FMRP inhibits translation elongation independent of mRNA G-quadruplexes

MaKenzie R. Scarpitti[1,2,3], Julia E. Warrick[1,2,3], and Michael G. Kearse[1,2,3*]

[1]The Biomedical Sciences Graduate Program, [2]Department of Biological Chemistry and Pharmacology, [3]Center for RNA Biology, The Ohio State University, Columbus, OH 43210 USA

*To whom correspondence should be addressed. Tel: +1 614 685 3709; Email: michael.kearse@osumc.edu

ABSTRACT

Loss of functional fragile X mental retardation protein (FMRP) causes fragile X syndrome, the leading form of inherited intellectual disability and the most common monogenic cause of autism spectrum disorders. FMRP is an RNA-binding protein that controls neuronal mRNA localization and translation. Notably, FMRP is thought to inhibit translation elongation after being recruited to target transcripts via binding RNA G-quadruplexes (G4s) within the coding sequence. Here we directly tested this model and report that FMRP inhibits translation elongation independent of mRNA G4s. Furthermore, we found that the RGG box motif together with its natural C-terminal domain forms a non-canonical RNA-binding domain (ncRBD) that binds reporter mRNA and all four polymeric RNA sequences. The ncRBD is essential for FMRP to inhibit translation. Transcripts that are bound by FMRP through the ncRBD co-sediment with heavy polysomes, which is consistent with stalling elongating ribosomes and a subsequent accumulation of slowed polysomes. Together, this work shifts our understanding of how FMRP inhibits translation elongation and supports a model where repression is driven by local FMRP and mRNA concentrations rather than target mRNA sequence.

INTRODUCTION

Loss of functional fragile X mental retardation protein (FMRP) causes fragile X syndrome (FXS) (1-4), the leading form of inherited intellectual disability and the most common monogenic cause of autism spectrum disorders (ASD) (5). FXS affects around one in 4,000 males and one in 8,000 females (5). Approximately one third of FXS patients are also diagnosed with ASD (6). The vast majority of FXS cases is caused by a CGG trinucleotide repeat expansion in the 5' untranslated region (UTR) of the *FMR1* gene. The expanded repeat is subsequently hypermethylated which causes transcriptional silencing of the locus (1-4). As a result, most FXS patients express little to no functional FMRP.

FMRP is an RNA-binding protein (RBP) with multiple RNA-binding domains (RBDs) including three K homology (KH) domains (KH0, KH1, and KH2) and a positively charged RGG box motif (7-9). In general, KH domains canonically bind to short stretches of RNA to provide binding specificity to particular transcripts (10-14). However, the KH domains of FMRP are unable to bind strongly to any combination of five ribonucleotides (15). *In vitro* selection experiments have found that the KH2 domain of FMRP does have high affinity for an artificial RNA pseudoknot (e.g., Δ kissing complex 2; Δkc2) (16). Reports using *in vitro* selection, filter binding, and fluorescence anisotropy have concluded that the RGG box motif of FMRP has preference for RNA G-quadruplexes (G4s) (15,17). Using these multiple binding domains and motifs, FMRP is thought to bind mRNAs with higher order RNA structure (e.g., pseudoknots and G4s). Importance of the KH domains and RGG box motif in FMRP function is highlighted by independent point mutations found in rare FXS cases. Enigmatically, a G266E mutation in the KH1 domain (18), a I304N mutation in the KH2 domain (19), and a guanine insertion that causes a frameshift early within the RGG box motif (20) all cause FXS. This suggests that FMRP may have multiple functions within neurons dictated by interactions between specific protein RBDs or motifs and mRNA sequences or structures.

FMRP is known to regulate translation. Consistent with this idea, FMRP in mouse brain tissue is found to primarily bind the coding sequence (CDS) of mRNAs (21). Most previous reports (21-25), but not all (23,26,27), suggest that FMRP represses translation of its target mRNAs by stalling bound elongating ribosomes. Thus, at least one facet of the FXS phenotype is thought to result from aberrant and unregulated protein synthesis of dendritic mRNAs at synapses. Because FMRP mutants that lack the RGG box motif do not repress translation (24), it is commonly postulated that FMRP binds to target mRNAs at RNA G4s and subsequently blocks translation elongation. In this model, the G4-bound FMRP acts as a physical roadblock and directly contacts the elongating ribosome near the A site to sterically hinder delivery of aminoacyl-tRNAs to the ribosome.

Here we directly test the ability of FMRP to inhibit translation of G4-lacking and G4-containing mRNAs. Our data show that FMRP represses translation independent of RNA G4s and that the RGG box works synergistically with the positively charged C-terminal domain (CTD) to form a non-canonical RNA-binding domain (ncRBD) that is critical for repression. The RGG box motif + CTD ncRBD is able to bind all four polymeric RNA sequences with slightly less preference for polymeric A RNA, providing FMRP the ability to inhibit translation of multiple mRNAs. Consistent with inhibiting elongation, our data show that FMRP harboring the ncRBD, but not a mutant that is missing the ncRBD, promotes accumulation of ribosomes on mRNAs. Taken together, our data indicate that FMRP inhibits translation elongation independent of mRNA G4s and supports a model where FMRP repression is driven by local FMRP and mRNA concentrations rather than binding to specific sequences in target transcripts.

MATERIALS AND METHODS

Plasmids

The nLuc coding sequence from pNL1.1 (Promega) was analyzed by Quadruplex forming G-Rich Sequences (QGRS) Mapper (https://bioinformatics.ramapo.edu/QGRS/index.php) and manually codon optimized to eliminate predicted G4 motifs. The final nLuc coding sequence harboring the P2A ribosome skipping motif and human β-globin 5' UTR was synthesized by Integrated DNA Technologies and cloned into pcDNA3.1(+). The G4 sequence and larger 5' UTR (three human β-globin 5' UTR sequences in tandem) were inserted using the Q5 Site-Directed Mutagenesis Kit (NEB # E0552S). The pcDNA3.1-D/CrPV IGR IRES nLuc reporter was previously described (28) but was additionally modified by site-directed mutagenesis to harbor a strong hairpin upstream of the IRES element to block any scanning ribosomes.

An *E. coli* optimized coding sequence for human FMRP (isoform 1) was designed and synthesized by Genscript, and then subcloned into pET His6 MBP TEV LIC cloning vector (1M), which was a gift from Scott Gradia (Addgene plasmid # 29656), through ligation-independent cloning (LIC) using Novagen's LIC-qualified T4 DNA polymerase (Sigma # 70099-M) as described by Q3 Macrolab (http://qb3.berkeley.edu/macrolab/). The His6-tag was deleted from the N-terminus and inserted at the C-terminus. The NT-hFMRP sequence included a P451S mutation to prevent ribosome stalling at a poly-proline stretch and formation of truncated recombinant protein, as previously described (29). Point mutations and deletions were achieved using the Q5 Site-Directed Mutagenesis Kit.

All plasmids were propagated in TOP10 *E. coli* (Thermo Fisher # C404006), purified using the PureYield Plasmid Miniprep or Midiprep Systems (Promega # A1222 and A2495), and validated by Sanger sequencing at The Ohio State University Comprehensive Cancer Center Genomics Shared Resource (OSUCCC GSR). Nucleotide sequences of the reporters and recombinant proteins are provided in the Supplementary Data. All reporter and expression plasmids are available on request and will be deposited to Addgene.

Reporter mRNA *in vitro* transcription

All plasmids were linearized with XbaI and purified using a Zymo DNA Clean & Concentrator 25 (Zymo Research # D4065). DNA was transcribed into mRNA which was co-transcriptionally capped with the Anti-Reverse Cap Analog (ARCA) 3′-O-Me-m7G(5′)ppp(5′)G (NEB # S1411) using the HiScribe T7 High Yield RNA Synthesis Kit (NEB # E2040). Our standard 10 µL reactions used 0.5 µg of linear plasmid template and an 8:1 ARCA:GTP ratio. Reactions were incubated at 30°C for 2 hrs and then purified using a Zymo RNA Clean & Concentrator 25 (Zymo Research # R1018). Reporter mRNA was eluted in 75 µL RNase-free water, aliquoted in single use volumes, and stored at -80°C. The CrPV IGR IRES nLuc reporter was synthesized identically but instead used the G(5′)ppp(5′)A cap analog (NEB # S1406) to generate A-capped mRNAs. Reporter mRNA integrity was confirmed by denaturing formaldehyde agarose gel electrophoresis and ethidium bromide visualization. We consistently found the 30°C incubation resulted in less observable truncated products and did not significantly affect yield for our purposes.

Recombinant protein expression and purification

All recombinant proteins were expressed in Rosetta 2(DE3) *E. coli* (Sigma # 71397-4) using MagicMedia *E. coli* Expression Medium (Thermo Fisher # K6803) supplemented with 50 µg/mL kanamycin and 35 µg/mL chloramphenicol for auto-induction. A 5 mL starter culture in LB media supplemented with 50 µg/mL kanamycin, 35 µg/mL chloramphenicol, and 1% glucose (w/v) was inoculated with a single colony and grown overnight at 37°C, 250 rpm. 1 mL of a fresh overnight starter culture was then used to inoculate 50 mL of room temperature MagicMedia

and incubated for 48 hrs at 18°C, 160 rpm in a 250 mL baffled flask. After auto-induction, cultures were pelleted and stored at -20°C for purification later. Recombinant proteins were purified using a dual affinity approach, first using the C-terminal His6-tag, then the N-terminal MBP-tag. Cell pellets were resuspended and lysed with BugBuster Master Mix (Sigma # 71456) using the recommended 5 mL per 1 g wet cell pellet ratio for 10 min at room temperature with gentle end-over-end rotation (10-15 rpm). Lysates were placed on ice and kept cold moving forward. Lysates were cleared by centrifugation for 20 min at 18,000 rcf in a chilled centrifuge (4°C). Lysates were then incubated with HisPur Cobalt Resin (Thermo Fisher # 89965) in a Peirce centrifugation column (Thermo # 89897) for 30 min at 4°C with gentle end-over-end rotation. Columns were centrifuged in a pre-chilled (4°C) Eppendorf 5810R for 2 min at 700 rcf to eliminate the flow through and then were washed 5X with two resin-bed volumes of ice-cold Cobalt IMAC Wash Buffer (50 mM Na_3PO_4, 300 mM NaCl, 10 mM imidazole; pH 7.4) in a pre-chilled (4°C) Eppendorf 5810R for 2 min at 700 rcf. His-tagged proteins were then eluted in a single elution step with two resin-bed volumes of ice-cold Cobalt IMAC Elution Buffer (50 mM Na_3PO_4, 300 mM NaCl, 150 mM imidazole; pH 7.4) by gravity flow. Eluates were then incubated with Amylose resin (NEB # E8021) in a centrifugation column for 2 hrs at 4°C with gentle end-over-end rotation. Columns were washed 5X with at least two bed-volumes of ice-cold MBP Wash Buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA; pH 7.4) by gravity flow. MBP-tagged proteins were then eluted by a single elution step with two resin-bed volumes of ice-cold MBP Elution Buffer (20 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA, 10 mM maltose; pH 7.4). Recombinant proteins were then desalted and buffer exchanged into Protein Storage Buffer (25 mM Tris-HCl, 125 mM KCl, 10% glycerol; pH 7.4) using a 7K MWCO Zeba Spin Desalting Column (Thermo Fisher # 89892) and, if needed, concentrated using 10K NMWL Amicon Ultra-4 (EMD Millipore # UFC803024). Recombinant protein concentration was determined by Pierce Detergent Compatible Bradford Assay Kit (Thermo Fisher # 23246) with BSA standards diluted in Protein Storage Buffer before aliquoting in single use volumes, snap freezing in liquid nitrogen, and storage at -80°C.

mRNA folding, mRNP formation, and *in vitro* translation

28 nM *in vitro* transcribed reporter mRNA in RNA folding buffer (10 mM Tris-HCl, 100 mM KCl; pH 7.4) was heated for 5 min at 70°C then gradually cooled for 30 min at room temperature on the bench. 5 mM $Mg(OAc)_2$ (final) was then added, gently mixed, and allowed to cool for an additional 30 min on the bench. In a total of 4 µL, 28 fmol of folded reporter mRNA was mixed with 0-10 picomol of recombinant protein and 100 picomol of UltraPure BSA (Thermo Fisher # AM2618) on ice for 1 hr. UltraPure BSA stock was diluted in protein storage buffer and its addition was necessary to prevent non-specific binding of the reporter mRNA to the tube. For *in vitro* translation reactions, 6 µL of a Rabbit Reticulocyte Lysate (RRL) master mix was added to each 4µL mRNP complex. 10 µL *in vitro* translation reactions were performed in the linear range using 2.8 nM mRNA in the Flexi RRL System (Promega # L4540) with final concentrations of reagents at 30% RRL, 10 mM amino acid mix minus leucine, 10 mM amino acid mix minus Methionine, 0.5 mM $Mg(OAc)_2$, 100 mM KCl, 8 U murine RNase inhibitor (NEB # M0314), 0-1 µM recombinant protein, and 10 µM UltraPure BSA. Reactions were incubated for 30 min at 30°C, terminated by incubation on ice and diluted 1:5 in Glo Lysis Buffer (Promega # E2661). 25 µL of prepared Nano-Glo reagent (Promega # N1120) was mixed with 25 µL of diluted reaction and incubated at room temperature for 5 min in the dark (with gentle shaking during the first minute), and then read on a Promega GloMax Discover Multimode Microplate Reader.

IC_{50} measurements were calculated using non-liner regression analysis following the [inhibitor] versus normalized response analysis in GraphPad Prism 9.1.2 The 95% confidence intervals were plotted along the mean of three biological replicates.

Denaturing PAGE, native PAGE, and nucleic acid staining

Denaturing TBE-Urea 6% PAGE gels (Thermo Fisher # EC68652BOX) were run with 1X TBE-Urea Sample Buffer (Thermo Fisher # LC6876) and 1X TBE Running Buffer (Thermo Fisher # LC6675). Gels were pre-run at 180 volts (constant) for 20 minutes, then samples were loaded and run at 180 volts (constant) for 3 hrs. Native TBE 6% PAGE gels (Thermo Fisher # EC62652BOX) were run with 1X Hi-Density TBE Sample Buffer (Thermo Fisher # LC6678) and 1X TBE Running Buffer. Gels were pre-run at 180 volts (constant) for 1 hr, then samples were loaded and run at 180 volts (constant) for 3 hrs. Total RNA was stained with 1X SYBR Green II RNA Gel Stain (Thermo Fisher # S7568) diluted in milliQ water for 10 min in the dark. G4s were selectively stained with 0.1 mg/mL N-methyl mesoporphyrin IX (NMM) (Frontier Scientific # NMM58025MG) in milliQ water for 10 min in the dark. Stained gels were imaged on a Bio-Rad GelDoc Go Gel Imaging System using the SYBR Green setting. NMM stock was made at 10 mg/mL in DMF and stored in single use aliquots at -20°C. For Native TBE PAGE, the mRNA was folded as described above. Due to the different sensitivity, 100 ng and 500 ng were loaded for staining with SYBR Green II and NMM, respectively.

