6222–6231 *Nucleic Acids Research, 2013, Vol. 41, No. 12* *Published online 24 April 2013*
doi:10.1093/nar/gkt286

Stability of RNA quadruplex in open reading frame determines proteolysis of human estrogen receptor α

Tamaki Endoh[1], Yu Kawasaki[2] and Naoki Sugimoto[1,2,*]

[1]Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojimaminamimachi, Kobe 650-0047, Japan and [2]Faculty of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojimaminamimachi, Kobe 650-0047, Japan

Received January 11, 2013; Revised March 26, 2013; Accepted April 29, 2013

ABSTRACT

mRNAs encodes not only information that determines amino acid sequences but also additional layers of information that regulate the translational processes. Notably, translational halt at specific position caused by rare codons or stable RNA structures is one of the potential factors regulating the protein expressions and structures. In this study, a quadruplex-forming potential (QFP) sequence derived from an open reading frame of human estrogen receptor α (*hERα*) mRNA was revealed to form parallel G-quadruplex and halt the translation elongation *in vitro*. Moreover, when the full-length *hERα* and variants containing synonymous mutations in the QFP sequence were expressed in cells, translation products cleaved at specific site were observed in quantities dependent on the thermodynamic stability of the G-quadruplexes. These results suggest that the G-quadruplex formation in the coding region of the *hERα* mRNA impacts folding and proteolysis of hERα protein by slowing down or temporarily stalling the translation elongation.

INTRODUCTION

During translation elongation, the ribosome decodes mRNA sequence and converts the information into the amino acid sequence of protein. Ribosome progression is arrhythmic rather than uniform because of features of mRNAs. Rare codons and specific secondary structures in open reading frames (ORFs) are known to slow or stall translation elongation (1,2). Previous studies have revealed that stalling of translation elongation at specific position regulates protein expression by affecting translation initiation of a second ORF or by inducing ribosomal frameshifts (2–5). In addition, the positions of rare codons and stable secondary structures of mRNA correlate with flexible linker regions in protein structures (6–9). As proteins partially fold during translation elongation (10–12), the rare codons and secondary structures of mRNA likely affect the structures of nascent proteins (8,9,13–18). These observations suggest that mRNA sequence and structure might impact not only expression levels but also the processing of proteins.

G-quadruplex is one of the non-canonical structures formed by guanine-rich (G-rich) sequence in DNA and RNA stabilized by stacking of planes consisting of four guanine bases that interact through Hoogsteen-type hydrogen bonds (19,20). The G-quadruplexes formed by RNAs are generally more stable than those formed by equivalent DNA sequences (21–23). Our previous studies indicate that stabilities of G-quadruplexes are higher in the presence of high concentration of cosolute, which might mimic cell-like molecular crowding condition, and encapsulated conditions than in diluted solution (24,25), whereas canonical secondary structures are destabilized in the presence of cosolute, especially solutes decreasing water activity (26,27). In addition, we recently performed that RNA G-quadruplex formed in ORF of mRNA suppressed translation elongation *in vitro* and reduced protein expression in cells (28). Thus, the G-rich sequence conserved in mRNAs likely forms RNA G-quadruplex in cells and contributes to biological procedures exhibiting an effect as strong as the canonical secondary structures based on Watson–Crick-type hydrogen bonds.

Human estrogen receptor α (hERα) is a ligand-inducible nuclear transcription factor that activates expression of genes responsible for growth, development and maintenance of the reproductive, skeletal, neuronal and immune systems in response to estrogens (29,30). The *hERα* mRNA contains a sequence with quadruplex-forming potential (QFP) in the region of the ORF that codes for domain D, a hinge region of the protein (Figure 1A) (29,31). Preceding DNA-binding domain (domain C) and after ligand-binding domain (domain E/F) form

*To whom correspondence should be addressed. Tel: +81 78 303 1416; Fax: +81 78 303 1495; Email: sugimoto@konan-u.ac.jp

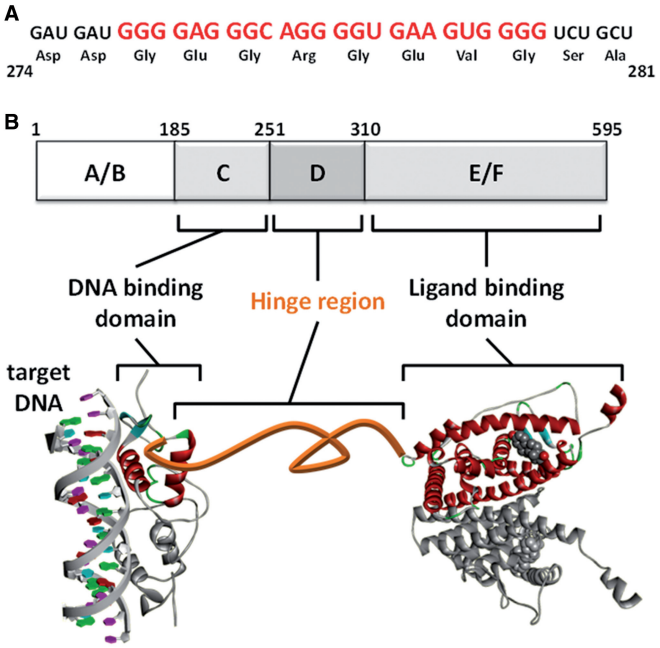

Figure 1. QFP sequence derived from *hERα* mRNA. (**A**) QFP sequence in the ORF of *hERα* mRNA. Amino acids coded on the QFP sequence are indicated. (**B**) Domains of hERα protein. The hinge region (domain D), which contains the QFP sequence, connects the DNA-binding domain (domain C) and ligand-binding domain (domain E/F). Structures of the DNA-binding domain (PDB: 1HCQ) and ligand-binding domain (PDB: 1A52) are shown.

respective domain structures and contribute the ligand-dependent gene expressions (Figure 1B). If the QFP sequence adopts a G-quadruplex structure and slows or stalls the translation elongation, it might affect the co-translational folding of the domain structures (Figure 1B). In this study, we investigated effects of the QFP sequence derived from the *hERα* mRNA on protein expression.

MATERIALS AND METHODS

Preparation of RNA

RNA oligonucleotides used in this study were purchased from Japan Bio Services Co., Ltd and were purified by high-performance liquid chromatography (HPLC). Concentrations of the RNA oligonucleotides were determined by measurement of the absorbance at 260 nm at a high temperature using a Shimadzu-1800 ultraviolet (UV)/Vis spectrophotometer connected to a thermoprogrammer. Extinction coefficients were calculated from mononucleotide and dinucleotide data using nearest-neighbor approximations.

Reporter mRNAs were prepared by *in vitro* transcription. A DNA template for transcription of a reporter mRNA without the QFP sequence was amplified by polymerase chain reaction (PCR) from a plasmid vector, previously constructed for synchronized translation (32) using a T7 promoter primer and an antisense primer. This DNA has the sequence 5′-TAAATGAATTTTTTT ATTAATAAATAAGATTTCATAGAAAGCATTTTG T-3′. DNA templates for reporter mRNAs containing QFP sequence variants were amplified from the first PCR product by using T7 promoter primer and the respective antisense primer. Antisense primers containing wild-type, A-mutant, C-mutant, G-mutant, and U-mutant QFP sequences are 5′-TTTTTTAGACCCCACT TCACCCCTGCCCTCCCCATCTTTACCTAAATGAA TTTTTTTATTAATAAATAA-3′, 5′-TTTTTTAGATC CTACTTCACCTCTGCCTTCTCCATCTTTACCTAA ATGAATTTTTTTATTAATAAATAA-3′, 5′-TTTTTT AGACCCGACTTCACCCCTGCCCTCCCCATCTTTA CCTAAATGAATTTTTTTATTAATAAATAA-3′, 5′-T TTTTTTAGACCCCACTTCCCCCCTCCCCTCCCCAT CTTTACCTAAATGAATTTTTTTATTAATAAATA A-3′, and 5′-TTTTTTAGAACCAACTTCACCCCTG CCCTCACCATCTTTACCTAAATGAATTTTTTTATT AATAAATAA-3′, respectively. mRNAs were prepared by using ScriptMAX Thermo T7 Transcription Kit (ToYoBo) and purified on denaturing polyacrylamide gel. All the mRNAs were refolded by cooling from 90°C to 4°C at 1°C min^{-1} in a buffer containing 30 mM HEPES (pH 6.8) and 100 mM KCl.

