Biochimica et Biophysica Acta 1840 (2014) 3503–3510

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Role of G-quadruplex located at 5′ end of mRNAs

Prachi Agarwala [a,b], Satyaprakash Pandey [a,b], Souvik Maiti [a,b,c,*]

[a] Proteomics and Structural Biology Unit, CSIR-Institute of Genomics and Integrative Biology, Mall Road, Delhi 110007, India
[b] Academy of Scientific and Innovative Research (AcSIR), Anusandhan Bhawan, 2 Rafi Marg, New Delhi 110001, India
[c] CSIR-National Chemical Laboratory, Dr. Homi Bhabha Road, Pune 411008, India

ARTICLE INFO

Article history:
Received 18 April 2014
Received in revised form 27 August 2014
Accepted 28 August 2014
Available online 16 September 2014

Keywords:
G-quadruplex
UTR
AKTIP
CTSB
FOXE3
APOA1BP

ABSTRACT

Background: Secondary structures in 5′ UTR of mRNAs play a critical role in regulating protein synthesis. Though studies have indicated the role of secondary structure G-quadruplex in translational regulation, position-specific effect of G-quadruplex in naturally occurring mRNAs is still not understood. As a pre-initiation complex recognises 5′ cap of the mRNA and scans along the untranslated region (UTR) before initiating translation, the presence of G-quadruplex in 5′ region may have a significant contribution in regulating translation. Here, we investigate the role of G-quadruplex located at the 5′ end of an mRNA.
Methods: Biophysical characterisation of putative G-quadruplexes was performed using UV and CD spectroscopy. Functional implication of G-quadruplex in the context of their location was assessed in cellulo using qRT-PCR and dual luciferase assay system.
Results: PG4 sequences in 5′ UTR of AKT interacting protein (AKTIP), cathepsin B (CTSB) and forkhead box E3 (FOXE3) mRNAs form G-quadruplex whereas it is unable to form G-quadruplex in apolipoprotein A-I binding protein (APOA1BP). Our results demonstrated diverse roles of G-quadruplex located at 5′ end of mRNAs. Though G-quadruplex in AKTIP and CTSB mRNA act as inhibitory modules, it activates translation in FOXE3 mRNA.
Conclusions: Our works suggests that G-quadruplex present at the 5′ terminal of an mRNA behaves differently in a different gene context. It can activate or inhibit gene expression.
General significance: This study demonstrated that it is difficult to predict the role of G-quadruplex on the basis of its position in 5′ UTR. The neighbouring nucleotide sequence, the intracellular milieu and the interacting partners might render diverse functions to this secondary structure.

1. Introduction

The level of mRNA in a cell may not necessarily correlate with the corresponding protein level. Several mechanisms regulate and maintain precise levels of proteins in a cell [1]. A growing body of evidence has revealed that regulation of gene expression at translational level is important for the homeostasis and its dysfunction is associated with the pathophysiology of many diseases [1,2]. The translation regulating ability of mRNAs primarily resides in its UTRs (Untranslated Regions) located at 5′ and 3′ ends [3]. Molecular profiling of many diseases has shown that UTRs of mRNAs play a pivotal role in disease progression and susceptibility. Anomaly in this region is responsible for manifestations of several diseases such as Breast Cancer, Hereditary hyperferritinemia cataract syndrome (HHCS) and Alzheimer's disease [2,4]. Many salient features of UTRs — upstream open reading frames, internal ribosome entry sites, GC-richness and structuredness influence the translation rates of mRNAs [2]. Additionally, various secondary structures such as hairpins, stem-loops, G-quadruplexes in UTR influence the biological function of the gene [2,5]. The information content of G-quadruplex in UTR of mRNAs is still poorly understood.

G-quadruplex is a four stranded secondary structure found in DNA and RNA. It consists of stacks of G-quartet, a square planar arrangement formed by a contiguous array of Hoogsteen-bonded guanine residues. G-quadruplex is stabilised in the presence of monovalent cations. After years of debate, the existence of RNA G-quadruplexes in mammalian cells has been visually demonstrated [6]. RNA G-quadruplex is involved in a myriad of biological processes ranging from telomere homeostasis, mRNA localisation, 3′ end processing and alternative splicing to translational regulation [7–10]. 5′ UTR G-quadruplex is implicated to play a regulatory role in translation [11–20].

Studies on G-quadruplex in 5′ UTR of NRAS, Zic-1, Bcl-2, TRF2, MT3-MMP, estrogen receptor and ADAM1 cellular mRNAs demonstrate their translational repressive activity [11–14,16,17,19,21]. Beaudoin et al. indicated the inhibitory role of G-quadruplex in several cellular mRNAs

Abbreviations: AKTIP, AKT interacting protein; CTSB, cathepsin B; FOXE3, forkhead box E3; APOA1BP, apolipoprotein A-I binding protein; UTR, untranslated region; CD, circular dichroism; TDS, thermal difference spectrum; T_m, melting temperature; PG4, putative G-quadruplex; Nts, nucleotides; FP, forward primer; RP, reverse primer

* Corresponding author at: Proteomics and Structural Biology Unit, Institute of Genomics and Integrative Biology, CSIR, Mall Road, Delhi, India. Tel.: +91 11 27666156; fax: +91 11 27667437.
E-mail address: souvik@igib.res.in (S. Maiti).

http://dx.doi.org/10.1016/j.bbagen.2014.08.017
0304-4165/

[15]. However, Basu and group were the first to report that 5′ UTR G-quadruplex is required for cap independent translation initiation in VEGF [18]. They showed that G-quadruplex disruption in VEGF leads to the reduction in gene activity, suggesting its translational enhancer role. Recently, our group established the activating role of G-quadruplex in modulating TGFβ2 gene expression, providing further evidences of G-quadruplexes' ability to augment translation [20]. Given the contrasting roles of G-quadruplexes as either activators or repressor motifs in UTRs, it would be appropriate to infer the role of this secondary structure considering its position in the UTR.

Studies examining consequences of artificially shifting the position of G-quadruplex sequence in the 5′ UTR have been made in cellulo [22,23]. However, these studies did not explore the positional effect of G-quadruplex in naturally occurring mRNAs. The composition of bases surrounding the G-quadruplex forming sequence, GC richness, number of G-quartets and loop composition may influence the modularity of G-quadruplex on gene expression [24–26]. It would be more applicable to study the role of UTRs harbouring naturally occurring G-quadruplex at the same given position. This approach would provide an insight about the positional role of G-quadruplex in different UTRs.