Fluorescent primer extension

mRNP complexes were scaled up 4-fold as compared to the values described in the mRNA folding section. To each reaction, 20 µL 5X AMV RT buffer (final: 50mM Tris-HCl, 50 mM KCl, 10 mM $MgCl_2$, 0.5 mM spermidine, 10 mM DTT; pH 8.3), 10 µL 10 µM 5'-FAM labeled-reverse primer (20 nt; synthesized by Integrated DNA technologies), 2 µL 25mM complete dNTP mix (Promega # U1432; final of each dNTP of 0.5 mM), 2 µL AMV Reverse Transcriptase (Promega # M5108; stock at 20-25 U/µL; final of 0.5 U/µl), and nuclease-free water up to 100 µL was added. Reverse transcription (RT) was allowed to progress for 35 min at 30°C. A known 258 nt labeled cDNA fragment was spiked-in for downstream normalization. FAM-labeled cDNA was extracted by transferring the 100 µL RT reaction to a new microcentrifuge tube with 150 µL nuclease-free water and adding 250 µL saturated Phenol:Chloroform:Isoamyl Alcohol (25:24:1), pH 8. After vigorous mixing for 1 min, samples were centrifuged at room temperature for 5 min at 18,000 rcf. The top aqueous phase was transferred and re-extracted with saturated Phenol:Chloroform:Isoamyl Alcohol (25:24:1), pH 8. The final aqueous supernatant was then concentrated using the Zymo DNA Clean and Concentrator-5 using a 7:1 ratio following the manufacture's recommendation. FAM-labeled cDNA was eluted in 7 µL nuclease-free water. 5 µL of each eluate was mixed with 10 µL Hi-Di Formamide (Thermo Fisher # 4440753), spiked with a GeneScan 500 LIZ dye Size Standard (Thermo Fisher # 4322682), and subjected to fragment analysis using an Applied Biosystems 3130xl Genetic Analyzer (at the OSUCCC GSR) with POP-7 polymer and all fragments being reported. Full length extension was represented by the expected 701 nt labeled cDNA fragment. The 5' FAM-labeled primer sequence is provided in the Supplementary Data.

Electrophoretic mobility shift assays

For EMSAs with mRNAs to assess mRNP formation, 4% PAGE gels made with 0.5X TBM (45 mM Tris, 45 mM borate, 2.5 mM $MgCl_2$) and acrylamide/bis-acrylamide, 37.5:1 (2.7% crosslinker) were poured between glass plates and allowed to polymerize for at least 1 hr. Gels were pre-run for 20 min at 100 volts (constant) with 0.5X TBM as the running buffer. In 18 µL, 1 picomol of folded reporter mRNA was mixed with 20 picomol of recombinant protein and 200 picomol of UltraPure BSA (Thermo Fisher # AM2618) in binding buffer (10 mM Tris-HCl, 100 mM KCl, 5 mM $Mg(OAc)_2$; pH 7.4) on ice for 1 hr. 2 µL of 20% Ficoll 400 (Sigma # F5415-25ML) was then added and the entire sample was loaded immediately. After loading, gels were run for 45 min at 100 volts (constant) at room temperature, stained with 1X SYBR Green II RNA Gel Stain, and visualized as described above.

For EMSAs with 5' FAM-labeled RNA oligos, 4% PAGE gels were made as described above with 0.5X TBM and acrylamide/bis-acrylamide, 37.5:1. 5' FAM-labeled RNA olgios were first diluted in nuclease-free water and not heated. In 18 µL, 4.8 picomol of 5' FAM-labeled RNA oligo was mixed with 54 picomol of recombinant protein and 542 picomol of UltraPure BSA in binding buffer (10 mM Tris-HCl, 100 mM KCl, 5 mM $Mg(OAc)_2$; pH 7.4) on ice for 1 hr in the dark. 2 µL of 20% Ficoll 400 was then added, and the entire sample was loaded immediately. For the FAM-U(G)17 RNA oligo, 24 picomol per reaction was used due to FAM being slightly quenched by proximal guanosines (a known caution provided by the RNA oligo manufacture). Gels were pre-run for 20 min at 100 volts (constant) with 0.5X TBM as the running buffer. Samples were loaded and gels were run for 40 min at 100 volts (constant) at room temperature in the dark. Gels in glass plates were then directly imaged using an Azure Sapphire Biomolecular Imager. The 5' FAM-labeled RNA oligo sequences are provided in the Supplementary Data. A single uridine was added as a spacer between the 5' FAM label and the polymeric guanosines to avoid quenching and was kept for consistency in the other labeled polymeric RNA oligos. We found flexible linkers did not help further avoid quenching by proximal guanosines.

Polysome analysis and RT-qPCR

In vitro translation reactions were scaled up 10-fold to 100 µL but were performed identically as described above. After 30 min at 30°C, reactions were kept on ice, diluted two-fold with ice-cold Polysome Dilution Buffer (10 mM Tris-HCl, 140 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 100 µg/mL cycloheximide; pH 7.4), and layered on top of a linear 10-50% buffered sucrose gradient (10 mM Tris-HCl, 140 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 100 µg/mL cycloheximide; pH 7.4) that was formed using a Biocomp Gradient Master. Gradients were centrifuged at 35K rpm for 120 min at 4°C in a SW-41Ti rotor (Beckman) with maximum acceleration and no brake using a Beckman Optima L-90 Ultracentrifuge. Gradients were subsequently fractionated into 0.9 mL volumes using a Biocomp piston fractionator with a TRIAX flow cell (Biocomp) recording a continuous $A_{260\ nm}$ trace. Total RNA was extracted from 400 µL of each fraction (spiked with 0.2 ng exogenous control FFLuc mRNA; Promega # L4561) by adding 600 µL TRIzol (Thermo Fisher # 15596018) and following the manufacturer's protocol. Glycogen (Thermo Fisher # R0561) was added at the isopropanol precipitation step. The resulting RNA pellet was resuspended in 30 µL nuclease-free water. 16 µL of extracted RNA was converted to cDNA using iScript Reverse Transcription Supermix for RT-qPCR (Bio-Rad # 1708841). cDNA reactions were then diluted 10-fold with nuclease-free water and stored at -20°C or used immediately. RT-qPCR was performed in 15 µL reactions using iTaq Universal SYBR Green Supermix (Bio-Rad # 1725124) in a Bio-Rad CFX Connect Real-Time PCR Detection System with 1.5 µL diluted cDNA and 250 nM (final concentration) primers. For each fraction, nLuc reporter mRNA abundance was normalized to the spiked-in control FFLuc mRNA using the Bio-Rad CFX Maestro software (ΔΔCt method). Abundance of total signal in each fraction was calculated using $Q_n = 2^{\Delta\Delta Ct}$ and $P = 100 \times Q_n/Q_{total}$ as previously described (30).

RESULTS

FMRP binds mRNA first to inhibit translation

For FMRP to inhibit translation elongation on select mRNAs, it is logical that it must bind target mRNAs before acting on a translating ribosome. High throughput sequencing with crosslinking immunoprecipitation of FMRP from mouse brain tissue revealed that FMRP binds to mRNA predominately in the CDS (21). However, these data are derived from multiple RBDs within FMRP, and it remains unclear which binding sites represent true FMRP translational repression targets (as opposed to mRNA transport or localization). Furthermore, recombinant FMRP can bind 80S ribosomes near the A site and 60S ribosomal subunits alone (24,25,31),

raising the possibility that FMRP can directly inhibit the ribosome independent of the mRNA sequence.

To determine if FMRP requires binding to target mRNA or the ribosome first to inhibit translation, we performed *in vitro* translation assays of a nanoLuciferase (nLuc) reporter mRNA with recombinant human FMRP and rabbit reticulocyte lysate (RRL) using different pre-incubation protocols. We chose to use the N-terminally truncated human FMRP (NT-hFMRP) (**Figure 1A, B**) because it is more stable than the full-length isoform and it retains translational repression activity (24,29). As a negative control, we programed RRL with reporter mRNA and protein storage buffer. FMRP was then either not pre-incubated, pre-incubated with RRL, or pre-incubated with reporter mRNA (**Figure 1C**). FMRP pre-incubation with RRL did inhibit translation to some degree (consistent with previous studies showing FMRP can bind to 60S subunits directly *in vitro*). However, FMRP pre-incubation with mRNA inhibited translation of reporter mRNA the greatest (**Figure 1C**), in agreeance with work from others (32). To further dissect this mechanism, we solely used the mRNA:FMRP pre-incubation strategy for all remaining experiments.

FMRP inhibits translation independent of mRNA G4s

Previous reports have shown that deletion of the RGG box motif from FMRP abolishes translational repression and ribosome binding *in vitro* (24). Dependence of the RGG box motif provided support to others that FMRP must target mRNAs and/or ribosomes by binding to intramolecular RNA G4s (24,25). When tested in isolation, the FMRP RGG box motif has high affinity for G-rich sequences that can form RNA G4 structures (15,17,33,34). Together, these data shaped the leading model that FMRP binds to RNA G4s in the CDS via the RGG box motif and then sterically blocks the A site of an elongating ribosome (24,25). This model has yet to be directly tested as most reporters used by the field harbor predicted G4 sequences and altering the sequence would mutate the reporter protein. To our knowledge, it has yet to be shown experimentally that RNA G4s are present and required in coding sequences for FMRP to inhibit translation.

To directly test this model, we generated a specialized nLuc reporter that either lacked or harbored an RNA G4 structure within the CDS (**Figure 2A**). Importantly, this nLuc nucleotide sequence was customized to lack predicted G4 structure without altering the amino acid sequence of the reporter. We also included a P2A ribosome skipping motif, which releases the nascent peptide but allows the ribosome to stay bound and continue elongation, directly upstream of the nLuc CDS. This allows uniform luciferase detection across both reporters. To experimentally confirm this reporter design, we took advantage of the selective G4 staining properties of N-methyl mesoporphyrin IX (NMM). As expected, the total RNA stain SYBR Green II detected both reporter mRNAs (**Figure 2B**, **Supplementary Figure S1**). However, NMM staining only detected the G4 reporter mRNA (**Figure 2B**). This same selective staining pattern of NMM for the G4 reporter mRNA was also seen in native PAGE (**Supplementary Figure S1**). These data support that the G4 reporter mRNA forms an intramolecular RNA G4 structure (as depicted in **Figure 2A**).

We next tested to what extent recombinant NT-hFMRP represses translation of control and G4 reporter mRNAs. If FMRP did in fact require RNA G4s on target transcripts to inhibit translation, we would expect enhanced repression of the G4 reporter mRNA. However, FMRP repressed translation of the G4 reporter mRNA slightly less than the control reporter mRNA (**Figure 2C**). These data indicate that human FMRP represses translation independent of mRNA G4s in the CDS.

The RGG box motif and CTD of FMRP together, but not independently, inhibit translation

We next sought to identify the critical RBD required for FMRP translational repression. FMRP can bind mRNA with at least three canonical RBDs (KH0, KH1, KH2) and a single RGG box motif. FXS patient mutations suggest that multiple regions of FMRP are critical for RNA binding-dependent function and contribute to pathology if mutated. The I304N patient mutation in the KH2 domain abolishes FMRP binding to polysomes in human cells (16). A guanosine insertion (ΔRGG+CTD) within the sequence that encodes the RGG box motif causes a frameshift and results in a truncated FMRP that lacks most of the RGG box motif and the entire C-terminal domain (CTD).

To further define the domain(s) that are critical for translational repression by FMRP, we purified recombinant NT-hFMRP harboring the I304N and ΔRGG+CTD mutations (**Figure 3A, B**) and tested their ability to inhibit translation. Multiple attempts were made to purify a KH1 domain G266E mutant (18), but we were consistently unable to recover soluble protein. We then determined the half-maximal inhibitory concentration (IC_{50}) of WT and each mutant NT-hFMRP (**Figure 3C-I**). As expected, the His6-MBP tag alone did not inhibit translation (**Figure 3C**). WT and I304N NT-hFMRP both inhibited translation in our assay, with the I304N mutant having a ~2-fold more potent IC_{50} than that of the WT isoform (**Figure 3D, E**). This suggests that although the I304N mutation alters FMRP binding to an optimal RNA pseudoknot substrate (i.e., Δkc2) and causes FMRP to dissociate from polysomes *in vivo* (16), the mutation does not interfere with translational repression.

Consistent with previous reports (24,25), the ΔRGG+CTD mutant did not repress translation (**Figure 3F**), but robust repression was observed with the isolated RGG+CTD region (**Figure 3G**). However, the RGG box alone nor the CTD alone were able to effectively inhibit translation (**Figure 3H, I**). Together, these data suggest that the RGG box must be present with the CTD to inhibit translation.

The RGG+CTD region forms a non-canonical RBD that binds reporter mRNA and multiple RNA sequences

We next sought to determine the RNA binding capability of the RGG+CTD region. We used electrophoretic mobility shift assays (EMSAs) to confirm recombinant NT-hFMRP (**Figure 3A**) can bind control reporter mRNA. WT and I304N NT-hFMRP, as well as the RGG+CTD fusion, do bind the reporter mRNA as they caused a complete gel shift (**Figure 4A**). This result is consistent with *in vitro* translational repression assays that show these three FMRP isoforms inhibit translation (**Figure 3D, E, G**). The RGG box motif alone and the CTD alone, neither of which inhibited translation (**Figure 3H, I**), were incapable of binding control reporter mRNA and did not cause a gel shift (**Figure 4A**). However, to our surprise, the ΔRGG+CTD mutant that did not inhibit translation (**Figure 3F**) caused a complete gel shift (**Figure 4A**). These data suggest that FMRP binding through the KH domain(s) does not elicit translational repression and the multiple RNA-binding domains and motifs provide a platform for FMRP to have multiple functions.