UV measurements

UV absorbance was measured with a Shimadzu-1800 UV/Vis spectrophotometer equipped with a temperature controller. Melting curves of 5 μM RNA oligonucleotides were obtained by measuring the UV absorbance at 260 and 295 nm in a buffer containing 50 mM MES–LiOH (pH 7.0) and 3 mM KCl in the presence or absence of 40 wt% PEG200. The heating rate was 0.2°C min^{-1}. The thermodynamic stabilities at 37°C (ΔG°_{37}) were calculated from the fit of the melting curves to a theoretical equation for an intramolecular association as described previously (25,33). Before the measurement, the samples were heated to 95°C, cooled at a rate of 0.2°C min^{-1} and incubated at 4°C for 10 min.

Non-denaturing polyacrylamide gel electrophoresis

RNA oligonucleotide (5 μM) was heated to 95°C and cooled to 4°C at a rate of 0.2°C min^{-1} in a buffer containing 50 mM MES–LiOH (pH 7.0), 3 mM KCl and 40 wt% PEG200. Aliquots (10 pmol) were separated on a 15% polyacrylamide gel containing K3-TB buffer (89 mM Tris, 89 mM borate and 3 mM KCl) at 37°C. K3-TB was used as a running buffer. After gel electrophoresis, RNA oligonucleotides were stained by ethidium bromide, and gels were imaged using FLA5100 fluorescence image scanner (Fuji Film) with 532 nm excitation and 575 nm emission.

CD measurements

After UV measurements in a buffer containing 50 mM MES–LiOH (pH 7.0) and 3 mM KCl in the presence or absence of 40 wt% PEG200, RNA oligonucleotides were cooled to 20°C at a rate of 1°C min^{-1}. Circular dichroism (CD) spectra were collected on a JASCO J-820 spectropolarimeter at 37°C in 1.0-mm path length cuvettes. The CD spectra shown are averages of three scans. The temperature of the cell

holder was regulated by a JASCO PTC-348 temperature controller.

Synchronized translation

Synchronized translation was performed according to a previously published experimental procedure (32) with slight modification. An *in vitro* translation mixture (PUREfrex; GeneFrontier) without release factor (RF)1, RF2 and tyrosyl-tRNA synthetase (TyrRS) was pre-incubated for 10 min. An amber suppressor tRNA (CR110-X-AF-tRNAamber; Protein Express) was added at final concentration of 1.6 μM. The first step was started by addition of reporter mRNAs at a final concentration of 3.6 μM. Simultaneously with the first reaction, reaction mixture of PUREfrex containing all requisite proteins except ribosomes was pre-incubated. After 10 min, the second step was started by addition of an equal volume of the pre-incubated reaction mixture to the reaction mixture of the first reaction. Aliquots of the reaction mixture were quickly frozen in liquid nitrogen at indicated times. Both of steps were performed at 37°C using tRNA concentrations specified by the manufacturer; the absorbance at 260 nm of the tRNA mix (A260/ml tRNAmix) was 20. Reaction mixtures were separated on a 16% Tris–tricine sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (34). Before electrophoresis, an equal volume of 4 M urea was added to the reaction mixture at 4°C followed by an addition of RNase A and incubation at 37°C for 10 min. Fluorescence signals of CR110-X-AF on the gel were imaged by FLA5100 with 473 nm excitation and 510 nm emission.

Mass spectrometry analysis

Reaction mixtures treated with RNase A were purified using anti-T7-tag agarose (MBL) as described previously (32). The mass spectrum of the purified products was analyzed by using a MALDI-TOF mass spectrometer (Autoflex III; Bruker Daltonics).

Construction of plasmids

Total RNA was purified from MCF7 cells using an RNeasy kit (Qiagen). A cDNA library was constructed from the RNA by reverse transcription using oligo-dT primer. The coding region of wild-type *hERα* was amplified by PCR and cloned into BamHI and XhoI sites of pcDNA5/FRT (Invitrogen, USA) to construct phERα. Coding sequences of *hERα* containing the QFP sequence variants were constructed by two-step PCR and cloned into same sites of restriction enzymes to construct phERα variants. All primers are listed in Supplementary Table S1. DNA fragments of *hERα* containing QFP sequence variants were digested by BamHI and XhoI and ligated into BglII and SalI sites of pAcGFP-C1 and pAcGFP-N3 (Clontech) to make pAcGFP–hERα and phERα–AcGFP, respectively.

Cell culture

Media for culturing MCF7, HepG2 and Flp-In 293 cells were Eagle's minimum essential medium containing non-essential amino acids (EMEM), EMEM and Dulbecco's modified Eagle's medium, respectively. Media was supplemented with 10% fetal bovine serum and antibiotics (penicillin 100 U/ml and streptomycin 100 μg/ml). Cells were maintained at 37°C under 5% CO_2.

Western blotting

HepG2 cells were plated on collagen-coated six-well plates at a density of 5×10^5 cells per well and incubated overnight. One microgram of phERα containing the QFP sequence variants was transfected into HepG2 cells using 6 μl of X-treamGENE 9 (Roche) according to the manufacturer's protocol. After incubation at 37°C for 48 h, cells were harvested and lysed by sonication (total of 1.5 min) in a lysis buffer containing 0.2% Nonidet P-40 (NP-40) and 0.05% SDS, and cellular debris were removed by centrifugation. Aliquots of cell lysates were separated on a 10% SDS–PAGE, and hERα was detected by western blotting using an anti-hERα mouse monoclonal antibody (Clone 1D5; Thermo Scientific) and anti-mouse IgG antibody conjugated with alkaline phosphatase (Promega).

Fluorescence imaging of AcGFP

MCF7, HepG2 and Flp-In 293 cells were plated on collagen-coated 24-well plate at densities of 5×10^4, 1×10^5 and 1×10^5 cells per well, respectively, and incubated overnight. Cells were transfected with 250 ng of pAcGFP–hERα or phERα–AcGFP containing the QFP sequence variants using 1.5 μl of X-treamGENE 9 according to the manufacturer's protocol. After incubation at 37°C for 48 h, cells were lysed by sonication (total of 1.5 min) in a lysis buffer containing 0.2% NP-40 and 0.05% SDS, and cellular debris were removed by centrifugation. Aliquots of cell lysates were separated on an 8% SDS–PAGE without denaturing. Fluorescence signals were imaged by FLA5100 with 473 nm excitation and 510 nm emission, and signal intensities were calculated by Multi Gauge software (Fuji Film).