With this objective, we studied the function of a G-quadruplex located at the 5′ end of an mRNA. As pre-initiation complex binds 5′ cap region of an mRNA, we chose to study G-quadruplex situated at the first nucleotide position of mRNA in view to answer the following questions: Does G-quadruplex's presence at the beginning of an mRNA affect the translational efficiency of ribosome? If yes, then to what extent this secondary structure modulates the translational efficiency? Is the position-specific effect of G-quadruplex similar for all mRNAs harbouring them?

Our studies indicated that G-quadruplex present at the beginning of an mRNA acts differently in a different gene context. It can increase or suppress gene expression. A predefined function cannot be assigned to G-quadruplex on the basis of its position in an mRNA. The neighbouring nucleotide sequence, the intracellular environment and the interacting partners might affect the function of G-quadruplex in cell, rendering diverse functions to this secondary structure.

2. Materials and methods

2.1. Oligonucleotides

For biophysical studies, HPLC purified RNA oligonucleotides were ordered from Sigma-Aldrich (Table 2 and Table S1). The concentration of oligonucleotides was measured via a UV spectrophotometer and was calculated using their molar extinct coefficient values.

DNA oligonucleotides for cloning purpose were also purchased from Sigma.

2.2. UV spectroscopy

2 μM RNA samples were prepared in 10 mM sodium cacodylate (pH 7.4) containing 1 mM KCl. UV-melting studies were carried out using a Cary 100 UV–visible spectrophotometer (Varian) equipped with a Peltier temperature controller. The samples were heated at 95 °C and subsequently slowly cooled till 20 °C at the temperature gradient of 0.2 °C/min. During melting and annealing steps, the absorbance values of oligonucleotides were recorded at 295 nm of wavelength. UV melting experiments were performed using RNA samples in a concentration range of 1 μM to 10 μM prepared in 10 mM sodium cacodylate buffer (pH 7.4) containing 1 mM KCl. The melting curves were analysed using Origin 7.0. The absorbance values obtained at 295 nm were further analysed for calculating thermodynamic parameters (within 10% error). This method involved the contribution from pre- and post-transition baselines and thermodynamic data was obtained using equations described previously [27].

$$Au = bu + (mu * T)$$

$$Al = bl + (ml * T)$$

$$Keq = \alpha/(1-\alpha)$$

$$A(T) = \alpha(Au - Al) + Al$$

$$Keq = \exp(\Delta G^{\circ}/RT) = \exp((\Delta H^{\circ}/RT) + (\Delta S^{\circ}/R)).$$

Au andAl are linear equations for the upper and lower baselines, respectively, where bu and bl describes fitted parameters for the intercepts for the upper and lower baselines with mu and ml as their respective slopes. Keq indicates the equilibrium constant for the unstructured–structured transition for an intramolecular system and α stands for the folded fraction. A(T) being the dependent variable is an experimentally determined absorbance at each temperature (T). These equations were used to calculate van't Hoff enthalpy (ΔH_{vH}) and entropy (ΔS_{vH}).

2.3. Circular dichroism

RNA samples at 2 μM concentration were heated at 95 °C for 5 min and slowly cooled down to 20 °C. CD experiments were carried out in a Jasco J-810 spectropolarimeter (Jasco Hachioji, Tokyo, Japan) equipped with a Peltier temperature controller and scans were taken

Table 2
RNA oligonucleotides used for biophysical studies.

Oligonucleotide name	Sequence
AKTIP	GGGGUGGGGCGGGGCGGGG
CTSB	GGGGCGGGGCCGGGAGGG
FOXE3	GGGAAGGGGUGAGUCUGGGUCUGGG
APOA1BP	GGGCCGGGCCGGGCCGGGGG

Table 1
Details of the genes used for the study.

Name	Accession number	Putative G-quadruplex sequence	Function
AKTIP(AKT interacting protein)	NM_001012398.1	GGGGUGGGGCGGGGCGGGG	Interacts with protein kinase B (PKB)/Akt and modulates its activity by enhancing the phosphorylation of PKB's regulatory sites.
CTSB (cathepsin B)	NM_001908.3	GGGGCGGGGCCGGGAGGG	Lysosomal cysteine proteinase is involved in the proteolytic processing of amyloid precursor protein (APP).
FOXE3 (forkhead box E3)	NM_012186.2	GGGAAGGGGUGAGUCUGGGUCUGGG	Functions as a lens-specific transcription factor and plays an important role in vertebrate lens formation.
APOA1BP (apolipoprotein A-I binding protein)	NM_144772.2	GGGCCGGGCCGGGCCGGGGG	Interacts with apolipoprotein A-I (apoA-I), the major apolipoprotein of high-density lipoproteins (HDLs).

In all these genes, putative G-quadruplex sequence is located at the start of 5′ UTR i.e. from the first nucleotide.

at 200–350 nm wavelength at 20 °C. After the subtraction of buffer alone, CD scans were performed in triplicates and CD spectrum obtained is the resultant average. All samples were prepared in 10 mM sodium cacodylate buffer (pH 7.4) containing 1 mM KCl.

2.4. Thermal difference spectrum

Thermal Difference Spectrum (TDS) was obtained by recording the absorbance of RNA oligonucleotides at 20 °C and 95 °C at wavelengths of 200–350 nm and then subtracting their absorbance values. All samples were prepared in 10 mM sodium cacodylate buffer (pH 7.4) containing 1 mM KCl.

2.5. Cloning

The dual luciferase vector psiCHECK2 was used for the cloning purpose. The sequences of interest were cloned at the Nhe I restriction site located at the 5′ end of Renilla luciferase gene of psiCHECK2 vector (Fig. 1). Briefly, the primers against 212 bp AKTIP 5′ UTR, 163 bp CTSB 5′ UTR, 245 bp FOXE3 5′ UTR and 39 bp APOA1BP 5′ UTR were designed according to the GenBank sequences NM_001012398.1, NM_001908.3, NM_012186.2 and NM_144772.2, respectively. The entire 5′ UTR of AKTIP, CTSB and FOXE3 were PCR amplified from MCF-7 cells cDNA using Phusion DNA polymerase (Thermo Scientific) and thereby gel eluted using the gel extraction kit (Qiagen). Both the vector and the amplicon were digested by Nhe I restriction enzyme (NEB) and again purified by means of PCR purification kit (Qiagen). The Nhe I digested vector was dephosphorylated using Calf intestinal dephosphatase (NEB) in order to prevent self ligation. The dephosphorylated vector and the amplicon were ligated at 22 °C for 16 h using T4 DNA ligase (Fermentas). The clones were digested with Nhe I, gel run and checked for insert release. The positive clones were then confirmed by sequencing.

