The RNA secondary structures located at the interchromosomal region of human ACAT1 chimeric mRNA are required to produce the 56-kD isoform

Jia Chen[1,#], **Xiao-Nan Zhao**[1,#], **Li Yang**[1], **Guang-Jing Hu**[1], **Ming Lu**[1], **Ying Xiong**[1], **Xin-Ying Yang**[1], **Catherine C. Y. Chang**[2], **Bao-Liang Song**[1], **Ta-Yuan Chang**[2], and **Bo-Liang Li**[1,*]

[1]State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

[2]Department of Biochemistry, Dartmouth Medical School, Hanover, NH 03756; USA

Abstract

We have previously reported that human ACAT1 gene produces a chimeric mRNA through an interchromosomal processing of two discontinuous RNAs transcribed from chromosomes 1 and 7. The chimeric mRNA uses $AUG_{1397-1399}$ and $GGC_{1274-1276}$ respectively as translation initiation codons to produce the normal 50-kD ACAT1 and a novel enzymatically active 56-kD isoform, which is authentically present in human cells including human monocyte-derived macrophages. In this work, we report that RNA secondary structures located in the vicinity of the $GGC_{1274-1276}$ codon are required for producing the 56-kD isoform. The effects of the three predicted stem-loops (nt 1255–1268, 1286–1342 and 1355–1384) were tested individually by transfecting expression plasmids, which contain the wild-type, deleted or mutant stem-loop sequences linked with the partial ACAT1 AUG-open reading frame (ORF) or with the ORFs of other genes, into cells. The expression patterns were monitored by Western blot analyses. We found that the upstream stem-$loop_{1255-1268}$ from chromosome 7 and downstream stem-$loop_{1286-1342}$ from chromosome 1 were needed for production of the 56-kD isoform, whereas the last stem-$loop_{1355-1384}$ from chromosome 1 was dispensable. The results of experiments using both the monocistronic and bicistronic vectors with a stable hairpin showed that the translation initiation from the $GGC_{1274-1276}$ codon was mediated by internal ribosome entry site (IRES). Further experiments revealed that the translation initiation from the $GGC_{1274-1276}$ codon required the upstream RNA secondary structure with AU-constitution and the downstream one with GC-richness. This mechanistic work further supports the biological significance that the chimeric human ACAT1 mRNA is expressed from two different chromosomes.

Keywords

human ACAT1 isoform; chimeric human ACAT1 mRNA; interchromosomal region; RNA secondary structure; internal ribosome entry site

---

[*]Corresponding author, Professor Bo-Liang Li, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China, Tel: (86-21)-5492-1278, Fax: (86-21)-5492-1279 blli@sibs.ac.cn.
[#]These authors contributed equally to this manuscript

Introduction

Acyl-coenzyme A:cholesterol acyltransferase (ACAT) is a key enzyme in cellular cholesterol metabolism. It catalyzes the formation of cholesteryl ester from cholesterol and long-chain fatty acyl-coenzyme A in various cell types [1]. ACAT controls the dynamic interconversion between cellular free cholesterol and cholesterol ester. Two ACAT genes (encoding ACAT1 and ACAT2) have been identified; ACAT1, an allosteric tetramer enzyme [2], is expressed ubiquitously in all human tissues examined, whereas ACAT2 is mainly expressed in the embryonic liver and in intestine [3–5]. Under pathological condition, ACAT1 plays a central role in the accumulation of cholesterol ester in macrophages during early stage of the atherosclerosis disease [6–8]. Surprisingly, the genomic DNAs that encode the human ACAT1 cDNA K1 [9] is located in two different chromosomes (1 and 7): exons 1–16 are located in chromosome 1, while the optional long exon Xa (1279 bp) is located in chromosome 7 [10]. Northern blot analyses revealed the presence of four ACAT1 mRNAs (7.0, 4.3, 3.6 and 2.8 knt) in almost all of the human tissues and cells examined. These mRNAs share the same sequence located at chromosome 1, but only the 4.3-knt mRNA contains the long exon Xa, which is located at chromosome 7. Thus, this mRNA is produced from two different chromosomes by a novel RNA recombination event that presumably involves interchromosomal trans-splicing [10]. The human ACAT1 cDNA K1 is derived from this mRNA.

RNA secondary structures are involved in regulation of protein expression in multiple manners. For examples, in the cap-dependent ribosome scanning model, RNA secondary structures located at the 5'-untranslated region (5'-UTR) of mRNA are inhibitory for translation, while those downstream to the initiation codon can increase the translational efficiency [11]. On the other hand, in the case of cap-independent internal ribosome entry, RNA secondary structures locating at the long 5'-UTR can facilitate the internal translation initiation by recruiting various translation initiation factors, especially when the translation is initiated from non-AUG codons [12, 13]. Additionally, at RNA splicing level, it has been reported that RNA secondary structures can recruit several positive (B52, SRp55, and NOVA-1) and negative (hnRNP A1) regulatory protein factors [14, 15]. We had previously shown that, for the ACAT1 mRNAs without the optional exon Xa, an upstream stem-loop structure can enhance the selection of the adjacent downstream AUG to produce the 50-kD human ACAT1 protein [16]. For the chimeric 4.3-knt ACAT1 mRNA, we note that various predicted RNA secondary structures exist within the optional long 5'-UTR, but whether they are related to the production of the 56-kD isoform has not been tested.

Previously, the chimeric 4.3-knt ACAT1 mRNA was expressed in mutant CHO cells lacking endogenous ACAT1, and it was shown that this mRNA produces a 50-kD ACAT1 protein from the initiation codon $AUG_{1397–1399}$ and a novel ACAT1 isoform, with an apparent molecular mass of 56 kD [17]. The 56-kD isoform alone is enzymatically active and its normalized activity is 30% of the activity of the 50-kD ACAT1 [17]. When the 56-kD isoform and 50-kD ACAT1 were co-expressed in the same cell, the normalized ACAT1 activity was about 50% of the activity of the 50-kD ACAT1 protein alone [17]. Deletion analysis demonstrated that sequence of the human ACAT1 gene located at both chromosomes 7 and 1 is required to produce the 56-kD isoform. We also showed that the partial sequence (nt 1243–1786) of human ACAT1 mRNA is sufficient to produce a 25-kD N-terminal truncation of the 56-kD ACAT1 isoform. Further mutation and mass spectrometry analysis showed that $GGC_{1274–1276}$ was used as the initiation codon to produce the 56-kD ACAT1 isoform [17]. Importantly, the native 56-kD ACAT1 isoform can be expressed in human THP-1 macrophages and in human monocyte-derived macrophages examined [17]. Furthermore, limited proteolysis of the native 56-kD isoform from macrophages and the recombinant one from transfected cells by V8 protease confirmed that

the native 56-kD ACAT1 isoform are the same as those expressed in transfected cells [17]. Given that the optional long 5'-UTR in the chimeric human ACAT1 mRNA will prevent the scanning of 40S ribosome complex and the translation of the native 56-kD isoform is initiated from a non-AUG codon, we suppose that the translation initiation from the $GGC_{1274-1276}$ codon may be mediated by some special RNA secondary structures.

In the current work, we analyzed whether the three predicted stem-loops (nt 1255–1268, 1286–1342 and 1355–1384) located either upstream and downstream of the $GGC_{1274-1276}$ codon could affect the production of proteins initiated from this codon. Our results revealed that both the upstream stem-loop$_{1255-1268}$ and downstream stem-loop$_{1286-1342}$ from two different chromosomes are required for producing the 56-kD human ACAT1 isoform.

Materials and Methods

Materials

Cell culture reagents and T4 DNA ligase were purchased from Invitrogen (Carlsbad, USA). All the restriction enzymes and agarose were from Promega (Madison, USA). Anti-rabbit and anti-mouse antibodies (IgGs) conjugated with HRPs were from Pierce (Rockford, USA). Taq DNA polymerase and dNTPs were from Sino-American Biotech (Shanghai, China). All the oligonucleotides were synthesized with an automated DNA synthesizer at Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences.

Cell culture and Transfection

AC29, the mutant Chinese hamster ovary (CHO) cells lacking the endogenous ACAT1 [18], was maintained in a basal Ham's F12 medium, supplemented with 10% fetal bovine serum (FBS) and antibiotics, in a humid atmosphere of 5% CO_2 and 95% air at 37°C. The constructed expression plasmids were individually transfected into AC29 cells by using FuGENE 6™ transfection reagent (Roche) according to the manufacturer's instructions and the transfected cells were cultured for another 48 h for harvesting.

Expression plasmids

The sequence encoding 3×Flag was amplified from p3×FLAG-CMV-14 expression vector (Sigma) by PCR with the primer set (3FlagF, 5'-AGTGAACCGTCAGAATTAAGC-3'/ 3FlagR, 5'-AAAGGGCCCATCACTACTTGTCATCGTC-3). The amplified fragment was then purified, digested with Xba I and Apa I, and inserted into the Xba I and Apa I sites of pcDNA3 (Invitrogen) to generate vector p3Flag for the tagged expression by fusing 3×Flag at C-terminal.

The partial human ACAT1 cDNA K1 sequence (bp 1243–1786) containing the vicinity of $GGC_{1274-1276}$ codon followed with the partial AUG-ORF sequence, which can encode 25-kD N-terminus of 56-kD isoform (ACAT1-NT25) from the $GGC_{1274-1276}$ codon and 17-kD N-terminus of 50-kD isoform (ACAT1-NT17) from $AUG_{1397-1399}$ codon, was amplified by PCR with the primer set of (1243F, 5'-AAAGGTACCTAGTTAAATAGCTATATTTAT-3'/ 1786R, 5'-AAATCTAGAATCTAAGAGAGAGCGCCT-3') and inserted into the Kpn I and Xba I sites of p3Flag to generate expression plasmid pNTF for productions of 26-kD ACAT1 N-terminus with 3×Flag (ACAT1-NT-Flag26) from the $GGC_{1274-1276}$ codon and 18-kD ACAT1 N-terminus with 3×Flag (ACAT1-NT-Flag18) from $AUG_{1397-1399}$ codon. Then, the sequential deletions of the vicinity of $GGC_{1274-1276}$ codon (bp 1243–1396) upstream to $AUG_{1397-1399}$ codon were obtained by using the constructed plasmid pNTF. The 5'-region deletion (Δ1243–1268) was gained by the above one-step PCR with the forward primer D1F (Table 1) and reverse primer 1786R. Other three deletions (Δ1280–

1288, Δ1289–1339 and Δ1340–1396) were obtained by a two-step PCR method described by Higuchi [19] with the common external primer set (1243F/1786R, shown above) and individual internal primer sets (D2F/D2R, D3F/D3R and D4F/D4R, shown in Table 1) with partial complementary sequences. All the amplified fragments with the sequential deletions of the vicinity of $GGC_{1274-1276}$ codon upstream to $AUG_{1397-1399}$ codon were respectively inserted into the Kpn I and Xba I sites of p3Flag to generate expression plasmids pNTF-D1 to -D4.

By using pNTF and pNTF-D1 to -D4, the partial ACAT1 AUG-ORF was replaced with the whole Rluc AUG-ORF (encoding Renilla luciferase, 316 amino acids with apparent molecular mass of 36 kD) to obtain expression plasmids pNTR and pNTR-D1 to -D4. The coding sequence of Rluc was amplified from pRL-CMV (Promega) by PCR with primer set (RlucF1, 5'-AAAGGCGGCCGCGACTTCGAAAGTTTATGATCCAGAACAAAGG-3'/ RlucR1, 5'-AAAGTCTAGATTGTTCATTTTTGAGAACTCGCTCAACGAA-3') and inserted into the Not I and Xba I sites of pcDNA3 to generate vector pRL. The upstream sequences of $AUG_{1397-1399}$ in pNTF, pNTF-D2 and pNTF-D3 were amplified with the forward (1243F shown above) and reverse (1396R, 5'-TCTGCGGCCGCCATTGTATTGTCTGAGGCCCGCGCCC-3') primers. The upstream sequences of $AUG_{1397-1399}$ in pNTF-D1 and pNTF-D4 were respectively amplified with the above primer sets 1269F/1396R and 1243F/1396R-2 (5' AAGCGGCCGCCATGGAAGCG 3'). These amplified fragments were individually digested with Kpn I and Not I and inserted to the Kpn I and Not I sites of pRL to generate the expression plasmids pNTR, pNTR-D1 to -D4.

