

Translation of the *FMR1* mRNA is not influenced by AGG interruptions

Anna L. Ludwig¹, Christopher Raske¹, Flora Tassone^{1,2}, Dolores Garcia-Arocena¹, John W. Hershey¹ and Paul J. Hagerman^{1,*}

¹Department of Biochemistry and Molecular Medicine, School of Medicine, University of California Davis, 4303 Tupper Hall, Davis, CA 95616 and ²MIND Institute, University of California Davis, Health System, Sacramento, CA, USA

Received July 15, 2009; Revised August 12, 2009; Accepted August 13, 2009

ABSTRACT

The fragile X mental retardation 1 (*FMR1*) gene contains a CGG-repeat element within its 5' untranslated region (5'UTR) which, for alleles with more than ~40 repeats, increasingly affects both transcription (up-regulation) and translation (inhibition) of the repeat-containing RNA with increasing CGG-repeat length. Translational inhibition is thought to be due to impaired ribosomal scanning through the CGG-repeat region, which is postulated to form highly stable secondary/tertiary structure. One striking difference between alleles in the premutation range (55–200 CGG repeats) and those in the normal range (<~40 repeats) is the reduced number/absence of 'expansion stabilizing' AGG interruptions in the larger alleles. Such interruptions, which generally occur every 9–11 repeats in normal alleles, are thought to disrupt the extended CGG-repeat hairpin structure, thus facilitating translational initiation. To test this hypothesis, we have measured the translational efficiency of CGG-repeat mRNAs with 0–2 AGG interruptions, both *in vitro* (rabbit reticulocyte lysates) and in cell culture (HEK-293 cells). We demonstrate that the AGG interruptions have no detectable influence on translational efficiency in either a cell-free system or cell culture, indicating that any AGG-repeat-induced alterations in secondary/tertiary structure, if present, do not involve the rate-limiting step(s) in translational initiation.

INTRODUCTION

The fragile X mental retardation 1 (*FMR1*) gene (1,2) [OMIM + 309550] codes for an RNA-binding protein (FMRP) that plays an important role in neurodevelopment through its regulatory functions in synaptogenesis

and synaptic plasticity. The 5' untranslated region (5'UTR) of the *FMR1* gene harbors a variable-length trinucleotide (CGG) repeat element that, when expanded into the premutation range (55–200 CGG repeats), is known to both up-regulate transcription and inhibit translation (3–10). For larger expansions (> 200 CGG repeats; full mutation range), the promoter region of the gene generally becomes hypermethylated and transcriptionally silenced, with loss of FMRP resulting in the neurodevelopmental disorder fragile X syndrome (11,12), the leading heritable form of cognitive impairment and the leading known disorder associated with autism.

For CGG-repeat expansions within the premutation range, *FMR1* mRNA levels are elevated from 2- to 8-fold (5–8,13–15). The expanded-CGG-repeat mRNA *per se* is currently believed to cause the late-adult-onset neurodegenerative disorder, fragile X-associated tremor/ataxia syndrome (FXTAS) through a direct 'toxic RNA gain-of-function' (16–21, reviews: 22,23). The core features of FXTAS include intention tremor and gait ataxia, with associated features of cognitive decline/dementia, parkinsonism, neuropathy and dysautonomia. Further implicating the expanded-repeat mRNA is its presence within the intranuclear neuronal and astrocytic inclusions throughout the brain in affected individuals (24–26), and the induction of inclusion formation in cultured neural cells upon expression of the expanded-CGG-repeat element upstream of a reporter (GFP) gene (19).

One of the long-standing puzzles with respect to the CGG-repeat element is the functional significance of periodic C-to-A transversions, often referred to as AGG 'interruptions' or 'anchors'. These AGG interruptions typically appear every 9–11 tri-nucleotide repeats (27–32) within normal length CGG-repeat tracts; however, they are diminished in number or absent from larger repeat tracts in the premutation and full mutation ranges. On the basis of structure-prediction algorithms and *in vitro* cleavage/protection studies of the structure of the CGG-repeat region (33,34), the AGG interruptions were

*To whom correspondence should be addressed. Tel: +1 530 754 7266; Fax: +1 530 754 7269; Email: pjhagerman@ucdavis.edu

postulated to prevent the formation of long CGG-repeat hairpin structures that would otherwise strongly impede translational initiation. However, the influence of AGG interruptions on translational initiation has never been tested at the functional level. In the current work, we have examined the influence of AGG interruptions on the efficiency of translation both in a cell-free system and in cultured cells; we conclude that the AGG interruptions have no influence on translation.

MATERIALS AND METHODS

Construction of AGG-interruption constructs for *in vitro* studies

The firefly luciferase (FL) reporter plasmid, pBR-SP6-5'-*FMRI*-UTR (CGG)₉₉-FL-polyA (35), was modified to include *FMRI* exon 1 by digesting the plasmid with the restriction enzymes *NheI* and partial *NarI* (NEB, New England Biolabs, Inc, Ipswich, MA) followed by insertion of the linker, 5'-CTATACCATGGAAGACGCCAAAAA CATAAAGAAAGGCCCGG-3', creating pBR-SP6-*FMRI*5'UTR(99)-FL-Linker. *FMRI* exon 1 was then removed from the plasmid pE5.1 (1) by digestion with *NheI* and *RsaI* (NEB), followed by insertion of the exon-containing fragment into *NheI*/*NcoI* linearized pBR-SP6-*FMRI*5'UTR(99)-FL-Linker, creating pBR-SP6-*FMRI* 5'UTR(99)-*FMRI*exon1-FL. The CGG(99) element was then replaced by 3 [CGG + *n*(AGG)] elements in the following manner. Genomic DNAs from premutation carriers were amplified by PCR with primers f (5'-AGCCCCGCAC TTCCACCACCAGCTCCTCCA-3') (36) and c3 (5'-TGT TTACACCCGCAGCGGGCCGGGGTTC-3') by using the method previously described (37), followed by digestion of the amplicon with *BlnI* and *XhoI* (NEB). The resulting fragments, which included the CGG/AGG repeat element and a majority of the *FMRI* 5'UTR, were inserted into *BlnI*/*XhoI*- linearized pBR-SP6-*FMRI*5'UTR(99CGG)-*FMRI*exon1-FL, creating pBR-SP6-*FMRI*5'UTR[65–66CGG/(AGG)_{0–2}]-*FMRI*exon1-FL-polyA, designated pSP6-*FMRI*(*n*AGG)-FL. The resulting plasmid constructs were confirmed both by sequencing and by *EciI* digestion (NEB) of the amplicon produced by PCR of the cloned region using standard *FMRI* primers c and f (36); *EciI* cleavage occurs 12 base pairs downstream of the A position in AGG interruptions.