Site directed mutagenesis at specific positions in the 5′ UTR of AKTIP, CTSB, FOXE3 (namely pAKTIP mUTR, pAKTIP mUTR1, pCTSB mUTR, pFOXE3 mUTR) were performed using Kappa Ready Mix (Saflabs). Briefly, wild type constructs (pAKTIP UTR, pCTSB UTR, pFOXE3 UTR) were PCR amplified with primers containing mutations at the annealing temperature of 58 °C. Then wild type vector backbone was digested with Dpn 1 at 37 °C for 6 h. These PCR products were used to transform DH5α *Escherichia coli* cells. The positive clones obtained were confirmed by sequencing.

In order to obtain APOA1BP plasmid constructs (APOA1BP PG4, APOA1BP mPG4, APOA1BP UTR, APOA1BP mUTR), both the forward and reverse strands for each construct were annealed at 37 °C for 1 h. Then the annealed products were phosphorylated by T4 polynucleotide kinase. These phosphorylated products were ligated with dephosphorylated psiCHECK2 vector using T4 DNA ligase. The positive clones were confirmed by sequencing.

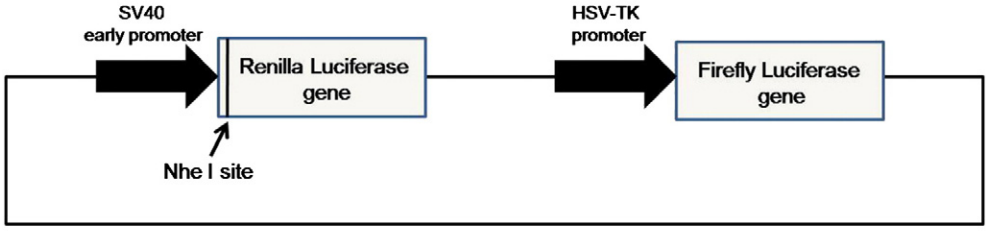
2.6. Cell culture

HEK 293T cell line was maintained at the humidified atmosphere containing 5% CO_2 incubator at 37 °C. The cells were grown in high glucose DMEM supplemented with 10% fetal bovine serum and antibiotic–antimycotic (Gibco).

2.7. Dual luciferase assay

The plasmid constructs (pAKTIP UTR, pAKTIP mUTR, pAKTIP mUTR1, pCTSB UTR, pCTSB mUTR, pFOXE3 UTR, pFOXE3 mUTR, APOA1BP PG4,

A



B

pAKTIP UTR	212 nts long 5′ UTR of AKTIP
pAKTIP mUTR1	212 nts long 5′ UTR of AKTIP with mutated PG4 AAGGTGGGGCGGGGCGGAA
pAKTIP mUTR	212 nts long 5′ UTR of AKTIP with mutated PG4 AAAGTGGGGCGGGGCGAAA
pCTSB UTR	163 nts long 5′ UTR of CTSB
pCTSB mUTR	163 nts long 5′ UTR of CTSB with mutated PG4 AAAGCGGGGCCGGGAGAA
pFOXE3 UTR	245 nts long 5′ UTR of FOXE3
pFOXE3 mUTR	245 nts 5′ UTR of FOXE3 with mutated PG4 GAGAAGAAGTGAGTCTGAGTCTGAG
pAPOA1BP UTR	39nts long 5′ UTR of APOA1BP
pAPOA1BP mUTR	39 nts 5′ UTR of APOA1BP with mutated PG4 GAGCCGAGCCGAGCCGGGGG
pAPOA1BP PG4	GGGCCGGGCCGGGCCGGGGG
pAPOA1BP mPG4	GAGCCGAGCCGAGCCGGGGG

Fig. 1. (A) Schematic representation of vector psiCHECK2 used for cloning and (B) details of the constructs generated. In each construct, sequence of interest was cloned at the Nhe I restriction site located at the 5′ end of Renilla luciferase gene of psiCHECK2 vector.

APOA1BP mPG4, APOA1BP UTR, APOA1BP mUTR) were transfected into HEK 293T cell line in 24-well plate using Lipofectamine 2000 (Invitrogen). The dual luciferase assay was performed after 48 h of incubation. Following treatment of the transfected cells with passive lysis buffer, Renilla and firefly luciferase readings of cell lysates were obtained using Dual Luciferase Assay kit (Promega) and TECAN plate reader. The ratio of Renilla luciferase activity to the firefly luciferase activity was subsequently calculated and normalised against the corresponding mutant plasmid readings.

2.8. Quantitative real time PCR

The abovementioned plasmids were transfected into HEK 293T cell line in 12-well plate using Lipofectamine 2000 reagent. The Renilla and firefly luciferase mRNA levels were quantified used Real Time PCR. Briefly, the total RNA of the transfected cells was isolated using a TRIzol reagent. Subsequently, cDNA was prepared using MMuLV Reverse Transcriptase enzyme and Oligo dT primers (Fermentas). The primers for Renilla and firefly luciferase genes were then used to amplify the transcripts of these genes using SYBR green master mix (Genecopoiea) and Roche detection system (Table S2). The luciferase mRNA levels were normalised against mutant plasmid readings using Pfaffl method [28].

3. Results

Using Quadfinder algorithm, we performed comprehensive in silico survey of a UTR database and obtained 10,788 hits with putative G-quadruplex sequence (PG4) in their 5′ UTR [29]. Based on the position of PG4, we sorted the mRNAs and identified 80 candidates (0.74%) having PG4 just at the beginning of the mRNAs (Supporting information). Apart from this, we also identified 685 (6.3%) and 3232 (30%) candidates possessing the start of PG4 within 10 nucleotides and 50 nucleotides from the 5′ end of mRNA, respectively. For our study, we chose mRNAs with small UTR size and PG4 present at their beginning (Table 1).

3.1. 3.1.G-quadruplex forming sequence in AKTIP RNA, CTSB RNA and FOXE3 RNA forms stable G-quadruplex in vitro

UV melting of AKTIP RNA, CTSB RNA and FOXE3 RNA was performed. As shown in Fig. 2A, UV spectra of AKTIP RNA, CTSB RNA and FOXE3 RNA displays a hypochromic shift with increasing temperature at 295 nm wavelength, a characteristic signature of G-quadruplex [30]. Thermal melting was performed in 10 mM sodium cacodylate buffer (pH 7.4) containing 1 mM KCl. UV spectra in 1 mM KCl containing buffer was chosen as the representative figure because all RNAs' UV melting profiles show the upper and lower baselines at this salt concentration. These baselines are required for proper estimation of the melting temperature and thermodynamic parameters. AKTIP RNA, CTSB RNA and FOXE3 RNA show the T_m of 79 ± 1 °C, 68.3 ± 1 °C and 40.5 ± 1 °C, respectively at 1 mM KCl salt concentration. AKTIP RNA is highly stable with T_m of 79 ± 1 °C at 1 mM KCl salt concentration. Even in the absence of salt, AKTIP RNA shows a T_m of 65.4 °C in 10 mM sodium cacodylate buffer (Fig. S1). All the RNAs showed greater stability with increasing salt concentration (data not shown). The thermodynamic parameters calculated from the UV melting profiles of AKTIP, CTSB, FOXE3 RNA are given in Table 3. Though formation of G-quadruplex in these sequences is entropically unfavourable, it is compensated by highly favourable enthalpy; resulting in the formation of a thermodynamically favourable secondary structure. The thermodynamic parameter free energy change (ΔG) of AKTIP (-11.68 kcal/mol) is much higher in magnitude than that of CTSB (-5.47 kcal/mol) and FOXE3 (-1.91 kcal/mol) at 1 mM KCl salt concentration, confirming the higher stability of AKTIP RNA G-quadruplex than CTSB and FOXE3 RNA G-quadruplex.