The fragment involving a stable hairpin that can block the scanning of 40S ribosome-factor complex [20] was annealed from two complement primers (HF1, 5'-AGCTTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCGGTAC-3'/HR1, 5'-CGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCA-3') and inserted into the Hind III and Kpn I sites of p3Flag to generate the stable hairpin containing vector ph3Flag. The partial human ACAT1 cDNA K1 sequence (bp 1243–1786) from pNTF was inserted into the Kpn I and Xba I sites of ph3Flag to generate the expression plasmid phNTF. The partial ACAT1 AUG-ORF sequence (1396–1786) was amplified with the forward (D5F, shown in Table 1) and reverse (1786R, shown above) primers and inserted into the Kpn I and Xba I sites of p3Flag and ph3Flag respectively to generate expression plasmids pNTF-D5 and phNTF-D5.

The Nhe I site was introduced into the pcDNA3 with the primer set (pcDNA3-Nhe-F, 5'-CCCGAAAAGTGCCACCTGACGTCGACGGAT-3'/ pcDNA3-Nhe-R, 5'-AAAAAGCTTACGCTAGCGGGTCTCCCTATAGTGAGTCGTATTAATTTC-3') to generate the vector pcDNA3-NheI. The whole ORF of Fluc was amplified with the primer set (FlucF1, 5'-AAATCTAGAATGGAAGACGCCAAAAACATA-3'/FlucR, 5'-AAAGGGCCCTTACACGGCGATCTTTCC-3') and inserted into the Xba I and Apa I sites of pcDNA3-NheI to generate the plasmid pFL. The whole ORF of Rluc was amplified with the primer set (RlucF2, 5'-AAAAAGCTTATGACTTCGAAAGTTTATGAT-3'/RlucR2, 5'-AAAGCGGCCGCTTATTGTTCATTTTTGAGAACT-3') and inserted into the Hind III and Not I sites of the plasmid pFL to generate the plasmid pRnF. The vicinity (1243–1396 bp) of the $GGC_{1274-1276}$ was amplified with the primer set (1243-NotI-F, 5'-AAGCGGCCGCTAGTTAAATAGCTATATTTAT-3'/1396-XbaI-R, 5'-AAATCTAGACATGGAAGCGGTCACAGA-3') and insert into the Not I and Xba I sites of pRnF to generate the plasmid pRAF. The annealed stable hairpin (HF2, 5'-CTAGCGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCA-3'/HR2, 5'-AGCTTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCG-3') was inserted into the Nhe I and Hind III sites of pRAF and pRnF to generate the plasmids

phRAF and phRnF. The annealed stable hairpin (HF3, 5'-
GGCCGCGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCA-3'/HR3, 5'-
AGCTTGGGGCGCGTGGTGGCGGCTGCAGCCGCCACCACGCGCCCCGC-3') was
inserted into the Hind III and Not I sites of pFL, then the coding sequence of Rluc was
amplified with the primer set (RlucF3, 5'-
AAAGCTAGCATGACTTCGAAAGTTTATGAT-3'/RlucR3, 5'-
AAAAAGCTTTTATTGTTCATTTTTGAGAACT-3') and also inserted into the pFL to
generate the plasmid pRhnF. The vicinity of the $GGC_{1274-1276}$ (1243–1396 bp) was
amplified with the primer set (1243-NotI-F, shown above/1396-XbaI-R, shown above) was
inserted into the Not I and Xba I site to generate the expression plasmid pRhAF.

The whole ORFs of Rluc and Fluc were amplified with the primer sets (RlucF4, 5'-
AAAAGATCTATGACTTCGAAAGTTTATGAT-3'/RlucR2, shown above; FlucF2, 5'-
AAAAGATCTATGGAAGACGCCAAAAACATA-3'/FlucR, shown above) and inserted
into the Bgl II and Not I sites, which does not contain the eukaryotic CMV promoter, of
pRAF to generate the plasmids pRAF-ΔCMV and pFl-ΔCMV. The IRES sequence of HCV
was amplified with the primer set (HCVIF, 5'-
AAAGCGGCCGCGCCAGCCCCCGATTGGGG-3'/HCVIR, 5'-
AAATCTAGAGGTTTTTCTTTGAGGTTTAGGA-3') from the plasmid pC1b (from Prof.
Akio Nomoto) and inserted into the Not I and Xba I sites, which contains the eukaryotic
CMV promoter, of pRAF to generate the plasmid pRHIF.

The partial ACAT1 cDNA K1 sequences (ACAT1 1243–1786) in the plasmids pNTF-D1 to
-D4 were inserted into the Kpn I and Xba I sites of ph3Flag to generate expression plasmids
phNTF-D1 to -D4. The vicinity of the $GGC_{1274-1276}$ (1243–1396 bp) in pNTF-D2 and
pNTF-D3 were amplified with the primer set (1243-NotI-F, shown above/1396-XbaI-R,
shown above). The vicinity of the $GGC_{1274-1276}$ (1243–1396 bp) in pNTF-D1 and pNTF-
D4 were respectively amplified with the primer sets (1269-NotI-F, 5'-
AAGCGGCCGCTCCAGGGCACCCCGAAT-3'/1396-XbaI-R, shown above) and (1243-
NotI-F, shown above/1396-XbaI-R, 5'-AAATCTAGACATGGAAGCGGTCACAGA-3').
These amplified fragments were inserted to the Not I and Xba I sites of pRAF to generate
the expression plasmids pRAF-D1 to -D4.

The sub-deletions (Δ1289–1306, Δ1307–1324, Δ1325–1339) of the stem-loop II region (nt
1289–1339) were performed by two-step PCR with the above common external primer set
(1243F/1786R) and individual internal primer sets (D6F/D6R, D7F/D7R and D8F/D8R,
shown in Table 1). The amplified fragments were respectively inserted into the Kpn I and
Xba I sites of p3Flag to generate expression plasmids pNTF-D6 to -D8.

The mutations were introduced into the stem-loop I by one-step PCR with individual
forward primers M1F to M8F with mutated nucleotides shown in Table 2 and common
reverse primer 1786R. By using vector p3Flag, the expression plasmids pNTF-M1 to -M8
containing the desired mutations were constructed by the same method described above.

The mutations were introduced to the regions of RNA secondary structures from plasmid
pNTF-D6 to -D8 also by two-step PCR with the above common external primer set (1243F/
1786R) and individual internal primer sets (M9F/M9R and M10F/M10R for pNTF-D5,
M11F/M11R and M12F/M12R for pNTF-D6, M13F/M13R and M14F/M14R for pNTF-D7,
shown in Table 3). The amplified fragments were respectively inserted into the Kpn I and
Xba I sites of p3Flag to generate expression plasmids pNTF-M9 to -M14.

Restriction enzyme digestion and DNA sequencing confirmed all the constructed plasmids.

Preparation of protein samples and Western blot analysis

Cells were washed twice, scraped in ice-cold PBS and extracted on ice for 30 min in RIPA buffer containing protease inhibitor mixture (Sigma). Cell debris was removed by a spin at 16,000×g. The protein concentrations were determined using the BCA protein assay kit (Bio-Rad). For Western blot analysis, 50 µg of total protein per lane was resolved in SDS-PAGE. The filters were probed with an ACAT1 antibody DM10 [21], anti-Flag (M2, Sigma), anti-Fluc (ab7358, Abcam) or anti-Rluc (MAB4400, Chemicon) antibody. After incubation with horseradish-conjugated secondary antibody, the signals were developed using ECL Western Blotting detection reagent (Pierce).

Prediction of RNA secondary structures

All the predicted secondary structures with folding Gibbs free energy (ΔG) in this manuscript were determined by the mFold program (version 3.2) [22].

Reverse transcription-PCR (RT-PCR) and -quantitive PCR (RT-qPCR) analysis

Total RNA isolated by Trizol (Invitrogen) from transfected cells was pretreated with RQ DNase I (Promega) and analysed by RT-PCR using two sets of primers: F1 (AAAGGAAACGGATGATAACTGGTC)/R1 (TCTCTTCATAGCCTTATGCAGTTGC), located at the 5' regions of Rluc and Fluc genes, respectively, and F2 (TCAAATCGTTCGTTGAGCGAGT)/R2 (CTTGCGTCGAGTTTTCCGGTAA), located at the 3' regions of Rluc and Fluc genes, respectively.

The quantification of ACAT1 transcripts was done by qPCR, using Brilliant SYBR Green qPCR Master Mix and a Light Cycler apparatus (Stratagene). The primer sets for GAPDH (glyceraldehyde-3- phosphate dehydrogenase gene) and ACAT1 were GAPDHF (TGGCTACAGCAACAGAGTGG)/GAPDHR (GGGGTTATTGGACAGGGACT) and ACAT1F (CGGGCCTCAGACAATACAAT)/ACAT1R (TCAATTCCTCTGCCTCTGCT). The human ACAT1 mRNA levels were normalized to the human GAPDH mRNA in each sample.

Other methods

Standard molecular biology techniques were performed according to the methods described by Sambrook et al [23].

Results

The predicted stem-loops from two different chromosomes are required for the production of protein initiated from the GGC$_{1274–1276}$ codon

Our previous study has indicated that a native 56-kD isoform can be produced from the chimeric human ACAT1 mRNA by using GGC$_{1274–1276}$ as the initiation codon, and the minimal region required for this non-AUG codon initiated translation has been narrowed to nt 1243–1786.[17] The in-frame codon GGC$_{1274–1276}$ is located at the sequence (nt 1243–1396) upstream to the AUG-ORF for normal 50-kD human ACAT1 protein (Figure 1A). In the regions upstream and downstream to the GGC$_{1274–1276}$ codon, there are three predicted stem-loops. The stem-loop$_{1255–1268}$ (stem-loop I) is transcribed from the chromosome 7, while most of the stem-loop$_{1286–1342}$ (stem-loop II) and the whole stem-loop$_{1355–1384}$ (stem-loop III) are transcribed from the chromosome 1 (Figure 1B). The stem-loop I is constituted with A and U nucleotides only; the stem-loop II downstream to the GGC$_{1274–1276}$ codon is more complex than other two, and contains a GC-rich stem with multiple bulges; the stem-loop III downstream to the stem-loop II and adjacent to the AUG-ORF also has a high GC content, and contains only one bulge in the stem. Previous work showed that the stem-loop

III could modulate the selection of its downstream AUG as the translation initiation codon to produce the normal 50-kD human ACAT1 protein [16].

In order to test the effect of predicted interchromosomal stem-loops on the production of protein initiated from the $GGC_{1274–1276}$ codon, the plasmid was constructed to contain the nt 1243–1786 of human ACAT1 cDNA K1 (ACAT1 1243–1786), which is the minimal sequence identified for the proteins translated from the $GGC_{1274–1276}$ codon [17], with 3×Flag sequence (3×Flag) at the 3'-end. Then, the sequential deletions (Δ1243–1268, Δ1280–1288, Δ1289–1339 and Δ1340–1369) were done (Figure 1C, left panel). After transfection with these expression plasmids, the productions of both proteins initiated from the in-frame $GGC_{1274–1276}$ and $AUG_{1397–1399}$ codons were observed. As shown in right panel of Figure 1C, deletion of stem-loop I or II dramatically reduced the production of protein initiated from the $GGC_{1274–1276}$ codon, but had little effect on the production of protein initiated from the $AUG_{1397–1399}$ codon (No. 2 and 4). Deletion of stem-loop III did not have any significant effect on production of protein initiated from the $GGC_{1274–1276}$ codon (Figure 1C, right panel, No. 5). The same results were obtained when anti-Flag antibody (M2) was employed for Western blot analysis (data not shown).

To study the effect of the partial ACAT1 AUG-ORF sequence (nt 1397–1786) on the translation initiation from the in-frame $GGC_{1274–1276}$ and $AUG_{1397–1399}$ codons, we replaced this sequence with the whole AUG-ORF of Renilla luciferase (Rluc) to generate the related expression plasmids (Figure 1D, left panel). After transfection, Western blotting with anti-Rluc antibody was performed to examine the protein expression of these plasmids. All the results in right panel of Figure 1D similar to those in right panel of Figure 1C indicated that the partial ACAT1 AUG-ORF sequence has no effect on production of protein initiated from either $GGC_{1274–1276}$ or $AUG_{1397–1399}$ codon. Additionally, due to the deletions in the coding region of the protein initiated from the $GGC_{1274–1276}$ codon, the proteins shorter than the control one in right panels of Figure 1C and 1D were observed.

These results indicate that both stem-loop I (nt 1255–1268, from chromosome 7) and stem-loop II (nt 1286–1342, from chromosome 1), but not stem-loop III (nt 1355–1384, from chromosome 1), are required for the production of proteins initiated from the $GGC_{1274–1276}$ codon. In conclusion, the optimal production of proteins from the $GGC_{1274–1276}$ codon needs two different upstream and downstream RNA secondary structures transcribed from two different chromosomes.