In addition to these three premutation-length, SP6-promoter FL plasmids, an analogous construct with 30 CGG repeats [designated pSP6-*FMRI*(30CGG)-FL] was designed by *XhoI*, *BlnI* digestion of pSP6-*FMRI* (65CGG)-FL, to remove the premutation-length repeat, followed by insertion of a similar digest containing 30 CGGs. The AGG interruption pattern of this mRNA is as follows: (CGG)₁₀(AGG)(CGG)₉(AGG)(CGG)₉.

An *FMRI* mRNA with a 5' hairpin, HP-*FMRI* (30CGG)-FL, was created by inserting the following sequence between the SP6 promoter and the *FMRI* 5'UTR of an *FMRI* 5'UTR(30CGG)-FL-polyA construct: 5'-GAACATTTGTAGGGGCGCGTGGTGGCGG CTGCAGCCGCCACCACGCGCCCC-3'. The italicized portion is a 40-base hairpin, which lies 12 bases from

the 5'-end of the resulting mRNA. *In vitro* transcription and translation of this mRNA, as well as *FMRI*(30CGG)-FL, were as per *FMRI*(*n*AGG)-FL mRNAs.

Bacterial maintenance

Top 10 *Escherichia coli* cells (Invitrogen Corp; Carlsbad, CA) were maintained in LB containing 50 µg/ml ampicillin. For minipreps, liquid cultures were grown in 4 ml LB overnight at 33°C; the reduced temperatures were found to improve stability of the CGG repeat during cloning.

In vitro translation of FL mRNAs

Each pSP6-*FMRI*(*n*AGG)-FL, pSP6-*FMRI*(30CGG)-FL and pSP6-HP-*FMRI*(30CGG)-FL plasmid was linearized downstream of the polyA tail with *EcoRI* (NEB), followed by recovery of the linearized plasmids using the Qiagen MinElute reaction cleanup kit (Qiagen NV; Valencia, CA). 5'-capped messages were obtained by *in vitro* transcription using the mMachine SP6 kit (Ambion Inc, Foster City, CA) according to the manufacturer's protocol (2 h, 37°C), followed by RQ1 DNase digestion (Promega Corp., Madison, WI). The resulting mRNAs were cleaned with RNeasy kit (Qiagen) and quantified with a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, DE). For each *in vitro* reaction, 0.4 pmol of each FL reporter was translated in 50 µl of nuclease-treated rabbit reticulocyte lysate (RRL) reaction mix (Promega) for 20 min at 30°C. Reactions were terminated by adding 12.5 µl of Passive Lysis Buffer (Promega).

Luciferase measurements of FL and RL protein levels *in vitro* and *in vivo*

FL and *Renilla* luciferase (RL) reporter control, pRL-CMV (Promega), activities were measured using the Dual-Luciferase Reporter Assay System according to the manufacturer's protocol (Promega). Briefly, 10 µl of the *in vitro* translation reaction or 5 µl of whole cell lysate was pipetted into each well in a 96-well format, followed by the addition of 100 µl of luciferase assay reagent. After a 2 s delay, FL luminescence was read for 5 s. Lysates from *in vivo* reactions were followed by the addition of 100 µl Stop-and-Glo and a further 2 s delay, and finally a 5 s luminescence reading of RL activity. An Lmax luminometer (Molecular Devices Corp, Sunnyvale, CA) running SOFTmax PRO software automated the procedure. Relative FL luminescence was achieved by dividing the measured FL luminescence by its corresponding RL value. Three aliquots of each *in vitro* or cell-based translation reaction were used for each set of conditions.

AGG-interruption constructs for *in vivo* (cell culture) studies

In converting SP6 to CMV promoter-based plasmids, the CMV promoter of pRL-CMV (Promega) was first inserted into a plasmid with an opposite-orientation Ori (opposite to the coding region sense strand), followed by transfer of each *FMRI*(*n*AGG)-FL cassette into this invOri plasmid; both transfers required intermediate linker plasmids. First, the Tet gene of pBR322 (Invitrogen) was removed by

HindIII/PvuII (NEB) double digestion, followed by insertion of the linker, 5'-AGCTTAGATCTTGATCAGGATCCCAG-3', which contains BglII and BamHI sites, to create pBR-Linker. Second, the BamHI and BglII fragment of pRL-CMV, which contains the CMV promoter, was inserted into the corresponding sites in pBR-Linker, resulting in pBR-CMV. pBR-CMV was then digested with SacI and PstI (near the 3'-end of the CMV promoter) followed by insertion of the linker, 5'-CGTTTAGTGAACCGTCAGATCAGTCAGGCGCTCAGCCTGCA-3', which restores the 3'-end of CMV promoter and the first few bases of the *FMRI* 5'UTR (to the BlnI site), creating pBR-CMV-Linker. This plasmid was then digested with BlnI and XbaI to remove the entire coding region. Each pSP6-*FMRI*(nAGG)-FL plasmid was likewise digested with BlnI and XbaI, and the fragment containing *FMRI* 5'UTR-*FMRI*exon1-FL was inserted into pBR-CMV-Linker, generating pBR-CMV-*FMRI*-5'UTR[65-66CGG/(AGG)₀₋₂]-*FMRI*exon1-FL, designated pCMV-*FMRI*(nAGG)-FL. The *FMRI* 5'UTR begins directly after the CMV promoter, eight bases after the transcription start site at base 741 of pRL-CMV (38). All constructs were confirmed by sequencing.

Mammalian cell plasmid transfection and RNA/protein extraction

Seven hundred and fifty-thousand HEK-293 cells (*American Type Culture Collection*, ATCC; CRL-1573) were seeded into each well of six-well plates in DMEM media (Gibco, Grand Island, NY) supplemented with 10% FBS (JR Scientific, Woodland, CA). Cells were grown without antibiotics in 5% CO₂. After 24 h, cells were transfected with 1 μg pCMV-*FMRI*(nAGG)-FL and 100 ng of pRL-CMV DNA, using 4 μl Lipofectamine 2000 (Invitrogen) in 500 μl Opti-MEM (Gibco/Invitrogen). After an additional 24 h, cells were lysed with two separate protocols to obtain RNA and protein. For protein, cells were washed once in PBS (Gibco) and aspirated, followed by addition of 250 μl of Passive Lysis Buffer (Promega) to each well. Plates were rocked at room temperature for 15 min and then frozen at -80°C for 1 h, followed by dual luciferase measurement. For RNA extractions, 350 μl RNeasy Lysis Buffer (Qiagen) was added to each aspirated well followed by 15 min rocking at room temperature. Plates were then frozen at -80°C for 1 h, followed by RNeasy column purification according to manufacturer's protocol (Qiagen).

mRNA reverse transcription

Purified RNAs were quantified on a NanoDrop 1000. Each RNA sample (10-15 μg) was then digested with ScaI (12 U) and BsrFI (12 U) for 45 min (to further reduce residual DNA contamination) followed by RQ1 DNase (1 U) (Promega) for 20 min. Digestions were cleaned using the RNeasy reaction cleanup protocol. Three concentrations of each RNA (500, 250 and 125 ng/100 μl reaction) were reverse transcribed as described earlier (5), including minus-reverse-transcriptase control reactions at the 500 ng/100 μl dilution.