As an alternative approach to assess how FMRP binds mRNA, we used primer extension to determine if WT or mutant FMRP inhibits full length extension of reverse transcriptase from an annealed oligonucleotide (**Figure 4B**). Tightly bound FMRP would not be displaced by the reverse transcriptase, resulting in less full-length primer extension. Consistent with our translational repression results (**Figure 3**), only WT, I304N, and RGG+CTD blocked the formation of full-length extension products (**Figure 4C**). Although we calculated I304N to have the most potent IC_{50} (**Figure 3**), RGG+CTD inhibited translation the most of any recombinant FMRP isoform at the highest concentration tested (**Figure 3**). Similarly, RGG+CTD blocked full-

length primer extension the most of any FMRP isoform tested (**Figure 4C**). The RGG box motif alone and the CTD alone only marginally reduced full-length primer extension (**Figure 4C**). These data suggest that the RGG box motif and CTD must be together bind mRNA and form a ncRBD.

To further define how RGG+CTD acts as a ncRBD, we tested its ability to bind polymeric RNA sequences using EMSAs. As additional comparisons, we also tested the RGG box motif alone and the CTD alone. The first identified RGG box motif, belonging to heterogenous nuclear ribonucleoprotein (RNP) U, was found to bind both polymeric G and U RNA sequences *in vitro*, with a higher preference for polymeric G RNA (35). Consistent with the FMRP RGG box motif favoring G-rich sequences and G4s (15,36), the RGG box motif alone had some observable binding to $U(G)_{17}$ RNA, but little to no binding to $U(A)_{17}$, $U(C)_{17}$, or $(U)_{18}$ RNA (**Figure 5A-E**). Similar results were seen with the positively charged CTD (**Figure 5A-E**). Conversely, the RGG+CTD ncRBD robustly bound all four polymeric RNAs, with slightly less preference for $U(A)_{17}$ RNA (**Figure 5A-E**).

We next pursued deletion analysis to determine if the entire ncRBD was required for RNA binding and translational repression (**Figure 5F, G**). Similar to the RGG+CTD, the RGG+CTD Δ19 and RGG+CTD Δ39 truncation mutants bound all polymeric RNAs with slightly less preference for $U(A)_{17}$ RNA (**Figure 5H-L**). The RGG+CTD Δ55 truncated mutant did not stably bind $U(A)_{17}$ RNA and had reduced binding to $U(C)_{17}$, $(U)_{18}$, or $U(G)_{17}$ RNA (**Figure 5H-L**). The RGG+CTD Δ55 truncated mutant also bound less reporter mRNA compared to longer RGG+CTD isoforms and did not robustly inhibit translation (**Supplementary Figure S3**). It should be noted that we routinely did not see the RNP enter the gel for most of the polymeric RNAs tested, suggesting the high positive charge of the RGG+CTD ncRBD neutralized the negative charge of the RNA oligo and/or the FAM label was quenched. The RGG box motif and CTD have theoretical isoelectric points of 12.1 and 10.0, respectively. Nevertheless, to be as consistent as possible across all the samples, unbound polymeric RNA was quantified (**Figure 5E, L**). RGG+CTD bound to $U(G)_{17}$, and to some extent $(U)_{18}$, noticeably formed complexes that entered the gel. It is possible that the $U(G)_{17}$ formed intermolecular G4s, providing a high net negative charge to the complex that facilitated its entry into the gel if RGG+CTD was sub-stoichiometric. Together, these data support that the RGG box motif with the CTD forms a ncRBD that enables FMRP to bind mRNA independent of a consensus sequence and inhibit translation.

FMRP inhibits translation elongation when bound to mRNA with the non-canonical RBD

FMRP bound to mRNA could directly inhibit translation by either blocking the scanning pre-initiation complex (PIC) in the 5' UTR or stalling the elongating ribosome in the CDS. In general, scanning PICs are more susceptible and sensitive to obstacles in their path (i.e., RNA structure or bound RBPs) than elongating ribosomes. Most mRNA-bound FMRP is mapped to the CDS of mRNAs *in vivo*, not to the 5' UTR (21). However, in our assay, it is possible that some of the recombinant FMRP is bound to the 5' UTR and is simply blocking the scanning PIC. To confirm that FMRP inhibits translation elongation, we used the following three distinct strategies.

First, we rationalized that if 5' UTR-bound recombinant FMRP was blocking scanning PICs, extending the 5' UTR length would enhance repression. In this case, extending the 5' UTR would provide FMRP increased opportunity to bind the 5' UTR instead of the CDS. To achieve this, we mutated the control reporter that harbors the 50 nt human β-globin 5' UTR and extended the 5' UTR three-fold by adding two additional β-globin 5' UTR sequences (resulting in a 150 nt 5' UTR). Nevertheless, FMRP inhibited translation of the control and long 5' UTR

reporter mRNAs were equally (**Figure 6A**). This result suggests that FMRP does not inhibit scanning PICs but rather inhibits elongating ribosomes.

Second, we took advantage of translation initiation mechanisms that do not use scanning PICs. The Cricket Paralysis Virus Intergenic Region (CrPV IGR) Internal Ribosome Entry Site (IRES) directly recruits the 40S small ribosomal subunit to form an initiation complex independent of any initiation factors or the initiator tRNA (37-40) . FMRP robustly repressed translation of the PIC scanning-independent CrPV IGR IRES reporter mRNA (**Figure 6B**), providing further support that FMRP inhibits elongation and not scanning PICs in the 5' UTR.

Lastly, we used sucrose density gradient ultracentrifugation to confirm the expected distribution of mRNAs with stalled elongating ribosomes. For example, elongation inhibitors (e.g., cycloheximide or emetine) cause an increase in polysomes as they do not prevent initiation, but slow down and stabilize elongating ribosomes on mRNAs (41). This is typically seen as an increase in polysome signal in the heavier fractions of sucrose gradients. If FMRP inhibits elongating ribosomes when bound to mRNA via its ncRBD, we would predict that polysomes accumulate on inhibited transcripts only when FMRP harbored the RGG+CTD ncRBD. To test this prediction, we used the ΔRGG+CTD mutant as a negative control as it does not inhibit translation (**Figure 3F**) but still binds the reporter mRNA (**Figure 4A**) in our experiments. After *in vitro* translation and ultracentrifugation, we quantified nLuc reporter mRNA abundance in each fraction. The monosome peak was consistently found in fractions 4 and 5 (**Supplementary Figure S3A**), indicating that polysomes would sediment in fractions 6 through 12. Compared to the negative control, WT NT-hFMRP increased the abundance of nLuc mRNA in the heavy polysomes (**Figure 6C-D**, **Supplementary Figure S3B-E**). Specifically, nLuc mRNA abundance in the heaviest polysomes in fractions 10-12 increased ~2-fold (**Figure 6E**). In total, these data support that FMRP uses the ncRBD formed by the RGG box motif and CTD to bind mRNA independent of RNA G4s and stalls elongating ribosomes (**Figure 7**).

DISCUSSION

FMRP is a multifunctional protein that harbors many RNA-binding domains (RBD). These RBDs may function independently or work synergistically in combination(s) to provide FMRP with different functions and binding preferences. The complexity of FMRP has led to much debate surrounding how it targets transcripts and what functions it may possess once bound to mRNA. Most early data support FMRP as a translational repressor (21-25), but recent work has provided compelling evidence that FMRP may in fact be a translational activator (26,27). It remains largely unclear which RNA motifs or structures FMRP targets as few common enriched sequences were found across multiple transcriptome-wide FMRP binding studies (42). Additionally, a single consensus sequence for FMRP may be an oversimplification since FMRP has multiple RNA-binding domains and motifs. The complex nature of FMRP has led to a longstanding deficit in understanding its function(s) when bound to mRNA (42).

In this study, we address these areas of debate surrounding FMRP-mediated translational control. Our work here provides evidence that FMRP inhibits translation elongation after binding to target transcripts using a ncRBD formed by the RGG box motif and CTD (**Figure 7**). The ncRBD binds all four polymeric RNA sequences, with slightly less preference for poly-A sequences. Conversely, the individual RGG box motif and CTD alone have some ability to bind poly-G RNA (**Figure 5**) but are unable to inhibit translation (**Figure 3**). Taken together, these data provide some clarity about how FMRP controls translation and how FMRP function may be dependent on which domain(s) are used to bind to transcripts (**Figure 7**). Cryo-EM determination of Drosophila FMRP•80S ribosome complexes show that the KH1 domain contacts RPL5 and that the KH1 and KH2 domains at least partially overlap with where the

anticodon stem of a bound tRNA would be in the P site (24). In this structure, the RGG+CTD region of FMRP is near the A site and leading edge of the ribosome (24). However, the exact inhibitory role of the KH1 and KH2 domains when bound to the ribosome is unclear as our data (**Figure 3**) and others show the RGG+CTD is sufficient for translational repression (25). Future structural determination of FMRP-inhibited ribosomes on mRNA will be critical for fully understanding FMRP-mediated translational repression.

A common successful approach used throughout RNA biology is to isolate conserved, structured RBDs to determine their binding preferences *in vitro* or *in vivo*. This approach has been used to parse through how FMRP targets transcripts. For example, *in vitro* selection experiments with recombinant forms of the FMRP RGG box motif identified the G4-forming sc1 RNA as a high affinity binding substrate (17,34). Filter binding assays and fluorescence anisotropy experiments support these conclusions (15,17,33). However, it remains rather ambiguous as subsequent structural determinations have shown that the RGG box motif binds to the major groove of the duplexed RNA region near the dsRNA:G4 junction and to the mixed tetrad of the artificial sc1 RNA but not to the G4 itself (33,34). Nevertheless, recombinant FMRP is unable to repress translation if the RGG box motif is deleted, suggesting that RNA G4s are required for translational control (24). Our data show that FMRP, when binding mRNA through the ncRBD formed by the RGG box motif + CTD, represses translation elongation independent of RNA G4s within the target transcript CDS. In support of our conclusions, recent ribosome profiling studies that incorporated RNA G4 prediction analyses show that transcripts that are derepressed in FMRP knockout cell lines are not enriched for RNA G4s (43). In this study, RNA G4s were instead correlated with FMRP-mediated mRNA localization, not translational repression. Together, our data highlight the inherent limitation of identifying the function of small RNA-binding motifs in isolation and that unstructured RNA-binding motifs like the FMRP RGG box motif are influenced by their surrounding protein context.

It is known that FMRP facilitates the transport of dendritic mRNAs to synapses *in vivo* and this transport can be dependent on its ability to bind mRNA (43,44). However, our data show that FMRP can bind multiple RNA sequences and when bound to mRNA can block translation elongation. Together, these data support a model where FMRP regulates translation of dendritic mRNAs based on local protein concentration, not mRNA sequence. Since FMRP and dendritic mRNAs are transported to the synapses, often together, both are found at relatively high concentrations compared to other translationally active sites of neurons. In FXS where FMRP is essentially lost, multiple cellular phenotypes will be present as FMRP functions differently depending on the cell type, its subcellular localization, how it binds mRNA (e.g., with which domain/motif), and its post-translational modification state (42).

FUNDING

MRS was supported by The Ohio State University Fellowship and The Ohio State University Center for RNA Biology Graduate Fellowship. This work was supported by NIH grant R00GM126064 to MGK. The Ohio State University Comprehensive Cancer Center Genomics Shared (OSUCCC GSR) is supported by NIH grant P30CA016058.

ACKNOWLEDGEMENTS

Experiments were conceived by MRS and MGK and were performed by MRS, JEW, and MGK. The manuscript was written by MRS and MGK with input from JEW. We thank members of the Kearse lab for critically reading the manuscript. We also thank Brian Scarpitti for his help

with Image J analyses and Christine Daugherty at OSUCCC GSR for her critical help with the fluorescent primer extension analysis.

REFERENCES

1. 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 (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, **65**, 905-914.
2. Kremer, E.J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S.T., Schlessinger, D., Sutherland, G.R. and Richards, R.I. (1991) Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science*, **252**, 1711-1714.
3. Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M.F. and Mandel, J.L. (1991) Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science*, **252**, 1097-1102.
4. Yu, S., Pritchard, M., Kremer, E., Lynch, M., Nancarrow, J., Baker, E., Holman, K., Mulley, J.C., Warren, S.T., Schlessinger, D. *et al.* (1991) Fragile X genotype characterized by an unstable region of DNA. *Science*, **252**, 1179-1181.
5. Niu, M., Han, Y., Dy, A.B.C., Du, J., Jin, H., Qin, J., Zhang, J., Li, Q. and Hagerman, R.J. (2017) Autism Symptoms in Fragile X Syndrome. *J Child Neurol*, **32**, 903-909.
6. Kelleher, R.J., 3rd and Bear, M.F. (2008) The autistic neuron: troubled translation? *Cell*, **135**, 401-406.
7. Ashley, C.T., Jr., Wilkinson, K.D., Reines, D. and Warren, S.T. (1993) FMR1 protein: conserved RNP family domains and selective RNA binding. *Science*, **262**, 563-566.
8. Siomi, H., Siomi, M.C., Nussbaum, R.L. and Dreyfuss, G. (1993) The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell*, **74**, 291-298.
9. Myrick, L.K., Hashimoto, H., Cheng, X. and Warren, S.T. (2015) Human FMRP contains an integral tandem Agenet (Tudor) and KH motif in the amino terminal domain. *Hum Mol Genet*, **24**, 1733-1740.
10. Hollingworth, D., Candel, A.M., Nicastro, G., Martin, S.R., Briata, P., Gherzi, R. and Ramos, A. (2012) KH domains with impaired nucleic acid binding as a tool for functional analysis. *Nucleic Acids Res*, **40**, 6873-6886.
11. Nicastro, G., Taylor, I.A. and Ramos, A. (2015) KH-RNA interactions: back in the groove. *Curr Opin Struct Biol*, **30**, 63-70.
12. Jensen, K.B., Musunuru, K., Lewis, H.A., Burley, S.K. and Darnell, R.B. (2000) The tetranucleotide UCAY directs the specific recognition of RNA by the Nova K-homology 3 domain. *Proc Natl Acad Sci U S A*, **97**, 5740-5745.
13. Lewis, H.A., Musunuru, K., Jensen, K.B., Edo, C., Chen, H., Darnell, R.B. and Burley, S.K. (2000) Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. *Cell*, **100**, 323-332.
14. Nicastro, G., Candel, A.M., Uhl, M., Oregioni, A., Hollingworth, D., Backofen, R., Martin, S.R. and Ramos, A. (2017) Mechanism of beta-actin mRNA Recognition by ZBP1. *Cell Rep*, **18**, 1187-1199.
15. Athar, Y.M. and Joseph, S. (2020) RNA-Binding Specificity of the Human Fragile X Mental Retardation Protein. *J Mol Biol*, **432**, 3851-3868.
16. Darnell, J.C., Fraser, C.E., Mostovetsky, O., Stefani, G., Jones, T.A., Eddy, S.R. and Darnell, R.B. (2005) Kissing complex RNAs mediate interaction between the Fragile-X mental retardation protein KH2 domain and brain polyribosomes. *Genes Dev*, **19**, 903-918.

17. Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T. and Darnell, R.B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell*, **107**, 489-499.
18. Myrick, L.K., Nakamoto-Kinoshita, M., Lindor, N.M., Kirmani, S., Cheng, X. and Warren, S.T. (2014) Fragile X syndrome due to a missense mutation. *Eur J Hum Genet*, **22**, 1185-1189.
19. De Boulle, K., Verkerk, A.J., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., Van den Bos, F., de Graaff, E., Oostra, B.A. and Willems, P.J. (1993) A point mutation in the FMR-1 gene associated with fragile X mental retardation. *Nat Genet*, **3**, 31-35.
20. Okray, Z., de Esch, C.E., Van Esch, H., Devriendt, K., Claeys, A., Yan, J., Verbeeck, J., Froyen, G., Willemsen, R., de Vrij, F.M. *et al.* (2015) A novel fragile X syndrome mutation reveals a conserved role for the carboxy-terminus in FMRP localization and function. *EMBO Mol Med*, **7**, 423-437.
21. Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W. *et al.* (2011) FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell*, **146**, 247-261.
22. Shah, S., Molinaro, G., Liu, B., Wang, R., Huber, K.M. and Richter, J.D. (2020) FMRP Control of Ribosome Translocation Promotes Chromatin Modifications and Alternative Splicing of Neuronal Genes Linked to Autism. *Cell Rep*, **30**, 4459-4472 e4456.
23. Liu, B., Li, Y., Stackpole, E.E., Novak, A., Gao, Y., Zhao, Y., Zhao, X. and Richter, J.D. (2018) Regulatory discrimination of mRNAs by FMRP controls mouse adult neural stem cell differentiation. *Proc Natl Acad Sci U S A*, **115**, E11397-E11405.
24. Chen, E., Sharma, M.R., Shi, X., Agrawal, R.K. and Joseph, S. (2014) Fragile X mental retardation protein regulates translation by binding directly to the ribosome. *Mol Cell*, **54**, 407-417.
25. Athar, Y.M. and Joseph, S. (2020) The Human Fragile X Mental Retardation Protein Inhibits the Elongation Step of Translation through Its RGG and C-Terminal Domains. *Biochemistry*, **59**, 3813-3822.
26. Greenblatt, E.J. and Spradling, A.C. (2018) Fragile X mental retardation 1 gene enhances the translation of large autism-related proteins. *Science*, **361**, 709-712.
27. Shu, H., Donnard, E., Liu, B., Jung, S., Wang, R. and Richter, J.D. (2020) FMRP links optimal codons to mRNA stability in neurons. *Proc Natl Acad Sci U S A*, **117**, 30400-30411.
28. Kearse, M.G., Green, K.M., Krans, A., Rodriguez, C.M., Linsalata, A.E., Goldstrohm, A.C. and Todd, P.K. (2016) CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Mol Cell*, **62**, 314-322.
29. Edwards, M., Xu, M. and Joseph, S. (2020) A simple procedure for bacterial expression and purification of the fragile X protein family. *Sci Rep*, **10**, 15858.
30. Pringle, E.S., McCormick, C. and Cheng, Z. (2019) Polysome Profiling Analysis of mRNA and Associated Proteins Engaged in Translation. *Curr Protoc Mol Biol*, **125**, e79.
31. Siomi, M.C., Zhang, Y., Siomi, H. and Dreyfuss, G. (1996) Specific sequences in the fragile X syndrome protein FMR1 and the FXR proteins mediate their binding to 60S ribosomal subunits and the interactions among them. *Mol Cell Biol*, **16**, 3825-3832.
32. Li, Z., Zhang, Y., Ku, L., Wilkinson, K.D., Warren, S.T. and Feng, Y. (2001) The fragile X mental retardation protein inhibits translation via interacting with mRNA. *Nucleic Acids Res*, **29**, 2276-2283.
33. Vasilyev, N., Polonskaia, A., Darnell, J.C., Darnell, R.B., Patel, D.J. and Serganov, A. (2015) Crystal structure reveals specific recognition of a G-quadruplex RNA by a beta-turn in the RGG motif of FMRP. *Proc Natl Acad Sci U S A*, **112**, E5391-5400.

34. Phan, A.T., Kuryavyi, V., Darnell, J.C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D. *et al.* (2011) Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nat Struct Mol Biol*, **18**, 796-804.
35. Kiledjian, M. and Dreyfuss, G. (1992) Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. *EMBO J*, **11**, 2655-2664.
36. Brown, V., Small, K., Lakkis, L., Feng, Y., Gunter, C., Wilkinson, K.D. and Warren, S.T. (1998) Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein. *J Biol Chem*, **273**, 15521-15527.
37. Pestova, T.V., Lomakin, I.B. and Hellen, C.U. (2004) Position of the CrPV IRES on the 40S subunit and factor dependence of IRES/80S ribosome assembly. *EMBO Rep*, **5**, 906-913.
38. Pestova, T.V. and Hellen, C.U. (2003) Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes Dev*, **17**, 181-186.
39. Wilson, J.E., Pestova, T.V., Hellen, C.U. and Sarnow, P. (2000) Initiation of protein synthesis from the A site of the ribosome. *Cell*, **102**, 511-520.
40. Jan, E., Kinzy, T.G. and Sarnow, P. (2003) Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. *Proc Natl Acad Sci U S A*, **100**, 15410-15415.
41. Juszkiewicz, S., Slodkowicz, G., Lin, Z., Freire-Pritchett, P., Peak-Chew, S.Y. and Hegde, R.S. (2020) Ribosome collisions trigger cis-acting feedback inhibition of translation initiation. *Elife*, **9**.
42. Richter, J.D. and Zhao, X. (2021) The molecular biology of FMRP: new insights into fragile X syndrome. *Nat Rev Neurosci*, **22**, 209-222.
43. Goering, R., Hudish, L.I., Guzman, B.B., Raj, N., Bassell, G.J., Russ, H.A., Dominguez, D. and Taliaferro, J.M. (2020) FMRP promotes RNA localization to neuronal projections through interactions between its RGG domain and G-quadruplex RNA sequences. *Elife*, **9**.
44. Ramos, A., Hollingworth, D. and Pastore, A. (2003) G-quartet-dependent recognition between the FMRP RGG box and RNA. *RNA*, **9**, 1198-1207.

FIGURE LEGENDS

Figure 1. FMRP first binds mRNA to inhibit translation. A) Schematic of MBP and His6-tagged WT N-terminally truncated human FMRP (NT-hFMRP). B) Coomassie stain of recombinant WT NT-hFMRP. C) *In vitro* translation of nLuc reporter mRNA with protein storage buffer as a negative control, with WT NT-hFMRP without a pre-incubation step, with WT NT-hFMRP and RRL pre-incubated together, or with WT NT-hFMRP and nLuc mRNA pre-incubated together. Pre-incubations were performed on ice for 60 min. Data are shown as mean ± SD. n = 3 biological replicates.

Figure 2. FMRP inhibits translation independent of mRNA G-quadruplexes in the CDS. A) Schematic of custom nLuc reporters either lacking a G4 (control reporter) or harboring a G4 in the coding sequence (G4 reporter). A P2A ribosome skipping motif was included immediately upstream of the nLuc coding sequence to ensure consistent nLuc activity. B) Denaturing PAGE of control and G4 reporters stained for total RNA with SYBR Green II or for G4 structures with NMM. C) *In vitro* translation of control and G4 reporter mRNA with protein buffer or WT NT-hFMRP. Data are shown as mean ± SD. n = 3 biological replicates.

Figure 3. The RGG box motif and CTD together are essential and sufficient to inhibit translation. A) Schematic of recombinant WT and mutant NT-hFMRP. Mutated/truncated domains are highlighted in red. B) Coomassie stain of recombinant proteins. C-I) *In vitro* translation of nLuc mRNA with a titration of recombinant wildtype and mutant NT-hFMRP isoforms. IC_{50} values were determined for His6-MBP tag (negative control) (C), WT (D), I304N KH2 domain patient-derived mutant (E), and ΔRGG+CTD mutant (F) NT-hFMRP. IC_{50} values were also determined for the RGG box motif + CTD fusion (G), the RGG box motif alone (H), and the CTD alone (I). Data are shown as the mean, n=3 biological replicates. A non-linear regression was used to calculate the IC_{50} and is shown as the line with the 95% confidence interval (CI) included as a watermark. The IC_{50} is reported ± 95% CI.

Figure 4. The RGG box motif requires its natural CTD to bind reporter mRNA. A) Native PAGE of control reporter mRNA incubated with recombinant protein and stained with SYBR Green II. B) Schematic of primer extension assay. After incubating control reporter mRNA and recombinant protein, a 5' FAM-labeled primer anneals to the 3' end of the mRNA and AMV-RT is used for reverse transcription. Stable RNA:protein interactions will block the formation of the full length 701 nt product. C) Primer extension assay of the control reporter mRNA incubated with the indicated recombinant protein. Data are shown as mean ± SD. n = 3 biological replicates.

Figure 5. The RGG+CTD non-canonical RBD binds all four polymeric RNAs. A-E) EMSA of 5' FAM-labeled polymeric RNA oligonucleotides with the indicated recombinant proteins. F-G) Schematic (F) and Coomassie stain (G) of RGG+CTD truncation mutants with 19, 39, or 55 amino acids deleted from the CTD. H-L) EMSA of 5' FAM-labeled polymeric RNA oligonucleotides with the indicated recombinant proteins. Due to the suspected net neutral charge of the RNA oligo:FMRP complex causing the RNP to not enter the gel, the intensity of unbound RNA was quantified and set relative to the buffer only sample (E, L).

Figure 6. Human FMRP inhibits translation elongation when binding mRNA via the non-canonical RBD. A) *In vitro* translation of the control reporter mRNA that harbors the 50 nt human β-globin 5' UTR and a long 5' UTR reporter mRNA that harbors three β-globin 5' UTRs in tandem (150 nt total). mRNPs were formed with protein storage buffer or WT NT-hFMRP. Data are shown as mean ± SD. n = 3 biological replicates. B) *In vitro* translation of a control reporter mRNA and a Cricket Paralysis Virus Intergenic Region (CrPV IGR) Internal Ribosome Entry Site (IRES) reporter mRNA. The CrPV IGR IRES is cap- and scanning-independent and directly recruits the ribosome for initiation. mRNPs were formed with protein storage buffer or WT NT-hFMRP. Data are shown as mean ± SD. n = 3 biological replicates. C) Distribution of nLuc reporter mRNA across sucrose gradients to assess polysome formation when pre-incubated with ΔRGG+CTD (Control) or WT NT-hFMRP. Abundance of reporter mRNA in each gradient fraction was determined by RT-qPCR. D) Cumulative nLuc abundance in heavy polysomes in fractions 10-12 from ΔRGG+CTD (Control) and WT NT-hFMRP samples. E) Fold change of nLuc mRNA abundance in heavy polysomes. Bar represents the mean of datum from three biological replicates.

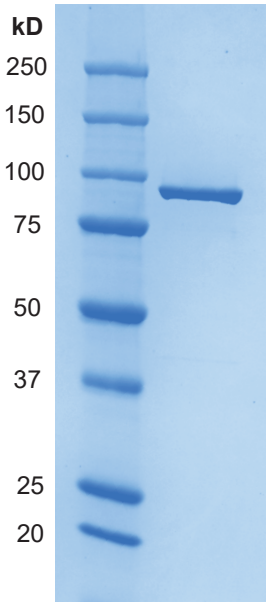
Figure 7. Model of FMRP-mediated inhibition of translation elongation. FMRP can bind mRNA through canonical and non-canonical RBDs. When bound to mRNA through its canonical KH domain(s), FMRP does not robustly inhibit translation. The RGG box motif and CTD of FMRP together form a ncRBD that allows FMRP to bind multiple mRNA sequences and subsequently inhibit translation elongation.