In order to ascertain the number of molecules involved in G-quadruplex formation, UV melting was performed with different strand concentrations of AKTIP, CTSB and FOXE3 RNA in 10 mM sodium

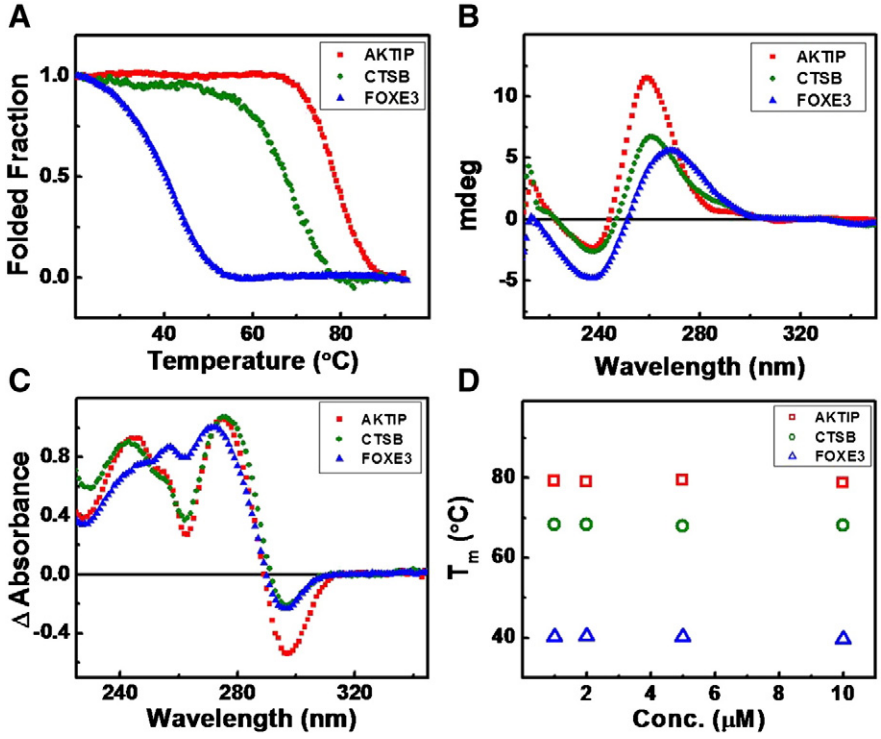


Fig. 2. Biophysical characterisation of putative RNA G-quadruplex sequences of AKTIP, CTSB and FOXE3. (A) UV melting of 2 μM AKTIP (red), 2 μM CTSB (dark green) and 2 μM FOXE3 (blue). (B) Circular dichroism spectra of 2 μM AKTIP (red), 2 μM CTSB (dark green) and 2 μM FOXE3 (blue). (C) Thermal difference spectra of 2 μM AKTIP (red), 2 μM CTSB (dark green) and 2 μM FOXE3 (blue). (D) T_m of AKTIP (red), CTSB (dark green) and FOXE3 (blue) at different strand concentrations. All experiments were performed in 10 mM sodium cacodylate buffer (pH 7.4) containing 1 mM KCl.

Table 3
T_m values and thermodynamic parameters calculated from the UV melting of RNA oligonucleotides.

Oligonucleotide name	T_m (°C)[a]	ΔH (kcal/mol)	ΔS (cal mol^{-1} deg^{-1})	ΔG (kcal/mol)[b]
AKTIP	79	−76.28	−216.78	−11.68
CTSB	68.3	−44.17	−129.87	−5.47
FOXE3	40.5	−39.37	−125.71	−1.91

All thermodynamics parameters calculated are within 10% error.
Thermodynamic parameters of mutated sequences AKTIP mut, CTSB mut and FOXE3 mut cannot be determined as they did not give hypochromic UV melting profile.
[a] T_m values reported with associated error of ±1 °C.
[b] ΔG is calculated at 25 °C.
The experiment was performed in 10 mM sodium cacodylate buffer (pH 7.4) containing 1 mM KCl.

cacodylate (pH 7.4) containing 1 mM KCl (Fig. 2D). In each of these oligonucleotides, a plot of T_m (in °C) versus concentration (in μM) clearly reveals similar T_m of various strand concentrations indicating the formation of intramolecular G-quadruplexes.

In contrary, the mutated version of these sequences (AKTIP mut, CTSB mut and FOXE3 mut) did not display any hypochromic UV shift with increasing temperature, indicating their inability to form G-quadruplex (Fig. S2A).

To corroborate the UV studies, we performed CD of the same RNA oligonucleotides. Circular Dichroism (CD) is the standard technique for studying the secondary structure of proteins and nucleic acids. As shown in Fig. 2B, 2 μM AKTIP RNA and 2 μM CTSB RNA in 10 mM Na cacodylate buffer (pH 7.4) containing 1 mM KCl salt gives a positive peak at around 261 nm of wavelength and a negative peak at 235 nm wavelength. This CD spectrum is the characteristic signature of parallel G-quadruplex. The CD spectrum of FOXE3 RNA gave a positive signal at a slightly higher wavelength which may be due to its comparatively lower stability than AKTIP and CTSB RNA in 1 mM KCl containing buffer. However, the CD spectrum of FOXE3 RNA at 10 mM KCl (Fig. S3) shows a similar positive signal at 262 nm wavelength and a negative signal at 241 nm wavelength indicating the formation of parallel G-quadruplex. Further, the intensity of positive peaks of these three RNAs indicates the stability of G-quadruplex in the order of AKTIP RNA > CTSB RNA > FOXE3 RNA, which again correlates well with the UV melting results of these RNAs. On the other hand, AKTIP mut, CTSB mut and FOXE3 mut showed a positive peak of very less intensity at around 264 nm, suggesting that they are unable to form G-quadruplex (Fig. S2B).

