# A novel method for gene-specific enhancement of protein translation by targeting 5'UTRs of selected tumor suppressorss

## Authors

Adam Master[1,2*], Anna Wójcicka[1,3,4], Kamilla Giżewska[2,], Piotr Popławski[1], Graham R. Williams[5], Alicja Nauman[1,3*]

## Affiliations

[1] The Centre of Postgraduate Medical Education, Department of Biochemistry and Molecular Biology, ul. Marymoncka 99/103, 01-813 Warsaw, Poland

[2] BioTe21, Laboratory of Molecular Medical Biology, ul. Krolowej Jadwigi 33/3b, 30-209 Cracow, Poland

[3] Centre of New Technologies, University of Warsaw, Banacha 2c, 02-089 Warsaw, Poland

[4] Genomic Medicine, Medical University of Warsaw, Zwirki i Wigury 61, 02-091 Warsaw, Poland

[5] Molecular Endocrinology Group, Department of Medicine, Imperial College London, Hammersmith Campus, London W12 0NN, United Kingdom

*Corresponding authors

E-mail: alicja.nauman@uw.edu.pl (AN), adam.master@gene-factory.com (AM)

Keywords

Translational control, translation-enhancing oligonucleotides, cis-acting elements, untranslated regions (UTR), THRB, cancer gene therapy.

# Abstract

**Background**: Translational control is a mechanism of protein synthesis regulation emerging as an important target for new therapeutics. Naturally occurring microRNAs and synthetic small inhibitory RNAs (siRNAs) are the most recognized regulatory molecules acting via RNA interference. Surprisingly, recent studies have shown that interfering RNAs may also activate gene transcription via the newly discovered phenomenon of small RNA-induced gene activation (RNAa). Thus far, the small activating RNAs (saRNAs) have only been demonstrated as promoter-specific transcriptional activators.

**Findings**: We demonstrate that oligonucleotide-based *trans*-acting factors can also specifically enhance gene expression at the level of protein translation by acting at sequence-specific targets within the messenger RNA 5'-untranslated region (5'UTR). We designed a set of short synthetic oligonucleotides (dGoligos), specifically targeting alternatively spliced 5'UTRs in transcripts expressed from the *THRB* and *CDKN2A* suppressor genes. The *in vitro* translation efficiency of reporter constructs containing alternative TRβ1 5'UTRs was increased by up to more than 55-fold following exposure to specific dGoligos. Moreover, we found that the most folded 5'UTR has higher translational regulatory potential when compared to the weakly folded TRβ1 variant. This suggests such a strategy may be especially applied to enhance translation from relatively inactive transcripts containing long 5'UTRs of complex structure.

**Significance:** This report represents the first method for gene-specific translation enhancement using selective *trans*-acting factors designed to target specific 5'UTR *cis*-acting elements. This simple strategy may be developed further to complement other available methods for gene expression regulation including gene silencing. The dGoligo-mediated translation-enhancing approach has the potential to be transferred to increase the translation efficiency of any suitable target gene and may have future application in gene therapy strategies to enhance expression of proteins including tumor suppressors.

# Introduction

Translational control is one of the most important mechanisms of post-transcriptional regulation of gene expression, determining final protein levels [1]. Initiation of translation [2] is a rate-limiting phase of protein synthesis, controlled by translation -silencing or -enhancing *cis*-acting elements located in the 5' and 3' untranslated regions (5'UTR, 3'UTR) of mRNAs [3]. The best studied *cis*-acting elements within the UTRs are the upstream open reading frames (uORFs) [4] and internal ribosomal entry sites (IRESs) [5]. These regulatory sequences may be organized in secondary and tertiary RNA structures that are recognized by *trans*-acting factors [6] such as translation factors [7], naturally occurring microRNAs (microRNAs) [8] as well as synthetic small interfering RNAs (siRNAs) [9] and antisense-like oligonucleotides (ASOs) usually lowering final protein levels [10].

Translation of most human mRNAs occurs via a 5'-cap-dependent mechanism [11]. Certain physiological or pathological factors may switch the classic mechanism to an alternative one that can be controlled by an mRNA element such as uORF, IRES, iron responsive element (IRE), RNA hypoxia response element (rHRE), differentiation-control element (DICE) or cap-independent translational enhancer (CITE) [12, 13]. An alternate cap-independent IRES-dependent translation [5, 13] is activated in some cellular phases such as cell division [14] and during integrated stress responses (ISR) [15] caused by heat shock, serum or amino-acid deprivation and in hypoxia, as frequently observed in solid tumors [16]. Expression of specific genes involved in the stress responses can be also controlled by uORFs [4, 13]. ISR-enhanced synthesis of ATF4 (Activating Transcription Factor 4) protein is an extensively-studied model of the translational control [4]. This mechanism involves the differential contribution of two uORFs: the 5' proximal uORF1 that is a positive *cis*-acting element and the inhibitory uORF2 overlapping correct ATF4 ORF in an out of frame manner. Non-stressed, normal conditions allow cells for fast translation of the short uORF1 and ribosome reinitiation at the uORF2, that results in synthesis of truncated proteins. In contrast, stress conditions increase the time required for more accurate scanning that allows the ribosomes to bypass the inhibitory uORF2 and reinitiate at the downstream ATF4-coding region [4]. Translation initiation can also be slowed down by various interfering *trans*-acting factors [1] or highly-ordered RNA structures [17], which require RNA helicases

to be scanned through [3, 13]. Moreover, 5'UTR structures, recruiting RNA helicase eIF4A2, have now been demonstrated to play a crucial role in 3'UTR-dependent, microRNA-mediated gene silencing [18]. Therefore, efficiency of various mechanisms involved in translation initiation has been thought to be dependent on the folding state of 5'UTRs, determined by the Gibbs energy (ΔG) [17].

Many genes have several alternative 5'UTR splice variants that can differentially regulate translation of downstream coding sequences [6]. One example of such a complex gene is *THRB* (GeneID 7068), which encodes β isoforms of human thyroid hormone receptors (TRβ1, TRβ2 and TRβ4) [19] and contains numerous alternatively spliced exons that generate various alternate 5'UTRs in mRNAs from which the TRβ tumor suppressor protein is expressed [19, 20]. Multiple strongly folded 5'UTRs can also be expressed by another tumor suppressor – *CDKN2A* (GeneID 1029) [21]. Both genes encode 5'UTRs containing numerous uORFs [21, 22] and IRES-like sequences [21, 23]. These 5'UTRs vary in length, GC-content and secondary structure and have been shown to influence the efficiency of protein translation [21, 23].

Recent studies have revealed that some naturally occurring microRNAs, considered traditionally as inhibitory *trans*-acting factors that bind to 3'UTR sequences, can also up-regulate protein synthesis [24]. Moreover, it has been discovered that several mRNAs contain similar microRNA targets termed miBridges present in both 3'UTR and 5'UTR regions that can bind the same microRNA molecule [25]. Further supporting the hypothesis of microRNA binding to 5'UTRs, a liver-specific microRNA, miR-122 has been shown to stimulate synthesis of hepatitis C virus (HCV) protein by direct interaction with two target sites in the 5'UTR of the HCV genome [26]. Even though a single microRNA usually targets possibly hundreds of cellular mRNAs [27], showing low selectivity towards transcripts of a single gene [28], these findings demonstrate a new role of short interfering RNAs that may lead not only to gene repression, but also to protein synthesis enhancement.

Recently, a new type of RNA interference has been shown to result from promoter-specific activation of gene transcription (RNAa, RNA activation) that is triggered by a novel class of interfering RNAs termed small activating RNAs (saRNAs), which target discrete promoter sequences [29]. The saRNAs were used for promoter-specific upregulation of gene transcription [30]. On the other hand, the saRNAs were alternatively reported to represent siRNAs that bind to and inhibit long naturally

occurring antisense transcripts (NATs) overlapping complementary promoter regions of target genes, which play a role in transcriptional repression [31, 32]. Thus, silencing of the NATs could indirectly lead to transcriptional activation of the genes [33, 34]. Both mechanisms of gene regulation, however, have been shown to control only the levels of mRNA expression [31, 33].

Here we describe a novel method for 5'UTR-specific enhancement of translation. Protein overexpression is triggered by synthetic, translation-enhancing oligonucleotides, termed dGoligos (dGs), which were originally designed to alter Gibbs energy-dependent secondary structure formation of specific sequences of TRβ1 5'UTRs. Although ΔG is a well-known measure of the stability of higher-order structures of nucleic acid molecules, we used this parameter in a new way, defined in a bespoke dGenhancer calculator. This tool allowed us to determine *cis*-acting elements within TRβ1 5'UTRs that were recognized by dGs. Then, the translation-enhancing effects were successfully confirmed by the use of dGs design to target $p16^{INK4a}$ 5'UTR encoded by the *CDKN2A* gene. dGoligos thus offer the potential for a novel and specific gene-therapy approach to re-express or over-express individual proteins such as tumor suppressors.

# Results

## Translation regulated by differentially folded TRβ1 5'UTRs

TRβ1 5'UTR splice variants A-G subcloned upstream of the luciferase reporter gene in pKS plasmids [22] were tested for their basic expression characteristics by coupled *in vitro* transcription-translation (RTS 100 Wheat Germ CECF). The basic luciferase levels served as starting points for further studies on translation-enhancing elements of the 5'UTRs. Initial results demonstrated that variants A-G differently regulate the reporter protein translation efficiency (Figs 1a and 1b). The measurements were shown in relation to control plasmid (pKS-control) containing an irrelevant synthetic vector-based leader sequence (ΔG=−6.8 kcal/mol) lacking a TRβ1 UTR (see Materials and Methods). We found that the basic luciferase expression rates were the highest (24.09% of the control, $p<0.001$) when driven by the most weakly folded variant A, possessing the lowest (negative) value of Gibbs energy (ΔG=−69.0 kcal/mol, Fig 1b). In contrast, luciferase expression from plasmids containing variant G

(ΔG=−127.0 kcal/mol) and the most folded variant F (ΔG=−128.9 kcal/mol) was strongly inhibited (3.00% and 4.03% of the control for variant G and F respectively, p<0.001). Similar effects were previously reported in human placental JEG-3 choriocarcinoma cells [22] and in Caki-2, a human clear cell renal cell carcinoma line [23]. To check whether the different luciferase protein levels resulted from changes in levels of particular transcripts, we quantified luciferase mRNAs after 6h of the coupled transcription-translation reaction. Real-time PCR revealed no significant differences in luciferase transcription rates driven from the tested variants A-G (Fig 1a). These results are consistent with previous observations in Caki-2 cells [23] and in another *in vitro* translation system based on rabbit reticulocyte lysates [22].

**Fig 1. Correlation between Gibbs energy and basic TRβ1 5'UTR-mediated translation efficiency.** (**a**) Luciferase mRNA levels from *in vitro* wheat germ-based coupled transcription-translation assay performed on plasmids containing TRβ1 5'UTR variants A-G are shown relative to control plasmid. (**b**) Effects of 5'UTR variants A-G on luciferase activities after 6h coupled transcription-translation. Three independent experiments were performed in triplicate and shown as mean % mRNA or luciferase activity ± SD. Data were analyzed by ANOVA followed by Dunnett's multiple comparison test; *p< 0.001 vs. control. **c**, Correlation between the calculated Gibbs energies (X axis) of each 5'UTR variant (S1 Table) and UTR-mediated luciferase translation efficiency. The correlation is shown as the exponential trend-line $y=127.29*e^{0,0248*X}$, where x = calculated Gibbs energy; Pearson product-moment correlation coefficient $r^2=0,9746$. Logarithmically transformed data of translation efficiency (Y axis) were analyzed together with Gibbs energies by linear regression; p=0.0123.

## Correlation between Gibbs energy and translation efficiency

Although multiple bioinformatic tools for the analysis of higher-order structures of RNA are available, their utility in predicting the effects of translation -silencing or -enhancing *cis*-acting elements on the levels of protein expression is limited [35]. These elements may require specific secondary and tertiary folding to exert their normal function and may regulate the translation of

downstream sequences independently of their nucleotide composition and Gibbs energy (ΔG) status [4, 36]. Therefore, we investigated whether the Gibbs energy of TRβ1 5'UTRs (S1 Table) could correlate with 5'UTR-controlled translation efficiency of a downstream ORF. High Pearson's correlation coefficient $r^2$=0.9746 ($p<0.005$) showed that protein levels are strictly dependent on Gibbs energies of the TRβ1 5'UTRs (Fig 1). The correlation also resulted in an exponential equation ($y=140.46*e^{0,0307*x}$, Fig 1c) that could serve for theoretical prediction of the translation rate of any TRβ1 5'UTR variant. An example application of this equation is shown in Table S2. Finally, the correlation allowed us to use ΔG-based algorithm derived from the dGenhancer calculator for an automatic design of oligonucleotide *trans*-acting factors (S1 Appendix).

## Prediction of *cis*-acting-elements of high regulatory importance

Since most of the alternatively spliced variants of TRβ1 5'UTRs were shown to have strongly folded, translation-inhibiting sequences [22, 23], further study was performed to estimate their translational potential and find a method that could release the *potential* to enhance protein synthesis. We aimed to identify sequences within TRβ1 5'UTRs that could be of particular importance in this process. At first, structures of the TRβ1 5'UTR variants A and F (S1 Fig) were drawn with RNAstructure version 5.2 [37] to determine the most stable secondary structures accompanied by the most negative ΔG. These folding predictions allowed us to identify elements that are likely to be required for secondary and tertiary folding of the 5'UTRs. Then, the elements were compared with publicly known *cis*-acting sequences of IRESite database [38] that allowed us to identify common sequence motifs of possible functional importance. We selected a hairpin sequence within a previously reported domain containing a putative IRES, which has been identified before in the TRβ1-5'UTRs [23], and a sequence conserved among all TRβ1 5'UTR variants with multiple alternate AUGs [22] (Fig 2). To check functional properties of the putative IRES site, we performed a simple test in Caki-2 cells, cultured in serum-deprived medium, which has been reported to switch between cap-dependent and cap-independent (IRES-mediated) translation [5, 12]. We found that protein synthesis rates after serum starvation resulted in higher luciferase activity from plasmid containing the TRβ1 IRES site (pGL3-A) [22] compared to the control (pGL3-control) [22] without the IRES sequence (S2 Fig).

Finally, the manually selected translation -enhancing element e1 (IRES) and translation -inhibiting element e3 (uORFs, Fig 2) were compared with automatically determined elements identified by the dGenhancer. The calculator works on the basis of ΔG changes observed among *in silico* generated 5'UTR sequence variants that differ in a single nucleotide substitution (SNP) altering overall ΔG of the sequences. These artificial variants were created by substitution (base by base) in each nucleotide position of the 5'UTRs (S1 Appendix). Comparing these two approaches we found that the manually and automatically determined elements of the 5'UTRs are fairly similar with one exception of the strongest signal of dGenhancer showing an additional translation-inhibiting element marked as e2 (Fig 2), located in exon 2a, which is present in all TRβ1 5'UTR variants. Identification of these *cis*-acting elements allowed us to design and synthesize specific oligonucleotide-based *trans*-acting factors, termed dGoligos (dGs, Fig 2, S3 Table), designed to recognize and bind the predicted 5'UTR sequences.

Fig 2. dGoligo recognition sites. (**a**) *Cis*-acting elements (e1, e2, e3) of variant A of TRβ1 5'UTR determined by dGenhancer, which indicates signal maxima (a.1, a.2 and a.3) corresponding to the 5'UTR fragments with the highest translational regulatory potential. The signal intensity represents transformed mean of 6 consecutive changes in Gibbs energy (ΔG) observed among 5'UTR sequences containing artificial SNPs. The SNPs were used as a theoretical model to calculate which sequence fragments (within the UTR) can change ΔGs (of the UTR) the most, thereby affecting the translational potential of the 5'UTR. (**b**) dGoligo (dGs 1-10) targets (e1, e2, e3) in TRβ1 5'UTR are shown as underlined sequences. dGs are presented above (sense) and below (antisense) the TRβ1 5'UTR A. Each dG shares homology with the TRβ1 5'UTR, targeting one of the indicated sequences: a putative IRES site on ex1c/ex2a junction (underlined) targeted by dG1, dG2, dG5, dG6, a sequence containing multiple alternate AUGs (uORFs-rich region) and located on ex2a/ex3 junction targeted by dG3, dG4 or a sequence in the middle of exon 2 targeted by dG7 dG8, dG9 and dG10. All dG*s* were designed as complementary pairs of antisense strand (dG2, dG4, dG6, dG8, dG10) directly recognizing the indicated region and sense strand (dG1, dG3, dG5, dG7, dG9) that can bind to distant sequences that interact through complimentary base-pairing with the indicated region (S1 Fig). Oligonucleotides dG5, dG6,

dG9 and dG10 were synthesized as microRNA-like oligonucleotides with 3-nt insertion in the middle of their sequences.

## dGoligo design and synthesis

We next evaluated whether we could selectively alter the Gibbs energy-dependent secondary structure formation of TRβ1 5'UTRs using oligonucleotide–based *trans*-acting factors. We synthesized a set of DNA oligonucleotides (dGoligos) directed to interfere with previously identified TRβ1 5'UTR *cis*-acting elements. High translational regulatory potential was defined as the potential of the translation-regulating elements to enhance protein synthesis from low to high rates. This regulation is thought to be controlled by distant *cis*- or *trans*-acting factors specifically binding to the regulatory elements (Fig 3d). A putative IRES [23] site and a sequence containing multiple alternate AUGs [22] were targets for dGoligos (dGs) dG1, dG2, dG5, dG6 and dG3, dG4 respectively (Fig 2). dG7 and dG8 were designed to target a sequence located in the middle of exon 2a, identified by the dGenhancer to have the highest regulatory potential. For *in vitro* assays, dG*s* were synthesized as sense-, antisense- or microRNA-like DNA oligonucleotides (S3 Table). 2'-O-methyl RNA-modified derivatives were synthesized for *in vivo* assays.

dG-mediated linearization of 5'UTRs was predicted to disturb inhibitory structures and/or liberate translation-enhancing elements. These proposed functions of the synthetic oligonucleotides were likely to be required for structural rearrangement of the 5'UTRs into their translationally active conformation (Fig 3) that facilitates interaction with naturally occurring elements directly controlling final protein levels.

**Fig 3. Folding states of TRβ1 5'UTRs.** Translation efficiency of TRβ1 is dependent on folding states of its 5'UTR, which is proposed to be: strongly folded (**a**), partially unfolded (**b**) or fully unfolded following interaction with dG1 and dG4 (**c**). The 5'UTR is shown as curve ended by an arrow at AUG translation start codon. Two linked ovals assigned by letter R represent ribosome that may be blocked by distant *cis*-acting element (*cis*-a.e) or *trans*-acting factor (dG1 or dG4). AUG start codon is preceded by selected translation-regulating elements (e1 and e3). Translation-enhancing element e1 contains putative Internal Ribosome Entry Site (IRES) that may be involved in enhancement of cap-independent

translation initiation. Translation-silencing element e3 contains upstream Open Reading Frames (uORFs)–rich region that can reduce translation initiation from correct AUG start codon due to simultaneous synthesis of truncated proteins originated from upstream AUGs (shown by inverted ribosome). **a**, Theoretical folding state of TRβ1 5'UTR characterized by the presence of highly-structured sequence that can block both: translation-enhancing e1 and translation-silencing e3, finally leading to only basal protein synthesis. (**b**) Another theoretical folding state described by partially unfolded 5'UTR with blocked e1 and unblocked e3, resulting in basal translation rates as well. (**c**) Proposed model of dG1 and dG4 -mediated enhancement of translation efficiency, wherein antisense dG4 could lead to repression of uORFs within e3, whereas binding of sense dG1 to a distant sequence (usually folded with e1) could release this translation-enhancing region, allowing for protein over-expression (additional description is given in S4 Fig). (**d**) dGoligo (dG) targets on mRNA. Locations of dG binding sites are shown in the context of typical targets of the most known small interfering RNAs. microRNA (2, 5), siRNA (3), saRNA (1) and ASO (4), are shown by short grey arrows. Newly described interactions that may result in up-regulation of gene expression are indicated by asterisk*.

## Translation-enhancing dGoligos targeting TRβ1 5'UTRs

The influence of dGoligos on translation efficiency was studied in coupled *in vitro* transcription-translation reactions using plasmids containing the least (A) and the most (F) folded variant of TRβ1 5'UTR cloned upstream of luciferase reporter. Effects of each dGoligo on protein synthesis were assessed in a translation-enhancing assay. Levels of luciferase mRNA and protein (luciferase activity) expressed from plasmid without or with dGoligo supplementation were measured by Real-Time PCR and luminometry. Maximum luciferase activity was observed after 6h (S3 Fig) and this time point was chosen for subsequent analyses. No statistically significant differences in luciferase mRNA levels were observed between control and plasmid constructs.

To eliminate the effects of possible non-specific dG-plasmid interactions, all transcription and translation measurements for both pKS-A and pKS-F plasmids were standardized to mRNA levels driven from control plasmid with a short irrelevant vector-based leader sequence that contained no specific dGoligo binding sites (pKS-control).