Figure 1

A

WT NT-hFMRP
MBP
KH1
KH2
RGG
C-terminal domain
His6

B


kD
250
150
100
75
50
37
25
20

C

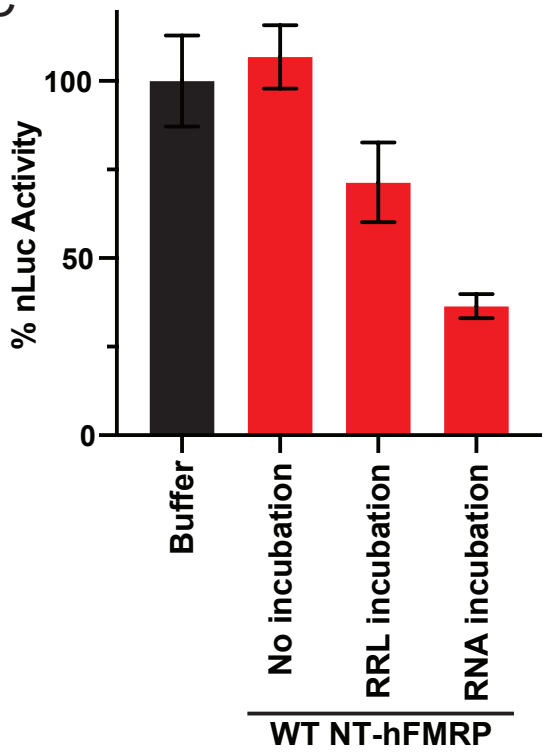

% nLuc Activity
100
50
0
Buffer
No incubation
RRL incubation
RNA incubation
WT NT-hFMRP

Figure 2

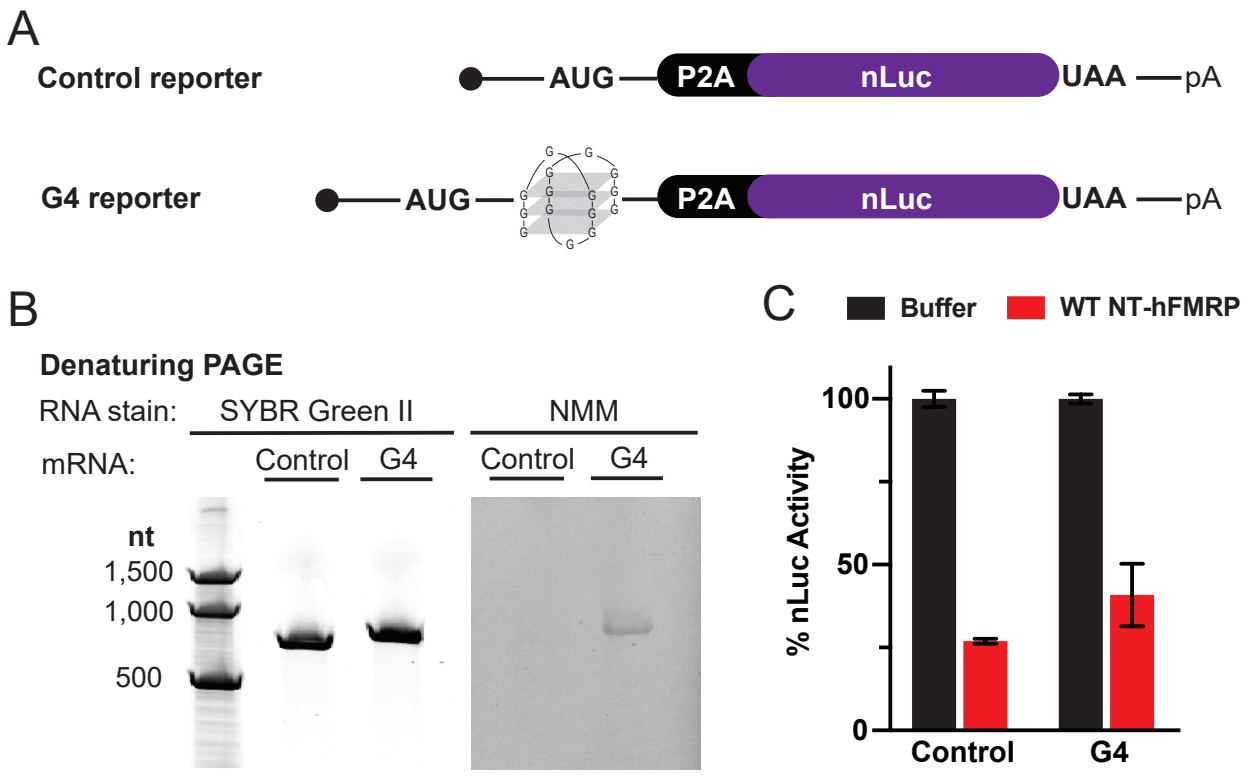

A
Control reporter
AUG
P2A
nLuc
UAA
pA
G4 reporter
AUG
P2A
nLuc
UAA
pA
B
Denaturing PAGE
RNA stain:
SYBR Green II
NMM
mRNA:
Control
G4
Control
G4
nt
1,500
1,000
500
C
Buffer
WT NT-hFMRP
% nLuc Activity
100
50
0
Control
G4

Figure 3

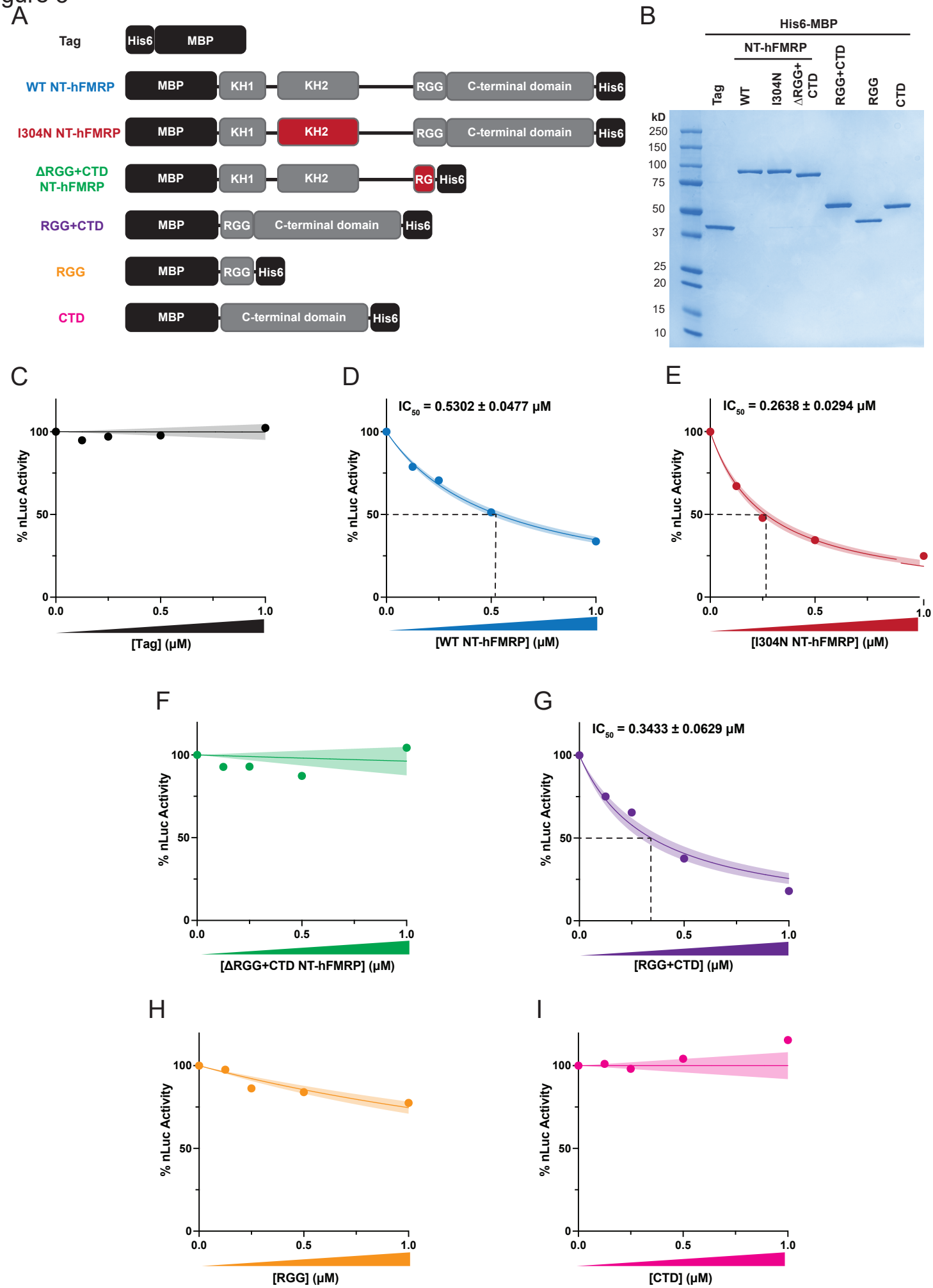
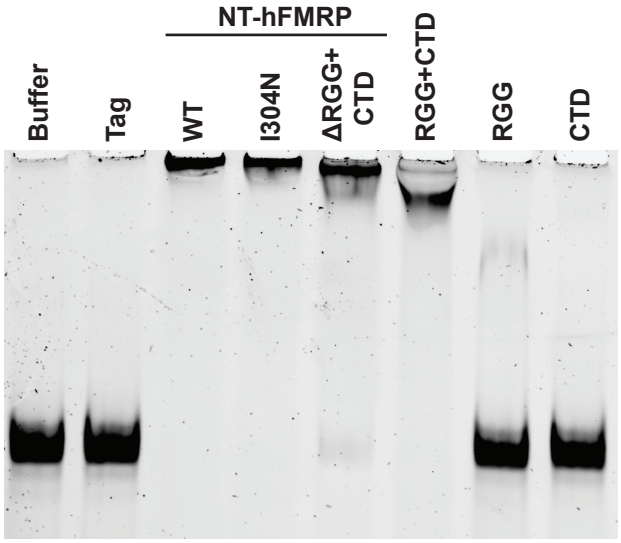
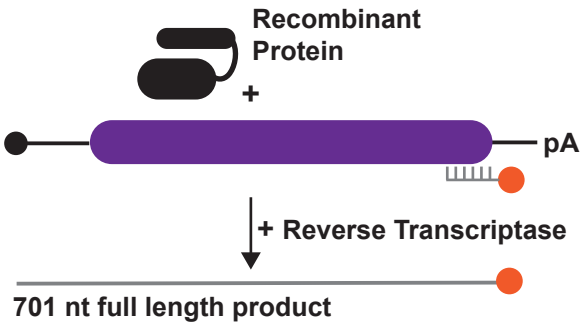

A
Tag
His6
MBP
WT NT-hFMRP
MBP
KH1
KH2
RGG
C-terminal domain
His6
I304N NT-hFMRP
ΔRGG+CTD NT-hFMRP
RG
RGG+CTD
RGG
CTD
B
His6-MBP
NT-hFMRP
Tag
WT
I304N
ΔRGG+ CTD
RGG+CTD
RGG
CTD
kD
250
150
100
75
50
37
25
20
15
10
C
D
E
F
G
H
I
% nLuc Activity
[Tag] (μM)
IC50 = 0.5302 ± 0.0477 μM
[WT NT-hFMRP] (μM)
IC50 = 0.2638 ± 0.0294 μM
[I304N NT-hFMRP] (μM)
[ΔRGG+CTD NT-hFMRP] (μM)
IC50 = 0.3433 ± 0.0629 μM
[RGG+CTD] (μM)
[RGG] (μM)
[CTD] (μM)

Figure 4

A


NT-hFMRP
Buffer
Tag
WT
I304N
ΔRGG+ CTD
RGG+CTD
RGG
CTD

B


Recombinant Protein
+
pA
+ Reverse Transcriptase
701 nt full length product

C

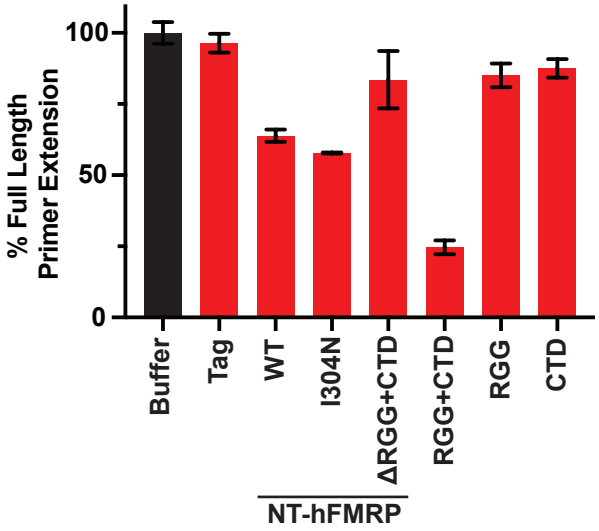

% Full Length Primer Extension
100
50
0
Buffer
Tag
WT
I304N
ΔRGG+CTD
RGG+CTD
RGG
CTD
NT-hFMRP

Figure 5

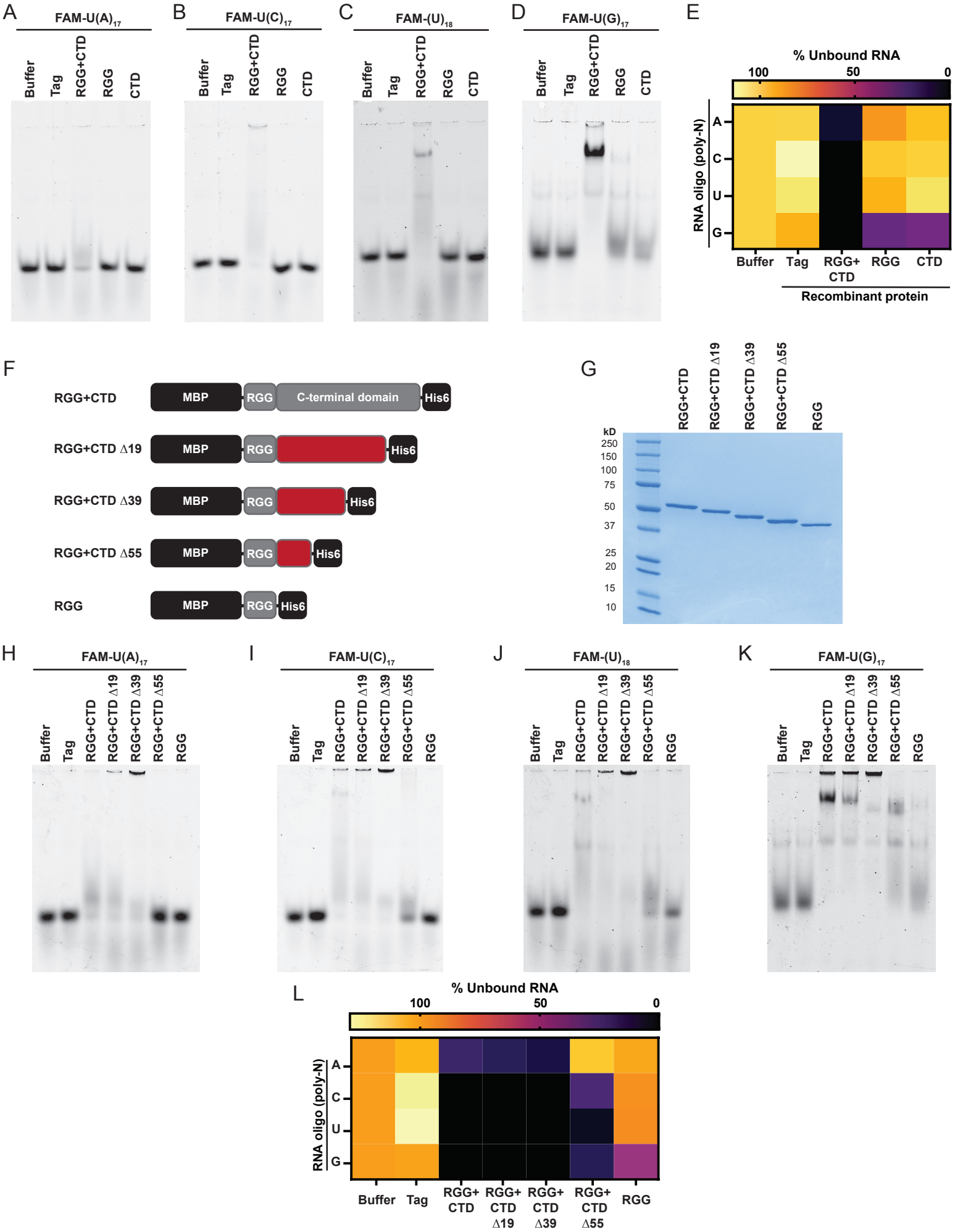

A
FAM-U(A)17
Buffer
Tag
RGG+CTD
RGG
CTD
B
FAM-U(C)17
C
FAM-(U)18
D
FAM-U(G)17
E
% Unbound RNA
100
50
0
RNA oligo (poly-N)
A
C
U
G
Buffer
Tag
RGG+ CTD
RGG
CTD
Recombinant protein
F
RGG+CTD
MBP
RGG
C-terminal domain
His6
RGG+CTD Δ19
RGG+CTD Δ39
RGG+CTD Δ55
RGG
G
RGG+CTD
RGG+CTD Δ19
RGG+CTD Δ39
RGG+CTD Δ55
RGG
kD
250
150
100
75
50
37
25
20
15
10
H
FAM-U(A)17
I
FAM-U(C)17
J
FAM-(U)18
K
FAM-U(G)17
L
% Unbound RNA
100
50
0
RNA oligo (poly-N)
A
C
U
G
Buffer
Tag
RGG+ CTD
RGG+ CTD Δ19
RGG+ CTD Δ39
RGG+ CTD Δ55
RGG

