ChemPubSoc Europe

DOI: 10.1002/cbic.201900336

CHEMBIOCHEM Full Papers

Presence of Potential G-Quadruplex RNA-Forming Motifs at the 5′-UTR of *PP2Acα* mRNA Repress Translation

Sushree Prangya P. Pany,[a] Mahak Sapra,[b] Jitendar Sharma,[a] V. Dhamodharan,[a] Swati Patankar,[c] and P. I. Pradeepkumar*[a]

Dedicated to Professor Jyoti Chattopadhyaya on the occasion of his 70th birthday

RNA G-quadruplex (G4)-forming motifs present at the 5′-UTR of the protein phosphatase (*PP2Ac*) gene are the regulatory targets of the fragile X mental retardation protein (FMRP), which is weakly expressed in Fragile X patients. Herein, we report that the existence of such G4-forming sequence represses the translation of the *PP2Acα* gene. This study opens therapeutic avenues to design small molecule ligands that mimic the function of the FMRP.

Introduction

Tandem guanine-rich sequences can fold to form four-stranded secondary structures in the presence of metal cations known as G-quadruplexes (G4s).[1] Potential G4-forming sequences are present in the human genome with a high frequency in the promoter and telomeric regions.[2] Also, they are present in the various regulatory regions of 5′- and 3′-untranslated regions (UTRs) and play a crucial role in the control of gene expression.[3] Though DNA G4 structures have been studied extensively, RNA G4 structures have also gained significant attention in the past few decades.[4] As RNA is single-stranded, the RNA G4 structure does not have to compete with the complementary strand to form G4 structures. RNA G4 structures are involved in various cellular functions like translation modulation, alternative splicing regulation, mRNA localization, and polyadenylation.[5] Inhibition of translation by RNA G4 structure was first reported in the *NRAS* proto-oncogene[6] and was subsequently verified in cellular systems for the *Zic-1* gene.[7] Successively, the functional relevance of such G4 structures has been reported in the 5′-UTR of *MT3-MMP*,[8] *TRF2*,[9] *BCL2*,[10] *CCND3*,[11] and *ADAM10*[12] all of which have an inhibitory effect on the gene expression. Moreover, a G4 structure located in the 3′-UTR of proto-oncogene *PIM1* was shown to repress protein synthesis.[13] Further suppression of the gene expression has been achieved by various G4-stabilizing small molecule ligands, such as RR110, 360A, PhenDC3, and PhenDC6.[14] Though a recent report has shown that G4 structures are mostly unfolded inside the cells, the effect of transient formation of these structures along with the role of G4-stabilizing proteins may contribute to their regulatory role reported in several studies.[15] In addition to the suppression of translation achieved by G-quadruplex motifs, their presence is also found to be essential for the initiation of the translation process of genes through a cap-independent process.[16]

Recently, the role of RNA-binding proteins, such as FMRP (fragile X mental retardation protein) and FMR2P (fragile X mental retardation 2 protein), on gene expression by interacting with G4 structures are reported.[17] Impaired development of FMRP is linked to an inherited intellectual disability known as fragile X syndrome.[18] One of the targets of FMRP is the G4-forming motif in the catalytic subunit of the protein phosphatase (*PP2Ac*).[19] FMRP acts as a negative regulator of the *PP2Ac* mRNA translation by binding with the G4 structure formed in the 5′-UTR of *PP2Ac*.[19b] Weak expression of FMRP increases the expression of *PP2Ac* and thereby alters the actin remodeling, which is a key process in morphogenesis of dendritic spines in FXS patients.[19b]

Herein, we show the presence of a putative G4-forming sequence in the 5′-UTR of the *PP2Acα* gene, which can fold to form a G4 structure and thereby suppress translation. Bioinformatics analysis revealed the existence of five G4-forming sequences upstream to the AUG start codon in the 5′-UTR of the *PP2Acα* mRNA (Figure S1 in the Supporting Information). We report the biophysical and functional characterization of the 22-mer G-rich sequence present four bases upstream to the translation start codon (Figure 1 A). We have also evaluated the effects of small molecules on this G4 structure that had been previously reported by us as specific stabilizing ligands for parallel G4 DNAs (Figure 1 B).[20]

[a] S. P. P. Pany, J. Sharma, Dr. V. Dhamodharan, Dr. P. I. Pradeepkumar
Department of Chemistry, Indian Institute of Technology Bombay
Mumbai 400076 (India)
E-mail: pradeep@chem.iitb.ac.in

[b] M. Sapra
Centre for Research in Nanotechnology and Science
Indian Institute of Technology Bombay
Mumbai 400076 (India)

[c] Prof. S. Patankar
Department of Bioscience and Bioengineering
Indian Institute of Technology Bombay
Mumbai 400076 (India)

Supporting information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201900336.

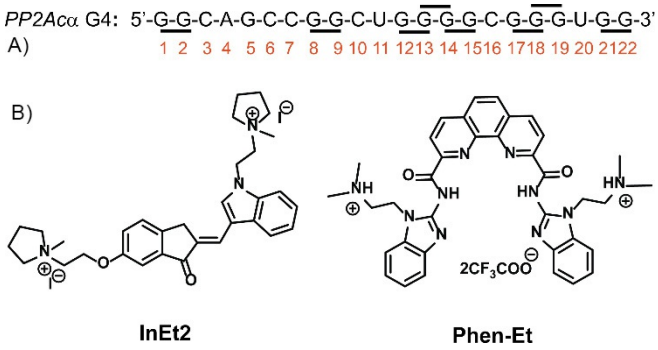

Figure 1. A) The 22-mer *PP2Acα* G4 RNA with marked G stretches, which can form G4 structures. B) Structures of the G4-stabilizing ligands InEt2 and Phen-Et.

Results and Discussion

To investigate the influence of the *PP2Acα* G4-forming RNA on translation, the sequence was cloned in plasmid psiCHECK-2, transfected into HeLa cells, and translational activity was assessed by using a dual luciferase reporter assay. Both the 22-mer *PP2Acα* G4-forming sequence and its corresponding mutant (Table S1) were cloned at the Nhe1 restriction site upstream to the *Renilla* luciferase gene (*Rluc*) translation start site (Figure S2).[21,22] After 48 h of transient transfection, the plasmid containing the *PP2Acα* G4 RNA exhibited decreased gene expression (≈70%) as compared to the positive control plasmid psiCHECK-2 (Figure 2). However, the insertion of mutant *PP2Acα* G4-forming RNA showed only marginal decrease in gene expression (Figure 2). Effects of the addition of the G4-stabilizing compound InEt2 were also assessed by using the dual reporter assay (data not shown). InEt2 being a G4-stabilizing compound should further decrease the gene expression from a plasmid containing *PP2Acα* G4, however no significant inhibitory effect was observed. Though this was surprising, but could probably be attributed to various reasons, such as poor cell permeability of the ligand, off-target effects, protection of G4 structure with G4-interacting proteins etc. Similarly, the effects of other two nearest guanine-rich short sequences (13- and 14-mer) and their mutant sequences on translation were also tested (Figure S3). These two sequences did not show any significant reduction in the gene expression. Overall, the reporter assay shows that only the 22-mer G4-forming sequence can efficiently repress the gene expression.