The translation initiation from the $GGC_{1274–1276}$ codon is mediated by internal ribosome entry site

Considering the long 5'-UTR in the chimeric human ACAT1 mRNA and the translation initiation from a non-AUG codon for the 56-kD isoform, we proposed that the translation initiation from the $GGC_{1274–1276}$ codon, which needed the upstream stem-loop I and downstream stem-loop II, might be mediated by internal ribosome entry site (IRES). So far, two strategies of constructing monocistronic or bicistronic vectors involving a stable hairpin had been used for identification of different IRESes [24–26]. In our case, these two strategies were used to test if translation initiation from the $GGC_{1274–1276}$ codon was mediated by IRES. On one hand, the nt 1243–1786 of human ACAT1 cDNA K1 (ACAT1 1243–1786) with 3 ×Flag sequence (3 ×Flag) at the 3'-end was constructed into the monocistronic vector with a 5'-stable hairpin (ΔG = −57 kcal/mol) that can impair the cap-dependent ribosome scanning (Figure 2A, No. 11). Two negative control plasmids (Figure 2A, No. 12 and 13) were further constructed by deleting the vicinal sequence of the $GGC_{1274–1276}$ codon (Δ1243–1396) containing the predicted stem-loops in the constructed monocistronic plasmids with or without a 5'-stable hairpin. The plasmid phNTF (No. 11) was expected to express the protein initiated from the $GGC_{1274–1276}$ codon only if there was

IRES in the vicinity of this codon (nt 1243–1396). On the other hand, three bicistronic vectors (pRnF, phRnF and pRhnF) without any intercistronic sequence were constructed to contain the first (whole AUG-ORF of Renilla luciferase, Rluc) and second (whole AUG-ORF of Firefly luciferase, Fluc) cistrons without (Figure 2C, No.15) or with (Figure 2C, No. 17 and 19) the stable hairpin ($\Delta G = -57$ kcal/mol) (see "Materials and Methods"). Next, with these bicistronic vectors, three expression plasmids (pRAF, phRAF and pRhAF) were further constructed to comprise the second cistron by fusing the vicinity of $GGC_{1274–1276}$ codon (nt 1243–1396) with the 5'-end of the whole AUG-ORF of Firefly luciferase (Fluc) (Figure 2C, No.14, 16 and 18), and used for detecting IERS in the vicinal sequence of the $GGC_{1274–1276}$ codon (nt 1243–1396). If the fused proteins (ACAT1-Fluc) initiated from the $GGC_{1274–1276}$ codon of the fused second cistron were detected by using all the expression plasmids, both the ribosomal read-through over the first cistron (Rluc) and reinitiation of the fused second cistron (ACAT1-Fluc) will be ruled out to further confirm that there is IRES in the vicinity of the $GGC_{1274–1276}$ codon (nt 1243–1396).

After transfection with these expression plasmids, Western blotting was performed to detect the expressed proteins. The results of the experiments by using the monocistronic plasmids (Figure 2A) showed that the amount of proteins initiated from the $GGC_{1274–1276}$ codon were apparently unaltered whether the 5'-stable hairpin existed or not (Figure 2B, No. 1 and 11), while no target protein can be examined for the negative controls without the vicinity of this codon (Figure 2B, No. 12 and 13). It was evidently indicated that there was IRES in the vicinity of the $GGC_{1274–1276}$ codon (nt 1243–1396), which mediated the translation from this initiation codon in a cap-independent manner (Figure 2B, No. 1 and 11). Further results obtained with three kinds of bicistronic expression plasmids demonstrated that the level of the fused proteins (ACAT1-Fluc) translated from the $GGC_{1274–1276}$ codon of the fused second cistron were similar without (Figure 2D, top panel, No. 14) or with the stable hairpin at the position upstream or downstream to the first cistron (Rluc) (Figure 2D, top panel, No. 16 and 18). Whereas, the 5'-stable hairpin potently impaired the translation of the first cistron (Rluc) in the cap-dependent manner (Figure 2D, bottom panel, No. 16 and 17), and the second cistron (Fluc) in the controls without any intercistronic sequence did not be translated (Figure 2D, top panel, No. 15, 17 and 19). Conclusively, these further evidences illustrate that the translation from the $GGC_{1274–1276}$ codon is mediated by IRES in the cap-independent manner. In addition, the proteins initiated from the downstream in-frame $AUG_{1397–1399}$ codon were also observed both from the monocistron with the 5'-stable hairpin (Figure 2B, No. 11) and from the fused second cistron (Figure 2D, top panel, No. 14, 16 and 18) only when the upstream vicinity of the $GGC_{1274–1276}$ codon (nt 1243–1396) was present.

Although the foregoing data have provided strong evidences that the translation from the $GGC_{1274–1276}$ codon is mediated by IRES, another two possibilities still cannot be excluded through the above experiments. One is that the cryptic promoter activity may result in transcription of the RNAs from the vicinity of the $GGC_{1274–1276}$ codon (nt 1243–1396), which might translate the proteins. The other is that the presence of splicing sites may lead to the expression of proteins. To rule out the first one, the transfection was performed with the monocistronic and bicistronic plasmids with or without the eukaryotic CMV promoter (Figure 3A). Western blotting showed that protein expression could not be detected in absence of the eukaryotic CMV promoter (Fig 3B, No. 20 and 22). Obviously, it is indicated that there is no cryptic promoter in the vicinity of the $GGC_{1274–1276}$ codon (nt 1243–1396) from this result. To rule out the second one, the integrity of the bicistronic mRNAs was examined from cells transfected with bicistronic plasmids (Figure 3C). RT-PCR analysis of the total RNA from pRAF (No. 14) transfected cells showed that only the full-length bicistronic mRNAs (Figure 3B, lanes b and d) were present *in vivo*, similar to the results from the cells transfected with the control plasmid pRHIF (No. 23) containing the HCV

IRES (Figure 3B, lanes f and h). The evident results demonstrate that there is no splicing site in the vicinity of the $GGC_{1274-1276}$ codon (nt 1243–1396).

Collectedly, all of these data showed that the translation initiation from the $GGC_{1274-1276}$ codon was mediated by IRES. Then, the above sequential deletions in the vicinity of the $GGC_{1274-1276}$ codon (nt 1243–1396) shown in Figure 1C and 1D were also performed to generate the monocistronic plasmids with the 5'- stable hairpin and the bicistronic plasmids containing the first (Rluc) and second (Fluc) cistrons (Figure 4A and 4C). Western blotting showed the similar results that the deletion of stem-loop I (Δ1243–1268) or II (Δ1289–1339) evidently reduced the production of protein initiated from the $GGC_{1274-1276}$ codon (Figure 4B, No. 24 and 26; Figure 4D, No. 28 and 30). It is confirmed that this translation initiation from the $GGC_{1274-1276}$ codon mediated by IRES requires both upstream stem-loop I and downstream stem-loop II from two different chromosomes.

AU-constitution of upstream RNA secondary structure from chromosome 7 is important for the translation initiation from the $GGC_{1274-1276}$ codon

To characterize the effect of the upstream RNA secondary structure from chromosome 7 on translation initiation from the $GGC_{1274-1276}$ codon, various mutations were introduced into the stem-loop I (Figure 5A). At first, twelve A-U interconversions (nt 1255–1259, nt 1261–1263 and 1264–1268, pNTF-M1) that do not disrupt the secondary structure were introduced into the stem-loop I (Figure 5C, No. 32), and the production of both proteins initiated from the $GGC_{1274-1276}$ and $AUG_{1397-1399}$ codons was similar as wild-type one (Figure 5B, No. 1 and 32). When three U to A substitutions were introduced (pNTF-M2) to disrupt the step-loop I (Figure 5C, No. 33), the production of protein initiated from the $GGC_{1274-1276}$ codon was dramatically reduced, while the production of protein initiated from the $AUG_{1397-1399}$ codon did not changed obviously (Figure 5B, No. 33). Interestingly, two additional A to U substitutions (pNTF-M3) to rebuild this stem-loop (Figure 5C, No. 34) can completely rescue the production of protein initiated from the $GGC_{1274-1276}$ codon (Figure 5B, No. 34). However, disruption of stem-loop I by mutating UAUU to CGCG (pNTF-M4) reduced the production of protein initiated from the $GGC_{1274-1276}$ codon (Figure 5B, No. 35), which can not be rescued by rebuilding the stem-loop with GC-rich stem (Figure 5B, No. 36).

According to the above results, we proposed that the stability of upstream stem-loop (Figure 5C, No. 36) was related to the production of protein initiated from the $GGC_{1274-1276}$ codon (Figure 5B, No. 36). In further experiments, the A-U base pairs in the stem were progressively changed to G-C base pairs (Figure 6A) to increase the stability of this stem-loop (Figure 6C). Western blotting displayed that the production of protein initiated from the $GGC_{1274-1276}$ codon gradually decreased accompanying the increase of the GC content in the stem-loop (Figure 6B, No. 1 and 37–39), indicating that the production of protein initiated from the $GGC_{1274-1276}$ codon is inversely correlated with the stability of the upstream RNA secondary structures (Figure 6D).

Taken together, these results demonstrate that the upstream AU-constituted RNA secondary structure from chromosome 7 is important for the translation initiation of ACAT1 56-kD isoform from the $GGC_{1274-1276}$ codon.

GC-richness of downstream RNA secondary structure from chromosome 1 is essential for the translation initiation from the $GGC_{1274-1276}$ codon

Since the earlier result (Figure 1C) demonstrated that the production of protein initiated from the $GGC_{1274-1276}$ codon was significantly decreased when the whole stem-loop II was deleted, we performed sub-deletions (Δ1289–1306, Δ1307–1324 and Δ1325–1339) of this stem-loop (pNTF-D6 to -D8 in Figure 7A). Western blotting (Figure 7B) indicated that none

of these sub-deletions inhibited the production of protein initiated from the $GGC_{1274-1276}$ codon (No. 40–42) to the same extent as the complete deletion did (No. 4). By RNA secondary structure prediction, we found that the sequence (nt 1274–1354) with the complete deletion could hardly form a stable secondary structure (Figure 7C, No. 4), while the three sub-deletions could still form RNA secondary structures with relatively strong stabilities (comparing to the complete deletion) and with high GC contents (Figure 7C, No. 40–42). So, we proposed that a downstream GC-rich RNA secondary structure could be essential for the translation initiation from the $GGC_{1274-1276}$ codon.

To validate the hypothesis, we replaced most of the GC nucleotides with AU nucleotides in the downstream RNA secondary structure formed after sub-deletions of the stem-loop II region (Figure 8A). The mutated nucleotides in the mRNA produced from the constructed plasmids pNTF-M9, -M11 and -M13, disrupted the formed secondary structures, but for pNTF-M10, -M12 and -M14, the mutated nucleotides maintained these secondary structures with AU-richness but not GC-richness (Figure 8B). Western blotting showed that the production of proteins initiated from the $GGC_{1274-1276}$ codon were completely disappeared (Figure 8C, No. 43–44, No. 45–46 and No. 47–48) when most of the GC nucleotides were replaced by AU nucleotides, whether the downstream RNA secondary structures were disrupted or maintained (Figure 8B). It should be noticed that the alteration of the production of protein initiated from the $AUG_{1397-1399}$ codon was also observed when the AU nucleotides were introduced (Figure 8C, No. 43–48). Thus, to check if this alternation is caused by the change of the RNA, the RT-qPCR experiment was performed and the results showed no significant alteration among the related mRNAs (Figure 8D).

In summary, besides the upstream AU-constituted RNA secondary structure from chromosome 7, the downstream GC-rich RNA secondary structure from chromosome 1 is another essential property for the translation initiation from the $GGC_{1274-1276}$ codon.

