Cell Biology

Insufficiency of 40S ribosomal proteins, RPS26 and RPS25, negatively affects biosynthesis of polyglycine-containing proteins in fragile-X associated conditions

Reviewed Preprint
v2 • February 19, 2025
Revised by authors

Reviewed Preprint
v1 • July 1, 2024

Katarzyna Tutak, Izabela Broniarek, Andrzej Zielezinski, Daria Niewiadomska, Tomasz Skrzypczak, Anna Baud, Krzysztof Sobczak

Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland • Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland • Center of Advanced Technology, Adam Mickiewicz University, Poznan, Poland

eLife Assessment

In this **valuable** study, Tutak and colleagues set out to identify factors that mediate Repeat Associated Non-AUG (RAN) translation of CGG repeats in the FMR1 mRNA which are implicated in toxic protein accumulation that underpins ensuing neurological pathologies. The authors provide **solid** evidence that RPS26 may be implicated in mediating the RAN translation of FMR1 mRNA. This article should be of broad interest to researchers in the variety of disciplines including post-transcriptional regulation of gene expression and neurobiology.

https://doi.org/10.7554/eLife.98631.2.sa3

Abstract

Expansion of CGG repeats (CGGexp) in the 5' untranslated region (5'UTR) of the *FMR1* gene underlies the fragile X premutation-associated conditions including tremor/ataxia syndrome, a late-onset neurodegenerative disease and fragile X-associated primary ovarian insufficiency. One common pathomechanism of these conditions is the repeat-associated non-AUG-initiated (RAN) translation of CGG repeats of mutant *FMR1* mRNA, resulting in production of FMRpolyG, a toxic protein containing long polyglycine tract. To identify novel modifiers of RAN translation we used an RNA-tagging system and mass spectrometry-based screening. It revealed proteins enriched on CGGexp-containing *FMR1* RNA in cellulo, including a ribosomal protein RPS26, a component of the 40S subunit. We demonstrated that depletion of RPS26 and its chaperone TSR2, modulates FMRpolyG production and its toxicity. We also found that the RPS26 insufficiency impacted translation of limited number of proteins, and 5'UTRs of mRNAs encoding these proteins were short and guanosine and cytosine-rich. Moreover, the silencing of another component of the 40S subunit, the

ribosomal protein RPS25, also induced repression of FMRpolyG biosynthesis. Results of this study suggest that the two 40S ribosomal proteins and chaperone TSR2 play an important role in noncanonical CGGexp-related RAN translation.

Introduction

Fragile X chromosome-associated syndromes are rare genetic diseases caused by dynamic mutations of the *fragile X messenger ribonucleoprotein 1* (*FMR1*) gene located on the X chromosome. The gene typically contains 25–30 CGG repeats in the 5' untranslated region (5'UTR). However, these triplet repeats are highly polymorphic and prone to expand, resulting in either a full mutation (FM; over 200 CGG repeats) or premutation (PM; 55–200 CGG repeats). On the molecular level, FM causes methylation of the *FMR1* promoter, leading to transcriptional silencing, loss of *FMR1* mRNA, and a lack of the main protein product, fragile X messenger ribonucleoprotein (FMRP), which is involved in modulating synaptic plasticity. An FM causes early onset neurodevelopmental fragile X syndrome (FXS), while PM is linked to many fragile X-associated conditions (FXPAC) including fragile X-associated tremor/ataxia syndrome (FXTAS), fragile X-associated primary ovarian insufficiency (FXPOI), and fragile X-associated neuropsychiatric disorders (FXAND). The estimated prevalence of PM is 1 in 150–300 females and 1 in 400–850 males. However, due to incomplete penetrance, approximately 1 in 5,000–10,000 men in their fifties or later will develop FXTAS. In female PM carriers, random X-inactivation lowers the risk of FXTAS development (Hagerman & Hagerman, 2016; Tassone *et al*, 2012; Jacquemont *et al*, 2004). FXTAS is a late-onset neurodegenerative disease. Its pathology includes neuropathy, white matter loss, mild brain atrophy, and ubiquitin-positive inclusions in neurons and glia (Hagerman & Hagerman, 2016; Greco *et al*, 2002, 2006). Patients suffer from cognitive decline, dementia, parkinsonism, imbalance, gait ataxia, and tremors accompanied by psychological difficulties such as anxiety or depression (Hagerman *et al*, 2018; Hagerman & Hagerman, 2016). To date, no effective treatment targeting the cause, rather than the symptoms, has been proposed for any PM-linked disorders.

Three main molecular pathomechanisms are believed to contribute to FXTAS, FXPOI, and FXAND development (Glineburg *et al*, 2018; Malik *et al*, 2021a; Hagerman *et al*, 2018). First, high-content guanosine and cytosine nucleotides in the 5'UTR of *FMR1* cause co-transcriptional DNA:RNA hybrid formations (R-loops), which trigger the DNA damage response, thus compromising genomic stability, leading to cell death (Loomis *et al*, 2014; Abu Diab *et al*, 2018). Second, RNA gain-of-function toxicity induces nuclear foci formation by mRNA containing expanded CGG repeats (CGGexp) which form stable hairpin structures, and sequester proteins leading to their functional depletion (Sellier *et al*, 2010, 2013, 2017). Finally, mRNA containing CGGexp can act as a template for noncanonical protein synthesis called repeats-associated non-AUG initiated (RAN) translation. Protein production from expanded nucleotide repeats is initiated at different near-cognate start codons in diverse reading frames. The resultant toxic proteins contain repeated amino acid tracts, such as polyglycine (FMRpolyG), polyalanine (FMRpolyA), polyarginine (FMRpolyR), or hybrids of them produced as a result of frameshifting (Todd *et al*, 2013; Wright *et al*, 2022; Glineburg *et al*, 2018; Kearse *et al*, 2016). Notably, the open reading frame for FMRP starting from the AUG codon downstream to the repeats is canonically synthesized. In FXTAS and FXPOI, the most abundant RAN protein is the toxic FMRpolyG, which aggregates and forms characteristic intranuclear or perinuclear inclusions observed in patient cells and model systems (Greco *et al*, 2002, 2006; Ariza *et al*, 2016; Todd *et al*, 2013; Sellier *et al*, 2017; Ma *et al*, 2019; Derbis *et al*, 2018).

According to the current RAN translation model of RNA with CGGexp, the eIF4F complex and 43S pre-initiation complex (PIC) bind to the 5'-cap of *FMR1* mRNA and scan through the 5'UTR until encountering a steric hindrance, a hairpin structure formed by CGGexp, or a nearby 5'UTR sequence. This blockage increases the dwell time of PIC and lowers initiation codon fidelity. As a

result, the stalled 40S ribosome initiates RAN translation at less favored, near-cognate start codons (ACG or GUG) upstream of repeats (in the FMRpolyG reading frame) or within CGG repeats (in the FMRpolyA reading frame) (Kearse *et al*, 2016; Green *et al*, 2016). Unwinding stable RNA secondary structures appears to be crucial for initiating RAN translation, as several RNA helicases such as ATP-dependent RNA helicase DDX3X, ATP-dependent DNA/RNA helicase DHX36, and eukaryotic initiation factor 4A (eIF4A) are involved in its regulation (Linsalata *et al*, 2019; Kearse *et al*, 2016; Tseng *et al*, 2021). Proteins indirectly involved in RAN translation can also contribute to RAN-mediated toxicity via pathways related to stress response and nuclear transport (Green *et al*, 2017; Zu *et al*, 2020; Malik *et al*, 2021b).

New insights into ribosome heterogeneity have explained rearrangements within ribosome components at different developmental stages or responses to environmental stimuli. This has led to a better understanding of events that shape local ribosome homeostasis and affect the translatome (Genuth & Barna, 2018; Shi & Barna, 2015). For example, although ribosomes depleted of small ribosomal subunit protein eS26 (RPS26) stays functional in the cell, they are translating preferentially selected mRNAs (Ferretti *et al*, 2017; Yang & Karbstein, 2022; Li *et al*, 2022). Moreover, ribosomes containing small ribosomal subunit protein eS25 (RPS25) and Large ribosomal subunit protein uL1 RPL10A translate different pools of mRNA including those encoding key components of cell cycle process, metabolism, and development (Shi *et al*, 2017).

Recently, different proteins involved in RAN translation regulation reviewed in Baud *et al*, 2022)were uncovered. However, mechanistic insight into this process remains unresolved. Given the toxicity of RAN translation products, identification of its regulating factors, which may serve as potential therapeutic targets to combat RAN proteins-related toxicity in fragile X-associated conditions, is essential.

In this study, we adapted an RNA-tagging technique to identify proteins natively bound to the RNA of the 5'UTR of *FMR1* with CGGexp that mimics mutated transcripts in PM carriers. Among tens of identified proteins, we focused on the RPS26 and investigated its involvement in CGGexp-related RAN translation. Previously it was shown that RPS26 can interact with mRNA associated with PIC and regulate the translation of selected transcripts depending on their sequence, 5'UTR's length, and stress conditions (Ferretti *et al*, 2017; Havkin-Solomon *et al*, 2023; Li *et al*, 2022; Yang & Karbstein, 2022; Pisarev *et al*, 2008). By regulating RPS26 at the cellular level and its incorporation into the assembling 40S subunit mediated by escortin TSR2 (Schütz *et al*, 2014; Yang & Karbstein, 2022), we found that insufficiency of this ribosomal protein has negative effect on the level and toxicity of FMRpolyG with no impact on encoding mRNA and FMRP. As FMRP is the primary protein product of the *FMR1* gene, it indicates that RPS26 depletion selectively modulates CGG-related RAN translation. Furthermore, by using a proteomic approach, we found that the number of proteins sensitive to RPS26 insufficiency was limited.

Results

Mass-spectrometry-based screening revealed numerous proteins interacting with the 5'UTR of FMR1 mRNA

To uncover new modifiers of CGGexp-related RNA toxicity, we used an RNA-tagging method coupled with mass spectrometry (MS) that allows for the in cellulo capture and identification of native RNA-protein complexes. The RNA bait, *FMR1 RNA* (used to pull down interacting proteins) consisted of the entire sequence of the 5'UTR of *FMR1*, which contained expanded, 99-times repeated, CGG tracts tagged with three times repeated MS2 RNA stem-loop aptamers (**Figure 1A**, **Supplementary Figure 1A**). These aptamers did not affect the RAN translation process as the synthesis of FMRpolyG was initiated from the near-cognate ACG start codon located upstream to CGGexp and terminated upstream to the MS2 aptamers (**Supplementary Figure 1B**). Given that

the sequence of the 5'UTR of *FMR1* was guanosine and cytosine (GC)-rich (ca. 90% of GC content), we used an RNA sequence of the same length with GC content higher than 70%, *GC-rich RNA* (**Figure 1A**, **Supplementary Figure 1A**), as a control. Importantly, both of these sequences contained open reading frames similar in length but with differing start codons (ACG in *FMR1 RNA* and AUG in *GC-rich RNA*). Thus, they served as templates for protein synthesis (**Supplementary Figure 1A**). Together with RNA baits, the MS2 protein— showing high affinity towards RNA MS2 stem-loop aptamers—was co-expressed in HEK293T cells to immunoprecipitate natively formed RNA bait-protein complexes.

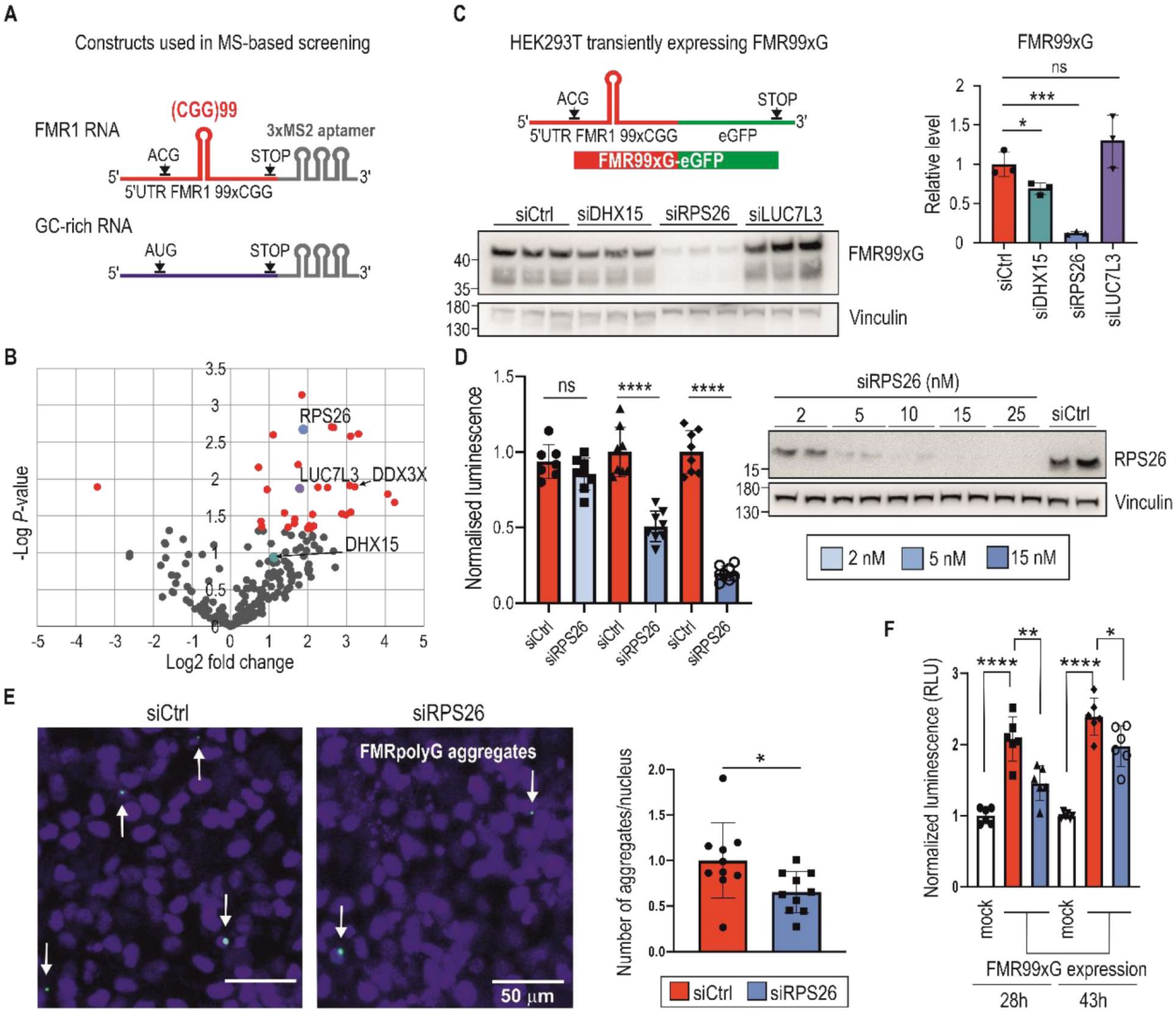

A
Constructs used in MS-based screening
FMR1 RNA
(CGG)99
3xMS2 aptamer
ACG
STOP
5'
3'
5'UTR FMR1 99xCGG
GC-rich RNA
AUG
STOP
B
-Log P-value
Log2 fold change
RPS26
LUC7L3
DDX3X
DHX15
C
HEK293T transiently expressing FMR99xG
5'UTR FMR1 99xCGG
eGFP
FMR99xG-eGFP
siCtrl
siDHX15
siRPS26
siLUC7L3
FMR99xG
Vinculin
Relative level
D
Normalised luminescence
ns
****
siRPS26 (nM)
2
5
10
15
25
siCtrl
RPS26
Vinculin
2 nM
5 nM
15 nM
E
siCtrl
siRPS26
FMRpolyG aggregates
50 µm
Number of aggregates/nucleus
F
Normalized luminescence (RLU)
mock
FMR99xG expression
28h
43h

Figure 1.

Mass-spectrometry (MS) - based screening revealed several proteins bound to *FMR1* RNA with expanded CGG repeats; some affect the yield of polyglycine-containing proteins, alleviating their toxicity.

(A) Scheme of two RNA molecules used to perform MS-based screening. The *FMR1 RNA* contains the entire length of the 5′ untranslated region (5′UTR) of *FMR1* with expanded CGG (x99) repeats forming a hairpin structure (red). The open reading frame for the polyglycine-containing protein starts at a repeats-associated non-AUG initiated (RAN) translation-specific ACG codon. RAN translation can also be initiated from a GUG near-cognate start codon, which is not indicated in the scheme. The GC-rich RNA contains *TMEM107* mRNA enriched with G and C nucleotide residues (GC content > 70%; similar to *FMR1* RNA) with the open reading frame starting at the canonical AUG codon (blue). Both RNAs are tagged with three MS2 stem-loop aptamers (grey) interacting with an MS2 protein tagged with an in vivo biotinylating peptide used to pull down proteins interacting with the RNAs.

(B) The volcano plot representing proteins captured during MS-based screening showing the magnitude of enrichment (log2 fold change) and the statistical significance (log *P*-value); red dots indicate proteins significantly enriched ($P < 0.05$) on *FMR1 RNA* compared to GC-rich RNA. Three proteins, RPS26, LUC7L3, and DHX15, tested in a subsequent validation experiment are marked. DDX3X is also indicated as this protein has been previously described in the context of interaction with *FMR1* mRNA.