Real-time PCR

Quantitative PCR reactions were performed using the Taqman Universal PCR Master Mix, No AmpErase UNG (Applied Biosystems, Carlsbad, CA), a pre-mix that contains polymerase, dNTPs and buffer. Each PCR reaction consisted of 5 μl cDNA, 4.3 pmol of each primer, 0.8 pmol probe and 6 μl of pre-mix in a total of 12 μl. Primers and probes are as follows. For the FL amplicon (202 bp): FL primers 5'-TACGCCCTGGTTCCTGGAA C-3', 5'-CTCCGATAAATAACGCGCCC-3'; FL probe 5'-FAM-CAGCCCATATCGTTTCATAGCTTCTGCCA A-3' TAMRA. For the RL amplicon (146 bp): RL primers 5'-TGAGGCACTGGGCAGGTGT-3', 5'-ACTGCGGAC CAGTTATCATCC-3'; RL probe 5'-FAM-TAAGGCT AGAGTACTTAATACGACTCACT-3' TAMRA. Each RT dilution (four dilutions including minus RT reactions) was PCR-amplified twice with each primer/probe set, making a total of 16 reactions per mRNA from a single transfection. Real-time PCR was performed in an Applied Biosystems 7900HT (Foster City, CA) with the following parameters: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 s, 60°C for 1 min. FL/RL mRNA ratios were determined for all initial RNA dilutions using the following formula (5):

$$\log_2(FL/RL) = C_T(RL) - C_T(FL),$$

where C_T values are the cycle numbers at which fluorescence levels (FL or RL probes) reach a defined threshold value.

Statistics

Standard errors for the FL/RL ratios were determined for each initial RNA sample from the replicate RT-PCR measurements at each concentration and for the dilution series, according to Tassone *et al.* (5), resulting in a single mean value and standard error. Subsequent statistical analyses were performed using these mean value determinations. Normalized FL values for each translation, protein, and mRNA measurement were additionally normalized to 1 by dividing each value by the mean of all measurements in an experiment, thus setting the average value to 1. Unpaired *t*-tests between each AGG-interruption pair were performed for each analysis. In addition, a one-way, independent analysis of variance (ANOVA) was performed between all three groups (0, 1 and 2 AGG) for each translation efficiency experiment.

RESULTS

Construction of AGG-interruption plasmids

For normal *FMRI* alleles (<~40 CGG repeats), the CGG-repeat element is typically interrupted twice, with spacings of 9-11 repeats, by a C-to-A transversion (AGG interruption). However, for repeat expansions in the premutation range, alleles are found that possess 0, 1 or 2 AGG interruptions (10,27,28,31,39,40). Thus, both for the purpose of examining repeats within the premutation range, which is associated with various forms of clinical involvement as well as reductions in

translational efficiency, and for the purpose of generating approximately equal CGG-repeat lengths with varying number of AGG interruptions, alleles in the ~60-repeat range were pre-selected for cloning based on varying number of AGG interruptions. Selected alleles were PCR-amplified and subsequently cloned into the 5'UTR of a FL reporter plasmid with an SP6 promoter driving an *FMRI* 5'UTR-*FMRI*exon 1-FL expression cassette and a poly A tail (Figure 1A). This procedure generated a series of plasmids, pSP6-*FMRI*(0–2AGG)-FL for *in vitro* studies to test the effect of the AGG interruptions on translational efficiency.

The *FMRI* 5'UTR luciferase reporter system allows for the rapid measurement of the production of protein product (luciferase), reflecting translational initiation, the rate-limiting step of eukaryotic translation (41,42). Since the efficiency of initiation is not strictly determined by the 5'UTR of a message, possibly also involving the N-terminal portion of the coding region, the AGG constructs incorporated the rest of the exon 1 of the *FMRI* gene, for an additional 51 nucleotides. The resulting protein is a chimera of the first 17 amino acids of FMRP and FL. No loss in FL activity was found following addition of the extra N-terminal residues (data not shown).

Upon sequencing, the plasmids were found to have pre-mutation expansions of either 65 or 66 total repeats (CGG + AGG) and with the following interruption patterns: (CGG)₈AGG(CGG)₉AGG(CGG)₄₆, (CGG)₉AGG(CGG)₅₅ and (CGG)₆₆. The AGG-interruption pattern of each DNA was confirmed with *Eci*I restriction digestion of PCR products, which cuts 12 bases downstream of the A in the sequence GGCGGA (Figure 1B), thus digesting the DNA amplifications 0, 1 or 2 times.

The AGG interruptions have no effect on translational efficiency of the pre-mutation CGG repeat *FMRI* 5'UTR *in vitro*

Following production of the *FMRI*(*n*AGG)-FL and *FMRI*(30CGG)-FL mRNAs by *in vitro* transcription,

equimolar amounts of capped mRNAs were translated in a cell-free, RRL system. A series of timed translation reactions were performed with the RRL to determine its linear range for production of protein (data not shown), and care was taken to translate the messages within the linear (time) range of the lysate. Accordingly, all reactions were run for 20 min at 30°C, followed by quick termination with lysis buffer.

Following *in vitro* translation, FL activities in each reaction were measured by adding FL substrate and measuring the resulting chemiluminescence. The readout (relative light units, RLUs) is an arbitrary unit for reporting light integrated over time. In replicate experiments, the means of translation efficiencies of *FMRI*-FL mRNAs with differing numbers of AGG interruptions did not differ from one another (*t*-test: 0 versus 1 AGG, *P* = 0.31; 0 versus 2 AGG, *P* = 0.84; 1 versus 2 AGG, *P* = 0.39; one-way ANOVA: *P* = 0.51, *F* = 0.73) (Figure 2). However, an increase in repeat length from 30 to 65 CGGs results in a moderate, ~25% decrease in FL activity from translations of the same molar amount of message (65–66CGG/(AGG)_{0–2} mRNAs versus *FMRI*(30CGG)-FL; each *P* < 0.005). All experiments were performed at least in triplicate, with sets of three measurements for each replicate.