To explore the structural properties of G4-forming RNAs, we have carried out biophysical studies. To confirm the formation of G4 structure, circular dichroism (CD) titration experiments were performed with different concentrations of monovalent metal ions. A positive spectrum at $\lambda = 264$ nm and a negative spectrum at $\lambda = 240$ nm were noticed in the absence of any added metal ions, which indicate the parallel topology[23] adopted by 22-mer *PP2Acα* G4 RNA sequences (Figure 3A). An increase in the peak ellipticity of the CD signal was observed with the 22-mer RNA upon addition of increasing concentration of K^+ ions (Figure 3A). This indicates further induction of the parallel topology adopted by the 22-mer sequence. On the other hand, the 13-mer G-rich sequence even in the presence of 100 mM K^+ ions showed only a positive peak at $\lambda = 264$ nm, which does not attribute towards any characteristic peak for a

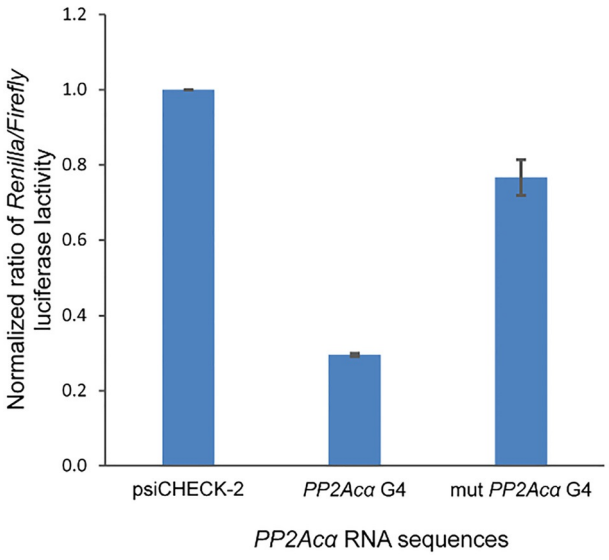

Figure 2. Luciferase activity of plasmid containing *PP2Acα* G4 and mutant *PP2Acα* after 48 h of transfection in HeLa cells. The Renilla luciferase activity was divided by the firefly luciferase activity, and the ratio was normalized to the positive control psiCHECK-2. The error bar represents the standard deviation (±SD) from three independent experiments.

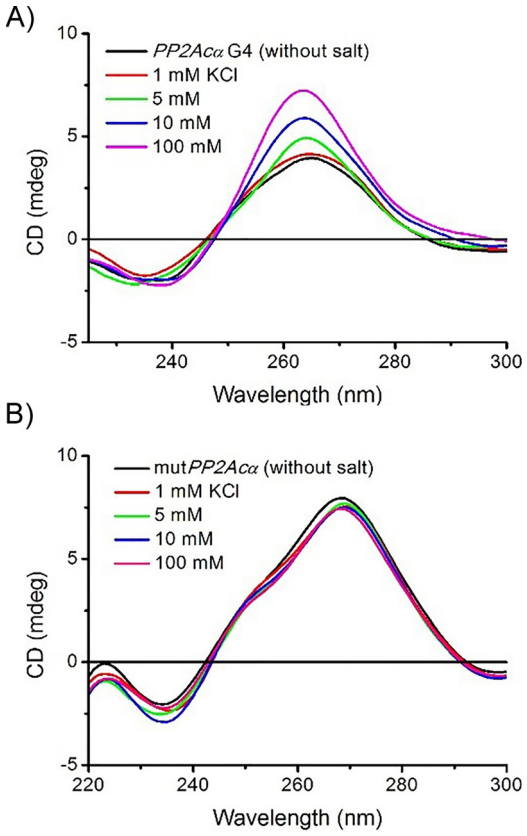

Figure 3. CD spectra of A) the *PP2Acα* G4 RNA and B) the mutant *PP2Acα* RNA sequence (12.5 μM in 100 mM lithium cacodylate buffer, pH 7.2) with an increasing concentration of KCl (0–100 mM).

parallel G4 topology. However, it indicates the formation of a possible hairpin structure, which is further supported by secondary structure analysis. Under similar salt conditions, the 14-mer G4 RNA showed characteristic CD signals of parallel G4 topology (Figure S4). CD titration by using a mutant G4 RNA showed a positive peak at $\lambda = 265$ nm and a negative peak at $\lambda = 235$ nm, which could be attributable towards the possible hairpin structure adopted by the sequence (Figure 3 B).[24]

To determine the thermal stability of the structure formed by *PP2Acα* G4-forming RNA, CD a melting experiment was performed.[25] The thermal melting experiment was carried out by monitoring the changes in the CD intensity at $\lambda = 260$ nm in the presence of 10 mM KCl, which afforded a melting temperature of 63 °C (Figure 4) at 10 μM strand concentration of the

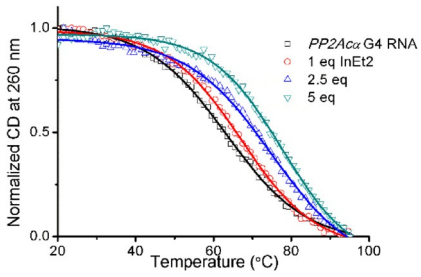

Figure 4. Normalized CD melting curve of the *PP2Acα* G4 RNA sequence (10 μM in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate buffer, pH 7.2) in the absence and in the presence of varying concentrations (0–5 equiv) of ligand InEt2.

22-mer G4 RNA. Under identical physiological conditions, 10 μM of the 13-mer and 14-mer G4 RNA yielded melting temperatures of 61.1 and 67.1 °C, respectively (Figure S5). Considering that the effective repression of the gene expression was observed with the 22-mer G4 RNA, the sequence was taken for further biophysical experiments. To investigate the effect of G4-stabilizing ligands on G4 RNA, we have performed melting experiments with the addition of five equivalent of InEt2 and Phen-Et. Adding five equivalent of the G4-stabilizing ligand InEt2 provided a significant increase in the melting temperature up to 15 °C (Figure 4 and Table 1). But, in the case of Phen-Et, only a 10 °C increase in the T_m was observed (Figure S6, and Table 1). To demonstrate the molecularity of G4

Table 1. Thermal stability of the *PP2Acα* G4 DNAs with varying concentration of ligands measured by CD melting experiments.

Ligand	ΔT_m[a] [°C]		
	1 equiv	2.5 equiv	5 equiv
InEt2	4.5 ± 0.1	12.1 ± 0.9	15.2 ± 0.1
Phen-Et	3.0 ± 0.3	5.3 ± 0.5	10.0 ± 0.4

[a] ΔT_m represents the difference in thermal melting [$\Delta T_m = T_m$(RNA+ligand) − T_m(RNA)]. All experiments were carried out with 10 μM *PP2Acα*RNA concentration in 10 mM KCl with 90 mM LiCl and 10 mM lithium cacodylate buffer pH 7.2. T_m in the absence of ligands is (63.2 ± 0.5) °C. The reported ΔT_m values are the average values with standard deviations from three independent experiments.

RNA, we have performed CD melting experiments by varying RNA strand concentrations (5–20 μM). These studies did not show any change in the T_m values, which indicates an intramolecular folding pattern by the G4 RNA (Figure S7). The same experiment was employed with the mutant RNA sequence, which exhibited a melting temperature of 73.9 °C (Figure S8). Further to confirm the formation of G4 structure, UV melting experiments were carried out with the 22-mer and the mutant sequence in the presence of 100 mM KCl by monitoring at $\lambda =$ 295 nm.

The results revealed that the 22-mer *PP2Acα* RNA forms a G4 structure with a melting temperature of 68 °C, whereas the mutant sequence forms a hairpin structure under similar conditions (Figure S9 A and B).

Overall, data obtained from CD and UV experiments reveal that under near physiological conditions the *PP2Acα* G-rich sequence can fold into a stable, parallel intramolecular G4 structure, which can be further be stabilized by G4-stabilizing ligands.

