

# Repeat-Associated Non-AUG Translation from Antisense CCG Repeats in Fragile X Tremor/Ataxia Syndrome

Amy Krans, MS,<sup>1,2</sup> Michael G. Kearse, PhD,<sup>1</sup> and Peter K. Todd, MD, PhD<sup>1,2</sup>

**Objective:** Repeat-associated non-AUG (RAN) translation drives production of toxic proteins from pathogenic repeat sequences in multiple untreatable neurodegenerative disorders. Fragile X-associated tremor/ataxia syndrome (FXTAS) is one such condition, resulting from a CGG trinucleotide repeat expansion in the 5' leader sequence of the *FMR1* gene. RAN proteins from the CGG repeat accumulate in ubiquitinated inclusions in FXTAS patient brains and elicit toxicity. In addition to the CGG repeat, an antisense mRNA containing a CCG repeat is also transcribed from the *FMR1* locus. We evaluated whether this antisense CCG repeat supports RAN translation and contributes to pathology in FXTAS patients.

**Methods:** We generated a series of CCG RAN translation-specific reporters and utilized them to measure RAN translation from CCG repeats in multiple reading frames in transfected cells. We also developed antibodies against predicted CCG RAN proteins and used immunohistochemistry and immunofluorescence on FXTAS patient tissues to measure their accumulation and distribution.

**Results:** RAN translation from CCG repeats is supported in all 3 potential reading frames, generating polyproline, polyarginine, and polyalanine proteins, respectively. Their production occurs whether or not the natural AUG start upstream of the repeat in the proline reading frame is present. All 3 frames show greater translation at larger repeat sizes. Antibodies targeted to the antisense FMR polyproline and polyalanine proteins selectively stain nuclear and cytoplasmic aggregates in FXTAS patients and colocalize with ubiquitinated neuronal inclusions.

**Interpretation:** RAN translation from antisense CCG repeats generates novel proteins that accumulate in ubiquitinated inclusions in FXTAS patients.

ANN NEUROL 2016;80:871–881

Nucleotide repeat expansions are among the most common inherited causes of neurodegeneration and neurological disease.<sup>1</sup> These progressive disorders are currently without any effective treatment. Repeat expansions elicit toxicity through a range of different pathogenic mechanisms, including polyglutamine-mediated aggregation and associated impairment of protein quality control pathways, RNA repeat-mediated sequestration of protein-bound factors, and DNA-mediated transcriptional silencing of the genes in which they reside.<sup>2–4</sup>

Work by a number of groups has recently demonstrated that expanded nucleotide repeats can also support an unusual form of protein translational initiation known as repeat-associated non-AUG (RAN) translation.<sup>5,6</sup> RAN translation allows for production of aberrant homopolymeric or dipeptide repeat proteins in the absence of an AUG start codon. It has now been

described at CAG, CUG, GGGGCC, CCCC GG, and CGG repeats.<sup>5,7–13</sup> RAN proteins can be generated from multiple reading frames of the same repeat, and in disorders where bidirectional transcription through the repeat is present, RAN translation can also occur on the antisense transcript, leading to a series of potentially toxic products that accumulate in patient neurons.<sup>5,6,9,10,12</sup>

Fragile X-associated tremor/ataxia syndrome (FXTAS) is an age-related neurodegenerative disease caused by a CGG trinucleotide repeat expansion in the 5' untranslated region (hereafter referred to as the 5' leader sequence) of the *FMR1* gene.<sup>14</sup> FXTAS is under-recognized clinically, with a predicted incidence of 1:3,000 men older than 50 years.<sup>15</sup> Clinical characteristics include progressive gait ataxia, intention tremor, parkinsonism, and dementia.<sup>16</sup> FXTAS patients and animals models of disease are characterized by elevated *FMR1* mRNA levels, decreased levels of the fragile X

View this article online at [wileyonlinelibrary.com](http://wileyonlinelibrary.com). DOI: 10.1002/ana.24800

Received Aug 23, 2016, and in revised form Oct 6, 2016. Accepted for publication Oct 17, 2016.

Address correspondence to Dr Todd, 4005 BSRB, 109 Zina Pitcher Place, Ann Arbor, MI 48109-2200. E-mail: [petertod@umich.edu](mailto:petertod@umich.edu)

From the <sup>1</sup>Department of Neurology, University of Michigan, Ann Arbor, MI; and <sup>2</sup>VA Medical Center, Ann Arbor, MI

protein, FMRP, and intranuclear ubiquitinated neuronal inclusions in the cerebellum and other brain regions.<sup>14,17,18</sup>

Our group previously demonstrated that RAN translation occurs on CGG repeats in FXTAS to produce homopolymeric polyglycine (FMRpolyG) and polyalanine (FMRpolyA) proteins.<sup>7</sup> FMRpolyG proteins aggregate in cellular, fly, and mouse models of FXTAS and in ubiquitinated inclusions in patient brain tissue, and FMRpolyG production drives CGG repeat toxicity in simple model systems.<sup>7,19</sup> The *FMRI* locus also produces an antisense transcript, *ASFMRI*, that contains an antisense CCG repeat.<sup>20,21</sup> This antisense repeat lies within a putative open reading frame, that if utilized would generate a short polyproline-containing protein, ASFMRP.<sup>20</sup> *ASFMRI* mRNA is ubiquitously expressed in human tissue, with highest expression in the brain, and its expression is elevated in FXTAS patients and models.<sup>20</sup> Because the CCG repeat in the *ASFMRI* transcript is predicted to form a stable secondary structure,<sup>22</sup> we hypothesized that it might support RAN translation and potentially contribute to disease pathogenesis in FXTAS, similar to the CGG repeat in the sense transcript.<sup>23</sup> Here, we provide evidence from cellular models that the CCG repeat can support RAN translation in all 3 reading frames to produce homopolymeric proteins. We further demonstrate that antibodies generated against the predicted proline (ASFMRpolyP) and alanine (ASFMRpolyA) RAN products stain ubiquitinated neuronal inclusions in FXTAS patients. These data suggest a role for CCG repeat RAN translation in neuronal inclusion formation in FXTAS.

## Materials and Methods

### Constructs

The base constructs and cloning strategies used were previously described.<sup>24</sup> Briefly, each frame of *ASFMRI* was cloned into GGG-NL-3xFLAG pcDNA by 2 rounds of Q5 site-directed mutagenesis (New England BioLabs, Ipswich, MA) followed by 2 rounds of annealing primer ligation using XhoI and EcoRV, and NarI and EcoRV restriction sites, respectively. All constructs were verified by Sanger sequencing. Expanded repeats were inserted using XhoI and NarI from FMRpolyG<sub>100</sub> green fluorescent protein (GFP).<sup>7</sup> Repeat sizes were determined by restriction digest. Primer and construct sequences are available upon request.

### Western Blotting

COS-7 cells were transfected with reporter constructs using Lipofectamine LTX with Plus reagent (Thermo Fisher Scientific, Waltham, MA) according to the manufacturer's protocol. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis and Western blotting were done at 24 hours after transfection as described previously.<sup>7,24</sup> For analysis of tagged constructs,

membranes were incubated with mouse monoclonal FLAG-M2 (1:1,000; Sigma, St Louis, MO; F1804) and mouse monoclonal glyceraldehyde-3-phosphate dehydrogenase (1:1,000; Santa Cruz Biotechnology, Santa Cruz, CA; sc32233) overnight at 4°C. Detection was performed using Western Lightning ECL-Plus (PerkinElmer, Waltham, MA) on autoradiography film. AUG-nanoluciferase-3xFLAG (AUG-NL) was loaded at one-tenth the amount of the other reporters to avoid overexposure on the film.