**dG1 and dG2** were synthesized as a pair of complementary, sense- and antisense- DNA oligonucleotides. dG1 shares sequence with the most stable secondary structure of the translation-enhancing element e1 (Fig 2) containing a putative IRES site [23] (S4g.1 Fig) while sequence of dG2 is complementary to this region. As a result, dG1 increased translation efficiency over 1.29-fold when using pKS-A and 2.90-fold in case of the use of pKS-F ($p<0.001$, Figs 4b and 4d), while dG2 decreased the protein levels by 1.80-fold for pKS-A and did not alter translation for pKS-F, probably due to the lack of 3'-end of exon 1c in the variant F (S1b Fig). Since the sense dG1 has the same sequence as e1 element of the 5'UTR, it can interact with and block the homologous distant mRNA sequences (*cis*-acting elements) (Fig 3a) that can fold with the e1 domain [23] and lead to its repression. Thus, dG1 was designed to release the domain allowing for appropriate folding of this sequence that appears to be required for efficient translation. Antisense dG2, complementary to e1 sequence, was designed to bind this region directly, preventing formation of an active sequence conformation (S4 Fig).

**dG3** (sense) and complementary **dG4** (antisense) were designed to target sequence at ex2a/ex3 junction of TRβ1 5'UTR (e3 in Fig 2) that contains numerous upstream translation start cordons - uAUGs (S4h.1 Fig). The sense dG3 has the same sequence as the uORFs-rich domain of the 5'UTR, and thus may interact with distant *cis*-acting elements (Fig 3a), which normally can fold with uAUG-rich domains [22] and act as uORF inhibitors allowing for more efficient translation from the canonical start codon. Thus, dG3 was designed to release the uORFs-rich domain of the mRNA, and was expected to facilitate initiation of translation from upstream AUGs (S4h.1 Fig) resulting in reduced initiation of protein synthesis from the correct AUG start codon (Figs 4b and 4d). By contrast, antisense dG4 was designed to bind the mRNA sequence containing the uAUGs to render the uAUG-rich region inaccessible for the translation machinery (S4b.1 Fig), resulting in enhanced translation initiation from the correct AUG start codon. These predictions were supported by results showing that dG3 decreased translation efficiency by 2.40-fold for pKS-A and 7.25-fold for pKS-F ($p<0.001$), whereas addition of dG4 increased translation efficiency by 1.33-fold for pKS-A and 1.86-fold for pKS-F ($p<0.001$). The findings suggest that blockade of alternate uAUGs is important for efficient protein translation and are consistent with results showing that initiation codons located upstream of the correct start codon of the TRβ1 can markedly affect the efficiency of protein synthesis [4, 22]. The translation-enhancing

action of dG4 could also be explained using a model of enhancement of ATF4 translation in stress conditions, which can *switch off* inhibitory uORFs by increasing the time of 5'UTR scanning [4]. This allows for ribosomes to *bypass* the uORFs and find the correct ATF4 start codon in the Kozak consensus sequence [4, 13]. Our results show that binding of dG4 to TRβ1 uORFs-rich region forms a double stranded sequence that possibly slows down the scanning machinery. Thus, the use of dG4 may delay translation initiation, as it is observed in stress conditions, leading to enhanced levels of correct proteins. Moreover, the uORF-regulated translation initiation in stress conditions is found to be accompanied by higher translation rates of IRES-containing mRNAs [4, 13]. Indeed, our *in vitro* experiments showed that combined addition of sense dG1 and antisense dG4 increased luciferase activity by 1.77-fold ( $p<0.001$, Fig 4b) from pKS-A and 6.58-fold from pKS-F ($p<0.001$, Fig 4d). The translation-enhancing effect could result from simultaneous release of the translation-enhancing element (e1) [23] and block of the uORFs-rich region [22] (Fig 3c). These results may also suggest that strongly folded variant F could be characterized by a higher translational regulatory potential (S2 Table and S5 Fig).

Furthermore, analysis of **dG7** and **dG8**, designed on the basis of a *cis*-acting element detected by dGenhancer (S3 Table), revealed that **dG8** enhanced translation by 6.02-fold and 8.30-fold for pKS-A and pKS-F respectively, whereas sense **dG7** had no significant effect (Fig 4a and 4b). Interestingly, a combination of antisense dG8 and sense dG1 enhanced luciferase activity over 28.1- (pKS-A) and 55.8-fold (pKS-F) ($p<0.001$). These effects reinforce the finding that exon 2a is conserved in all TRβ1 alternatively spliced 5'UTR variants and suggest an important role in translation control from this locus. Thus, blocking of exon 2 with complementary antisense **dG8** resulted in the strongest enhancement of translation, indicating that the *cis*-acting element at this site (e2) is not affected by other distant sequences of the 5'UTRs and has a key inhibitory role in translational control of TRβ1. These findings support the hypothesis that dGenhancer may be used to identify ΔG-dependent, translation-regulating elements in 5'UTRs that could be targeted by dGs to alter their Gibbs folding energy and regulate the translation efficiency. Finally, the data suggest a role for the multiple alternatively spliced TRβ1 5'UTRs. Strongly folded variants (including variant F) may serve as a reservoir of less-active mRNAs that could be recruited to increase translation efficiency at times of cellular stress, for example, by the use of

specific *trans*-acting factors such as ncRNAs. Interestingly, bioinformatic analysis of microRNA target sites within TRβ1 untranslated regions revealed that hsa-miR-211 could potentially target both TRβ1 3'UTR and 5'UTR (S6a and S6b Figs) and binding of this non-selective ncRNA could at least affect secondary structures of the UTRs. Indeed, 2'-O-methyl RNA modified hsa-miR-211 enhanced TRβ1 5'UTR-mediated translation by 1.95-fold in Caki-2 cells (S6c Fig).

**Fig 4. dGoligo-mediated gene expression changes under *in vitro* conditions.** Effects of each DNA oligonucleotides dG1-dG10 (S3 Table), on *in vitro* transcription of luciferase reporter constructs (panels **a**, **c**) and translation efficiency (**b**, **d**), using pKS-A (a, b) or pKS-F plasmid (c, d). Data normalized to control (dG-) containing pKS-A or pKS-F without addition of dGoligo. Scrambled control (dGsc) had no effect on transcription or translation. The strongest enhancing effects on luciferase activity were obtained by combining dG1+dG8 (28.10-fold from pKS-A and 55.80-fold from pKS-F). Results from three independent experiments performed in triplicates are shown as mean % mRNA (a, c) or luciferase activity (b, d) ± SD. Data analyzed by ANOVA followed by Dunnett's multiple comparison test. *p<0.001 vs. control.

## dGoligo controls and binding capacity

All the results presented above show that, in contrast to translation-enhancing dGs, their complementary control partners (antisense dG2, sense dG3 and sense dG7) had no or opposite effects (Figs 4b and 4d), thus confirmed target site-specific action of sense dG1, antisense dG2 and antisense dG8. The fact that both sense and antisense oligonucleotides directed to bind 5'UTRs seriously altered translation levels gives a new insight into the nature of these molecules and indicates that this action may depend on specific properties of a target *cis*-acting element. Interestingly, these results also suggest that sense oligonucleotides, used in numerous studies as a control to antisense nucleotides (ASOs), could actually interact with distant *cis*-acting elements, significantly changing translation efficiency as it was shown in our study (Fig 4d).

To check whether the sequence structure of the dG*s* has an impact on their function we synthesized **mismatched control dG5** and **control dG6** sharing the same sequence with dG1 and dG2, respectively, but containing a 2-3 nucleotide long insertion in the middle of both oligonucleotides (Fig

2 and S3 Table). Upon binding target sequence, these additional nucleotides should form a loop that mimics metazoan microRNA structure and prevents perfect base pairing with target TRβ1 5'UTR. By mutating the dGs instead of their target sequences, we avoided problems with undesirable loss of functional properties of investigated 5'UTR *cis*-acting elements [3, 17]. Since numerous studies suggest that translationally active conformation of the UTR variants is of greatest importance for the UTR-mediated translational control, it seems our strategy was the best choice for subsequent control reactions. In addition, *in vitro* transcription-translation assays were performed in wheat germ extract and, as reported, plant microRNAs require nearly perfect base pairing with the target RNA to exert RNAi related effects [39]. Therefore, mismatched dG5 and dG6 served as mutated controls for other dG*s* (S4e.1 and S4f.1 Figs) and were expected not to exert any possible RNAi effects in the wheat germ translation system. As a result, neither the control sense dG5 nor control antisense dG6 altered translation levels (Figs 4b and 4d) that may provide a proof for selectivity of other fully complementary dG*s*. Similar microRNA-like controls were designed on the basis of another pair of dG*s* (dG7, dG8) and termed **dG9** and **dG10** (S3 Table). The use of these oligonucleotides revealed no effects on translation, supporting the observation that in the used plant-derived translation system, antisense-like dG*s* need full complementarity to affect gene expression [39].

Additional **scrambled control** (dGsc) with a random sequence (S3 Table) was also shown to have no effect on luciferase activity (Fig 4). dG binding assays revealed high binding capacity of all tested antisense-like dGs that were complementary to pKS-A transcripts (S7 and S8 Figs). Although sense-like dGs shared the same sequence with the variant A of the TRβ1 5'UTR (pKS-A), they were able to bind RNA as well, however, with a lower capacity when compared to the antisense dGs. At the same time, the binding of the scrambled control was undetectable (S7 Fig). These results may confirm our assumption that sense dGs can bind, at least partially, to the distant inhibitory sequences of the TRβ1 5'UTR, releasing translation-enhancing elements normally blocked by secondary structures in a translationally less active transcripts (Fig 3c).

## Protein up-regulation induced by p16 5'UTR-specific dGoligos

To test whether our approach could be applied to enhance expression of another gene, we used published sequence data [21] as well as dGenhancer calculator to design dG*s*, specifically targeting $p16^{INK4a}$ 5'UTR (NCBI Ac.: NM_000077.4), a transcript of *CDKN2A* tumor suppressor. In this *in vitro* study, dG-mediated regulation of protein synthesis was tested using PCR-amplified linear expression construct containing T7 promoter, $p16^{INK4a}$ 5'UTR and the coding sequence of luciferase allowing for fast and reliable measurements of protein levels (S1 Appendix). The effects of each DNA-based dG*s* dG1p16-dG6p16 (S3 Table) were measured using coupled *in vitro* transcription-translation assay (Fig 5). Results from semi-quantitative real-time PCR, performed with luciferase specific primers (S4 Table), and measurements of luciferase activity revealed that negative control (dG-), scrambled control (dGscp16), dG5p16 and microRNA-like dG3p16 and dG4p16 had no effects neither on transcription nor translation efficiency. These results are in agreement with our data, including those showing no effects of microRNA-like DNA dGs in wheat germ lysates (Fig 4). Sense dG1p16 and antisense dG2p16 were designed on the basis of an element e1 (S1c Fig) containing an IRES sequence [21]. In samples supplemented with sense dG1p16 we observed unchanged transcription that was accompanied by strong translation-enhancing effect (4.78-fold, $p<0.001$). Similarly, dG2p16 elevated protein levels by 2.56-fold ($p<0.001$), however, this particular result could be a consequence of higher mRNA levels (1.3-fold, $p<0.001$). These results may indicate that apart from the explicit dG-mediated translation-enhancing effects, confirming findings obtained with TR$\beta$1 5'UTRs, some dG*s* can influence transcription machinery as well, thereby resembling the action of saRNAs [29, 30]. Using a combination of dG1p16 and dG6p16 (Fig 5) we observed over 12.30-fold increase in luciferase activity that is in accordance with previously observed effects triggered by a mixture of sense and antisense dGs: dG1+dG8 or dG1+dG4 targeting TR$\beta$1 5'UTRs. All the results were normalized to control (dG-). Data from three independent experiments were performed in triplicate and analyzed by ANOVA followed by Dunnett's multiple comparison test. *$p< 0.001$ vs. control.

**Fig 5. dGoligo-mediated upregulation of *CDKN2A* expression.** PCR-amplified linear luciferase expression construct containing 5'UTR of $p16^{INK4a}$ (*CDKN2A*) was generated (S1 Appendix) and used as a template in coupled transcription-translation assay that was performed as described in experiments

399 with TRβ1 5'UTRs. (**a**) Luciferase mRNA levels after 6-hour *in vitro* reaction of the linear construct with
400 a DNA-based dGoligos dG1p16 - dG6p16, dG1p16+dG6p16 or dGscp16 (S3 Table) targeting the $p16^{INK4a}$
401 5'UTR. (**b**) Luciferase activity as a measure of dG-mediated translational control. All data are shown as
402 mean % mRNA (a) or luciferase activity (b) ± SD. Data were analyzed by ANOVA followed by Dunnett's
403 multiple comparison test. *p< 0.001 vs. control.

## 404 dGoligo-mediated translation-enhancing effects in Caki-2 cells

405 To test our *in vitro* data in a cellular context, similar experiments were performed in Caki-2
406 cells using TRβ1 5'UTR A (pGL3-A) and appropriate dG*s* (Fig 6). Although unmodified
407 deoxyoligonucleotides can display some *in vivo* activity, they are subject to rapid degradation by
408 nucleases and are of limited utility in mammalian cells [40]. Therefore, we synthesized nuclease-
409 resistant, 2'-O-methyl RNA modified oligonucleotides, which do not activate the RNase H pathway [41].
410 Figure 6b shows that each dG differently regulated reporter protein synthesis. After transfection the
411 cells with the DNA-based dG*s* targeting variant A of TRβ1 5'UTR there was no significant effect on
412 translation of luciferase reporter protein (S9 Fig). By contrast, RNA-based dG*s* showed increased
413 translation efficiency between 1.7-2.1-fold (dG1, dG3), while the action of dG2 and dG4 resulted in
414 1.3-1.4-fold decrease in the reporter protein levels (Fig 6b). Surprisingly, antisense microRNA-like dG6,
415 which was previously used as a mismatched control in *in vitro* assay, resulted in 2.6-fold increase in
416 luciferase activity, whereas sense dG5 had no significant effect on translation when compared to
417 control without any dG ($p<0.01$). The similar effects were observed when using microRNA-like dG9 and
418 dG10 (1.09- and 4.8-fold respectively). These results showed a difference between the *in vitro* and *in*
419 *vivo* studies, wherein the TRβ1 5'UTR targeting microRNA-like dG*s* exerted the strongest enhancing
420 effects on translation in Caki-2 cells. The observed difference compelled us to introduce an additional
421 2'-O-methyl RNA modified scrambled control (dGsc) with an irrelevant (random) sequence that was
422 shown to produce no change in luciferase activity, thus, confirming the specificity of the *in vivo* dG
423 action. All the results of the luciferase activity after dG supplementation were normalized to control
424 pGL3-A plasmid (mock transfected group). Data from three independent experiments were performed

in triplicate. The Shapiro–Wilk test was used to determine normality of data distribution. Normally distributed data were analyzed using ANOVA followed by Dunnett's multiple comparison test.

**Fig 6. dGoligo-mediated gene expression changes in Caki-2 cells.** Effects of 2'-O-methyl RNA oligonucleotides on luciferase transcription (panel **a**) and translation (**b**) in Caki-2 cells transfected with pGL3-A. MicroRNA-like dG10 and microRNA-like dG6 exerted the strongest translation-enhancing effects in Caki-2 cells (4.83-fold and 2.60-fold respectively). Results from three independent experiments performed in triplicates are shown as mean % mRNA (a) or luciferase activity (b) ± SD. Data analyzed by ANOVA followed by Dunnett's multiple comparison test. *p<0.001 vs. control.

## Over-expression of endogenous TRβ1 proteins in dGoligo-treated cells

Translation-enhancing properties of selected dGs were confirmed in Caki-2 cells, where transfection with dGs resulted in increased levels of endogenous TRβ1 protein and its downstream target - type 1 iodothyronine deiodinase DIO1 [8, 23] (Fig 7). In this part of the study, Caki-2 cells were transfected with 2'-O-methyl-modified RNA-based dG6, dG10 or scrambled control – dGsc and cultured (without any plasmid) according to the procedure used in transcription and translation assay. dGs were selected on the basis of previously obtained results (Fig 6). Semi-quantitative real-time PCR for TRβ1 (exon 2-3) and DIO1, as well as western blot for TRβ1 and β-actin (Abcam plc, Cambridge, UK) were performed as described before [23]. Relative density of bands was quantified by densitometry and TRβ1 protein levels were normalized to β-actin. Figs 7b and 7d show that the most efficient enhancement of translation was achieved by action of antisense, microRNA-like dG10, which upregulated the endogenous TRβ1 protein synthesis by over 2.3-fold, whereas TRβ1 mRNA levels remained unchanged (p<0.001, Fig 7a). These results may provide evidence that translation of endogenous TRβ1 can be enhanced by dGs resulting in modification of the functional response, as evidenced by over 2.5-fold increase in expression of the DIO1 target gene (p<0.001, Fig 7c). Data from three independent experiments were performed in triplicate and shown as mean values ± SD. Statistics were calculated using t-test to compare cells transfected with dGs vs. dGsc. *p<0.001.

**Fig 7. Effects of selected dGs on expression of endogenous TRβ1 in Caki-2 cells.** Caki-2 cells were transfected with 2'-O-methyl-modified RNA-based dG6, dG10 or scrambled control – dGsc and cultured (without any plasmid) according to the procedure used in translation-enhancing assay (Materials and Methods). dGs were selected on the basis of previously obtained results (Fig 6). Expression of TRβ1 mRNA (**a**), protein (**b**) and DIO1 mRNA (**c**) are shown in upper panel. Semi-quantitative real-time PCR was performed for TRβ1 (exon 2-3) and DIO1, as described before [23]. Data from three independent experiments were performed in triplicate and shown as mean values ± SD. Statistics were calculated using t-test to compare cells transfected with dGs vs. dGsc. *p<0.001. (**d**) An example western blot for TRβ1 and β-actin is shown in lower panel. Each band (dGsc, dG10, dG6) represents sample combined from nine protein lysates. Relative density of bands was quantified by densitometry and TRβ1 protein levels were normalized to β-actin. (**e)** A simplified model of dG-mediated upregulation of endogenous TRβ1 protein, which has been demonstrated before to act as a transcription factor activating transcription of multiple genes including type 1 iodothyronine deiodinase (DIO1).

# Discussion

These studies demonstrate that specific enhancement of gene expression can be achieved at the level of protein translation. We found this phenomenon to be triggered in a specific manner by an exogenous synthetic small enhancing oligonucleotide - dGoligo (dG) targeting a specific 5'UTR *cis*-acting element.

## Targeting 5'UTRs could be a novel way to control protein levels

As previously demonstrated, alternative splicing of TRβ1 5'UTR variants is impaired in human clear cell renal cell carcinoma (ccRCC) and differential expression of multiple mRNA variants is accompanied by varying levels of TRβ1 protein expression [23]. Although the functional significance of these observations is not known in ccRCC, aberrant expression of alternative 5'UTRs has been shown to contribute to carcinogenesis mediated by silenced tumor suppressors [42] or activated oncogenes

[43]. In the light of complex secondary structures of low copy number TRβ1 5'UTRs including variant F [23] and evidence for selective protein synthesis of some alternatively spliced mRNA variants in oxygen deprived tumors or metastatic cancers [13, 16], it has been suggested that the sequence diversification of TRβ1 5'UTRs could play an important role in controlling *THRB* gene expression and this may influence tumor progression [23, 44]. Thus, the reported lack of correlation between the mRNA and TRβ1 protein levels in ccRCC [23] raised the hypothesis that the observed impairment may result, at least in part, from differing translational efficiencies of the TRβ1 5'UTR variants. This hypothesis is supported by the correlation observed between the *in vitro* translation efficiency of each 5'UTR and its Gibbs energy (Fig 1c), resulting in the aim to evaluate whether translation efficiency could be altered specifically by affecting folding Gibbs energy (ΔG). To investigate further, we used oligonucleotide–based *trans*–acting factors termed dGoligos (S3 Table) to selectively target TRβ1 5'UTRs and change the Gibbs energy-dependent secondary structure formation of the 5'UTRs (Fig 2). Since a misfolded conformation of mRNA *cis*-acting domains could result in either enhanced or reduced protein translation [4, 5], direct binding to these domains (in case of antisense-like dG*s*) or binding to distant *cis*-acting sequences folding with these domains (sense-like dGs) may enhance or repress protein synthesis (Fig 3). To find ΔG-dependent, translation-regulating domains we used a bespoke dGenhancer calculator, which allowed us to design the most effective, translation-enhancing dGs (S1 Appendix). However, this version of the calculator is unable to show ΔG-independent, functionally-active elements including IRESs, therefore it should be used together with other available databases of *cis*-acting elements.