To investigate the ligand–G4 RNA interaction and to find out the binding constant UV titration experiments were conducted. The ligand InEt2 was selected for the experiment because of its high stabilization effect on G4 RNA. The UV/visible spectrum of InEt2 exhibited intense absorbance maxima at $\lambda = 280$ and 411 nm (Figure 5 A). The band at $\lambda = 411$ nm showed hypochromicity and a red shift of 15 nm upon addition of an in-

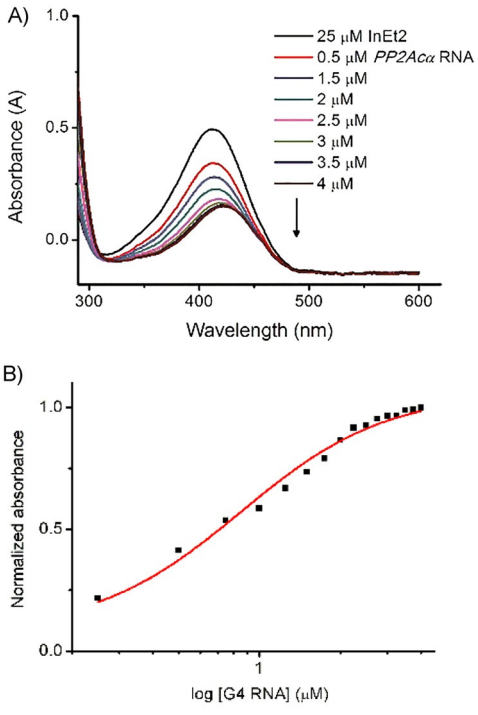

Figure 5. Absorption spectra of InEt2 complexed with *PP2Acα* G4 RNA and the corresponding binding plot. A) Absorption spectra of InEt2 (25 μM in 100 mM KCl and10 mM lithium cacodylate buffer, pH 7.2) with the addition of an increasing concentration of *PP2Acα* G4 RNA (0–4 μM under identical salt and buffer conditions). B) Plot of the normalized absorbance versus the logarithm of [G4 RNA]. The curve was fitted by using the Hill-1 equation with $n = 1.6$.

creasing concentration of *PP2Acα* G4 RNA (0–4 μM, Figure 5A). These indicates a strong interaction between InEt2 and the *PP2Acα* G4 RNA. A plot of the normalized absorbance against the logarithm of the concentration of G4 RNA provides the binding constant $K_a = 1.19 \times 10^6\ M^{-1}$ (Figure 5B). To determine the binding stoichiometry between the ligand and *PP2Acα* G4 RNA, a continuous variation binding analysis (Job plot) was performed. The Job plot analysis revealed a 2:1 mode of the binding interaction between the ligand InEt2 and *PP2Acα* RNA (Figure S10).

To elucidate the Gs involved in the G4 structure formation ribonuclease T1 (RNaseT1) footprinting experiments were performed. As marked in Figure 1A, the 22-mer *PP2Acα* G-rich RNA has five G stretches with four Gs flanked on both sides. Theoretically, such a sequence can adopt more than fifteen two quartet intramolecular G4 structures. The Gs, which are involved in the formation of the G4 structure, should be more protective from cleavage than those present in the loop region.[26] The footprinting results show that in the presence of 150 mM K^+ ions, the Gs toward the 3′-end (G8–G22) remain protected from cleavage. As expected in the case of Li^+ the protection was not very strong (Figure 6A). Further to elucidate the structure formed by the mutant sequence, we have treated the sequence with RNaseT1 in the presence of 150 mM K^+ and Li^+ ions, respectively. The results are in good agreement with the UV and CD experiments and validate the formation of the hairpin structure by the mutant sequence, which is independent of the nature of the monovalent ions used unlike in the case of the G-rich sequence (Figure S11).

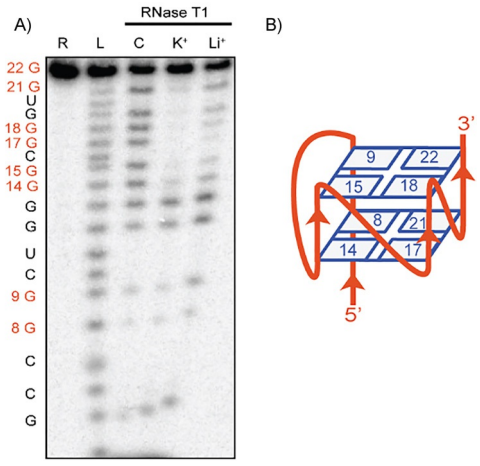

Figure 6. A) Polyacrylamide gel electrophoresis (PAGE) (20%, 7 M urea) of RNaseT1 footprinting of the 22-mer *PP2Acα* G4 RNA with different monovalent ions. R = labeled RNA, L = alkaline hydrolysis ladder, C = Control without added monovalent cation, K^+ = with 150 mM KCl, Li^+ = with 150 mM LiCl. B) Schematic representation of the G4 RNA structure.

Overall, the footprinting assay reveals the formation of a major G4 structure by the 22-mer RNA, which consists of two G tetrads formed by four G stretches, that is, G8–G9, G14–G15, G17–G18, and G21–G22, whereas G12, G13, and G19 are accommodated in the loops (Figure 6B).

Conclusion

In summary, we have shown that the presence of a putative G4-forming sequence in the 5′-UTR of the *PP2Acα* gene represses gene expression. Various biophysical and biochemical assays unambiguously revealed the formation of a stable intramolecular G4 structure. The stabilizing effects of G4-binding ligands were also elucidated. This study opens possible avenues to harness RNA G4 structures, which are targets of FMRP for drug discovery.

Experimental Section

Oligonucleotides: All DNA oligonucleotides used for cloning (Table S1) were purchased from Integrated DNA Technology (Belgium). The oligonucleotides were further purified by 20% denaturing PAGE (7 M urea) by using standard protocols. The RNA oligonucleotides (HPLC purified, Table S2) used for biophysical and biochemical experiments were purchased from Eurogentec (Belgium). The concentration of the oligomers was calculated by using appropriate molar extinction coefficients (Tables S1 and S2) obtained from the absorbance measured at $\lambda = 260$ nm by using a Perkin–Elmer Lambda Bio$^+$ UV spectrophotometer.

Ligand stock solutions: Ligands were prepared by using previously reported protocols.[20] The stock solution of InEt2 (5 mM) was made in water and the stock solution of ligand Phen-Et was made in 40% DMSO in water (*v*/*v*).

CD titration experiments: CD titration experiments were performed by using a Jasco 815 CD spectrometer. The strand concentration of the RNA used was 12.5 μM in 10 mM lithium cacodylate buffer, pH 7.2. Varying concentrations (0–100 mM) of monovalent K^+ ions were used for titrations. Typically, a 250 μL sample was taken in a quartz cuvette with an optical path length of 1 mm, placed in the spectrometer, and allowed to equilibrate at 20 °C for 10 min. Three CD scans, over the wavelength range $\lambda = 220$–320 nm were performed at 100 nm min^{-1} with a 2 s response time. For each experiment, the CD spectrum of the buffer was recorded and subtracted from the spectrum obtained for the RNA-containing solution.