RESULTS

G-quadruplex formation by *hERα* mRNA sequence

RNA oligonucleotides corresponding to the QFP sequence derived from the ORF of the *hERα* mRNA and variants containing mutations that encode the same amino acid sequence (Figure 2A) were synthesized, and their thermodynamic stabilities and structures were evaluated *in vitro*. As the QFP sequences had melting temperature higher than 90°C in the presence of a physiological concentration (100 mM) of potassium (data not shown), UV melting data were obtained in buffer containing 3 mM potassium. Figure 2B shows normalized UV melting profiles of the QFP sequence variants analyzed from absorbance at 295 nm. Wild-type, C-mutant and G-mutant oligonucleotides had curves typical of G-quadruplex. The A-mutant and U-mutant oligonucleotides showed melting transitions at 260 nm both in the presence of 3 and 100 mM potassium (Supplementary

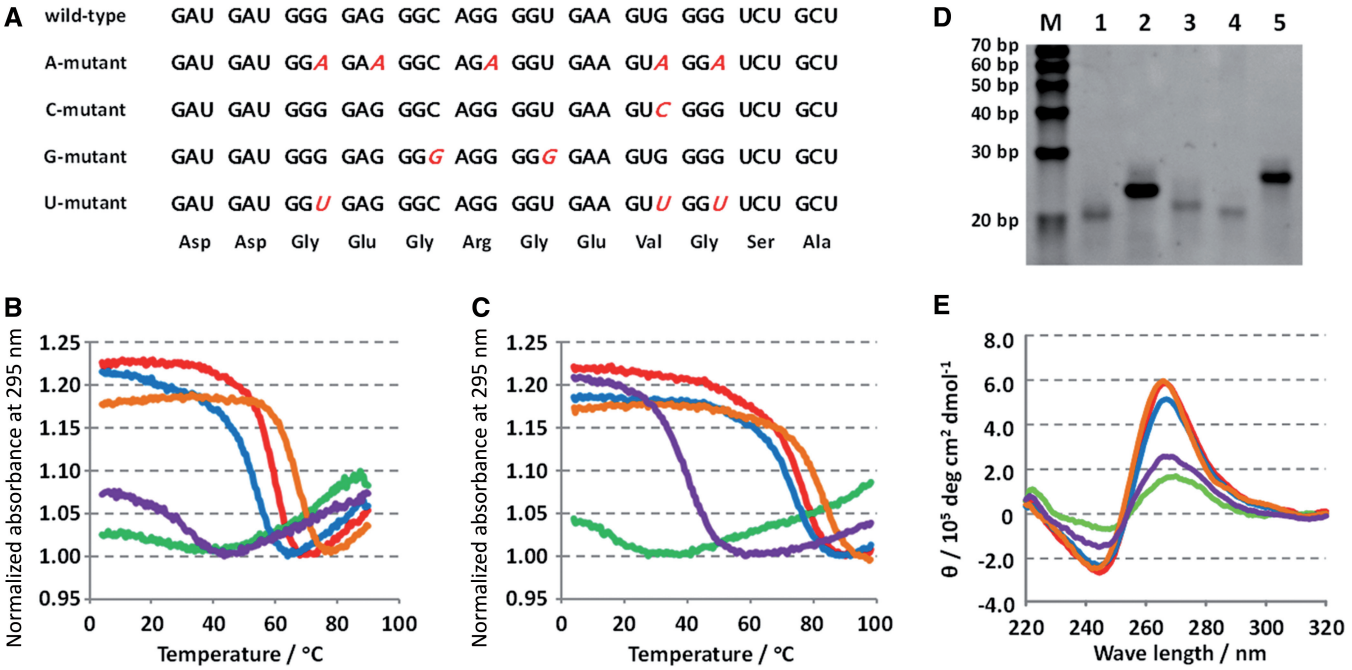

Figure 2. Formation of G-quadruplexes by QFP sequence variants. (**A**) RNA oligonucleotides corresponding to the QFP sequence derived from ORF of *hERα* mRNA and variants that mutate the G-rich regions (indicated in red and italic). Amino acid sequences encoded by the QFP sequence variants are indicated. All mutations are synonymous. (**B** and **C**) Normalized absorbance at 295 nm of 5 μM QFP sequence variants, wild-type (red), A-mutant (green), C-mutant (blue), G-mutant (orange) and U-mutant (purple), in a buffer containing 50 mM MES–LiOH, pH 7.0, and 3 mM KCl in the absence (B) and presence (C) of 40 wt% PEG200. (**D**) Non-denaturing 12% PAGE of QFP sequence variants. Ten picomoles of wild-type (lane 1), A-mutant (lane 2), C-mutant (lane 3), G-mutant (lane 4) and U-mutant (lane 5) was incubated at 95°C and cooled to room temperature at 0.2°C min^{-1} before loading. Electrophoresis was performed at 37°C with 3 W of constant electricity. Lane M shows 10 bp ladder marker. (**E**) CD spectra of 5 μM QFP sequence variants, wild-type (red), A-mutant (green), C-mutant (blue), G-mutant (orange) and U-mutant (purple), at 37°C in a buffer containing 50 mM MES–LiOH, pH 7.0, 3 mM KCl and 40 wt% PEG200.

Figure S1A), suggesting formation of secondary structures containing Watson–Crick-type base pairs in diluted solution as predicted by the mfold program (Supplementary Figure S1B) (35). However, the U-mutant also showed little melting transition at 295 nm in the presence of 100 mM potassium, although the melting temperature at 295 nm seemed less than that at 260 nm (Supplementary Figure S1C). This transition indicates that the U-mutant potentially forms G-quadruplex. To focus on the stability of G-quadruplexes formed by the QFP sequence variants, the UV melting profiles were recorded in the presence of 40 wt% polyethylene glycol (PEG) having averaged molecular weight of 200 (PEG200) as a cosolute because PEG200 has previously shown to stabilize quadruplexes but destabilize duplexes (21,25–27). Figure 2C shows normalized UV melting profiles of the QFP sequence variants at 295 nm in the presence of 3 mM potassium and 40 wt% PEG200. The melting temperatures of wild-type, C-mutant and G-mutant oligonucleotides shifted to higher temperatures under these conditions. The U-mutant also showed a clear melting transition at 295 nm. The U-mutant G-quadruplex may have mix-tetrad consists of guanines and uracils like U-tetrad shown to stabilize the human telomeric RNA G-quadruplex (Supplementary Figure S2) (36). When analyzed by native PAGE at 37°C, all the QFP sequence variants migrated as single bands (Figure 2D). Wild-type, C-mutant and G-mutant oligonucleotides migrated faster than other two mutants, suggestive of intramolecular G-quadruplex structures. The weak signal intensities of wild-type, C-mutant and G-mutant compared with the U-mutant and A-mutant signals likely reflect poorer binding of ethidium bromide to the quadruplexes.

CD spectra of the QFP sequence variants at 37°C were also obtained in the presence of 3 mM potassium and 40 wt% PEG200 (Figure 2E). Wild-type, C-mutant, G-mutant and U-mutant spectra were characterized by positive and negative peaks near 265 and 240 nm, respectively, which were larger than those of observed in the A-mutant spectra. These results indicated that all QFP sequence variants with the exception of the A-mutant formed parallel G-quadruplex structures. Wild-type, C-mutant, G-mutant and U-mutant indicated melting transitions at 265 nm that corresponded to the UV melting transitions in Figure 2C (Supplementary Figure S3A). In addition, the CD spectra had isosbestic point near 250 nm and linear correlation between signal intensities at 265 and 242 nm during the melting transition (Supplementary Figure S3B). Thus, the RNA G-quadruplexes formed by the QFP sequence variants ware suggested to be melted by two-state transition. CD spectra of wild-type, C-mutant and G-mutant were virtually identical in the presence and absence of 40 wt% PEG200 (Supplementary Figure S4). The monophonic properties of RNA G-quadruplex irrespective of solution

Table 1. Thermodynamic stabilities of QFP sequence variants in the presence of 3 mM potassium and 40 wt% PEG200

QFP sequence variants	ΔG°_{37}[a]/kcal mol^{-1}	Tm/°C
wild-type	−5.15 ± 0.84	76.4 ± 0.9
A-mutant	ND	ND
C-mutant	−4.77 ± 0.37	72.6 ± 1.2
G-mutant	−5.86 ± 0.55	82.4 ± 1.0
U-mutant	−0.47 ± 0.03	39.9 ± 0.5

[a]Values of ΔG°_{37} were calculated from five melting profiles of the QFP sequence variants at the heating rate of 0.2°C min^{-1}.

conditions (21,37) might be a cause of the two-state transitions. Thermodynamic stabilities at 37°C (ΔG°_{37}) and melting temperatures (Tm) of the QFP sequence variants in the presence of 3 mM potassium and 40 wt% PEG200 were calculated from the UV melting profiles by van't Hoff equation (Table 1).