### Luciferase Assay

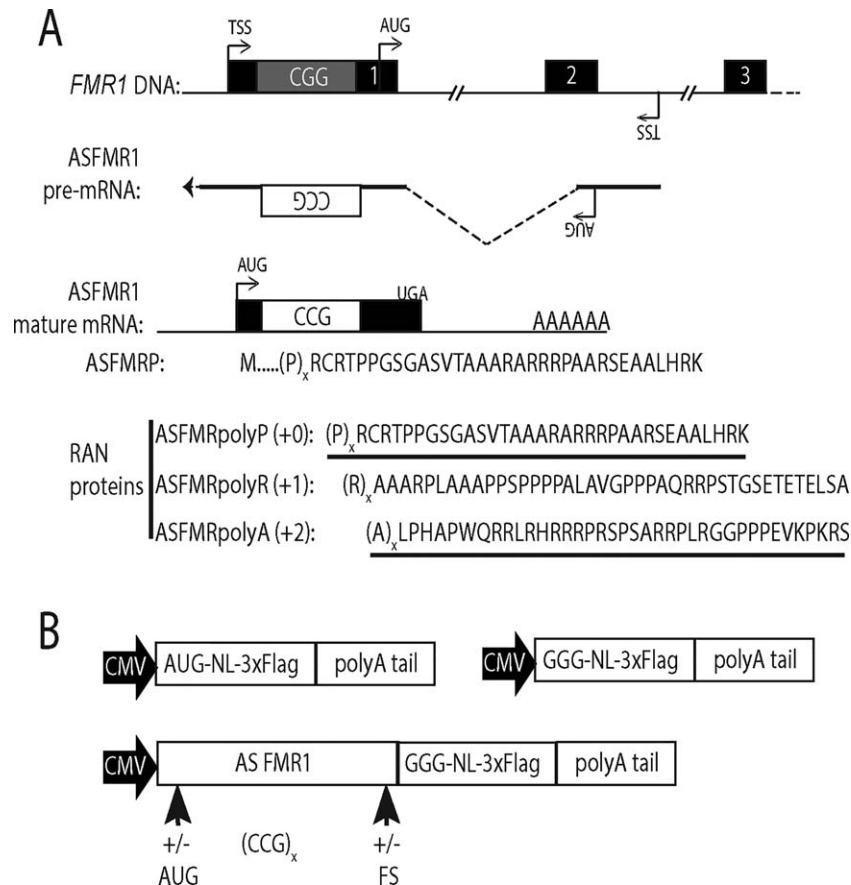
COS-7 cells were plated on a 96-well tissue culture plate and were cotransfected with *ASFMRI* reporter constructs and pGL4.13 (Firefly luciferase; Promega, Madison, WI) using Viatect (Promega) according to the manufacturer's protocol. Luciferase assays were performed at 24 hours after transfection as previously described.<sup>24</sup> Nanoluciferase (NL) signal was normalized to the internal Firefly luciferase control. Each experiment included 3 technical replicates averaged to generate a single "n". All experiments were carried out at least in triplicate. Experimental results were analyzed by 1-way analysis of variance using Prism (GraphPad, San Diego, CA). Significance of the difference between individual constructs was determined by Fisher least significant difference with Bonferroni correction for multiple comparisons.

### Generation of Polyclonal Antibodies

Rabbit polyclonal antibodies were generated by NeoScientific (Cambridge, MA) to synthetic peptides corresponding to the repeat and C-terminal sequence of the predicted proteins (exact epitopes are shown in Fig 1). Antisera were affinity purified using the respective peptide immunogens. Preimmune sera were used as a negative control.

### Immunocytochemistry

COS-7 cells were grown on 4-well chamber slides and transfected with Lipofectamine LTX with Plus reagent. Twenty-four hours after transfection, cells were fixed with 4% paraformaldehyde, permeabilized with 0.1% Triton-X in phosphate-buffered saline, 1mM MgCl<sub>2</sub>, 0.1mM CaCl<sub>2</sub> (PBS-MC), and then blocked in 5% normal goat serum (NGS) in PBS-MC for 1 hour at room temperature. Incubation with primary antibodies (FLAG-M2, 1:100, Sigma; ASpolyP, 1:50, NeoScientific; ASpolyA, 1:10, NeoScientific; Nucleolin, 1:500, Abcam [Cambridge, MA], ab22758) diluted in 5% NGS in PBS-MC was done overnight at 4°C. After rinsing cells with PBS-MC, slides were incubated with goat antimouse and goat antirabbit IgG antibodies conjugated with Alexa 488 and Alexa 555, respectively (1:500 each, Thermo Fisher Scientific, A11029 and A21428, respectively) for 1 hour at room temperature. Slides were washed, and coverslips were mounted with Prolong Gold with DAPI (Thermo Fisher Scientific). Images were captured on an inverted Olympus (Tokyo, Japan) IX71 microscope at the same exposure and processed using SlideBook 5.5 software, with changes in brightness and contrast applied to the whole



**FIGURE 1: ASFMR1 transcript and putative repeat-associated non-AUG (RAN) translation products. (A)** The *FMR1* locus is bidirectionally transcribed with a start site (TSS) between exons 2 and 3 in the antisense orientation. *ASFMR1a* mRNA includes an AUG start codon upstream of the CCG repeat, creating a polyproline-containing open reading frame (ASFMRP). RAN translation of *ASFMR1a* mRNA could result in proline (+0 reading frame, ASFMRpolyP), arginine (+1, ASFMRpolyR), and alanine (+2, ASFMRpolyA) repeat proteins. Underlined regions are epitopes used for antibody generation. **(B)** *ASFMR1* RAN-specific reporters contain the *ASFMR1* mRNA sequence including the repeat upstream of a C-terminally 3xFLAG-tagged nanoluciferase lacking a start codon (GGG-NL). Expression constructs were generated for each reading frame by addition of nucleotide frameshifts (+/-FS) at different CCG repeat lengths and with or without the AUG start codon for ASFMRP (+/-AUG). CMV = cytomegalovirus promoter.

image and identically to all images used in a given figure according to published standards.<sup>25</sup>

### Immunohistochemistry and Coimmunofluorescence

Control and FXTAS autopsy tissue was obtained from the University of Michigan Brain Bank and the New York Brain Bank with informed consent of patients or their relatives and approval of local institutional review boards. Two of the FXTAS cases were previously described.<sup>26</sup> The third case had parkinsonism, progressive gait difficulties, and dementia with onset in his 70s. Autopsy showed ubiquitinated inclusions in his brainstem, hippocampus, and cortex with rare Lewy bodies in the substantia nigra (data not shown). Tissue sections were processed as previously described.<sup>7</sup> For immunohistochemistry, primary antibodies (ASpolyA, 1:10; ASpolyP, 1:100; ubiquitin, 1:250, Dako [Carpinteria, CA], Z0458) were diluted in 5% NGS in Tris, pH 7.6, 0.1% Triton X-100, 0.5% bovine serum albumin and incubated overnight at 4°C. Antigen retrieval was required for some antibodies

(ubiquitin: 0.1mM sodium citrate, pH 8, 10 minutes at 80°C; ASpolyA: 0.1mM sodium citrate, pH 6, 5 minutes at 80°C). For coimmunofluorescence studies with ubiquitin, primary antibodies (ubiquitin, 1:250, Millipore [Billerica, MA], MAB1510 with ASpolyA, 1:10 or ASpolyP, 1:50) were incubated overnight at 4°C with 5% NGS in PBS-MC, and incubated with goat anti-rabbit and goat antimouse IgG antibodies conjugated with Alexa 488 and Alexa 635, respectively (1:500, Thermo Fisher Scientific, A11008 and A31574, respectively). Images were captured on an Olympus confocal microscope, compiled using ImageJ, and analyzed as previously described.<sup>7</sup>