TDS is a spectrum obtained by subtracting the absorbance values of a structure in unfolded and folded states [31]. Absorbances of AKTIP RNA, CTSB RNA and FOXE3 RNA were recorded at 20 °C (folded) and 85 °C (unfolded). The TDS spectra of AKTIP RNA, CTSB RNA and FOXE3 RNA showed positive peaks at 275 nm and 244 nm, 276 nm and 243 nm, 273 nm and 257 nm, respectively and all displayed a negative peak at around 297 nm of wavelength (Fig. 2C). These two positive peaks and a negative signal at 297 nm of wavelength are the characteristic signature of G-quadruplex. Again, this G-quadruplex signature was not demonstrated by mutated oligonucleotides AKTIP mut, CTSB mut and FOXE3 mut upon performing TDS (Fig. S2C).

3.2. Putative G-quadruplex sequence of APOA1BP RNA is unable to form G-quadruplex in vitro

Conversely, the UV melting profile of APOA1BP RNA did not show a hypochromic shift at 295 nm wavelength suggesting its inability to form G-quadruplex (Fig. S4A). Even performing UV melting of APOA1BP in the presence of 20% PEG 200 (polyethylene glycol) and 40% PEG 200, no hypochromic UV profile was obtained indicating that APOA1BP cannot form G-quadruplex in the presence of molecular crowding agents such as PEG (Fig. S5). APOA1BP mut UV profile also displayed the same UV profile (Fig. S2A).

Similarly, the CD spectrum of APOA1BP (Fig. S4B) shows a positive peak at 261 nm wavelength and negative peaks at 232 nm and 210 nm of wavelengths. A negative peak at 210 nm suggests the formation of RNA duplex by this sequence. Again, APOA1BP mut oligonucleotide CD spectrum demonstrated its inability to form G-quadruplex (Fig. S2B).

Although the APOA1BP RNA shows two positive peaks at 278 nm and 247 nm in TDS, it did not reveal a negative peak at 297 nm wavelength which is an essential signal for showing G-quadruplex formation (Fig S4C). Also, the TDS of APOA1BP mut showed a positive peak at 263 nm, suggesting its inability to form G-quadruplex (Fig. S2C). Altogether, APOA1BP putative G-quadruplex sequence is unable to form G-quadruplex in vitro.

3.3. G-quadruplex executes diverse regulatory functions in 5′ UTR of mRNAs

In order to assess the role of G-quadruplex in the 5′ UTR of the mRNA, we cloned the 5′ UTR of four genes AKTIP, CTSB, FOXE3 and APOA1BP in the Nhe I cloning site of psiCHECK2 vector (pAKTIP UTR, pCTSB UTR, pFOXE3 UTR, pAPOA1BP UTR). These plasmids and their mutants (pAKTIP mUTR, pAKTIP mUTR1, pCTSB mUTR, pFOXE3 mUTR, pAPOA1BP mUTR where the putative G-quadruplex sequence is mutated) were transfected into HEK-293T cell line and dual luciferase assay was performed after 48 h of incubation. The presence of wtPG4 sequence in pAKTIP UTR and pCTSB UTR decreased the reporter gene expression to 80% and 70% respectively (Fig. 3). Further, substitution of guanine residues to adenines at specific positions in PG4 sequence in AKTIP 5′ UTR (pAKTIP mUTR1) so as to disrupt G4 G-quadruplex and favour G2 G-quadruplex formation, did not lead to any changes in gene expression between pAKTIP mUTR and pAKTIP mUTR1. This shows that G4 G-quadruplex is much more potent than G2 G-quadruplex in executing its gene repressing function.

Interestingly, Renilla luciferase activity normalised with firefly luciferase activity increased significantly to 160% in the presence of PG4 in FOXE3 5′ UTR suggesting the activating role of FOXE3 5′ UTR G-quadruplex.

In order to assess whether the gene regulation is occurring at transcriptional or translational level, we quantified the mRNA levels of reporter genes in the transfected cells using real time PCR. A similar transcriptional profile was obtained for all the abovementioned constructs (Fig. 4). No significant changes in the mRNA level of reporter genes suggest that the role of 5′ UTR G-quadruplex in all the studied genes is at the translational level and not at the transcriptional level.

3.4. Putative G-quadruplex sequence does not modulate gene expression when present in 5′ UTR of APOA1BPS

However, potential G-quadruplex in 5′ UTR of APOA1BP demonstrated no significant changes in reporter gene expression levels in the presence and absence of PG4 (Fig. S6). Thus, to see whether PG4 sequence alone brings any changes in reporter gene expression, we cloned the APOA1BP PG4 and its mutant mPG4 into the psiCHECK-2 (pAPOA1BP PG4 and pAPOA1BP mPG4) and performed dual luciferase assay. Again, there was no significant difference in the reporter gene expression in these constructs indicating the inability of PG4 to perform any regulating function.

4. Discussion and conclusion

The untranslated regions of mRNA dictate the extent of gene expression post-transcriptionally. GC richness and length of 5′ UTR modulate the scanning rate of ribosome and alter the time required to reach the initiation codon [1]. The presence of the secondary structures regulate efficacy of ribosome to scan the mRNA [3,32]. The secondary structure of G-quadruplex with its inherent dynamicity and flexibility has been studied extensively in gene regulation, however, its role in 5′ UTR is in