Discussion

There are three stem-loops in the vicinity of the $GGC_{1274-1276}$ codon, which is located at the interchromosomal region of the chimeric human ACAT1 mRNA (Figure 1B). We find that two of these RNA secondary structures are required for the production of proteins initiated from the $GGC_{1274-1276}$ codon. The upstream stem-loop I from chromosome 7 is most effective for the production of protein from the $GGC_{1274-1276}$ codon when it is constituted with AU nucleotides, and this structure can tolerate some changes of the nucleotide sequences unless the structure is disrupted (Figure 5 and 6). The downstream GC-rich stem-loop II from chromosome 1 is the largest and the most complicated one of the three predicted stem-loops. Experimental results demonstrate that the production of protein initiated from the $GGC_{1274-1276}$ codon needs a downstream RNA secondary structure with GC-richness. The downstream secondary structure can also be formed after certain sub-deletions, even if their stabilities are relatively weaker than the wild type one (Figure 7). Further results show that the GC-richness of this downstream RNA secondary structure is indispensable (Figure 8). Detailed experiments show that the translation initiation from the $GGC_{1274-1276}$ codon is mediated by IRES (Figure 2 and 3). Forward evidences elucidate that this translation initiation from the $GGC_{1274-1276}$ codon mediated by IRES requires the upstream RNA secondary structure with AU-constitution and the downstream one with GC-richness (Figure 4). In addition, less protein initiated from the downstream in-frame $AUG_{1397-1399}$ codon in the cap-independent manner were also observed from both the monocistron with 5'-stable hairpin (Figure 2B, No. 11) and the fused second cistron (Figure 2D, No. 14, 16 and 18), only when the upstream vicinity of the $GGC_{1274-1276}$ codon (nt 1243–1396) was present. It implies that there might be another IRES in the nt 1243–1396 region, which can mediate the translation initiated from the $AUG_{1397-1399}$ codon. This

mechanistic work further supports the biological significance that the chimeric human ACAT1 mRNA is expressed from two different chromosomes.

Our previous studies showed that the 56-kD isoform was enzymatically active and its activity is about 30% that of the 50-kD ACAT1.[17] When the 56-kD isoform and 50-kD ACAT1 were co-expressed in the same cell, the normalized ACAT1 activity was obviously lower than that of 50-kD ACAT1 protein alone.[17] Considering that the 50-kD ACAT1 protein forms homotetramers in intact cells and *in vitro*,[2] it is possible that the 56-kD isoform might serve as an endogenous inhibitor of the 50-kD ACAT1 by forming hetero-oligomers to reduce the allosteric enzyme activity when the excessive high ACAT activity should be down-regulated. On the other hand, in some cell-stressed conditions that inhibit the cap-dependent translations, it is postulated that the dynamic interconversion between cellular free cholesterol and cholesterol ester could be still maintained by the 56-kD isoform initiated from the $GGC_{1274-1276}$ codon, which is mediated by IRES in the cap-independent manner. Conclusively, this 56-kD ACAT1 isoform could be produced under various conditions including the cellular stresses and provide a more elaborate regulation of this allosteric tetramer enzyme activity for the cellular cholesterol homeostasis.

Other investigators showed that most of the RNA secondary structures located at the long 5'-UTR could be functioned as IRES [27, 28] and in a few case, the downstream sequence of translation initiation codon was required for IRES activity [29]. In our case, both the upstream and downstream RNA secondary structures transcribed from two different chromosomes are required for the translation initiation from the $GGC_{1274-1276}$ codon mediated by IRES. It has been reported that RNA secondary structures perform their IRES functions by recruiting various proteins including translation initiation factors. For examples, the EMCV IRES needs to recruit eIF4A and eIF4G to initiate translation [30, 31]. Poliovirus and rhinovirus IRESes both bind to PTB and to the poly(rC) binding protein 2 (PCBP2) [32, 33]. In addition, RNA secondary structures are also involved in regulating the gene expression at multiple levels, such as splicing, polyadenylation and editing, etc [34–36]. We note that the upstream RNA secondary structure is AU-constituted and the downstream one is GC-rich, suggesting that the upstream and downstream RNA secondary structures may be needed to recruit and bind certain special proteins, for producing 56-kD human ACAT1 isoform initiated from the $GGC_{1274-1276}$ codon. In our oncoming studies, the proteins bound to the upstream and downstream RNA secondary structures from two different chromosomes will be identified, and how they regulate this translation initiation will be further tested.

Acknowledgments

References

1. Chang TY, Chang CC, Cheng D. Acyl-coenzyme A:cholesterol acyltransferase. Annu Rev Biochem. 1997; 66:613–638. [PubMed: 9242919]

2. Yu C, Chen J, Lin S, et al. Human acyl-CoA:cholesterol acyltransferase-1 is a homotetrameric enzyme in intact cells and in vitro. J Biol Chem. 1999; 274:36139–36145. [PubMed: 10593897]

3. Rudel LL, Lee RG, Cockman TL. Acyl coenzyme A: cholesterol acyltransferase types 1 and 2: structure and function in atherosclerosis. Curr Opin Lipidol. 2001; 12:121–127. [PubMed: 11264983]

4. Chang TY, Chang CC, Lin S, et al. Roles of acyl-coenzyme A:cholesterol acyltransferase-1 and −2. Curr Opin Lipidol. 2001; 12:289–296. [PubMed: 11353332]

5. Song BL, Wang CH, Yao XM, et al. Human acyl-CoA:cholesterol acyltransferase 2 gene expression in intestinal Caco-2 cells and in hepatocellular carcinoma. Biochem J. 2006; 394:617–626. [PubMed: 16274362]

6. Yang JB, Duan ZJ, Yao W, et al. Synergistic transcriptional activation of human Acyl-coenzyme A: cholesterol acyltransterase-1 gene by interferon-gamma and all-trans-retinoic acid THP-1 cells. J Biol Chem. 2001; 276:20989–20998. [PubMed: 11399774]

7. Yang L, Yang JB, Chen J, et al. Enhancement of human ACAT1 gene expression to promote the macrophage-derived foam cell formation by dexamethasone. Cell Res. 2004; 14:315–323. [PubMed: 15353128]

8. Chang C, Dong R, Miyazaki A, et al. Human acyl-CoA:cholesterol acyltransferase (ACAT) and its potential as a target for pharmaceutical intervention against atherosclerosis. Acta Biochim Biophys Sin. 2006; 38:151–156. [PubMed: 16518538]

9. Chang CC, Huh HY, Cadigan KM, Chang TY. Molecular cloning and functional expression of human acyl-coenzyme A:cholesterol acyltransferase cDNA in mutant Chinese hamster ovary cells. J Biol Chem. 1993; 268:20747–20755. [PubMed: 8407899]

10. Li BL, Li XL, Duan ZJ, et al. Human acyl-CoA:cholesterol acyltransferase-1 (ACAT-1) gene organization and evidence that the 4.3-kilobase ACAT-1 mRNA is produced from two different chromosomes. J Biol Chem. 1999; 274:11060–11071. [PubMed: 10196189]

11. Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene. 2005; 361:13–37. [PubMed: 16213112]

12. Huez I, Creancier L, Audigier S, et al. Two independent internal ribosome entry sites are involved in translation initiation of vascular endothelial growth factor mRNA. Mol Cell Biol. 1998; 18:6178–6190. [PubMed: 9774635]

13. Sasaki J, Nakashima N. Methionine-independent initiation of translation in the capsid protein of an insect RNA virus. Proc Natl Acad Sci U S A. 2000; 97:1512–1515. [PubMed: 10660678]

14. Burge, CB.; Tuschl, T.; Sharp, PA. Splicing of precursors to mRNAs by the spliceosome. New York: Cold Spring Harbor Laboratory Press; 1999.

15. Krainer, AR. Eukaryotic mRNA processing. New York: Oxford University Press; 1997.

16. Yang L, Chen J, Chang CC, et al. A stable upstream stem-loop structure enhances selection of the first 5'-ORF-AUG as a main start codon for translation initiation of human ACAT1 mRNA. Acta Biochim Biophys Sin. 2004; 36:259–268. [PubMed: 15253151]

17. Yang L, Lee O, Chen J, et al. Human acyl-coenzyme A:cholesterol acyltransferase 1 (acat1) sequences located in two different chromosomes (7 and 1) are required to produce a novel ACAT1 isoenzyme with additional sequence at the N terminus. J Biol Chem. 2004; 279:46253–46262. [PubMed: 15319423]

18. Cadigan KM, Heider JG, Chang TY. Isolation and characterization of Chinese hamster ovary cell mutants deficient in acyl-coenzyme A:cholesterol acyltransferase activity. J Biol Chem. 1988; 263:274. [PubMed: 3335499]

19. Higuchi R, Krummel B, Saiki RK. A general method of in vitro preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. Nucleic Acids Res. 1988; 16:7351–7367. [PubMed: 3045756]

20. Kozak M. Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. Mol Cell Biol. 1989; 9:5134–5142. [PubMed: 2601712]

21. Chang CC, Chen J, Thomas MA, et al. Regulation and immunolocalization of acyl-coenzyme A: cholesterol acyltransferase in mammalian cells as studied with specific antibodies. J Biol Chem. 1995; 270:29532–29540. [PubMed: 7493995]

22. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003; 31:3406–3415. [PubMed: 12824337]

23. Sambrook, J.; Fritsch, EF.; Maniatis, T. Molecular Cloning. 2nd Edition. New York: Cold Spring Harbor Laboratory Press; 1989.

24. Vagner S, Waysbort A, Marenda M, et al. Alternative translation initiation of the Moloney murine leukemia virus mRNA controlled by internal ribosome entry involving the p57/PTB splicing factor. J Biol Chem. 1995; 270:20376–20383. [PubMed: 7657611]

25. Yang DQ, Halaby MJ, Zhang Y. The identification of an internal ribosomal entry site in the 5'-untranslated region of p53 mRNA provides a novel mechanism for the regulation of its translation following DNA damage. Oncogene. 2006; 25:4613–4619. [PubMed: 16607284]

26. Dhar D, Roy S, Das S. Translational control of the interferon regulatory factor 2 mRNA by IRES element. Nucleic Acids Res. 2007; 35:5409–5421. [PubMed: 17698501]

27. Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev. 1998; 12:67–83. [PubMed: 9420332]

28. Gebauer F, Hentze MW. Molecular mechanisms of translational control. Nat Rev Mol Cell Biol. 2004; 5:827–835. [PubMed: 15459663]

29. Rijnbrand R, Bredenbeek PJ, Haasnoot PC, et al. The influence of downstream protein-coding sequence on internal ribosome entry on hepatitis C virus and other flavivirus RNAs. RNA. 2001; 7:585–597. [PubMed: 11345437]

30. Pestova TV, Hellen CU, Shatsky IN. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. Mol Cell Biol. 1996; 16:6859–6869. [PubMed: 8943341]

31. Pestova TV, Shatsky IN, Hellen CU. Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. Mol Cell Biol. 1996; 16:6870–6878. [PubMed: 8943342]

32. Hellen CU, Witherell GW, Schmid M, et al. A cytoplasmic 57-kDa protein that is required for translation of picornavirus RNA by internal ribosomal entry is identical to the nuclear pyrimidine tract-binding protein. Proc Natl Acad Sci U S A. 1993; 90:7642–7646. [PubMed: 8395052]

33. Walter BL, Nguyen JH, Ehrenfeld E, Semler BL. Differential utilization of poly(rC) binding protein 2 in translation directed by picornavirus IRES elements. RNA. 1999; 5:1570–1585. [PubMed: 10606268]

34. Buratti E, Baralle FE. Influence of RNA secondary structure on the pre-mRNA splicing process. Mol Cell Biol. 2004; 24:10505–10514. [PubMed: 15572659]

35. Wu C, Alwine JC. Secondary structure as a functional feature in the downstream region of mammalian polyadenylation signals. Mol Cell Biol. 2004; 24:2789–2796. [PubMed: 15024068]