## Strongly folded 5'UTRs have higher regulatory potential

In this study, the strong enhancement of translation was achieved by coupled action of sense dG1, designed to unblock a translation-enhancing element (e1), and antisense dG4, directly binding to an inhibitory region (e3, Fig 2). When both dG*s* were added, 1.77-fold and 6.58-fold increases in translation efficiency from weakly folded 5'UTR variant A and strongly folded variant F respectively were observed (Figs 4b and 4d). At the same time, the basal translation level (control without any dG) (Fig 1b) of variant A (24.09% of control) was 5.96-fold higher when compared to variant F (4.03% of control), suggesting that the folded variant F could possess higher translational regulatory potential

505 that was triggered by dG1 and dG4 (S5 Fig). These results suggested the hypothesis that mRNAs
506 containing strongly folded 5'UTRs may constitute a pool of translationally non-active or less-active
507 transcripts that could be recruited to translation through interaction with naturally occurring small
508 RNAs [18, 34], which may interfere with mRNA 5'UTRs in the same way as our dG*s*. This hypothesis is
509 supported by the previously reported observation that cellular microRNA miR-10a can interact with
510 the 5'UTR of mRNAs encoding ribosomal proteins that results in enhancement of their translation and
511 may be implicated in tumor invasion and metastasis [45]. Here, we showed translation-enhancing
512 effects of synthetic, TRβ1 5'UTR-specific, microRNA-like dGs (Fig 6), however, we also found naturally
513 occurring microRNA hsa-miR-211 to have target sites in both TRβ1 3'UTR and 5'UTR (exon 2/3, S6 Fig).
514 Furthermore, recent studies on 3'UTR-mediated gene silencing showed a correlation between
515 microRNAs targeting 3'UTRs and 5'UTR structures, which can recruit RNA helicase eIF4A2, a key factor
516 of eIF4F through which microRNAs function [18]. The authors have demonstrated that the eIF4A2
517 activity in 5'UTRs are critical for microRNA-mediated gene regulation as well as that mRNAs with
518 weakly folded 5'UTRs are refractory to microRNA repression [18]. This report and our current results
519 show that, in spite of low basal translation efficiency of mRNAs containing highly-structured 5'UTRs
520 (Fig 1b), these regions alone or together with 3'UTRs have higher translational regulatory potential
521 compared to unfolded 5'UTR variants (Fig 4b and 4d). It seems, therefore, that UTR-controlled
522 translation-enhancing or -silencing phenomena could be triggered in response to exposure to available
523 *trans*-acting factors that may lead not only to gene repression [9, 27] but also gene activation [30, 31].

524 ## dGoligos can lead to over-expression of selected proteins

525 *In vitro* results revealed that both, sense and antisense dG*s* can trigger translation-enhancing
526 effects that appear to be mostly dependent on a specific function of a 5'UTR *cis*-acting element. The
527 action of sense dG1 was thought to increase protein synthesis by releasing translation-enhancing
528 element e1 (Fig 3c) containing a putative IRES domain that has been identified before in the TRβ1-
529 5'UTRs [23] and tested in Caki-2 cells (S2 Fig). The translation-enhancing action of antisense dG4 and
530 dG8, which are complementary to a highly structured region containing multiple uAUGs, could be
531 explained by linearization of their target sites and blocking the uAUGs-rich region to prevent from
532 translation of alternative polypeptides (Fig 3c). This explanation is in agreement with a well-studied

533 model of the selective enhancement of ATF4 protein synthesis during integrated stress response (ISR)
534 [4, 13]. ISR can delay cap-dependent translation that makes uAUGs less attractive as start codons and
535 allows ribosomes to scan through the inhibitory uAUGs to find the correct codon of ATF4 [4]. Similarly,
536 antisense dG4 or dG8 could serve as a *trans*-acting factor making the TRβ1 uAUG-rich domain
537 inaccessible for the translation machinery, thus, facilitating the ribosomes to start the synthesis at the
538 correct AUG. These antisense-mediated effects might be supported by cap-independent translation,
539 initiated at the IRES domain [12] that is released by sense dG1. Indeed, the most efficient translation
540 was observed in the presence of sense dG1 and antisense dG8 that enhanced *in vitro* luciferase activity
541 over 28.1 and 55.8-fold for variant A and F, respectively (Figs 4b and 4d). At the same time,
542 transcription levels were noted to be unchanged (Fig 4a and 4b), suggesting that this regulation may
543 differ from recently described RNAa phenomenon resulting in up-regulation of gene transcription,
544 induced via promoter-specific activation [29, 30] or by promoter-directed antigene RNAs [31, 46].
545 In contrast to translation-enhancing dGs, their complementary partners (dG2, dG3 and dG7)
546 had no or opposite effects (Fig 4). Moreover, neither mismatched control dG5 nor control dG6 altered
547 protein levels in significant way (Fig 4b and 4d). Scrambled control with a random sequence was shown
548 to have no effect on luciferase activity as well (Figs 4b and 4d), thus, confirming target site-specific
549 action of dG1, dG4 and dG8.
550 In studying the translational regulatory potential of TRβ1 5'UTR variants we raised the question
551 whether the observed translation-enhancing effects triggered by dGs are universal or TRβ1-specific.
552 To check this out we designed dGs against the IRES identified within the $p16^{INK4a}$ 5'UTR (*CDKN2A*) [21]
553 and used the dG*enhancer* calculator to design dGs specifically targeting the ΔG-dependent, translation-
554 regulating elements within this 5'UTR (Appendix S1). The *CDKN2A* gene is frequently mutated or
555 deleted in a wide range of tumors and produces at least three alternatively spliced variants encoding
556 four distinct proteins [21]. An analysis of translation under the control of the $p16^{INK4a}$ 5'UTR, which was
557 incorporated into a PCR-amplified linear luciferase expression construct (Appendix S1) revealed a 4.78-
558 fold increase in protein levels and unchanged transcription rates after addition of dG1p16 (Fig 5). As
559 was found for the dG1 unblocking IRES oligo in TRβ1 5'UTR (Fig 3), the sense dG1p16 can enhance
560 translation via binding to distant sequences that may interact through complimentary base-pairing

with the IRES region of the $p16^{INK4a}$ 5'UTR. The strongest enhancing effects on luciferase activity (12.30-fold) were obtained by combining dG1p16 and dG6p16 (Fig 5) that is in accordance with previously observed in TRβ1 5'UTRs reactions translation-enhancing effects triggered by a mixture of sense dG1+dG8 or dG1+dG4 unblocking an IRES region (e1) and blocking translation-inhibitory element (e2 or e3). Although different constructs were used in this study, these results clearly confirm findings obtained *in vitro* with TRβ1 5'UTRs and show that dGs could be used as an universal tool controlling levels of selected proteins.

## microRNA-like dGoligos are more effective to enhance *in vivo* translation

These experiments were designed to investigate whether dG*s* can regulate protein translation in a cellular context. We used Caki-2 cells transfected with pGL3-derived plasmid [23] carrying TRβ1 5'UTR and downstream luciferase that allows for fast and reliable assessment of quickly changing translation rates in these cells after treatment with dGs. In contrast to results obtained *in vitro* with RNase H deficient wheat germ extracts, transfection of Caki-2 with DNA based dG*s* did not alter luciferase activity (S9 Fig), likely because unmodified deoxy-oligonucleotides are rapidly degraded by cellular nucleases [41], which can also switch off the translation in a non-specific way [15]. The use of 2'-O-methyl RNA -modified dG*s*, however, influenced the translation efficiency in these cells (Fig 6). Surprisingly, 2'-O-methyl modified, antisense, microRNA-like dGs: dG6 and dG10, containing a 3 nucleotide long insertion (loop) in the middle of their sequences (S3 Table), resulted in a greater than 2.6-fold and 4.8-fold increase in luciferase activity, respectively, whereas microRNA-like sense dG5 (complementary to dG6), sense dG9 (complementary to dG10) and scrambled control dGsc had no significant effect on the translation (Fig 6). All tested dGs did not affect mRNA levels suggesting that they could be involved specifically in translational control. These results are consistent with reports showing that some naturally occurring microRNAs can bind to 5'UTRs and regulate translation initiation [26, 45], however, their selectivity toward a single mRNA is thought to be low [27]. In contrast, synthetic micro-RNA like dGs with almost full complementarity to a target sequence and

reduced positions of potential G:U wobble base-pairing were shown to have high binding capacity and selectivity toward the complementary sequence (S7c Fig).

## dGoligo-treated cells can enhance translation of a native protein

Translation-enhancing properties of selected 2'-O-methyl-modified dGs were confirmed in Caki-2 cells on translation of endogenous TRβ1 protein that has been reported to be a transcription factor controlling transcription rates of type 1 iodothyronine deiodinase DIO1 [8]. DIO1 transcript, therefore, served as an estimate for TRβ1 transcription factor activity, which was expected to be dependent on the TRβ1 protein levels [8, 23]. Our experiments showed that the cells (without any plasmid) transfected with microRNA-like dG10 over-expressed the DIO1 mRNA by 2.5-fold that was accompanied by 2.3-fold enhancement in translation of the endogenous TRβ1 protein (Fig 7). It has also been shown that the levels of this protein can be elevated even more using alternative methods of the dGs delivery [47]. All tested dGs had no impact on TRβ1 mRNA levels, and treatment with scrambled control (dGsc) unchanged transcription and translation rates. Therefore, the elevated levels of DIO1 mRNA may indicate higher transcription factor activity of TRβ1 [8, 23] in the dG-treated cells (Fig 7e) and may provide evidence that dGs can affect the functional response of the living cells.

## dGoligo may interfere with machinery of translational control

Although the exact action of dGs remains unknown, it is clear that binding of these oligonucleotides can affect secondary and tertiary structures of a target sequence that may result in altering its translation regulating properties (Fig 3). This action is considered to trigger subsequent mechanisms leading to translation-enhancing or -silencing effects [18, 29].

Antisense DNA oligonucleotides (ASOs) are widely used to suppress gene expression by inducing RNase H-mediated mRNA degradation of the target mRNA [48]. The DNA/RNA heteroduplexes are subsequently targeted for endonucleolytic cleavage by the RNase H, however, previous observations suggest that ASOs, which are usually used to target a coding sequence, may result in RNase H-dependent generation of stable mRNA cleavage fragments without 5'-cap, followed by expression of truncated proteins. The lack of the 5'-cap structure could further be bypassed by the

613 cap-independent but 5’ end-dependent translation, initiated from an AUG start codon located a few
614 nucleotides downstream of the 5’ end of the RNA fragment [48]. This mechanism of translation was
615 observed *in vitro* and *in vivo*, albeit with severely reduced efficiency [48]. Translation of the cleavage
616 fragments may also occur via direct binding of ribosomes to internal RNA secondary structures (IRESs)
617 present on various cellular mRNAs, however, the IRES-mediated translation efficiency is condition-
618 dependent [5, 13] (S2 Fig). These findings provide a rationale for understanding the translation of
619 mRNA fragments generated by RNase H and could be considered *in vivo* as a potential mechanism of
620 action of small enhancing oligonucleotides. They, as other ASOs, may interfere with the RNAse H
621 pathway and subsequently generate RNA cleavage fragments [48] including transcripts with shorter,
622 less folded 5’UTRs. However, it was also elucidated, that 2’-O-methyl sugar modifications result in an
623 increased resistance to nuclease degradation [41, 49]. In addition, RNase H activity in wheat germ
624 lysates has been reported to be markedly reduced in comparison to other mammalian-based
625 translation systems [49]. Moreover, in our *in vitro* coupled transcription/translation experiments with
626 dG*s*, the levels of transcripts after 6-hour reactions were unchanged (S3 Fig), suggesting that, indeed,
627 RNAse H could not induce cleavage of dGoligo target sites and probably do not have strong impact on
628 the observed over 58-fold (dG1 and dG8) enhancement of translation efficiency in the used *in vitro*
629 system.

630 Comparing results from two different transcription-translation assays performed in the plant
631 cell-free lysates and human cells (Fig 4 and Fig 6), we considered whether dGs could be involved in
632 RNAi/RNAa related phenomena. Unlike mammalian microRNAs, plant microRNAs require nearly
633 perfect base pairing to induce the RNAi machinery [39]. Our results showed that neither microRNA-
634 like dG5 nor dG6 altered *in vitro* protein levels in significant way (Figs 4b and 4d), indicating that when
635 the assay is performed in the plant extract, a microRNA-like sequence loop introduced in the synthetic
636 dGs can block their action. On the contrary to fully complementary sense/antisense-like dGs that we
637 found to be the most effective in the plant system (Fig 4), the antisense microRNA-like dGs exerted the
638 strongest translation-enhancing effects in Caki-2 cells (Fig 6). These findings are in agreement with
639 distinct mechanisms of RNA interference in mammals and plants and could serve as an argument for
640 involvement of dG-5'UTR dimmers in some elements of this machinery. Although our assumption

641 needs to be studied in details, it can be supported by the known action of non-selective translation-
642 enhancing microRNAs including miR-122 [26] or miR-10a [45] and a link between microRNA targets in
643 3'UTRs and 5'UTR structures that are thought to play an essential role in RNAi [18]. Recently discovered
644 small activating RNAs (saRNAs) [28] can also trigger mechanisms leading to similar gene-enhancing
645 effects, however, unlike our single stranded translation-enhancing dG*s*, saRNAs have been shown to
646 be effective as double stranded transcription-activating molecules targeting promoter regions [29].
647

# 648 Conclusion

649 In summary, this work presents the first evidence for gene-specific translation-enhancing
650 effects triggered by small selective oligonucleotides termed dGoligos (dGs). These synthetic *trans*-
651 acting factors were originally designed to alter Gibbs energy-dependent secondary structure formation
652 of TRβ1 5'UTRs encoded by *THRB* suppressor gene. The applied approach allowed us for over 55.8-fold
653 translational enhancement of reporter protein when dG1 and dG8 were used in coupled *in vitro*
654 translation-transcription assay. Complementary *in vivo* study showed that dGs can enhance TRβ1-
655 5'UTR -mediated translation up to 4.8-fold. Interestingly, this assay showed that protein can be more
656 effectively synthesized when microRNA-like, 2'-O-methyl RNA antisense dG*s* were used. Furthermore,
657 dGenhancer calculator, which allowed us to determine targets within TRβ1 5'UTRs, was also
658 successfully used to design dGs enhancing translation of another *CDKN2A* tumor suppressor transcript,
659 thus confirming the universality and potential of dGs to over-express selected proteins. The concept
660 of this approach was based on our discovery that the most folded 5'UTR variants have higher
661 translational regulatory potential that can be released to enhance translation efficiency by the use of
662 specific dGs. They served as a molecular switch to translationally active conformation of the folded
663 5'UTRs. Taking together, this report would be the first showing a method for specific activation of
664 translation-enhancing elements of high regulatory potential. This strategy may complement other
665 available methods for gene expression regulation including gene silencing and may find its use in
666 enhancement of genes frequently silenced in cancers or even in biotechnology of recombinant
667 proteins.

# Materials and Methods

**Luciferase reporter constructs.** *In vitro* studies were performed with pBluescript-KS(+)-derived plasmid vectors containing different TRβ1 5'UTR variants (pKS-A,-B,-C,-D,-E,-F,-G) or irrelevant leader sequence lacking any TRβ1 UTR (pKS-control) [22]. 5'UTRs were subcloned upstream of the luciferase reporter gene [22]. For *in vivo* analyses, we used pGL3-derived plasmid, carrying variant A of TRβ1 5'UTR (pGL3-A) [22], which was found to be the most predominant in kidney cells [23]. pGL3-control (without TRβ1 5'UTR) served as a control plasmid [22].

**Prediction of translation-enhancing elements.** Two methods were used. *Manual* method allowed us to identify higher-order structures within 5'UTR *cis*-acting sequences (IRESs or uORFs stretches). Folding predictions from RNAstructure version 5.2, together with sequence analysis using NCBI tools were combined to select putative *cis*-acting elements containing the most stable secondary structures (the most negative $\Delta$G). As a second method, dGenhancer - an excel-based calculator was used to automatically identify putative ΔG-dependent translation-regulating elements within 5'UTR sequences (S1 Appendix). The algorithms of the calculator were constructed to visualize ΔG changes after *in silico* introduced single nucleotide substitutions (SNPs) of the 5'UTR sequences. These artificial SNPs differently affected overall sequence ΔGs (Gibbs energies) that were drawn by the dGenhancer to show regions where substitution can alter ΔGs the most, indicating putative *cis*-acting elements with the highest translational regulatory potential. The software that implements the calculations can be accessed here: http://www.serwer1448847.home.pl/biotechnology/dGenhancer.xlsx

**dGoligo synthesis.** Sense-, antisense- or microRNA-based DNA oligonucleotides were designed (S3 Table) to target *cis*-acting elements of TRβ1 5'UTRs (S1 Appendix). For *in vivo* studies nuclease-resistant 2'-O-methyl modified RNA oligonucleotides were synthesized. Oligonucleotides were performed with ABI 3900 High-Throughput DNA Synthesizer (Applied Biosystems, Foster City, CA) using standard DNA or 2'-O-methyl-modified phosphoramidites (Link Technologies, Lanarkshire, UK).

**Coupled *in vitro* transcription and translation assay.** 500ng of each plasmid were simultaneously transcribed and translated in 0.2mL-PCR tubes using RTS 100 Wheat Germ CECF Kit (Roche Diagnostics, Mannheim, Germany). The translation assay was conducted in 20μL of Reaction Solution, supplemented with 20uL of Feeding Solution after initial 3h-incubation. All reactions were maintained at 37°C for 6h with shaking at 600 rpm, using the RTS ProteoMaster Instrument (Roche Applied Science, Mannheim, Germany). After reaction, DNA levels of appropriate pKS plasmids (plasmid copy number per each reaction) were measured by semi-quantitative Real-Time PCR and served as internal controls of transcription and translation efficiency (S1 Appendix). mRNA levels were determined by semi-quantitative measurement of luciferase transcripts using **Real-Time PCR** (Quanti-Fast SYBR Green PCR Kit, Qiagen, Hilden, Germany) and two pairs of PCR primers (S4 Table). The reactions were performed with LightCycler® 480 (Roche, Germany) under standard conditions shown in Materials and Methods in SM. *In vitro* **translation-enhancing assay** was performed with 500ng of pKS-A, pKS-F and pKS-control constructs were expressed as above in the presence of 0,25μM each tested dGoligo (S3 Table) or in the absence of dGoligo (dG-). For normalization, the results were divided by corresponding results obtained for pKS-control, to eliminate any possible non-specific dGoligo effects. Translation efficiency was determined by the use of **Luciferase Reporter Gene Assay** (Promega, Madison, WI) with the Synergy2 luminometer (BioTek, Winooski, VT) in conditions recommended by the manufacturers.

**Cell-culture based, *in vivo* transcription and translation assay.** The human clear cell renal carcinoma cell line (Caki-2) was used (American Type Culture Collection, Manassas, VA). Caki-2 cells were grown in McCoy's 5A medium with L-glutamine (Gibco/Invitrogen, Carlsbad, Ca) with 10% fetal bovine serum (FBS; Sigma-Aldrich, Saint Louis, MO) and 1x penicillin-streptomycin solution (Sigma-Aldrich, Saint Louis, MO). The cells were maintained at 37°C in 5% $CO_2$ atmosphere. For all the experiments, Caki-2 cells were seeded into 75cm$^2$ bottles, 6- or 12-well culture plates at density 13x10$^3$ cells/cm$^2$, 24h before transfection. Three independent *in vivo* experiments were performed in triplicate.

**Luciferase expressing plasmids and dGoligo transfection**. 24 hours after seeding, cells were transfected with 100 ng of control pRL-TK (Promega, Madison, WI) and 1ug of pGL3-A plasmid [22], using 1μg/ul PEI (Linear Polyethylenimine, Polysciences Inc., Warrington, PA) and 150mM NaCl in FBS-

free McCoy's medium. Five hours after transfection, the medium was replaced with McCoy's medium plus 10% FBS. PEI-mediated transfection reactions contained 36nM of each dG and was carried overnight. The medium was then replaced with McCoy's medium plus 10% FBS and 1x penicillin-streptomycin solution. 24h after the last medium replacement, cells were harvested. The cells were divided into two equal parts for isolation of total RNA and luciferase protein. The RNA was processed as described below. **Dual-luciferase assay**. The protein measurements were performed using dual-luciferase assay (Promega, Madison, WI) in the Synergy2 luminometer (BioTek, Winooski, VT), according to the manufacturer's instructions.

**Cellular RNA isolation.** Total RNA for real-time PCR was purified from the second part of the collected cells as it was described for *in vitro* assay.

**Reverse transcription and Semi-Quantitative Real-time PCR.** Reverse transcription and Real-time PCR of luciferase pGL3-A and pRL-TK control was performed according to the protocol used for *in vitro* study. The transcript levels of Firefly luciferase were compared with *Renilla* using specific primers (S4 Table). Relative changes in gene expression were calculated using the $2^{(-\Delta\Delta Ct)}$.

**dGoligo controls.** All dG*s* were tested as complementary sense and antisense sequences (S3 Table). dG5, dG6, dG7 and dG8 were synthesized as mismatched controls containing a 3 nucleotide-long mismatched insertion in the middle of the oligonucleotides (Fig 2). An additional scrambled control oligonucleotide (dGsc) with an irrelevant (random) sequence was as designed with GeneScript software (S3 Table).