### Results

The *FMR1* locus is bidirectionally transcribed under both normal and pathological conditions (see Fig 1).<sup>20,21</sup> The antisense transcript, *ASFMR1*, has multiple promoters and alternative splice isoforms. One isoform, *ASFMR1a*, includes a region of the second intron and exon of *FMR1* fused by alternative splicing to ~500 nucleotides of exon

1 extending past the annotated transcription start site of *FMRI*.<sup>20</sup> *ASFMR1a* transcripts include the 5' leader sequence of *FMRI* and the repeat in a CCG orientation. This transcript has a predicted open reading frame with an AUG start codon that includes the repeat in the polyproline reading frame (ASFMRP; see Fig 1A).<sup>20</sup> If RAN translation of *ASFMR1a* mRNA were to occur upstream of the repeat, it would produce 3 different repeat proteins: ASFMRpolyP, a polyproline protein derived from the same CCG (+0) reading frame as ASFMRP; ASFMRpolyR, a polyarginine protein from the CGC (+1) reading frame; and ASFMRpolyA, a polyalanine protein from the GCC (+2) reading frame.

To evaluate whether CCG repeats in the context of the *ASFMR1a* transcript can support RAN translation, we generated a series of CCG RAN translation-specific NL reporters tagged with a carboxy-terminal 3xFLAG epitope to allow for easy detection by Western blot and immunocytochemistry and quantification by luciferase activity (see Fig 1B).<sup>24</sup> To assure that the reporter only provided signal if initiation occurred within the *ASFMR1* sequence, we mutated the AUG start codon of NL to GGG, which eliminated most of its translation and luciferase activity.<sup>24</sup> We then cloned the *ASFMR1a* sequence upstream of the modified NL coding sequence. Two sets of constructs were generated. One set retained the native AUG start codon naturally present in *ASFMR1*, and a second set was generated where this AUG was removed (see Fig 1B). In addition, frameshifts were introduced below the repeat such that the NL coding sequence would be in frame for each of the 3 different potential RAN products, creating repeat protein–luciferase fusions. We also generated constructs with a range of repeat sizes, from the normal range in humans (~25 repeats) up into the pathologic repeat size in FXTAS patients (>55 repeats).

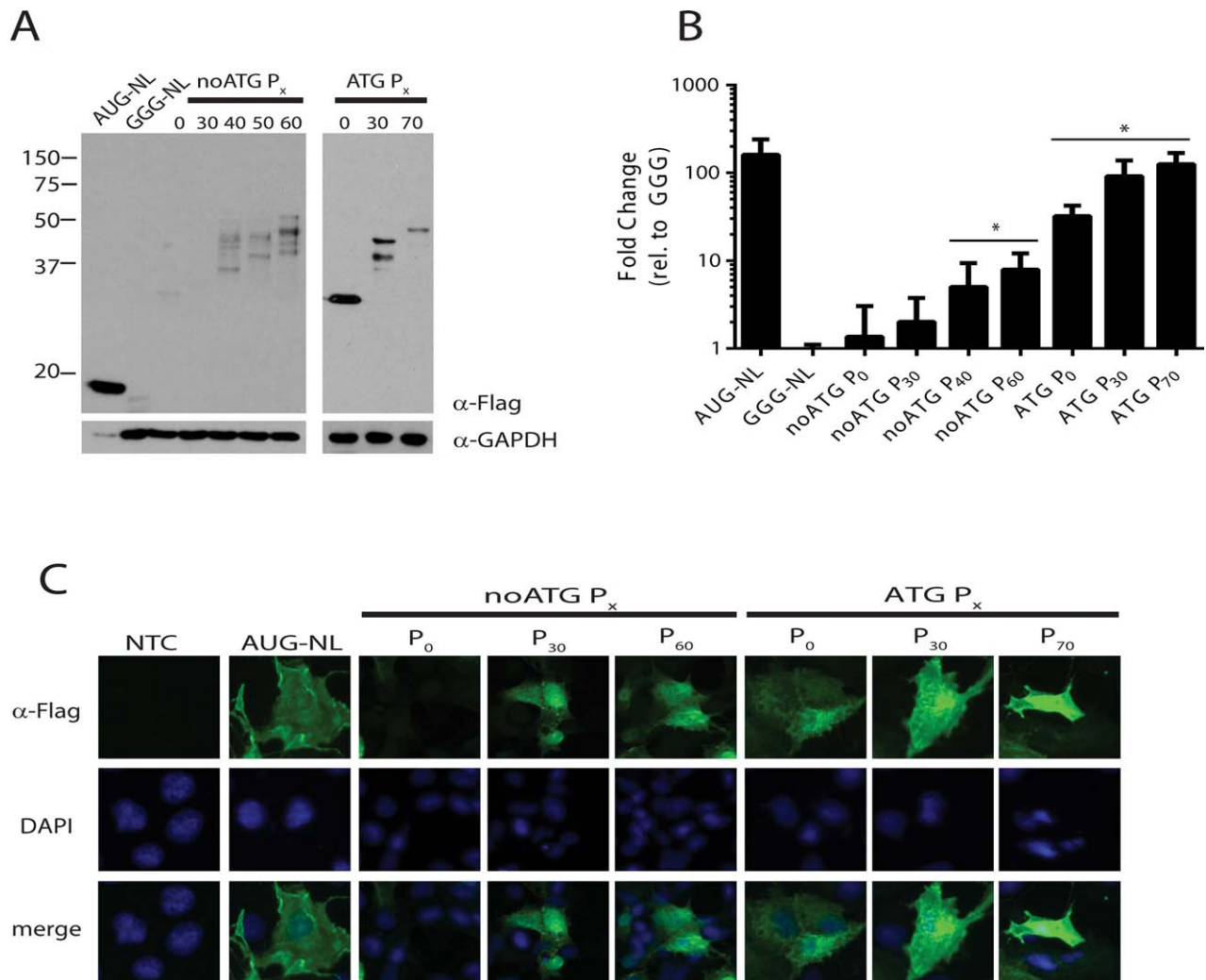
AUG-NL expressed by transient transfection in COS-7 cells ran as a single band of ~18kD by Western blot analysis of the 3X-FLAG tag (Fig 2). Mutation of this start codon to GGG led to a marked reduction in NL production and luciferase signal.<sup>24</sup> When the *ASFMR1* sequence was introduced above GGG NL in the +0 (proline) reading frame, higher molecular weight products were detected that increased in size with increasing CCG repeat length, consistent with translation initiating in *ASFMR1* above the repeat. To determine whether RAN translation could occur, the natural upstream AUG start codon initiating ASFMRP was removed. This change markedly decreased the higher molecular weight protein products in constructs lacking the CCG repeat or with a normal repeat size (30 CGG repeats). However, at larger repeat sizes (>40 CGG repeats) these higher molecular weight species persisted in the absence of the AUG codon,

with enhanced abundance with increasing repeat size as measured by luciferase activity, consistent with RAN translational initiation.

To determine the subcellular distribution of these novel polyproline-containing proteins, we performed immunofluorescence against the FLAG tag epitope in transfected cells (see Fig 2C). AUG-initiated ASFMRP fused to NL was distributed throughout the cytoplasm and nucleus in a pattern that was similar to AUG-NL alone. Neither increasing repeat size nor the removal of the AUG start codon significantly altered this distribution.

