl (lane MgK) and 150 mM LiCl (lane Li). Protected guanines are marked in the gel and denoted in bold letters in Quad-pre-miR-27a sequence.

Figure 3. MCR-ALS analysis of Quad-pre-miR-27a and Stem-loop-pre-miR-27a. Panels a) and c) Distribution profiles for Quad-pre-miR-27a (panel a) and Stem-loop-pre-miR-27a (panel c) where three different structural species namely G-quadruplex (blue), duplex (green) and single stranded (pink) structures are represented. The relative concentration of each of the conformation at respective temperature is shown. Panels b) and d) are pure spectra obtained from CD melting after resolution of different structural species.

Figure 4. *In cellulo* analysis of miR-27a. A] qRT-PCR analysis shows less mature miR-27a production in Quad-pre-miR-27a (putative quadruplex) than pre-miR-27a (no quadruplex). U6 was used as normalization gene. Control denotes untransfected sample with endogenous levels of miR-27a. B] Luciferase assay shows more reporter signals in Quad-pre-miR-27a than pre-miR-27a transfected sample. Prohibitin 3' UTR was cloned under Renilla gene in dual luciferase system wherein Firefly served as normalization control. Plasmid denotes sample containing only miR-27a target vector. Data was analysed using one-way ANOVA, $n=4$, * p value < 0.05 and *** p value < 0.001 . Error bars represent standard error of the samples.

Figure 5. A] Pie chart distribution of G-quadruplexes in human pre-miRNAs with G2 ($N_{\text{sum}} \leq 7$) and G3 ($N_{\text{sum}} \leq 22$) as consensus motif. B] Bar graph denoting the GC content (in percentage) against Mean folding energy (MFE) [kcal/mol]. Bar graphs indicate average GC content to be higher in G-quadruplex harbouring pre-miRNAs while average MFE remains more or less similar for all pre-miRNAs. C] and D] Loop distribution profiles for G2 ($N_{\text{sum}} \leq 7$) and G3 ($N_{\text{sum}} \leq 22$) predicted G-quadruplexes. Mononucleotide and dinucleotide loops are enriched suggesting presence of stable G-quadruplexes in pre-miRNAs.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 6. *In cellulo* analysis of G-quadruplex in pre-let-7e. A] qRT-PCR analysis shows less mature let-7e production in pre-let-7e (putative quadruplex) than pre-let-7e MQ (no quadruplex). Control denotes untransfected sample with endogenous levels of miR-27a. B] Luciferase assay shows more reporter signals in pre-let-7e than pre-let-7e MQ transfected sample. IL-13 UTR was cloned under Firefly reporter gene and Renilla plasmid was used for normalisation of reporter signals. Plasmid denotes sample transfected only with let-7e target vector. Data was analysed using one-way ANOVA, n=4, **p value< 0.01 and *** p value <0.001. Error bars represent standard error of the samples. Error bars represent standard error.

Figure 7. TmPyP₄ destabilizes RNA G-quadruplexes in pre-miRNAs. A] Quad-pre-miR-27a showed a three-fold increase and pre-let-7e showed an increase of four fold in let-7e levels as compared to non-quadruplex harbouring premiRNAs (Stem-loop-pre-miR-27a and pre-let-7e MQ) upon 10 µM TmPyP₄ treatment. B] Luciferase assay shows more reporter signals in Quadruplex harbouring pre-miRNA (Quad-pre-miR-27a and pre-let-7e) as compared to non-quadruplex premiRNAs (Stem-loop-pre-miR-27a and pre-let-7e MQ) transfected sample. Control denotes untransfected sample with endogenous levels of miR-27a and let-7e respectively. Plasmid denotes sample transfected only with miR-27a and let-7e target vector respectively. Data was analysed using two-way ANOVA, n=4, *p value< 0.05 and *** p value <0.001. Error bars represent standard error of the samples. Error bars represent standard error.

Figure 8. Heat map with fold change differences in the mature miRNA levels of G-quadruplex pre-miRNAs upon TmPyP₄ treatment. Expression levels changes are plotted as log₁₀ fold changes of levels of miRNA upon 10 µM TmPyP₄ treatment normalized to control or untreated sample. A strong correlation in fold change values was observed in four different treatment sets (T1-T4). Colour key in the top left column denotes no change in expression (red) and up-

regulated miRNAs as black and green depending on change in their levels upon TmPyP₄ treatment. miRNAs denoted in the box served as control miRNAs which did not possess a putative G-quadruplex sequence in their premature forms.