36. Linnstaedt SD, Kasprzak WK, Shapiro BA, Casey JL. The role of a metastable RNA secondary structure in hepatitis delta virus genotype III RNA editing. RNA. 2006; 12:1521–1533. [PubMed: 16790843]
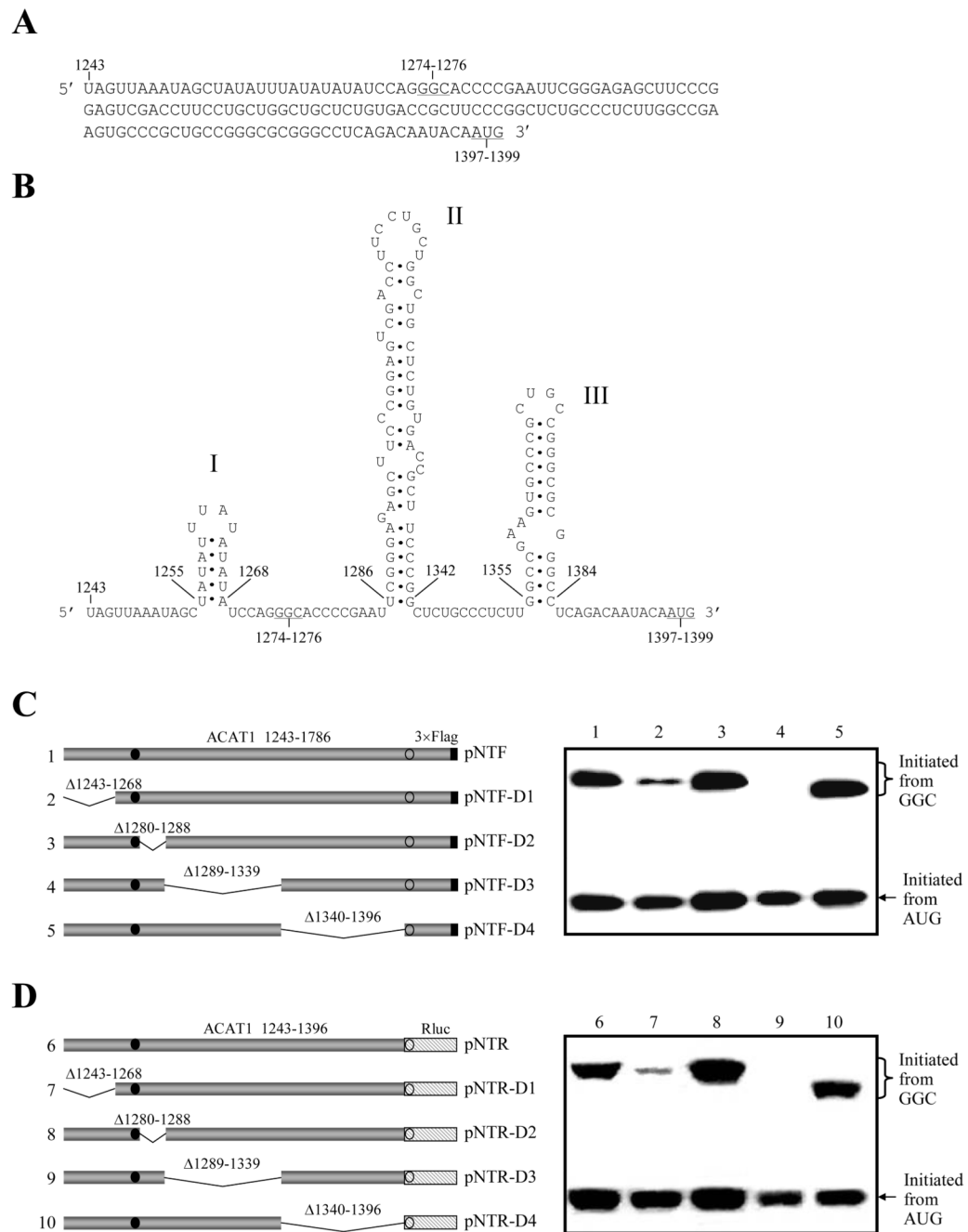
A

```
      1243                                    1274-1276
   5′ UAGUUAAAUAGCUAUAUUUAUAUAUAUCCAGGGCACCCCGAAUUCGGGAGAGCUUCCCG
      GAGUCGACCUUCCUGCUGGCUGCUCUGUGACCGCUUCCCGGCUCUGCCCUCUUGGCCGA
      AGUGCCCGCUGCCGGGCGCGGGCCUCAGACAAUACAAUG  3′
                                             1397-1399
```

B

C

D

Figure 1. Predicted stem-loops in the vicinity of GGC$_{1274-1276}$ codon are required for the production of ACAT1 isoforms

(A) Vicinal sequence of the GGC$_{1274-1276}$ codon (nt 1243–1396). The initiation codons GGC$_{1274-1276}$ for the 56-kD human ACAT1 isoform and AUG$_{1397-1399}$ for the 50-kD one are underlined.

(B) Predicted RNA secondary structures in the vicinity of GGC$_{1274-1276}$ codon (nt 1243–1396). The three successive stem-loops are respectively labeled with I, II and III.

(C) Schematic representation of the partial ACAT1 mRNA sequence (nt 1243–1786) and its truncated forms on the left panel. The deleted regions (Δ1243–1268, Δ1280–1288, Δ1289–1339 and Δ1340–1369) are marked on the top of each bar. Gray bar, ACAT1 mRNA

sequence (ACAT1 1243–1786); black bar, 3×Flag coding sequence (3×Flag); filled circle, $GGC_{1274-1276}$ initiation codon; hollow circle, $AUG_{1397-1399}$ initiation codon. The expression plasmids depicted on left are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-ACAT1 antibodies (DM10). The curly bracket and the arrow indicate the positions of ACAT1-NT-Flag proteins respectively initiated from $GGC_{1274-1276}$ and $AUG_{1397-1399}$. The experiments are repeated three times with similar results.

(D) Schematic representation of the replacement of partial ACAT1 AUG-ORF with the whole Renilla luciferase AUG-ORF on the left panel. Gray bar, vicinity of $GGC_{1274-1276}$ codon (ACAT1 1243–1396) (ACAT1 1243–1396); hatched bar, the whole AUG-ORF of Renilla luciferase (Rluc); others representing the same in (C). The expression plasmids depicted on left are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-Rluc antibody. The curly bracket and the arrow indicate the positions of the fused ACAT1-Rluc protein initiated from the $GGC_{1274-1276}$ and Rluc protein initiated from $AUG_{1397-1399}$. The experiments are repeated three times with similar results.
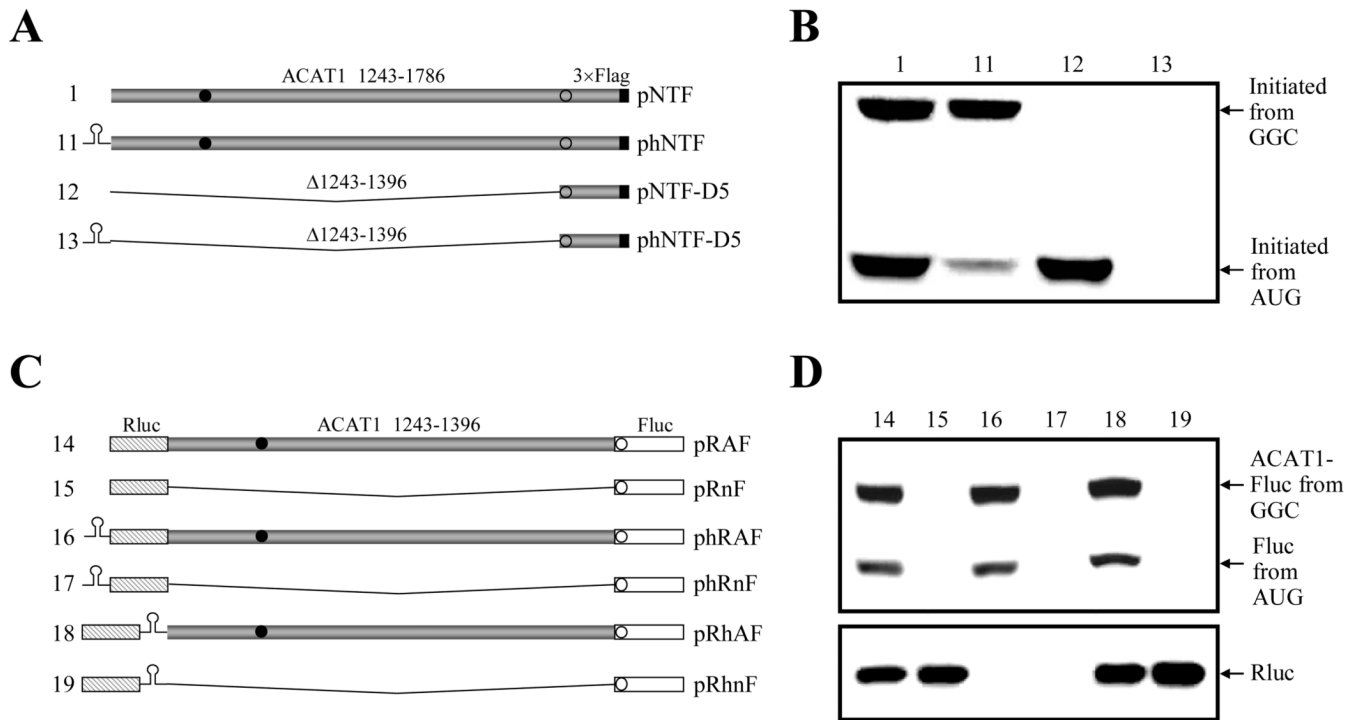
Figure 2. Translation initiation from GGC$_{1274-1276}$ codon is mediated by IRES

(A) Schematic representation of the partial ACAT1 mRNA sequences (nt 1243–1786) and its truncated form (nt 1397–1786) without or with a 5'-stable hairpin. The stable hairpin (ΔG = −57 kcal/mol) is located at 5'-end of the partial ACAT1 mRNA sequences and its truncated form in the plasmids phNTF and phNTF-D5. The vicinity of the GGC$_{1274-1276}$ codon was deleted (Δ1243-1396) in the negative control plasmids pNTF-D5 and phNTF-D5. Gray bar, ACAT1 mRNA sequence (ACAT1 1243–1786); black bar, 3×Flag coding sequence (3×Flag); filled circle, GGC$_{1274-1276}$ initiation codon; hollow circle, AUG$_{1397-1399}$ initiation codon.

(B) The expression plasmids depicted in (A) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-ACAT1 antibodies (DM10). Arrows indicate the positions of ACAT1-NT-Flag proteins respectively initiated from GGC$_{1274-1276}$ and AUG$_{1397-1399}$. The experiments are repeated three times with similar results.

(C) Schematic representation of the two cistrons without or with the stable hairpin. The first and second cistrons are the whole AUG-ORFs of Renilla luciferase and Firefly luciferase in the plasmids pRnF, phRnF and pRhnF as the negative controls. The expression plasmids pRAF, phRAF and pRhAF contain the second cistrons by fusing the vicinity of GGC$_{1274-1276}$ codon (nt 1243–1396) with the 5'-end of whole AUG-ORF of Firefly luciferase. The stable hairpin depicted in (A) is located at 5'- or 3'-end of the first cistron. Gray bar, the vicinity of GGC$_{1274-1276}$ codon (ACAT1 1243–1396); hatched bar, whole AUG-ORF of Renilla luciferase (Rluc); white bar, whole AUG-ORF of Firefly luciferase (Fluc); others representing the same in (A).

(D) The expression plasmids depicted in (C) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-Fluc antibody and anti-Rluc antibody respectively. The immunoblotting result with anti-Fluc antibody is shown on the top panel and arrows indicate the positions of the fused ACAT1-Fluc protein initiated from the GGC$_{1274-1276}$ and Fluc protein initiated from AUG$_{1397-1399}$. The immunoblotting result

with anti-Rluc antibody is shown on the bottom panel and an arrow indicates the position of Rluc protein. The experiments are repeated three times with similar results.
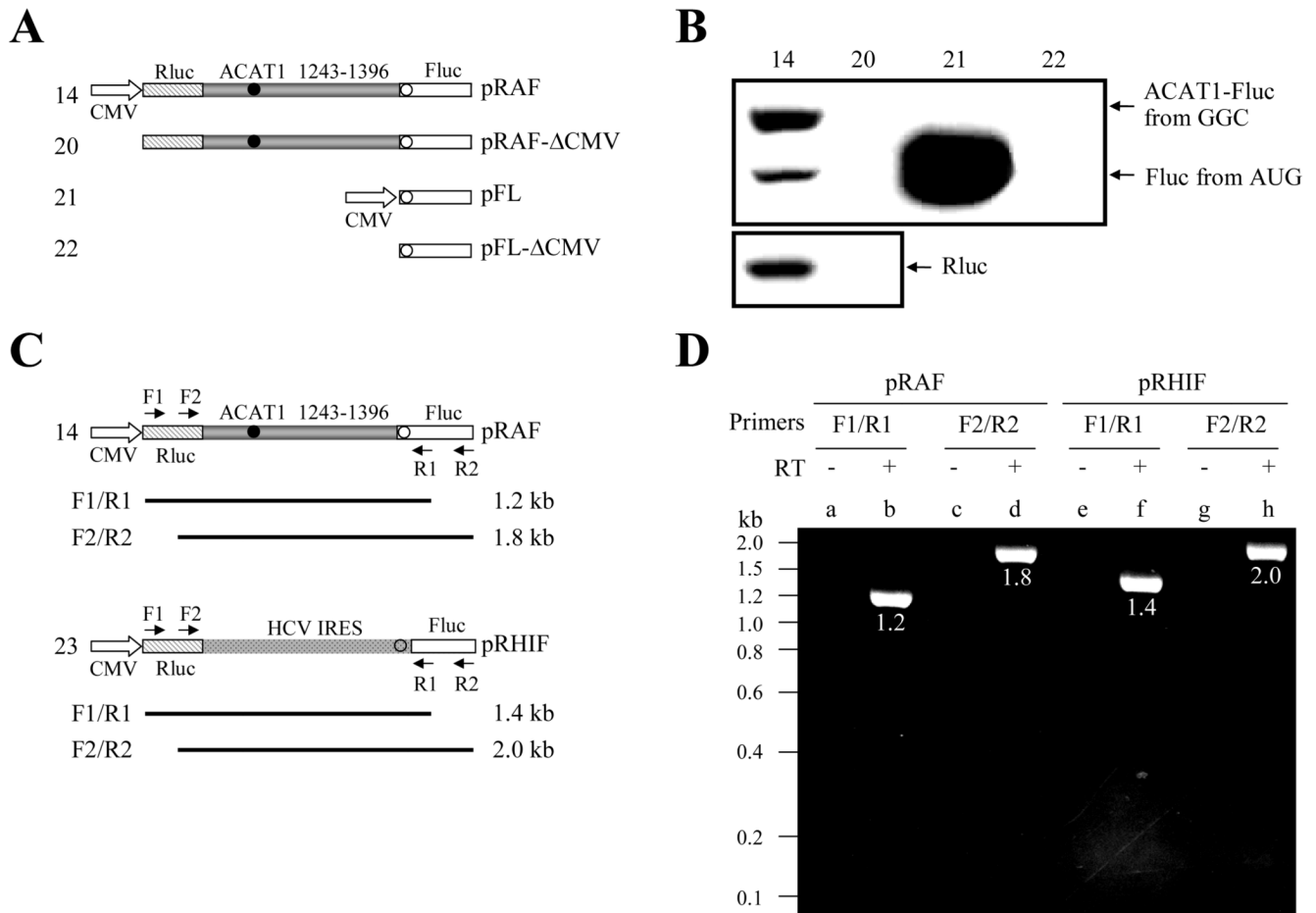
A