We next tested whether RAN translation could occur in the other 2 potential repeat reading frames (CGC, arginine and GCC, alanine). When the GGG-NL reporter was placed in the +1 (arginine) reading frame, we observed a higher molecular weight species by Western blot that increased with increasing repeat number and formed large (>150kD) complexes at higher repeat sizes, consistent with published reports of arginine-containing RAN proteins (Fig 3).<sup>8,11</sup> This product was generated and present regardless of whether the AUG codon normally in the proline reading frame was present or removed (see Fig 3A, noATG). As with ASFMRpolyP, translation of ASFMRpolyR was enhanced at increasing repeat sizes. Unlike the staining pattern of ASFMRpolyP, ASFMRpolyR exhibited a clear redistribution in its intracellular staining pattern that was dependent on the arginine repeat, such that at normal and expanded repeat numbers, the protein localized to the nucleolus. This change in localization was confirmed by coimmunofluorescence with the nucleolar marker nucleolin and is consistent with findings reported for dipeptide repeat-containing RAN products generated in C9orf72.<sup>27,28</sup>

In the third (+2, alanine, GCC) potential reading frame, a similar pattern was observed (Fig 4). Translation of higher molecular weight species were observed in this reading frame in both the presence and absence of the repeat. Mutational analysis demonstrated that in the absence of any repeat, initiation occurred predominantly below the repeat at 2 different near-AUG codons found in the human sequence (see Fig 4A and data not shown). Initiation at these sites did not significantly increase the luciferase signal above that of our negative control construct. However, at increasing repeat sizes, an additional higher molecular weight product appeared that increased in size in line with repeat number. At > 50 CCG repeats, products accumulated as a large complex (>150kD) near the top of the gel. ASFMRpolyA abundance increased with addition of repeats and was greatest at the largest repeat sizes. ASFMRpolyA production was significantly suppressed by inclusion of the upstream AUG codon in the proline reading frame in the absence of any repeat or

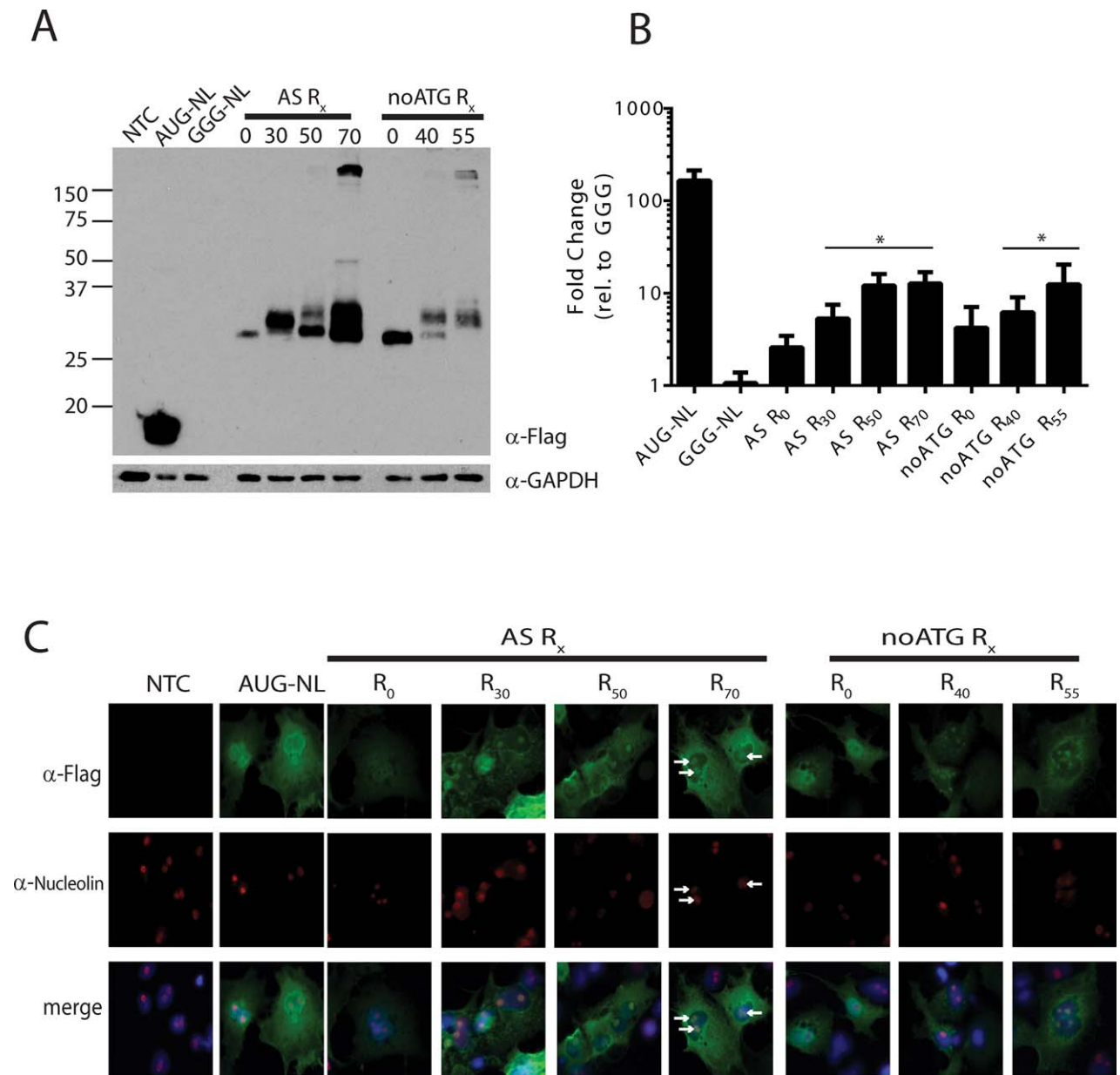


**FIGURE 2:** Repeat-associated non-AUG (RAN) translation from CCG repeats in the proline reading frame of *ASFMR1*. (A) Anti-FLAG Western blot of whole cell lysates from COS-7 cells transfected with the indicated reporters for ASFMRP and ASFMRpolyP. Molecular weight of *ASFMR1*-derived proteins increased with expanded repeats (ATG P<sub>x</sub>). Removal of the AUG start codon for ASFMRP (noATG P<sub>x</sub>, left) eliminated these proteins in the absence of a repeat, but did not prevent their generation at larger repeat sizes. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) served as a loading control and AUG-nanoluciferase (NL) and GGG-NL served as positive and negative controls, respectively. (B) NL activity from indicated constructs. For all figures, the Y axis shows mean ± standard deviation expressed as fold change above GGG-NL (n > 3). \*p < 0.05 by Fisher's least significant difference with Bonferroni correction for individual comparisons to GGG-NL and by analysis of variance for repeat length-dependent differences among RAN reporters. (C) Localization of ASFMRP and ASFMRpolyP in transfected COS-7 cells stained for FLAG (green). There was no change in distribution with increasing repeat size, and FLAG-positive inclusions were not observed. DAPI (blue) was used to counterstain nuclei. NTC = Non-transfected control.

at normal repeat sizes, but not at expanded repeats. This alanine translation product, ASFMRpolyA, exhibited a clear change in its cellular distribution in transfected cells compared to AUG-NL alone, with a repeat length-dependent redistribution into the nucleus.