Suppression of translation by QFP sequence

To evaluate the effect of the G-quadruplex formed by the QFP sequence derived from the *hERα* mRNA on translation elongation, reporter mRNAs containing the QFP sequence variants at 3′ terminal of the mRNA were constructed (Figure 3A). Formation of G-quadruplex in the mRNAs was confirmed by using fluorescence signal of *N*-methyl mesoporphyrin (NMM), which specifically binds to G-quadruplex structure and increase its fluorescence (Supplementary Figure S5A) (38). Higher fluorescence intensities with wild-type, C-mutant, G-mutant and U-mutant mRNAs than that with A-mutant mRNA or without mRNA suggest formation of G-quadruplexes in the reporter mRNAs containing wild-type and mutant QFP sequences except A-mutant. It should be noted that the mRNAs were refolded in the presence of 100 mM potassium and absence of PEG200 because PEG200 would disturb *in vitro* translation (39). The U-mutant mRNA was considered to partially formed G-quadruplex as expected from UV melting profile (Supplementary Figure S1C) and slightly increased the fluorescence signal of NMM.

Time course of translation elongation was analyzed *in vitro* by using a synchronized translation that we recently developed to evaluate a single turnover of the elongation reaction (32). In the first step of the reaction, translation was performed in a reconstructed *in vitro* translation system (PURExpress; GeneFrontier) in which RF1, RF2 and TyrRS were left out of the reaction mixture; this artificially halted translation before the 30th codon encoding tyrosine (Figure 3A). During this reaction, a non-natural fluorescent amino acid (CR-110-X-AF) was incorporated into the protein by using an amber suppressor tRNA aminoacylated with CR110-X-AF (CR110-X-AF-tRNAamber; Protein Express). In the second step, translation elongation was restarted by addition of a reaction mixture containing all requisite proteins except the ribosome. This second reaction was stopped at 20, 60, 120, 300 and 600 s from the restart. Fluorescence signal of CR110-X-AF, which was incorporated into translated product, was imaged after separation by Tris–tricine SDS–PAGE. Figure 3B shows the time course of the translation elongation using the reporter mRNAs containing QFP sequence variants. Strong signals near the bottom of the gel are due to unincorporated CR110-X-AF (32). Before the restart, the fluorescence signal of a 29-residue peptide because of the artificial halt of translation before the tyrosine codon was similar in reaction with all the reporter mRNAs (Figures 3B, indicated by H). In the case of wild-type, C-mutant and G-mutant mRNAs, which form relatively stable G-quadruplexes, the main translation products were observed at a slightly higher position (Figure 3B, indicated by I) comparing with the product before the restart. The fluorescence signal indicated by I corresponds translation product in which translation elongation was stalled at an intermediate position of mRNA because the main translation products with U-mutant and A-mutant mRNAs were observed at further higher position (Figure 3B, indicated by F) despite the amino acid sequence encoded by all of the reporter mRNAs are the same (Figure 3A). The fluorescence signal indicated by F corresponds the final product in which the translation elongation progressed to the 3′-end of the mRNA. The intermediate product indicated by I was temporarily observed at 20 s from the restart in the case of the U-mutant mRNA, and the signal was gradually disappeared with increasing the reaction time. In addition, the final product indicated by F was gradually increased with increasing the reaction time in the case of wild-type and C-mutant mRNAs. From these results, it was suggested that the RNA G-quadruplex formed in the reporter mRNAs stalled the translation elongation, and time range of the stalling depends on the stability of the RNA G-quadruplex (Table 1).

To determine the position, where the translation elongation was stalled, the mass spectrum of the reaction at 20 s after the translation restart was analyzed using wild-type mRNA (Figure 3D). The signal at 5591.8 kDa corresponded to the mass of a 42-residue peptide with an adenosine moiety at the C terminus (calculated mass was 5592.2 with protonation). The position of the 42nd codon corresponds to a site on the mRNA six nucleotides before the QFP sequence. The stalling position corresponded to the position that we observed using mRNA-containing RNA G-quadruplexes derived from *Escherichia coli* genes (28). The mass of the 42-residue peptide was not observed in the translation product with A-mutant mRNA (data not shown). These data also indicated that translation elongation was specifically stalled by the G-quadruplex structure formed on the mRNA that block the mRNA entry into ribosome.

hERα expression depends on QFP sequence

The potent translation suppression mediated by the RNA G-quadruplex formed by QFP sequence of *hERα* mRNA may impact *hERα* expression inside cells. The coding sequence of *hERα* was cloned into plasmid vector, and variants containing synonymous mutations in the QFP sequence were constructed. To confirm whether the QFP sequence in natural *hERα* mRNA form RNA