(C) The scheme of RNA used for the transient overexpression of FMR99xG (mutant, long, 99 polyglycine tract-containing protein). The construct contains the entire length of the 5′UTR of *FMR1* with expanded CGG repeats forming a hairpin structure (red) tagged with enhanced green fluorescent protein (eGFP) (green). Western blot analysis of FMR99xG and Vinculin for HEK293T cells with insufficient DHX15, RPS26, and LUC7L3 induced by specific short interfering RNA (siRNA) treatment. To detect FMR99xG, the 9FM antibody was used. The upper bands were used for quantification. The graph presents the mean signal for FMR99xG normalized to Vinculin from $N = 3$ biologically independent samples with the standard deviation (SD). An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$; ***, $P < 0.001$; ns, non-significant.

(D) Results of luciferase assay for cells with overexpression of FMR16xG fused with nanoluciferase pretreated with different concentrations of siCtrl or siRPS26 (see legend). The graph presents luminescence signals (means from $N = 8$ with SDs) normalized to siCtrl treated cells. Western blot represents the siRNA-concentration dependent effect on RPS26 downregulation. An unpaired Student's t-test was used to calculate statistical significance: ****, $P < 0.0001$; ns, non-significant.

(E) Results of microscopic quantification of FMR99xG-positive aggregates (white arrows) in HeLa cells transfected with the FMR99xG construct on the background of normal (siCtrl) or reduced amount of RPS26 (siRPS26). Representative images were pseudo-colored and merged; green, GFP-positive inclusions; blue, nuclei stained with Hoechst 33342; scale bars, 50 μm. The histogram presents a number of FMRpolyG-GFP inclusions per nucleus in cells treated with siCtrl (blue) or siRPS26 (red) in $N = 10$ biologically independent samples with SDs. In average, 1100 cells were identified per single image and final value were normalized to 1 for control siRNA treated cells. An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$.

(F) The influence of RPS26 silencing on apoptosis evoked in HeLa cells after 28h or 43h of FMR99xG overexpression or in mock treated cells. Apoptosis was measured as luminescence signals (relative luminescence units; RLU) derived from Annexin V fusion protein bound to phosphatidylserine (PS) exposed on the outer leaflet of cell membranes, an indicator for early apoptosis. The graph presents relative mean values from $N = 6$ biologically independent samples treated with either siCtrl (red) or siRPS26 (blue) with the SD normalized to mock controls (cells transfected only with the delivering reagent and cultured for 28h or 43h). An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. Note, that significantly higher apoptosis was observed in cells expressing FMR99xG compare to control (mock), and that pretreatment with siRPS26 significantly reduced this phenotype.

The MS-based screening identified over 500 proteins binding to two prepared RNA baits. Of those 172 proteins were identified in all technical replicates for FMR1 RNA bait (Supplementary Table 3). The most enriched Gene ontology (GO) terms indicated that the majority of proteins binding to *FMR1 RNA* had RNA binding properties and were involved in ribosome biogenesis, translation, and regulating mRNA metabolic processes (**Supplementary Figure 1C**, a complete list of GO terms is in the Supplementary Table 4).

While searching for novel RAN translation modifiers, we focused on proteins enriched on *FMR1 RNA*. We applied a label-free quantification analysis, which allowed us to elucidate 32 significantly enriched interactors of *FMR1 RNA* compared to *GC-rich RNA* (**Figure 1B**, Supplementary Table 5). Next, we preselected eight proteins based on their biological function and possible role in modulating translation, and using short interfering RNA (siRNA)-based silencing, we investigated the effect of insufficiency of these proteins on the level of toxic FMRpolyG, which could indicate their role in modulating RAN translation. These experiments were conducted in HEK293T cells transiently expressing FMRpolyG containing 99 glycine residues tagged with GFP, named FMR99xG (**Figure 1C**). Silencing of mRNA encoding for the ATP-dependent RNA helicase DHX15 and RPS26, but not six other analyzed proteins, resulted in a significant decrease in steady-state level of FMR99xG, with no effect on its mRNA (**Figure 1C**, **Supplementary Figure 1D** &E). Moreover, using other reporter system expressing polyglycine protein from short CGG repeats (16xCGG) fused with nanoluciferase tag, we showed that partial knock down of RPS26 resulted in significant FMRpolyG downregulation (**Figure 1D**). This indicates that the impact of RPS26 insufficiency on polyglycine level is independent of CGG repeat length, type of reporter used, and that even partial silencing of RPS26 is sufficient to significantly decrease RAN translation from *FMR1* mRNA.

We also investigated whether RPS26 depletion affected the efficiency of FMRpolyG aggregates formation and cell toxicity. In HeLa cells transiently expressing FMR99xG, the frequency of GFP-positive aggregates was reduced upon RPS26 silencing (**Figure 1E**). Moreover, cells with depleted RPS26 exhibited significantly lower apoptosis tendencies evoked by toxic FMRpolyG (**Figure 1F**). These results suggest that decreasing level of RPS26 helps to alleviate FXPAC-related phenotype in cell models.

An orthogonal biochemical assay was used to confirm coprecipitation of RPS26 with the 5'UTR of *FMR1*. We applied an in vitro RNA-protein pull-down using three biotinylated RNAs: 5'UTR of *FMR1* with 99xCGG repeats, synthetic RNA with 23xCGG repeats, and GC-rich RNA as a control. All three RNAs were incubated with an extract derived from HEK293T cells. The first two RNAs, but not GC-rich RNA, were enriched with RPS26, however, the anticipated interaction between RNAs containing CGG repeats and RPS26 was not solely dependent on the triplet number, as RPS26 was pulled down with similar efficiency by RNAs containing more or fewer repeats (**Supplementary Figure 1F**).

In sum, we identified 32 proteins enriched on mutant *FMR1 RNA*. RPS26 and DHX15 insufficiency hindered FMR99xG RAN translation efficiency in a transient expression system. Notably, among proteins identified in the screening there were the ATP-dependent RNA helicase (DDX3X)—the protein previously described as RAN translation modifier (Linsalata *et al*, 2019), and the Src-associated in mitosis 68 kDa protein (SAM68)—the splicing factor sequestered on RNA containing CGGexp (Sellier *et al*, 2010) (**Figure 1B**, **Supplementary Figure 1G**). Identified interactors can be involved not only in RAN translation, but potentially also in different metabolic processes of mutant *FMR1* RNA, such as transcription, protein sequestration, RNA transport, localization, and stability. Importantly, our data showed that silencing of RPS26 alleviated the pathogenic effect of toxic FMRpolyG produced from *FMR1* mRNA containing CGGexp. These facts encouraged us to further investigate the role of RPS26 in the context of CGG-exp-related RAN translation in more natural models.

RPS26 depletion affects cells proliferation but does not inhibit global translation

The RPS26 is a component of the 40S ribosomal subunit and its inclusion or depletion from this subunit affects translation of selected mRNA but does not impact overall translation rate (Schütz *et al*, 2018; Ferretti *et al*, 2017; Havkin-Solomon *et al*, 2023; Yang & Karbstein, 2022; Gaikwad *et al*, 2021). For instance, in yeast cultured in stress conditions, Rps26 is disassociated from 40S subunit, which results in translation of different mRNA pool, especially those mRNAs encoding proteins implicated in stress response pathway (Ferretti *et al*, 2017; Yang & Karbstein, 2022). Moreover, it was shown that RPS26 is involved in translational regulation of selected mRNA contributing to the maintenance of pluripotency in murine embryonic stem cells (Li *et al*, 2022). Its C-terminal domain interacts with mRNA sequences upstream to the E-site of an actively translating ribosome (Pisarev *et al*, 2008; Hussain *et al*, 2014; Anger *et al*, 2013). This localization may be responsible for the regulation of efficient translation initiation (Pisarev *et al*, 2008; Havkin-Solomon *et al*, 2023), especially in a context of non-canonical RAN translation. Hence, these facts encouraged us to further investigate this protein in the context of CGGexp-related RAN translation.

It was recently demonstrated that RPS26 insufficiency affects cells viability (Havkin-Solomon *et al*, 2023), hence we asked if RPS26 silencing affects the proliferation rate of HEK293T cells. Forty-eight hours post RPS26 silencing, cells divided less frequently as their number was reduced by 34% in comparison to control siRNA treated cells (**Figure 2A**). Importantly, almost no dead cells were detected indicating that RPS26 depletion is not detrimental for the cells (**Figure 2A**). Given that RPS26 is crucial for 40S maturation (Plassart *et al*, 2021), we verified if RPS26 depletion affects global translation. Surface sensing of translation (SUnSET) assay (Schmidt *et al*, 2009) demonstrated that RPS26 insufficiency did not elicit significantly the global translation inhibition, as opposed to cycloheximide (CHX) treatment used as a positive control (**Figure 2B**).

RPS26 acts as a RAN translation modulator of mRNA with short and long CGG repeats

To monitor FMRpolyG RAN translation efficiency and test the modulatory properties of preselected proteins, we generated two cell lines stably expressing a fragment of an *FMR1* gene with expanded (95xCGG) or short, normal CGG repeats (16xCGG), named *S-95xCGG* and *S-16xCGG*, respectively. These two models were generated by taking advantage of the Flp-In™ T-Rex™ 293 system (Szczesny *et al*, 2018), which expresses transgenes encoding for either longer (FMR95xG) or shorter (FMR16xG) polyglycine tract-containing proteins tagged with EGFP under the control of a doxycycline-inducible promoter (**Figure 3A**). In these cell lines, the expression of transgenes is detectable a few hours after promoter induction (**Supplementary Figure 2A**); however, even in prolonged doxycycline exposure (up to 35 days), we did not observe FMR95xG positive aggregates (**Supplementary Figure 2B**). A lack of aggregation is advantageous for monitoring RAN translation efficiency, as it allows the entire pool of RAN proteins present in cellular lysate to be measured by western blot, which is impossible if part of the proteins is trapped in insoluble aggregates (Derbis *et al*, 2021, 2018). Moreover, single copy integration of the transgene containing a fragment of the *FMR1* gene mimics natural situation. Notably, the level of transgene expression was comparable to *FMR1* endogenous expression (**Supplementary Figure 2C**) and was homogeneous between cells (**Figure 3A**, **Supplementary Figure 2B**). The second model, named *L-99xCGG*, was generated using lentiviral transduction of HEK293T cells (**Figure 3D**). Lentiviral particles were prepared based on genetic construct used in the transient transfection system (**Figure 1C**). This stable cell line constantly expresses FMR99xG tagged with GFP. In the *L-99xCGG* model, FMR99xG was present in soluble and, to a lesser extent, aggregated form (**Figure 3D**, **Supplementary Figure 2D**).

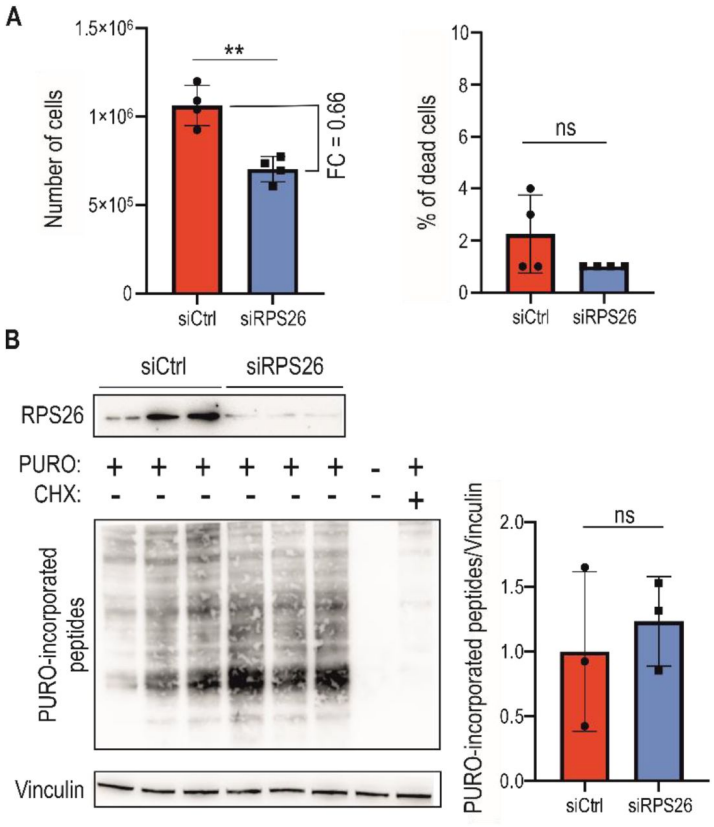


Figure 2.

RPS26 depletion affects proliferation of cells but does not inhibit global translation.

(A) The left graph represents the number of cells and right graph represents percentage (%) of dead cells positively stained with trypan blue after 48h post siRNA-induced silencing. Prior siRNA transfection equal amount of HEK293T cells were seeded (ca. 1×10^4). The graphs present means from $N = 4$ with SDs. FC, fold change. An unpaired Student's t-test was used to calculate statistical significance: **, $P < 0.01$; ns, non-significant. Note, that the number of cells in samples treated with siRPS26 was lowered by 34% compared to siCtrl treated samples.

(B) Western blot analysis of the surface sensing of translation (SUnSET) assay 48h post RPS26 silencing (siRPS26) in HEK293T cells, with antibody against puromycin (PURO), to monitor efficiency of PURO incorporation into peptides during translation. Cells were incubated with siRPS26 or siCtrl and 10 min before protein isolation PURO was added. As a positive control for efficient translation inhibition cells were treated with cycloheximide (CHX), an inhibitor of translation elongation (CHX was added 5 min prior PURO treatment). As a negative control cells treated with DMSO only were used (negative for both PURO and CHX). The graph represents mean signal from whole lanes from $N = 3$ independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: ns, non-significant.

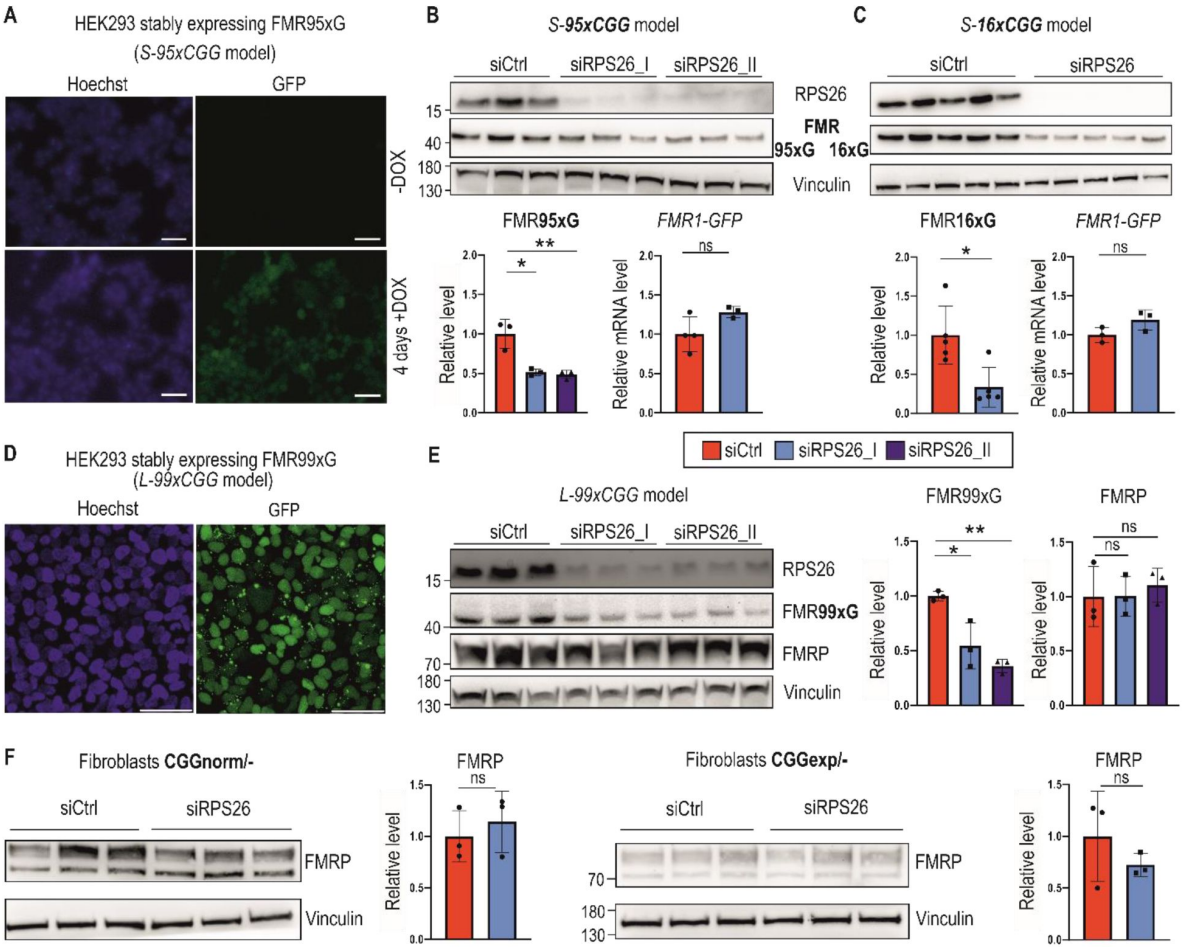


Figure 3.