Figure 6

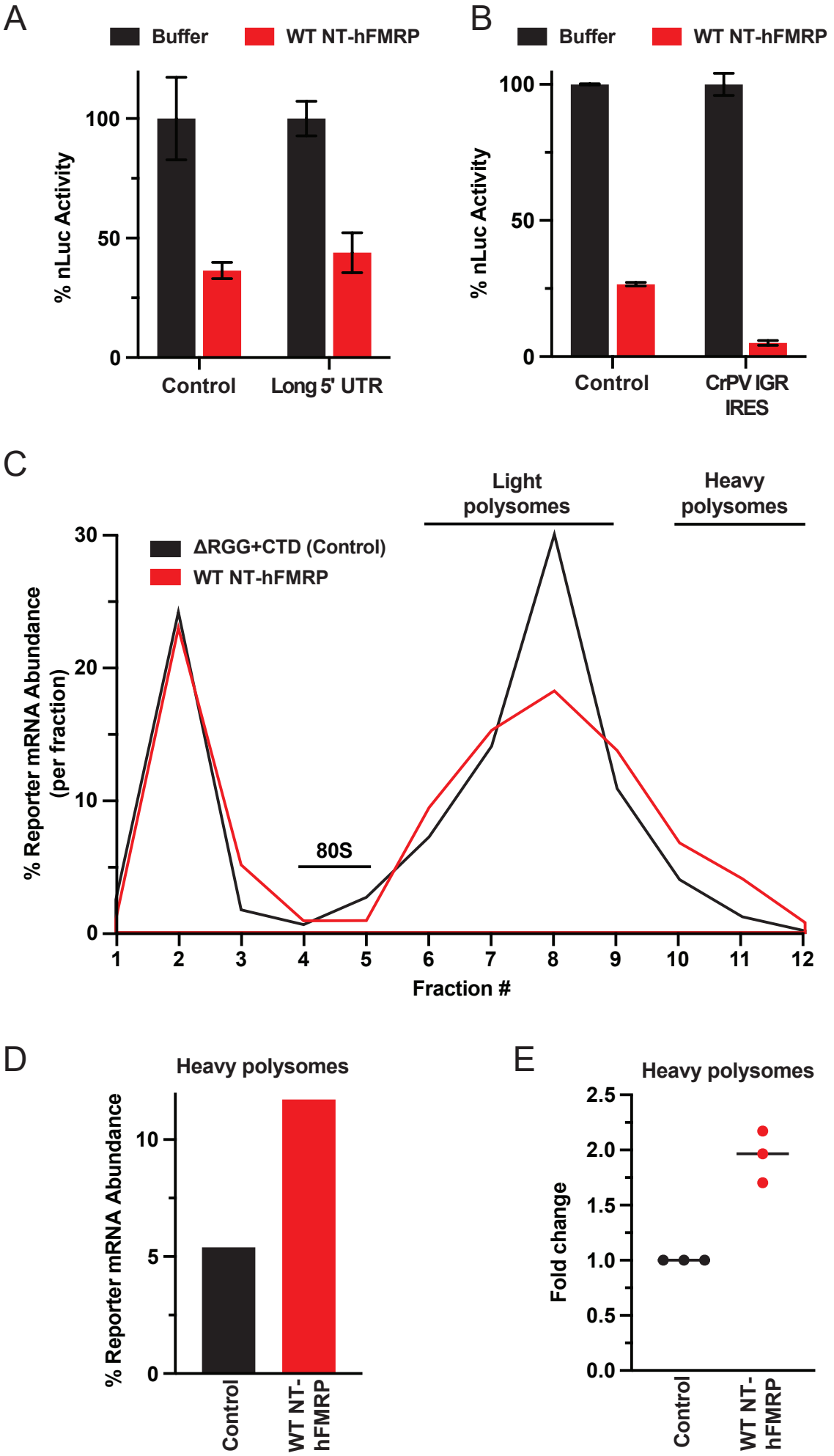
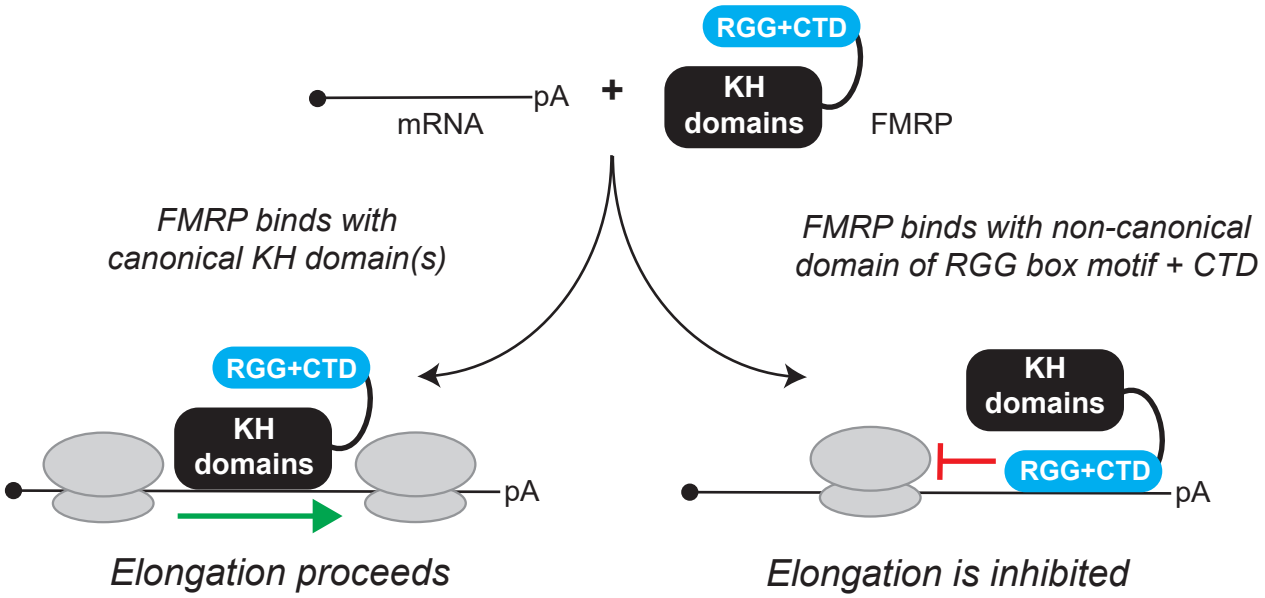

A
Buffer
WT NT-hFMRP
% nLuc Activity
100
50
0
Control
Long 5' UTR
B
Buffer
WT NT-hFMRP
% nLuc Activity
100
50
0
Control
CrPV IGR IRES
C
Light polysomes
Heavy polysomes
ΔRGG+CTD (Control)
WT NT-hFMRP
% Reporter mRNA Abundance (per fraction)
30
20
10
0
80S
1
2
3
4
5
6
7
8
9
10
11
12
Fraction #
D
Heavy polysomes
% Reporter mRNA Abundance
10
5
0
Control
WT NT-hFMRP
E
Heavy polysomes
Fold change
2.5
2.0
1.5
1.0
0.5
0.0
Control
WT NT-hFMRP

Figure 7


RGG+CTD
KH domains
FMRP
pA
mRNA
+
FMRP binds with canonical KH domain(s)
FMRP binds with non-canonical domain of RGG box motif + CTD
RGG+CTD
KH domains
pA
Elongation proceeds
KH domains
RGG+CTD
pA
Elongation is inhibited

SUPPLEMENTARY DATA

Table S1. Oligonucleotides used in this study.

Name	Sequence (5'-3')
R_5' FAM nLuc	/56-FAM/TTACTTGTCATCGTCATCCT
F_nLuc RT-qPCR	ATCAAGTGCTCGAACAAGGC
R_nLuc RT-qPCR	CAATTTTCAGGCCGTTTTCGC
F_Control FFLuc RT-qPCR	AACGCTTCCATCTTCCAGGG
R_Control FFLuc RT-qPCR	CCAGATCCACAACCTTCGCT

Supplementary Figure Legends

Supplementary Figure S1. G4 reporter mRNA, but not control reporter mRNA, harbors a G-quadruplex. Native PAGE of control and G4 reporter mRNAs stained for total RNA with SYBR Green II or stained for G4 structures with NMM.

Supplementary Figure S2. The RGG+CTD Δ55 truncation mutant binds less reporter mRNA compared to longer RGG+CTD isoforms and does not inhibit translation. A) EMSA of control reporter mRNA with the indicated recombinant proteins. B) Signal intensity of samples in panel A. C) *In vitro* translation of the control reporter mRNA when pre-incubated with protein storage buffer or the indicated recombinant protein. Data are shown as mean ± SD. n = 3 biological replicates.

Supplementary Figure S3. WT NT-hFMRP causes heavy polysomes to accumulate. A) A representative $A_{260\ nm}$ trace of sucrose gradient fractionation (polysome analysis) of *in vitro* translation reactions. The 80S monosome sediments in fractions 4 and 5. Thus, fractions 6-12 contain polysomes. Robust polysome curves are not visualized as mRNA input amounts were limited to ensure responses were in the linear dynamic range (see Materials and Methods). B-E) Separate biological replicates of experiments shown in **Figure 6E**. B & D) Distribution of reporter mRNA across sucrose gradients to assess polysome formation with ΔRGG+CTD (Control) or WT NT-hFMRP. Abundance of reporter mRNA in each gradient fraction was determined by RT-qPCR. C & E) Cumulative nLuc abundance in heavy polysomes in fractions 10-12 from ΔRGG+CTD (Control) and WT NT-hFMRP samples.

Reporter Sequences

T7 promoter
Human beta globin 5' UTR
Heterologous 5' UTR
Start codon
P2A ribosome skipping motif
nLuc
3xFLAG
Stop codon
G4 insert
CrPV IGR IRES

Control nLuc reporter

*TAATACGACTCACTATAG*GGACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAA
CAGACACCATGGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAaGCTGGAGACGTcG
AaGAaAACCCTGGACCTGTGTTTACCCTtGAAGATTTTGTcGGCGATTGGCGCCAGACCGCG
cGCTATAACCTcGATCAaGTGCTcGAACAaGGCGGCGTGAGCAGCCTGTTTCAGAACCTGG
GCGTGAGCGTGACCCCGATTCAGCGCATTGTGCTGAGCGGCGAAAACGGCCTGAAAATTG
ATATTCATGTGATTATTCCGTATGAAGGCCTGAGCGGCGATCAGATGGGCCAGATTGAAAA
AATTTTTAAAGTcGTaTATCCtGTcGATGATCATCATTTTAAAGTGATTCTGCATTATGGCACC
CTGGTGATTGATGGCGTGACCCCGAACATGATTGATTATTTTGGCCGCCCGTATGAAGGCA
TTGCtGTtTTTGATGGCAAgAAgATTACCGTGACCGGCACCCTGTGGAACGGCAACAAAATTA
TTGATGAACGCCTGATTAACCCaGATGGCAGCCTGCTGTTTCGCGTGACCATTAACGGCGT
GACCGGCTGGCGCCTGTGCGAACGCATTCTcGCaGACTACAAAGACCATGACGGTGATTAT
AAAGATCATGACATCGATTACAAGGATGACGATGACAAGTAA

G4 nLuc reporter
*TAATACGACTCACTATAG*GGACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAA
CAGACACCATGCAACAACAACAACAAGGGGGGGGGGGGGGCAACAACAACAACAAGGA
AGCGGAGCTACTAACTTCAGCCTGCTGAAGCAaGCTGGAGACGTcGAaGAaAACCCTGGAC
CTGTGTTTACCCTtGAAGATTTTGTcGGCGATTGGCGCCAGACCGCGcGCTATAACCTcGAT
CAaGTGCTcGAACAaGGCGGCGTGAGCAGCCTGTTTCAGAACCTGGGCGTGAGCGTGACC
CCGATTCAGCGCATTGTGCTGAGCGGCGAAAACGGCCTGAAAATTGATATTCATGTGATTA
TTCCGTATGAAGGCCTGAGCGGCGATCAGATGGGCCAGATTGAAAAAATTTTTAAAGTcGT
aTATCCtGTcGATGATCATCATTTTAAAGTGATTCTGCATTATGGCACCCTGGTGATTGATGG
CGTGACCCCGAACATGATTGATTATTTTGGCCGCCCGTATGAAGGCATTGCtGTtTTTGATG
GCAAgAAgATTACCGTGACCGGCACCCTGTGGAACGGCAACAAAATTATTGATGAACGCCT
GATTAACCCaGATGGCAGCCTGCTGTTTCGCGTGACCATTAACGGCGTGACCGGCTGGCG
CCTGTGCGAACGCATTCTcGCaGACTACAAAGACCATGACGGTGATTATAAAGATCATGAC
ATCGATTACAAGGATGACGATGACAAGTAA