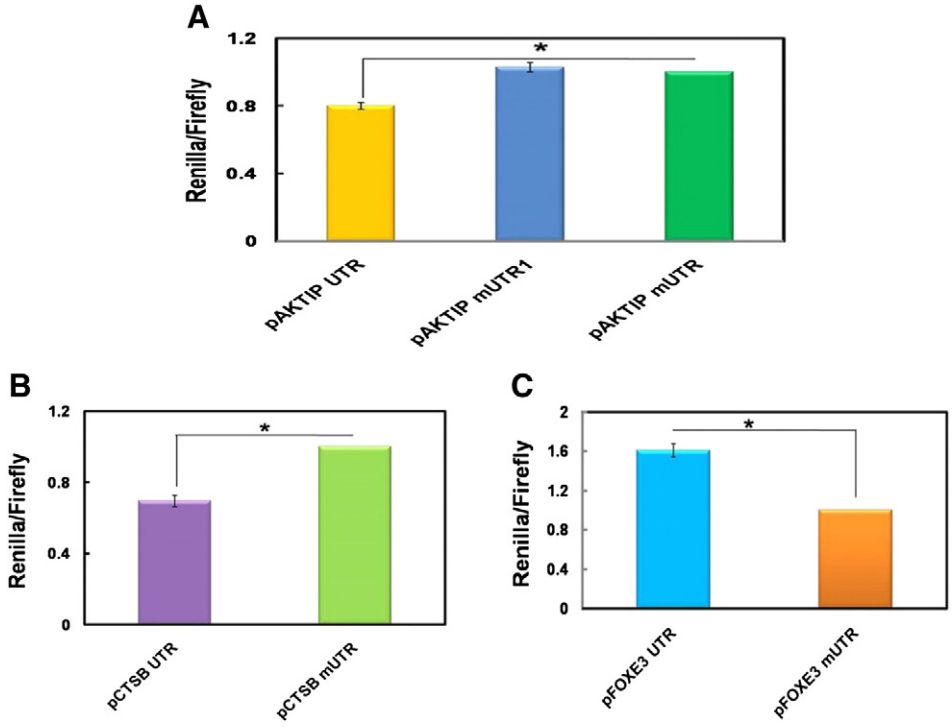


Fig. 3. Modulation of reporter gene expression in the presence of G-quadruplex motif. HEK 293T cells were transfected with various constructs and dual luciferase assay was performed. (A) G-quadruplex in AKTIP 5′UTR decreases gene expression. (B) G-quadruplex in CTSB 5′UTR suppresses gene expression. (C) G-quadruplex in FOXE3 5 UTR augments gene expression. The experiments are performed in triplicate and results are expressed in Mean ± SEM. Asterisks indicate the statistical significance (Student's t-test) relative to mUTR transfected cells (*, $p < 0.05$).

infancy [33]. Although studies individually addressed the role of G-quadruplex in 5′ UTR, there is a need to explore the position specific effect of G-quadruplex in naturally occurring mRNAs.

Thus, employing in silico search for putative G-quadruplex sequence at start site of mRNA, we located AKTIP, CTSB, FOXE3 and APOA1BP genes. Biophysical characterisation of putative G-quadruplex sequences (PG4) in AKTIP mRNA, CTSB mRNA and FOXE3 RNA depicted the formation of highly stable G-quadruplexes. AKTIP G-quadruplex was found to be more stable than CTSB and FOXE3 G-quadruplex. At 1 mM KCl salt concentration, it gives a higher magnitude of free energy change (-11.68 kcal/mol) than CTSB G-quadruplex (-5.47 kcal/mol). Their stability are in an order of AKTIP RNA > CTSB RNA > FOXE3 RNA. The most probable reason for highest stability of AKTIP G-quadruplex is its ability to form G4 quadruplex in contrast to G3 quadruplexes forming ability of

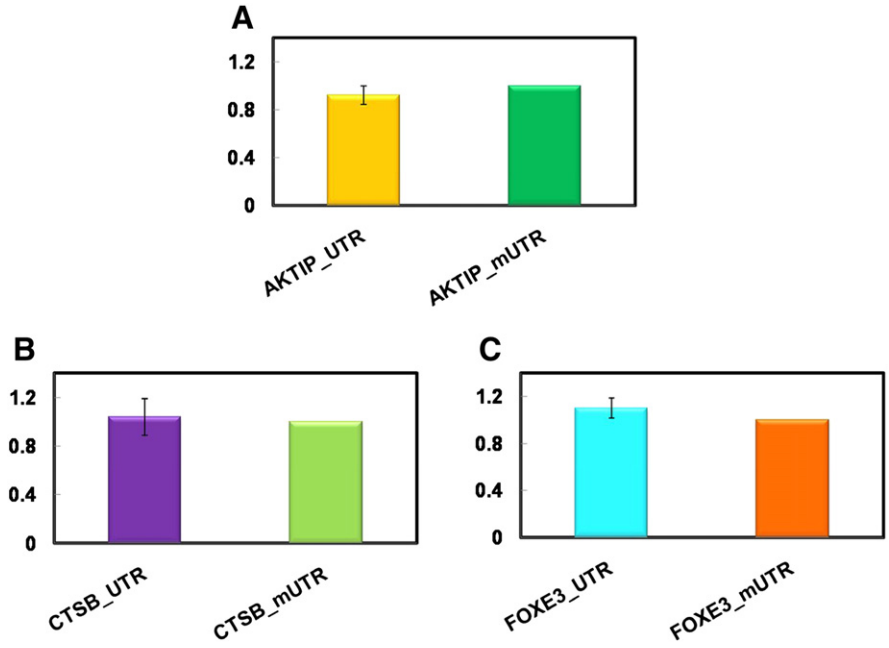


Fig. 4. Quantification of transcript levels of reporter genes. Real time PCR showed similar mRNA levels of reporter gene transfected with (A) pAKTIP UTR, pAKTIP mUTR1 and pAKTIP mUTR (B) pCTSB UTR and pCTSB mUTR (C) pFOXE3 UTR and pFOXE3 mUTR. The experiments are performed in triplicate and results are expressed in Mean ± SEM.

CTSB and FOXE3 PG4. Even in the absence of salt, AKTIP G-quadruplex shows a T_m of 65.4 °C in 10 mM sodium cacodylate buffer (pH 7.4). These T_m values were in agreement with the current existing literature [11–13,19,20]. CD signals of these PG4s are also in accord with their UV melting determined stability parameters. Their CD spectra indicated parallel topology of G-quadruplex as previously shown. Riboguanines tend to attain anti-conformation (and not able to adopt syn-conformation), a prerequisite for parallel G-quadruplex formation.

Albeit APOA1BP PG4 abides in the G-quadruplex forming sequence $G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+}$, it fails to show any structural existence in in vitro conditions. UV and CD studies clearly demonstrated its inability to form G-quadruplex in vitro. As cytosine nucleotide is in abundance in this PG4 sequence, the free guanines may not be available to base pair among themselves to form G-quadruplex. Instead guanine Watson Crick base pair with cytosine to form a duplex [25]. This was well supported by a negative CD peak at 210 nm of wavelength, a signature for duplex formation.

On evaluating the role of G-quadruplex located just at the beginning of 5′ UTR using dual luciferase assay, we observed the execution of a wide range of functions by this secondary structure. While G-quadruplex in 5′ UTR of AKTIP and CTSB RNA exhibited a repressive role, the presence of G-quadruplex in FOXE3 5′ UTR increased gene expression. Moreover, the extent of regulation varied in different genes. In the presence of G-quadruplex, there was 30% gene repression in CTSB and 20% gene suppression in AKTIP. G-quadruplex increased gene expression by 60% in FOXE3 suggesting the versatility of G-quadruplex in modulating gene function. Additionally, when we mutated G-quadruplex forming sequence in AKTIP 5′ UTR such as to inhibit G4 G-quadruplex and favour G2 quadruplex formation, it was unable to suppress gene expression in AKTIP indicating the requirement of G4 G-quadruplex in AKTIP to modulate gene activity. Furthermore, RT-PCR indicated similar transcript levels of reporter genes in transfected cells, negating out the chances of gene regulation at the transcriptional level.