RPS26 insufficiency induces a lower production of polyglycine, but not FMRP, in multiple FXTAS cellular models.

(A) Representative microscopic images showing inducible expression of FMR95xG fused with eGFP in stable transgenic cell line containing a single copy of 5'UTR FMR1 99xCGG-eGFP transgene under the control of a doxycycline-inducible promoter: *S-95xCGG model*. Blue are nuclei stained with Hoechst; the green signal is derived from FMR95xG tagged with eGFP; scale bar 50 µm; +DOX and DOX indicate cells treated (or not) with doxycycline to induce transcription of the transgene from the doxycycline-dependent promoter.

(B & C) Results of western blot analyses of FMR95xG (with long, mutant polyglycine stretches) or FMR16xG (with short, normal polyglycine stretches), both fused with GFP, normalized to Vinculin and an RT-qPCR analysis of *FMR1-GFP* transgene expression normalized to *GAPDH* upon RPS26 silencing in *S-95xCGG* and *S-16xCGG* models, respectively. siRPS26_I and siRPS26_II indicate two different siRNAs used for RPS26 silencing. The graphs present means from $N = 3$ (**B**) or $N = 5$ biologically independent samples (**C**) with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$; **, $P < 0.01$; ns, non-significant.

(D) Representative microscopic images of cells stably expressing FMR99xCGG obtained after transduction with lentivirus containing the 5'UTR FMR1 99xCGG-eGFP transgene: *L-99xCGG* model. Nuclei are stained with Hoechst and green fluorescence signal is derived from FMR99xG tagged with GFP. Scale bar 50 µm.

(E) Results of western blot analyses of FMR99xG and FMRP normalized to Vinculin upon RPS26 silencing in the *L-99xCGG* model. The graphs present means from $N = 3$ biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$; **, $P < 0.01$; ns, non-significant.

(F) Results of western blot analyses of FMRP levels normalized to Vinculin upon RPS26 silencing in fibroblasts derived from two males, either from healthy individual (CGGnorm – XY, 31xCGG) or FXTAS patient (CGGexp – XY, 81xCGG). The graphs present means from $N = 3$ biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: ns, non-significant.

siRNA-induced silencing of RPS26 in the stable *S-95xCGG* cell model expressing mRNA with long 95xCGG repeats resulted in a significantly decreased level of steady-state FMR95xG, with no effect on its mRNA (**Figure 3B**, **Supplementary Figure 2E**). Moreover, the RPS26 silencing in the *S-16xCGG* model also decreased the level of RAN protein product derived from short 16xCGG repeats without affecting its mRNA level (**Figure 3C**). This suggests that RPS26 depletion affects RAN protein levels independently of CGG repeat content. Similarly to the *S-95*/*16xCGG* models, RPS26 silencing in the lentivirus integration-based *L-99xCGG* model also significantly downregulated FMR99xG biosynthesis (**Figure 3E**).

Previously, several proteins known as RPS26 responders, e.g., murine Polycomb protein (Suz12) and Histone H3.3, were shown to be negatively affected by RPS26 depletion (Li *et al*, 2022). Here, we showed that they also negatively responded to the RPS26-specific siRNAs used in this study (**Supplementary Figure 2F**).

We further investigated whether the translation of FMRP might be affected by RPS26 depletion, as an open reading frame for FMRpolyG–which appears to be under the control of RPS26-sensitive translation–is derived from the same mRNA. To test this hypothesis, we treated human fibroblasts derived from a FXTAS patient and a healthy control with RPS26-specific siRNA. Our results indicate that RPS26 depletion does not affect FMRP levels, regardless of CGG repeat length in *FMR1* mRNA (**Figure 3F**). Similarly, the FMRP level was not affected by RPS26 depletion in other cell models (**Figure 3E**).

Together, these results imply that although the presence of RPS26 in 40S subunit has a positive effect on RAN translation of FMRpolyG and other previously identified RPS26 responders, it does not affect the canonical translation of FMRP produced from the same mRNA. Moreover, a lack of significant differences in translation efficiency of long and short polyglycine-containing proteins in cells with insufficiency of RPS26 suggests that observed effect may depend on certain RNA sequences or structures within the 5'UTR of *FMR1* mRNA or other features of this mRNA, rather than the CGG repeat length.

RPS26 depletion affects only a small subset of the human proteome

To assess globally which proteins are sensitive to RPS26 insufficiency, we used stable isotope labeling with amino acids in cell culture (SILAC). The protein expression level in control and RPS26-depleted HEK293T cells was determined via quantitative mass spectrometry (SILAC-MS, Supplementary Table 6). Using this approach, we identified over 2,600 proteins, with more than 80% identifications based on at least 2 unique peptides. Differential data analysis of quantified proteins (N = 1,887) indicated that most (ca. 80%) proteins identified by SILAC-MS were not sensitive to RPS26 deficiency; we named these proteins non-responders (N = 1,506) (Supplementary Table 7). We also identified set of proteins that negatively (negative responders; N = 223 if the P-value was < 0.05) or positively responded (positive responders; N = 158 if the P-value < 0.05) to RPS26 deficiency (**Figure 4A**, Supplementary Table 7). Non-responders were used as a background list in the GO analysis. An analysis of positive responders found that the proteins in this group were mainly components of translation initiation complexes or formed large ribosomal subunits (**Figure 4B**, Supplementary Table 8). A similar analysis for negative responders did not reveal any significantly enriched GO terms (Supplementary Table 8), partially due to the small number of identified proteins. To further validate this data, we performed western blots for selected negative responders (PDCD4, ILF3, RPS6, and PCBP2), non-responders (FMRP and FUS) or positive responders (EIF5 and EIF3J) from independent cell samples collected 48h post siRPS26 treatment (**Figure 4C**, **Supplementary Figure 3A**). Most tested proteins (6 out of 8) aligned with the quantitative SILAC-MS data (**Figure 4C**, **Supplementary Figure 3A**).

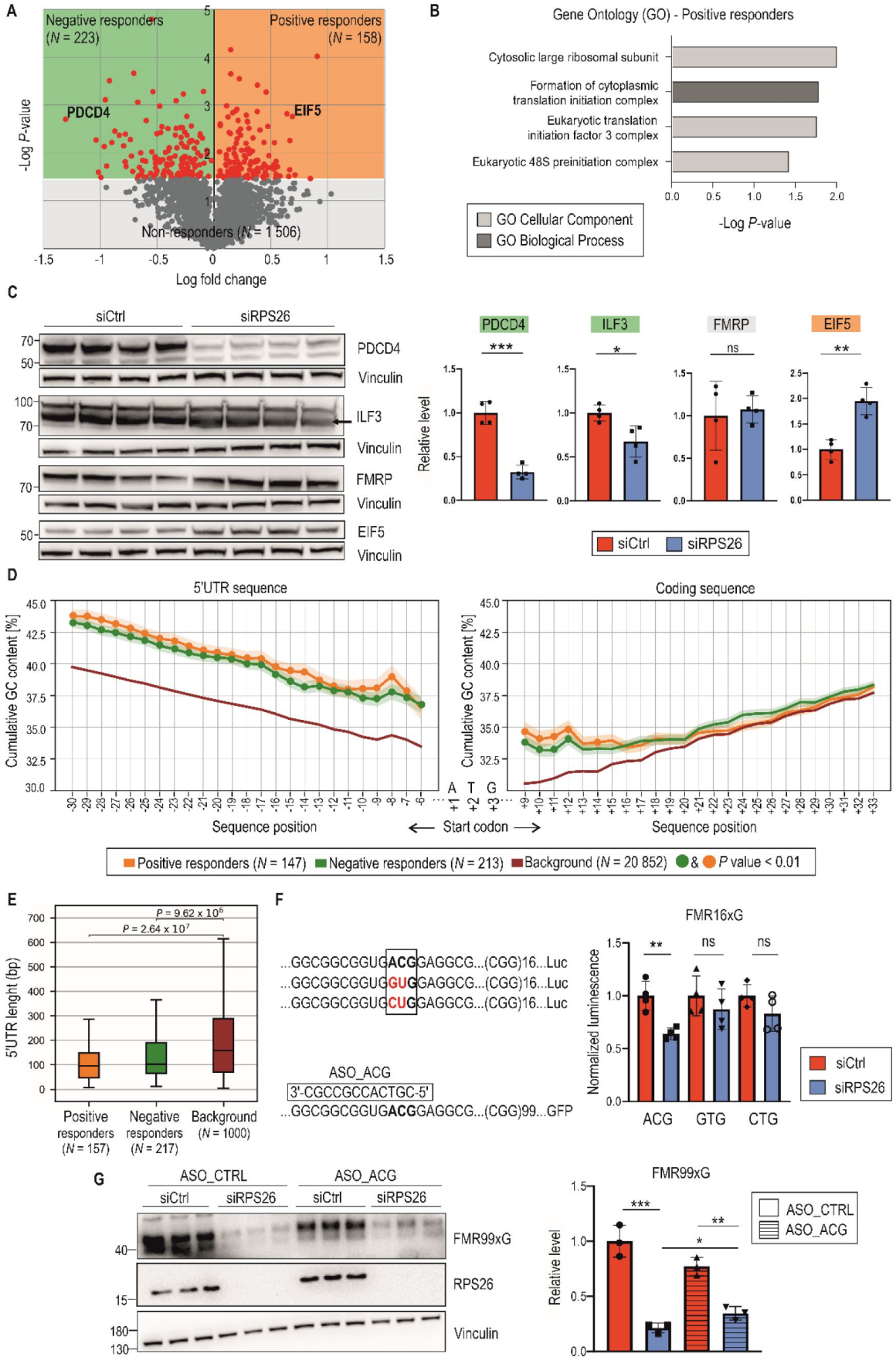

A
Negative responders (N = 223)
Positive responders (N = 158)
PDCD4
EIF5
Non-responders (N = 1 506)
-Log P-value
Log fold change
B
Gene Ontology (GO) - Positive responders
Cytosolic large ribosomal subunit
Formation of cytoplasmic translation initiation complex
Eukaryotic translation initiation factor 3 complex
Eukaryotic 48S preinitiation complex
GO Cellular Component
GO Biological Process
-Log P-value
C
siCtrl
siRPS26
PDCD4
Vinculin
ILF3
FMRP
EIF5
Relative level
siCtrl
siRPS26
D
5'UTR sequence
Coding sequence
Cumulative GC content [%]
Sequence position
Start codon
Positive responders (N = 147)
Negative responders (N = 213)
Background (N = 20 852)
& P value < 0.01
E
P = 9.62 x 10^6
P = 2.64 x 10^7
5'UTR lenght (bp)
Positive responders (N = 157)
Negative responders (N = 217)
Background (N = 1000)
F
...GGCGGCGGUGACGGAGGCG...(CGG)16...Luc
...GGCGGCGGUGGUGGAGGCG...(CGG)16...Luc
...GGCGGCGGUGCUGGAGGCG...(CGG)16...Luc
ASO_ACG
3'-CGCCGCCACTGC-5'
...GGCGGCGGUGACGGAGGCG...(CGG)99...GFP
FMR16xG
Normalized luminescence
ACG
GTG
CTG
siCtrl
siRPS26
G
ASO_CTRL
ASO_ACG
siCtrl
siRPS26
FMR99xG
RPS26
Vinculin
FMR99xG
Relative level
ASO_CTRL
ASO_ACG

Figure 4.

Changes in the proteome of RPS26-deficient cells are not robust.

(A) The volcano plot represents a stable isotope labeling using amino acids in cell culture (SILAC)-based quantitative proteomic analysis identifying proteins sensitive to RPS26 insufficiency. It shows the magnitude of protein-level changes (log2 fold change) vs. the statistical significance (log *P*-value) 48h post siRPS26 treatment. Data were collected from three independent biological replicates for each group. Grey dots indicate proteins non-responding to RPS26 depletion (*N* = 1506). Red dots indicate proteins responding to RPS26 insufficiency (*P* < 0.05); EIF5 and PDCD4 are examples of Negative (*N* = 223; green) and Positive (*N* = 158; orange) responders, respectively. Protein groups were further analyzed in B, C, and D.

(B) Gene ontology (GO) analysis performed for positive responders of RPS26 insufficiency from the proteomic experiment shown in A. The graph presents significantly enriched GO terms (*P* < 0.05); statistical significance was calculated using Fisher's Exact test with the Bonferroni correction. Note that for negative responders, no GO terms were significantly enriched.

(C) Data validation from the proteomic experiment described in A. Western blot analyses of PDCD4, ILF3, FMRP, and EIF5 proteins normalized to Vinculin for HEK293T cells treated with siRPS26. The graphs present means from *N* = 4 biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001; ns, non-significant.

(D) The percentage of GC content in 5'UTRs and coding sequences across extending sequence windows initiated from the start codon within three groups of transcripts: Negative responders (green), Positive responders (orange), and Background (BG) understood as the total transcriptome (red). To avoid biases in the GC-content mean, transcripts with 5'UTR sequences shorter than 20 nucleotides were excluded from the analysis, yielding the following sample sizes: Negative responders (*N* = 213), Positive responders (*N* = 147), and BG (*N* = 20,862). For example, position 6 in 5' UTR sequences corresponds to a 6-nucleotide fragment (window from 6 to 1 positions upstream of ATG), while position 7 corresponds to a 7-nucleotide fragment (window from 7 to 1 positions). The solid line shows the mean GC content at a given position (i.e., within the window), and the shade indicates the standard error of the mean. *P*-values < 0.01 are denoted by green or yellow dots. These reflect pairwise comparisons of GC content between transcript groups (compared to BG) and were determined using a two-tailed paired t-test with Bonferroni correction.

(E) Box plot showing the results of a one-tailed Mann-Whitney U test comparing 5'UTR sequence lengths between positive and negative responders relative to the background transcriptome. To ensure a balanced comparison, the background transcriptome is represented by 1000 randomly selected transcripts from the complete human transcriptome. The plot highlights the median (indicated by the central line), the upper and lower quartiles (represented by the boxes), and the highest and lowest values, which are determined within a range of 1.5 times the interquartile range (IQR) (whiskers).

(F) Results of luciferase assay post RPS26 silencing showing differences in expression of transgene containing 5'UTR of *FMR1* with 16xCGG fused with nanoluciferase from which RAN translation was initiated from either ACG (wild type) or GTG or CTG near cognate codons. The signals for activity of nanoluciferase fused with FMR16xG were normalized to firefly luciferase signals (means from *N* = 4 with SDs are presented on the graph). Note that firefly luciferase was translated from canonical ATG start codon. An unpaired Student's t-test was used to calculate statistical significance: **, *P* < 0.01; ns, non-significant.

(G) Western blot analysis of FMR99xG post RPS26 silencing and treatment with control ASO (ASO_CTRL) or ASO targeting ACG codon, RAN translation initiation codon in 5'UTR of FMR1 (ASO_ACG). The graph represents the means of *N* = 3 with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001; ns, non-significant.

Previously published results showed that specialized ribosomes are formed to selectively translate mRNAs with specific features (Genuth & Barna, 2018; Shi *et al*, 2017). In yeast, mRNAs translated by Rps26-depleted ribosomes lack conservation of all Kozak sequence elements, while mRNAs translated by Rps26-containing ribosomes present a full Kozak consensus (Ferretti *et al*, 2017). RPS26 is localized next to the E-site of the translating ribosome and was found to interact with template mRNAs (Pisarev *et al*, 2008; Anger *et al*, 2013). For instance, if the 43S recognizes the start codon, the RPS26 contacts position −4 from the start codon in yeast (Ferretti *et al*, 2017) and from −11 up to −16 of attached mRNA in mammals (Havkin-Solomon *et al*, 2023). Given that RPS26 may be involved in start codon fidelity (Ferretti *et al*, 2017; Havkin-Solomon *et al*, 2023), we searched for specific sequence motifs in mRNAs encoding proteins responding to RPS26 deficiency. We investigated sequences containing 5'UTRs and coding sequences (CDSs) close to the start codon of mRNAs encoding positive and negative RPS26 responders. We used the total human transcriptome, named background (BG; N = 22,160), as a reference. We did not observe any significant differences in the frequency of individual nucleotide positions in the 20-nucleotide vicinity of the start codon relative to the expected distribution in the BG (**Supplementary Figure 3B**, Supplementary Table 9). However, we identified a significantly higher GC content in the 5'UTR sequences in the positive and negative responder groups relative to the BG at all analyzed positions from −6 to −30 (**Figure 4D**). Although the difference in the GC content in CDSs of responders relative to the BG was much smaller than for the 5'UTRs, we observed a significantly higher (P < 0.05) GC content in the close vicinity of the start AUG codon at upstream positions from +9 up to +14 (**Figure 3D**). Moreover, when we applied more stringent selection criteria (P < 0.01), the number of analyzed transcripts was significantly reduced (positive responders; N = 42, negative responders; N = 54), but the GC richness appeared to be more predominant for mRNAs encoding negative RPS26 responders (**Supplementary Figure 3C**). Additionally, we found that 5'UTRs of proteins encoding positive and negative RPS26 responders were significantly shorter in comparison to reference represented by 1,000 randomly selected transcripts from the complete human transcriptome (**Figure 4E**).