Long 5' UTR nLuc reporter
*TAATACGACTCACTATAG*GGACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAA
CAGACACCACATTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCACA
TTTGCTTCTGACACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGGAAGCGGAGC
TACTAACTTCAGCCTGCTGAAGCAaGCTGGAGACGTcGAaGAaAACCCTGGACCTGTGTTTA
CCCTtGAAGATTTTGTcGGCGATTGGCGCCAGACCGCGcGCTATAACCTcGATCAaGTGCTc
GAACAaGGCGGCGTGAGCAGCCTGTTTCAGAACCTGGGCGTGAGCGTGACCCCGATTCA
GCGCATTGTGCTGAGCGGCGAAAACGGCCTGAAAATTGATATTCATGTGATTATTCCGTAT
GAAGGCCTGAGCGGCGATCAGATGGGCCAGATTGAAAAAATTTTTAAAGTcGTaTATCCtGT
cGATGATCATCATTTTAAAGTGATTCTGCATTATGGCACCCTGGTGATTGATGGCGTGACCC
CGAACATGATTGATTATTTTGGCCGCCCGTATGAAGGCATTGCtGTtTTTGATGGCAAgAAgA
TTACCGTGACCGGCACCCTGTGGAACGGCAACAAAATTATTGATGAACGCCTGATTAACCC
aGATGGCAGCCTGCTGTTTCGCGTGACCATTAACGGCGTGACCGGCTGGCGCCTGTGCGA
ACGCATTCTcGCaGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACA
AGGATGACGATGACAAGTAA

Control nLuc reporter (only used in Figure 6B)
*TAATACGACTCACTATAG*GG**AGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTTGGCAAT**
CCGGTACTGTTGGTAAATAAGCCACCATGGTCTTCACACTCGAAGATTTCGTTGGGGACT
GGCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTT
TGTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAA
TGGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAAT

GGGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTG
ATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCG
GACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCC
TGTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGT
TCCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGAC
TACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACA
AGTAA

CrPV IGR IRES nLuc reporter

*TAATACGACTCACTATAG*GG**AGACCCAAGCTGGTTGGGGCGCGTGGTGGCGGCTGCAGC**
CGCCACCACGCGCCCCGGCTAGTTAAGCTTGGTACCGAGCTCGGATCCAGTACCCTTCA
CCAAAGCAAAAATGTGATCTTGCTTGTAAATACAATTTTGAGAGGTTAATAAATTACAAGTA
GTGCTATTTTTGTATTTAGGTTAGCTATTTAGCTTTACGTTCCAGGATGCCTAGTGGCAGCC
CCACAATATCCAGGAAGCCCTCTCTGCGGTTTTTCAGATTAGGTAGTCGAAAAACCTAAGA
AATTTACCTGCTACATTTCAAGATACCATGGTCTTCACACTCGAAGATTTCGTTGGGGACTG
GCGACAGACAGCCGGCTACAACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTT
GTTTCAGAATCTCGGGGTGTCCGTAACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAAT
GGGCTGAAGATCGACATCCATGTCATCATCCCGTATGAAGGTCTGAGCGGCGACCAAATG
GGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATCATCACTTTAAGGTGA
TCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGATCGACTATTTCGG
ACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGATCACTGTAACAGGGACCCT
GTGGAACGGCAACAAAATTATCGACGAGCGCCTGATCAACCCCGACGGCTCCCTGCTGTT
CCGAGTAACCATCAACGGAGTGACCGGCTGGCGGCTGTGCGAACGCATTCTGGCGGACT
ACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAA
GTAA

Recombinant Protein Coding Sequences

- Start codon
- MBP
- His6
- NT-hFMRP (WT and mutant)
- Stop codon

His6-MBP

ATGGGTTCTTCTCACCATCACCATCACCATGGTTCTTCTATGAAAATCGAAGAAGGTAAACT
GGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTC
GAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCC
CACAGGTTGCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTG
GTGGCTACGCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACA
AGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAACGGCAAGCTGATTGCTTACCCGAT
CGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACC
TGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGAT
GTTCAACCTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGC
GTTCAAGTATGAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGC
GAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGAATGCAGACACC
GATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCC
CGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGA
CCTTCAAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCG

CCAGTCCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGG
TCTGGAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGA
AGAGTTGGCGAAAGATCCACGTATTGCCGCCACTATGGAAAACGCCCAGAAAGGTGAAAT
CATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAA
CGCCGCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATGGGA
TCGAGGAAAACCTGTACTTCCAATCCAATATTGGAAGTGGATAA

MBP-(WT NT-hFMRP)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGCTTCCACGAACAATTTATTGTTCGCGAGGACCTGATGGGCTTGGCGATTGGCAC
CCACGGCGCGAACATTCAGCAGGCGCGTAAAGTCCCAGGTGTTACCGCTATTGATCTGGA
CGAAGACACGTGCACCTTTCATATTTATGGTGAAGACCAGGATGCTGTGAAAAAGGCGCGT
TCCTTCCTGGAATTTGCGGAAGACGTGATCCAGGTTCCGCGTAACCTGGTTGGTAAAGTCA
TCGGCAAAAACGGTAAGTTAATCCAAGAGATCGTGGACAAAAGCGGGGTAGTTCGCGTTC
GGATTGAAGCGGAAAATGAGAAGAACGTTCCGCAGGAGGAAGAAATTATGCCGCCAAATA
GCCTGCCGAGCAACAACTCACGTGTCGGTCCGAACGCTCCGGAAGAGAAGAAGCACCTG
GATATTAAAGAGAACAGCACCCATTTCAGCCAACCAAACTCCACTAAGGTGCAGCGTGTTT
TGGTAGCCAGCTCCGTTGTTGCCGGTGAGTCGCAAAAGCCGGAACTGAAAGCGTGGCAG
GGTATGGTGCCGTTCGTCTTTGTGGGCACCAAGGACAGCATCGCCAACGCAACGGTTCTG
CTGGACTACCATCTGAATTACCTGAAAGAAGTCGATCAGCTTCGTTTGGAACGCTTGCAAA
TCGATGAGCAACTGCGCCAGATCGGTGCGAGCTCTCGCCCGTCTCCGAACCGTACCGACA
AAGAGAAGAGCTACGTGACCGACGACGGTCAAGGTATGGGCCGTGGCAGCCGTCCGTAT
CGTAATAGGGGACATGGCCGCCGTGGTCCGGGTTACACCTCAGGTACGAACTCTGAGGC
GTCTAACGCCTCCGAAACCGAGTCGGATCATCGTGATGAGCTGAGCGACTGGTCACTGGC
GCCGACCGAAGAGGAGCGCGAGAGCTTTCTGCGTCGCGGTGATGGTCGCCGCAGAGGC
GGTGGCGGTCGTGGCCAGGGCGGCCGTGGCAGAGGCGGCGGTTTTAAAGGTAATGATGA
TCACAGCCGCACTGACAACCGTCCGCGTAATCCGCGTGAGGCGAAGGGCAGAACTACAG
ATGGTAGCTTGCAAATCCGTGTGGACTGTAATAACGAACGCAGCGTGCATACGAAAACCCT
ACAAAACACCAGTAGCGAGGGTAGCCGCCTTCGTACCGGTAAAGACCGCAACCAGAAGAA
AGAGAAACCGGACAGCGTTGATGGCCAACAACCGTTGGTGAATGGTGTTCCGGGTTCTTC
TCACCATCACCATCACCATTAA

MBP-(I304N NT-hFMRP)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGCTTCCACGAACAATTTATTGTTCGCGAGGACCTGATGGGCTTGGCGATTGGCAC
CCACGGCGCGAACATTCAGCAGGCGCGTAAAGTCCCAGGTGTTACCGCTATTGATCTGGA
CGAAGACACGTGCACCTTTCATATTTATGGTGAAGACCAGGATGCTGTGAAAAAGGCGCGT
TCCTTCCTGGAATTTGCGGAAGACGTGATCCAGGTTCCGCGTAACCTGGTTGGTAAAGTCA
TCGGCAAAAACGGTAAGTTAAACCAAGAGATCGTGGACAAAAGCGGGGTAGTTCGCGTTC
GGATTGAAGCGGAAAATGAGAAGAACGTTCCGCAGGAGGAAGAAATTATGCCGCCAAATA
GCCTGCCGAGCAACAACTCACGTGTCGGTCCGAACGCTCCGGAAGAGAAGAAGCACCTG
GATATTAAAGAGAACAGCACCCATTTCAGCCAACCAAACTCCACTAAGGTGCAGCGTGTTT
TGGTAGCCAGCTCCGTTGTTGCCGGTGAGTCGCAAAAGCCGGAACTGAAAGCGTGGCAG
GGTATGGTGCCGTTCGTCTTTGTGGGCACCAAGGACAGCATCGCCAACGCAACGGTTCTG
CTGGACTACCATCTGAATTACCTGAAAGAAGTCGATCAGCTTCGTTTGGAACGCTTGCAAA
TCGATGAGCAACTGCGCCAGATCGGTGCGAGCTCTCGCCCGTCTCCGAACCGTACCGACA
AAGAGAAGAGCTACGTGACCGACGACGGTCAAGGTATGGGCCGTGGCAGCCGTCCGTAT
CGTAATAGGGGACATGGCCGCCGTGGTCCGGGTTACACCTCAGGTACGAACTCTGAGGC
GTCTAACGCCTCCGAAACCGAGTCGGATCATCGTGATGAGCTGAGCGACTGGTCACTGGC
GCCGACCGAAGAGGAGCGCGAGAGCTTTCTGCGTCGCGGTGATGGTCGCCGCAGAGGC
GGTGGCGGTCGTGGCCAGGGCGGCCGTGGCAGAGGCGGCGGTTTTAAAGGTAATGATGA
TCACAGCCGCACTGACAACCGTCCGCGTAATCCGCGTGAGGCGAAGGGCAGAACTACAG
ATGGTAGCTTGCAAATCCGTGTGGACTGTAATAACGAACGCAGCGTGCATACGAAAACCCT
ACAAAACACCAGTAGCGAGGGTAGCCGCCTTCGTACCGGTAAAGACCGCAACCAGAAGAA
AGAGAAACCGGACAGCGTTGATGGCCAACAACCGTTGGTGAATGGTGTTCCGGGTTCTTC
TCACCATCACCATCACCATTAA

MBP-(ΔRGG+CTD NT-hFMRP)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA

CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGCTTCCACGAACAATTTATTGTTCGCGAGGACCTGATGGGCTTGGCGATTGGCAC
CCACGGCGCGAACATTCAGCAGGCGCGTAAAGTCCCAGGTGTTACCGCTATTGATCTGGA
CGAAGACACGTGCACCTTTCATATTTATGGTGAAGACCAGGATGCTGTGAAAAAGGCGCGT
TCCTTCCTGGAATTTGCGGAAGACGTGATCCAGGTTCCGCGTAACCTGGTTGGTAAAGTCA
TCGGCAAAAACGGTAAGTTAATCCAAGAGATCGTGGACAAAAGCGGGGTAGTTCGCGTTC
GGATTGAAGCGGAAAATGAGAAGAACGTTCCGCAGGAGGAAGAAATTATGCCGCCAAATA
GCCTGCCGAGCAACAACTCACGTGTCGGTCCGAACGCTCCGGAAGAGAAGAAGCACCTG
GATATTAAAGAGAACAGCACCCATTTCAGCCAACCAAACTCCACTAAGGTGCAGCGTGTTT
TGGTAGCCAGCTCCGTTGTTGCCGGTGAGTCGCAAAAGCCGGAACTGAAAGCGTGGCAG
GGTATGGTGCCGTTCGTCTTTGTGGGCACCAAGGACAGCATCGCCAACGCAACGGTTCTG
CTGGACTACCATCTGAATTACCTGAAAGAAGTCGATCAGCTTCGTTTGGAACGCTTGCAAA
TCGATGAGCAACTGCGCCAGATCGGTGCGAGCTCTCGCCCGTCTCCGAACCGTACCGACA
AAGAGAAGAGCTACGTGACCGACGACGGTCAAGGTATGGGCCGTGGCAGCCGTCCGTAT
CGTAATAGGGGACATGGCCGCCGTGGTCCGGGTTACACCTCAGGTACGAACTCTGAGGC
GTCTAACGCCTCCGAAACCGAGTCGGATCATCGTGATGAGCTGAGCGACTGGTCACTGGC
GCCGACCGAAGAGGAGCGCGAGAGCTTTCTGCGTCGCGGTGATGGTCGCCGCAGAGGC
GGTGGCGGGTCGTGGCCAGGGCGGCCGTGGCAGAGGCGGCGGTTTGGTTCTTCTCACCA
TCACCATCACCATTAA

MBP-(RGG+CTD)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGTCGCGGTGATGGTCGCCGCAGAGGCGGTGGCGGTCGTGGCCAGGGCGGCCGT

GGCAGAGGCGGCGGTTTTAAAGGTAATGATGATCACAGCCGCACTGACAACCGTCCGCGT
AATCCGCGTGAGGCGAAGGGCAGAACTACAGATGGTAGCTTGCAAATCCGTGTGGACTGT
AATAACGAACGCAGCGTGCATACGAAAACCCTACAAAACACCAGTAGCGAGGGTAGCCGC
CTTCGTACCGGTAAAGACCGCAACCAGAAGAAAGAGAAACCGGACAGCGTTGATGGCCAA
CAACCGTTGGTGAATGGTGTTCCGGGTTCTTCTCACCATCACCATCACCATTAA

MBP-(RGG)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGTCGCGGTGATGGTCGCCGCAGAGGCGGTGGCGGTCGTGGCCAGGGCGGCCGT
GGCAGAGGCGGCGGTTTTAAAGGTAATGATGATCACAGCCGCGGTTCTTCTCACCATCAC
CATCACCATTAA

MBP-(CTD)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCAACTGACAACCGTCCGCGTAATCCGCGTGAGGCGAAGGGCAGAACTACAGATGGTAG

CTTGCAAATCCGTGTGGACTGTAATAACGAACGCAGCGTGCATACGAAAACCCTACAAAAC
ACCAGTAGCGAGGGTAGCCGCCTTCGTACCGGTAAAGACCGCAACCAGAAGAAAGAGAAA
CCGGACAGCGTTGATGGCCAACAACCGTTGGTGAATGGTGTTCCGGGTTCTTCTCACCAT
CACCATCACCATTAA