AGG interruptions do not alter translation efficiency of the pre-mutation CGG-repeat alleles in HEK-293 cells

One concern with the RRL system is that it represents an *in vitro* assay in which the reaction mixture is enriched with various translation factors. Thus, the absence of any dependence of translation efficiency on the number of AGG repeats might reflect the specific features of the RRL milieu. Therefore, to further test the effect of AGG interruptions on translation initiation of the expanded CGG repeat in the context of the *FMRI* 5'UTR, the CGG-repeat-containing constructs were moved into CMV-promoter polyA-signal plasmids, followed by their transfection into human embryonic kidney (HEK-293) cells. In addition to the *FMRI*(*n*AGG)-FL plasmids,

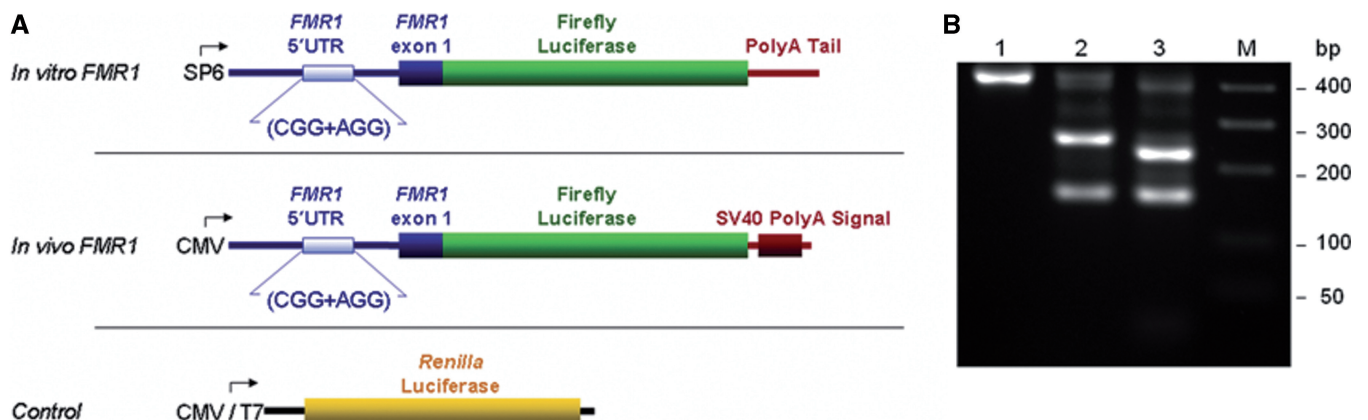


Figure 1. (A) FL and RL constructs. CGG repeats are located within *FMRI* 5'UTRs and have the following repeats: 63 CGG + 2 AGG; 64 CGG + 1 AGG; 66 CGG + 0 AGG. Promoters are indicated by arrows. (B) *Eci* I digests of *FMRI*(*n*AGG)-FL DNAs after PCR amplification of *FMRI* 5'UTRs with primers c and f. Lane 1: *FMRI*(0AGG), lane 2: *FMRI*(1AGG), lane 3: *FMRI*(2AGG), lane 4: DNA marker. Note the 30-base pair band in *FMRI*(2AGG)-FL, the result of two AGG interruptions 10 repeats apart, is not visible in this image.

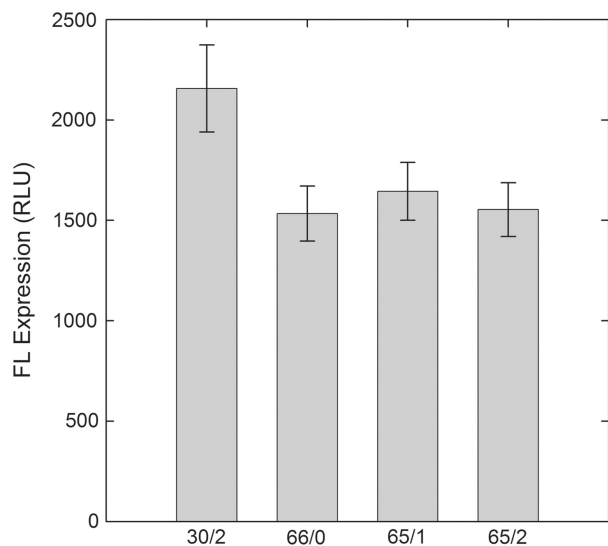


Figure 2. FL values of *in vitro* translations of equimolar *FMRI*(30CGG)-FL and *FMRI*(*n*AGG)-FL mRNAs. Since protein measurements are the result of translating equal molar amounts of mRNAs, they are also *in vitro* translation efficiencies of *FMRI* mRNAs with 0, 1 or 2 AGG interruptions in 65–66 total repeats (designated 66/0, 65/1, 65/2, respectively) and 2 AGGs within 30 total repeats (designated 30/2). Error bars are standard deviations of four replicate experiments.

a RL transfection control was added at a 10:1 ratio of FL:RL, which allows for normalization of FL protein product in a dual luciferase experiment. Twenty-four hours after transfection, the cells were lysed and analyzed for both protein and mRNA levels of FL and RL.

The transfected HEK-293 cells were lysed with passive lysis buffer and FL and RL activities measured. FL values were normalized to the transfection control RL (Figure 3A). To establish the efficiency of translation (i.e. protein produced per unit of mRNA), FL and RL mRNA levels were also determined from the transfected HEK-293 cells. Relative FL mRNA levels for each AGG interruption are shown in Figure 3B. Finally, the relative translation efficiencies were determined (FL protein/relative FL mRNA), as displayed in Figure 3C. As with the *in vitro* studies, no significant differences were found in translation efficiencies among the *FMRI* mRNAs with differing AGG interruptions (*t*-test: 0 versus 1 AGG, $P = 0.14$; 0 versus 2 AGG, $P = 0.40$; 1 versus 2 AGG, $P = 0.62$; one-way ANOVA: $P = 0.30$, $F = 1.25$). Thus, we find that the number of AGG interruptions (or their absence) within a given CGG-repeat length has no appreciable effect on the efficiency of translation, either *in vitro* or *in vivo*.

Translational initiation proceeds from the 5'-end of the reporter mRNAs

The above translation experiments were performed under the assumption that translation initiation of *FMRI* occurs via ribosomal scanning, in which the 43S ribosome binds as a ribonucleoprotein complex at the 5' cap of a message