Although bioinformatic analyses did not reveal any importance of position −4 from the start codon in any group of RPS26 responders (**Supplementary Figure 3B**), we experimentally verified whether changes in this position in human cells would affect RAN translation initiation. Using mutagenesis, we substituted G to A in the −4 position from ACG near cognate codon of *FMR1*, as A has previously been shown to be an enhancer of translation initiation in yeast 43S containing Rps26 (Ferretti *et al*, 2017, 2018). We did not observe significant differences in efficiency of FMRpolyG biosynthesis between the two tested mRNAs in human HEK293T cells, confirming our bioinformatic data (**Supplementary Figure 3D**). Additionally, to verify if some of the near-cognate codons are more sensitive to RPS26 depletion, we substituted ACG for GTG or CTG codons in 5'UTR of *FMR1* with 16xCGG in frame with nanoluciferase reporter. We found that RPS26 silencing influenced only the level of FMRpolyG initiated at ACG near cognate start codon (**Figure 4F**). To further investigate importance of codon selection by RPS26, we simultaneously knocked down RPS26 and delivered short antisense oligonucleotides (ASO) blocker targeting ACG codon to cells expressing FMRpolyG. We found that the effect of RPS26 depletion on FMR99xG synthesis was hindered by ASO treatment and the depth of the effect is different than in siCtrl treated cells (**Figure 4G**), suggesting that accessibility of ACG start codon in *FMR1* mRNA is required for RPS26 level-sensitive translation.

We also searched for specific, short sequence motifs, *k*-mers, which might have been enriched in the 5'UTR of mRNAs encoding responders to RPS26 deficiency. We identified a list of 6 and 14 significantly over-represented hexamers ($P < 10^{-5}$) in positive and negative responder groups, respectively. These predominantly comprised G(s) and C(s). For example, the most over-represented hexamers in the 5'UTRs of the positive responders ($P < 10^{-10}$) included GCCGCC, CCGCTG, and CCGGTC, and for negative responders, the highest over-representation ($P < 10^{-6}$) had CGCCGC, GCCGCC, and GCGGCG (**Supplementary Figure 3E**, Supplementary Table 10).

Altogether, these data indicate that RPS26 depletion may impact the translation rate of mRNAs containing GC-rich sequences in shorter 5'UTRs and in close proximity downstream to the AUG start codon (up to 14 nucleotides). It may suggest that thermodynamic stability of RNA structure formed upstream or downstream of a scanning 43S PIC and perhaps the dynamics of translocation of this complex could be a factor modulating sensitivity to RPS26 insufficiency. We did not see any importance of A or G in –4 position from ACG initiation codon on FMRpolyG biosynthesis in cells having or lacking RPS26, however we provided the evidence that cells with downregulated RPS26 display near-cognate start codon selection regulatory properties. Our data showed also that the translation of most human proteins, including FMRP, was unaffected by the impairment of RPS26 level.

TSR2 mediates the RAN translation of FMRpolyG, perhaps via RPS26

The pre-rRNA-processing protein TSR2 (Tsr2) is a chaperone-acting protein that regulates the Rps26 cellular level and is responsible for incorporating Rps26 into 90S pre-ribosome in yeast nuclei (Schütz *et al*, 2018, 2014). RPS26 inclusion into pre-40S together with other factors facilitates final step of 40S subunit maturation (Plassart *et al*, 2021). Moreover, Tsr2 mediates the disassembly of Rps26 from mature 40S subunit in cytoplasm in high salt and pH conditions (Yang & Karbstein, 2022) or when Rps26 is oxidized (Yang *et al*, 2023).

Assuming that depletion of TSR2 in mammalian cells may affect the level of RPS26 loaded on 40S in human cells, we hypothesized that silencing TSR2 would affect the biosynthesis of FMRpolyG. Indeed, in stable *S-95xCGG model* (**Figure 5A**) and COS7 cell line (**Supplementary Figure 4A**) treated with siRNA against TSR2 (siTSR2) we observed lower level of both RPS26, FMR95xG or transiently expressed FMR99xG,respectively,but the level of FMRP was unchanged (**Figure 5A**). The effect of TSR2 silencing on RAN translation was independent of CGG repeat length, as similar results were obtained for smaller FMR16xG proteins (**Supplementary Figure 4B**). Silencing TSR2 did not affect the level of other components of the 40S subunits such as RACK1, RPS6 and RPS15, however, the level of a known responder of RPS26 insufficiency, Histone H3.3 (Li *et al*, 2022), was significantly reduced (**Figure 5A** &B). The observed reduction of FMR95xG and FMR16xG protein by TSR2 depletion did not stem from global translation inhibition, as TSR2 silencing did not impact the overall translation rate (**Figure 5C**). Moreover, to confirm selective impact of RPS26 on polyglycine biosynthesis, we tested the effect of silencing of other small ribosomal subunit protein, RPS6, on RAN translation efficiency and found, that RPS6 depletion did not impact FMR99xG level (**Figure 5C**). Additionally, we verified how TSR2 and RPS6 depletion affects cell proliferation rate. Upon TSR2 knockdown, cell number decreased just by around 15% when comparing to control, but RPS6 insufficiency induced decrease of cell number by ca. 50% (**Figure 5D**).

Obtained results could be explained by the fact that insufficiency of TSR2 may affect the incorporation of RPS26 during nuclear maturation (Schütz *et al*, 2018) or cytoplasmic regeneration of 40S (Yang *et al*, 2023), thus having a negative effect on FMRpolyG biosynthesis, without affecting the global translation. Overall, this data strengthens our conclusion regarding the positive and selective effect of RPS26 and TSR2 on RAN translation initiated from the near-cognate ACG codons located in the 5'UTR of *FMR1*.

The other 40S subunit component, RPS25, affects CGGexp-related RAN translation

Considering data concerning heterogenous ribosomes and their diverse roles in translation regulation (Genuth & Barna, 2018; Shi *et al*, 2017), we hypothesized that other component of the 40S subunit, the ribosomal protein RPS25, which localizes near RPS26 on 40S structure (Pisarev

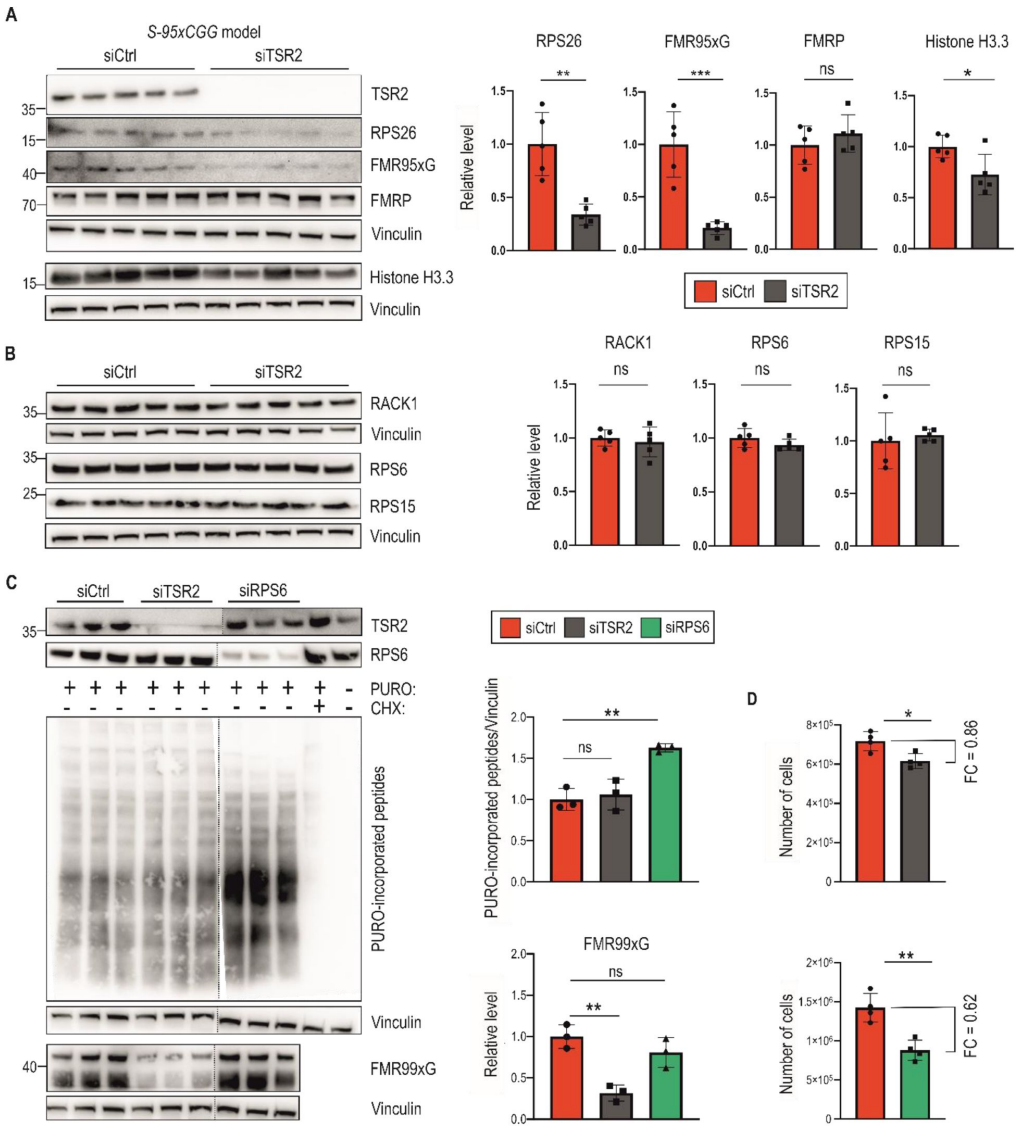


Figure 5.

Insufficiency of the TSR2 chaperone protein lowers the level of RPS26 and FMRpolyG but not FMRP and selected 40S components.

(A) Western blot analyses of RPS26, FMR95xG, FMRP but also Histone H3.3, a sensor of ribosomes depleted with RPS26, normalized to Vinculin in stable *S-95xCGG* cells treated with siTSR2. Graphs represent means from *N* = 5 biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: **, *P* < 0.01; ***, *P* < 0.001; ns, non-significant.

(B) Western blot analyses of selected 40S ribosomal proteins RACK1, RPS6, and RPS15, normalized to Vinculin upon TSR2 silencing in the *S-95xCGG* model. The graphs present means from *N* = 5 biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, *P* < 0.05; ns, non-significant.

(C) Western blot analysis of the results of SUnSET assay (upper) and the level of FMR99xG (lower) 48h post TSR2 or RPS6 silencing in HEK293T cells. CHX treatment served as a positive control for translation inhibition. Negative control (PURO negative and CHX negative) was protein extract from cells treated with DMSO only (see **Figure 2** for more details). Graphs represent means form *N* = 3 independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: **, *P* < 0.01; ns, non-significant. Note that presented blots are cropped. Full size images are deposited on Zenodo server under the DOI: 10.5281/zenodo.13860369.

(D) Graphs represent the number of cells post 48h of TSR2 or RPS6 silencing. Prior siRNA transfection equal amount of HEK293T cells were seeded (~1 × 10^4). The graphs present means from *N* = 4 with SDs. FC, fold change. An unpaired Student's t-test was used to calculate statistical significance: *, *P* < 0.05; **, *P* < 0.01.

et al, 2008; Anger *et al*, 2013), may regulate CGGexp-related RAN translation. Importantly, this protein was already identified as a modifier of RAN translation of mRNAs containing other types of expanded repeats–GGGGCCexp and CAGexp (Yamada *et al*, 2019).

Upon silencing of RPS25 in the stable *S-95xCGG* model we observed significant decline of FMRpolyG level with no effect on its encoding mRNAs (**Figure 6A**). However, we found that RPS25 silencing, like RPS26 silencing, negatively affected HEK293T cells proliferation (**Figure 6B**). Therefore, we also performed similar experiment in human neuroblastoma SH-SY5Y cells, which viability was moderately affected by silencing of all investigated factors (**Figure 6C**). We demonstrated that in SH-SY5Y with transient expression of FMR99xG, the reduction of RAN translation product was comparable after depletion of RPS25 or RPS26 (**Figure 6D**, **Supplementary Figure 5A**). We also found that RPS25 coprecipitates with the 5'UTR of *FMR1* that lacks or harbors expanded 99xCGG repeats via biochemical assay using the biotinylated RNA-protein pull-down assay (**Supplementary Figure 5B**). Altogether, these results suggest that insufficiency of RPS25, like RPS26, is a factor which negatively affects CGGexp-related RAN translation in polyglycine frame.

Discussion

RAN translation was first described in 2011 and reported for Spinocerebellar Ataxia type 8 (SCA8) linked to CAG triplet expansion (Zu *et al*, 2011) and in other repeat expansion-related disorders (REDs), such as Huntington's disease (HD) (Bañez-Coronel *et al*, 2015), C9orf72-linked amyotrophic lateral sclerosis (C9-ALS), and frontotemporal dementia (FTD) (Mori *et al*, 2013; Ash *et al*, 2013), which correlate to CAG and GGGGCC repeats expansions, respectively. RAN translation contribute to the development and progression of many REDs (Zu *et al*, 2011; Mori *et al*, 2013; Ash *et al*, 2013; Bañez-Coronel *et al*, 2015), including FXTAS (Todd *et al*, 2013; Sellier *et al*, 2017) and FXPOI (Buijsen *et al*, 2016); however, no effective therapy targets this pathomechanism. Currently, some RAN translation modifiers of *FMR1* mRNA have been described. For instance, RNA helicases such as DDX3X (Linsalata *et al*, 2019) or DHX36 (Tseng *et al*, 2021), and others extensively reviewed in Baud et al. (Baud *et al*, 2022)) have been identified as affecting CGG repeats-related RAN translation by facilitating ribosomal scanning via unwinding the structured RNA. Components of the endoplasmic reticulum ER-resident kinase (PERK) signaling pathway were shown to modulate RAN translation under stress conditions (Green *et al*, 2017), and the activity of another kinase, SRPK1, retains mutated *FMR1* mRNA containing CGGexp in the nucleus blocking its transport to the cytoplasm and subsequent translation (Malik *et al*, 2021a). Recently, it was demonstrated that RAN translation activates ribosome-associated quality control (RQC) pathway, which prevents accumulation of RAN misfolded proteins by ubiquitination and subsequent proteasomal degradation (Tseng *et al*, 2024). Nevertheless, mechanistic insights into RAN translation remain elusive.

Our research aimed to reveal proteins coprecipitating with the 5'UTR of *FMR1* mRNA containing expanded CGG repeats and discover novel modifiers of CGGexp-related RAN translation. The RNA-tagging approach allowed us to reveal 172 proteins potentially binding to *FMR1* RNA including factors with translation regulatory properties. Some identified proteins overlapped with ones already described as binding to mutant *FMR1* mRNA such as SRPK1 or TAR DNA-binding protein 43 (TDP-43) (Malik *et al*, 2021b; Rosario *et al*, 2022; Sellier *et al*, 2010; Jin *et al*, 2007; Cid-Samper *et al*, 2018; Sellier *et al*, 2017). Further, we described three new modifiers of RAN translation of this mRNA. Depletion of DHX15 helicase and two components of the 40S subunit, RPS26 and RPS25, significantly decreased FMRpolyG levels (**Figure 1C** & **Figure 6A&D**). Importantly, silencing RPS26 alleviated the aggregation phenotype and slowed the apoptosis process caused by FMRpolyG expression (**Figure 1E&F**).

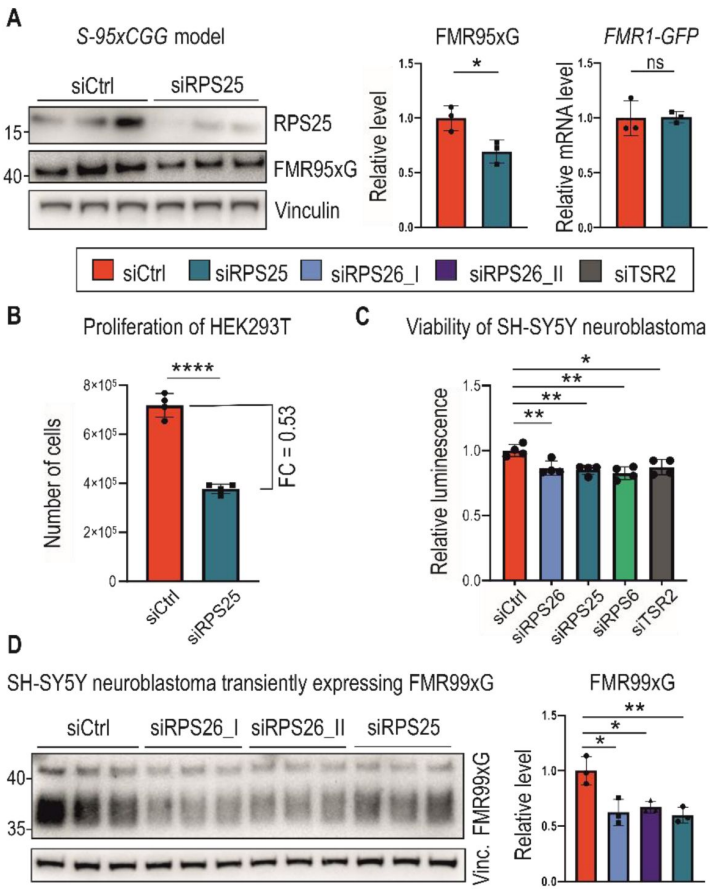


Figure 6.