CD and UV melting experiments: Melting experiments were performed by using a Jasco 815 CD spectrometer. For the melting studies, various RNA G4 strand concentrations (5, 10, 20 μM) were taken in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate, pH 7.2. The RNA was annealed by heating at 95 °C for 5 min, followed by gradual cooling to room temperature over 3–4 h. Ligands (1, 2.5, and 5 equivalents) were added to the annealed RNA (10 μM) and the mixtures were kept at 4 °C overnight. Thermal melting was monitored at $\lambda = 260$ nm with a heating rate of 1 °C min^{-1}. UV melting experiments were performed by using a Varian Cary UV spectrophotometer. RNA (10 μM) in salt (100 mM) and lithium cacodylate (10 mM, pH 7.2) was annealed by heating at 95 °C for 5 min, followed by gradual cooling to room temperature over 3–4 h. Melting was monitored at $\lambda = 295$ nm at a ramp rate of 1 °C min^{-1}. The CD intensities and UV absorbance were normalized and the melting temperatures were determined from the sigmoidal curve fit by using the Boltzmann function in the Origin 8.0 software. All experiments were triplicated and values are reported with estimated standard deviations.

UV/Vis absorption studies: UV/visible absorption experiments were carried out by using a Perkin–Elmer (Lambda Bio +) UV spec-

trophotometer. The sample were prepared as 25 μM ligand solutions in 100 mM KCl and 10 mM lithium cacodylate buffer, pH 7.2. The absorption spectra were measured in the range $\lambda = 200$–600 nm by using quartz cuvette of 1 mm path length. The ligand solution was titrated against an increasing concentration of RNA (0–4 μM) after 10 min of equilibration, and the experiment was terminated after reaching saturation. The normalized absorbance was plotted against the log of the concentration of RNA. The binding constant K_b was derived from the curve fitting by using the Hill-1 equation [Eq. (1)].

$$A_N = A_0 + (A_0 - A_s)\frac{[DNA]^n}{[K_d]^n + [DNA]^n} \quad (1)$$

Where A_0 is the absorbance of the ligand without RNA, A_S is the absorbance at saturation in the presence of RNA, n is the number of binding sites, and K_d is the dissociation constant ($1/K_b$).

Job plot analysis: Job plot analysis was performed by using the UV absorption study. The intensity of the absorbance maxima was monitored at $\lambda = 412$ nm. The total concentration of ligand and RNA was kept constant at 10 μM in 100 mM KCl and 10 mM lithium cacodylate buffer, pH 7.2. The RNA solution was pre-annealed under the same salt and buffer condition. All eleven samples were prepared by varying the mole fraction of the ligand (0–1) and incubated for 30 min at 20 °C. A plot of the normalized absorbance and the mole fraction of the ligand gave rise the binding stoichiometry.

5′-End radiolabeling: The 5′-end of the RNA (10 pmol) was labeled with γ^{32}-P ATP (10 μCi) by using T4 polynucleotide kinase (PNK, 5 U) in 1× PNK buffer (50 mM Tris·HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine and 0.1 mM ethylenediaminetetraacetic acid (EDTA)). The total volume of the reaction mixture was 10 μL. The reaction mixture was incubated at 37 °C for 1 h and the enzyme activity was stopped by heating at 70 °C for 3 min. Salts and excess of γ^{32}-P ATP were removed by using the QIAquick nucleotide removal kit with the protocol provided by the manufacturer. The full-length labeled RNA was purified by 20% denaturing polyacrylamide gel electrophoresis (7 M urea). The gel slice was crushed and soaked in a solution of TEN buffer (1 mM Tris·HCl, pH 8.0, 0.5 M EDTA, pH 8.0, 3 M NaCl). The RNA pellet was extracted and desalted by ethanol precipitation and washed two times with cold 70% ethanol. The pellet was air dried and dissolved in an appropriate amount of RNase-free water.

RNase T1 footprinting assay: The 5′-end radiolabeled RNA in the presence of 150 mM KCl/LiCl was folded by heating for 5 min at 90° C and then allowed to slow cooling to room temperature over a period of 3–4 h. The annealed labeled RNA was treated with 0.01 U of the RNase T1 enzyme over a period of 5 min at 37 °C. The reaction was stopped by adding four times the volume of stop solution (80% formamide) to the reaction mixture. The cleaved products were analyzed by 20% denaturing PAGE (7 M urea) by using 1× TBE (89 mM of each Tris and boric acid, 2 mM EDTA, pH 8.2) as running buffer. The gels were kept inside the cassette for 6 h and analyzed by autoradiography using Storm 825 phosphorimager.

Cloning: The plasmid psiCHECK-2 (Promega) was used as the reporter vector for cloning the G4-forming and mutant sequences. The 22-mer *PP2Acα*G4-forming sequence (Table S1) was inserted eleven bases upstream to the translation start through site by site-directed mutagenesis (SDM).[21] The SDM polymerase chain reaction (PCR) was performed in a total volume of 25 μL, containing 20 ng psiCHECK-2 as template DNA, 0.6 μM phosphorylated forward primer and reverse primer each (Table S1), 0.2 mM dNTPs, 0.5 U Kapa DNA polymerase, 1× Kapa HiFi buffer containing 2 mM $MgCl_2$ and 3% *v/v* DMSO. PCR conditions: 96 °C for 5 min., then 35 cycles of 94 °C for 1 min., 66 °C for 30 s, and 72 °C for 1.40 min. followed by 72 °C for 10 min. The PCR product was purified by using a standard Gene JET extraction kit. The purified product was ligated by using 0.5 U of T4 DNA ligase enzyme in 1× T4 DNA ligase buffer (40 mM Tris·HCl, 10 mM $MgCl_2$, 10 mM EDTA, 1.5 mM ATP, pH 7.8) for 4.5 h at 22 °C. The purified product was transformed into DH5α competent cells and the observed colony (single colony) was inoculated in 5 mL lysogeny broth (LB) medium containing a final concentration of 100 μg mL^{-1} ampicillin. The minipreparation of plasmid DNA was performed with all the transformed DH5α clones by using a Gene JET plasmid miniprep kit and verified with digestion by 10 U of EcoRI in 2× Tango buffer (66 mM Tris·acetate, pH 7.9, 20 mM Mg acetate, 132 mM K acetate, and 0.2 mg mL^{-1} BSA) for 5 h at 37 °C. Insertion of the correct sequence was confirmed by sequencing.

For the introduction of the 22-mer mutant and the 14- and 13-mer sequences into the psiCHECK2 plasmid restriction digestion followed by a ligation method was used.[22] The respective oligonucleotides (Table S1) were annealed at 90 °C for 5 min followed by slow cooling to room temperature by using 1× annealing buffer (5 mM Tris, pH 7.5, 15 mM NaCl, 0.1 mM EDTA, pH 8.0) prior to insertion at the Nhe1 restriction site upstream to the renilla luciferase translation start site. The plasmid was digested by using 10 U of Nhe1 at 37 °C with 1× Tango buffer (33 mM Tris·acetate, pH 7.9, 10 mM Mg acetate, 66 mM K acetate, 0.1 mg mL^{-1} bovine serum albumin (BSA)) at 37 °C for 12 h. The digested product was then purified by a Gene JET purification kit and ligated by using 2.5 U of T4 DNA ligase in 1× T4 DNA ligase buffer (40 mM Tris·HCl, 10 mM $MgCl_2$, 10 mM EDTA, 1.5 mM ATP, pH 7.8) for 5.5 h at 22 °C followed by incubation at 4 °C for 10 h. The engineered plasmids were transformed, individual clones were subsequently processed, and their sequences were confirmed as described above for the 22-mer G4 sequence.

Cell culture: HeLa (human cervical cancer) cells were maintained in eagle's minimum essential media (MEM) supplemented with 10% fetal bovine serum (FBS) and 1% antibiotic solution containing streptomycin and penicillin. Cells were cultured in a humidified atmosphere containing 5% of CO_2 at 37 °C.