If RAN translation products from *ASFMR1a* mRNA are generated in vivo, then we would predict that we should be able to identify them in FXTAS patients. To test this hypothesis, we generated a series of polyclonal antibodies against short stretches of the predicted repeats and the C-terminal portion of the predicted proteins (see Fig 1A, underlined region of protein sequences). To assess the

specificity of these antibodies, we performed Western blotting and immunofluorescence on cells expressing the appropriate RAN translation product or control constructs lacking the 5' leader sequence. Both the ASFMRpolyP-specific and ASFMRpolyA-specific antibodies exhibited specificity for their cognate RAN protein by Western blot (Fig 5A, C) and by immunofluorescence (see Fig 5B, D), respectively. Unfortunately, despite multiple attempts and use of different epitopes, antibodies generated against the predicted ASFMRpolyR protein failed to exhibit sufficient specificity in these validation assays to support further testing in human tissues (data not shown).

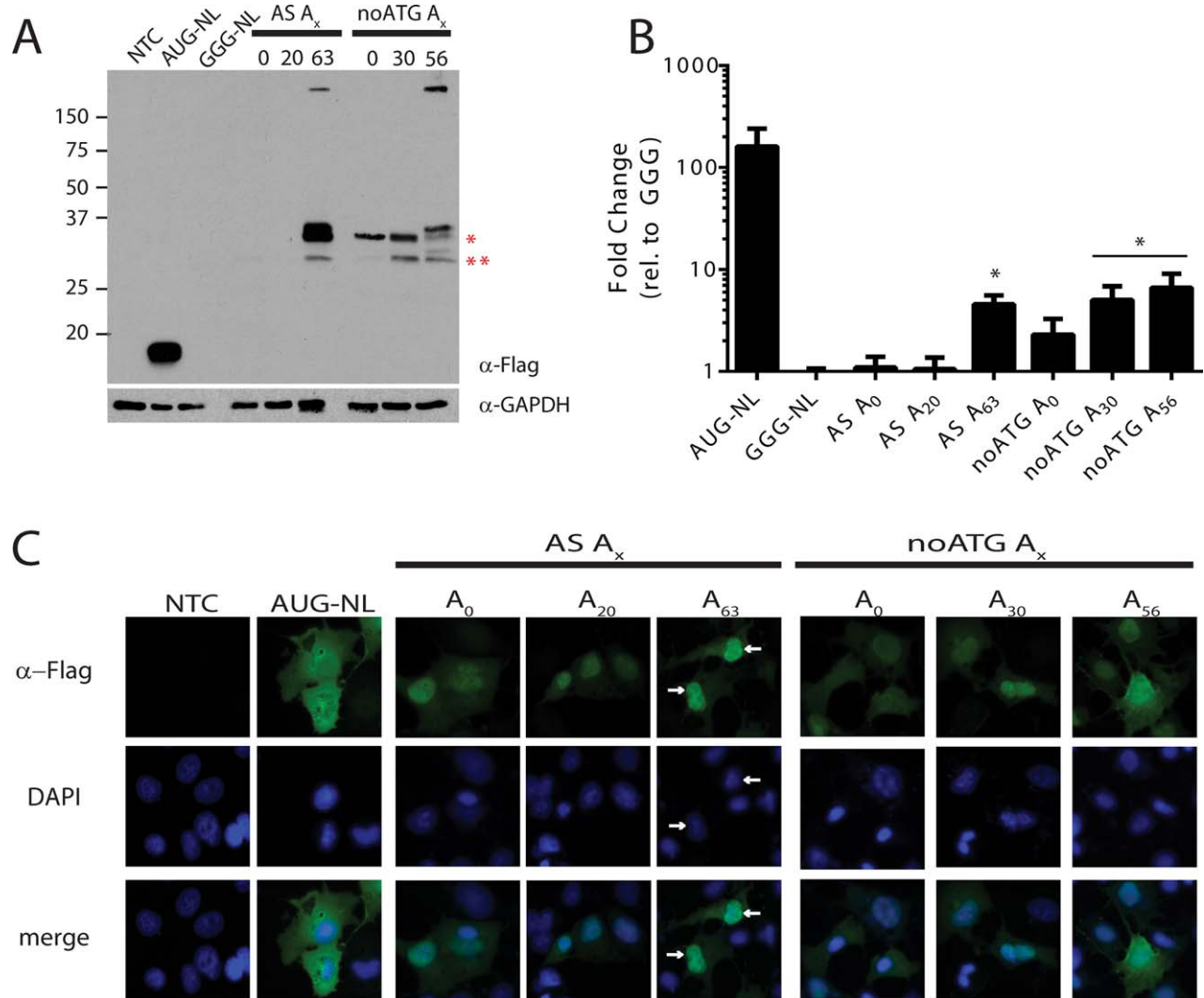


**FIGURE 3:** Repeat-associated non-AUG (RAN) translation from CGC repeats in the arginine reading frame of *ASFMR1*. (A) Western blot against FLAG in cells expressing the indicated *ASFMRpolyR* reporters. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a loading control. “noATG” indicates the ATG codon in the proline reading frame was removed. (B) Nanoluciferase (NL) activity derived from the indicated constructs 24 hours after transfection. \* $p < 0.05$  by Fisher least significant difference with Bonferroni correction for individual comparisons to GGG-NL and by analysis of variance for repeat length-dependent differences among RAN reporters. (C) FLAG staining of *ASFMRpolyR* constructs transfected into COS-7 cells. *ASFMRpolyR* staining (green) shifted to the nucleus in the presence of the expanded CCG repeats and colocalized (arrows) with the nucleolar marker nucleolin (red). DAPI (blue) was used to counterstain nuclei. NTC = no transfection control, AS = antisense.

With the validated antibodies, we looked for the presence of the *ASFMR1* protein products in human brain tissue from FXTAS patients. Preimmune sera for both antibodies showed minimal background staining in both controls and FXTAS cases (Figs 6C and 7C). When FXTAS tissues were stained with *ASFMRpolyP* antibodies, we observed staining in both FXTAS and control tissues in the hippocampus, cortex, and cerebellum (see Fig 6). Staining was most intense in

the perinuclear region of neurons and was more robust in most tissues in FXTAS cases than in controls. In addition, *ASFMRpolyP* antibodies reliably stained intranuclear neuronal aggregates in FXTAS tissues that were not observed in control tissues. These aggregates were present in multiple brain regions and were primarily neuronal. To better characterize these aggregates, we performed coimmunofluorescence with *ASFMRpolyP* and ubiquitin followed by confocal