FIGURES

Scheme 1

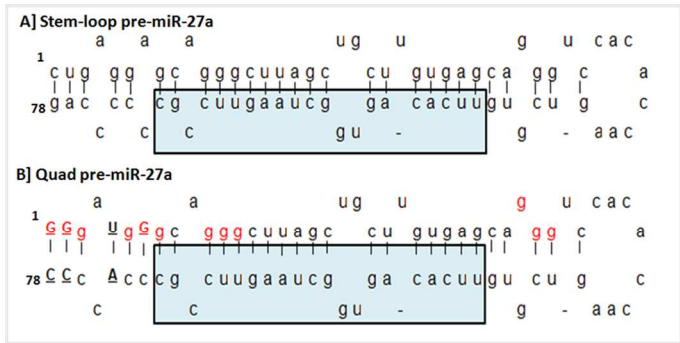


Figure 1

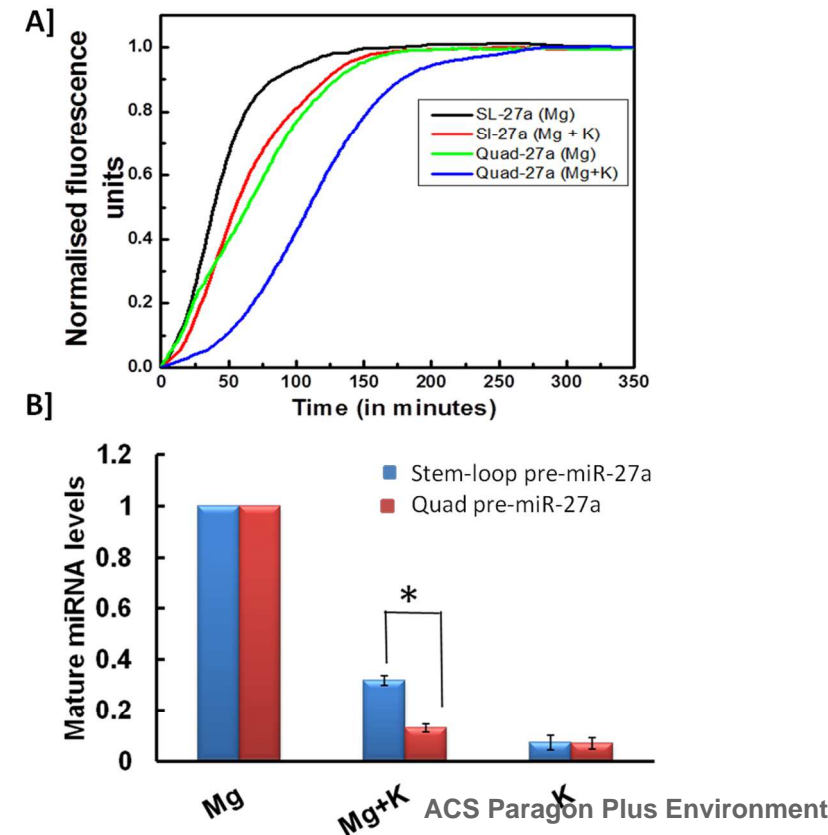


Figure 2

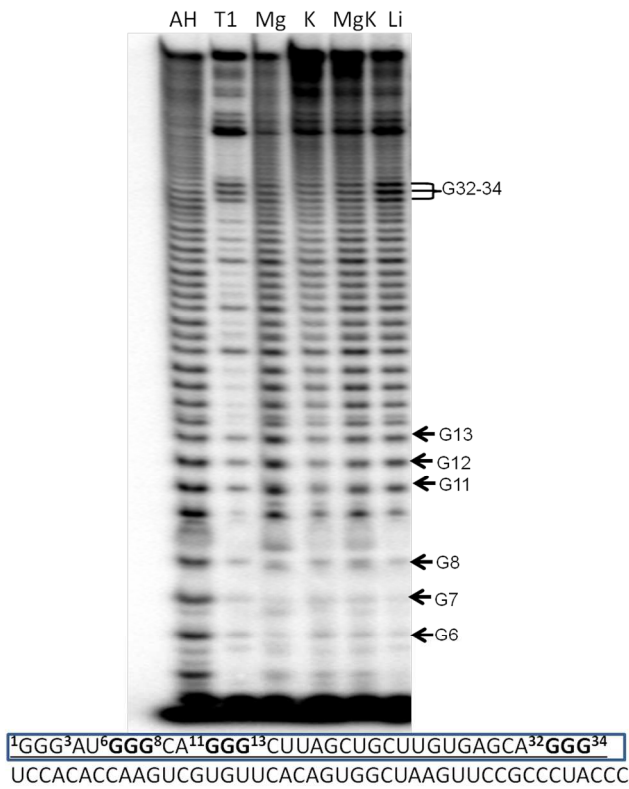


Figure 3

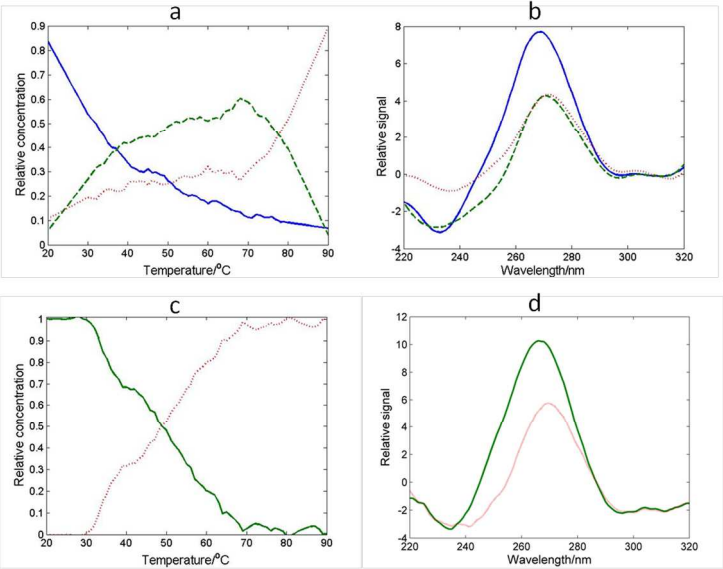


Figure 4

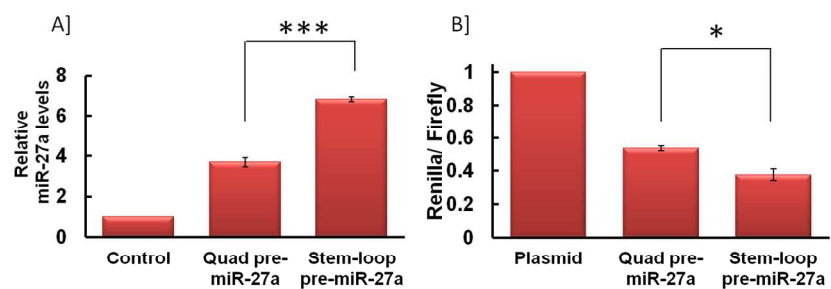


Figure 5

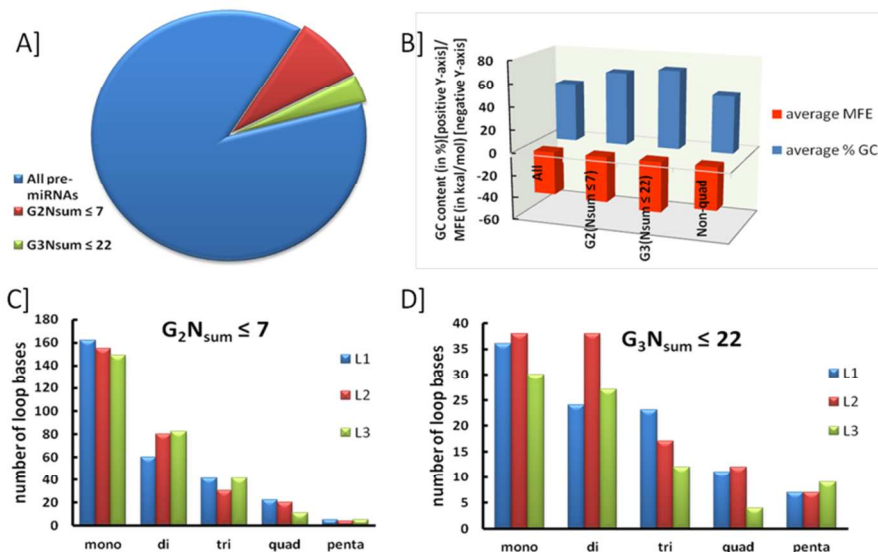