G-quadruplex forming sequence in APOA1BP 5′ UTR did not bring about any change in reporter gene activity. Even PG4 sequence per se did not alter gene expression, clearly supporting the biophysical results of this sequence. The mere presence of G-quadruplex forming sequence in 5′ UTR of an mRNA, thus, does not support its prevalence in vivo [25,34]. Its overall GC content acts as the deciding factor for its existence in vivo. Watson–Crick duplex formation possibly outcompetes Hoogsteen G-quadruplex formation diminishing the existence of G-quadruplex in the 5′ UTR of APOA1BP mRNA.

We have also verified the folding of the AKTIP, CTSB, FOXE3, APOA1BP sequences using the new scoring system [25], and found the following scoring values (cG/cC score) of PG4 sequences with 100 nts flank: AKTIP- 8.53, CTSB- 3.00, FOXE3- 8.23, APOA1BP- 3.95. A threshold value of 3.05 was selected by Beaudoin et al. to increase specificity of cG/cC score, with high scorers supporting G-quadruplex folding and low scorers were non-folding. AKTIP and FOXE3 depicted high scoring values indicating their high propensity to form G-quadruplex; also experimentally validated in our studies. However, CTSB and APOA1BP with scoring values lying between 3 and 4 demonstrated discrepancies. Although APOA1BP gave a score of 3.95, it demonstrated its inability to form G-quadruplex in vitro and in cellulo conditions. On the other hand, even on having a score of 3.0, CTSB was able to form G-quadruplex in vitro and in cellulo. While demonstrating the efficiency of this scoring system in predicting G-quadruplex folding, Beaudoin et al. have also pointed that among 14 PG4 candidates selected for the study, this scoring methodology with a threshold of 3.05 gave two discrepancies with one false positive and one false negative. Our candidate genes CTSB and APOA1BP are also showing such divergence as they are lying very close to the threshold value.

Though biophysical studies indicated higher stability of G-quadruplex in AKTIP than CTSB, the extent of gene regulation in AKTIP is found to be lower than that in CTSB. In contrast to 30% repression in gene expression in CTSB, G-quadruplex reduced the gene expression of AKTIP only by 20%. These observations clearly suggest that the stability of this structure cannot be directly correlated with its ability to regulate gene expression. Moreover, it suggests that the intracellular milieu and the interacting partners might affect G-quadruplex stability and the extent of its gene regulating function.

Although the mechanism of gene regulation by G-quadruplex is not known, several conjectures can be made. Since G-quadruplex in 5′ UTR of AKTIP and CTSB mRNA represses gene expression, we hypothesise that G-quadruplex structure itself may impede the binding and subsequent movement of the pre-initiation complex along the mRNA. It may interfere with the binding of the cap binding protein eIF4E, delay the helicase activity of eIF4A or affect the 40S ribosomal subunit recruitment onto the mRNA [1,3]. In the process, the cell has to expend some amount of energy to resolve this structure [3]. Alternatively, G-quadruplex may likely decoy some inhibitors that interfere with the binding of the pre-initiation complex. The group of Tinoco Jr performed an elegant study in which they translated single messenger RNA hairpins tethered by the ends to optical tweezers using single ribosome and concluded that overall translation process greatly depends on the pause durations spent in destabilising secondary structures present in mRNA [35]. The helicase activity of ribosome applies force and reduces pause time periods without influencing the translocation time period. G-quadruplex structure may act as a mechanical roadblock to the ribosomal protein and provides an activation barrier which needs to be overcome by the force applied by mRNA helicase activity of the ribosome [36]. By estimating mechanical stability of RNA G-quadruplex and corresponding rupture force applied by ribosomal motor proteins, one can calculate the pause durations as well as success/failure of ribosomal complex to unfold this structure. In case mechanical stability of G-quadruplex far exceeds the rupture force of ribosome, it can recruit other helicase enzymes to resolve this structure. More detailed single molecule studies are required to shed light on this tug of war force between G-quadruplex and motor proteins. On the other hand, the translational enhancer role of 5′ UTR G-quadruplex in FOXE3 can be attributed to the recruitment of activators onto 5′ cap region of mRNA. These activating factors may facilitate the assemblage of the pre-initiation complex onto the mRNA, thereby increasing gene expression.

To conclude, our study highlighted the role of G-quadruplex at 5′ terminal of cellular mRNAs. We found that this structure exhibits contrasting regulatory roles in different genes. It activates or represses gene function. Even the extent or magnitude of gene modulation by this structure varied. Furthermore, our studies suggested that in many instances, the stability of G-quadruplex cannot be extrapolated to the level of its functional activity. The unknown factors such as molecular partners, spatio-temporal requirements and intracellular milieu greatly influence the role of G-quadruplex.

Funding source statement

This work was financially supported by the BSC0123 project (Project: Genome Dynamics in Cellular Organisation, Differentiation and Enantiostasis) from the Council of Scientific and Industrial Research (CSIR), Government of India.

Appendix A. Supplementary data

Sequences of mutated RNA oligonucleotides for biophysical studies and DNA primers for RT-PCR are given in Table S1 and S2. More biophysical studies on RNA oligonucleotides are given in Fig. S1, S2, S3, S4 and S5. Dual Luciferase Assay results for APOA1BP are provided in Fig. S6. Supplementary data to this article can be found online at doi: http://dx.doi.org/10.1016/j.bbagen.2014.08.017.