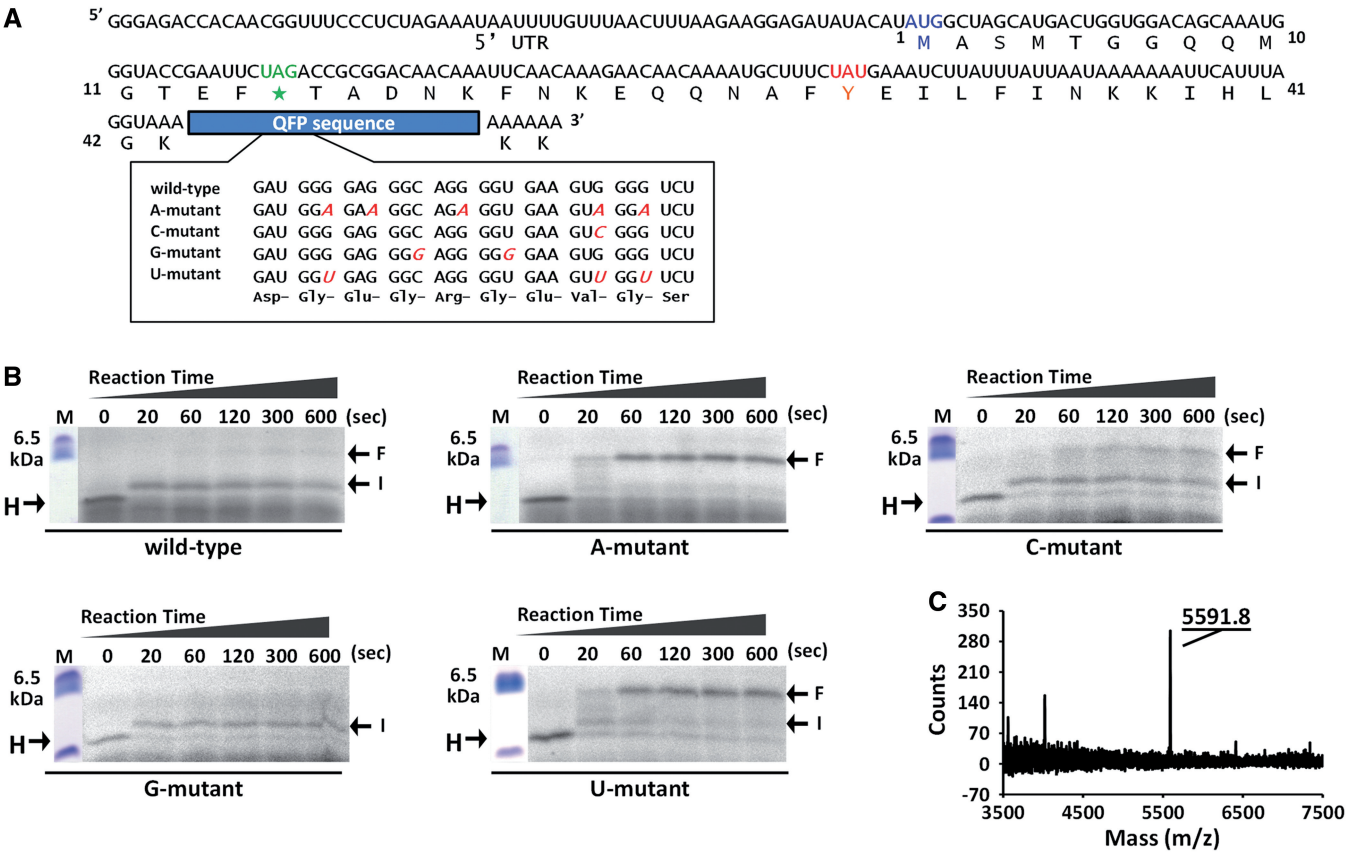

Figure 3. Translational halt before QFP sequence *in vitro*. (**A**) Nucleotide and amino acid sequences of reporter mRNAs containing QFP sequence variants derived from *hERα*. Positions of start codon (blue), amber codon for incorporation of non-natural fluorescent amino acid (green, asterisk) and tyrosine codon to artificially halt translation (orange) are indicated. Nucleotides mutated relative to wild-type QFP sequence are indicated in red. (**B** and **C**) Time course of the synchronized translation of reporter mRNAs. Samples taken from the synchronized translation at indicated reaction times were resolved on 16% Tris–tricine SDS–PAGE. The fluorescence signal of incorporated non-natural amino acid (CR110-X-AF) was imaged using 473 nm excitation and 510 nm emission. Translated products because of ribosomal stalling (I) and complete elongation (F) are indicated. (**D**) Mass spectrometry analysis of translated products of wild-type mRNA after 20 s of synchronized translation. Observed mass of main signal is indicated.

G-quadruplex, as well as reporter mRNA, full-length mRNA encoding *hERα* variants were prepared by *in vitro* transcription and mixed with NMM (Supplementary Figure S5B). The results were similar to those of the reporter mRNAs for synchronized translation. Although slightly higher fluorescence signal with A-mutant *hERα* mRNA compared with that without mRNA made it difficult to find significant difference between the A-mutant and U-mutant *hERα* mRNAs, other three mRNAs containing wild-type, C-mutant and G-mutant QFP sequences were confirmed to form G-quadruplex because the fluorescence intensities were further increased comparing with the A-mutant and U-mutant *hERα* mRNAs. It is considered that NMM non-specifically bound to the A-mutant *hERα* mRNA because the length of the *hERα* mRNAs (1946 nt) was much longer than that of reporter mRNAs (228 nt) used in Supplementary Figure S5A.

The plasmid vectors were transfected into HepG2 human liver carcinoma cells in which the expression of endogenous hERα is extremely low (40). Figure 4 shows results of western blotting using anti-hERα antibody specific for N-terminal domain of the protein (Thermo Scientific). It should be noted that the hERα expressed from the vectors contains an additional seven amino acids at C-terminus relative to the endogenous form, and the full-length hERα obtained from the expression vectors should be 67 kDa. In cells transfected with all the QFP sequence variants, a signal corresponding to full-length hERα was observed located between the 50- and the 75-kDa marker proteins (Figure 4, indicated by F). In cells transfected with wild-type and G-mutant, a strong signal was observed from a band that migrated slightly slower than the 25-kDa marker protein (Figure 4, indicated by N). Weaker signals from this region were also observed in C-mutant and U-mutant transfected cells, whereas the signal was vanishingly low in A-mutant transfected cells. The amino acid sequence of hERα expressed from each of these vectors is the same. As the signal intensities of the short hERα correlated with the thermodynamic stability of the G-quadruplex formed by the QFP sequence variants (Table 1), the production of the short hERα was likely caused by the differences in the G-quadruplexes formed within the ORF of mRNAs.

As shown in Figure 3, if the translation elongation was stalled before the QFP sequence and truncated at the

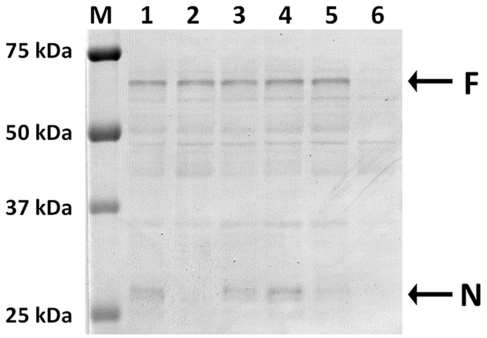

Figure 4. Expression of hERα in HepG2 cells. Vectors for expression of hERα containing the QFP sequence variants, wild-type (lane 1), A-mutant (lane 2), C-mutant (lane 3), G-mutant (lane 4) and U-mutant (lane 5), or control plasmid (pcDNA5/FRT, Invitrogen) (lane 6) were transfected into HepG2 cells. Cell lysates obtained after 48-h incubation at 37°C were resolved on 8% SDS–PAGE. Expression of hERα was analyzed by western blotting using an anti-hERα antibody specific for the N-terminal domain. Lane M shows marker proteins. Signals of hERα correspond to full-length protein, and shorter forms are indicated by F and N, respectively.

position during the stalling, the product would include 273 amino acids and have a molecular weight of 30.0 kDa, which is close to that of the product observed on the gel (Figure 4 indicated by N). The G-quadruplex formed by the QFP sequence might cause the production of the short hERα by inducing an mRNA surveillance mechanism termed no-go mRNA decay, which was recently discovered in yeast and is thought to be caused by an artificially introduced stable duplex (41,42). A previous study demonstrated that the hERα translated *in vitro* using rabbit reticulocyte and wheat germ extracts exhibited different properties against proteolysis by chymotrypsin (43). This observation suggests an alternative possibility that the production of the short hERα might due to the proteolysis of the full-length product.

Proteolysis of hERα depends on QFP sequence

To quantitatively evaluate the effect of the G-quadruplex stabilities on the production of the short hERα, the coding sequence of the fluorescent protein (AcGFP) was inserted either into the 5′ or the 3′ region of hERα to express AcGFP fused to hERα at the N-terminus (AcGFP–hERα) or the C-terminus (hERα–AcGFP), respectively. Vectors were transfected into MCF7 human breast carcinoma cells that consecutively express hERα. AcGFP preserves its fluorescence during SDS–PAGE unless it has been denatured before the electrophoresis. Figure 5A and B show fluorescence signals of AcGFP–hERα and hERα–AcGFP, respectively, in 8% SDS–PAGE. AcGFP–hERα showed similar expression patterns when analyzed by western blotting. Fluorescence signals of the full-length protein were observed near the 75-kDa marker protein in all the QFP sequence variants, and for certain QFP sequence variants a signal was observed between the 37- and the 50-kDa maker proteins (Figure 5A, indicated by N). In the case of hERα–AcGFP, although several signals appeared uniformly in all the QFP sequence variants, a fluorescence signal near the 50-kDa marker was observed for certain QFP sequence variants (Figure 5B, indicated by C). Because hERα–AcGFP contains AcGFP at C-terminus, the fluorescence signal depending on the QFP sequence variants would not be observed if the cause of the short hERα observed in Figure 4 was truncation of translation during the stalling of the translation elongation. Thus, the expression patterns of both AcGFP–hERα and hERα–AcGFP suggested that the short hERα results from proteolysis.