**Bioinformatic analysis.** Total Gibbs energy prediction ($\Delta G$=ΔH–TΔS) of 5'UTR secondary structures was performed using RNAstructure version 5.2 [37]. NCBI-BLASTN program and IRESite database [38] were used for comparative sequence analysis towards evolutionary conserved 5'UTR domains such as IRES consensus sequences. The dGenhancer calculator was used to determine translation regulating elements (S1 Appendix).

**Statistics:** At least three independent experiments were carried out for each assay and measured in triplicate. Normality of data distribution was estimated using Shapiro-Wilk test and in each case data

were analyzed by ANOVA followed by Dunnett's multiple comparison test. $p<0.001$ was considered statistically significant. Correlation of Gibbs energy and translation efficiency (Fig 1c) was estimated by r-squared value of the Pearson product-moment correlation coefficient. Logarithmically transformed data of translation efficiency were analyzed with the Gibbs energies by linear regression. $p<0.05$ was considered statistically significant.

# Acknowledgments

The authors thank the Polish State Committee for Scientific Research and the Centre of Postgraduate Medical Education in Warsaw for financial support of the research.

**Author Contributions**

Conceived and designed the experiments: AM, AN. Performed the experiments: AM, AW, KG, PP, AN. Analyzed the data: AM, AW, AN. Contributed reagents/materials/analysis tools: AM, GRW, AN. Wrote the paper: AM, AW, GRW, AN.

# References

1. Van Der Kelen K, Beyaert R, Inzé D, De Veylder L. Translational control of eukaryotic gene expression. Crit Rev Biochem Mol Biol. 2009 Jul-Aug;44(4):143-68.
2. Myasnikov AG, Simonetti A, Marzi S, Klaholz BP. Structure-function insights into prokaryotic and eukaryotic translation initiation. Curr Opin Struct Biol. 2009 Jun;19(3):300-9.
3. Mahen EM, Watson PY, Cottrell JW, Fedor MJ. mRNA secondary structures fold sequentially but exchange rapidly in vivo. PLoS Biol. 2010 Feb 9;8(2):e1000307.
4. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci U S A. 2004 Aug 3;101(31):11269-74.
5. Wellensiek BP, Larsen AC, Stephens B, Kukurba K, Waern K, Briones N, et al. Genome-wide profiling of human cap-independent translation-enhancing elements. Nat Methods. 2013 Aug;10(8):747-50. doi: 10.1038/nmeth.2522.
6. Chatterjee S, Pal JK. Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. Biol Cell. 2009 May;101(5):251-62.

7. Loreni F, Mancino M, Biffo S. Translation factors and ribosomal proteins control tumor onset and progression: how? Oncogene. 2014 Apr 24;33(17):2145-56.

8. Boguslawska J, Wojcicka A, Piekielko-Witkowska A, Master A, Nauman A. MiR-224 targets the 3'UTR of type 1 5'-iodothyronine deiodinase possibly contributing to tissue hypothyroidism in renal cancer. PLoS One. 2011;6(9):e24541.

9. Kenski DM, Butora G, Willingham AT, Cooper AJ, Fu W, et al. siRNA-optimized Modifications for Enhanced In Vivo Activity. Mol Ther Nucleic Acids. 2012 Jan 24;1:e5.

10. Sud R, Geller ET, Schellenberg GD. Antisense-mediated Exon Skipping Decreases Tau Protein Expression: A Potential Therapy for Tauopathies. Mol Ther Nucleic Acids. 2014 Oct 21;3:e204.

11. Richter JD, Sonenberg N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature. 2005 Feb 3;433(7025):477-80.

12. Komar AA, Hatzoglou M. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle. 2011 Jan 15;10(2):229-40.

13. Master A, Nauman A. Molecular mechanisms of protein biosynthesis initiation - biochemical and biomedical implications of a new model of translation enhanced by the RNA hypoxia response element (rHRE). Postepy Biochem. 2014; 60(1):39-54.

14. Pyronnet S, Dostie J, Sonenberg N. Suppression of cap-dependent translation in mitosis. Genes Dev. 2001 Aug 15;15(16):2083-93.

15. Clemens MJ. Translational regulation in cell stress and apoptosis. Roles of the eIF4E binding proteins. J Cell Mol Med. 2001 Jul-Sep;5(3):221-39.

16. Van den Beucken T, Koritzinsky M, Wouters BG. Translational control of gene expression during hypoxia. Cancer Biol Ther. 2006 Jul;5(7):749-55. Epub 2006 Jul 1.

17. Biro JC. Correlation between nucleotide composition and folding energy of coding sequences with special attention to wobble bases. Theor Biol Med Model.2008 Jul 29;5:14.

18. Meijer HA, Kong YW, Lu WT, Wilczynska A, Spriggs RV, Robinson SW, et al. Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. Science. 2013 Apr 5;340(6128):82-5.

19. Master A, Nauman A. THRB (Thyroid Hormone Receptor, Beta). Atlas Genet Cytogenet Oncol Haematol. 2014; 18(6):400-433.

20. Martínez-Iglesias O, Garcia-Silva S, Tenbaum SP, Regadera J, Larcher F, Paramio JM, et al. Thyroid hormone receptor beta1 acts as a potent suppressor of tumor invasiveness and metastasis, Cancer Res. 69 (2009) 501–509.

21. Bisio A, Latorre E, Andreotti V, Bressac-de Paillerets B, Harland M, Scarra GB, Ghiorzo P, Spitale RC, Provenzani A, Inga A. The 5'-untranslated region of p16INK4a melanoma tumor suppressor acts as a cellular IRES, controlling mRNA translation under hypoxia through YBX1 binding. Oncotarget. 2015 Nov 24;6(37):39980-94.

22. Frankton S, Harvey CB, Gleason LM, Fadel A, Williams GR. Multiple messenger ribonucleic acid variants regulate cell-specific expression of human thyroid hormone receptor beta1. Mol Endocrinol. 2004 Jul;18(7):1631-42.

23. Master A, Wojcicka A, Piekielko-Witkowska A, Boguslawska J, Poplawski P, Tanski Z, et al. Untranslated regions of thyroid hormone receptor beta 1 mRNA are impaired in human clear cell renal cell carcinoma. Biochim Biophys Acta. 2010 Nov;1802(11):995-1005.

24. Vasudevan S, Tong Y, Steitz JA. Switching from repression to activation: microRNAs can up-regulate translation. Science. 2007 Dec 21;318(5858):1931-4.

25. Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, et al. New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. Genome Res. 2009 Jul;19(7):1175-83.

26. Henke JI, Goergen D, Zheng J, Song Y, Schüttler CG, Fehr C, et al. microRNA-122 stimulates translation of hepatitis C virus RNA. EMBO J. 2008 Dec 17;27(24):3300-10.

27. Birmingham A, Selfors LM, Forster T, Wrobel D, Kennedy CJ, Shanks E, et al. Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods. 2009 Aug;6(8):569-75.

28. Jayaswal V, Lutherborrow M, Ma DD, Yang YH. Identification of microRNA-mRNA modules using microarray data. BMC Genomics. 2011 Mar 6;12:138.

29. Li LC, Okino ST, Zhao H, Pookot D, Place RF, Urakami S, et al. Small dsRNAs induce transcriptional activation in human cells. Proc Natl Acad Sci U S A. 2006 Nov 14;103(46):17337-42.

30. Huang V, Qin Y, Wang J, Wang X, Place RF, Lin G, et al. RNAa is conserved in mammalian cells. PLoS One. 2010 Jan 22;5(1):e8848.

31. Schwartz JC, Younger ST, Nguyen NB, Hardy DB, Monia BP, Corey DR, et al. Antisense transcripts are targets for activating small RNAs. Nat Struct Mol Biol. 2008 Aug;15(8):842-8.

32. Master A, Nauman A. Gene expression regulation by long naturally occurring antisense transcripts. Post. Biol. Kom. 2014;41(1):3-28.

33. Schwartz JC, Younger ST, Nguyen NB, Hardy DB, Monia BP, Corey DR, Janowski BA. Antisense transcripts are targets for activating small RNAs. Nat Struct Mol Biol. 2008 Aug;15(8):842-8.

34. Master A, Nauman A. Genomic context and expression regulation of nuclear thyroid hormone receptors by long naturally occurring antisense transcripts. Post. Biol. Kom. 2014; 41(1):29-58.

35. Xia X, Holcik M. Strong eukaryotic IRESs have weak secondary structure. PLoS One. 2009;4(1):e4136.

36. Bugaut A, Balasubramanian S. 5'-UTR RNA G-quadruplexes: translation regulation and targeting. Nucleic Acids Res. 2012 Jun;40(11):4727-41.

37. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics. 2010 Mar 15;11:129.

38. Mokrejs M, Masek T, Vopálensky V, Hlubucek P, Delbos P, et al. IRESite--a tool for the examination of viral and cellular internal ribosome entry sites. Nucleic Acids Res. 2010 Jan;38(Database issue):D131-6.

39. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell. 2002 Aug 23;110(4):513-20.

40. Wu H, Lima WF, Zhang H, Fan A, Sun H, Crooke ST. Determination of the role of the human RNase H1 in the pharmacology of DNA-like antisense drugs. J Biol Chem. 2004 Apr 23;279(17):17181-9.

41. Schneider PN, Olthoff JT, Matthews AJ, Houston DW. Use of fully modified 2'-O-methyl antisense oligos for loss-of-function studies in vertebrate embryos. Genesis. 2011 Mar;49(3):117-23.

42. Kim WG, Zhu X, Kim DW, Zhang L, Kebebew E, Cheng SY. Reactivation of the silenced thyroid hormone receptor β gene expression delays thyroid tumor progression. Endocrinology. 2013 Jan;154(1):25-35.

43. Huret JL, Ahmad M, Arsaban M, Bernheim A, Cigna J, Desangles F, et al. Atlas of genetics and cytogenetics in oncology and haematology in 2013. Nucleic Acids Res. 2013 Jan;41(Database issue):D920-4.

44. Kim WG, Zhao L, Kim DW, Willingham MC, Cheng SY. Inhibition of tumorigenesis by the thyroid hormone receptor β in xenograft models. Thyroid. 2014 Feb;24(2):260-9.

45. Ørom UA, Nielsen FC, Lund AH. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell. 2008 May 23;30(4):460-71.

46. Morris KV, Santoso S, Turner AM, Pastori C, Hawkins PG. Bidirectional transcription directs both transcriptional gene activation and suppression in human cells. PLoS Genet. 2008 Nov;4(11):e1000258.

47. Master A, Wojcicka A, Nauman A. The 5'UTR-dependent enhancement of protein translation efficiency triggered by self-transfecting 3'-aminoallyl-containing oligonucleotides (aa-dGoligos) targeting a pool of strongly folded transcript variants of the THRB suppressor gene. Mol Ther. 2014 May; 22: S31-S38; 10.1038/mt.2014.64.

48. Wu H, Lima WF, Zhang H, Fan A, Sun H, Crooke ST. Determination of the role of the human RNase H1 in the pharmacology of DNA-like antisense drugs. J Biol Chem. 2004 Apr 23;279(17):17181-9.

49. Cazenave C, Frank P, Büsen W. Characterization of ribonuclease H activities present in two cell-free protein synthesizing systems, the wheat germ extract and the rabbit reticulocyte lysate. Biochimie. 1993;75(1-2):113-22.

# Supporting Information

**S1 Appendix. Supporting Materials and Methods.**

**S1 Table. Basic characteristics of selected TRβ1 5'UTR variants A-G.**

**S2 Table. Prediction of translational regulatory potential of 5'UTRs.**

868 **S3 Table. List of dGoligos (dGs) used in the study.**

869 **S4 Table. List of primers used in Real-Time and classic PCR.**

870 **S1 Fig. Folding of TRβ1 5'UTRs.**

871 **S2 Fig. Translation-enhancing, IRES-like element in TRβ1 5'UTR.**

872 **S3 Fig. Time-course of protein synthesis rates in RTS 100 Wheat Germ CECF system.**

873 **S4 Fig. Proposed folding patterns of TRβ1 5'UTR after dGoligo supplementation.**

874 **S5 Fig. Change in translation efficiency after dG1 and dG4 supplementation.**

875 **S6 Fig. Potential hsa-miR-211 target sites within TRβ1 3'UTR and 5'UTR.**

876 **S7 Fig. dGoligo binding capacity.**

877 **S8 Fig. Binding selectivity confirmed by dGoligo-primed reverse transcription.**

878 **S9 Fig. DNA-based dGoligo-mediated effects in Caki-2 cells.**

# Supporting Information

## S1 Appendix. Supporting Materials and Methods.

**Prediction of translation-regulating elements.** An excel-based calculator - dGenhancer was used to search for putative 5'UTR *cis*-acting elements, which functional activity could be determined by Gibbs energy-dependent secondary structure formation. Prediction of total Gibbs energies (ΔG=ΔH–TΔS) of the 5'UTR structures was performed using RNAstructure version 5.2 [s1]. These ΔGs were treated as input data for dGenhancer calculations showing the strongest translation-regulating signal (high peak) at nucleotides 130 and 133 located in the middle of exon 2a of TRβ1 5'UTRs (see print screens below).

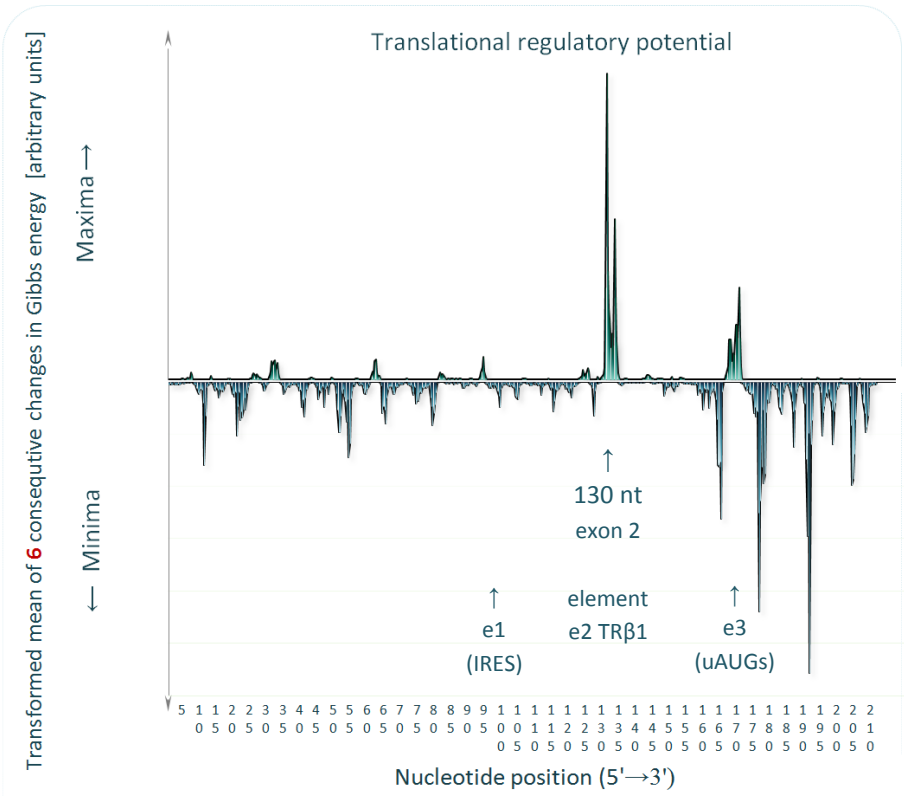
All annotations and formulae are included in the calculator available under the following link: http://www.serwer1448847.home.pl/biotechnology/dGenhancer.xlsx

The dGenhancer can show ΔG changes observed among 5'UTR sequences containing *virtual* SNPs (red) that were substituted base by base *in silico* in each nucleotide position of the 5'UTRs, as it is shown below for two exemplary 5'UTR bases (green).

1 2 3 4 5 6 7 8 9 ... <--- nucleotide positions ($nt_s$) within a fragment sequence of TRβ1 variant A

5'UTR-...**A** G A G C C C G C ...-3' ΔG=-68,30 [kcal/mol]
5'UTR-...C G A G C C C G C ...-3' ΔG=-69,30 [kcal/mol]
5'UTR-...T G A G C C C G C ...-3' ΔG=-66,30 [kcal/mol]
5'UTR-...G G A G C C C G C ...-3' ΔG=-70,30 [kcal/mol]
5'UTR-...A **G** A G C C C G C ...-3' ΔG=-71,30 [kcal/mol]
5'UTR-...A A A G C C C G C ...-3' ΔG=-70,30 [kcal/mol]
5'UTR-...A C A G C C C G C ...-3' ΔG=-70,10 [kcal/mol]
5'UTR-...A T A G C C C G C ...-3' ΔG=-65,30 [kcal/mol] ... and so on ... for all 5'UTR bases

As a result, the calculator makes a graph presenting nucleotide stretches (elements), which substitution can change the total 5'UTR Gibbs energy the most, thereby indicating regions that could be characterized by the highest potential to regulate protein synthesis (translational regulatory potential). Oligonucleotide-based *trans*-acting factors (termed here dGoligos, dGs), which are designed to selectively bind to these 5'UTR regions, could block or release their translation -silencing or -enhancing elements. As chemically synthesized siRNAs and ASOs, dGs are highly sequence-specific nucleic acid molecules, but on the contrary to the gene-silencing oligonucleotides, allow for specific binding to their target sequence followed by selective enhancement of protein synthesis.

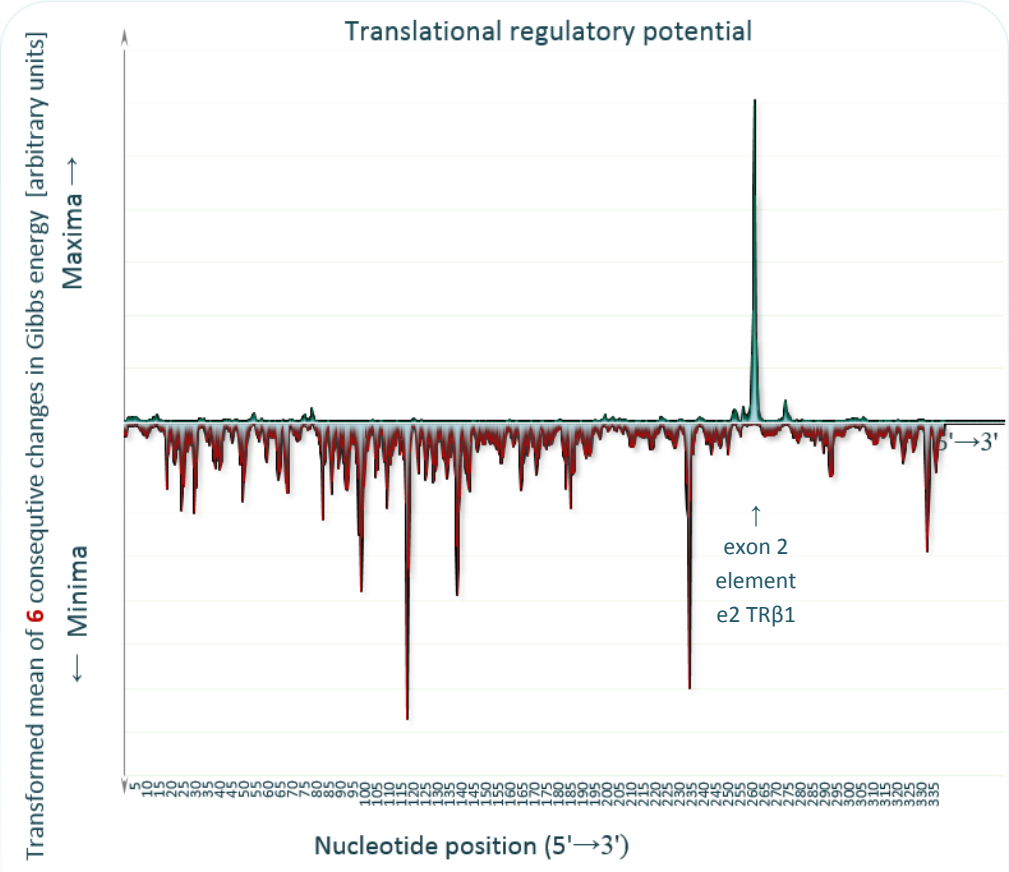



← (**a**) a graph generated by dGenhancer. Variant A of TRβ1 5'UTR ΔG= -67,8 [kcal/mol].

Maxima and Minima may indicate translation-regulating elements that are specifically dependent on a sequence folding state.

Maxima ~ putative translation-silencing elements that can inhibit translation in a normal folding state of a 5'UTR (...should be blocked to enhance translation).