Silencing RPS25, the other component of 43S PIC, reduces the level of polyglycine produced from mutant *FMR1* mRNA.

(A) A western blot analysis of FMR95xG normalized to Vinculin and an RT-qPCR analysis of the *FMR1-GFP* transgene expression normalized to *GAPDH* upon RPS25 silencing in the *S-95xCGG* model. The graphs present means from $N = 3$ biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$; ns, non-significant.

(B) The graph represents the number of cells post 48h of RPS25 silencing. Prior siRNA transfection equal amount of HEK293T cells were seeded (~1×10^4). The graph presents means from $N = 4$ with SDs. FC, fold change. An unpaired Student's t-test was used to calculate statistical significance: ****, $P < 0.0001$.

(C) The graph represents viability of SH-SY5Y neuroblastoma cells measured by luminescence based on the reducing potential of viable cells. SH-SY5Y cells were treated with all siRNAs tested in this study. The graph presents means from $N = 4$ biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$; **, $P < 0.01$.

(D) A western blot analysis of FMR99xG normalized to Vinculin from the human neuroblastoma cell line (SH-SY5Y) upon silencing of either RPS26 or RPS25. siRPS26_I and siRPS26_II indicate two different siRNAs used for RPS26 silencing. The upper bands were used for quantification. The graph presents means from $N = 3$ biologically independent samples with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$; **, $P < 0.01$; ns, non-significant.

In yeast, the presence or absence of Rps26 in the ribosomal structure leads to the expression of different protein pools influencing temporal protein homeostasis in response to environmental stimuli (Ferretti *et al*, 2017; Yang & Karbstein, 2022). For instance, high-salt and high-pH stress induce the release of Rps26 from mature ribosomes by its chaperone Tsr2, enabling the translation of mRNAs engaged in stress response pathways (Yang & Karbstein, 2022). A previous study indicated that RAN translation in human cells is selectively enhanced by activating stress pathways in a feed-forward loop (Green *et al*, 2017). Therefore, differences in concentration of ribosomal components may influence FMRpolyG production under certain stress conditions.

In eukaryotic cells, RPS26 is involved in stress responses, however on the contrary to yeast model, RPS26 remains associated to the ribosome under energy stress (Havkin-Solomon *et al*, 2023). It was demonstrated that cells with mutated C-terminus of RPS26 were more resistant to glucose starvation, than the wild type cells (Havkin-Solomon *et al*, 2023). Moreover, RPS26 was shown to be involved in other cellular processes such as the DNA damage response (Cui *et al*, 2013), activation of the mTOR signaling pathway (Havkin-Solomon *et al*, 2023) and cellular lineage differentiation by preferential translation of certain transcripts (Piantanida *et al*, 2022; Li *et al*, 2022). We evaluated global changes in the human cellular proteome under RPS26 depletion and found that the expression level was significantly changed in approximately 20% of human proteins, while most proteins' expression levels remained intact (**Figure 3A**, Supplementary Table 7). This suggests that many proteins, including FMRP, are not negatively affected by depleting RPS26, although FMRpolyG biosynthesis appears to be RPS26 level-sensitive. Our observation aligns well with current knowledge about how deficiency of different ribosomal proteins alters translation of some classes of mRNAs (Luan *et al*, 2022; Cheng *et al*, 2019). It was shown that depletion of RPS26 affects translation rate of different mRNAs compared to depletion of other proteins of small ribosomal subunit. We also verified whether RPS26 silencing would affect FMRP endogenously expressed in FXTAS patient-derived and control fibroblasts, as they differ in CGG repeats content in natural locus of *FMR1*, which potentially might influence the effect of RPS26 deficiency. However, we did not observe any changes in FMRP levels upon RPS26 silencing in either genetic variant with the PM or normal *FMR1* allele (**Figure 3F**).

Given the role of TSR2 in incorporating or replacing RPS26 into the ribosome, we investigated whether the depletion of TSR2 modulates FMRpolyG production. In line with previous studies (Schütz *et al*, 2014), we demonstrated that the RPS26 level decreases upon TSR2 silencing. We subsequently showed that TSR2 silencing incurs the same effect on RAN translation efficiency as the direct depletion of RPS26 and the observed regulation does not stem from global translation inhibition (**Figure 5A** &C). However, it remains unclear whether this is an effect of hampering RPS26 loading to the 40S subunit by TSR2 depletion, or a consequence of decrease of RPS26 level. The silencing of this chaperone also exhibited selectivity towards FMRpolyG over the FMRP reading frame without affecting the other 40S ribosomal proteins, such as RACK1, RPS6, and RPS15 (**Figure 5A&B**). Previously, it was shown that Tsr2 is responsible for replacing damaged Rps26, which undergoes oxidation on mature 80S ribosomes under stress conditions (Yang *et al*, 2023). Therefore, the activity of this chaperone may play an important role in modulating RAN translation *via* Rps26 assembly or disassembly from ribosomal subunits, also during different stresses (Yang & Karbstein, 2022). Importantly, we provided additional evidence that RPS26 specifically affects CGGexp-related RAN translation, as the depletion of another 40S ribosomal protein, RPS6, did not affect the FMRpolyG biosynthesis (**Figure 5C**). Moreover, mutations in genes encoding RPS26 and TSR2 were associated with hematopoiesis impairment that underlies the genetic blood disorder Diamond-Blackfan anemia (DBA) (Piantanida *et al*, 2022; Li *et al*, 2023; Doherty *et al*, 2010). Previously, analyses of DBA patient cells and RPS26-depleted HeLa cells found alterations in ribosome biogenesis and pre-rRNA processing (Doherty *et al*, 2010). Similarly, we found that proteins implicated in the eukaryotic PIC and components of large ribosomal subunits were altered upon RPS26 silencing in HEK293T cells (**Figure 4B**), suggesting rearrangements in ribosomal composition. Similar observation was noted in case of RPS25 knockdown which resulted in changes in abundance of ribosomal components (Luan *et al*,

2022). Hence, our data is consistent with the fact that mutants of 40S components accumulates more 60S subunits (Cheng *et al*, 2019) and the knockdown of 40S ribosomal proteins elicits the enhanced translation of 60S ribosomal proteins (Luan *et al*, 2022). Altogether, our data supports the evidence regarding co-dependent regulation of stability and accumulation of 40S and 60S ribosomal proteins (Gregory *et al*, 2019).

Moreover, RPS26 is highly expressed in ovarian cells and was shown to be important for female fertility (Liu *et al*, 2018). It is necessary for oocyte growth and follicle development, and its depletion causes changes in transcription and chromatin configuration, leading to premature ovarian failure (Liu *et al*, 2018). Our discovery that RPS26 regulates the level of FMRpolyG in cellular models sheds new light on the potential role of RPS26-related RAN translation in FXPOI, where FMRpolyG aggregates are detected in ovarian cells and contribute to the development and progression of fertility issues (Buijsen *et al*, 2016; Shelly *et al*, 2021; Rosario *et al*, 2022).

Given these findings, it is likely that the depletion of RPS26 from 40S subunits contributes to its specialization and modulates the translation of specific mRNAs (Yang & Karbstein, 2022; Ferretti *et al*, 2017; Li *et al*, 2022; Gaikwad *et al*, 2021). Remarkably, our experiments identified nearly 400 proteins responding to RPS26 depletion, which is similar to ca. 500 mRNAs, for which translation rate was altered upon Rps26 insufficiency in yeast (Gaikwad *et al*, 2021). In fact, depletion or mutations in RPS26 resulted in the reduction of 40S subunits level, (Gaikwad *et al*, 2021; Havkin-Solomon *et al*, 2023), however, overall translation rate was not impacted by RPS26 impairment (Havkin-Solomon *et al*, 2023). We also observed that RPS26 depletion did not elicit global translation inhibition (**Figure 2B**), which additionally strengthens our conclusion that decrease of polyglycine production is not a byproduct of global translation shut down but rather the specific response to RPS26 deficiency. Alternatively, the observed effect may be explained by the decrease of translationally active ribosomes, which in return, impacts the translation rate of selected mRNAs, including upstream ORF of *FMR1*. In fact, it was shown that RPS26 insufficiency resulted in the decrease of mammalian cells viability (Havkin-Solomon *et al*, 2023), which is in line with our analysis regarding cells proliferation upon RPS26 silencing (**Figure 2A**). Moreover, RPS6 and RPS25 depletion also negatively impacted cells proliferation (**Figure 5D** & 6B), which is consistent with previously described data linking the ribosomal proteins deficiencies (both from 40S and 60S subunits) with a cellular growth rate inhibition (Cheng *et al*, 2019; Luan *et al*, 2022).

It has been shown that the C-terminal domain of RPS26 is essential for mRNA interaction (Havkin-Solomon *et al*, 2023), although whether RPS26 recognizes specific sequential or structural motifs within the mRNA remains unclear. Data derived from yeast models suggest that nucleotides in positions −1 to −10 upstream of the AUG codon, especially Kozak sequence elements, play an important role in Rps26 interactions and translation efficiency (Ferretti *et al*, 2017). According to the established ribosome-mRNA structure, RPS26 contacts with mRNA upstream to AUG codon and the C-terminus reaches into mRNA exit channel (Anger *et al*, 2013; Hussain *et al*, 2014). Recent findings indicate that in eukaryotic cells, positions −11 to −16 upstream of the start codon might be more significant for RNA recognition, stabilizing PIC, and translation initiation (Havkin-Solomon *et al*, 2023). Our bioinformatic analysis of proteins sensitive to RPS26 depletion revealed that their transcripts were GC-rich (especially in the 5'UTR region) and enriched with *k*-mers mainly consisting of Gs and Cs (**Figure 3D**, **Supplementary Figure 3E**), however we did not identify the importance of any specific sequence positions from AUG codon. These data and the fact that *FMR1* 5'UTR is a GC-rich sequence (~90%) suggest that RPS26 level-sensitive translation is selective for transcripts rich with GC nucleotides. Moreover, we demonstrated that RPS26 depletion affected RAN translation initiated from ACG but not CTG and GTG codons (**Figure 4F**), which indicates that the start codon selection step may constitute an important part of RPS26 level-sensitive RAN translation regulation. Such hypothesis can be supported by the fact that blockage of ACG codon with ASO caused less effective RAN translation downregulation evoked by RPS26 depletion (**Figure 4G**). This may point to the importance of specific RNA structural motifs localized within 5'UTRs

or the speed of PIC scanning, which depends on how effectively stable secondary/tertiary structures are resolved. Other explanation would be differences in dynamics of either scanning of PICs differing in the presence of RPS26 or assembling of 80S on the initiation codon.

The ribosomal protein RPS25, a component of 40S subunit was previously described in the context of GGGGCC and CAG repeats-related RAN translation corresponding to C9-ALS/FTD, HD, and SCA (Yamada *et al*, 2019). The depletion of RPS25 in the *Drosophila* C9orf27 model, and in induced motor neurons derived from C9-ALS/FTD patients, alleviated toxicity caused by RAN translation (Yamada *et al*, 2019). On the contrary, in FXTAS *Drosophila* model, RPS25 knockdown led to enhancement of CGG repeats-related toxicity; however, underlying mechanism was not determined (Linsalata *et al*, 2019). Here, we demonstrated that RPS25 coprecipitated with the 5'UTR of *FMR1* and its depletion negatively affected the biosynthesis of FMRpolyG (**Figure 5A&B**, **Supplementary Figure 5**), thus expanding current knowledge about RPS25 RAN translation modulatory properties in REDs.

Altogether, we have identified two ribosomal proteins, RPS26 and RPS25 as modifiers of CGGexp-related RAN translation of mutant *FMR1* mRNA, which imply that the downregulation of these two 40S ribosomal proteins affects FMRpolyG biosynthesis. Importantly, we demonstrated that RPS26 depletion alleviated toxicity caused by FMRpolyG but did not affect FMRP, the main product of the *FMR1* gene. This suggests that sequence/structure elements within *FMR1* mRNA, which make this transcript sensitive to the level of studied ribosomal proteins, may be potential therapeutic targets.

Materials and methods

Genetic constructs

For MS-based screening *FMR1 RNA* construct was generated based on backbone described in (Sellier *et al*, 2017) (see also Addgene plasmid #63091), which contains 5'UTR of *FMR1* with expanded 99xCGG repeats. Enhanced green fluorescent protein (EGFP) sequence was removed and replaced with 4 STOP codons in FMRpolyG frame. Three times repeated MS2 stem loop aptamers (3xMS2 stem loops) were amplified by Phusion polymerase (Thermo Fisher Scientific) with primers introducing EagI restriction site from the plasmid (Addgene, #35572, Tsai et al., 2011 (Tsai *et al*, 2011)). After gel-purification, PCR product was digested with EagI (New England Biolab) and inserted into EagI-digested and dephosphorylated backbone (CIAP; Thermo Fisher Scientific) downstream of 5'UTR of *FMR1* using T4 ligase (Thermo Fisher Scientific). To generate *GC-rich RNA* construct, GC rich sequence (corresponding to *TMEM170* mRNA, Supplementary file - sequences) was amplified by PCR introducing EagI restriction site (CloneAmp HiFi PCR Premix, TakaraBio). PCR product was digested by EagI (New England Biolabs) and ligated into *FMR1 RNA* construct backbone instead of 5'UTR of *FMR1* sequence using T4 ligase (Thermo Fisher Scientific). FMRpolyG-GFP (named here FMR99xG) transient expression was derived from 5'UTR CGG 99x FMR1-EGFP vector (Adggene plasmid #63091), a kind gift from N. Charlet-Berguerand. In case of 5'UTR of *FMR1* with 16xCGG repeats tagged with nanoluciferase, all mutants were obtained using inverse PCR with primers containing 15 nucleotide long overhangs (In-Fusion cloning, Takara Bio). PCR products were run on the agarose gel, appropriate bands were cut, and DNA was purified and used in a reaction with HiFi NEBuilder mix (New England Biolab) according to the manufacturer's instructions. Sequence of all constructs as well as content of CGG repeats were verified by Sanger sequencing.

Generation of cell lines stably expressing FMRpolyG

To generate cell lines with doxycycline-inducible, stable FMRpolyG expression we used the Flp-In™ T-REx™ system. Through this approach, we obtained integration of constructs containing 5'UTR FMR1 with either 95 or 16 CGG repeats fused to *EGFP* into genome of 293 Flp-In® T-Rex® cells. These cell lines were named *S-95xCGG* and *S-16xCGG*, respectively, and expressed RAN proteins

referred to as FMR95xG or FMR16xG. A detailed protocol concerning experimental procedure was described previously by Szczesny et al., 2018 (Szczesny *et al*, 2018). The constructs used in the procedure were generated by modyfing pKK-RNAi-nucCHERRYmiR-TEV-EGFP (Addgene plasmid #105814). The insert containing CGG repeats within the 5'UTR of FMR1 gene fused with *EGFP* sequence was derived from 5'UTR CGG 99x FMR1-EGFP (Addgene plasmid #63091). The insert was ligated into the vector instead of *EGFP* sequence. In the second approach, we used lentiviral transduction system using constructs containing 5'UTR of *FMR1* with 99 CGG repeats fused to EGFP followed by T2A autocleavage peptide and puromycin resistance (FMRpolyG-GFP_T2A_PURO). Cloning FMRpolyG-GFP_T2A_PURO sequence into TetO-FUW-AscII-Puro vector (Addgene plasmid #97329, Yang et al., 2017 (Yang *et al*, 2017)) as well as lentiviral particles production was performed by The Viral Core Facility, part of the Charité – Universitätsmedizin Berlin. For transduction procedure 0.15×10^6 HEK293T cells were plated in 6-well plates at ~60% confluency and incubated with lentiviral particles for 48h. Subsequently, in order to obtain the pool of cells with integrated transgene, selection with puromycin (in final concentration 1 µg/ml, Sigma) was initiated and lasted for 72h, until all transgene-negative cells died. After selection we obtained polyclonal cell line named *L-99xCGG* expressing FMR99xG.