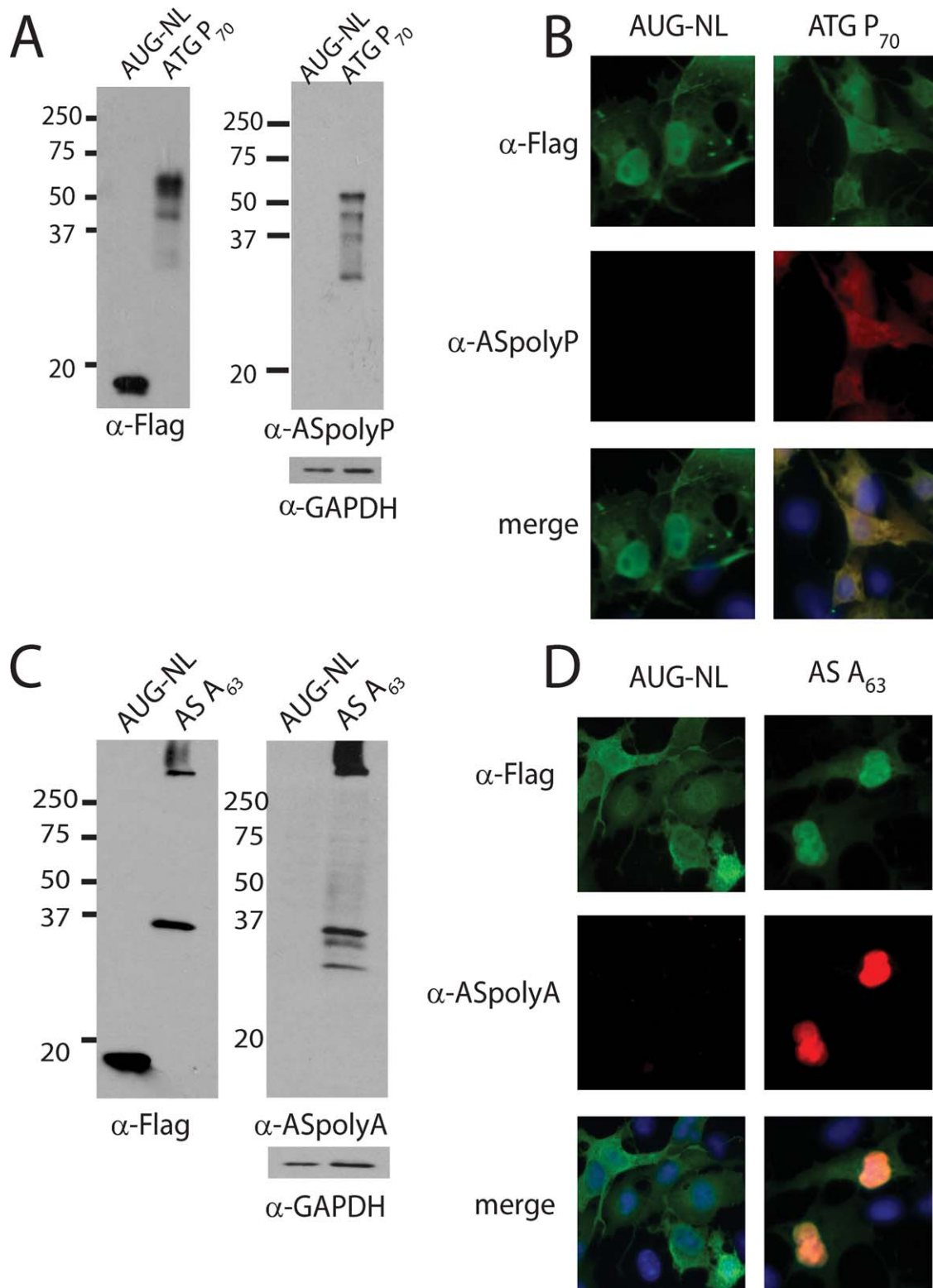
**FIGURE 4:** Repeat-associated non-AUG (RAN) translation from GCC repeats in the alanine reading frame of *ASFMR1*. (A) Western blot against FLAG on lysates from COS-7 cells transfected with indicated *ASFMRpolyA* reporters. Red asterisks indicate bands generated from initiation 3' to the repeat site but in the human sequence at a noncanonical start codon. "noATG" indicates the ATG codon in the proline reading frame was absent. (B) Nanoluciferase (NL) activity from *ASFMRpolyA* constructs compared to GGG-NL. \* $p < 0.05$  by Fisher least significant difference with Bonferroni correction for individual comparisons to GGG-NL and by analysis of variance for repeat length-dependent differences among RAN reporters. (C) Localization of *ASFMRpolyA* (green) was primarily nuclear (arrows) compared to AUG-NL, which was cytoplasmic. DAPI (blue) was used to counterstain nuclei. NTC = no transfection control, GAPDH = glyceraldehyde-3-phosphate dehydrogenase, AS = antisense.

microscopy. We observed numerous aggregates in neurons in FXTAS cases that were both ubiquitin and *ASFMRpolyP* positive.

To determine whether *ASFMRpolyA* also accumulated in FXTAS cases, we performed similar immunohistochemical and coimmunofluorescence studies. Like *ASFMRpolyP*, *ASFMRpolyA* was found extensively in FXTAS, with the greatest staining in the perinuclear regions of neurons in the hippocampus and cortex (see Fig 7A). *ASFMRpolyA* stained numerous intranuclear neuronal inclusions in multiple tissues, and these inclusions were ubiquitin positive by coimmunofluorescence (see Fig 7B, D).

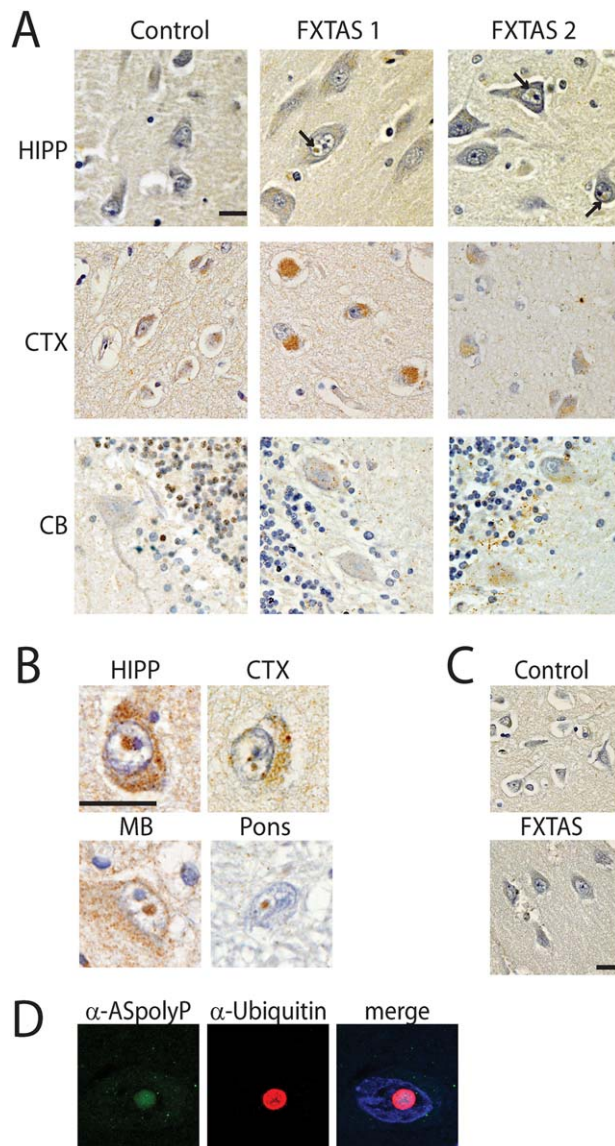
## Discussion

Aggregation of misfolded proteins is a hallmark in neurodegenerative disorders across a spectrum of etiologies.<sup>2,29</sup> Identification of the misfolded protein species in each disorder is a critical first step in elucidating the pathogenic cascades responsible for neurodegeneration in that condition. Here we demonstrate at least 2 new proteins that accumulate in the neurodegenerative disorder FXTAS: a polyproline-containing protein, *ASFMRpolyP*, and a polyalanine-containing protein, *ASFMRpolyA*. Using a series of reporter constructs, we further demonstrate that both of these proteins can be generated