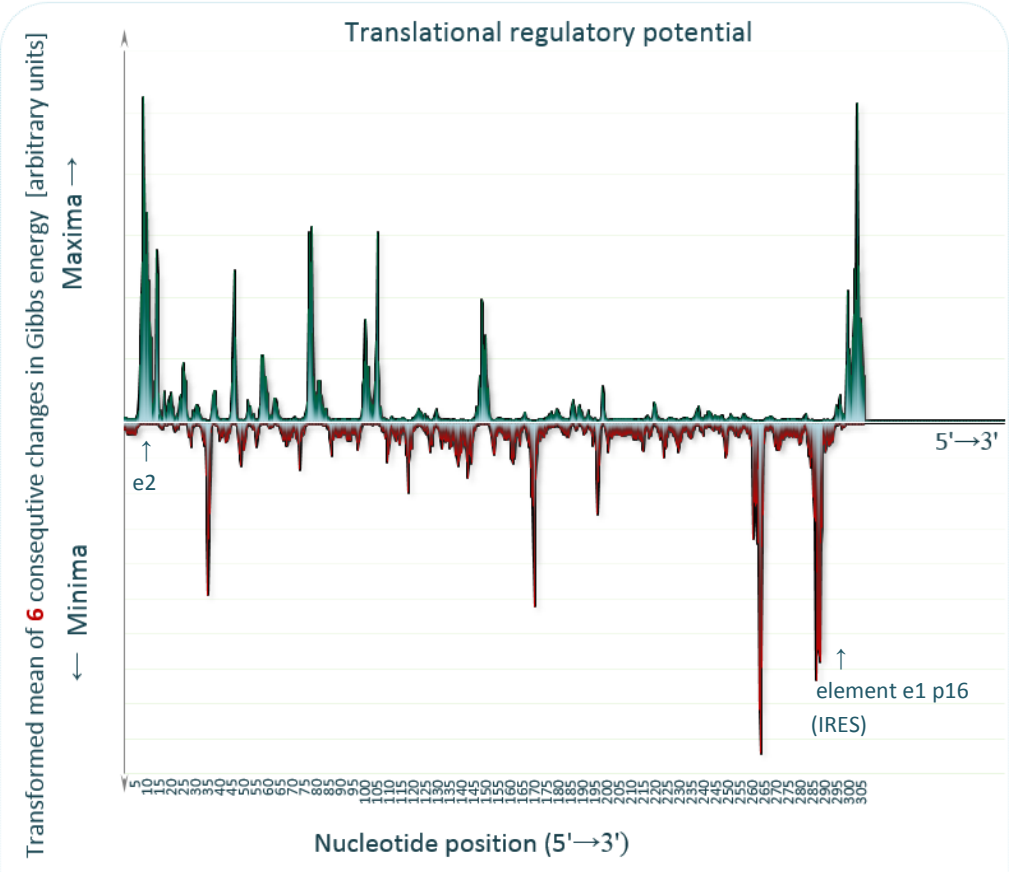
Minima ~ putative translation-enhancing elements that can elevate translation in a normal folding state of a 5'UTR (...should be released to enhance translation).




← (**b**) a graph generated by dGenhancer. Variant F of TRβ1 5'UTR.

ΔG= - 125,3[kcal/mol].

E = Max ΔG - n*ΔG = 6,40 [kcal/mol] ~ susceptibility to translation enhancement (more info. in section d).




← (**c**) a graph generated by dGenhancer.

5'UTR of variant p16INK4a (CDKN2A).

ΔG= - 146,4 [kcal/mol],

E = Max ΔG - n*ΔG = 7,10 [kcal/mol].

(**d**) Selected print screens of dGenhancer calculations (Variant A of TRβ1 5'UTR).

All annotations and formulae are included in the dGenhancer calculator.

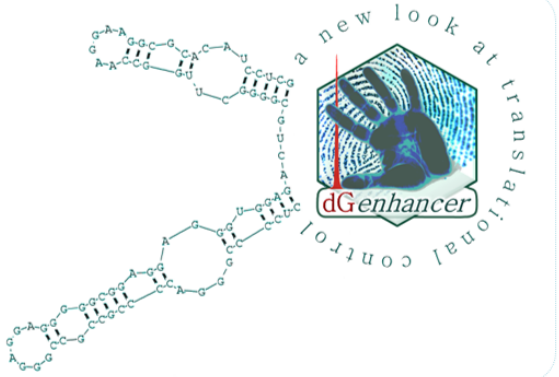
| Put your data in the table | Tested sequence (D3) | Sequence length [nt] |
|---|---|---|
| Paste your sequence as a text | AAGAGCCCGCAGGCTACCTTCCCCGGG CAGGGGCGCTCAACCCAACCGGCTCCA | 211 |
| Sample name | Variant A of TRbeta1 5'UTR | |
| Go to [2-sort the results] | OK | |

Short instruction (with internal links):

1. Paste the sequence of your interest into D3 cell. The sequence may have no gaps or other signs (only A G C and T are allowed)

2. Copy (as text) each sequence from column M and calculate value of Gibbs energy for each one (use any available software to calculate the gibbs energy), then clear old (exemplary) data in column P and paste your new data in cells of column P.

3. Sort the results (2-...) and take your results (3-...).




2016-03-26 22:41

| Analysis parameters: | | Reference parameters |
|---|---|---|
| First substituting nucleotide | A | A, G, C, or T |
| Second substituting nucleotide | C | A, G, C, or T |
| Third substituting nucleotide | T | A, G, C, or T |
| Fourth substituting nucleotide | G | A, G, C, or T |
| Number of substitutions per base | 3 | 3 or 4 |
| Exponent | 140 | 20-200 |
| Enhancing (E) or Silencing (S) effects | E | E or S |
| Length of designed dGoligo | 21 | any |
| SNP position within sense-like dGoligo | 11 | any |
| Loop position within microRNA-like dG | 11 | any |
| Sequence of the loop | ttt | any |

OK

Take your results

Nucleotide position ↓ ↓ Base of the tested seq. P = ΔGs ↓

↓ Artificial substitution

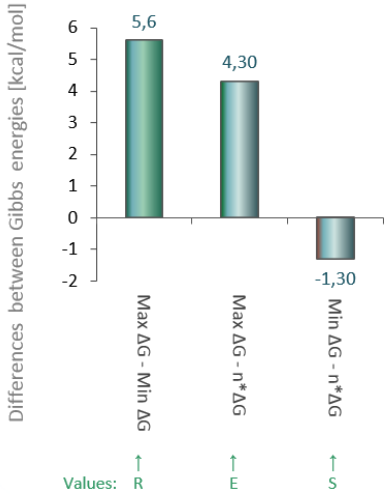
| | 5' | 5' | M | P |
|---|---|---|---|---|
| 1 | A | A | AAGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 1 | A | C | CAGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 1 | A | T | TAGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 1 | A | G | GAGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 2 | A | C | ACGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 2 | A | T | ATGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 2 | A | G | AGGAGCCCGCAGGCTACCTTCC ... | -69,70 |
| 3 | G | A | AAAAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 3 | G | C | AACAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 3 | G | T | AATAGCCCGCAGGCTACCTTCC ... | -68,00 |
| 4 | A | C | AAGCGCCCGCAGGCTACCTTCC ... | -67,50 |
| 4 | A | T | AAGTGCCCGCAGGCTACCTTCC ... | -68,40 |
| 4 | A | G | AAGGGCCCGCAGGCTACCTTCC ... | -68,40 |
| 17 | C | G | AAGAGCCCGCAGGCTAGCTTCC ... | -68,50 |
| 17 | C | A | AAGAGCCCGCAGGCTAACTTCC ... | -66,30 |
| 18 | C | T | AAGAGCCCGCAGGCTACTTTCC ... | -67,50 |
| 18 | C | G | AAGAGCCCGCAGGCTACGTTCC ... | -68,00 |
| 18 | C | A | AAGAGCCCGCAGGCTACATTCC ... | -68,10 |
| 19 | T | G | AAGAGCCCGCAGGCTACCGTCC ... | -69,40 |
| 19 | T | A | AAGAGCCCGCAGGCTACCATCC ... | -66,40 |
| 19 | T | C | AAGAGCCCGCAGGCTACCCTCC ... | -70,20 |
| 20 | T | G | AAGAGCCCGCAGGCTACCTGCC ... | -71,40 |
| 20 | T | A | AAGAGCCCGCAGGCTACCTACC ... | -67,80 |
| 20 | T | C | AAGAGCCCGCAGGCTACCTCCC ... | -68,80 |
| 21 | C | T | AAGAGCCCGCAGGCTACCTTTC ... | -68,50 |
| 21 | C | G | AAGAGCCCGCAGGCTACCTTGC ... | -71,20 |
| 21 | C | A | AAGAGCCCGCAGGCTACCTTAC ... | -67,80 |
| 22 | C | T | AAGAGCCCGCAGGCTACCTTCT ... | -68,80 |
| 23 | C | T | AAGAGCCCGCAGGCTACCTTCC ... | -69,10 |
| 23 | C | G | AAGAGCCCGCAGGCTACCTTCC ... | -70,40 |
| 23 | C | A | AAGAGCCCGCAGGCTACCTTCC ... | -66,10 |
| 24 | C | T | AAGAGCCCGCAGGCTACCTTCC ... | -70,60 |
| 24 | C | G | AAGAGCCCGCAGGCTACCTTCC ... | -69,10 |
| 24 | C | A | AAGAGCCCGCAGGCTACCTTCC ... | -65,50 |
| 25 | G | A | AAGAGCCCGCAGGCTACCTTCC ... | -64,90 |
| 25 | G | C | AAGAGCCCGCAGGCTACCTTCC ... | -69,10 |
| 25 | G | T | AAGAGCCCGCAGGCTACCTTCC ... | -66,90 |
| 26 | G | A | AAGAGCCCGCAGGCTACCTTCC ... | -65,80 |
| 26 | G | C | AAGAGCCCGCAGGCTACCTTCC ... | -68,40 |
| 26 | G | T | AAGAGCCCGCAGGCTACCTTCC ... | -66,50 |
| 27 | G | A | AAGAGCCCGCAGGCTACCTTCC ... | -66,60 |
| 27 | G | C | AAGAGCCCGCAGGCTACCTTCC ... | -66,60 |
| 27 | G | T | AAGAGCCCGCAGGCTACCTTCC ... | -68,40 |
| 28 | C | T | AAGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 28 | C | G | AAGAGCCCGCAGGCTACCTTCC ... | -68,00 |
| 28 | C | A | AAGAGCCCGCAGGCTACCTTCC ... | -67,80 |
| 29 | A | C | AAGAGCCCGCAGGCTACCTTCC ... | -71,20 |
| 29 | A | T | AAGAGCCCGCAGGCTACCTTCC ... | -69,40 |
| 29 | A | G | AAGAGCCCGCAGGCTACCTTCC ... | -68,30 |
| 30 | G | A | AAGAGCCCGCAGGCTACCTTCC ... | -65,10 |
| 30 | G | C | AAGAGCCCGCAGGCTACCTTCC ... | -67,40 |
| 30 | G | T | AAGAGCCCGCAGGCTACCTTCC ... | -66,40 |

## Sequences of the most importance

| | Sense dGoligo covering SNP (in the middle) | Sense, microRNA-like dGoligo containing a loop located in SNP | Nucleotide position |
|---|---|---|---|
| Maximum ΔG [kcal/mol]: | -63,5 | | |
| Forward: | GGCTGTCCTGCGTGGGTGCCA | GGCTGTCCTGCtttGTGGGTGCCA | 130 |
| Reversed: | TGGCACCCACGCAGGACAGCC | TGGCACCCACAAAGCAGGACAGCC | 130 |
| Minimum ΔG [kcal/mol]: | -69,1 | | |
| Forward: | AGTTCCACACATGATTTAATG | AGTTCCACACAtttTGATTTAATG | 151 |
| Reversed: | CATTAAATCATGTGTGGAACT | CATTAAATCAAAATGTGTGGAACT | 151 |

n*ΔG of "native" seq. [kcal/mol]: -67,80

## Change in Gibbs energy




Max ΔG = max. value of Gibbs energy.
Maximum of Gibbs energy after substitution of a nucleotide (virtual SNP). The least folded sequence.
Less negative value.
~ the state that can facilite the translation initiation.

Min ΔG = min. value of Gibbs energy.
Minimum of Gibbs energy after substitution of a nucleotide (virtual SNP). The most strongly folded sequence.
More negative value.
~ the state that can inhibit the translation initiation.

n*ΔG
Value of Gibbs energy of non-substituted, native sequence.

G = Max ΔG - Min ΔG
~ susceptibility to ΔG-dependent translation regulation.

E = Max ΔG - n*ΔG ≈ how much the SNP could enhance translation efficiency,
~ susceptibility to dG-mediated translation enhancement.

S = Min ΔG - n*ΔG ≈ how much the SNP could silence translation efficiency
~ susceptibility to dG-mediated translation silencing.

**dGoligo design and synthesis.** dGs were synthesized as a structurally diverse group of sense-, antisense- or microRNA-like DNA oligonucleotides (S3 Table). dGs were designed to target the most stable (showing the most negative ΔG) secondary structures of indicated *cis*-acting elements of TRβ1 5'UTRs, thus the primary function of synthetic dGs was to change the *Gibbs energy-dependent secondary structure formation* [s2, s3]. Except for a short 3-nt loop structure in microRNA-like dGs (dG5, dG6, dG9, dG10), the oligonucleotides share full homology with human TRβ1 mRNA sequence (NCBI GeneBank Acc. No. NM_000461), 5'UTR variant A (GeneBank Acc. No. AY286465.1) and 5'UTR variant F (GeneBank Acc. No. AY286470.1). dGs were expected to target one of the sequences: a) element e1 containing a putative IRES site (Master et al. 2010) located on exon 1c/2a junction (dG1, 2, 5, 6), b) element e3 - a sequence conserved among all TRβ1 5'UTR variants, containing multiple alternative AUGs (Fig 2), located on exon 2a/3 junction (dG3, 4) or c) a target site detected automatically with dGenhancer calculator in the middle of exon 2 (dG7, 8, 9 and 10). All dG*s* were designed as pairs of a) antisense strand (dG2, 4, 6, 8, 10) directly recognizing the indicated regulatory sequence (IRES, uAUG or dGenhancer-detected translation regulating element) on the TRβ1 5'UTR and b) sense strand (dG1, 3, 5, 7, 9) releasing the indicated region by binding to a sequence that folds with these regions. (S3 Table and Fig 2). All oligonucleotides were synthesized on ABI 3900 High-Throughput DNA Synthesizer (Applied Biosystems, Foster City, CA) using standard DNA phosphoramidites or 2'-O-methyl modified RNA phosphoramidites (Link Technologies, Lanarkshire, UK), deprotected by treatment with a 50:50 mixture of ammonium hydroxide and aqueous methylamine (AMA) (Sigma-Aldrich, Saint Louis, MO) and purified on HPLC using Transgenomic Wave System (Transgenomic Omaha, NE).

**dGoligo binding.** The direct dG binding to RNA targets was confirmed with a standard gel-electrophoresis technique and using an approach based on primer extension by reverse transcriptase. Proper length and quality of PCR products was confirmed in agarose gel electrophoresis (S7 Fig). Target RNA for dGoligo (dG) binding was obtained by in vitro T7 polymerase-mediated transcription of pKS-A or pKS-F plasmids. Before electrophoresis, RNA (containing TRβ1 5'UTR A or F and downstream coding sequence of luciferase) was treated with DNase I (Fermentas, Vilnius, Lithuania) to remove remnant plasmid DNA and purified with RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany). Then, 80ng RNA was denatured, co-hybridized with 20pmol of a single dG and stained with SYBR Green I (S7 Fig). **Binding selectivity** of dGs was assessed by measuring their ability to drive synthesis of specific cDNAs during reaction of reverse transcription, wherein each tested dG served as a specific primer for reverse transcriptase that requires complementarity between a target sequence and, at least, 3'-end of an oligonucleotide (S8 Fig). pKS-A and pKS-F transcripts served as a template for DNA-based antisense-like dGs. Sense-like dGs share the same sequence with matrix RNA, thus were expected to have no effects on transcription of the RNA. In case of the sense dGs, instead of RNA, we used purified first strand cDNA as a template. dG-primed products were synthesized by reverse transcription of pKSs' RNA with RevertAidTM H Minus First Strand cDNA Synthesis Kit (Fermentas, Vilnius, Lithuania). The RNAs were previously treated with DNase I (Fermentas, Vilnius, Lithuania) and purified with RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany). Then, standard PCR was performed to confirm the expected dG-primed products (S8 Fig). Due to a tendency of 2'-O-methyl groups to impede reverse transcriptase [s4], binding of dGs modified by this group were tested only by the standard gel-electrophoresis (S7 Fig).

**Genetic constructs containing 5'UTRs.** Preparation of luciferase reporter constructs containing different TRβ1 5'UTR variants is described by Francton et al [s5]. **Linear expression construct containing p16INK4a 5'UTR** was performed by assembling: T7 promoter, 5'UTR of p16INK4a (306nt) and luciferase reporter sequence with its 3'UTR. The construct was carried out using a three-step overlap extension PCR protocol [s6] that was elaborated on the basis of principles described by

Roche in the RTS Wheat Germ LinTempGenSet manual [s7, s8] (now distributed by 5 PRIME). In the first step, p16INK4a 5'UTR-specific PCR starters (SI.F and SI.R, S4 Table) were used to add overlap regions to the amplified sequence of p16INK4a 5'UTR (see scheme below). T7 promoter and a 5' fragment of luciferase coding sequence (CDS) were added to the flanking primers (cSIII.F-T7.p, cSII.F). The luciferase with 3'UTR was amplified in the second step, wherein luciferase-specific primers (SII.F, SII.R) were used to add overlap regions to the luciferase CDS. Both SI.R and SII.R contained overlap regions (cSIII.F, cSIII.R) for amplifying primers used in the third step. Human cDNA and pGL3 Luciferase Reporter Vector (pGL3-control vector, Promega) were used as a template for the first and second step, respectively. In the third step, overlap extension PCR, the products of the first and second PCR annealed with the added flanking primers (SIII.F, SIII.R) and the 5' and 3' ends were extended. Due to high GC-content in 3'-end of p16INK4a 5'UTR preamplification of the 5'UTR was performed (PCR 0), using shorter primers: S0.F and S0.R, which included one degenerated base to facilitate the PCR 0 (amplicon length = 306bp). Subsequent PCR reactions were performed using the following oligonucleotides: PCR-I (367bp): SI.F(that includes: cSIII.F – T7p. – p16 5'UTR) SI.R(cSII.F); PCR-II (2234bp) SII.F, SII.R(cSIII.R); PCR-III (2234bp) SIII.F, SIII.R (S4 Table). Finally, the linear expression construct (2234bp) was ready for subsequent coupled in vitro transcription-translation performed using RTS 100 Wheat Germ CECF system. This reaction was carried out in the same way as it was described in TRβ1 studies (see article). The following dGs were used: sense dG1p16, antisense dG2p16, microRNA-like sense dG3p16, microRNA-like antisense dG4p16 and scrambled control dGscp16 (S3 Table). MicroRNA-like loop was created by adding two non-complementary bases in the middle of dG1p16 and dG2p16. Reverse transcription and semi-quantitative Real-Time PCR was performed as described in Materials and Methods using the same primer pairs. Luc-rev-r, T7prom-f, Luc-RT-f, Luc-RT-r (S4 Table).


Scheme of PCR-amplified linear expression construct containing 5'UTR of p16INK4a (*CDKN2A*). This construct was generated to serve as a template in coupled *in vitro* transcription/translation assay. T7 promoter (T7.p), 5'UTR of p16INK4a (306nt), luciferase reporter sequence (CDS) together with its 3'UTR were assembled using a three-step overlap extension PCR protocol. SI.F, SI.R, SII.F, SII.R, cSII.F, cSIII.F and cSIII.R represent names of primers (S4 Table) that were used in the three-step PCR (PCRI, PCRII and PCRIII). Human cDNA and pGL3 Luciferase Reporter Vector (pGL3-control vector, Promega) were used as templates for the first (PCRI) and second step (PCRII) respectively.

**Analysis of translational regulatory potential of TRβ1 5'UTRs.** Since the Translation Regulatory Potential (TRP) was important for predicting the 5'UTR target sites for dGs, we tried to determine a numerical parameter that could assess the TRP of our mRNA variants. To determine the TRP of TRβ1 5'UTRs we used an exemplary single nucleotide polymorphism (SNP, refID: rs62255380) relating to C219T on a putative TRβ1 IRES domain located in exon 2. This SNP was the only one polymorphism of TRβ1 5'UTR, validated in NCBI SNP database that could alter *Gibbs energy-dependent secondary structure formation* of all TRβ1 5'UTR variants. In other words, we tried to determine the translation regulatory potential of various TRβ1 5'UTRs by assesing the effects of the C219T substitution on theoretical translation efficiency (TTE). The calculations and results are shown in Table S2.

**Translation-enhancing assay**. This experiment was performed to assess translation-enhancing effects triggered by dGoligos (dGs). TRβ1 5'UTR-specific, translation-enhancing assay was designed on the basis of a previous observation that one of transcript variants encoded by C*DKN2A* suppressor gene (NCBI Gene ID: 1029) can be efficiently enhanced in the presence of a PCR sense primer directed to its strongly folded 5'UTR. Universality of this approach was confirmed by the use of TRβ1 5'UTR- and p16INK4a 5'UTR-specific dGs. 500ng of the plasmids pKS-A, pKS-F and pKS-control were transcribed and translated in the presence of 0,25μM of tested dG (S3 Table) or in the absence of any dG (control), using RTS 100 Wheat Germ CECF system (Roche Diagnostics, Mannheim, Germany) in conditions described in the article. mRNA levels and luciferase activity measured in each experiment were divided by the corresponding results obtained for pKS-control lacking a TRβ1 or p16INK4a 5'-UTR. Reaction mixtures were collected for analysis by luciferase assay and real-time PCR. Reactions were performed in triplicate in three independent assays (Fig 4, Fig 5).