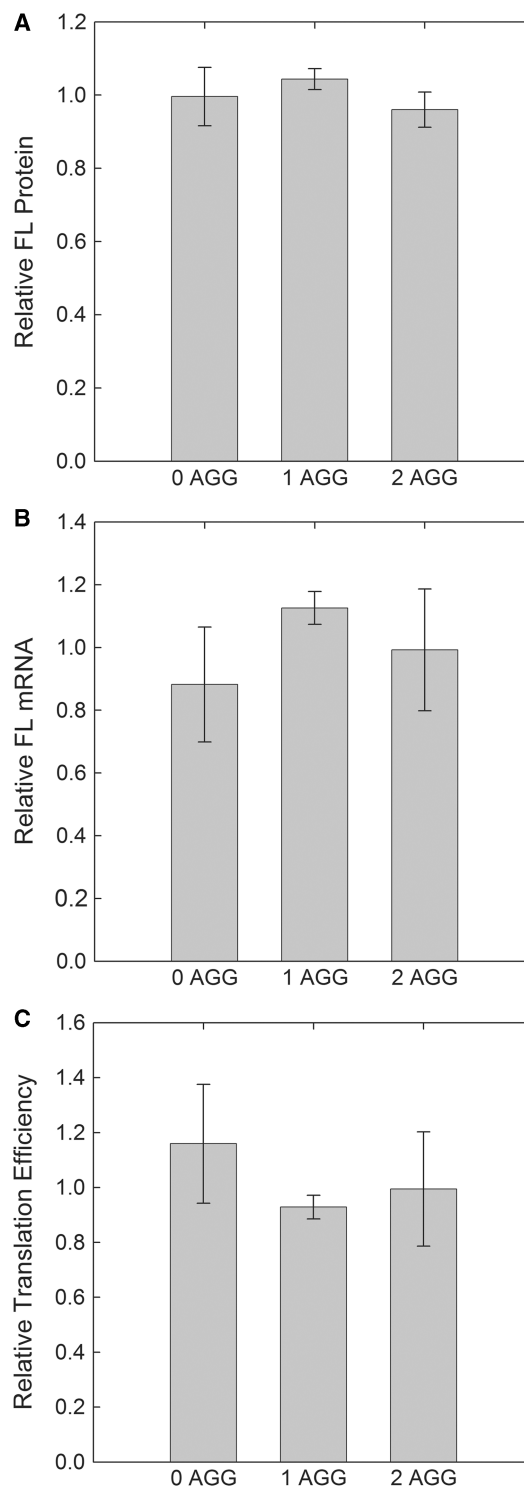


Figure 3. HEK-293 cell transfections of pRL-CMV and either pCMV-*FMRI*(0AGG)-FL, pCMV-*FMRI*(1AGG)-FL or pCMV-*FMRI*(2AGG)-FL. (A) Normalized FL protein values. Error bars are standard deviations of six replicate experiments. (B) Normalized FL mRNA values. Error bars are standard deviations of three replicate experiments. (C) Normalized FL protein:mRNA ratios. Error bars are standard deviations of three replicate experiments.

followed by movement along the 5'UTR until an AUG start codon is recognized. As a means of testing the mechanism of translation initiation, we designed an *FMRI* construct that would, by insertion of a perfect hairpin near the 5'-end of the mRNA (Figure 4A), inhibit scanning of the 5'UTR after cap binding. A strong, 40-base hairpin was added as an extension onto the 5'-end of an *FMRI*(30CGG)-FL mRNA (designated HP) so as not to disturb the *FMRI* sequence or secondary structure, which could be involved in alternative initiation mechanisms. The hairpin completely blocked *in vitro* translation of capped *FMRI* message in RRL (Figure 4B) [$P < 0.00001$ between HP and *FMRI*(30CGG)-FL mRNAs]. These data argue against internal ribosome entry site (IRES) initiated translation, whereby the ribosome binds the message at or near the AUG start codon without cap recognition or scanning. We cannot, however, rule out

ribosomal shunting, which also requires cap recognition, at this time (43,44).

DISCUSSION

CGG-repeat tracts in RNA are capable of forming stable hairpin structures (33,34,45,46) and tetraplex structures (35,47). Although the presence of such structures has not been established *in vivo*, it has been proposed that, due to the predicted stability of such structures in the 5'UTR of the *FMRI* message, their expansion into the premutation or full mutation ranges would likely impede translational initiation (33,34,47). Indeed, *FMRI* mRNAs with longer CGG repeats were found to have decreased translation efficiency, as well as reduced association with polysomes (3,4,7,8,35).

The prediction that CGG-repeat secondary structure impedes translation of *FMRI* is based on the assumption that initiation of *FMRI* translation occurs by ribosomal scanning from the 5'-end of the message, and not by an IRES mechanism. The latter, previously thought to occur as a result of *FMRI* internal initiation within dicistronic mRNAs (48,49) was effectively ruled out by using an *FMRI*-FL mRNA to which was added a 5' hairpin extension upstream of the native UTR, thereby avoiding disrupting the *FMRI* 5'UTR sequence. The 5'-blocked construct translates at <1% of *FMRI*(30CGG)-FL mRNA *in vitro*, thus arguing against an IRES-type initiation mechanism (Figure 4). Although shunt-type mechanisms (43,44) cannot be formally ruled out, the absence of an upstream open reading frame (ORF) renders such a possibility unlikely.

One noteworthy feature of the transition into the premutation range is the gradual loss of AGG interruptions in the 5'UTR region of the *FMRI* mRNA. Studies *in vitro* and *in silico* have raised the possibility that the presence of such interruptions would destabilize otherwise long CGG-repeat hairpin structures (33,34,47). By nuclease cleavage protection studies of *FMRI* 5'UTR mRNA, Napierala *et al.* (34) have found that AGG substitutions substantially modified the CGG-repeat structure, causing the formation of branch points in the CGG hairpin stem region and conversion to multiple, shorter hairpins. However, in a recent NMR study (46), we found that an AGG substitution did not appreciably change the structure of an r(CG) hairpin, and also saw little change in the melting temperature of a CGG-repeat hairpin with an AGG substitution at physiological magnesium. Moreover, in contrast to the Napierala *et al.* study, the NMR results clearly demonstrated the presence of A-G base-pairing.

If AGG interruptions were able to disrupt r(CG) secondary structures, one would expect the efficiency of *FMRI* translation to increase with increasing number of AGG interruptions; however, this idea has never been tested, either *in vitro* or *in vivo*. In the current work, we have tested this idea, both *in vitro* (RRL translation system) and *in vivo* (HEK-293 cell culture), and there was no evidence of altered translational efficiency in either system. The implications of the current observations

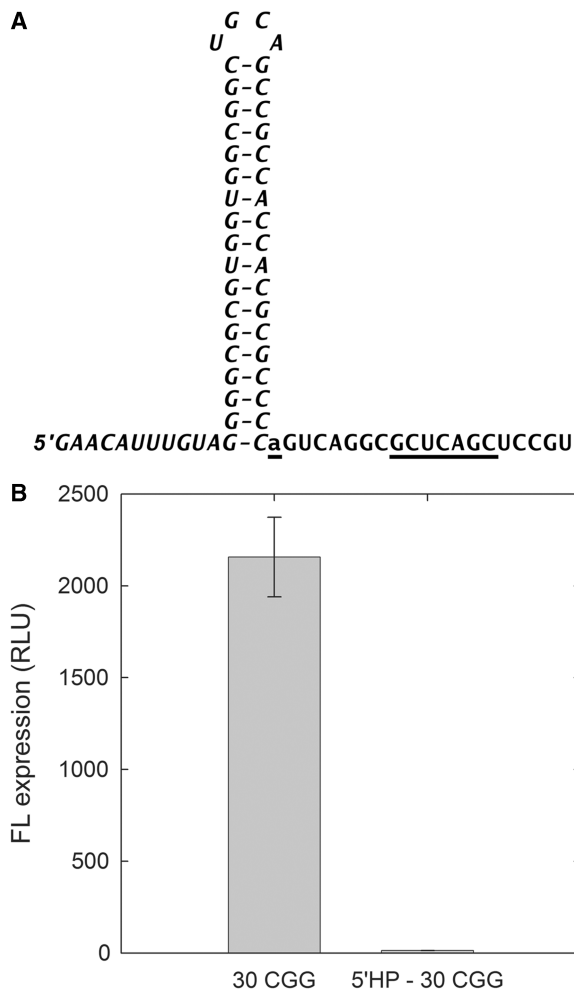


Figure 4. (A) The 5' hairpin lies within a sequence extension (italicized font) added to the 5'-end of the *FMRI* 5'UTR (normal font, begins at underlined, lower-case a), 12 bases from the 5'-end of the message. Shown are the hairpin, 11 bases upstream and 20 bases of the *FMRI* 5'UTR. The BspI site in the *FMRI* 5'UTR is underlined. (B) FL expression, in RLU, of *FMRI*(30CGG)-FL and HP-*FMRI*(30CGG)-FL mRNAs. Error bars are standard deviations of at least three replicate experiments.