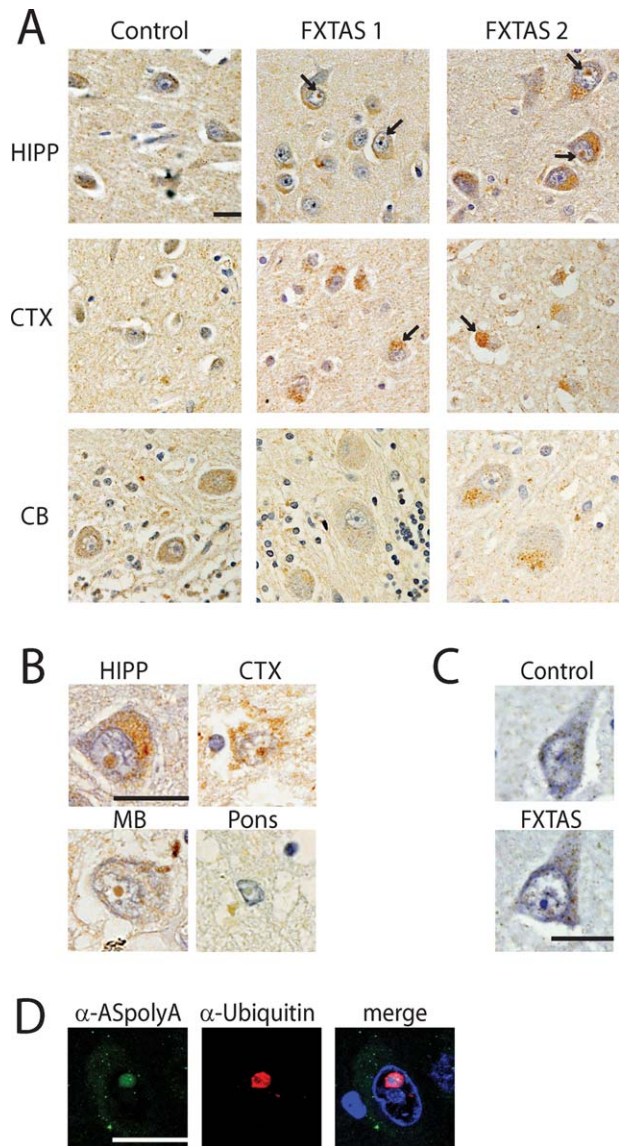
**FIGURE 5:** ASFMRpolyP and ASFMRpolyA antibody validation. (A, C) Western blots of constructs probed with FLAG or antibody generated against ASFMRpolyP ( $\alpha$ -ASpolyP; A), or against ASFMRpolyA ( $\alpha$ -ASpolyA; C). Glyceraldehyde-3-phosphate dehydrogenase GAPDH was used as a loading control. (B, D)  $\alpha$ -ASpolyP (red; B) recognized ASFMRpolyP but not AUG-nanoluciferase (NL) by coimmunofluorescence. Similarly,  $\alpha$ -ASpolyA (red; D) specifically recognized ASFMRpolyA but not AUG-NL. AS = antisense.





**FIGURE 6:** ASFMRpolyP accumulates in fragile X-associated tremor/ataxia syndrome (FXTAS) brain tissue and intranuclear inclusions. (A) Control and FXTAS tissue from the indicated brain regions stained with  $\alpha$ -ASpolyP. Hematoxylin (blue) was used as a counterstain to identify nuclei. In addition to strong perinuclear staining, nuclear aggregates (arrows) were observed in FXTAS cases that were not seen in control tissue. (B) Higher magnification of intranuclear neuronal inclusions from the indicated brain regions from FXTAS cases probed with  $\alpha$ -ASpolyP. (C) Preimmune sera for  $\alpha$ -ASpolyP did not show staining in control or FXTAS patient cortex. (D) Immunofluorescence of FXTAS hippocampus showed colocalization of ASFMRpolyP (green) with ubiquitin (red). DAPI (blue) was used to identify nuclei. Scale bars = 20  $\mu$ m. CB = cerebellum; CTX = frontal cortex; HIPP = hippocampus; MB = midbrain.

through RAN translation, a recently described unconventional form of initiation that occurs at multiple repeat expansions, including the CGG repeat in the sense strand of *FMRI* in FXTAS.<sup>7</sup>



**FIGURE 7:** ASFMRpolyA repeat-associated non-AUG proteins aggregate in fragile X-associated tremor/ataxia syndrome (FXTAS) brain tissue. (A) Control and FXTAS tissue from the indicated brain regions stained with  $\alpha$ -ASpolyA. Hematoxylin (blue) was used to identify nuclei. In addition to strong perinuclear staining, nuclear aggregates (arrows) were observed in FXTAS cases that were not seen in control tissue. (B) Higher magnification of intranuclear neuronal inclusions from the indicated brain regions from FXTAS cases probed with  $\alpha$ -ASpolyA. (C) Preimmune sera for  $\alpha$ -ASpolyA did not show specific staining in control or FXTAS patient cortex. (D) Coimmunofluorescence on FXTAS hippocampus showed colocalization of ASFMRpolyA (green) with ubiquitin (red). DAPI (blue) was used to identify nuclei. Scale bars = 20  $\mu$ m. CB = cerebellum; CTX = frontal cortex; HIPP = hippocampus; MB = midbrain.

RAN translation has now been reported for 6 different repeats, with 3 derived from sense strand mRNA transcripts (CAG, GGGGCC, and CGG repeats) and 3 derived primarily from antisense RNAs (CAG, CUG,

CCCCGG, and CCG repeats).<sup>5,7–13</sup> All of these transcripts are capable of forming strong secondary structures in vitro, either RNA hairpins or G-quadruplexes, and the ability to form these secondary structures appears important to the process underlying RAN translation.<sup>5,27,30,31</sup> That this CCG repeat resides within an open reading frame is also consistent with previous data on CAG repeats in Huntington disease and spinocerebellar ataxia type 8, both of which allow RAN translation to occur in all 3 potential reading frames within an open reading frame.<sup>13</sup> However, recent work on RAN translation at CCG repeats supports a model for initiation that retains a requirement for a 7-methylguanosine 5' cap on the mRNA and ribosomal scanning,<sup>24</sup> both of which superficially do not fit with a location of the repeat downstream within an open reading frame and would require RAN translation to bypass the canonical AUG start codon. Thus, studies of initiation mechanisms at CCG repeats in its native sequence context will be needed to delineate how this atypical process occurs in human cells and how it agrees or disagrees with findings at other repeats.

What role *ASFMR1* mRNA-derived protein products have in FXTAS disease pathogenesis is also unclear. Expression of CCG repeats in isolation in the 5' leader sequence of GFP was sufficient to elicit toxicity in a *Drosophila* model system.<sup>23</sup> However, whether this toxicity was driven by the repeats as RNA or as RAN-translated proteins is not known. Polyalanine-containing proteins and RAN products have previously been shown to be toxic in isolation,<sup>5,32</sup> and oculopharyngeal muscular dystrophy results from a polyalanine expansion in polyadenylate-binding protein 2.<sup>33,34</sup> In contrast, little is known about the potential for polyproline or polyarginine proteins to cause toxicity. Expression of dipeptide-repeat-containing proteins generated from GGGGCC repeats in C9orf72-containing arginine and/or proline (glycine–arginine, proline–alanine, and proline–arginine repeats) is toxic in cells and simple model systems when expressed at high levels in the absence of a structured RNA repeat.<sup>35–38</sup> Moreover, the repetitive arginine elements target these proteins to the nucleolus,<sup>38</sup> which we also observe with ASFMRpolyR in cell transfection studies (Fig. 3b). Thus, defining both the potential for each of these *ASFMR1* RAN proteins to elicit toxicity in isolation in model systems and their relative abundance in FXTAS tissues will be important next steps in determining their potential roles in disease pathogenesis.