References

[1] R.J. Jackson, C.U. Hellen, T.V. Pestova, The mechanism of eukaryotic translation initiation and principles of its regulation, Nature reviews, Mol. Cell Biol. 11 (2010) 113–127.
[2] S. Chatterjee, J.K. Pal, Role of 5′- and 3′-untranslated regions of mRNAs in human diseases, Biol. Cell /Eur. Cell. Biol. Org. 101 (2009) 251–262.
[3] A. Parsyan, Y. Svitkin, D. Shahbazian, C. Gkogkas, P. Lasko, W.C. Merrick, N. Sonenberg, mRNA helicases: the tacticians of translational control, Nature reviews, Mol. Cell Biol. 12 (2011) 235–245.
[4] B.M. Pickering, A.E. Willis, The implications of structured 5′ untranslated regions on translation and disease, Semin. Cell Dev. Biol. 16 (2005) 39–47.
[5] Y. Wan, M. Kertesz, R.C. Spitale, E. Segal, H.Y. Chang, Understanding the transcriptome through RNA structure, Nature reviews, Genetics 12 (2011) 641–655.
[6] G. Biffi, M. Di Antonio, D. Tannahill, S. Balasubramanian, Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells, Nat. Chem. 6 (2014) 75–80.
[7] A. Bugaut, S. Balasubramanian, 5′-UTR RNA G-quadruplexes: translation regulation and targeting, Nucleic Acids Res. 40 (2012) 4727–4741.
[8] T. Agarwal, G. Jayaraj, S.P. Pandey, P. Agarwala, S. Maiti, RNA G-quadruplexes: G-quadruplexes with "U" turns, Curr. Pharm. Des. 18 (2012) 2102–2111.
[9] S. Millevoi, H. Moine, S. Vagner, G-quadruplexes in RNA biology, Wiley interdisciplinary reviews, RNA 3 (2012) 495–507.
[10] T. Endoh, Y. Kawasaki, N. Sugimoto, Suppression of gene expression by G-quadruplexes in open reading frames depends on G-quadruplex stability, Angew. Chem. Int. Ed. Engl. 52 (2013) 5522–5526.
[11] S. Kumari, A. Bugaut, J.L. Huppert, S. Balasubramanian, An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation, Nat. Chem. Biol. 3 (2007) 218–221.
[12] A. Arora, M. Dutkiewicz, V. Scaria, M. Hariharan, S. Maiti, J. Kurreck, Inhibition of translation in living eukaryotic cells by an RNA G-quadruplex motif, RNA 14 (2008) 1290–1296.
[13] M.J. Morris, S. Basu, An unusually stable G-quadruplex within the 5′-UTR of the MT3 matrix metalloproteinase mRNA represses translation in eukaryotic cells, Biochemistry 48 (2009) 5313–5319.
[14] G.D. Balkwill, K. Derecka, T.P. Garner, C. Hodgman, A.P. Flint, M.S. Searle, Repression of translation of human estrogen receptor alpha by G-quadruplex formation, Biochemistry 48 (2009) 11487–11495.
[15] J.D. Beaudoin, J.P. Perreault, 5′-UTR G-quadruplex structures acting as translational repressors, Nucleic Acids Res. 38 (2010) 7022–7036.
[16] D. Gomez, A. Guedin, J.L. Mergny, B. Salles, J.F. Riou, M.P. Teulade-Fichou, P. Calsou, A G-quadruplex structure within the 5′-UTR of TRF2 mRNA represses translation in human cells, Nucleic Acids Res. 38 (2010) 7187–7198.
[17] R. Shahid, A. Bugaut, S. Balasubramanian, The BCL-2 5′ untranslated region contains an RNA G-quadruplex-forming motif that modulates protein expression, Biochemistry 49 (2010) 8300–8306.
[18] M.J. Morris, Y. Negishi, C. Pazsint, J.D. Schonhoft, S. Basu, An RNA G-quadruplex is essential for cap-independent translation initiation in human VEGF IRES, J. Am. Chem. Soc. 132 (2010) 17831–17839.
[19] S. Lammich, F. Kamp, J. Wagner, B. Nuscher, S. Zilow, A.K. Ludwig, M. Willem, C. Haass, Translational repression of the disintegrin and metalloprotease ADAM10 by a stable G-quadruplex secondary structure in its 5′-untranslated region, J. Biol. Chem. 286 (2011) 45063–45072.
[20] P. Agarwala, S. Pandey, K. Mapa, S. Maiti, The G-quadruplex augments translation in the 5′ untranslated region of transforming growth factor beta2, Biochemistry 52 (2013) 1528–1538.
[21] M.J. Morris, K.L. Wingate, J. Silwal, T.C. Leeper, S. Basu, 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 (2012) 4137–4145.
[22] S. Kumari, A. Bugaut, S. Balasubramanian, 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 (2008) 12664–12669.
[23] K. Halder, M. Wieland, J.S. Hartig, Predictable suppression of gene expression by 5′-UTR-based RNA quadruplexes, Nucleic Acids Res. 37 (2009) 6811–6817.
[24] M. Wieland, J.S. Hartig, RNA quadruplex-based modulation of gene expression, Chem. Biol. 14 (2007) 757–763.
[25] J.D. Beaudoin, R. Jodoin, J.P. Perreault, New scoring system to identify RNA G-quadruplex folding, Nucleic Acids Res. 42 (2014) 1209–1223.
[26] S. Pandey, P. Agarwala, S. Maiti, Effect of loops and G-quartets on the stability of RNA G-quadruplexes, J. Phys. Chem. B 117 (2013) 6896–6905.
[27] A. Arora, S. Maiti, Differential biophysical behavior of human telomeric RNA and DNA quadruplex, J. Phys. Chem. B 113 (2009) 10515–10520.
[28] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Res. 29 (2001) e45.
[29] V. Scaria, M. Hariharan, A. Arora, S. Maiti, Quadfinder: server for identification and analysis of quadruplex-forming motifs in nucleotide sequences, Nucleic Acids Res. 34 (2006) W683–W685.
[30] J.L. Mergny, A.T. Phan, L. Lacroix, Following G-quartet formation by UV-spectroscopy, FEBS Lett. 435 (1998) 74–78.
[31] J.L. Mergny, J. Li, L. Lacroix, S. Amrane, J.B. Chaires, Thermal difference spectra: a specific signature for nucleic acid structures, Nucleic Acids Res. 33 (2005) e138.
[32] N. Kolotева, P.P. Muller, J.E. McCarthy, The position dependence of translational regulation via RNA–RNA and RNA–protein interactions in the 5′-untranslated region of eukaryotic mRNA is a function of the thermodynamic competence of 40 S ribosomes in translational initiation, J. Biol. Chem. 272 (1997) 16531–16539.
[33] J.L. Huppert, A. Bugaut, S. Kumari, S. Balasubramanian, G-quadruplexes: the beginning and end of UTRs, Nucleic Acids Res. 36 (2008) 6260–6268.
[34] S. Saxena, D. Miyoshi, N. Sugimoto, Sole and stable RNA duplexes of G-rich sequences located in the 5′-untranslated region of protooncogenes, Biochemistry 49 (2010) 7190–7201.
[35] J.D. Wen, L. Lancaster, C. Hodges, A.C. Zeri, S.H. Yoshimura, H.F. Noller, C. Bustamante, I. Tinoco, Following translation by single ribosomes one codon at a time, Nature 452 (2008) 598–603.
[36] P.M. Yangyuoru, A.Y. Zhang, Z. Shi, D. Koirala, S. Balasubramanian, H. Mao, Mechanochemical properties of individual human telomeric RNA (TERRA) G-quadruplexes, Chembiochem: Eur. J. Chem. Biol. 14 (2013) 1931–1935.