Cell culture and transfection

The monkey COS7, human HEK293T, HeLa and SH-SY5Y cells were grown in a high glucose DMEM medium with L-Glutamine (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS; Thermo Fisher Scientific) and 1x antibiotic/antimycotic solution (Sigma). *S-95xCGG* and *S-16xCGG* cells were grown in DMEM medium containing certified tetracycline-free FBS (Biowest). Fibroblasts derived from FXTAS male patient (1022-07) with $(CGG)_{81}$ in *FMR1* gene, and control, non-FXTAS male individual (C0603) with $(CGG)_{31}$ in *FMR1* were cultured in MEM medium (Biowest) supplemented with 10% FBS (Thermo Fisher Scientific), 1% MEM non-essential amino acids (Thermo Fisher Scientific) and 1x antibiotic/antimycotic solution (Sigma). All cells were grown at 37°C in a humidified incubator containing 5% CO_2. FXTAS 1022-07 (XY, 81xCGG, age 68 years old) line was a kind gift from P. Hagerman (Garcia-Arocena *et al*, 2010) while control C0603 fibroblast line (XY, 31xCGG) were given by A. Bhattacharyya (Rovozzo *et al*, 2016). For the delivery of siRNA with the final concentration in culture medium ranging from 15 to 25 nM, reverse transfection protocol was applied using jetPRIME® reagent (Polyplus) with the exception of fibroblasts, which were plated on the appropriate cell culture vessels the day before the transfection. Plasmids were delivered 24h post siRNA treatment, and the DNA/jetPRIME® reagent (Polyplus) ratio 1:2 was applied. For the experiment from **figure 4G**, three hours post plasmid encoding FMR99xG delivery, cells were transfected with ASOs with the final concentration of 200 nM, composed of LNA units. ASO_ACG sequence: 5'-CGTCACCGCCCG-3' and ASO_CTRL 5'-TTGAACATAAG-3'. Cells were harvested 48h post siRNA silencing and 24h post transient plasmids expression. The list of all siRNAs used in the study is available in a Supplementary Table 1.

Mass spectrometry-based proteins screening

To capture RNA-protein complexes natively assembled within *FMR1 RNA* we adapted the MS2 *in vivo* biotin tagged RNA affinity purification (MS2-BioTRAP) technique, originally published by Tsai et al., 2011 (Tsai *et al*, 2011). The principle of this technique is to co-express bacteriophage MS2 protein fused to a HB tag which undergoes biotinylation *in vivo*, and RNAs tagged with so-called MS2 stem loop RNA aptamers, towards which MS2-HB protein represent high affinity. Natively assembled RNA-protein complexes can be then fixed and identified with HB-tag based affinity purification using streptavidin-conjugated beads. To perform the screening, 2×10^6 HEK293T cells were co-transfected with genetic constructs: 2 µg of MS2-HB plasmid (#35573, Tsai et al., 2011 (Tsai *et al*, 2011)) along with 8 µg of *FMR1 RNA* or *GC-rich RNA* encoding vectors (described in detail above). Three 10 cm plates were used per given mRNA with MS2 stem loop aptamers. 24h post co-transfections cells were washed 2 times with ice-cold phosphate buffered saline (PBS) and crosslinked with 0.1% formaldehyde (ChIP-grade, Pierce) for 10 min, followed by 0.5 M glycine quenching for 10 min in room temperature (RT). After 1 wash with ice-cold PBS cells

were lysed for 30 min on ice in cell lysis buffer (50 mM Tris-Cl pH 7.5, 150 mM NaCl, 1% Triton X-100, 0.1% Na-deoxycholate) with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific) and RNAsin Plus (Promega) and vortexed. Next, cells lysates were sonicated (15 cycles: 10 sec. on/10 sec. off) using sonicator (Bioruptor, Diagenode) and centrifuged at 12,000 g for 10 min at 4°C. Precleared protein extracts were transferred to Protein Lobind tubes (Eppendorf) and incubated with pre-washed 50 µl of magnetic-streptavidin beads (MyOne C1, Sigma) for 1.5h in cold room rotating. After immunoprecipitation (IP) procedure beads were washed 4-times for 5 min each on rotator: 1-time with 2% sodium dodecyl sulfate (SDS), 1-time with cell lysis buffer, 1-time with 500 mM NaCl, and last wash with 50 mM Tris-Cl pH 7.5. Finally, 20% of IP fraction was saved for western blot and remaining beads were submerged into digestion buffer (6 M Urea, 2 M Thiourea, 100 mM Tris, pH 7.8 and 2% amidosulfobetaine-14) and were shaken for 1h at RT. Then, samples were reduced using dithioerythritol for 1h at RT, and alkylated with iodoacetamide solution for 30 min in dark. Then, Trypsin/Lys-C (Promega) solution was added, and samples were incubated for 3h at 37°C, followed by adding fresh Milli-Q water to dilute Urea to ~1 M and samples were further incubated at 37°C overnight. Finally, beads were removed on a magnet, and peptides transferred to the clean tube, and desalted using C18 Isolute SPE columns (Biotage). Samples were analyzed in Mass Spectrometry Laboratory, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5a Street, 02-106 Warsaw, Poland, using LC-MS system composed of Evosep One (Evosep Biosystems, Odense, Denmark) directly coupled to a Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific, USA). Peptides were loaded onto disposable Evotips C18 trap columns (Evosep Biosystems, Odense, Denmark) according to manufacturer protocol with some modifications. Briefly, Evotips were activated with 25 µl of Evosep solvent B (0.1% formic acid in acetonitrile, Thermo Fisher Scientific, USA), followed by 2 min incubation in 2-propanol (Thermo Fisher Scientific, USA) and equilibration with 25 µl of solvent A (0.1% FA in H_2O, Thermo Fisher Scientific, Waltham, Massachusetts, USA) Chromatography was carried out at a flow rate 220 nL/min using the 88 min gradient on EV1106 analytical column (Dr Maisch C18 AQ, 1.9 µm beads, 150 µm ID, 15 cm long, Evosep Biosystems, Odense, Denmark). Data was acquired in positive mode with a data-dependent method. MS1 resolution was set at 60,000 with a normalized AGC target 300%, Auto maximum inject time and a scan range of 350 to 1,400 m/z. For MS2, resolution was set at 15,000 with a Standard normalized AGC target, Auto maximum inject time and top 40 precursors within an isolation window of 1.6 m/z considered for MS/MS analysis. Dynamic exclusion was set at 20 s with allowed mass tolerance of ±10 ppm and the precursor intensity threshold at 5e3. Precursor were fragmented in HCD mode with normalized collision energy of 30%.

Biotinylated RNA-protein pull down

Biotinylated RNA probes were produced by *in vitro* transcription using T7 RNA polymerase (Promega), and by adding 1:10 CTP-biotin analog:CTP to the reaction mix, to incorporate biotinylated cytidines in random manner. Cellular extracts were prepared by lysing 2.5×10^6 HEK273T cells in 200 µL of mammalian cell lysis buffer. Lysates were cleared by centrifugation and 200 µL of supernatant was incubated for 20 min at 21°C with 5 µg of RNA in 200 µL of 2xTENT buffer (100 mM Tris pH 7.8, 2 mM EDTA, 500 mM NaCl, 0.1% Tween) supplemented with RNAsin (Promega). RNA-protein complexes were then incubated with MyOne Streptavidin T1 DynaBeads (Thermo Fisher Scientific) for 20 min, followed by washing steps in 1x TENT buffer. Bound proteins were released in Bolt LDS Sample Buffer (Thermo Fisher Scientific) followed by heat denaturation at 95°C for 10 min. Samples were analyzed by SDS-PAGE and western blotting.

Cell proliferation assay

Cells were seeded at the density of ca. 1×10^5 cells per well and treated with siRNA using reverse transfection protocol. Forty eight or fifty hours post silencing, cells were harvested and incubated with trypan blue (Thermo Fisher Scientific). Subsequently, alive and dead cells were calculated using Countess 3 (Thermo Fisher Scientific) equipment.

Cell viability assay

Cells were seeded at the density of 1×10^5 cells per well and treated with siRNA using reverse transfection protocol. Subsequently, reagents from RealTime-Glo™ MT Cell Viability Assay (Promega) were added to the culture. Measurement of luminescence signal corresponding to cell viability was taken upon 24 and 48h post siRNA treatment using SPARK microplate reader (TECAN).

Apoptosis assay

For luminescent-based apoptosis assay 1×10^4 HeLa cells were seeded on 96-well plate and transfected with siRPS26 and siCtrl in final concentration of 15 nmol. In order to induce FMRpolyG derived toxicity, after 24h post silencing, cells well transfected with plasmid encoding FMR99xG or with jetPRIME® reagent (Polyplus) as a MOCK control. Subsequently, reagents from RealTime-Glo™ Annexin V Apoptosis Assay (Promega) were added to the culture. Measurement of luminescence signal corresponding to apoptosis progression was taken upon 28 and 43h post FMR99xG expression using SPARK microplate reader (TECAN).

Nanoluciferase assay

For luminescence-based analysis of RAN translation post RPS26 silencing, 1×10^4 of HEK293T cells were seeded on 94 well plate and transfected with siRPS26 using reverse transfection protocol. On the next day, cells were cotransfected with constructs expressing FMRpolyG (16xCGG) in frame with nanoluciferase together with the construct expressing firefly luciferase used for normalization purposes. 24h post plasmid delivery, cells were lysed for 1h on ice in RIPA buffer (Merck). Subsequently, 1/5 of the lysate was transferred to black 96 well plate (Thermo Fisher Scientific) and incubated with ONE-Glo^{TM} EX reagent (Promega). After 10 min incubation, firefly luciferase-based luminescence was measured by SPARK microplate reader (TECAN). Next, the NanoDLR^{TM} Stop & Glo (Promega) buffer containing substrate for nanoluciferase was added and after 10 min incubation, the nanoluciferase-based luminescence was measured by the same equipment.

Microscopic analysis

To detect aggregated form of FMRpolyG fluorescence microscopy experiments were performed as described previously (Derbis *et al*, 2018). Briefly, before analysis, HeLa cells were grown on 48-well plates in standard growth medium, siRNA and plasmid transfections were performed as described in Cell culture and transfection section. Prior microscopic analysis, final concentration of 5 µg/ml of Hoechst 33342 (Thermo Fisher Scientific) was added to cell culture and incubated for 10 min. Images were taken with Axio Observer.Z1 inverted microscope equipped with A-Plan 10×/0.25 Ph1 or LD Plan-Neofluar 20×/0.4 Ph2 objective (Zeiss), Zeiss Colibri 7 excitation band 385/30 nm, emission filter 425/30 nm (Hoechst) and Zeiss Colibri 7 excitation band 469/38 nm, emission filter 514/30 nm (GFP), Zeiss AxioCam 506 camera and ZEN 2.6 pro software, 48h post siRNA transfection. Presented values were quantified from 10 microscopic images for each condition, number of cells (blue signal) and aggregates (dense green signal) were calculated using ImageJ and AggreCount plugin (Klickstein *et al*, 2020), which localizes cellular aggregates in relation to the nuclei. Each image contained ca. 1,100 cells and the number of aggregates per nucleus were calculated as follows: nuclei were processed via fluorescent signal enhancement, converted to binary image and region of interest (ROI) were captured, size filter was set to 100-6×10^6 voxels; aggregates were detected used difference of gaussians methodology, size filter was set to 10-500 voxels. Cells were segmented using the watershed-based segmentation algorithm. For each image, the number of detected aggregates was divided per number of nuclei, and the results for all images were plotted on a histogram. To validate FMR95xG or FMR16xG expression in *S-95xCGG* and *S-16xCGG* models, microscopic analysis was performed at two time points: 4 days and 35 days post doxycycline (DOX) induction. 4 days post doxycycline induction cells were incubated

in a cell culture medium with Hoechst 33342 (Thermo Fisher Scientific) at a final concentration 5 µg/ml for 30 min at 37°C. Images were taken as described above. 35 days post doxycycline induction images were captured with Leica Stellaris 8 Inverted Confocal Microscope equipped with HC PL APO CS2 63x/1.20 water objective and an onstage incubation chamber controlling temperature and CO_2 concentration. GFP was excited with 489 nm laser and detected with Power HyD S detector (spectral positions: 494 nm - 584 nm). To visualize FMRpolyG positive aggregates the *L-99xCGG* model cells were incubated with Hoechst 33342 (Thermo Fisher Scientific) at a final concentration 5 µg/ml for 15 min at 37°C. Images were taken with Leica SP8 confocal microscope equipped with White Light Laser 2 PP, Diode Laser 405 nm, HyD S detectors, and high precision objective (PL APO 63x/1.20 W CORR CS2). GFP was excited with 489 nm and detected in range of 500-550 nm. Hoechst was excited with 405 nm laser lines and detected in 420-470 nm detector range.

RNA isolation and quantitative real-time RT-PCR

Cells were harvested in TRI Reagent (Thermo Fisher Scientific) and total RNA was isolated with Total RNA Zol-Out D (A&A Biotechnology) kit according to the manufacturer's protocol. 500-1,000 ng of RNA was reversely transcribed using GoScript™ Reverse Transcriptase (Promega) and random hexamers (Promega). Quantitative real-time RT-PCRs were performed in a QuantStudio 7 Flex System (Thermo Fisher Scientific) using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher Scientific) with 5 ng of cDNA in each reaction. Transgene *FMR1-GFP* mRNAs with expanded CGG repeats were amplified with primers: Forward: 5' GCAGCCCACCTCTCGGGG 3', Reverse: 5' CTTCGGGCATGGCGGACTTG 3' with a note that reverse primer was anchored in GFP sequence in order to distinguish endogenously expressed *FMR1* transcripts (amplified with primer pair: Forward: 5' TGTGTCCCCATTGTAAGCAA 3', Reverse: 5' CTCAACGGGAGATAAGCAG 3'). Reactions were run at 58°C annealing temperature and Ct values were normalized to *GAPDH* mRNA level (amplified with primer pair: Forward: 5' GAGTCAACGGATTTGGTCGT 3', Reverse: 5' TTGATTTTGGAGGGATCTCG 3'). Fold differences in expression level were calculated according to the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001).

SDS-PAGE and Western blot

Cells were lysed in cell lysis buffer supplemented with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific) for 30 min on ice, vortexed, sonicated and centrifuged at 12,000 g for 10 min at 4°C. Protein lysates were heat-denatured for 10 min at 95°C with the addition of Bolt LDS buffer (Thermo Fisher Scientific) and Bolt Reducing agent (Thermo Fisher Scientific) and separated in Bolt™ 4–12% Bis-Tris Plus gels (Thermo Fisher Scientific) in Bolt™ MES SDS Running Buffer (Thermo Fisher Scientific). Next, proteins were electroblotted to PVDF membrane (0.2 µM, GE Healthcare) for 1 h at 100 V in ice-cold Bolt™Transfer Buffer (Thermo Fisher Scientific). Membranes were blocked in room temperature for 1 h in 5% skim milk (Sigma) in TBST buffer (Tris-buffered saline [TBS], 0.1% Tween 20) and subsequently incubated overnight in cold room with primary antibody solutions diluted in blocking buffer (the list of all commercially available primary antibodies used in the study including catalog numbers is presented in a Supplementary Table 2, with the exception for home-made anti-FUS antibody (Raczynska *et al*, 2015)). The following day membranes were washed 3-times with TBST for 7 min and incubated with corresponding solutions of anti-mouse (A9044, Sigma; 1:15,000) or anti-rabbit (A9169, Sigma; 1:20,000) antibodies conjugated with horse radish peroxidase (HRP) for 1h in RT. For Vinculin and GFP detection, membranes were incubated with antibodies already conjugated with HRP (Supplementary Table 2) overnight in cold room. After final washing steps signals were developed using Immobilon Forte Western HRP substrate (Sigma) using G:Box Chemi-XR5 (Syngene) and ChemiDoc Imaging System (BioRad) and quantified using Multi Gauge 3.0 software (Fujifilm). Briefly, area of bands corresponding to proteins of interest and to background were manually selected for measurement. Then, the signal of background was subtracted and the percentage (gray-scale density) of each band was calculated. Relative protein level was normalized to Vinculin.

Surface sensing of translation (SuNseT) assay

48h post siRNA transfection HEK293T cells were incubated with puromycin (PURO; Sigma) at a final concentration of 10 µg/ml, which allowed to incorporate PURO into newly synthesized peptides to asses global translation rate (Schmidt *et al*, 2009). After 10 min incubation with PURO cells were subsequently harvested. For positive control of translation inhibition cells were incubated for 5 min with cycloheximide (CHX; 100 µg/ml) prior addition of puromycin. DMSO treated cells without puromycin incubation served as a negative control. Protein lysates were prepared as described above. Next, protein concentration was determined by BCA assay (Pierce) and 10 µg of protein lysate was loaded on the Bolt™ 4–12% Bis-Tris Plus gel (Thermo Fisher Scientific). Following steps of western blot were described as above. Anti puromycin antibody (1:5000 (clone 12D10, Merck), overnight incubation at 4°C) was used to detect puromycin-incorporated peptides.