B

C

D

Figure 3. The vicinity of the GGC$_{1274-1276}$ codon in the bicistronic plasmid does not show the cryptic promoter activity or contain the aberrant splicing site

(A) Schematic representation of the bicistrons and monocistron with or without the eukaryotic CMV promoter. The plasmids pRAF and pRAF-ΔCMV contain the first cistron of the whole AUG-ORF of Renilla luciferase and the second one by fusing the vicinity of GGC$_{1274-1276}$ codon (nt 1243–1396) with the 5'-end of whole AUG-ORF of Firefly luciferase. The plasmids pFL and pFL-ΔCMV only contain a cistron of the whole AUG-ORF of Firefly luciferase. The eukaryotic CMV promoter is respectively deleted from the 5'-ends of the whole AUG-ORF of Renilla luciferase and Firefly luciferase in the plasmids pRAF-ΔCMV and pFL-ΔCMV. Gray bar, the vicinity of GGC$_{1274-1276}$ codon (ACAT1 1243–1396); hatched bar, whole AUG-ORF of Renilla luciferase (Rluc); white bar, whole AUG-ORF of Firefly luciferase (Fluc); filled circle, GGC$_{1274-1276}$ initiation codon; hollow circle, AUG$_{1397-1399}$ initiation codon.

(B) The expression plasmids depicted in (A) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-Fluc antibody and anti-Rluc antibody respectively. The immunoblotting result with anti-Fluc antibody is shown on the top panel and arrows indicate the positions of the fused ACAT1-Fluc protein initiated from the GGC$_{1274-1276}$ and Fluc protein initiated from AUG$_{1397-1399}$. The immunoblotting result with anti-Rluc antibody is shown on the bottom panel and an arrow indicates the position of Rluc protein. The experiments are repeated three times with similar results.

(C) Schematic representation of the location of four PCR primers at the bicistronic plasmids pRAF and pRHIF. The primer F1 and R1 are respectively located at the 5'-ends of the whole

AUG-ORF of Renilla luciferase and Firefly luciferase, and the primer F2 and R2 are respectively located at the 3'-ends of the whole AUG-ORF of Renilla luciferase and Firefly luciferase. The calculated sizes of PCR products by using primer sets F1/R1 and F2/R2 are shown. Spots bar, the region of HCV IRES; hollow circle, AUG initiation codon; others representing the same in (A).

(D) RT-PCR analysis of total RNAs from the cells transfected with the plasmids pRAF and pRHIF is performed by using two sets of primers F1/R1 and F2/R2. Lanes b (1.2 kb), d (1.8 kb), f (1.4 kb) and h (2.0 kb) show the amplified products with the size depicted on left from the pRAF- or pRHIF-transfected cells when F1/R1 and F2/R2 primer sets are used. Lanes a, c, e and g show reverse transcription-negative controls for each sample. The experiments are repeated three times with similar results.
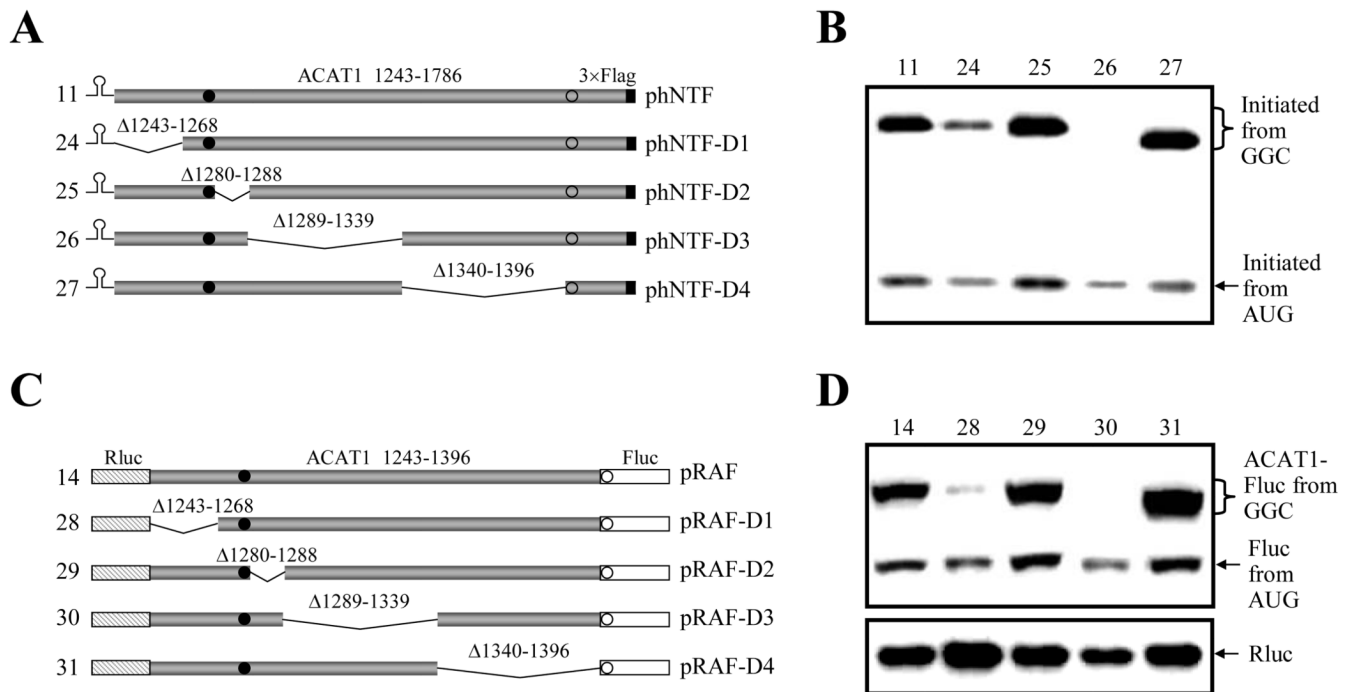
Figure 4. The translation initiation from the $GGC_{1274-1276}$ codon mediated by IRES requires both stem-loop I and II

(A) Schematic representation of the partial ACAT1 mRNA sequence (nt 1243–1786) and its truncated forms with the 5'-stable hairpin. The deleted regions (Δ1243–1268, Δ1280–1288, Δ1289–1339 and Δ1340–1369) are marked on top and the 5'-stable hairpin ($\Delta G = -57$ kcal/mol) is located at the 5'-end. Gray bar, ACAT1 mRNA sequence (ACAT1 1243–1786); black bar, 3×Flag coding sequence (3×Flag); filled circle, $GGC_{1274-1276}$ initiation codon; hollow circle, $AUG_{1397-1399}$ initiation codon.

(B) The expression plasmids depicted in (A) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-ACAT1 antibodies (DM10). The curly bracket and the arrow indicate the positions of ACAT1-NT-Flag proteins respectively initiated from $GGC_{1274-1276}$ and $AUG_{1397-1399}$. The experiments are repeated twice with similar results.

(C) Schematic representation of two cistrons containing the first cistron of the whole AUG-ORF of Renilla luciferase and the second one by fusing the vicinity of $GGC_{1274-1276}$ codon (nt 1243–1396) or its truncated forms with 5'-end of the whole AUG-ORF of Firefly luciferase. The deleted regions (Δ1243–1268, Δ1280–1288, Δ1289–1339 and Δ1340–1369) are marked on top. Gray bar, the vicinity of $GGC_{1274-1276}$ codon (ACAT1 1243–1396); hatched bar, whole AUG-ORF of Renilla luciferase (Rluc); white bar, whole AUG-ORF of Firefly luciferase (Fluc); others representing the same in (A).

(D) The expression plasmids depicted in (C) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-Fluc antibody and anti-Rluc antibody respectively. The immunoblotting result with anti-Fluc antibody is shown on the top panel, and the curly bracket and the arrow indicate the positions of the fused ACAT1-Fluc protein initiated from the $GGC_{1274-1276}$ and Fluc protein initiated from $AUG_{1397-1399}$. The immunoblotting result with anti-Rluc antibody is shown on the bottom panel and an arrow indicates the position of Rluc protein. The experiments are repeated twice with similar results.
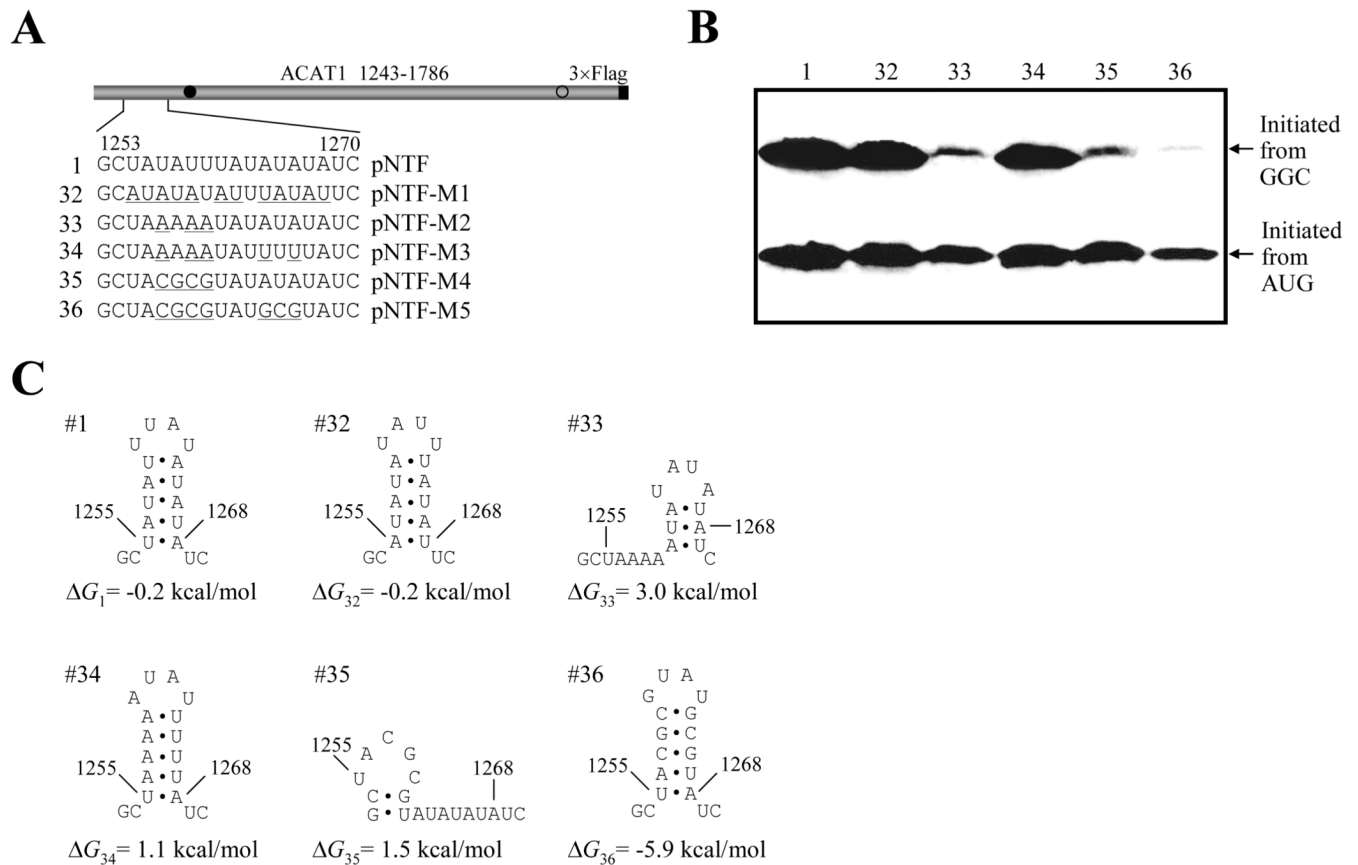
A


```
                    ACAT1 1243-1786              3×Flag

      1253                1270
  1   GCUAUAUUUAUAUAUAUC  pNTF
 32   GCAUAUAUAUUUAUAUUC  pNTF-M1
 33   GCUAAAAAUAUAUAUAUC  pNTF-M2
 34   GCUAAAAAUAUUUUUAUC  pNTF-M3
 35   GCUACGCGUAUAUAUAUC  pNTF-M4
 36   GCUACGCGUAUGCGUAUC  pNTF-M5
```