Transfection: Cells were seeded at a density of 2×10^5 cells per well in 24-well plates 24 h before transfection. Transfection was carried out when the cells reached a confluency of about 60–70%. For each transfection reaction, the plasmid psiCHECK-2 (Promega) containing firefly and renilla luciferase was used as positive control. The cells were transfected with the corresponding plasmids (500 ng per well) by using 3.5 μL Lipofectamine 3000 in a serum-free medium according to the manufacturer's protocol and incubated at 37 °C for 5 h. The medium was then replaced with MEM supplemented with 20% FBS. After 12 h of transfection, the cells could grow in culture media and were harvested 48 h after transfection for dual luciferase reporter assay.

Dual luciferase reporter assay: The firefly and renilla luciferase activity were measured sequentially from the cell lysate on a Luminometer (Berthold Junior LB 9509) by using the dual luciferase assay system (Promega) according to the manufacturer's protocol. The results were expressed as a ratio of renilla to firefly luciferase activity. The relative activity of cloned plasmids *PP2Acα*G4 and mut *PP2Acα*G4 with respect to psiCHECK-2 was obtained by normaliz-

ing the activity data to the positive control psiCHECK-2. Data were collected from at least three independent experiments and presented with corresponding standard deviations.

Acknowledgements

This work was financially supported by grants from the Science and Engineering Research Board (SERB-DST)-Government of India (grant no. EMR/2016/003268) and from the Department of Biotechnology (DBT)-Government of India (grant no: 6242-P4/RGCB/PMD/DBT/PKPI/2015) to P.I.P. We are thankful to the Department of Biosciences and Bioengineering, IIT Bombay for providing access to tissue culture facility and central facility supported by IRCC-IIT Bombay for MALDI-MS facility. We are also thankful to Prof. Ruchi Anand for providing access to her laboratory facilities. Dr. Bharathwaj Sathyamoorthy (IISER, Bhopal) for NMR spectra of RNA, and Dr. K. V. Diveshkumar for synthesizing the ligand InEt2. Dr. S. Harikrishna, Dr. Ganesh R. Nawale, and Saaz Sakrikar are also acknowledged for their initial contributions toward this project. S.P.P. thanks the Council of Scientific and Industrial Research (CSIR), Government of India for a Ph.D. fellowship.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: FMRP · G-quadruplexes · PP2Ac-alpha · RNA · translation

[1] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402 – 5415.
[2] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2007**, *35*, 406 – 413.
[3] a) J. L. Huppert, *Chem. Soc. Rev.* **2008**, *37*, 1375 – 1384; b) P. Agarwala, S. Pandey, S. Maiti, *Org. Biomol. Chem.* **2015**, *13*, 5570 – 5585; c) J. Song, J.-P. Perreault, I. Topisirovic, S. Richard, *Translation* **2016**, *4*, e1244031; d) J. Kim, C. Cheong, P. B. Moore, *Nature* **1991**, *351*, 331 – 332.
[4] A. Cammas, S. Millevoi, *Nucleic Acids Res.* **2017**, *45*, 1584 – 1595.
[5] a) A. Bugaut, S. Balasubramanian, *Nucleic Acids Res.* **2012**, *40*, 4727 – 4741; b) M. M. Fay, S. M. Lyons, P. Ivanov, *J. Mol. Biol.* **2017**, *429*, 2127 – 2147; c) S. Millevoi, H. Moine, S. Vagner, *RNA* **2012**, *3*, 495 – 507.
[6] S. Kumari, A. Bugaut, J. L. Huppert, S. Balasubramanian, *Nat. Chem. Biol.* **2007**, *3*, 218 – 221.
[7] A. Arora, M. Dutkiewicz, V. Scaria, M. Hariharan, S. Maiti, J. Kurreck, *RNA* **2008**, *14*, 1290 – 1296.
[8] M. J. Morris, S. Basu, *Biochemistry* **2009**, *48*, 5313 – 5319.
[9] D. Gomez, A. Guédin, J.-L. Mergny, B. Salles, J.-F. Riou, M.-P. Teulade-Fichou, P. Calsou, *Nucleic Acids Res.* **2010**, *38*, 7187 – 7198.
[10] R. Shahid, A. Bugaut, S. Balasubramanian, *Biochemistry* **2010**, *49*, 8300 – 8306.
[11] H.-Y. Weng, H.-L. Huang, P.-P. Zhao, H. Zhou, L.-H. Qu, *RNA Biol.* **2012**, *9*, 1099 – 1109.
[12] S. Lammich, F. Kamp, J. Wagner, B. Nuscher, S. Zilow, A.-K. Ludwig, M. Willem, C. Haass, *J. Biol. Chem.* **2011**, *286*, 45063 – 45072.
[13] A. Arora, B. Suess, *RNA Biol.* **2011**, *8*, 802 – 805.
[14] a) A. Bugaut, R. Rodriguez, S. Kumari, S.-T. D. Hsu, S. Balasubramanian, *Org. Biomol. Chem.* **2010**, *8*, 2771 – 2776; b) K. Halder, E. Largy, M. Benzler, M.-P. Teulade-Fichou, J. S. Hartig, *ChemBioChem* **2011**, *12*, 1663 – 1668.
[15] a) J. U. Guo, D. P. Bartel, *Science* **2016**, *353*, 929; b) T. Serikawa, J. Eberle, J. Kurreck, *FEBS Lett.* **2017**, *591*, 3649 – 3659.
[16] a) M. Morris, Y. Negishi, C. Pazsint, J. Schonhoft, S. Basu, *J. Am. Chem. Soc.* **2010**, *132*, 17831 – 17839; b) S. Bonnal, C. Schaeffer, L. Creancier, S. Clamens, H. Moine, A. C. Prats, S. Vagner, *J. Biol. Chem.* **2003**, *278*, 39330 – 39336.
[17] a) M. Melko, B. Bardoni, *Biochimie* **2010**, *92*, 919 – 926; b) J. C. Darnell, K. B. Jensen, P. Jin, V. Brown, S. T. Warren, R. B. Darnell, *Cell* **2001**, *107*, 489 – 499; c) M. Bole, L. Menon, M.-R. Mihailescu, *Mol. BioSyst.* **2008**, *4*, 1212 – 1219; d) C. Schaeffer, B. Bardoni, J. L. Mandel, B. Ehresmann, C. Ehresmann, H. Moine, *EMBO J.* **2001**, *20*, 4803.
[18] V. Brown, P. Jin, S. Ceman, J. C. Darnell, W. T. O'Donnell, S. A. Tenenbaum, X. Jin, Y. Feng, K. D. Wilkinson, J. D. Keene, R. B. Darnell, S. T. Warren, *Cell* **2001**, *107*, 477 – 487.
[19] a) R. Lu, H. Wang, Z. Liang, L. Ku, W. T. O'Donnell, W. Li, S. T. Warren, Y. Feng, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15201 – 15206; b) M. Castets, C. L. Schaeffer, E. Bechara, A. Schenck, E. W. Khandjian, S. Luche, H. Moine, T. Rabilloud, J.-L. Mandel, B. Bardoni, *Hum. Mol. Genet.* **2005**, *14*, 835 – 844.
[20] a) K. V. Diveshkumar, S. Sakrikar, F. Rosu, S. Harikrishna, V. Gabelica, P. I. Pradeepkumar, *Biochemistry* **2016**, *55*, 3571 – 3585; b) V. Dhamodharan, S. Harikrishna, A. C. Bhasikuttan, P. I. Pradeepkumar, *ACS Chem. Biol.* **2015**, *10*, 821 – 833.
[21] K. Halder, M. Benzler, J. S. Hartig, *Methods* **2012**, *57*, 115 – 121.
[22] M. D. Horwich, P. D. Zamore, *Nat. Protoc.* **2008**, *3*, 1537 – 1549.
[23] D. M. Gray, J. D. Wen, C. W. Gray, R. Repges, C. Repges, G. Raabe, J. Fleischhauer, *Chirality* **2008**, *20*, 431 – 440.
[24] A. Bugaut, P. Murat, S. Balasubramanian, *J. Am. Chem. Soc.* **2012**, *134*, 19953 – 19956.
[25] A. Guédin, L. Lacroix, J. L. Mergny, *Methods Mol. Biol.* **2010**, *613*, 25 – 35.
[26] G. Knapp, *Methods Enzymol.* **1989**, *180*, 192 – 212.