**Translation controlled by IRES-like element** in TRβ1 5'UTR. Since an alternate cap-independent, IRES-dependent translation is demonstrated to be activated by serum deprivation, which can initiate integrated stress response (ISR) [s9, s10], we performed a simple study to determine whether serum-starved Caki-2 cells (clear cell Renal Cell Cancer) can change 5'UTR-controlled translation efficiency of a downstream coding sequence. We used pGL3-A expression plasmid [s5] containing 5'UTR variant A, which has been reported to possess an IRES-like sequence located at exon 1c/2a boundary [s11]. The measurements were shown in relation to pGL3-control plasmid containing an irrelevant synthetic vector-based leader sequence lacking any TRβ1 5'UTR. Caki-2 cells were seeded at 5×105 cells per well using 12-well plates and cultured 24 hrs in McCoy's medium supplemented with 10% FBS. After 24 hrs the cells were transfected with 100ng pRL-TK and 1μg of pGL3-A or pGL3-control plasmids, using 1μg/μl PEI and 150 mM NaCl in FBS-free McCoy's medium. 5 hrs after, transfection the medium was replaced with fresh FBS-free medium to induce ISR caused by serum deprivation. At the same time, control cell cultures were supplemented with 10% FBS. Proliferation of the serum-starved Caki-2 cells but not FBS-supplemented cells was inhibited that was assessed by cell counting. The cells were maintained at 37°C in 5% CO2 atmosphere, harvested after 24 hrs and quickly divided into 2 equal parts – for isolation of total RNA and luciferase protein. Luciferase mRNA levels were assessed with Real-Time PCR and the protein measurements were performed using dual-luciferase assay in the Synergy2 luminometer. The levels of firefly luciferase activity (pGL3-A) were normalized to activity of constitutively expressed Renilla luciferase (pRL-TK). Materials used in this study are described in the article. Data from three independent experiments were performed in 12 repeats. The Shapiro–Wilk test was used to determine normality of data distribution. Normally distributed data were analyzed by ANOVA followed by Dunnett's multiple comparison test, *p< 0.01, **p<0.0001 vs. control (S2 Fig).

**Measurements of transcripts.** Control mRNA levels were determined using quantitative real-time PCR method (Q-PCR), performed with LightCycler® 480 (Roche, Germany). Reaction mixtures of coupled transcription-translation containing equal quantity of reporter constructs were purified using GeneMATRIX Universal RNA Purification Kit (EURx, Gdansk, Poland). Reverse transcription in experiments with luciferase-containing plasmids was performed directly on the purified reaction mixture, using specific primer Luc-rev-r (S4 Table and S8 Fig) and the RevertAidTM H Minus First Strand cDNA Synthesis Kit (Fermentas, Vilnius, Lithuania). 1μl of 5x diluted reverse transcription reaction was used for further Q-PCR reactions using Quanti-Fast SYBR Green PCR Kit (Qiagen, Hilden, Germany) and first pair of primers: Luc-RT-f and Luc-RT-r amplifying both luciferase DNA (plasmid vector) and cDNA (RNA reverse transcription product), under the following conditions: 95°C 5min; 50 cycles: 95°C 10s, 57°C 15s, 72°C 15s; melting curve analysis: 135 cycles: 50°C; 0.3°C increase in each cycle. Ct data were acquired after reaching the threshold in real-time module, usually between 18

and 36 cycle; cycle efficiency was corrected using LightCycler® 480 (Roche, Germany). Standard curve was prepared using serial dilutions of luciferase cDNA amplification products. Second Q-PCR reaction was performed using second pair of primers: T7prom-f and Luc-RT-r (S4 Table), specific only to the template vector DNA, serving as internal control for transcript levels. The final amount of each transcript was calculated by dividing quantity of the PCR products of first primer pair (amplifying both DNA and RNA) and the second primer pair (amplifying only DNA). Relative changes in gene expression were calculated using $2^{(-\Delta\Delta Ct)}$ [s12]. Levels of naturally occurring mRNAs in *in vivo* experiments were determined as described above, using transcript-specific primers (S4 Table).

**References** for supporting materials and methods:


s1 Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics. 2010 Mar 15;11:129.

s2 Lu ZJ, Gloor JW, Mathews DH. Improved RNA secondary structure prediction by maximizing expected pair accuracy. RNA. 2009 Oct;15(10):1805-13.

s3 Tinoco I Jr, Collin D, Li PT. The effect of force on thermodynamics and kinetics: unfolding single RNA molecules. Biochem Soc Trans. 2004 Nov;32(Pt 5):757-60.

s4 Maden BE. Mapping 2'-O-methyl groups in ribosomal RNA. Methods. 2001 Nov;25(3):374-82.

s5 Frankton S, Harvey CB, Gleason LM, Fadel A, Williams GR. Multiple messenger ribonucleic acid variants regulate cell-specific expression of human thyroid hormone receptor beta1. Mol Endocrinol. 2004 Jul;18(7):1631-42.

s6 Swartz JR. *Cell-Free Protein Expression*. Springer Science & Business Media, New York 2012,

s7 Mc Pherson MJ, Moller SG. *PCR, The basic from background to bench*. 1st published, BIOS Scientific Publishers Ltd., Oxford, UK. 2000.

s8 Singh H, Makino S, Endo Y, Li Y, Stephens AN, Nie G. Application of the wheat-germ cell-free translation system to produce high temperature requirement A3 (HtrA3) proteases. Biotechniques. 2012 Jan;52(1):23-8.

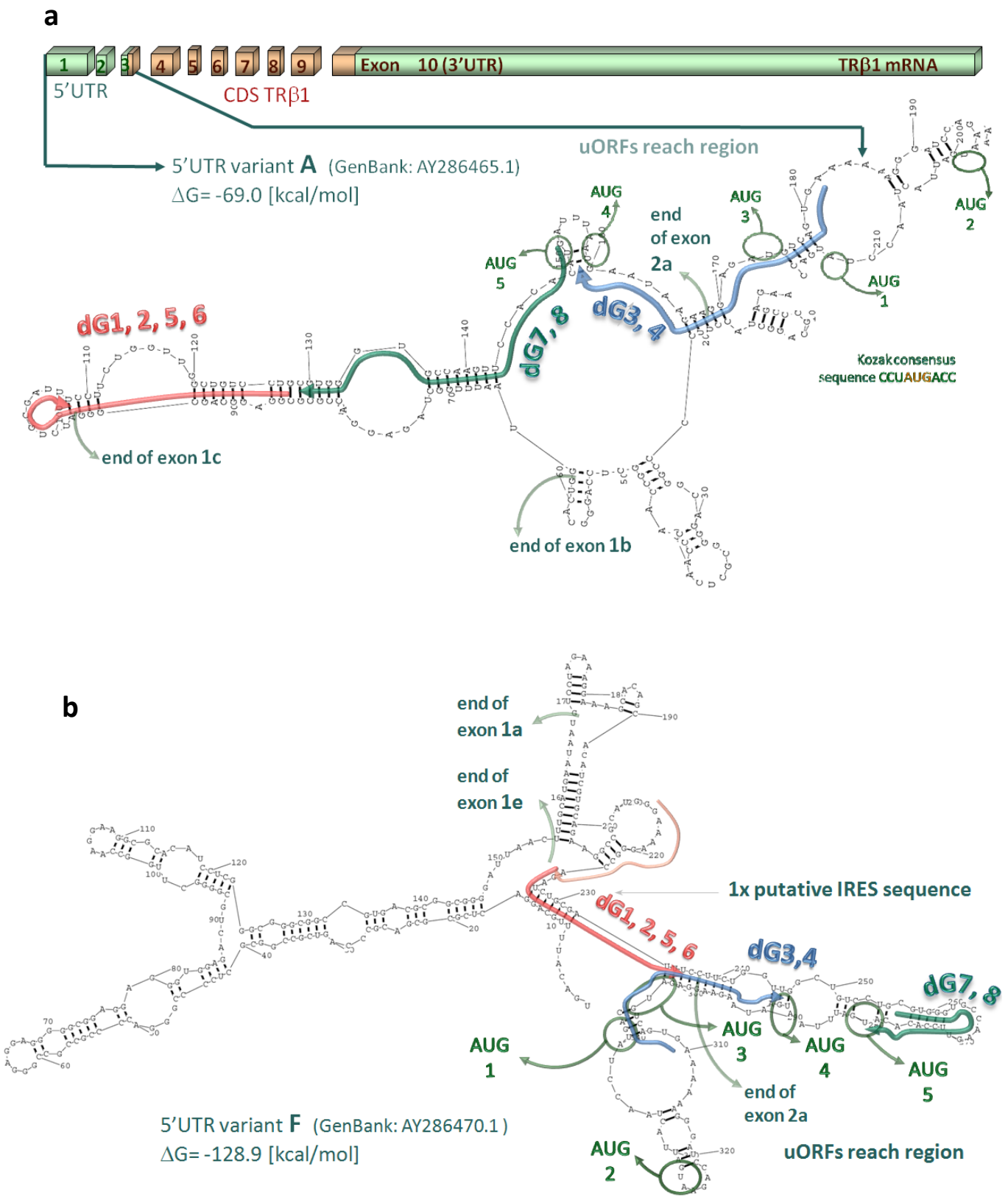
s9 Mokrejs M, Masek T, Vopálensky V, Hlubucek P, Delbos P, Pospísek M. IRESite--a tool for the examination of viral and cellular internal ribosome entry sites. Nucleic Acids Res. 2010 Jan;38(Database issue):D131-6.

s10 Komar AA, Hatzoglou M. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle. 2011 Jan 15;10(2):229-40.

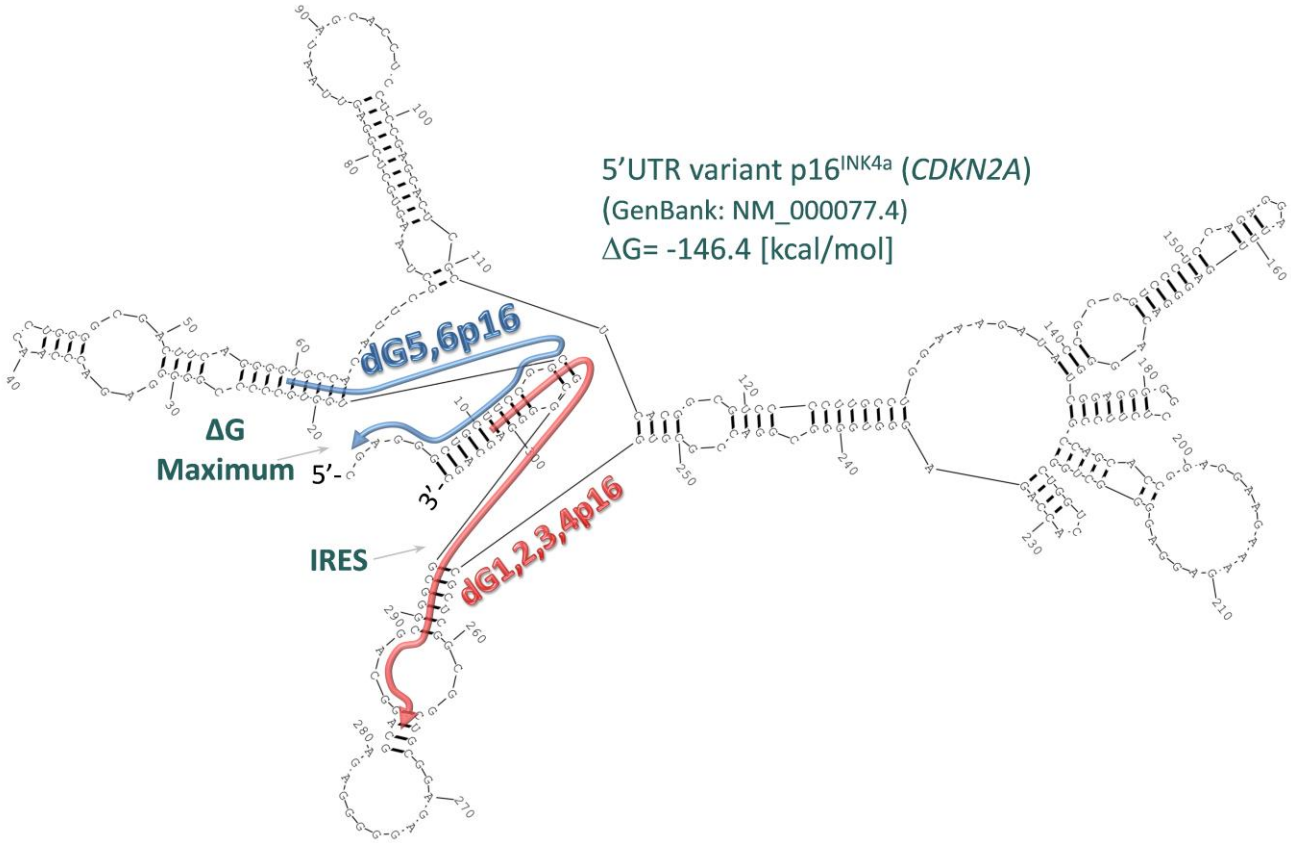
s11 Master A, Nauman A. Molecular mechanisms of protein biosynthesis initiation - biochemical and biomedical implications of a new model of translation enhanced by the RNA hypoxia response element (rHRE). Postepy Biochem. 2014; 60(1):39-54.

s12 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001 Dec; 25(4):402-8.

## Supporting figures



**c**




**S1 Fig. Folding of TRβ1 5'UTRs.**

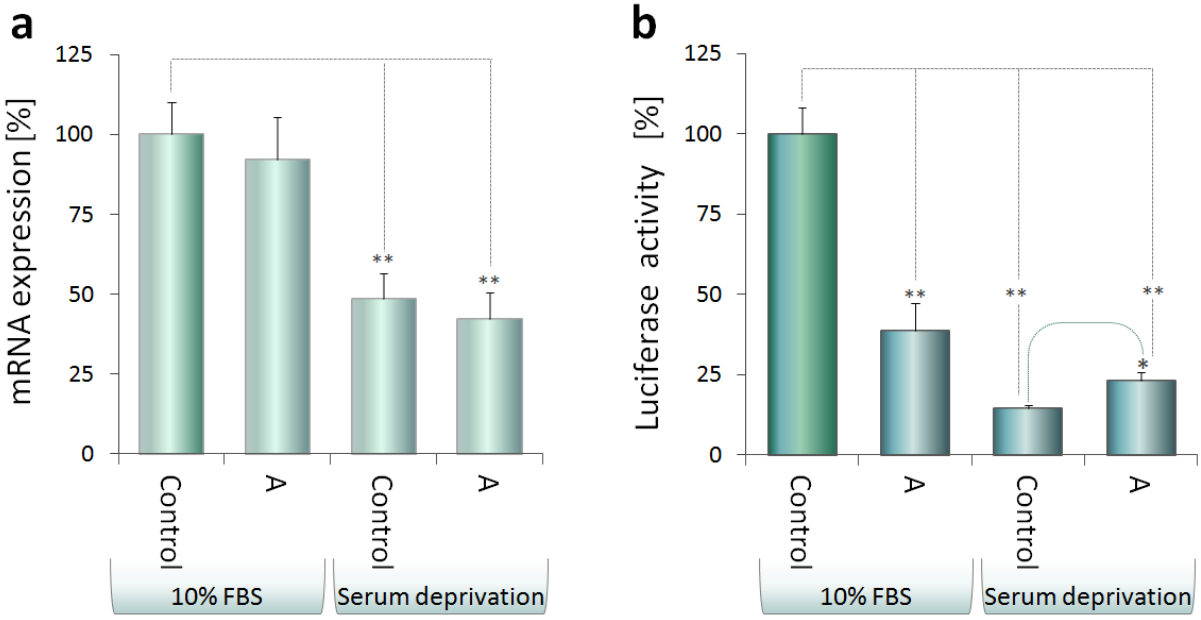
Secondary structures were modeled using RNAstructure version 5.2. (**a**) Weakly folded 5'UTR variant A (ΔG=-69.0 kcal/mol) and (**b**) Strongly folded 5'UTR variant F (ΔG=-128.9 kcal/mol) with indicated 5'UTR exons and coding sequence of TRβ1 mRNA (GeneBank Acc. No. NM_000461). Variant F lacks ex1c present in variant A what results in incomplete sequence homology at 3'-end of antisense dG2. Putative IRES sequences (Master et al. 2010), uAUGs (Frankton et al. 2004), exon-exon junctions and dG binding sites are indicated with arrows in both figures. This set of enhancing dG*s* can result in simultaneous unblocking of putative IRES sequence and blocking of uAUGs-rich region leading to significant enhancement of translation efficiency (see Fig 4). (**c**) Strongly folded 5'UTR variant $p16^{INK4a}$ encoded (*CDKN2A*) (ΔG=-146.4 kcal/mol, GeneBank Acc. No. NM_000077.4) with indicated dGp16 target sites, ΔG maximum (identified by dGenhancer*)* and IRES sequence reported by Bisio et al. 2015.

# Supporting tables

| 5'UTR variant* | GenBank Acc. No.: | Length [nt] | uAUGs [no.] | GC-content [%] | ΔG [kcal/ mol] | Translation** efficiency [%] |
|---|---|---|---|---|---|---|
| control | - | 29 | 0 | 55.17 | -6.3 | 100 |
| **A** | AY286465.1 | 211 | 4 | 54.98 | **-69.0** | 24.09 |
| B | AY286466.1 | 229 | 5 | 54.15 | -82.0 | 12.08 |
| C | AY286467.1 | 195 | 4 | 56.41 | -77.5 | 14.99 |
| D | AY286468.1 | 314 | 5 | 48.09 | -95.0 | 7.01 |
| E | AY286469.1 | 305 | 5 | 47.54 | -92.5 | 7.99 |
| **F** | AY286470.1 | 339 | 7 | 58.41 | **-128.9** | 4.02 |
| G | AY286471.1 | 388 | 6 | 52.84 | -127.0 | 3.00 |

**S1 Table. Basic characteristics of selected TRβ1 5'UTR variants A-G.**

Prediction of secondary structures of 5'UTRs was performed as described in bioinformatic analysis. The least and most folded TRβ1 variants are shown in bold. uAUGs represent start codons located upstream of the main start codon and Kozak consensus sequence. *TRβ1 5'UTR splice variants termed according to Frankton et al 2004. ** Translation efficiency measured with RTS 100 Wheat Germ CECF system.