are 2-fold. First, our results indicate that, were AGG repeats capable of modulating the overall structure of the CGG-repeat region, such structural transitions would not be affecting the rate-limiting steps in translation initiation, even for 5'-end-dependent initiation/scanning. Second, from the perspective of the energetics of the CGG repeat, the magnitude of the perturbations introduced by the AGG (actually, the single C-to-A transversions) are likely to be much less than had been proposed on the basis of the cleavage/protection experiments (34), since those authors saw no evidence for A-G base pair formation. Such interactions are well-known as conserved interactions in structured RNAs, and were in fact observed in the NMR studies (46). While the secondary structures of *FMRI* mRNAs with differing CGG-repeat lengths and AGG-interruption patterns may have important implications for certain biological functions, the aim of this study was to determine whether there is a penalty to translational efficiency with the loss of the AGG repeats; we find that there is no significant translation penalty for loss of AGGs either *in vitro* or *in vivo* under conditions of CGG-repeat expansion (~65 CGG repeats) where alleles with 0, 1 or 2 AGG repeats coexist in the general population.

The results of the current study raise the broader question of the functional role of the AGG interruptions in the *FMRI* 5'UTR. Deoxytrinucleotide d(CGG) repeats have also been shown to form stable secondary structure in both hairpin (50–54) and tetraplex forms (9,55–57). Such secondary structures have been implicated in CGG-repeat instability/expansions during DNA replication, possibly through strand slippage or polymerase stalling at the repeat (57–60). Since longer CGG repeats are associated with fewer (or absence of) AGG interruptions, such interruptions are thought by many to represent repeat-stabilizing 'anchors' (28,30,31,40,61). Although AGGs in principle could protect against repeat expansion by disrupting higher-order CGG structures, there is no evidence for significant repeat instability in differentiated somatic cells. Thus, whatever effect, if any, the AGGs might exert on repeat stability would necessarily be limited to the initial two to three week period in embryogenesis during which repeat size instability appears to be manifest. In addition, it is known that AGGs do not appear to have any effect on the rate of transcription of the *FMRI* gene (10).

The CGG-repeat lengths used in this study are within the premutation range, in which translation efficiencies have been found to be impeded *in vitro* (35), as well as in plasmid-transfected SK and HEK-293 cells (8) and patient-derived fibroblast and lymphoblastoid cells (3,4,7). To ensure that the mid-length CGG repeats used in this study do not simply block FL reporter translation of each AGG mRNA equally, we also included a normal CGG-repeat mRNA, *FMRI*(30CGG)-FL, in RRL experiments. This message is analogous to the premutation-length AGG mRNAs; however it has the mode human *FMRI* CGG-repeat length (30 CGGs). As can be seen in Figure 2, an increase in repeat number from 30 to 65 repeats leads to only a 25% reduction of total FL produced *in vitro*. We conclude that the moderate decrease

in translation efficiency seen for longer-repeat mRNAs is not sufficient to warrant concern. Indeed, this reduction roughly corresponds with previous studies (7,8) that found 20–50% decreases in translation efficiency of *FMRI* mRNAs between the normal and premutation CGG-repeat ranges.

Finally, we have noticed differing stability of the CGG repeat within *FMRI* 5'UTR plasmids for differing number of AGGs when propagated in *E. coli*. It is known that longer tandem repeats will contract or delete from plasmids in bacterial culture (8,60). However, the plasmids used in this study had similar numbers of CGG repeats but the repeats clearly deleted at different rates. After screening 35 plasmid minipreps of each AGG-interruption construct, we found that 0-AGG plasmids deleted significantly more often (54% deleted) than both 1- (22%) and 2-AGG (11%) interruption plasmids (0 versus 1 AGG, $P < 0.01$; 0 versus 2 AGG, $P < 0.005$; testing for the significance of the difference between two independent proportions) (data not shown). Plasmids having 1 AGG were not significantly different from those with 2 (1 versus 2 AGG, $P = 0.1$). No repeat expansions were found. Thus, although our studies found no difference in the translation efficiency of CGG-repeat mRNAs with or without AGG interruptions, it is apparent that having at least one AGG interruption offers plasmids some protection (in *E. coli*) against CGG-repeat deletion.

ACKNOWLEDGEMENTS

The authors wish to thank Wenting Zhang (M.I.N.D. Institute) for isolating genomic DNAs for cloning.

FUNDING

National Institute of Child Health and Development [R01 HD040661 to P.J.H.; R01HD055510 to F.T.]. Funding for open access charge: R01 HD040661.

Conflict of interest statement. None declared.

REFERENCES

- Verkerk,A.J., Pieretti,M., Sutcliffe,J.S., Fu,Y.H., Kuhl,D.P., Pizzuti,A., Reiner,O., Richards,S., Victoria,M.F., Zhang,F.P. *et al.* (1991) Identification of a gene (*FMRI*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, **65**, 905–914.
- Oberle,I., Rousseau,F., Heitz,D., Kretz,C., Devys,D., Hanauer,A., Boue,J., Bertheas,M. and Mandel,J. (1991) Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science*, **252**, 1097–1102.
- Primerano,B., Tassone,F., Hagerman,R.J., Hagerman,P., Amaldi,F. and Bagni,C. (2002) Reduced *FMRI* mRNA translation efficiency in fragile X patients with premutations. *RNA*, **8**, 1482–1488.
- Feng,Y., Zhang,F., Lokey,L.K., Chastain,J.L., Lakkis,L., Eberhart,D. and Warren,S.T. (1995) Translational suppression by trinucleotide repeat expansion at *FMRI*. *Science*, **268**, 731–734.
- Tassone,F., Hagerman,R.J., Taylor,A.K., Gane,L.W., Godfrey,T.E. and Hagerman,P.J. (2000) Elevated levels of *FMRI* mRNA in carrier males: a new mechanism of involvement in the fragile-X syndrome. *Am. J. Hum. Genet.*, **66**, 6–15.