B

C

$\Delta G_1 = -0.2$ kcal/mol $\Delta G_{32} = -0.2$ kcal/mol $\Delta G_{33} = 3.0$ kcal/mol

$\Delta G_{34} = 1.1$ kcal/mol $\Delta G_{35} = 1.5$ kcal/mol $\Delta G_{36} = -5.9$ kcal/mol

Figure 5. AU-constitution of the upstream stem-loop I is important for producing proteins initiated from the GGC$_{1274-1276}$ codon

(A) Schematic representation of the partial ACAT1 mRNA sequence (nt 1243–1786) and mutations of stem-loop I. The detailed sequence of nt 1253–1270 is listed without or with the mutant nucleotides (underlined). Gray bar, partial ACAT1 mRNA sequence (ACAT1 1243–1786); black bar, 3×Flag coding sequence (3×Flag); filled circle, GGC$_{1274-1276}$ initiation codon; hollow circle, AUG$_{1397-1399}$ initiation codon.

(B) The expression plasmids depicted in (A) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-ACAT1 antibodies (DM10). Arrows indicate the positions of ACAT1-NT-Flag proteins respectively initiated from GGC$_{1274-1276}$ and AUG$_{1397-1399}$. The experiments are repeated twice with similar results.

(C) Predicted RNA secondary structures of nt 1253–1270 without or with mutations of stem-loop I depicted in (A). The folding Gibbs free energy (ΔG) of predicted RNA secondary structure is listed beneath each one. These experiments are repeated three times with similar results.
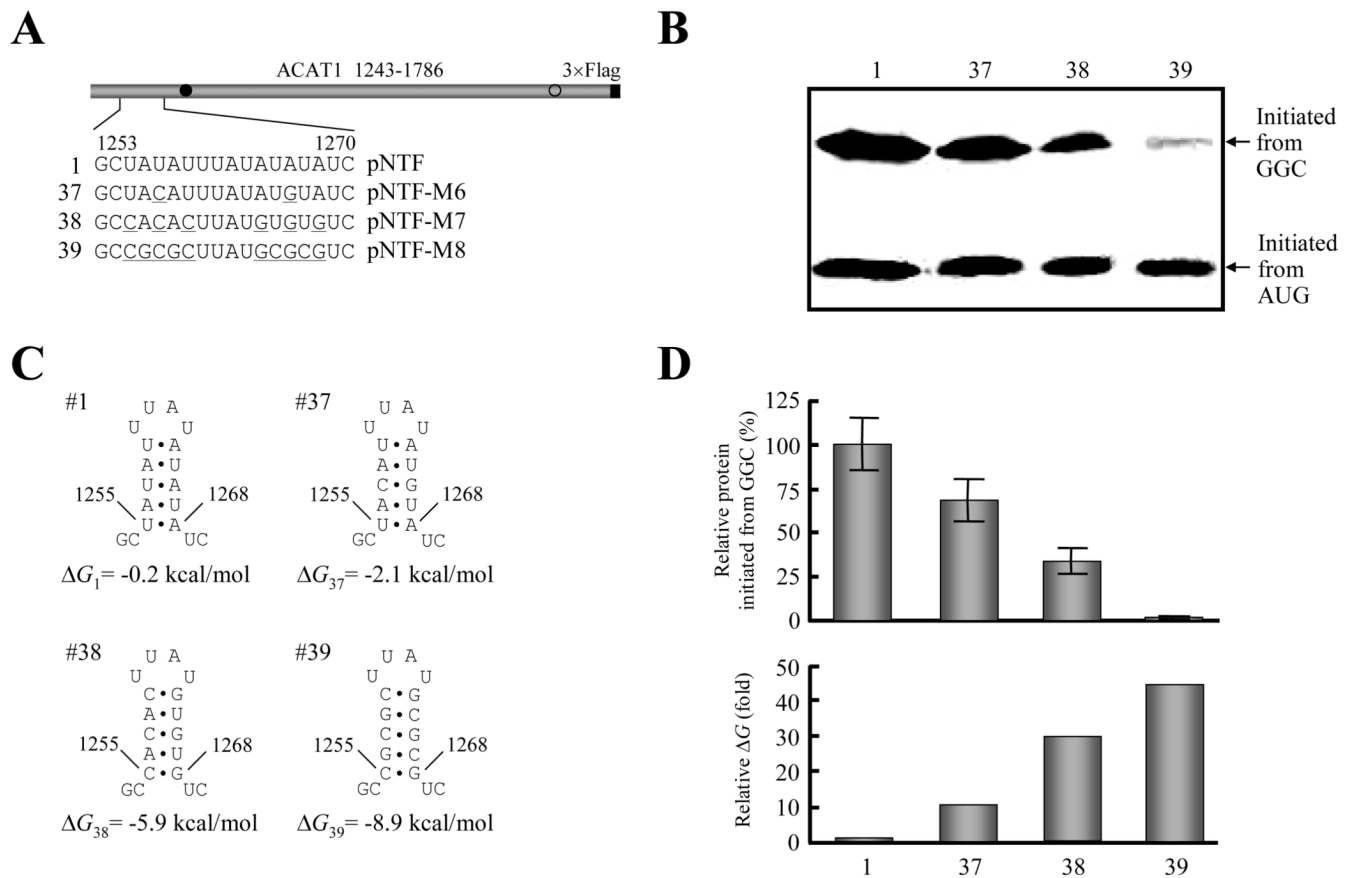
A

ACAT1 1243-1786   3×Flag

```
         1253              1270
 1  GCUAUAUUUAUAUAUAUC  pNTF
37  GCUACAUUUAUAUGUAUC  pNTF-M6
38  GCCACACUUAUGUGUGUC  pNTF-M7
39  GCCGCGCUUAUGCGCGUC  pNTF-M8
```

B

C

```
#1      U  A
      U      U
      U • A
      A • U
1255  U • A  1268
      A • U
      U • A
    GC      UC
```
$\Delta G_1 = -0.2$ kcal/mol

```
#37     U  A
      U      U
      U • A
      A • U
1255  C • G  1268
      A • U
      U • A
    GC      UC
```
$\Delta G_{37} = -2.1$ kcal/mol

```
#38     U  A
      U      U
      C • G
      A • U
1255  C • G  1268
      A • U
      C • G
    GC      UC
```
$\Delta G_{38} = -5.9$ kcal/mol

```
#39     U  A
      U      U
      C • G
      G • C
1255  C • G  1268
      G • C
      C • G
    GC      UC
```
$\Delta G_{39} = -8.9$ kcal/mol

D

Figure 6. Stability of the upstream RNA secondary structure is inversely correlated with production of protein initiated from the GGC$_{1274-1276}$ codon

(A) Schematic representation of the partial ACAT1 mRNA sequence (nt 1243–1786) and the replacement of AU nucleotides by GC nucleotides in stem-loop I. The detailed sequence of nt 1253–1270 is listed and the replaced nucleotides are underlined. Gray bar, partial ACAT1 mRNA sequence (ACAT1 1243–1786); black bar, 3×Flag coding sequence (3×Flag); filled circle, GGC$_{1274-1276}$ initiation codon; hollow circle, AUG$_{1397-1399}$ initiation codon.

(B) The expression plasmids depicted in (A) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-ACAT1 antibodies (DM10). Arrows indicate the positions of ACAT1-NT-Flag proteins respectively initiated from GGC$_{1274-1276}$ and AUG$_{1397-1399}$. The experiments are repeated three times with similar results.

(C) Predicted RNA secondary structures of nt 1253–1270 without or with mutations of stem-loop I depicted in (A). The folding Gibbs free energy (ΔG) of predicted RNA secondary structure is listed beneath each one.

(D) Relative production of ACAT1-NT-Flag protein initiated from the GGC$_{1274-1276}$ codon (up panel) and folding ΔG of RNA secondary structures (down panel). The intensity of products from the GGC$_{1274-1276}$ codon in (B) is quantified by using the UVP Labwork software (UVP Inc.) for densitometric analysis and normalized to the value of wild type one. The data mean ± SD from three independent experiments. The relative folding ΔG of RNA secondary structures in (C) is shown by using the wild type one as 1.0.
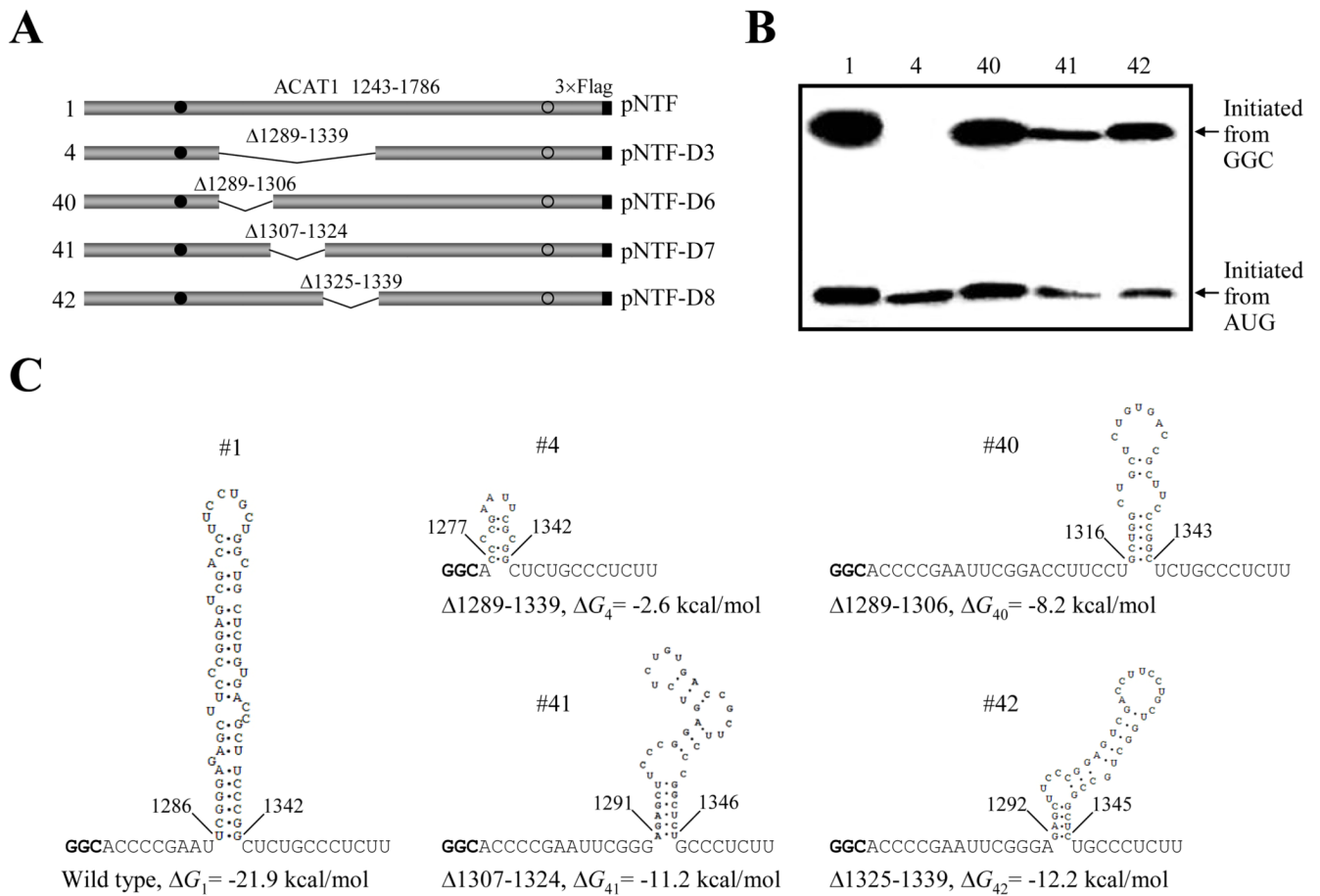
Figure 7. The downstream RNA secondary structure is required for producing proteins initiated from the GGC$_{1274-1276}$ codon

(A) Schematic representation of the partial ACAT1 mRNA sequence (nt 1243–1786) and the whole- or sub-deletions of stem-loop II. The deleted regions (Δ1289–1339, Δ1289–1306, Δ1307–1324 and Δ1325–1339) are marked on top. Gray bar, partial ACAT1 mRNA sequence (ACAT1 1243–1786); black bar, 3×Flag coding sequence (3×Flag); filled circle, GGC$_{1274-1276}$ initiation codon; hollow circle, AUG$_{1397-1399}$ initiation codon.