Manuscript received: May 18, 2019
Accepted manuscript online: June 17, 2019
Version of record online: September 26, 2019

CHEMBIOCHEM

Supporting Information

Presence of Potential G-Quadruplex RNA-Forming Motifs at the 5′-UTR of *PP2Acα* mRNA Repress Translation

Sushree Prangya P. Pany,[a] Mahak Sapra,[b] Jitendar Sharma,[a] V. Dhamodharan,[a] Swati Patankar,[c] and P. I. Pradeepkumar*[a]

cbic_201900336_sm_miscellaneous_information.pdf

Electronic Supplementary Information (ESI)

Presence of potential G-quadruplex RNA forming motifs at the 5'-UTR of *PP2Acα* mRNA repress translation

Sushree Prangya P. Pany,[a] Mahak Sapra,[b] Jitendar Sharma,[a] V.Dhamodharan,[a] Swati Patankar,[c] and Pradeepkumar P. I.[a]*

TABLE OF CONTENTS

Schematic representation of the 5'-UTR of *PP2Acα* mRNA

GCUUCAGUUACCAGCCGGCUACGUCGCGCCUGCGCUUUGACCCCCAGUUUGC
GCCCCAACUCCGGUCGUGCGGCCGCCCGGGGAGGGCUCUGCAGUUGCGCAGC
UUGCUCCCCGGCCCUUUUCCCCUCCGCUCCCCGCCGCCUCCUGACGCCGGCGU
GACGUCACCACGCCCGGCGGCCGCCAUUACAGAGAGCCGAGCUCUGGAGCCU
CAGCGAGCGGAGGAGGAGGCGCAGCGGCCGACGGCCGAGUACUGCGGUGAG
AGCCAGCGGGCCAGCGCCAGCCUCAACAGCCGCCAGAAGUACACGAGGAACC
GGCGGCGGCGUGUGCGUGUAGGCCCGUUGCGGGCGGCGGCGCGGGAGCAGC
GCGGAGCGGCAGCCGGCUGGGGCGGGUGGCAUCAUG

Figure S1. Schematic representation of 5'-UTR of *PP2Acα* mRNA; Red: possible G4 forming sequences; Green: translation start codon.

Schematic representation of the psiCHECK-2 plasmid with an insertion site for *PP2Acα* sequence

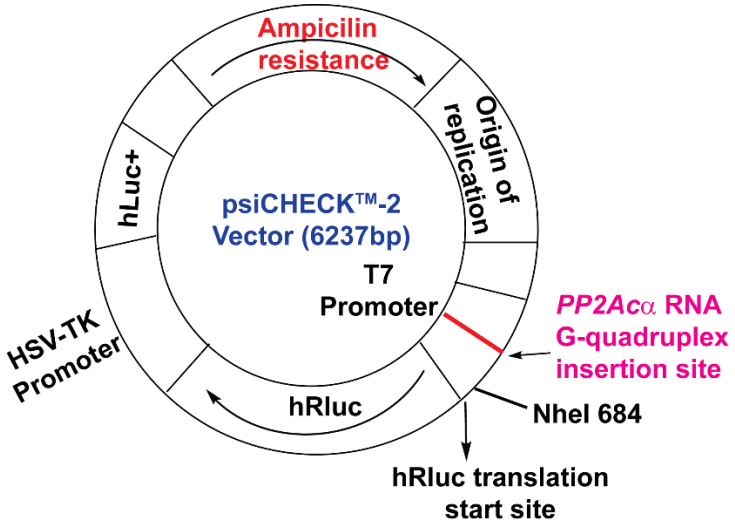

Figure S2. Schematic representation of the psiCHECK-2 reporter vector used for the cloning

Dual luciferase reporter assay with 13 mer, 14 mer and mutant G-rich sequences

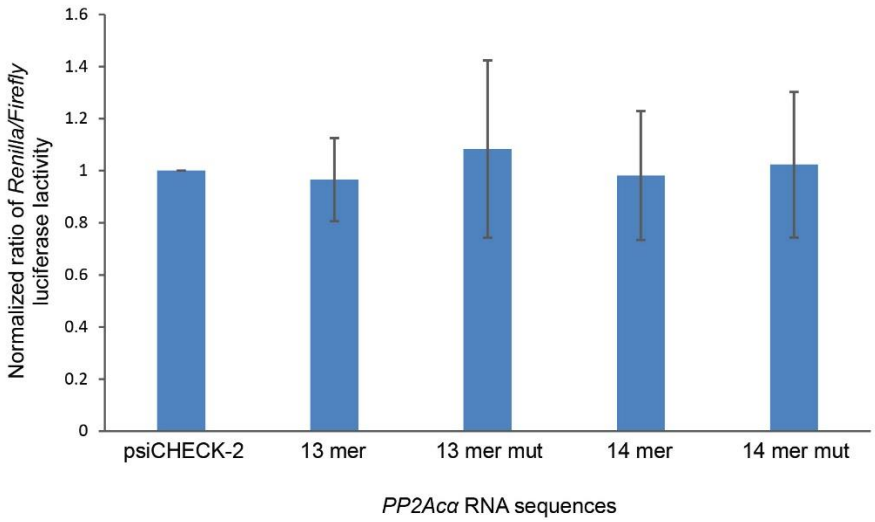

Figure S3. Normalized luciferase activity of plasmid containing *PP2Acα* 13 mer, 14 mer, and their corresponding mutant G4 sequences (Table S1) post 48 h of transfection in HeLa cells. *Firefly* luciferase activity was divided by *Renilla* luciferase activity, and the ratio was normalized to the positive control (psiCHECK-2). The error bars represent the standard deviation (±SD) from three independent experiments

CD spectra of the 13 and 14 mer *PP2Acα* G4 RNA with added metal ions

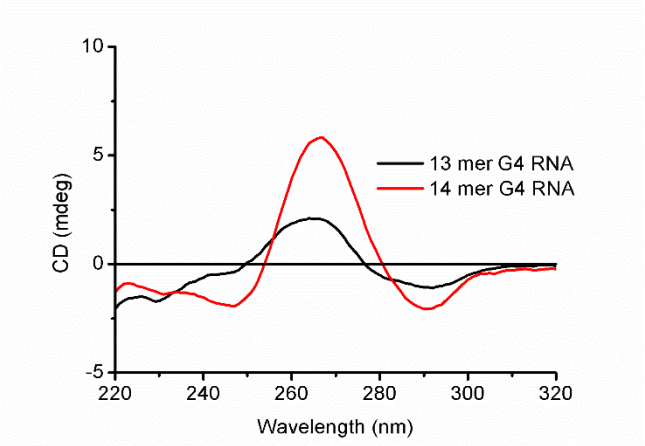

Figure S4. CD spectra of 13 mer and 14 mer *PP2Acα* G4 RNA (12.5 μM in 100 mM KCl, 10 mM lithium cacodylate buffer, pH 7.2). The RNA sequence is shown in Table S2.