6. Tassone, F., Hagerman, R.J., Chamberlain, W.D. and Hagerman, P.J. (2000) Transcription of the *FMRI* gene in individuals with fragile X syndrome. *Am. J. Med. Genet.*, **97**, 195–203.
7. Kenneson, A., Zhang, F., Hagedorn, C.H. and Warren, S.T. (2001) Reduced FMRP and increased *FMRI* transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum. Mol. Genet.*, **10**, 1449–1454.
8. Chen, L.S., Tassone, F., Sahota, P. and Hagerman, P.J. (2003) The (CGG)_n repeat element within the 5' untranslated region of the *FMRI* message provides both positive and negative cis effects on in vivo translation of a downstream reporter. *Hum. Mol. Genet.*, **12**, 3067–3074.
9. Lenzmeier, B.A. and Freudenreich, C.H. (2003) Trinucleotide repeat instability: a hairpin curve at the crossroads of replication, recombination, and repair. *Cytogenet. Genome Res.*, **100**, 7–24.
10. Tassone, F., Beilina, A., Carosi, C., Albertosi, S., Bagni, C., Li, L., Glover, K., Bentley, D. and Hagerman, P.J. (2007) Elevated *FMRI* mRNA in premutation carriers is due to increased transcription. *RNA*, **13**, 555–562.
11. Hagerman, R.J. (2006) Lessons from fragile X regarding neurobiology, autism, and neurodegeneration. *J. Dev. Behav. Pediatr.*, **27**, 63–74.
12. Bassell, G.J. and Warren, S.T. (2008) Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron*, **60**, 201–214.
13. Allen, E.G., He, W., Yadav-Shah, M. and Sherman, S.L. (2004) A study of the distributional characteristics of *FMRI* transcript levels in 238 individuals. *Hum. Genet.*, **114**, 439–447.
14. Hessler, D., Tassone, F., Loesch, D.Z., Berry-Kravis, E., Leehey, M.A., Gane, L.W., Barbato, I., Rice, C., Gould, E., Hall, D.A. *et al.* (2005) Abnormal elevation of *FMRI* mRNA is associated with psychological symptoms in individuals with the fragile X premutation. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **139B**, 115–121.
15. Cohen, S., Masyn, K., Adams, J., Hessler, D., Rivera, S., Tassone, F., Brunberg, J., DeCarli, C., Zhang, L., Cogswell, J. *et al.* (2006) Molecular and imaging correlates of the fragile X-associated tremor/ataxia syndrome. *Neurology*, **67**, 1426–1431.
16. Hagerman, R.J., Leehey, M., Heinrichs, W., Tassone, F., Wilson, R., Hills, J., Grigsby, J., Gage, B. and Hagerman, P.J. (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology*, **57**, 127–130.
17. Jacquemont, S., Hagerman, R.J., Leehey, M., Grigsby, J., Zhang, L., Brunberg, J.A., Greco, C., Des Portes, V., Jardini, T., Levine, R. *et al.* (2003) Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. *Am. J. Hum. Genet.*, **72**, 869–878.
18. Jacquemont, S., Hagerman, R.J., Leehey, M.A., Hall, D.A., Levine, R.A., Brunberg, J.A., Zhang, L., Jardini, T., Gane, L.W., Harris, S.W. *et al.* (2004) Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *JAMA*, **291**, 460–469.
19. Arocena, D.G., Iwahashi, C.K., Won, N., Beilina, A., Ludwig, A.L., Tassone, F., Schwartz, P.H. and Hagerman, P.J. (2005) Induction of inclusion formation and disruption of lamin A/C structure by premutation CGG-repeat RNA in human cultured neural cells. *Hum. Mol. Genet.*, **14**, 3661–3671.
20. Handa, V., Goldwater, D., Stiles, D., Cam, M., Poy, G., Kumari, D. and Usdin, K. (2005) Long CGG-repeat tracts are toxic to human cells: implications for carriers of Fragile X premutation alleles. *FEBS Lett.*, **579**, 2702–2708.
21. Berry-Kravis, E., Goetz, C.G., Leehey, M.A., Hagerman, R.J., Zhang, L., Li, L., Nguyen, D., Hall, D.A., Tartaglia, N., Cogswell, J. *et al.* (2007) Neuropathic features in fragile X premutation carriers. *Am. J. Med. Genet. A*, **143**, 19–26.
22. Amiri, K., Hagerman, R.J. and Hagerman, P.J. (2008) Fragile X-associated tremor/ataxia syndrome: an aging face of the fragile X gene. *Arch. Neurol.*, **65**, 19–25.
23. Brouwer, J.R., Willemsen, R. and Oostra, B.A. (2009) The *FMRI* gene and fragile X-associated tremor/ataxia syndrome. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **150B**, 782–798.
24. Greco, C.M., Hagerman, R.J., Tassone, F., Chudley, A.E., Del Bigio, M.R., Jacquemont, S., Leehey, M. and Hagerman, P.J. (2002) Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain*, **125**, 1760–1771.
25. Greco, C.M., Berman, R.F., Martin, R.M., Tassone, F., Schwartz, P.H., Chang, A., Trapp, B.D., Iwahashi, C., Brunberg, J., Grigsby, J. *et al.* (2006) Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS). *Brain*, **129**, 243–255.
26. Tassone, F., Iwahashi, C. and Hagerman, P.J. (2004) *FMRI* RNA within the intranuclear inclusions of fragile X-associated tremor/ataxia syndrome (FXTAS). *RNA Biol.*, **1**, 103–105.
27. Eichler, E.E., Macpherson, J.N., Murray, A., Jacobs, P.A., Chakravarti, A. and Nelson, D.L. (1996) Haplotype and interspersed analysis of the *FMRI* CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. *Hum. Mol. Genet.*, **5**, 319–330.
28. Eichler, E.E., Holden, J.J., Popovich, B.W., Reiss, A.L., Snow, K., Thibodeau, S.N., Richards, C.S., Ward, P.A. and Nelson, D.L. (1994) Length of uninterrupted CGG repeats determines instability in the *FMRI* gene. *Nat. Genet.*, **8**, 88–94.
29. Snow, K., Tester, D.J., Kruckeberg, K.E., Schaid, D.J. and Thibodeau, S.N. (1994) Sequence analysis of the fragile X trinucleotide repeat: implications for the origin of the fragile X mutation. *Hum. Mol. Genet.*, **3**, 1543–1551.
30. Zhong, N., Yang, W., Dobkin, C. and Brown, W.T. (1995) Fragile X gene instability: anchoring AGGs and linked microsatellites. *Am. J. Hum. Genet.*, **57**, 351–361.
31. Dombrowski, C., Levesque, S., Morel, M.L., Rouillard, P., Morgan, K. and Rousseau, F. (2002) Premutation and intermediate-size *FMRI* alleles in 10572 males from the general population: loss of an AGG interruption is a late event in the generation of fragile X syndrome alleles. *Hum. Mol. Genet.*, **11**, 371–378.
32. Nolin, S.L., Brown, W.T., Glicksman, A., Houck, G.E. Jr, Gargano, A.D., Sullivan, A., Biancalana, V., Brondum-Nielsen, K., Hjalgrim, H., Holinski-Feder, E. *et al.* (2003) Expansion of the fragile X CGG repeat in females with premutation or intermediate alleles. *Am. J. Hum. Genet.*, **72**, 454–464.
33. Handa, V., Saha, T. and Usdin, K. (2003) The fragile X syndrome repeats form RNA hairpins that do not activate the interferon-inducible protein kinase, PKR, but are cut by Dicer. *Nucleic Acids Res.*, **31**, 6243–6248.
34. Napierala, M., Michalowski, D., de Mezer, M. and Krzyzosiak, W.J. (2005) Facile *FMRI* mRNA structure regulation by interruptions in CGG repeats. *Nucleic Acids Res.*, **33**, 451–463.
35. Khateb, S., Weisman-Shomer, P., Hershco-Shani, I., Ludwig, A.L. and Fry, M. (2007) The tetraplex (CGG)_n destabilizing proteins hnRNP A2 and CBF-A enhance the in vivo translation of fragile X premutation mRNA. *Nucleic Acids Res.*, **35**, 5775–5788.
36. Fu, Y.H., Kuhl, D.P., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J., Holden, J.J., Fenwick, R.G. Jr, Warren, S.T. *et al.* (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell*, **67**, 1047–1058.
37. Saluto, A., Brussino, A., Tassone, F., Arduino, C., Cagnoli, C., Pappi, P., Hagerman, P., Migone, N. and Brusco, A. (2005) An enhanced polymerase chain reaction assay to detect pre- and full mutation alleles of the fragile X mental retardation 1 gene. *J. Mol. Diagn.*, **7**, 605–612.
38. Macias, M.P., Huang, L., Lashmit, P.E. and Stinski, M.F. (1996) Cellular or viral protein binding to a cytomegalovirus promoter transcription initiation site: effects on transcription. *J. Virol.*, **70**, 3628–3635.
39. Hirst, M.C., Grewal, P.K. and Davies, K.E. (1994) Precursor arrays for triplet repeat expansion at the fragile X locus. *Hum. Mol. Genet.*, **3**, 1553–1560.
40. Kunst, C.B. and Warren, S.T. (1994) Cryptic and polar variation of the fragile X repeat could result in predisposing normal alleles. *Cell*, **77**, 853–861.
41. Gebauer, F. and Hentze, M.W. (2004) Molecular mechanisms of translational control. *Nat. Rev. Mol. Cell Biol.*, **5**, 827–835.
42. Sonenberg, N. and Hinnebusch, A.G. (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell*, **136**, 731–745.
43. Mauro, V.P., Chappell, S.A. and Dresios, J. (2007) Analysis of ribosomal shunting during translation initiation in eukaryotic mRNAs. *Methods Enzymol.*, **429**, 323–354.