Figure 6

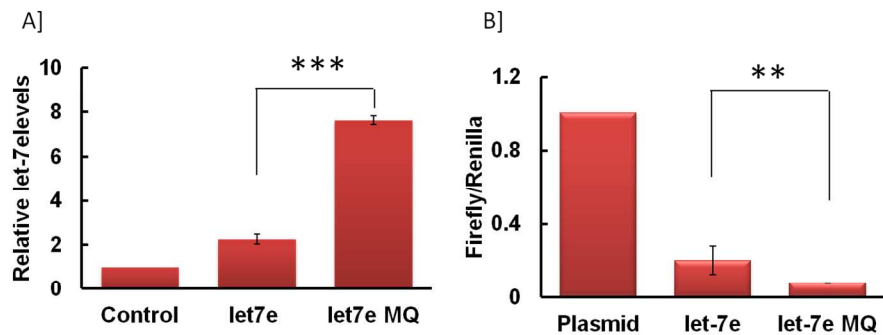
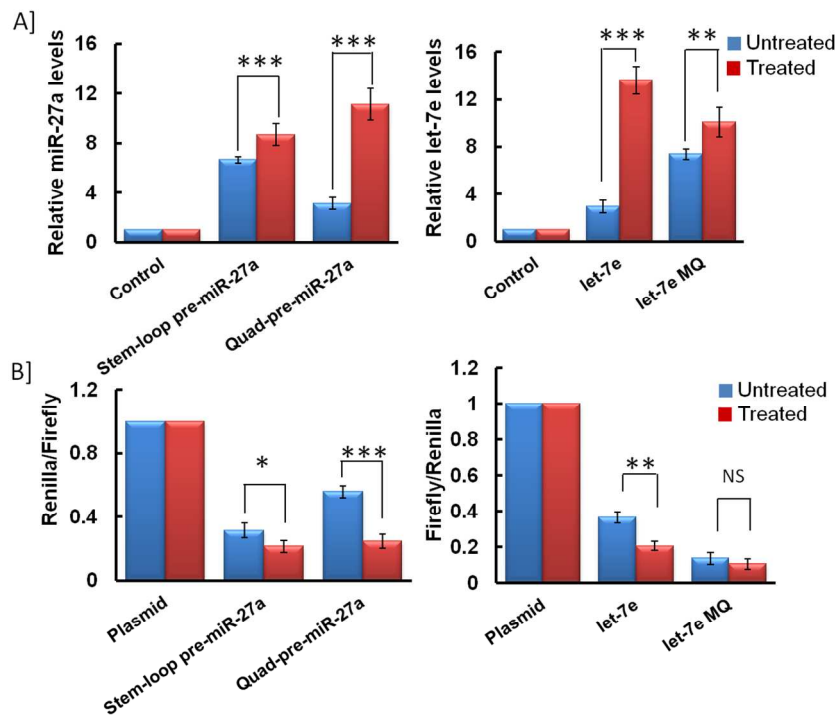
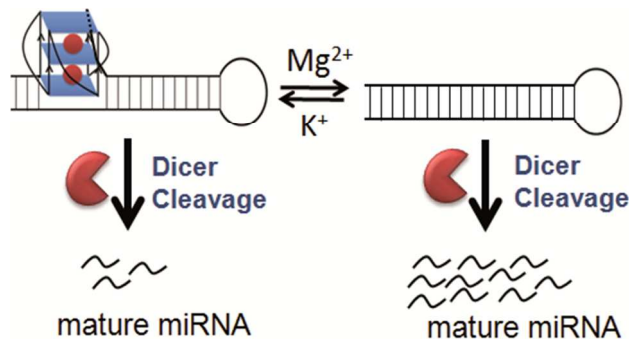


Figure 7



Insert Table of Contents Graphic and Synopsis Here



TOC. Putative quadruplex sequences present in pre-miRNA potentially competes with canonical stem-loop structures. In cellulo conditions favor G-quadruplex formation which results in interference with Dicer cleavage. Mature miRNA levels serve as readout for G-quadruplex formation. Less mature miRNA indicates G-quadruplex existence and provides a proof for equilibrium shift between G-quadruplex and stem-loop (duplex) RNA structures inside the cell.