CD melting curves of the 13 and 14 mer *PP2Acα* G4 RNA with varying strand concentration

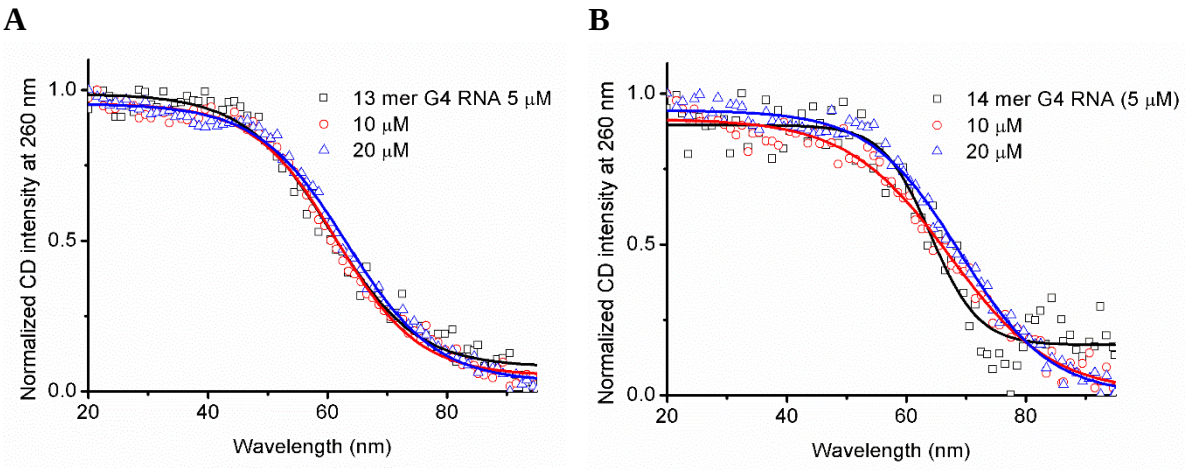

Figure S5. Normalized CD melting curve of 13 mer and 14 mer *PP2Acα* G4 RNA with different strand concentration (5, 10 and 20 μM in 10 mM KCl, 90 mM LiCl and 10 mM lithium cacodylate buffer, pH 7.2). A) 13 mer *PP2Acα* G4 RNA; B) 14 mer *PP2Acα* G4 RNA.

CD melting curves of the *PP2Acα* G4 RNA with the ligand Phen-Et

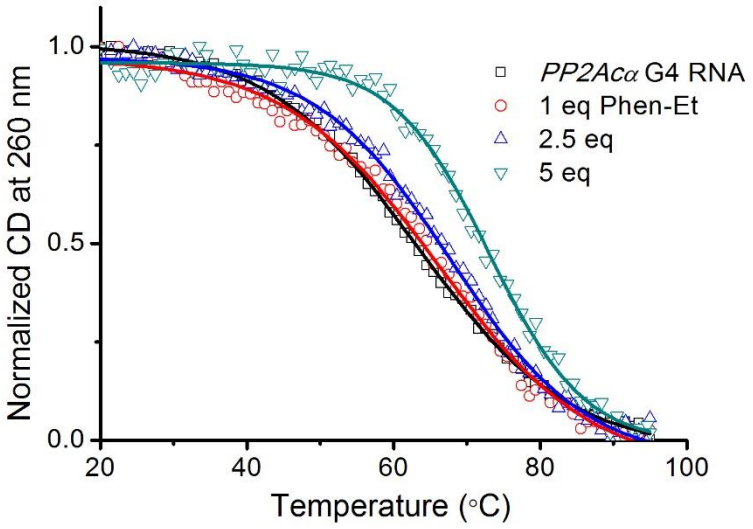

Figure S6. Normalized CD melting curve of *PP2Acα* G4 RNA (10 μM in 10 mM KCl, 90 mM LiCl and 10 mM lithium cacodylate buffer, pH 7.2) in the absence and in the presence of varying concentration (0-5 equivalent) of ligand **Phen-Et**.

CD melting curves of the *PP2Acα* G4 RNA sequence with varying strand concentration

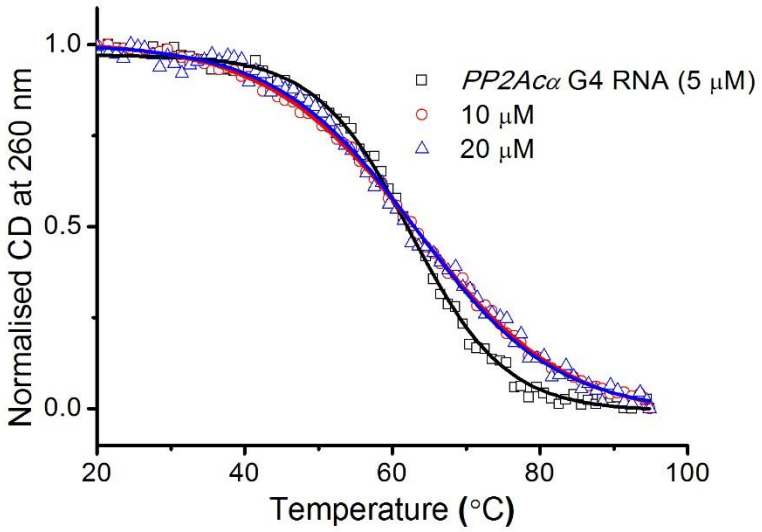

Figure S7. Normalized CD melting curve of 22mer *PP2Acα* G4 RNA with different strand concentration (5, 10, and 20 μM in 10 mM KCl, 90 mM LiCl and 10 mM lithium cacodylate buffer, pH 7.2).

CD melting curves of the mutant *PP2Acα* G4 RNA sequence

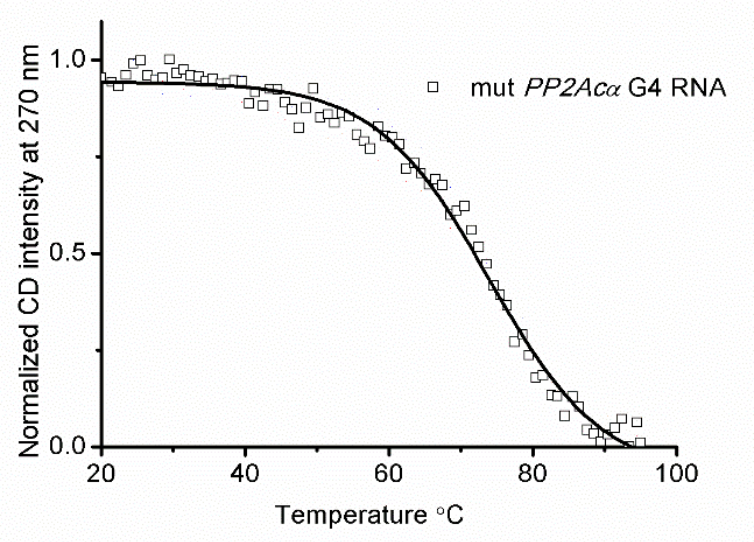

Figure S8. Normalized CD melting curve for mut *PP2Acα* RNA (10 μM in 10mM KCl and 90 mM LiCl and 10 mM lithium cacodylate, pH 7.2)