Our current study and antibodies cannot differentiate between polyproline products generated by AUG-initiated translation and RAN translation. Our cell-based luciferase reporter assays suggest that removal of the AUG start codon from the ASFMRP open reading frame

decreases the signal by approximately 10-fold (see Fig 2B). In addition, we observed staining in control tissues with this antibody (although no staining of intranuclear aggregates; see Fig 6A), but limited RAN products were detected based on reporter assays at normal repeat sizes (see Fig 2A, B). Thus, our antisense proline antibody is likely staining both the AUG-initiation derived ASFMRP protein and the RAN-derived ASFMRpolyP protein, and both are potentially contributing to aggregation formation and toxicity.

In summary, we provide evidence for RAN translation at CCG repeats derived from an antisense *FMR1* transcript. This work expands the list of potential pathogenic species at play in fragile X-associated tremor/ataxia syndrome and provides further support for noncanonical translation of microsatellite repeat expansions in the pathology of human neurodegenerative disease.

---

## Acknowledgments

This work was funded by grants from the NIH (R01NS086810, P.K.T.; F3211541507, M.G.K.) and Veterans Health Administration (1101BX001689, P.K.T.). Histological samples from FXTAS cases were provided by the New York Brain Bank at Columbia University and University of Michigan Brain Bank (Michigan Alzheimer Disease Center grant P30 AG053760).

We thank members of the Todd laboratory for technical suggestions and contributions; H. Paulson and A. Lieberman for comments on the manuscript; and Dr L. Ranum for supplying unpublished reagents that were used in preliminary work not included in this article.

## Author Contributions

All authors contributed to the conception, design of the study, and data analysis; A.K. performed the experiments; A.K. and P.K.T. wrote the manuscript with editorial input from M.G.K.

## Potential Conflicts of Interest

Nothing to report.

---

## References

1. Mason AR, Ziemann A, Finkbeiner S. Targeting the low-hanging fruit of neurodegeneration. *Neurology* 2014;83:1470–1473.
2. Williams AJ, Paulson HL. Polyglutamine neurodegeneration: protein misfolding revisited. *Trends Neurosci* 2008;31:521–528.
3. Todd PK, Paulson HL. RNA-mediated neurodegeneration in repeat expansion disorders. *Ann Neurol* 2010;67:291–300.
4. Nelson DL, Orr HT, Warren ST. The unstable repeats—three evolving faces of neurological disease. *Neuron* 2013;77:825–843.

5. Zu T, Gibbens B, Doty NS, et al. Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A* 2011;108:260–265.
6. Green KM, Linsalata AE, Todd PK. RAN translation—what makes it run? *Brain Res* 2016;1647:30–42.
7. Todd PK, Oh SY, Krans A, et al. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* 2013;78:440–455.
8. Mori K, Weng SM, Arzberger T, et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science* 2013;339:1335–1338.
9. Mori K, Arzberger T, Grasser FA, et al. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol* 2013;126:881–893.
10. Zu T, Liu Y, Banez-Coronel M, et al. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A* 2013;110:E4968–E4977.
11. Ash PE, Bieniek KF, Gendron TF, et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 2013;77:639–646.
12. Gendron TF, Bieniek KF, Zhang YJ, et al. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol* 2013;126:829–844.
13. Banez-Coronel M, Ayhan F, Tarabochia AD, et al. RAN translation in Huntington disease. *Neuron* 2015;88:667–677.
14. Hagerman RJ, Leehey M, Heinrichs W, et al. Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* 2001;57:127–130.
15. Jacquemont S, Hagerman RJ, Leehey MA, et al. Penetrance of the fragile X-associated tremor/ataxia syndrome in a premutation carrier population. *JAMA* 2004;291:460–469.
16. Berry-Kravis E, Abrams L, Coffey SM, et al. Fragile X-associated tremor/ataxia syndrome: clinical features, genetics, and testing guidelines. *Mov Disord* 2007;22:2018–2030; quiz 2140.
17. Hagerman P. Fragile X-associated tremor/ataxia syndrome (FXTAS): pathology and mechanisms. *Acta Neuropathol* 2013;126:1–19.
18. Greco CM, Hagerman RJ, Tassone F, et al. Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers. *Brain* 2002;125:1760–1771.
19. Oh SY, He F, Krans A, et al. RAN translation at CGG repeats induces ubiquitin proteasome system impairment in models of fragile X-associated tremor ataxia syndrome. *Hum Mol Genet* 2015;24:4317–4326.
20. Ladd PD, Smith LE, Rabaia NA, et al. An antisense transcript spanning the CGG repeat region of FMR1 is upregulated in premutation carriers but silenced in full mutation individuals. *Hum Mol Genet* 2007;16:3174–3187.
21. Khalil AM, Faghihi MA, Modarresi F, et al. A novel RNA transcript with antiapoptotic function is silenced in fragile X syndrome. *PLoS One* 2008;3:e1486.
22. Kiliszek A, Kierzek R, Krzyzosiak WJ, Rypniewski W. Crystallographic characterization of CCG repeats. *Nucleic Acids Res* 2012;40:8155–8162.
23. Sofola OA, Jin P, Botas J, Nelson DL. Argonaute-2-dependent rescue of a Drosophila model of FXTAS by FRAXE premutation repeat. *Hum Mol Genet* 2007;16:2326–2332.
24. Kearsse MG, Green KM, Krans A, et al. CGG repeat-associated non-AUG translation utilizes a cap-dependent scanning mechanism of initiation to produce toxic proteins. *Mol Cell* 2016;62:314–322.
25. Rossner M, Yamada KM. What's in a picture?. The temptation of image manipulation. *J Cell Biol* 2004;166:11–15.
26. Louis E, Moskowitz C, Friez M, et al. Parkinsonism, dysautonomia, and intranuclear inclusions in a fragile X carrier: a clinical-pathological study. *Mov Disord* 2006;21:420–425.
27. Haeusler AR, Donnelly CJ, Periz G, et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 2014;507:195–200.
28. Schludi MH, May S, Grasser FA, et al. Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathol* 2015;130:537–555.
29. Forman MS, Trojanowski JQ, Lee VM. Neurodegenerative diseases: a decade of discoveries paves the way for therapeutic breakthroughs. *Nat Med* 2004;10:1055–1063.
30. Reddy K, Zamiri B, Stanley SY, et al. The disease-associated r(GGGGCC)<sub>n</sub> repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J Biol Chem* 2013;288:9860–9866.
31. Sobczak K, de Mezer M, Michlewski G, et al. RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res* 2003;31:5469–5482.
32. van Eyk CL, McLeod CJ, O'Keefe LV, Richards RI. Comparative toxicity of polyglutamine, polyalanine and polyisoleucine tracts in Drosophila models of expanded repeat disease. *Hum Mol Genet* 2012;21:536–547.
33. Brais B, Bouchard JP, Xie YG, et al. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. *Nat Genet* 1998;18:164–167.
34. Becher MW, Kotzok JA, Davis LE, Bear DG. Intranuclear inclusions in oculopharyngeal muscular dystrophy contain poly(A) binding protein 2. *Ann Neurol* 2000;48:812–815.
35. Yamakawa M, Ito D, Honda T, et al. Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS. *Hum Mol Genet* 2015;24:1630–1645.
36. Mizielińska S, Lashley T, Norona FE, et al. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol* 2013;126:845–857.
37. Wen X, Tan W, Westergard T, et al. Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* 2014;84:1213–1225.
38. Tao Z, Wang H, Xia Q, et al. Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Hum Mol Genet* 2015;24:2426–2441.