**S2 Fig. Translation-enhancing, IRES-like element in TRβ1 5'UTR.**

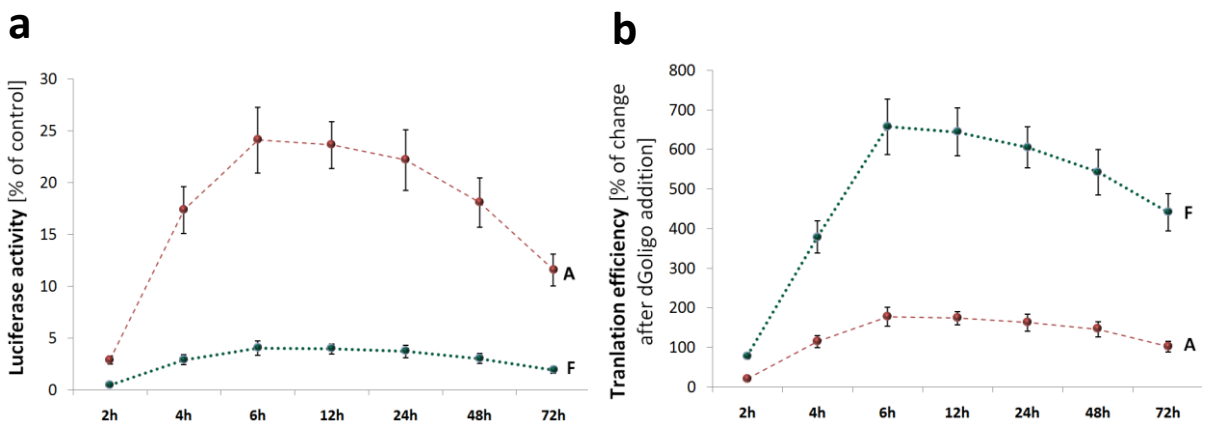
This figure shows functional properties of a putative IRES-like element of TRβ1 5'UTRs, activated in response to serum starvation in Caki-2 cell culture (S1 Appendix). We used pGL3-A expression plasmid containing 5'UTR variant A (A), which has been reported to possess a putative IRES-like sequence located at exon 1c/2a boundary (see Fig 2). (**a**) Luciferase mRNA levels of pGL3-A (A) are shown relative to control plasmid - pGL3-Control (Control). Transcript levels were significantly reduced in serum-deprived Caki-2 cells (A and Control). (**b**) Luciferase activity (protein) levels of pGL3-A are shown relative to pGL3-Control. 2.59-, 6.92- and 4.36-fold lower translation rates of variant A in 10% FBS, Control and variant A in serum deprived medium were noted relative to Control in 10% FBS. Simultaneously, 1.59-fold higher luciferase protein levels were detected in serum-deprived cells transfected with pGL3-A when compared to pGL3-Control showing that the TRβ1 5'UTR contains a *cis*-acting element allowing for relatively efficient translation under serum-deprived conditions. These results are consistent with previously reported putative IRES-site in the TRβ1 5'UTR and may support our mechanistic model of dG action (S4 Fig). Data from three independent experiments were performed in 12 repeats. The Shapiro–Wilk test was used to determine normality of data distribution. Normally distributed data were analyzed by ANOVA followed by Dunnett's multiple comparison test, *$p < 0.01$, **$p < 0.0001$ vs. control.

| Column → | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| [&]**Abbre-viation →** | Acc. No. | ΔG [kcal/mol] | ETE [%] | **TE** [%] | ΔG [kcal/mol] | **TE** [%] | ΔG shift [%] | TTE shift [%] |
| **Sequence →** | reference | reference | reference | reference | *substituted* | *substituted* | substituted/ reference | substituted/ reference |
| **Value →** | | theoretical | experimental | theoretical | theoretical | theoretical | theoretical | theoretical |
| **Equation →** | | determined using RNA-structure prog. | determined with translation assay | =127.24* e^(0.0284* column-3) | determined using RNA-structure prog. | =127.24* e^(0.0284* column-6) | =\|100%-column6/3 *100%\| | =\|100%-column7/5 *100%\| |
| Reference 5'UTR variant ↓ | GenBank Acc. No. – reference variants | ΔG of refe-rence variants | Experi-mental results of translation efficiency (TE) | Predicted TE of reference variants | ΔG of substituted variant | Predicted TE of substituted variants | Shift in Gibbs energy (between substituted and reference variants) | Shift in predicted theoretical TE (between substituted and reference variants) |
| Control | - | -6.3 | 100 | 106.44 | -6.3 | 106.44 | 0.00 | 0.00 |
| **A** | AY286465.1 | -69.0 | 24.09 | 17.94 | -67.5 | 18.72 | **2.17** | **4.35** |
| B | AY286466.1 | -82.0 | 12.08 | 12.40 | -78.8 | 13.58 | 3.90 | 9.51 |
| C | AY286467.1 | -77.5 | 14.99 | 14.09 | -75.2 | 15.04 | 2.97 | 6.75 |
| D | AY286468.1 | -95.0 | 7.01 | 8.57 | -90.3 | 9.80 | **4.95** | 14.28 |
| E | AY286469.1 | -92.5 | 7.99 | 9.20 | -88.4 | 10.34 | 4.43 | 12.35 |
| **F** | AY286470.1 | -128.9 | 4.02 | 3.27 | -123.0 | 3.87 | 4.58 | **18.24** |
| G | AY286471.1 | -127.0 | 3.00 | 3.45 | -121.4 | 4.05 | 4.41 | 17.24 |

## S2 Table. Prediction of translational regulatory potential of 5'UTRs.

This analysis was performed to determine translation regulatory potential (TRP) of various TRβ1 5'UTRs and was used in dGenhancer calculations (S1 Appendix). The TRP predictions are shown as numerical parameters such as shifts in *theoretical translation efficiency* (TTE shifts, column-9) and Gibbs energy (ΔG shifts, column-8). The TTE shifts were calculated by dividing predicted TTE of virtually *substituted* TRβ1 5'UTR variants (column-7) and reference non-*substituted* variants (column-5). The Gibbs energy shifts were calculated by dividing Gibbs energy values of the *substituted* (column-6) and reference variants (column-3). Computational prediction of the translation efficiency (TE) was performed on the basis of exponential trend-line equation correlating experimentally obtained values of translation efficiency and the Gibbs energies of the 5'UTRs ($y=127.24 \cdot e^{0.0284 \cdot x}$, where x means calculated Gibbs energy, number e - constant = 2.718, y - translation efficiency value). Extreme values (min., max.) of the shift in Gibbs energy and TTE shifts are shown in bold. *Substituted* variant D and F were predicted (by ΔG shifts and TTE shifts, respectively) to have the highest translational regulatory potential. In contrast to ΔG shifts, TTE shifts include calculations from a trend-line equation correlating experimental results of translation efficiency with 5'UTR Gibbs energies. In our experimentally obtained data variant F was found to have the highest dG-triggered TRP (Fig 4.d and S5 Fig), whereas variant D has been previously reported to drive efficient luciferase expression in kidney-derived COS-7 cells (Frankton et al. 2004) that is in concordance with the prediction in this table. These data may show that 5'UTR TRP should always be estimated in the context of a translation system involving additional parameters of translation machinery such as various *trans*-acting factors that, besides the Gibbs energy, may influence protein synthesis efficiency.

[&]**Abbreviations**: TRP – Translational Regulatory Potential; ETE – Experimentally determined Translation Efficiency; TTE – Theoretical Translation Efficiency, calculated on the basis of trend-line equation (see below) and values of ΔG – Gibbs Energy; Acc. No. – Accession Number of Gene Bank (NCBI); reference sequence – correct sequence; substituted sequence – containing Small Nucleotide Polymorphism (SNP); shift - change between calculated values (ΔG [kcal/mol] or translation efficiency [%]) of reference and substituted 5'UTR variant.




**S3 Fig. Time-course of protein synthesis rates in RTS 100 Wheat Germ CECF system.**

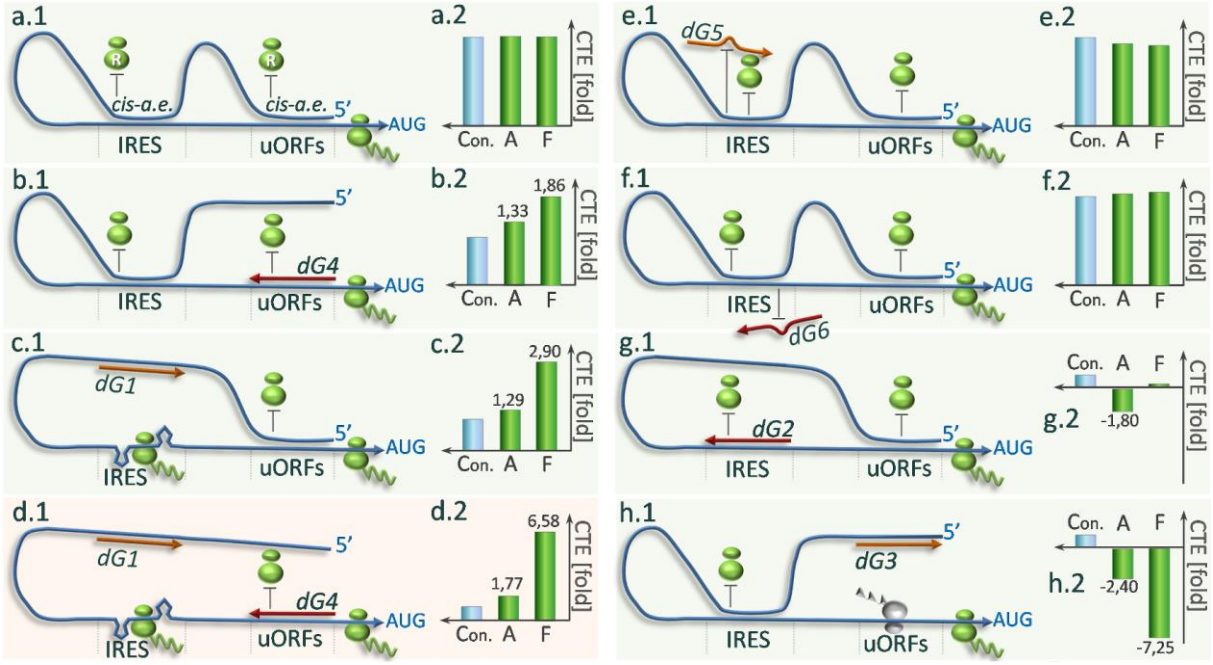
(**a**) Effects of 5'UTR variants A (red line) and F (green line) on luciferase activities (reporter protein levels) after 2, 4, 6, 12, 24, 48 and 72 hours of coupled transcription-translation assay are shown relative to the control plasmid (Control). (**b**) Effects of both dG1 and dG4 (S3 Table) on translation efficiency from pKS-A after 2, 4, 6, 12, 24, 48 and 72 hours (red line) or pKS-F (green line) are shown normalized to control (dG-). Experiments were performed in triplicate and shown as mean % luciferase activity ± SD. Data were analyzed by ANOVA, *p<0.001 vs. control.

| dGoligo full name and target site within 5'UTR | Short -cut | Position 5'→3' [nt] | Target sequence (element) | Structure / strand orientation | Design method | dGoligo Sequence 5'→3' |
|---|---|---|---|---|---|---|
| dGoligo1-TRβ1vA-IRES-ex1c/ex2a-F | dG1 | 86→108 | e1 TRβ1 | sense | manual | GGAGGCAGCGGGATCTGCGATTT |
| dGoligo2-TRβ1vA-anti-IRES-ex1c/ex2a-R | dG2 | 108→86 | e1 TRβ1 | antisense | manual | AAATCGCAGATCCCGCTGCCTCC |
| dGoligo3-TRβ1vA-uAUGs-ex3/ex2a-F | dG3 | 159→179 | e3 TRβ1 | sense | manual | TGAATAAGAAGGAGATGTCAG |
| dGoligo4-TRβ1vA-anti-uAUGs-ex3/ex2a-R | dG4 | 179→159 | e3 TRβ1 | antisense | manual | CTGACATCTCCTTCTTATTCA |
| dGoligo5-TRβ1vA-mirSenseIres- ex1c/ex2a-F | dG5 | 86→108 | e1 TRβ1 | miRNA-like sense control | manual | GGAGGCAGCGGGtagATCTGCGATTT |
| dGoligo6-TRβ1vA-anti-mirAntiSenseIres- ex1c/ex2a-R | dG6 | 108→86 | e1 TRβ1 | miRNA-like antisense control | manual | AAATCGCAGATctaCCCGCTGCCTCC |
| dGoligo7-TRβ1vA-cisAE-ex2a-F | dG7 | 129→151 | e2 TRβ1 | sense | automated dGenhancer | CGTGGGTGCCAAGTTCCACACAT |
| dGoligo8-TRβ1vA-anti-cisAE-ex2a-R | dG8 | 151→129 | e2 TRβ1 | antisense | automated | ATGTGTGGAACTTGGCACCCACG |
| dGoligo9-TRβ1vA-mirSenseCisAE-ex2a-F | dG9 | 129→151 | e2 TRβ1 | miRNA-like sense | automated dGenhancer | CGTGGGTGCCAActaGTTCCACACAT |
| dGoligo10-TRβ1vA-mirAntisense-cisAE-ex2a-R | dG10 | 151→129 | e2 TRβ1 | miRNA-like antisense | automated | ATGTGTGGAACTctcTGGCACCCACG |
| dGoligoSc-TRβ1vA | dGsc | 108→86 | - | nonsense / scrambled control | automated GeneScript | GCCATCACTACGCCTACCTCAGTCGA |
| dGoligo1p16-p16INK4a5'UTR-sense-U | dG1p16 | 280→301 | e1 p16 | sense | manual and automated dGenhancer | AGCAGGCAGCGGGCGGCGGGGA |
| dGoligo2p16- p16INK4a5'UTR-antisense-L | dG2p16 | 301→280 | e1 p16 | antisense | manual and automated dGenhancer | TCCCCGCCGCCCGCTGCCTGCT |
| dGoligo3p16-p16INK4a5'UTR-miRNA-U | dG3p16 | 280→301 | e1 p16 | miRNA-like sense | automated dGenhancer | AGCAGGCAGCGGtaGCGGCGGGGA |
| dGoligoG4p16-p16INK4a5'UTR-miRNA-L | dG4p16 | 301→280 | e1 p16 | miRNA-like antisense | automated dGenhancer | TCCCCGCCGCtaCCGCTGCCTGCT |
| dGoligo5p16-p16INK4a5'UTR-sense-U | dG5p16 | 2→22 | e2 p16 | sense | automated dGenhancer | GAGGGCTGCTTCCGGCTGGTG |
| dGoligo6p16- p16INK4a5'UTR-antisense-L | dG6p16 | 22→2 | e2 p16 | antisense | automated dGenhancer | CACCAGCCGGAAGCAGCCCTC |
| dGoligoGscp16-p16INK4a5'UTR | dGscp16 | 301→280 | - | nonsense / scrambled control | automated GeneScript | GTCCTCGCCTCCGCACCGGTCCTC |
| dGoligo-hsa-miR211 | dG211 | 192→168 | e3/e4 TRβ1 | miRNA | miRBase | TTCCCTTTGTCATCCTTCGCCT |
| dGoligo-hsa-miR-211-3p | dG211c | 168→192 | e3/e4 TRβ1 | Complem. miRNA | miRBase | GCAGGGACAGCAAAGGGGTGC |

## S3 Table. List of dGoligos (dGs) used in the study.

Position in 5'UTR indicates dG recognition site in 5'UTR variant A of TRβ1 (TRβ1vA, GeneBank acc. no.: AY286465.1) encoded by *THRB* gene and in p16INK4a 5'UTR (NM_000077.4) encoded by *CDKN2A*. 2-3-nucleotide, microRNA-mimicking loop is indicated with lowercase letters, dGsc-scrambled control.

dGenhancer (S1 Appendix)
GeneScript - https://www.genscript.com/ssl-bin/app/scramble
miRBase - http://www.mirbase.org/




**S4 Fig. Proposed folding patterns of TRβ1 5'UTR after dGoligo supplementation.**

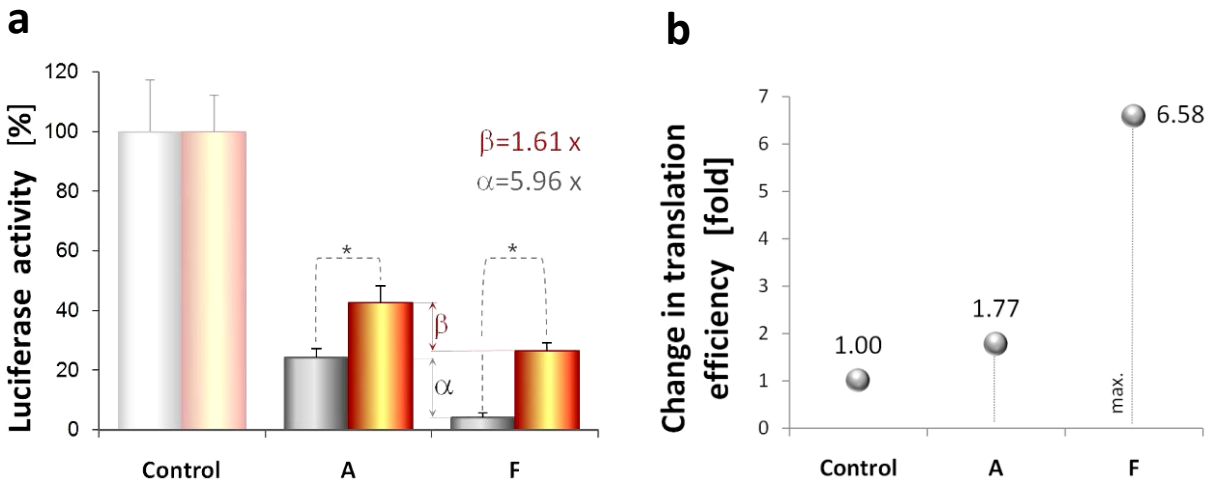
Simplified models of secondary structures that could be linearized by dGoligos (dGs) are shown (a.1-h.1) in relation to experimentally obtained data presenting changes in translation efficiency (CTE) after dG supplementation (a.2-h.2). Statistically significant CTEs ($p<0.001$) are indicated by fold of change in translation efficiency shown above or below green bars of 5'UTR variant A (A) and variant F (F) normalized to 100% of Control (Con. in blue) without supplementation of any dG. TRβ1 5'UTR is shown as blue curve ended by an arrow at AUG translation start codon. Two linked green ovals represent ribosome complex, that may be blocked by distant *cis*-acting element (*cis*-a.e) or *trans*-acting factor (*trans*-a.f., here dG). Putative Internal Ribosome Entry Site (IRES) involved in enhancement of cap-independent translation initiation (when free of distant *cis*-a.e.) and upstream Open Reading Frames (uORFs)–rich region, which may reduce translation initiation from the correct AUG start codon (when free of inhibitory *cis*-a.e. and *trans*-a.f.) are shown between doted vertical lines. (**a.1**) Theoretical state of naturally folded 5'UTR (without supplementation of any dG), with IRES and uORFs-rich domains are at least partially blocked by distant *cis*-acting elements resulting in basal translation level of correct protein. (**b.1**) Proposed model of dG4-mediated enhancement of translation efficiency, in which antisense dG4 can alter *Gibbs energy-dependent secondary structure formation* via direct binding to uORFs-rich region. This binding may block translation of truncated proteins originating from upstream AUGs, that finally may enhance translation initiation from correct AUG start codon. In the model, putative IRES domain stays at least partially blocked by distant *cis*-acting element. (**c.1**) Model of dG1-mediated enhancement of translation efficiency, where the sense dG1 can release e1 element containing putative IRES domain via binding to distant *cis*-acting sequences, normally interacting with the IRES sequence. This may allow for appropriate secondary structure formation of IRES domain needed for efficient cap-independent translation. In the model, uORFs-rich region stay at least partially blocked by naturally occurring distant *cis*-acting element of the 5'UTR, that finally may allow for translation initiation from correct AUG start codon. (**d.1**) Model of coupled action of dG1 and dG4 that mediate strong enhancement of translation efficiency (d.2). dG1 can release putative IRES domain via binding to distant *cis*-acting element and antisense dG4 can repress undesirable translation originated from uAUGs via direct binding to uORFs-rich region. This may allow for appropriate secondary structure formation of IRES domain needed for efficient cap-independent translation and blocking of uORFs-rich region required for efficient cap-dependent translation initiation from correct AUG start codon.