Ratios of the cleaved AcGFP–hERα (Figure 5A, indicated by N) and hERα–AcGFP (Figure 5B, indicated by C) relative to full-length protein (Figure 5A and B, indicated by F) were calculated from the fluorescence intensities (Figure 5C and D). Both ratios depended on the thermodynamic stabilities of the G-quadruplex formed by the QFP sequence variant; the variants able to form more stable G-quadruplexes resulted in more production of the short hERα. The ratios calculated for the A-mutant were used as standards, because the A-mutant did not form G-quadruplex even in the presence of 100 mM potassium and 40 wt% PEG200 (data not shown). Figure 5E and F show plots of relative ratio values of QFP sequence variants against the ratio value of A-mutant versus the thermodynamic stabilities of the QFP sequence variants calculated from UV melting transition in Figure 2B (Table 1). The data from both the cleaved AcGFP–hERα and hERα–AcGFP fit a single exponential function. In HepG2 cells and in cells from a human embryonic kidney cell-line (Flp-In 293, Invitrogen), the relative ratios of cleaved AcGFP–hERα also fit a single exponential function (Supplementary Figure S6).

DISCUSSION

During translation elongation, the ribosome must unwind any structure formed by an mRNA. The mRNA structures and stabilities would affect both efficacy and velocity of the translation elongation. In the case of G-quadruplexes, it was recently suggested that G-quadruplex formation in 5′-untranslated region (UTR) of mRNAs potentially suppress protein expression level by inhibiting small ribosomal subunit in scanning, especially in the UTRs of oncogene mRNAs (44–47). The effect of the G-quadruplexes in ORFs of mRNAs has not been studied, although it has been speculated that suppression of elongation might impact folding and processing of nascent proteins (13,15). The QFP sequence in the ORF of the *hERα* mRNA is located in the region that encodes the hinge between the DNA-binding domain and the ligand-binding domain (Figure 1) (29). It was previously shown that proteolysis of hERα by chymotrypsin depends on the *in vitro* translation system used (43). Thus, structure and processing of hERα are likely to be affected if the QFP sequence forms a G-quadruplex and inhibits ribosome in elongation reaction.

The QFP sequence derived from *hERα* mRNA was cloned into 3′ region of reporter mRNA, and effects of the QFP sequence on translation elongation were evaluated in a synchronized translation system (Figure 3). Translation elongation was suppressed at six

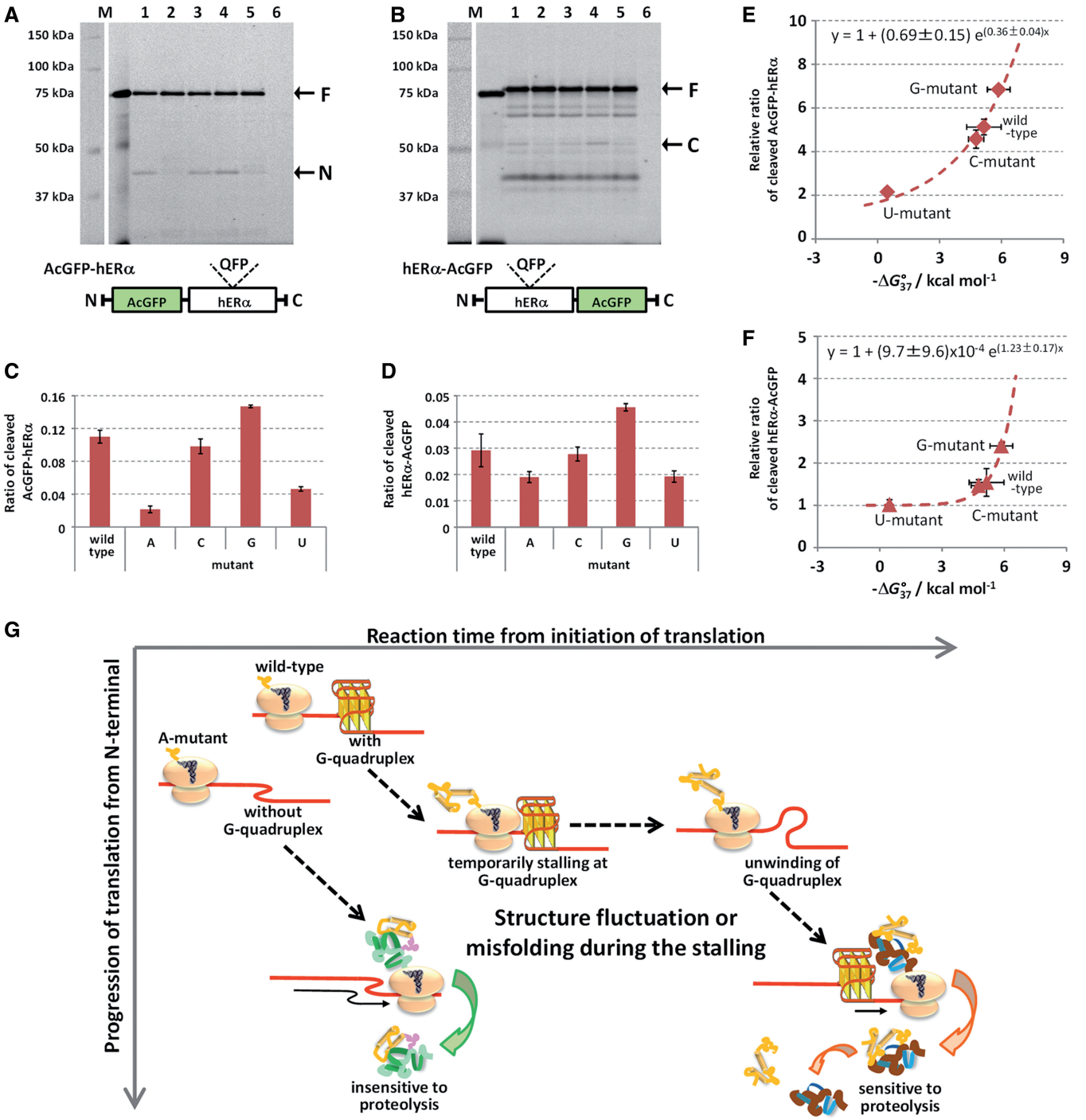

Figure 5. Proteolysis of hERα depending of thermodynamic stabilities of G-quadruplexes. (**A** and **B**) Protein expression patterns of AcGFP–hERα (A) and hERα–AcGFP (B) in MCF7 cells. Cells were transfected with vectors for expression of AcGFP–hERα or hERα–AcGFP containing the QFP sequence variants, wild-type (lane 1), A-mutant (lane 2), C-mutant (lane 3), G-mutant (lane 4) and U-mutant (lane 5) or were not transfected with plasmid (lane 6). Cells were lysed after 48-h. Cell lysates were resolved on 8% SDS–PAGE without denaturing, and fluorescence signals were imaged using 473 nm excitation and 510 nm emission. (**C** and **D**) Ratio of cleaved AcGFP–hERα (C) and hERα–AcGFP (D) relative to the full-length protein ($n = 3$). (**E** and **F**) Plots of the relative ratios of the cleaved AcGFP–hERα (E) and hERα–AcGFP (F) normalized to the ratio of A-mutant versus $-\Delta G^{\circ}_{37}$ of the QFP sequence variants calculated from the UV melting. Fits to a single exponential equation (interior and dashed line) are indicated. Error bars represent SD. (**G**) Scheme of translation procedures of wild-type mRNA (with G-quadruplex) and A-mutant mRNA (without G-quadruplex). Temporarily stalling of the ribosome before the G-quadruplex affects co-translational folding of nascent protein and makes the protein structure sensitive to proteolysis.

nucleotides before the wild-type QFP sequence. The parallel G-quadruplex formed by the QFP sequence (Figure 2) likely inhibits the helicase activity of ribosome (48) and blocks mRNA entry into the ribosome. The translation suppression was temporary rather than permanent because a full-length product was observed with increasing the reaction time depending on the thermodynamic stabilities of RNA G-quadruplexes formed by the QFP sequence variants (Figure 3B).