(B) The expression plasmids depicted in (A) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-ACAT1 antibodies (DM10). Arrows indicate the positions of ACAT1-NT-Flag proteins respectively initiated from GGC$_{1274-1276}$ and AUG$_{1397-1399}$. The experiments are repeated twice with similar results.

(C) Predicted RNA secondary structures of nt 1274–1354 without or with the whole- or sub-deletions depicted in (A). The folding Gibbs free energies (ΔG) of each predicted RNA secondary structure are listed beneath each one.
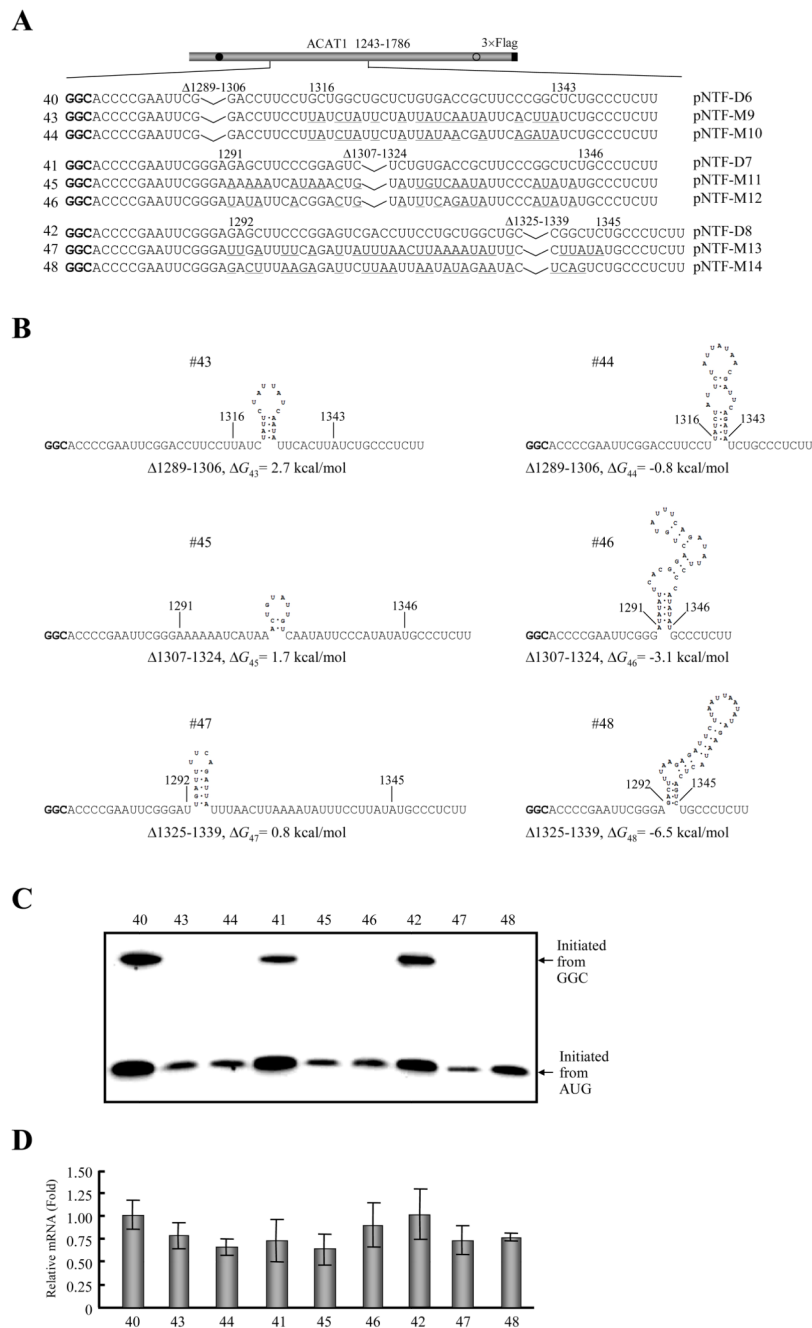
Figure 8. GC-richness of the downstream RNA secondary structures is essential for production of proteins initiated from the GGC$_{1274-1276}$ codon

(A) Schematic representation of the partial ACAT1 mRNA sequence (nt 1243–1786) and mutations introduced into the formed RNA secondary structures after sub-deletions of stem-loop II. Each sequence after sub-deletion of stem-loop II is listed without or with the mutated nucleotides (underlined) and the deleted regions (Δ1289–1339, Δ1289–1306, Δ1307–1324 and Δ1325–1339) are marked on top. Gray bar, partial ACAT1 mRNA sequence (ACAT1 1243–1786); black bar, 3×Flag coding sequence (3×Flag); filled circle, GGC$_{1274-1276}$ initiation codon; hollow circle, AUG$_{1397-1399}$ initiation codon.

(B) Predicted RNA secondary structures with introduced mutations depicted in (A). The folding Gibbs free energy (ΔG) of predicted RNA secondary structure is listed beneath each one.

(C) The expression plasmids depicted in (A) are transiently transfected into AC29, the lysates are prepared and immunoblotting is carried out with anti-ACAT1 antibodies (DM10). Arrows indicate the positions of ACAT1-NT-Flag proteins respectively initiated from $GGC_{1274-1276}$ and $AUG_{1397-1399}$. The experiments are repeated three times with similar results.

(D) RT-qPCR analysis of total RNA from the cells transfected with the plasmids depicted in (A) is performed according to procedures described in the Materials and Methods. The ACAT1 mRNA levels are normalized to the GAPDH mRNA levels for each sample and the relative mRNA of the cells transfected with pNTF-D6 is designated as the control one (1.0). The data mean ± SD from three independent experiments.

Table 1

Primers for deletions at the vicinity of the $GGC_{1274-1276}$ codon

Plasmid (deleted region)	Primer name	Primer sequence
pNTF-D1	D1F	5'-AAAGGTACCTCCAGGGCACCCCGAAT-3'
(Δ1243–1268)		
pNTF-D2	D2F	5'-ATCCAGGGCACCGGAGAGCTTCCCGGAG-3'
(Δ1280–1288)	D2R	5'-GGGAAGCTCTCCGGTGCCCTGGAT-3'
pNTF-D3	D3F	5'-ACCCCGAATTCGCGGCTCTGCCCTCTTG-3'
(Δ1289–1339)	D3R	5'-AGGGCAGAGCCGCGAATTCGGGGTGCCCTG-3'
pNTF-D4	D4F	5'-GTGACCGCTTCCATGGTGGGTGAAGAG-3'
(Δ1340–1396)	D4R	5'-TTCACCCACCATGGAAGCGGTCACAGAG-3'
pNTF-D5	D5F	5'-AAAGGTACCATGGTGGGTGAAGAGAAG-3'
(Δ1243–1396)		
pNTF-D6	D6F	5'-ACCCCGAATTCGGACCTTCCTGCTG-3'
(Δ1289–1306)	D6R	5'-AGCAGGAAGGTCCGAATTCGGGGTG-3'
pNTF-D7	D7F	5'-CTTCCCGGAGTCTCTGTGACCGCTTC-3'
(Δ1307–1324)	D7R	5'-AGCGGTCACAGAGACTCCGGGAAG-3'
pNTF-D8	D8F	5'-CCTGCTGGCTGCCGGCTCTGCCCTCTT-3'
(Δ1325–1339)	D8R	5'-AGGGCAGAGCCGGCAGCCAGCAGGAAG-3'

Table 2

Primers for mutations introduced into the upstream stem-loop I

Plasmid	Primer name	Primer sequence
pNTF-M1	M1F	5'-AAAGGTACCTAGTTAAATAGC<u>ATATAT</u>A<u>TT TATAT</u> TCCAGGGCACCCCGAATTC-3'
pNTF-M2	M2F	5'-AAAGGTACCTAGTTAAATAGCTA<u>AAAA</u>TATATATATCCAGGGCAC-3'
pNTF-M3	M3F	5'-AAAGGTACCTAGTTAAATAGCTA<u>AAAA</u>TAT<u>TT</u>TTATCCAGGGCACCCCGAATTC-3'
pNTF-M4	M4F	5'-AAAGAATTCTAGTTAAATAGCTA<u>CGCG</u>TATATATATCCAGGGCACCCCGAA-3'
pNTF-M5	M5F	5'-AAAGAATTCTAGTTAAATAGCTA<u>CGCG</u>TAU<u>GCG</u>TATCCAGGGCACCCCGAATTC-3'
pNTF-M6	M6F	5'-AAAGAATTCTAGTTAAATAGCTA<u>C</u>ATTTATAT<u>G</u>TATCCAGGGCACCCCGAATTC-3'
pNTF-M7	M7F	5'-AAAGAATTCTAGTTAAATAGC<u>CACAC</u>TTAT<u>GTGTG</u>TCCAGGGCACCCCGAATTC-3'
pNTF-M8	M8F	5'-AAAGAATTCTAGTTAAATAGC<u>CGCGC</u>TTAT<u>GCGCG</u>TCCAGGGCACCCCGAATTC-3'

Table 3

Primers for mutations introduced into the downstream RNA secondary structures

Plasmid (deleted region)	Primer name	Primer sequence
pNTF-M9 (Δ1289–1306)	M9F	5'-TATCTATTCTAT TATCAATA TTCACTTATCTGCCCTC TTGGCCGAAGT-3'
	M9R	5'-TAAGTGAATATTGATAATAGAATAGATAAGGAAGG TCCGAATTCGGGGT-3'
pNTF-M10 (Δ1289–1306)	M10F	5'-TATCTATTCTATTATAACGATTCAGATATCTGCCCT CTTGGCCGAAGT-3'
	M10R	5'-TATCTGAATCGTTATAATAGAATAGATAAGGAAGG TCCGAATTCGGGGT-3'
pNTF-M11 (Δ1307–1324)	M11F	5'-AAAAATCATAAACTGTATTGTCAATATTCCCATATA TGCCCTCTTGGCCGAAGTGCC-3'
	M11R	5'-TATATGGGAATATTGACAATACAGTTTATGATTTTT TCCCGAATTCGGGGTGCCCTG-3'
pNTF-M12 (Δ1307–1324)	M12F	5'-TATATTCACGGACTGTATTTCAGATATTCCCATATA TGCCCTCTTGGCCGAAGTGCC-3'
	M12R	5'-TATATGGGAATATCTGAAATACAGTCCGTGAATATA TCCCGAATTCGGGGTGCCCTG-3'
pNTF-M13 (Δ1325–1339)	M13F	5'-TTGATTTTCAGATTATTTAACTTAAAATATTTCCTTA TATGCCCTCTTGGCCGAAGTGC-3'
	M13R	5'-TATAAGGAAATATTTTAAGTTAAATAATCTGAAAAT CAATCCCGAATTCGGGGTGCCCT-3'
pNTF-M14 (Δ1325–1339)	M14F	5'-GACTTTAAGAGATTCTTAATTAATATAGAATACTCA GTCTGCCCTCTTGGCCGAAGTGC-3'
	M14R	5'-CTGAGTATTCTATATTAATTAAGAATCTCTTAAAGT CTCCCGAATTCGGGGTGCCCT-3'

The mutated nucleotides were underlined.