The observed enhancement of translation efficiency of folded variant F (d.2, F) may show that strongly folded 5'UTR variants have higher translational regulatory potential (TRP) when compared to weakly folded variants (d.2, A), which are efficiently translated even without addition of any *trans*-acting factors (here dGs). (**e.1**) Model of control dG5 action, that was blocked (e.2) by 3-nt insertion in the middle of the dG leading to insufficient similarity with 5'UTR mRNA sequence. (**f.1**) Model of control dG-6 (dG6) action, that was blocked (f.2) by 3-nt insertion mutation in the middle of the dG leading to insufficient complementarity with 5'UTR mRNA sequence and/or partial complementarity disturbing plant RNAi-related machinery, which could be involved in observed effects of completely complementary dG1 and dG4 (d.2). (**g.1**) Model of dG2 action leading to repression of translation efficiency (g.2 A) via direct binding of the dG to putative IRES domain. This may block cap-independent translation. uORFs-rich region stay at least partially blocked by naturally occurring distant *cis*-acting element of the 5'UTR that finally may allow for translation initiation from correct AUG start codon. (**h.1**) Model of dG3-mediated repression of translation, wherein sense dG3 can release uORFs-rich domain via binding to distant sequences, normally blocking the domain. This may allow for translation of truncated proteins originated from uAUGs reducing translation initiated from correct AUG start codon. In the model, putative IRES domain stay at least partially blocked by distant sequences inhibiting cap-independent translation initiation.

| Oligonucleotide name | Oligonucleotide Sequence 5′ → 3′ | Recognition site /target | Description | Strand orientation |
|---|---|---|---|---|
| T7prom-f | TATGTAATACGACTCACTATAGGG | T7 promoter, only DNA | Real Time PCR, PCR product B | sense, forward |
| Luc-RT-f | CACCATGGAAGACGCCAA | Luciferase Kozak cons. seq./coding sequence | Real Time PCR, PCR product A | sense, forward |
| Luc-RT-r | CCAGCGGTTCCATCTTCCAG | Luciferase sequence | Real Time PCR, PCR product A | antisense, reverse |
| Luc-rev-r | GATGTCCACCTCGATATGTGC | Luciferase sequence | Reverse Transcription | antisense, reverse |
| Ffir | ATCGTGGACCGCCTGAAGTC | Firefly luciferase | Real-Time-PCR | forward |
| Rfir | ACGACGGCGGCAGGCAGC | Firefly luciferase | Real-Time-PCR | reverse |
| Fren | TGAGGAGTTCGCTGCCTACC | *Renilla* luciferase | Real-Time-PCR | forward |
| Ren | TGCGGACAATCTGGACGACG | *Renilla* luciferase | Real-Time-PCR | reverse |
| P1 | GTAATTTGGCTAGAGGACC | 5′ end of TRβ1 ex. 1c (specific to variant A) | PCR | forward |
| P2 | GTCCTAGAAAGGAAAGCACAG | 5′ end of TRβ1 ex. 1e (specific to variant F) | PCR | forward |
| P3 | AGGACCGCGCGGAGGCAG | 3′ end of TRβ1 ex.1c (specific to variant A) | PCR | forward |
| P4 | TCGAAGCTTCAGTCAGTGG-CAACCAGAAGGAAATCGCAGAT | 5′end of TRβ1 2a with 5′overhang | PCR | reverse |
| P5 | TGACATTTTGCAGGACTCG | 3′end of TRβ1 ex. 1a (specific to variant F) | PCR | forward |
| P6 | CCAACCAGAAGGAAATCGCAG | TRβ1 2a | PCR | reverse |
| S0.F | CGAGGGCTGCTTCCGGCT | p16INK4a 5'UTR | PCR0 | forward |
| S0.R | GCTGCTCCCCGCTGCCCGCT | p16INK4a 5'UTR | PCR0 | reverse |
| SI.F (cSIII.F – T7p. – p16 5′UTR) | TCGAAGCTTCAGTCAGT-ATATGTAATACGACTCACTATAGGG-T-CGAGGGCTGCTTCCGGCT | p16INK4a 5'UTR linear construct | PCRI | forward |
| SI.R(cSII.F) | GTTTTTGGCGTCTTCCAT-GCTGCTCCCCGCCGCCCGCT | p16INK4a 5'UTR linear construct | PCRI | reverse |
| SII.F | TGGAAGACGCCAAAAACAT | p16INK4a 5'UTR linear construct | PCRII | forward |
| SII.R(cSIII.R) | TGACCCTGGTTGACCCTACT-CCGGAAGGAGCTGACTGG | p16INK4a 5'UTR linear construct | PCRII | reverse |
| SIII.F | TCGAAGCTTCAGTCAGT | p16INK4a 5'UTR linear construct | PCRIII | forward |
| SIII.R | TGACCCTGGTTGACCCTACT | p16INK4a 5'UTR linear construct | PCRIII | reverse |

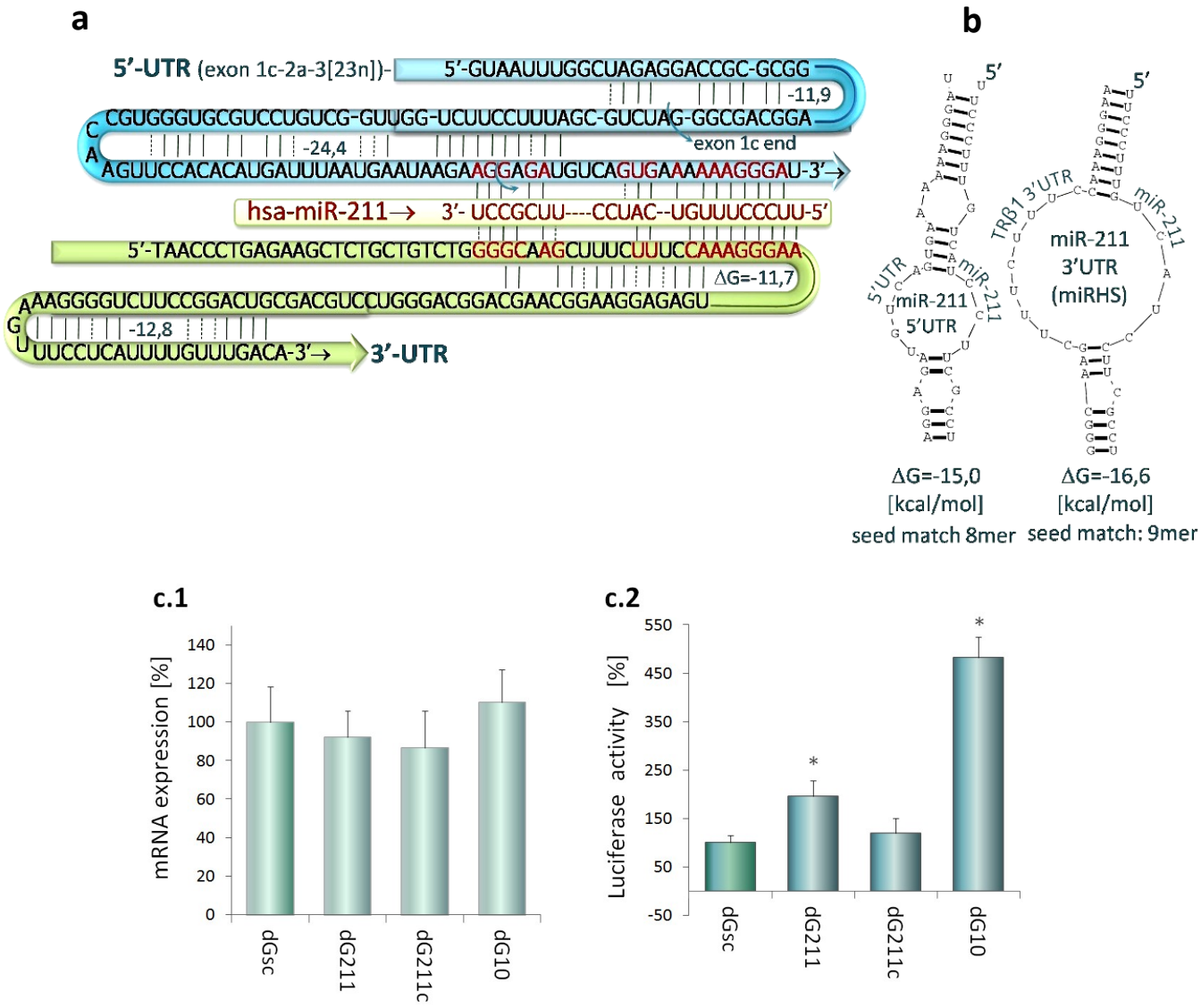
## S4 Table. List of primers used in Real-Time and classic PCR.

This table contains sequences of DNA oligonucleotides (primers) used in reverse transcription and PCR assays of TRβ1 (*THRB*) and p16INK4a (*CDKN2A*, S1 Appendix).



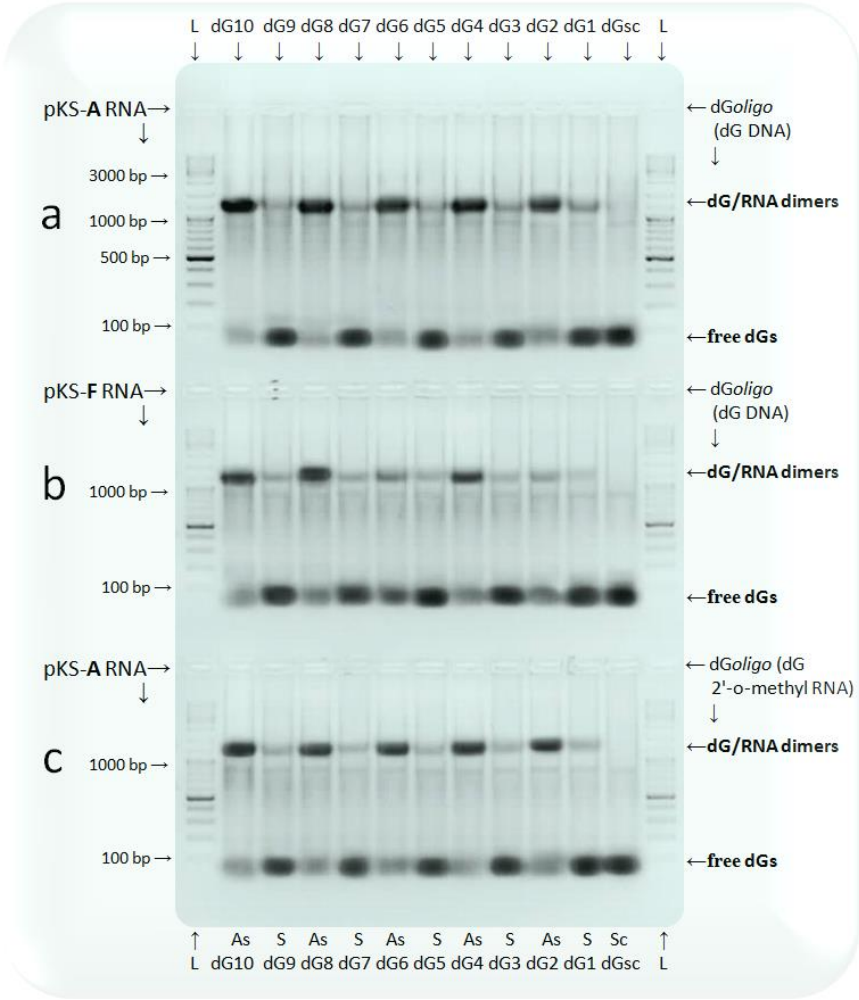

**S5 Fig. Change in translation efficiency after dG1 and dG4 supplementation.**

(**a**) Effects of dG1 and 4 on translation efficiency (luciferase activity) of 5'UTR variants A and F are shown in orange bars, whereas basal translation rates of the 5'UTRs are indicated with grey bars; all results normalized to pKS-control plasmid (Control). α value, representing 5.96-fold higher basal translation rate of variant A compared to variant F, was reduced after treatment with dG1 and 4 to value β, showing only 1.61-fold higher translation of variant A when compared to variant F. α and β, which are indicated by bidirectional arrows, are experimentally obtained values of translational regulatory potential (TRP) of TRβ1 variants A and F that is in agreement with our predictions (S2 Table). After dG1+dG4 supplementation, strongly folded variant F exceeded the basal translation level of weakly folded variant A, showing that the strongly folded variant served as a translationally inactive /less-active transcript, which was recruited to translation through interaction with a *trans*-acting factor (here dG1 + dG4). Three independent experiments were performed in triplicate and shown as luciferase activity ± SD. Results obtained for variant A and F were analyzed by ANOVA followed by Dunnett's multiple comparison test; *p< 0.001 vs. basal translation rate was considered statistically significant. (**b**) Translation-enhancing effects of dG1 and dG4 on variants A and F that were normalized to control plasmid (Control) and shown as grey dots. Coupled action of dG1 and dG4 enhanced translation efficiency over 1.77- and 6.58-fold for the variant A and F respectively.



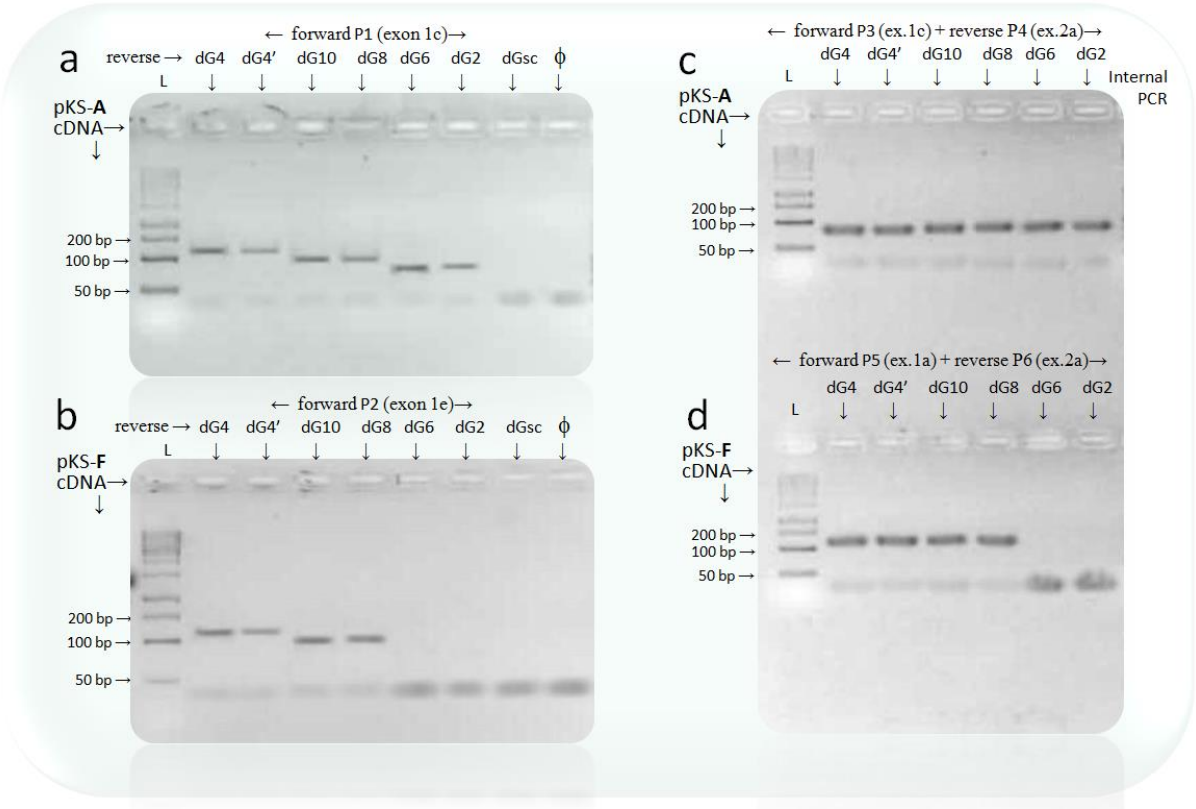

**S6 Fig. Potential hsa-miR-211 target sites within TRβ1 3'UTR and 5'UTR.**

(**a**) Target sites for a microRNA - hsa-miR-211 (miRBase Acc. no. MI0000287) in TRβ1 untranslated regions are highlighted in blue (5'UTR) and green (3'UTR). This non-selective microRNA binding may influence secondary structures of both UTRs and contribute to changes in Gibbs energy that finally may affect protein synthesis. (**b**) Target sites within TRβ1 UTR sequences were identified using miRBase (http://www.mirbase.org) and RNAhybrid (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/) that uses a Gibbs energy (ΔG) algorithm to calculate favorable binding interactions between a microRNA and potential target sites within mRNA. (**c**) Effects of 2'-O-methyl RNA modified hsa-miR-211 termed dGoligo-hsa-miR211 (dG211, S3 Table), dGoligo-hsa-miR211-3p (dG211c, ~complementary to dG211), dG10 and scrambled control (dGsc) on luciferase transcription (**c.1**) and translation (**c.2**) in Caki-2 cells transfected with pGL3-A (S1 Appendix), containing TRβ1 5'UTR variant A, luciferase coding sequence and irrelevant 3'UTR. miRBase and RNAhybrid - based analysis revealed no hsa-miR-211 and hsa-miR211-3p targets within pGL3-A 3'UTR, suggesting that the observed effects (c) could be mediated through TRβ1 5'UTR. Although our dG10 (designed on the basis of TRβ1 5'UTR) showed the strongest translation-enhancing effect in Caki-2 cells (see Fig 6), hsa-miR-211 (dG211) enhanced translation by 1.95-fold as well and had no effects on luciferase mRNA levels. Results from three independent experiments performed in triplicates are shown as mean % mRNA (a) or luciferase activity (b) ± SD. Data analyzed by ANOVA followed by Dunnett's multiple comparison test. *p<0.001 vs. control.



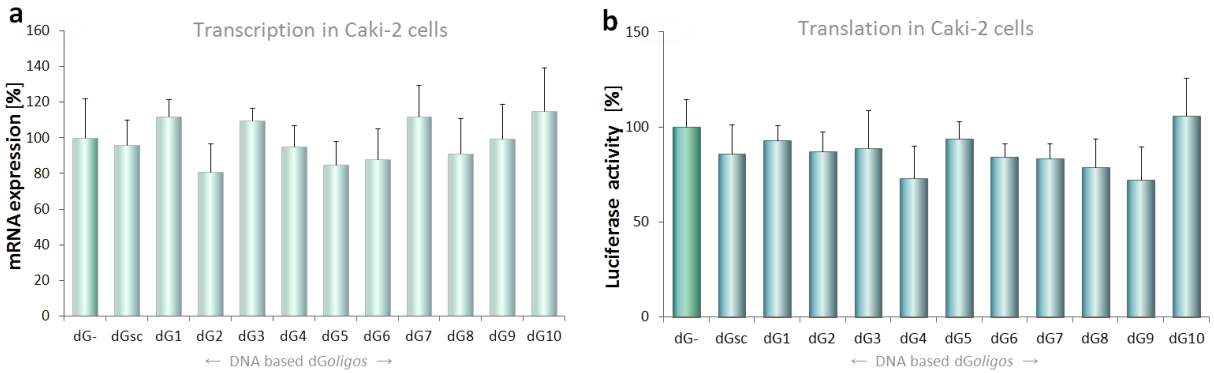

**S7 Fig. dGoligo binding capacity.**

SYBR Green I stained agarose gel electrophoretograms show bands formed by a hybridized dGoligo (dG) and a target RNA obtained by *in vitro* T7 polymerase-mediated transcription of pKS-A or pKS-F plasmids (S1 Appendix). (**a**) Binding of DNA-based dGs to the pKS-A RNA. (**b**) Binding of DNA-based dGs to the pKS-F RNA. (**c**) Binding of 2'-O-methyl modified RNA dGs to the pKS-A RNA. Arrowheads on the right of each panel indicate positions of dG*s*/RNAs dimmers (between 1000-3000bp) and free dGs (<100bp). Due to lower binding capacity of Sybr Green I to free RNA, it can be only slightly seen below the dG*s*/RNAs pairs. All dGs, which are shown here individually, were designed originally as antisenses (dG2, dG4, dG6, dG8, dG10) directly recognizing regulatory sequence within TRβ1 5'UTR b) senses (dG1, dG3, dG5, dG7, dG9) that could release homologous region by binding to a distant sequence folding within this region (S3 Table and S1 Fig). Antisense-like dGs (As) generate stronger band signals when compared to sense (S) dGs, which may share only partial complementarity with the distant 5'UTR sequence fragments. Although sense-like dGs exerted weak binding capacity, their translation-enhancing action could be released by partial complementarity with the distant 5'UTR sequences unfolding the homologous sequences or via interaction with other trans-acting factors. Scrambled control (dGsc) with an irrelevant (random) sequence revealed no interaction with the RNA, thus, confirming the specificity of binding by other dGs. Due to different exon 1e/2a boundary of variant F compared to variant A (ex1c/2a), only half of dG2 and dG6 shares sequence with 5'UTR variant F (pKS-F) that results in a weak binding capacity (electrophoretogram b). GeneRuler DNA Ladder Mix (Thermo Scientific) served as a marker ladder (L), shown on the right and left of each gel. The observed binding capacity and selectivity of dGs was also tested independently, by dG-primed reverse transcription (S8 Fig).




**S8 Fig. Binding selectivity confirmed by dGoligo-primed reverse transcription.**

Here we show PCR products obtained with dGoligos (dGs) used to drive synthesis of specific cDNAs in reaction of reverse transcription. This approach was based on primer extension by reverse transcriptase, which requires complementarity between a target sequence and, at least, 3'-end of an oligonucleotide (S1 Appendix). Panels a, b, c and d present selected SYBR Green I stained agarose gel electrophoretograms showing PCR products obtained on the basis of cDNA matrixes that were synthesized using antisense-like dG*s* including dG2, dG4, dG4' (dG4 from control synthesis), dG6, dG8, dG10 and control dGsc (scrambled). (**a**) PCR-amplified DNA fragments obtained on the basis of pKS-A cDNA, synthesized using one of the mentioned antisense-like dGs or dGsc. Common forward primer recognizing 5' end of exon 1c (P1) and one of the dGs (dG2, 4, 6, 8, 10) as a reverse primer were used in the PCR. (**b**) PCR fragments obtained on the basis of pKS-F cDNA, synthesized before using one of dGs or dGsc. Common forward primer recognizing 5' end of exon 1e (P2) and one of the dGs as a reverse primer were used. (**c**) Internal PCR fragments obtained on the basis of pKS-A cDNA, synthesized before using one of the dGs or dGsc. Common forward (P3) and reverse (P4) primer recognizing 3' end of exon 1c and exon 2a were used. (**d**) PCR fragments obtained on the basis of pKS-F cDNA, synthesized before using one of the dGs or dGsc. Common forward (P5) and reverse (P6) primer recognizing 3' end of exon 1a and exon 2a were used. ϕ indicates a control PCR-sample containing $H_2O$ instead of the cDNA. Arrowheads on the left of each panel indicate size of bands of marker ladder (L). Except for dGsc, all tested dGs mediated reverse transcription showing their binding selectivity (a, b, c, d). However, due to different exon 1e/2a boundaries of variant F compared to variant A (exon 1c/2a), only half of dG2 and dG6 shares sequence with 5'UTR variant F (pKS-F). In case of 5'UTR variant F, we observed lack of PCR products primed by dG2 and dG6 (b, d), confirming that 3'-end of these dGs do not form non-specific base-pairs with their targets.




**S9 Fig. DNA-based dGoligo-mediated effects in Caki-2 cells.**

Here we show that DNA dGoligos (dGs) have no significant effects on luciferase transcription (**a**) and translation (**b**) in Caki-2 cells transfected with pGL3-A. These data are consistent with previously reported findings that unmodified deoxyoligonucleotides can be rapidly degraded by nucleases and are of limited utility in mammalian cells (see results in the article). Results from three independent experiments performed in triplicates are shown as mean % mRNA (a) or luciferase activity (b) ± SD. Data analyzed by ANOVA followed by Dunnett's multiple comparison test. *$p<0.001$ vs. control.