Analysis of expression of hERα in HepG2 cells indicated production of short hERα with a molecular weight ~25–30 kDa that was recognized by anti-hERα antibody specific for N-terminal domain (Figure 4). The amount of short hERα produced depended on the synonymous mutations in the QFP sequence; none of this form was observed on translation of the A-mutant that could not form G-quadruplex. In addition, expression of mRNAs encoding AcGFP–hERα and hERα–AcGFP yielded products shorter than full-length protein in amounts that depended on the QFP sequence variants (Figure 5 and Supplementary Figure S5). These results indicated that the QFP sequence contributed proteolysis of hERα at a specific position. The relative ratios of the short to full-length products depended on the thermodynamic stabilities of the QFP sequence variants (Figure 5E and F), as well as suppression of translation elongation *in vitro* (Figure 3B). The ribosome should completely unwind G-quadruplex into single strand and import the mRNA to continue the translation. In addition, the ribosome progresses three nucleotides in every translocation reaction, which occurs within 0.1 s (1). This progression would be enough to unwind G-quadruplex consisting of three quartets at a time, and intermediate states during dissociation of G-quadruplex would be negligible. Thus, we assumed that the activation energy necessary for the ribosome to continue the translation elongation is proportional to the thermodynamic stability of the structure. In this case, the rate to overcome the stalling should depend on the G-quadruplex stability in a single exponential manner, and this is indeed what we observe. Similar results have been observed *in vitro* in correlation between mechanical strength of RNA pseudoknots and ribosomal frameshift, which is considered to be caused by temporary stalling of translation elongation before the pseudoknot (49). Thus, temporary stalling of translation elongation before the G-quadruplex (Figure 3B) is likely a main determinant of the proteolysis sensitivity of hERα by inducing fluctuation or misfolding of the nascent hERα (Figure 5G). The differences observed in the rabbit and wheat germ *in vitro* translation systems (43) might be due to differences in G-quadruplex stability in the two systems.

Although the proteolysis position has not been accurately determined, the short hERα had a molecular weight between 25 and 30 kDa (Figure 4, indicated by N). We assume that hERα is cleaved after the DNA-binding domain because the molecular weight of N-terminal 251 amino acids containing the activation function-1 domain (domain A/B) and the DNA-binding domain should be 27.6 kDa (29). The *hERα* pre-mRNA is known to be alternatively spliced resulting in hERα variants that confer dominant-positive and dominant-negative effects on the protein functions (50), and these variants are hypothesized to contribute to progression of certain cancers (51–54). The proteolysis products of wild-type hERα might retain certain functions, as well as products of the alternative splicing. In that case, the RNA G-quadruplex is considered to operate the protein polymorphism.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–6.

ACKNOWLEDGEMENTS

Customized PURExfrex was kindly provided by GeneFrontier Corp. (http://www.genefrontier.com/english/index.html). The authors thank Misa Kinoshita, Nobuaki Hattori and Takashi Tokisue for supporting experiments.

FUNDING

Grants-in-Aid for Scientific Research; MEXT-Supported Program for the Strategic Research Foundation at Private Universities (2009–2014), Japan. Funding for open access charge: Grants-in-Aid for Scientific Research, Japan.

Conflict of interest statement. None declared.

REFERENCES

1. Wen,J.D., Lancaster,L., Hodges,C., Zeri,A.C., Yoshimura,S.H., Noller,H.F., Bustamante,C. and Tinoco,I. (2008) Following translation by single ribosomes one codon at a time. *Nature*, **452**, 598–603.
2. Buchan,J.R. and Stansfield,I. (2007) Halting a cellular production line: responses to ribosomal pausing during translation. *Biol. Cell*, **99**, 475–487.
3. Giedroc,D.P. and Cornish,P.V. (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus Res.*, **139**, 193–208.
4. Farabaugh,P.J. (1996) Programmed translational frameshifting. *Microbiol. Rev.*, **60**, 103–134.
5. Fernandez,J., Yaman,I., Huang,C., Liu,H., Lopez,A.B., Komar,A.A., Caprara,M.G., Merrick,W.C., Snider,M.D., Kaufman,R.J. *et al.* (2005) Ribosome stalling regulates IRES-mediated translation in eukaryotes, a parallel to prokaryotic attenuation. *Mol. Cell*, **17**, 405–416.
6. Zhang,G. and Ignatova,Z. (2009) Generic algorithm to predict the speed of translational elongation: implications for protein biogenesis. *PLoS One*, **4**, e5036.
7. Watts,J.M., Dang,K.K., Gorelick,R.J., Leonard,C.W., Bess,J.W. Jr, Swanstrom,R., Burch,C.L. and Weeks,K.M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, **460**, 711–716.
8. Zhang,G., Hubalewska,M. and Ignatova,Z. (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.*, **16**, 274–280.
9. Komar,A.A. (2009) A pause for thought along the co-translational folding pathway. *Trends Biochem. Sci.*, **34**, 16–24.
10. O'Brien,E.P., Hsu,S.T., Christodoulou,J., Vendruscolo,M. and Dobson,C.M. (2010) Transient tertiary structure formation within the ribosome exit port. *J. Am. Chem. Soc.*, **132**, 16928–16937.
11. Kaiser,C.M., Goldman,D.H., Chodera,J.D., Tinoco,I. Jr and Bustamante,C. (2011) The ribosome modulates nascent protein folding. *Science*, **334**, 1723–1727.