UV melting curves of the *PP2Acα* G4 and mutant RNA sequence

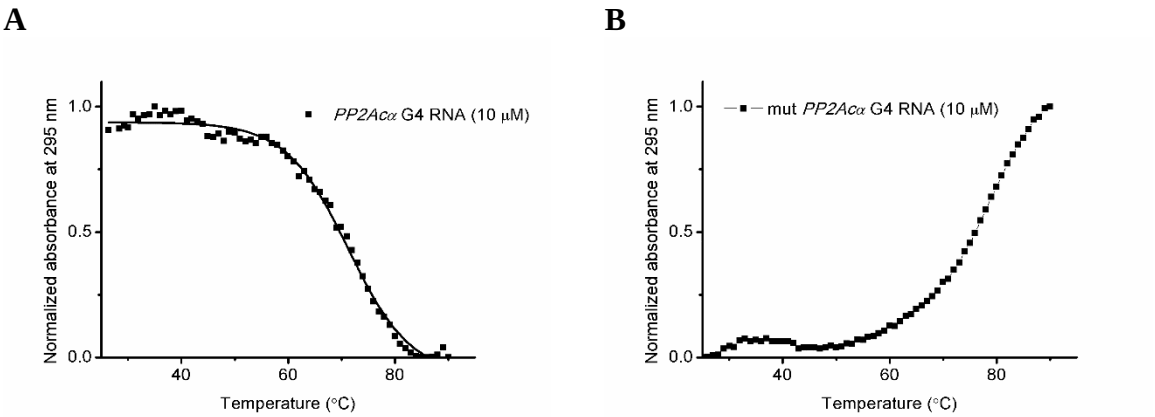

Figure S9. UV melting curve for *PP2Acα* G4 RNA and mutant sequence (10 μM in 100 mM KCl and 10 mM lithium cacodylate, pH 7.2) A) 22mer *PP2Acα* G4 RNA; B) mut *PP2Acα* RNA

Job plot analysis of the *PP2Acα* G4 RNA with the ligand InEt2

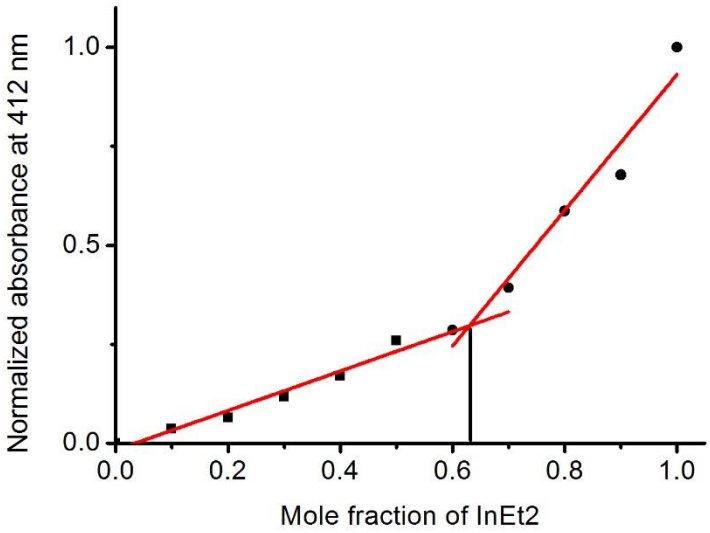

Figure S10. Job plot analysis of *PP2Acα* G4 RNA with **InEt2**. The molar concentration of the solution (ligand + RNA) remains constant (10 μM). Absorbance at 412 was measured for each data points. Point of the intersection at 0.62 indicates the 2:1 binding stoichiometry of **InEt2** with *PP2Acα* G4 RNA.

RNase T1 footprinting assay of mutant G4 RNA

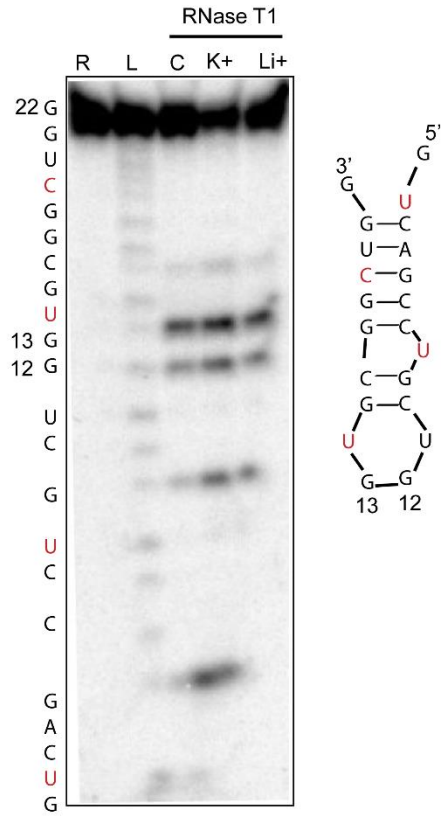

Figure S11. PAGE (20%, 7M urea) of RNaseT1 footprinting of mutant 22mer *PP2Acα* G4 RNA using different monovalent ions. R: labelled RNA (lane 1); L: alkaline hydrolysis ladder (lane 2); C: Control without added monovalent cation (lane 3); with 150 mM KCl (lane 4); with 150 mM LiCl (lane 5)

DNA oligonucleotide used for cloning

DNA	Sequences	Molar exaction coefficient (ε in L mol^{-1} cm^{-1})
22 mer forward primer	5'-GCTGGGGCGGGTGGAGCTAGCCACCATGGCTTC-3'	3.04 x 10^5
22 mer reverse primer	5'-CGGCTGCCGCTAGCTATAGTGAGTCGTATTAAGTA CTCTAGCC-3'	4.08 x 10^5
22 mer mutant sense	5'-CTAGTGTCAGCCTGCTGGTGCGGCTGGA-3'	2.88 x 10^5
22 mer mutant antisense	5'-CTAGTCCAGCCGCACCAGCAGGCTGACA-3'	2.96 x 10^5
14 mer sense	5'-CTAGTGGAACCGGCGGCGGA-3'	2.21 x 10^5
14 mer antisense	5'-CTAGTCCGCCGCCGGTTCCA-3'	1.90 x 10^5
14 mer mutant sense	5'-CTAGTGTAACCGTCGGCGGA-3'	2.15 x 10^5
14 mer mutant antisense	5'-CTAGTCCGCCGACGGTTACA-3'	2.06 x 10^5
13 mer sense	5'-CTAGTGGCGGCGGCGCGGA -3'	2.02 x 10^5
13 mer antisense	5'- CTAGTCCGCGCCGCCGCCA -3'	1.80 x 10^5
13 mer mutant sense	5'-CTAGTGTCGTCGGCGCGGA-3'	1.96 x 10^5
13 mer mutant antisense	5'-CTAGTCCGCGCCGACGACA-3'	1.96 x 10^5

Table S1. DNA sequences used for cloning by site-directed mutagenesis (SDM) or by insertion methods. G rich regions and their complementary bases are shown in red for oligos used for insertion methods. The blue color indicates the regions in the PCR primers, which contribute to the insertion of G rich sequence. The molar extinction coefficients were calculated by using the nearest neighbor method. (Biopolymers, 1970, 9, 1059-1077.)

RNA oligonucleotides used for various experiments

RNA	Sequences	Molar extinction coefficient (ε in L mol^{-1} cm^{-1})
13 mer	5'-GGCGGCGGCGCGG-3'	1.35 x 10^5
14 mer	5'-GGAACCGGCGGCGG-3'	1.54 x 10^5
22 mer	5'-GGCAGCCGGCUGGGGCGGGUGG-3'	2.36 x 10^5
22 mer mutant	5'-GUCAGCCUGCUGGUGCGGCUGG-3'	2.28 x 10^5

Table S2. *PP2Acα* G4 and mutant RNA sequences used in the biophysical and biochemical experiments. Mutations are underlined. The molar extinction coefficients were calculated using the nearest neighbor method.