44. Jackson, R.J. (2000) In Sonenberg, N., Hershey, J.W.B. and Mathews, M. (eds), *Cold Spring Harbor Monograph Series*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 127–183.
45. Sobczak, K., de Mezer, M., Michlewski, G., Krol, J. and Krzyzosiak, W.J. (2003) RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res.*, **31**, 5469–5482.
46. Zumwalt, M., Ludwig, A., Hagerman, P.J. and Dieckmann, T. (2007) Secondary structure and dynamics of the r(CGG) repeat in the mRNA of the fragile X mental retardation 1 (*FMRI*) gene. *RNA Biol.*, **4**, 93–100.
47. Weisman-Shomer, P., Cohen, E. and Fry, M. (2000) Interruption of the fragile X syndrome expanded sequence d(CGG)(n) by interspersed d(AGG) trinucleotides diminishes the formation and stability of d(CGG)(n) tetrahelical structures. *Nucleic Acids Res.*, **28**, 1535–1541.
48. Chiang, P.W., Carpenter, L.E. and Hagerman, P.J. (2001) The 5'-untranslated region of the *FMRI* message facilitates translation by internal ribosome entry. *J. Biol. Chem.*, **276**, 37916–37921.
49. Dobson, T., Kube, E., Timmerman, S. and Krushel, L.A. (2008) Identifying intrinsic and extrinsic determinants that regulate internal initiation of translation mediated by the *FMRI* 5' leader. *BMC Mol. Biol.*, **9**, 89.
50. Nadel, Y., Weisman-Shomer, P. and Fry, M. (1995) The fragile X syndrome single strand d(CGG)n nucleotide repeats readily fold back to form unimolecular hairpin structures. *J. Biol. Chem.*, **270**, 28970–28977.
51. Gacy, A.M., Goellner, G., Juranic, N., Macura, S. and McMurray, C.T. (1995) Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell*, **81**, 533–540.
52. Chen, X., Mariappan, S.V., Catasti, P., Ratliff, R., Moyzis, R.K., Laayoun, A., Smith, S.S., Bradbury, E.M. and Gupta, G. (1995) Hairpins are formed by the single DNA strands of the fragile X triplet repeats: structure and biological implications. *Proc. Natl Acad. Sci. USA*, **92**, 5199–5203.
53. Mitas, M., Yu, A., Dill, J., Kamp, T.J., Chambers, E.J. and Haworth, I.S. (1995) Hairpin properties of single-stranded DNA containing a GC-rich triplet repeat: (CTG)₁₅. *Nucleic Acids Res.*, **23**, 1050–1059.
54. Renciu, D., Zemanek, M., Kejnovska, I. and Vorlickova, M. (2009) Quadruplex-forming properties of FRAXA (CGG) repeats interrupted by (AGG) triplets. *Biochimie*, **91**, 416–422.
55. Fry, M. and Loeb, L.A. (1994) The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure. *Proc. Natl Acad. Sci. USA*, **91**, 4950–4954.
56. Kettani, A., Kumar, R.A. and Patel, D.J. (1995) Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat. *J. Mol. Biol.*, **254**, 638–656.
57. Usdin, K. and Woodford, K.J. (1995) CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Res.*, **23**, 4202–4209.
58. Kamath-Loeb, A.S., Loeb, L.A., Johansson, E., Burgers, P.M. and Fry, M. (2001) Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGG)n trinucleotide repeat sequence. *J. Biol. Chem.*, **276**, 16439–16446.
59. Kang, S., Ohshima, K., Shimizu, M., Amirhaeri, S. and Wells, R.D. (1995) Pausing of DNA synthesis in vitro at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *J. Biol. Chem.*, **270**, 27014–27021.
60. Hirst, M.C. and White, P.J. (1998) Cloned human *FMRI* trinucleotide repeats exhibit a length- and orientation-dependent instability suggestive of in vivo lagging strand secondary structure. *Nucleic Acids Res.*, **26**, 2353–2358.
61. Ashley-Koch, A.E., Robinson, H., Glicksman, A.E., Nolin, S.L., Schwartz, C.E., Brown, W.T., Turner, G. and Sherman, S.L. (1998) Examination of factors associated with instability of the *FMRI* CGG repeat. *Am. J. Hum. Genet.*, **63**, 776–785.