MBP-(RGG+CTD Δ19)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGTCGCGGTGATGGTCGCCGCAGAGGCGGTGGCGGTCGTGGCCAGGGCGGCCGT
GGCAGAGGCGGCGGTTTTAAAGGTAATGATGATCACAGCCGCACTGACAACCGTCCGCGT
AATCCGCGTGAGGCGAAGGGCAGAACTACAGATGGTAGCTTGCAAATCCGTGTGGACTGT
AATAACGAACGCAGCGTGCATACGAAAACCCTACAAAACACCAGTAGCGAGGGTAGCCGC
CTTCGTACCGGTAAAGACCGCAACCAGGGTTCTTCTCACCATCACCATCACCATTAA

MBP-(RGG+CTD Δ39)-His6

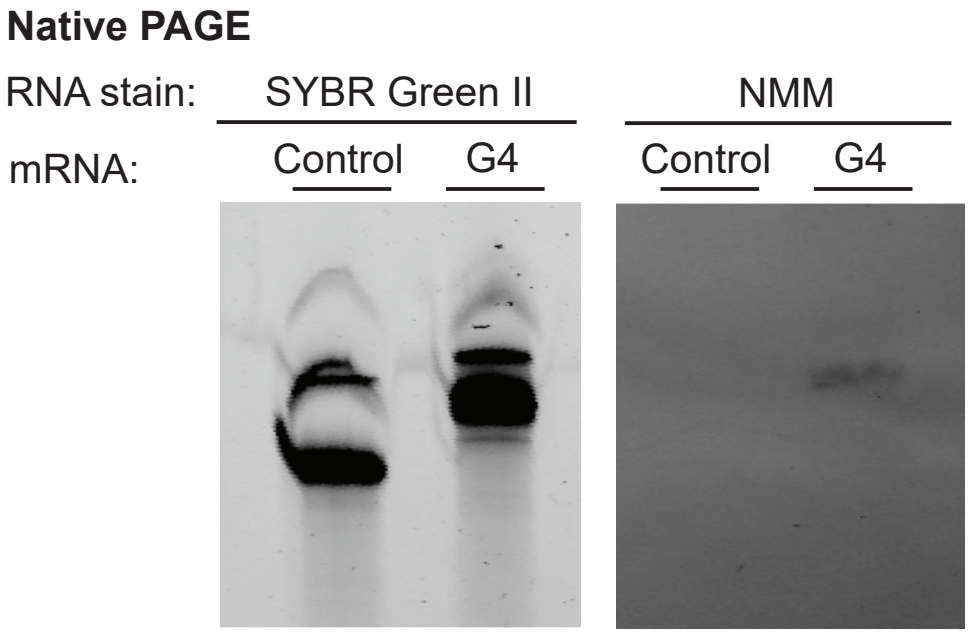
ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG

ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGTCGCGGTGATGGTCGCCGCAGAGGCGGTGGCGGTCGTGGCCAGGGCGGCCGT
GGCAGAGGCGGCGGTTTTAAAGGTAATGATGATCACAGCCGCACTGACAACCGTCCGCGT
AATCCGCGTGAGGCGAAGGGCAGAACTACAGATGGTAGCTTGCAAATCCGTGTGGACTGT
AATAACGAACGCAGCGTGCATACGAAAGGTTCTTCTCACCATCACCATCACCATTAA

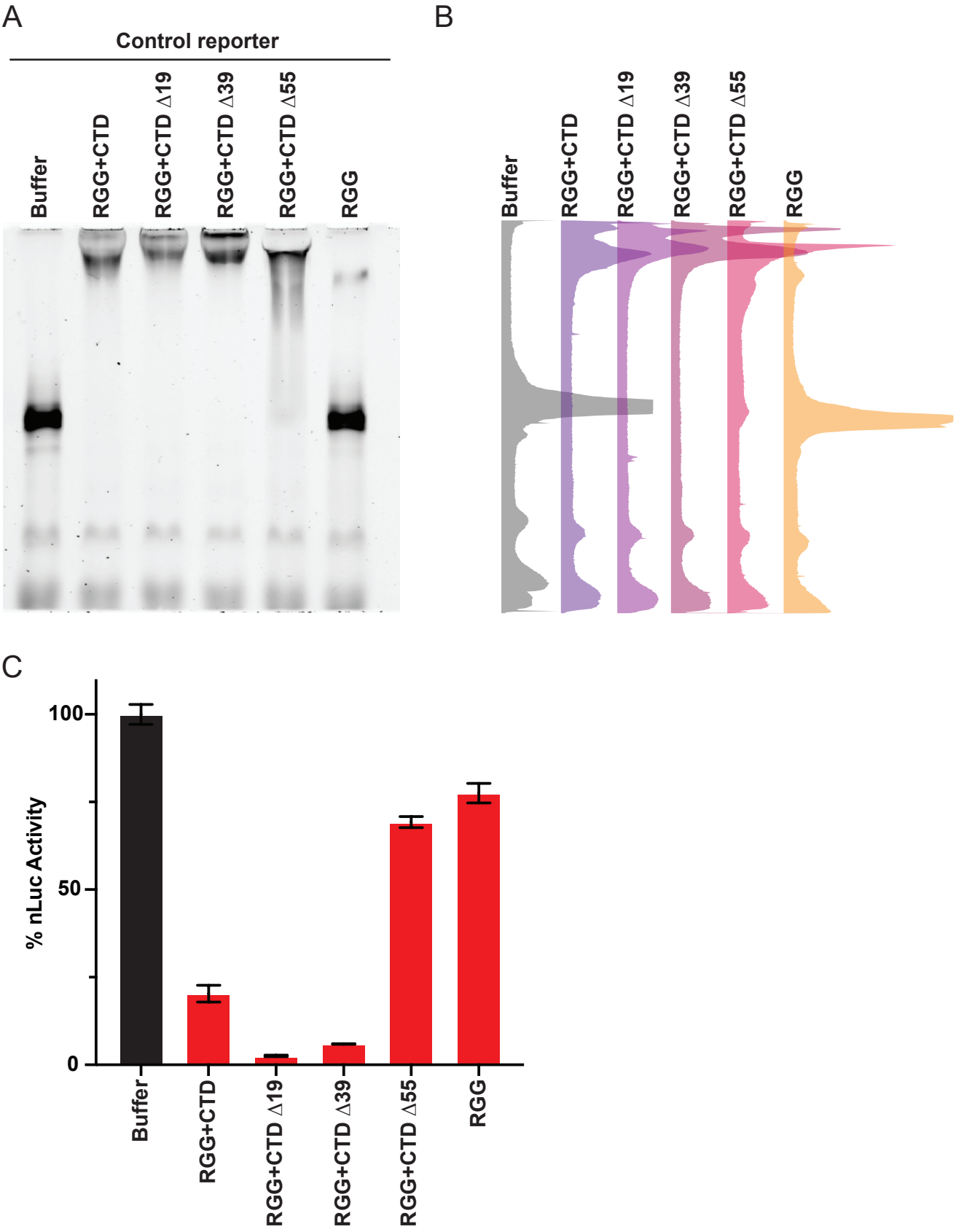
MBP-(RGG+CTD Δ55)-His6

ATGGGTTCTTCTATGAAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAG
GCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGATACCGGAATTAAAGTCAC
CGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTTGCGGCAACTGGCGATGG
CCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTGTTG
GCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCC
GTACGTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTT
ATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAAGAGATCCCGGCGCTGGATA
AAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAACCTGCAAGAACCGTACTTCA
CCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAGTACG
ACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTG
ACCTGATTAAAAACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTT
TAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGGGCATGGTCCAACATCGACAC
CAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTCAAGGGTCAACCATCCAAACCG
TTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTGGCAAAA
GAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAAC
CGCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTG
CCGCCACTATGGAAAACGCCCAGAAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCG
CTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCCGCCAGCGGTCGTCAGACTGTCG
ATGAAGCCCTGAAAGACGCGCAGACTAATGGGATCGAGGAAAACCTGTACTTCCAATCCAA
TGCACGTCGCGGTGATGGTCGCCGCAGAGGCGGTGGCGGTCGTGGCCAGGGCGGCCGT
GGCAGAGGCGGCGGTTTTAAAGGTAATGATGATCACAGCCGCACTGACAACCGTCCGCGT
AATCCGCGTGAGGCGAAGGGCAGAACTACAGATGGTAGCGGTTCTTCTCACCATCACCAT
CACCATTAA

Supplementary Figure S1

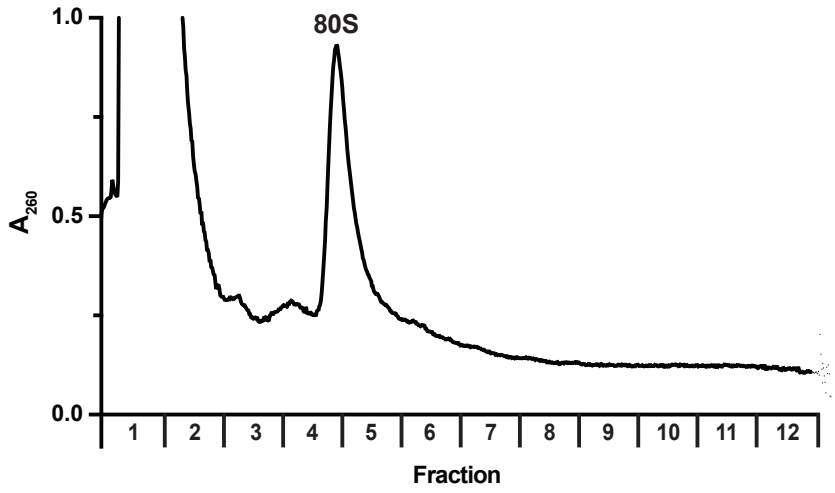


Supplementary Figure S2

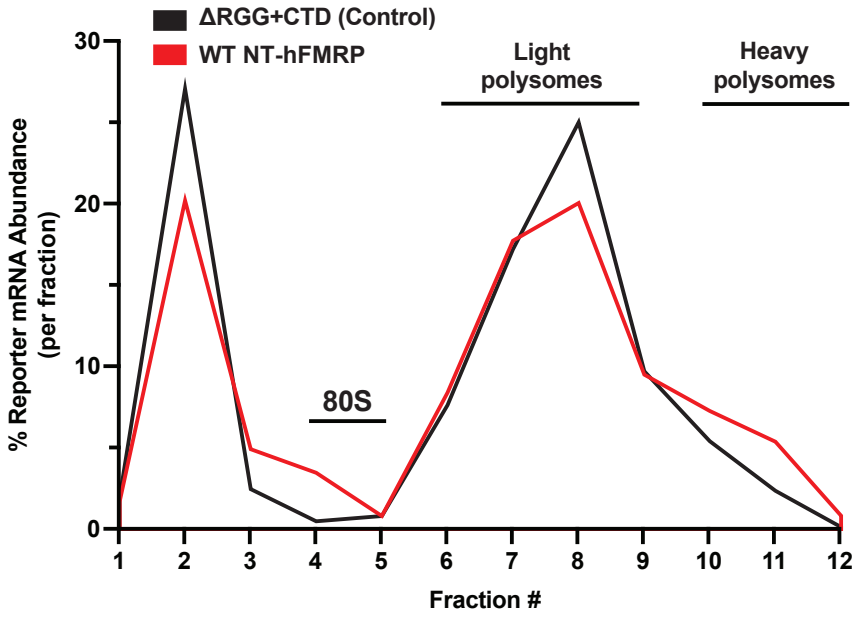


Supplementary Figure S3

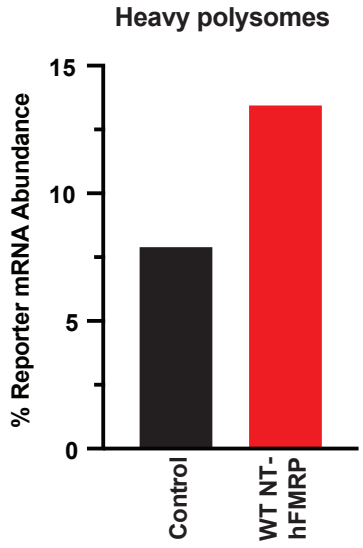
A



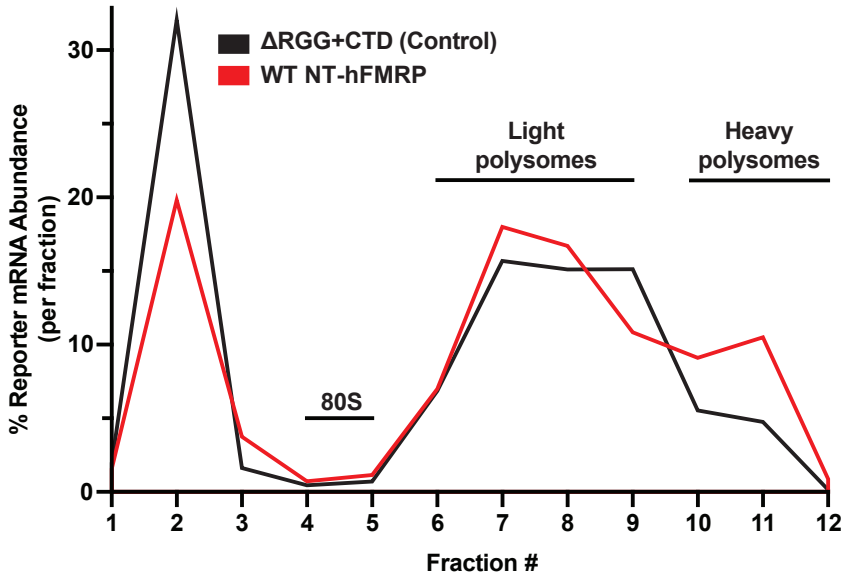
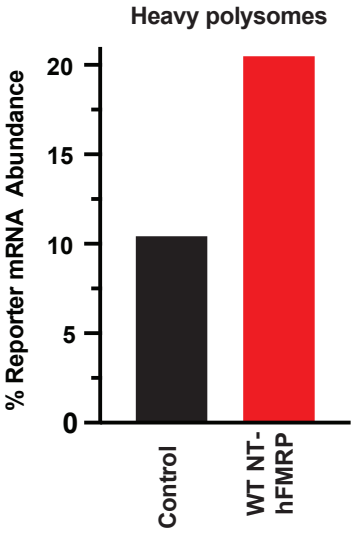
B



C



D



E