12. O'Brien,E.P., Christodoulou,J., Vendruscolo,M. and Dobson,C.M. (2011) New scenarios of protein folding can occur on the ribosome. *J. Am. Chem. Soc.*, **133**, 513–526.
13. Zhang,G. and Ignatova,Z. (2011) Folding at the birth of the nascent chain: coordinating translation with co-translational folding. *Curr. Opin. Struct. Biol.*, **21**, 25–31.
14. Tsai,C.J., Sauna,Z.E., Kimchi-Sarfaty,C., Ambudkar,S.V., Gottesman,M.M. and Nussinov,R. (2008) Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J. Mol. Biol.*, **383**, 281–291.
15. O'Brien,E.P., Vendruscolo,M. and Dobson,C.M. (2012) Prediction of variable translation rate effects on cotranslational protein folding. *Nat. Commun.*, **3**, 868.
16. Kimchi-Sarfaty,C., Oh,J.M., Kim,I.W., Sauna,Z.E., Calcagno,A.M., Ambudkar,S.V. and Gottesman,M.M. (2007) A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science*, **315**, 525–528.
17. Evans,M.S., Sander,I.M. and Clark,P.L. (2008) Cotranslational folding promotes beta-helix formation and avoids aggregation *in vivo*. *J. Mol. Biol.*, **383**, 683–692.
18. Thanaraj,T.A. and Argos,P. (1996) Ribosome-mediated translational pause and protein domain organization. *Protein Sci.*, **5**, 1594–1612.
19. Williamson,J.R. (1994) G-quartet structures in telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 703–730.
20. Collie,G.W. and Parkinson,G.N. (2011) The application of DNA and RNA G-quadruplexes to therapeutic medicines. *Chem. Soc. Rev.*, **40**, 5867–5892.
21. Zhang,D.H., Fujimoto,T., Saxena,S., Yu,H.Q., Miyoshi,D. and Sugimoto,N. (2010) Monomorphic RNA G-quadruplex and polymorphic DNA G-quadruplex structures responding to cellular environmental factors. *Biochemistry*, **49**, 4554–4563.
22. Joachimi,A., Benz,A. and Hartig,J.S. (2009) A comparison of DNA and RNA quadruplex structures and stabilities. *Bioorg. Med. Chem.*, **17**, 6811–6815.
23. Arora,A. and Maiti,S. (2009) Differential biophysical behavior of human telomeric RNA and DNA quadruplex. *J. Phys. Chem. B*, **113**, 10515–10520.
24. Pramanik,S., Nagatoishi,S. and Sugimoto,N. (2012) DNA tetraplex structure formation from human telomeric repeat motif (TTAGGG):(CCCTAA) in nanocavity water pools of reverse micelles. *Chem. Commun.*, **48**, 4815–4817.
25. Miyoshi,D., Karimata,H. and Sugimoto,N. (2006) Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions. *J. Am. Chem. Soc.*, **128**, 7957–7963.
26. Pramanik,S., Nagatoishi,S., Saxena,S., Bhattacharyya,J. and Sugimoto,N. (2011) Conformational flexibility influences degree of hydration of nucleic acid hybrids. *J. Phys. Chem. B*, **115**, 13862–13872.
27. Nakano,S., Karimata,H., Ohmichi,T., Kawakami,J. and Sugimoto,N. (2004) The effect of molecular crowding with nucleotide length and cosolute structure on DNA duplex stability. *J. Am. Chem. Soc.*, **126**, 14330–14331.
28. Endoh,T., Kawasaki,Y. and Sugimoto,N. (2013) Suppression of gene expression by G-quadruplexes in open reading frames depends on G-quadruplex stability. *Angew. Chem. Int. Ed.*, in press.
29. Gibson,D.A. and Saunders,P.T. (2012) Estrogen dependent signaling in reproductive tissues - a role for estrogen receptors and estrogen related receptors. *Mol. Cell. Endocrinol.*, **348**, 361–372.
30. Schultz-Norton,J.R., Ziegler,Y.S. and Nardulli,A.M. (2011) ERalpha-associated protein networks. *Trends Endocrinol. Metab.*, **22**, 124–129.
31. Balkwill,G.D., Derecka,K., Garner,T.P., Hodgman,C., Flint,A.P. and Searle,M.S. (2009) Repression of translation of human estrogen receptor alpha by G-quadruplex formation. *Biochemistry*, **48**, 11487–11495.
32. Endoh,T., Kawasaki,Y. and Sugimoto,N. (2012) Synchronized translation for detection of temporal stalling of ribosome during single-turnover translation. *Anal. Chem.*, **84**, 857–861.
33. Marky,L.A. and Breslauer,K.J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers*, **26**, 1601–1620.
34. Schagger,H. (2006) Tricine-SDS-PAGE. *Nat. Protoc.*, **1**, 16–22.
35. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
36. Xu,Y., Ishizuka,T., Kimura,T. and Komiyama,M. (2010) A U-tetrad stabilizes human telomeric RNA G-quadruplex structure. *J. Am. Chem. Soc.*, **132**, 7231–7233.
37. Zhang,A.Y., Bugaut,A. and Balasubramanian,S. (2011) A sequence-independent analysis of the loop length dependence of intramolecular RNA G-quadruplex stability and topology. *Biochemistry*, **50**, 7251–7258.
38. Arthanari,H., Basu,S., Kawano,T.L. and Bolton,P.H. (1998) Fluorescent dyes specific for quadruplex DNA. *Nucleic Acids Res.*, **26**, 3724–3728.
39. Ge,X., Luo,D. and Xu,J. (2011) Cell-free protein expression under macromolecular crowding conditions. *PLoS One*, **6**, e28707.
40. Harnish,D.C., Evans,M.J., Scicchitano,M.S., Bhat,R.A. and Karathanasis,S.K. (1998) Estrogen regulation of the apolipoprotein AI gene promoter through transcription cofactor sharing. *J. Biol. Chem.*, **273**, 9270–9278.
41. Harigaya,Y. and Parker,R. (2010) No-go decay: a quality control mechanism for RNA in translation. *Wiley Interdiscip. Rev. RNA*, **1**, 132–141.
42. Doma,M.K. and Parker,R. (2006) Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature*, **440**, 561–564.
43. Horjales,S., Cota,G., Senorale-Pose,M., Rovira,C., Roman,E., Artagaveytia,N., Ehrlich,R. and Marin,M. (2007) Translational machinery and protein folding: evidence of conformational variants of the estrogen receptor alpha. *Arch. Biochem. Biophys.*, **467**, 139–143.
44. Halder,K., Wieland,M. and Hartig,J.S. (2009) Predictable suppression of gene expression by 5′-UTR-based RNA quadruplexes. *Nucleic Acids Res.*, **37**, 6811–6817.
45. Beaudoin,J.D. and Perreault,J.P. (2010) 5′-UTR G-quadruplex structures acting as translational repressors. *Nucleic Acids Res.*, **38**, 7022–7036.
46. Bugaut,A. and Balasubramanian,S. (2012) 5′-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res.*, **40**, 4727–4741.
47. Kumari,S., Bugaut,A. and Balasubramanian,S. (2008) Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5′ UTR of the NRAS proto-oncogene. *Biochemistry*, **47**, 12664–12669.
48. Takyar,S., Hickerson,R.P. and Noller,H.F. (2005) mRNA helicase activity of the ribosome. *Cell*, **120**, 49–58.
49. Chen,G., Chang,K.Y., Chou,M.Y., Bustamante,C. and Tinoco,I. Jr (2009) Triplex structures in an RNA pseudoknot enhance mechanical stability and increase efficiency of −1 ribosomal frameshifting. *Proc. Natl Acad. Sci. USA*, **106**, 12706–12711.
50. Barone,I., Brusco,L. and Fuqua,S.A. (2010) Estrogen receptor mutations and changes in downstream gene expression and signaling. *Clin. Cancer Res.*, **16**, 2702–2708.
51. Bollig,A. and Miksicek,R.J. (2000) An estrogen receptor-alpha splicing variant mediates both positive and negative effects on gene transcription. *Mol. Endocrinol.*, **14**, 634–649.
52. Flouriot,G., Brand,H., Denger,S., Metivier,R., Kos,M., Reid,G., Sonntag-Buck,V. and Gannon,F. (2000) Identification of a new isoform of the human estrogen receptor-alpha (hER-alpha) that is encoded by distinct transcripts and that is able to repress hER-alpha activation function 1. *EMBO J.*, **19**, 4688–4700.
53. Chaidarun,S.S. and Alexander,J.M. (1998) A tumor-specific truncated estrogen receptor splice variant enhances estrogen-stimulated gene expression. *Mol. Endocrinol.*, **12**, 1355–1366.
54. Garcia Pedrero,J.M., Zuazua,P., Martinez-Campa,C., Lazo,P.S. and Ramos,S. (2003) The naturally occurring variant of estrogen receptor (ER) ERDeltaE7 suppresses estrogen-dependent transcriptional activation by both wild-type ERalpha and ERbeta. *Endocrinology*, **144**, 2967–2976.