Stable isotope labeling using amino acids in cell culture (SILAC) coupled with MS

To quantify changes in protein levels upon RPS26 silencing we applied SILAC in HEK293T cells grown in light and heavy amino acids ($^{13}C_6$ $^{15}N_2$ L-lysine-2HCl and $^{13}C_6$ $^{15}N_4$ L-arginine-HCl) containing media (Thermo Fisher Scientific). Cells were cultured in light or heavy media for more than 14 days to ensure maximum incorporation of labeled amino acids. Subsequently, cells were transfected with siRPS26 or siCtrl in final concentration of 15 nM and harvested 48h post silencing. Sample preparation was performed using modified filter-aided sample preparation method as described previously (Laakkonen *et al*, 2017; Wiśniewski *et al*, 2009). Briefly, 10 µg of proteins were washed 8-times with 8 M Urea, 100 mM ammonium bicarbonate in Amicon Ultra-0.5 centrifugal filters, and Lysine-C endopeptidase solution (Wako) in a ratio of 1:50 w/w was added to the protein lysates followed by incubation at room temperature overnight with shaking. The peptide digests were collected by centrifugation and trypsin solution was added in a ratio of 1:50 w/w in 50 mM ammonium bicarbonate and incubated overnight at room temperature. The peptide samples were cleaned using Pierce C18 reverse-phase tips (Thermo Fisher Scientific). Dried peptide pellets were re-suspended in 0.3% Trifluoroacetic acid and analyzed using nano-LC-MS/MS in Meilahti Clinical Proteomics Core Facility, University of Helsinki, Helsinki, Finland. Peptides were separated by Ultimate 3000 LC system (Dionex, Thermo Fisher Scientific) equipped with a reverse-phase trapping column RP-2TM C18 trap column (0.075 × 10 mm, Phenomenex, USA), followed by analytical separation on a bioZen C18 nano column (0.075 × 250 mm, 2.6 µm particles; Phenomenex, USA). The injected samples were trapped at a flow rate of 5 µl/min in 100% of solution A (0.1% formic acid). After trapping, peptides were separated with a linear gradient of 125 min. LC-MS acquisition data was performed on Thermo Q Exactive HF mass spectrometer with following settings: resolution 120,000 for MS scans, and 15,000 for the MS/MS scans. Full MS was acquired from 350 to 1400 m/z, and the 15 most abundant precursor ions were selected for fragmentation with 45 s dynamic exclusion time. Maximum IT were set as 50 and 25 ms and AGC targets were set to 3 e6 and 1 e5 counts for MS and MS/MS respectively. Secondary ions were isolated with a window of 1 m/z unit. The NCE collision energy stepped was set to 28 kJ mol–1.

MS Data analysis

Raw data obtained from the LC-MS/MS runs were analyzed in MaxQuant v2.0.3.0 (Cox & Mann, 2008) using either the label-free quantification (LFQ, for MS2 pull down samples) or stable isotope labeling-based quantification (for SILAC-MS samples) with default parameters. UniProtKB database for reviewed human canonical and isoform proteins of May 2023 was used. The false discovery rate (FDR) at the peptide spectrum matches and protein level was set to 0.01; variable peptide modifications: oxidation (M) and acetyl (N-term), fixed modification: carbamidomethyl (C), label Arg10, Lys8 (for SILAC-MS samples only), two missed cleavages were allowed. Statistical analyses were performed using Perseus software v2.0.3.0 (Tyanova *et al*, 2016) after filtering for

"reverse", "contaminant" and "only identified by site" proteins. The LFQ intensity was logarithmized (log2[x]), minimum valid values were set to 70% in order to remove proteins which were not quantified, and imputation of missing values was performed with a normal distribution (width = 0.3; shift = 1.8). Proteomes were compared using t-test statistics with a permutation-based FDR of 5% and *P*-values <0.05 were considered to be statistically significant. The data sets, the Perseus result files used for analysis, and the annotated MS/MS spectra were deposited at the ProteomeXchange Consortium (Deutsch *et al*, 2023) *via* the PRIDE partner repository (Perez-Riverol *et al*, 2022) with the dataset identifier PXD047400 - MS2 pull down data and PXD047397 – SILAC-MS data.

Gene ontology analysis

Gene ontology (GO) analysis performed on proteins that bind to *FMR1 RNA* (common between three replicates) was performed using an online tool - g:Profiler (Raudvere *et al*, 2019) with g:SCS algorithm, where at least 95% of matches above threshold are statistically significant. As reference proteome we used total human proteome. GO analysis performed on SILAC-MS data was performed with PANTHER 18.0 (Thomas *et al*, 2022). Statistical significance was calculated using Fisher's Exact test with Bonferroni correction and only GO terms with *P*-values < 0.05 were plotted. As reference proteome we used proteins named as non-responders (described in Results section).

Bioinformatic analyses

The sequence logo analysis represents the frequency of nucleotides across transcript sequence positions in the close vicinity of start codon within three groups of transcripts: positive responders, negative responders, and background (BG, total transcriptome). The analyzed sequence positions span the last 20 nucleotides of the 5'UTR sequence (from −20 to −1 downstream of the start codon) and the first 20 nucleotides of the coding sequence (from +1 to +20 upstream of the start codon), forming 40-nucleotide sequence fragments that were used to generate the sequence logos with WebLogo v2.8.2 (Crooks et al, 2004). To calculate the GC content of 5'UTRs and CDSs, gene annotations for all protein-coding human genes (GRCh38.p14), including coding and 5'UTR sequences of transcripts, were obtained from Ensembl release 110 (Nov2023) (Martin et al, 2023). UniProt protein accessions for positive and negative responders were mapped to corresponding Ensembl genes and transcripts using BioMart (Smedley et al, 2009). The reference BG dataset comprised all protein-coding genes, excluding positive and negative responder genes, with one randomly selected transcript per gene. P-values for pairwise comparisons of GC content between transcript groups were calculated using a two-tailed paired t-test with Bonferroni correction. For 5'UTRs lengths, a balanced comparison was ensured by randomly selecting 1,000 transcripts from the complete human transcriptome to represent background. *P*-values for comparing 5'UTR sequence lengths between positive and negative responders relative to the background transcriptome were calculated using a one-tailed Mann-Whitney U test. Hexamer frequencies in 5'UTR sequences of the three datasets (i.e., positive and negative responders, and BG) were calculated using Jellyfish v2.3.1 (Marçais & Kingsford, 2011). The *P*-value associated with each hexamer in the positive and negative responder datasets, representing its overrepresentation compared to BG, was calculated from the binomial distribution. All statistical analyzes related to the comparison of nucleotide composition among the BG, positive and negative responder datasets were performed using SciPy v1.10.1 (Virtanen et al, 2020).

Statistics

Group data are expressed as the means ± standard deviation (SD). Error bars represent SD. The statistical significance (if not indicated otherwise) was determined by an unpaired, two-tailed Student's t-test using Prism software v.8 (GraphPad): ∗, $P < 0.05$; ∗∗, $P < 0.01$; ∗∗∗, $P < 0.001$; ****, $P < 0.0001$; ns, non-significant. All experiments presented in this work were repeated at least two times with similar results with at least three independent biological replicates ($N = 3$).

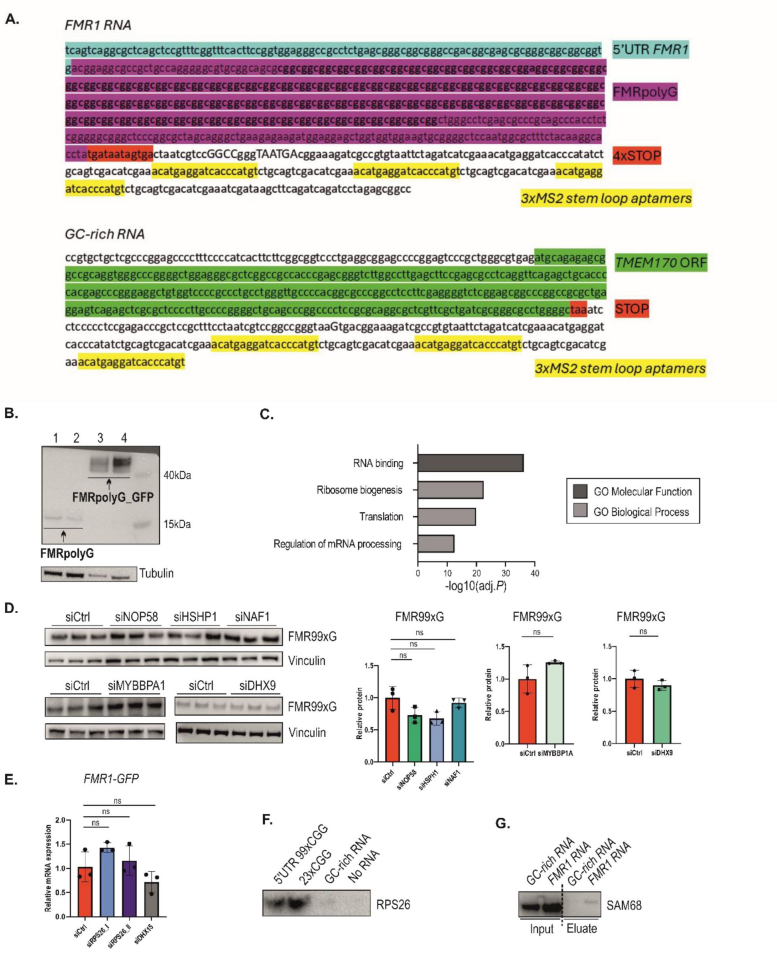
Data availability

The data underlying this article are available in ProteomeXchange Consortium via the PRIDE partner repository and can be accessed with the dataset identifiers:

- PXD047400 - MS2 pull down data
- PXD047397 - SILAC-MS data.

Supplementary data is deposited on Zenodo server under the DOI: *10.5281/zenodo.13860369*.

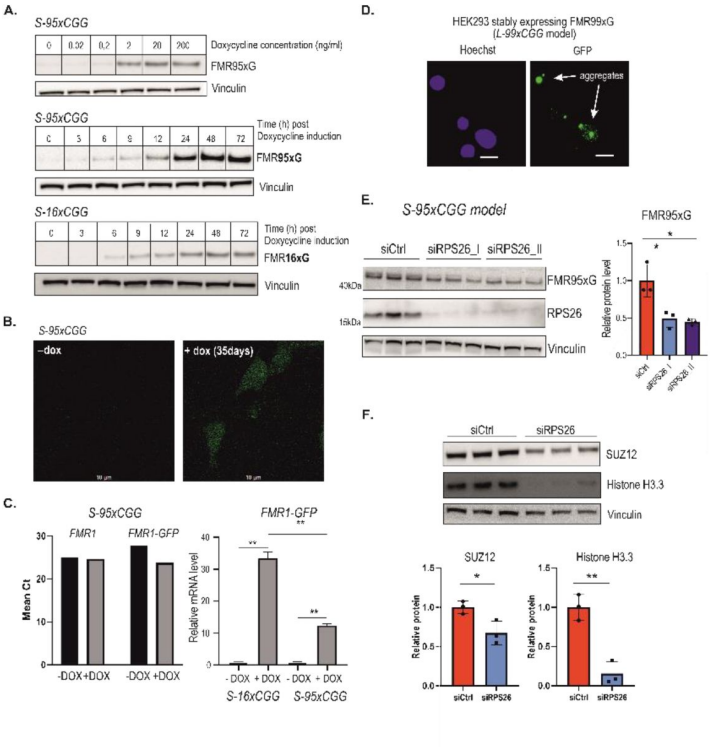
Supplementary figures



Supplementary Figure S1.

MS-based screening revealed proteins interacting with *FMR1 RNA*.

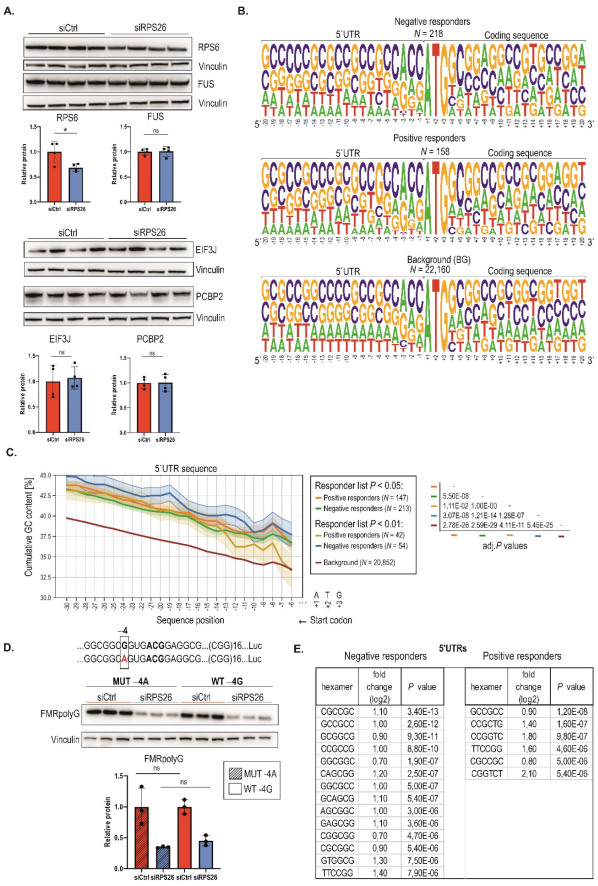
(A) Sequences of *FMR1 RNA* and *GC-rich RNA* constructs used in MS2-based screening. Note that the sequences presented are a fragment of the plasmids used in the experiment.
(B) Results of western blot representing FMRpolyG derived from *FRM1 RNA* construct (lane 1&2) in comparison to FMRpolyG tagged with GFP (lane 3&4) derived from template construct (Addgene #63089). Tubulin was used as a loading control.
(C) Gene Ontology (GO) analysis performed on proteins detected in *FMR1 RNA* sample group (common among three technical replicates). Graph presents significantly enriched, selected GO terms.
(D) Results of western blot analysis of FMR99xG normalized to Vinculin upon NOP58, HSHP1, NAF1, MYBBP1A and DHX9 siRNA-based silencing in HEK293T cells. Graphs present means from *N* = 3 biologically independent samples with SDs. Unpaired Student's t test was used to calculate statistical significance: ns, non-significant. Note that presented blots are cropped.
(E) RT-qPCR analysis of *FMR1-GFP* transgene expression derived from transient transfection system normalized to *GAPDH* upon RPS26 and DHX15 depletion in HEK293T cells. siRPS26_I and siRPS26_II indicate two independent siRNAs. Graphs present means from *N* = 3 biologically independent samples with SDs. Unpaired Student's t test was used to calculate statistical significance: ns, non-significant
(F) Results of western blot demonstrating enrichment of RPS26 detected in 5'UTR 99xCGG and 23xCGG eluates compared to GC-rich RNA in biochemical assay. 50% of eluate fraction was loaded on the gel. All RNA molecules in the experiment were biotinylated (b) in order to perform RNA-protein pull down procedure from HEK293T cell extract. No RNA sample served as a negative control.
(G) Results of western blot demonstrating enrichment of SAM68 protein detected in *FMR1 RNA* sample eluate after MS2-based, *in cellulo* RNA-protein pull down procedure; the percentage of following fractions were loaded on the gel: Input, 5% of total lysate; eluate, 20% of immunoprecipitated fraction. Note that presented blot is cropped.



Supplementary Figure S2.

Characterization of *S-95xCGG* and *S-16xCGG* models and RPS26-responders.

(A) Results of western blot of FMR95xG and FMR16xCGG expression in *S-95xCGG* and *S-16xCGG* models, respectively, after induction of doxycycline-inducible promoter. Upper image represents the titration of doxycycline (in ng/ml). Following images demonstrate time-dependent course of FMR95xG/FMR16xCGG expression post promoter induction. The same amount of protein sample was loaded on gels. Vinculin serves as a loading control.
(B) The representative confocal microscopic images showing homogenous expression of FMR95xG fused with eGFP in *S-95xCGG* model after 35 days of doxycycline treatment; scale bar 10 µm. Note that no GFP-positive aggregates of polyglycine were detected.
(C) RT-qPCR analysis of endogenous *FMR1* and *FMR1-GFP* transgene expression in *S-95xCGG* and *S-16xCGG* models with (+DOX) or without (-DOX) doxycycline induction. On the left, bars represent the mean cycle (Ct) for *S-95xCGG* model, and on the right, bars indicate relative *FMR1-GFP* mRNA level normalized to *GAPDH* for both models. Unpaired Student's t test was used to calculate statistical significance: **, $P < 0.01$.
(D) Images of *L-99xCGG* model were taken with Nikon A1 confocal microscope, equipped with GaAsP Detectors. GFP was excited with 488 nm ion argon Laser and detected with 525/50 Detector. Arrows indicate FMRpolyG-GFP-positive aggregates (green); blue are nuclei; scale bar 10 µm.
(E) Results of western blot analysis of FMR95xG normalized to Vinculin upon RPS26 silencing in *S-95xCGG*. siRPS26_I and siRPS26_II indicate two different siRNAs used in the study. The graphs present means from $N = 3$ with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, $P < 0.05$.
(F) Results of western blot analysis of SUZ12 and Histone H3.3 proteins normalized to Vinculin upon RPS26 silencing in COS7 cells. Graphs present means from $N = 3$ biologically independent samples with SDs. Unpaired Student's t test was used to calculate statistical significance: *, $P < 0.05$; **, $P < 0.01$.



hexamer	fold change (log2)	P value
CGCCGC	1,10	3,40E-13
GCCGCC	1,00	2,60E-12
CCGGCG	0,90	9,30E-11
CCGCCG	1,00	8,80E-10
GGCGGC	0,70	1,90E-07
CAGCGG	1,20	2,50E-07
GGCGCC	1,00	5,00E-07
GCAGCG	1,10	5,40E-07
AGCGGC	1,00	3,00E-06
GAGCGG	1,10	3,60E-06
CGGCGG	0,70	4,70E-06
CGCGGC	0,90	5,40E-06
GTGGCG	1,30	7,50E-06
TTCCGG	1,40	7,90E-06

hexamer	fold change (log2)	P value
GCCGCC	0,90	1,20E-08
CCGCTG	1,40	1,60E-07
CCGGTC	1,80	9,80E-07
TTCCGG	1,60	4,60E-06
CGCCGC	0,80	5,00E-06
CGGTCT	2,10	5,40E-06

Supplementary Figure S3.

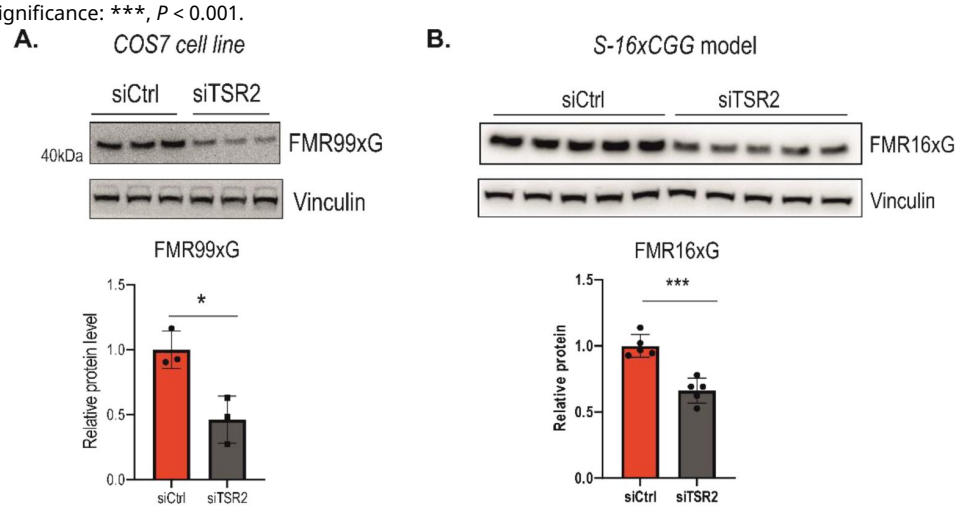
Validation of SILAC-MS results and bioinformatic analysis of mRNAs encoding RPS26-sensitive proteins.

(A) Validation of SILAC-MS data. Results of western blot analysis of RPS6 & FUS (results are in line with proteomic analysis) and PCBP2 & EIF3J (results are not in line with proteomic analysis) normalized to Vinculin upon RPS26 silencing in HEK293T cells. Graphs present means from *N* = 4 biologically independent samples with SDs. Unpaired Student's t test was used to calculate statistical significance: *, $P < 0.05$; ns, non-significant.
(B) Frequency of nucleotides across transcript sequence positions in the close vicinity of start codon within three groups of transcripts: negative responders (*N* = 218), positive responders (*N* = 158), and background transcripts (*N* = 22,160). The analyzed sequence positions span the last 20 nucleotides of the 5' UTR sequence (from 20 to 1 downstream of the start codon) and the first 20 nucleotides of the coding sequence (from +1 to +20 upstream of the start codon). Each nucleotide symbol's height on the graph is proportionate to its relative frequency at the corresponding sequence position. The cumulative height of the stack at each position sums to one, encompassing the frequencies of all four nucleotides.
(C) The percentage of GC content of 5'UTR across extending sequence windows initiated upstream from the start codon within groups of transcripts. The graph gathers analysis of transcripts determined with two *P*-value cut offs ($P < 0.05$ and $P < 0.01$) yielding different samples sizes for each group. Colors corresponding to given group as well as its size are indicated in the legend. The solid line shows the mean GC content at a given position (i.e., within the window), and the shade indicates the standard error of the mean. adj. *P* values (right panel) reflect pairwise comparisons of GC content between transcript groups and were determined using a two-tailed paired t-test with Bonferroni correction.
(D) Results of western blot analysis of FMRpolyG protein derived from wild type (WT) or mutated (4G<A) construct normalized to Vinculin upon RPS26 silencing in HEK293T cells. Graphs present means from *N* = 3 biologically independent samples with SDs. Unpaired Student's t test was used to calculate statistical significance: ns, non-significant.
(E) Comparison of overrepresented hexamers in 5'UTR full-length sequences within two transcript groups: negative (*N* = 218) and positive responders (*N* = 158). The frequency of each hexamer is compared to the corresponding hexamer frequency in 5'UTR sequences of background transcripts (BG, *N* = 22,160) (indicated as log2 fold change). The *P*-value associated with each hexamer signifies the degree of overrepresentation (enrichment) and is calculated from the binomial distribution. Only hexamers with *P*-values < 10^{-5} are displayed.

Supplementary Figure S4.

Silencing of TSR2 reduces FMR99xG and FMR16xG level

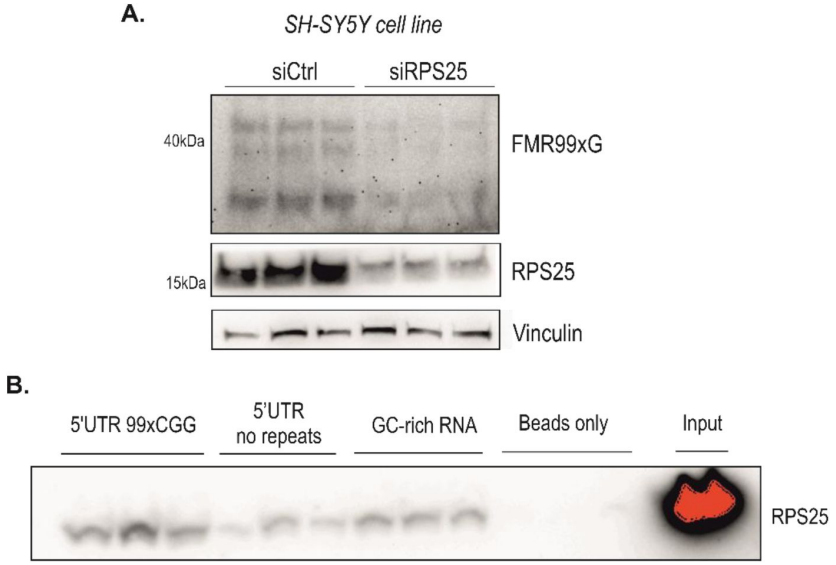
(A) Results of western blot analysis of FMR99xG normalized to Vinculin upon TSR2 silencing in COS7 cell line. The graphs present means from *N* = 3 with SDs. An unpaired Student's t-test was used to calculate statistical significance: *, *P* < 0.05.
(B) Results of western blot analysis of FMR16xG normalized to Vinculin upon TSR2 silencing in *S-16xCGG* model. Graphs present means from *N* = 5 biologically independent samples with SDs. Unpaired Student's t test was used to calculate statistical significance: ***, *P* < 0.001.



Supplementary Figure S5.

RPS25 co-precipitates with *FMR1* and affects FMR99xG level

(A) Results of western blot analysis of FMR99xG level upon RPS25 silencing in SH-SY5Y cell line.
(B) Results of western blot demonstrating enrichment of RPS25 detected in samples obtained with biotinylated 5'UTR 99xCGG RNA and to a lesser extend in 5'UTR (no repeats) & GC-rich RNA eluates (ca. 20% of eluate fraction was loaded on the gel). Input, 5% of total protein lysate was loaded. All RNA molecules in the experiment were biotinylated in order to perform RNA-protein pull down procedure. Beads only was a sample (with no RNA) which served as a negative control



Acknowledgements

We thank A. Bhattacharyya and P. Hagerman for FXTAS and control fibroblasts. The 5'UTR CGG 99x FMR1-EGFP construct was a gift from Nicolas Charlet-Berguerand (Addgene plasmid # 63091). We also thank Dominik Cysewski for participating in the preparation and analysis of MS samples from the MS2-based protein screening, Marc H. Baumann and Rabah Soliymani from Meilahti Clinical Proteomics Core Facility for SILAC-MS samples analysis, Dorota Raczyńska for the kind gift of antibodies (anti-FUS and anti-RPS6), Wojciech Kwiatkowski for the assistance with the microscopic analyses, and Roman Szczęsny for the pKK-RNAi-nucCHERRYmiR-TEV-EGFP genetic construct.

Additional information

Funding

This work was supported by the Polish National Science Center [2019/35/D/NZ2/02158 to A.B., 2020/38/A/NZ3/00498 to K.S.] and the European Union's Horizon 2020 Research and Innovation Program under the Marie Sklodowska-Curie grant agreement [No. 101003385 to A.B.] K.T. holds the Adam Mickiewicz University Foundation scholarship, awarded for the academic year 2023/24.

Author contributions

K.T., A.B., and K.S. conceptualized the study. K.T., I.B., A.Z., D.N., T.S., and A.B. performed the experiments and/or analyzed the data. A.B. performed the in vitro assay, the FMR99xG aggregation assay post RPS26 silencing, helped with the mass-spectrometry sample preparation, and performed the bioinformatics analysis of the MS-based protein screening and the SILAC-MS analysis. I.B. generated and characterized the *S-95xCGG* and *S-16xCGG* models. D.N. prepared nanoluciferase tagged and the mutated (−4G<A) *FMR1* constructs. A.Z. performed the bioinformatics analyses. T.S. performed some microscopic analyses. K.T. generated the *L-99xCGG* model, performed and analyzed all other experiments, and prepared the figures. K.T., A.B., and K.S. wrote the original draft, and the other authors reviewed the manuscript.

References

Abu Diab M, Mor-Shaked H, Cohen E, Cohen-Hadad Y, Ram O, Epsztejn-Litman S, Eiges R 2018) **The G-rich Repeats in FMR1 and C9orf72 Loci Are Hotspots for Local Unpairing of DNA** *Genetics* **210**:1239–1252

Anger AM, Armache JP, Berninghausen O, Habeck M, Subklewe M, Wilson DN, Beckmann R 2013) **Structures of the human and Drosophila 80S ribosome** *Nature* **497**:80–85

Ariza J, Rogers H, Monterrubio A, Reyes-Miranda A, Hagerman PJ, Martínez-Cerdeño V 2016) **A Majority of FXTAS Cases Present with Intranuclear Inclusions Within Purkinje Cells** *Cerebellum* **15**:546–551

Ash PEA, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, Van Blitterswijk MM, Jansen-West K, Paul JW, Rademakers R, et al 2013) **Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS** *Neuron* **77**:639–646

Bañez-Coronel M, Ayhan F, Tarabochia AD, Zu T, Perez BA, Tusi SK, Pletnikova O, Borchelt DR, Ross CA, Margolis RL, et al 2015) **RAN Translation in Huntington Disease** *Neuron* **88**:667–677

Baud A, Derbis M, Tutak K, Sobczak K 2022) **Partners in crime: Proteins implicated in RNA repeat expansion diseases** *Wiley Interdiscip Rev RNA* **13** https://doi.org/10.1002/wrna.1709

Buijsen RAM, Visser JA, Kramer P, Severijnen EAWFM, Gearing M, Charlet-Berguerand N, Sherman SL, Berman RF, Willemsen R, Hukema RK 2016) **Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a role in fragile X-associated primary ovarian insufficiency** *Hum Reprod* **31**:158–168

Cheng Z, Mugler CF, Keskin A, Hodapp S, Chan LYL, Weis K, Mertins P, Regev A, Jovanovic M, Brar GA 2019) **Small and Large Ribosomal Subunit Deficiencies Lead to Distinct Gene Expression Signatures that Reflect Cellular Growth Rate** *Mol Cell* **73**:36–47

Cid-Samper F, Gelabert-Baldrich M, Lang B, Lorenzo-Gotor N, Ponti RD, Severijnen LAWFM, Bolognesi B, Gelpi E, Hukema RK, Botta-Orfila T, et al 2018) **An Integrative Study of Protein-RNA Condensates Identifies Scaffolding RNAs and Reveals Players in Fragile X-Associated Tremor/Ataxia Syndrome** *Cell Rep* **25**:3422–3434

Cox J, Mann M 2008) **MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification** *Nat Biotechnol* **26**:1367–1372

Cui D, Li L, Lou H, Sun H, Ngai SM, Shao G, Tang J 2013) **The ribosomal protein S26 regulates p53 activity in response to DNA damage** *Oncogene* **33**:2225–2235

Derbis M, Konieczny P, Walczak A, Sekrecki M, Sobczak K 2018) **Quantitative evaluation of toxic polyglycine biosynthesis and aggregation in cell models expressing expanded CGG repeats** *Front Genet* **9**

Derbis M, Kul E, Niewiadomska D, Sekrecki M, Piasecka A, Taylor K, Hukema RK, Stork O, Sobczak K 2021) **Short antisense oligonucleotides alleviate the pleiotropic toxicity of RNA harboring expanded CGG repeats** *Nat Commun* **12**:1–17

Deutsch EW, Bandeira N, Perez-Riverol Y, Sharma V, Carver JJ, Mendoza L, Kundu DJ, Wang S, Bandla C, Kamatchinathan S, et al 2023) **The ProteomeXchange consortium at 10 years: 2023 update** *Nucleic Acids Res* **51**:D1539–D1548

Doherty L, Sheen MR, Vlachos A, Choesmel V, O'Donohue MF, Clinton C, Schneider HE, Sieff CA, Newburger PE, Ball SE, et al 2010) **Ribosomal protein genes RPS10 and RPS26 are commonly mutated in Diamond-Blackfan anemia** *Am J Hum Genet* **86**:222–228

Ferretti MB, Barre JL, Karbstein K 2018) **Translational Reprogramming Provides a Blueprint for Cellular Adaptation** *Cell Chem Biol* **25**:1372–1379

Ferretti MB, Ghalei H, Ward EA, Potts EL, Karbstein K 2017) **Rps26 directs mRNA-specific translation by recognition of Kozak sequence elements** *Nat Struct Mol Biol* **24**:700–707

Gaikwad S, Ghobakhlou F, Young DJ, Visweswaraiah J, Zhang H, Hinnebusch AG 2021) **Reprogramming of translation in yeast cells impaired for ribosome recycling favors short, efficiently translated mRNAs** *eLife* :10

Garcia-Arocena D, Yang JE, Brouwer JR, Tassone F, Iwahashi C, Berry-Kravis EM, Goetz CG, Sumis AM, Zhou L, Nguyen D V., et al 2010) **Fibroblast phenotype in male carriers of FMR1 premutation alleles** *Hum Mol Genet* **19**:299–312

Genuth NR, Barna M 2018) **The discovery of ribosome heterogeneity and its implications for gene regulation and organismal life** *Mol Cell* **71**:364

Glineburg MR, Todd PK, Charlet-Berguerand N, Sellier C 2018) **Repeat-associated non-AUG (RAN) translation and other molecular mechanisms in Fragile X Tremor Ataxia Syndrome** *Brain Res* **1693**:43–54 https://doi.org/10.1016/j.brainres.2018.02.006

Greco CM, Berman RF, Martin RM, Tassone F, Schwartz PH, Chang A, Trapp BD, Iwahashi C, Brunberg J, Grigsby J, et al 2006) **Neuropathology of fragile X-associated tremor/ataxia syndrome (FXTAS)** *Brain* **129**:243–255

Greco CM, Hagerman RJ, Tassone F, Chudley AE, Del Bigio MR, Jacquemont S, Leehey M, Hagerman PJ 2002) **Neuronal intranuclear inclusions in a new cerebellar tremor/ataxia syndrome among fragile X carriers** *Brain* **125**:1760–1771

Green KM, Glineburg MR, Kearse MG, Flores BN, Linsalata AE, Fedak SJ, Goldstrohm AC, Barmada SJ, Todd PK 2017) **RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response** *Nat Commun* :8

Green KM, Linsalata AE, Todd PK 2016) **RAN translation—What makes it run?** *Brain Res* **1647**:30–42 https://doi.org/10.1016/j.brainres.2016.04.003

Gregory B, Rahman N, Bommakanti A, Shamsuzzaman M, Thapa M, Lescure A, Zengel JM, Lindahl L 2019) **The small and large ribosomal subunits depend on each other for stability and accumulation** *Life Sci alliance* :2

Hagerman RJ, Hagerman P 2016) **Fragile X-associated tremor/ataxia syndrome-features, mechanisms and management** *Nat Rev Neurol* **12**:403–412 https://doi.org/10.1038/nrneurol

.2016.82

Hagerman RJ, Protic D, Rajaratnam A, Salcedo-Arellano MJ, Aydin EY, Schneider A 2018) **Fragile X-Associated Neuropsychiatric Disorders (FXAND)** *Front Psychiatry* **9**:564

Havkin-Solomon T, Fraticelli D, Bahat A, Hayat D, Reuven N, Shaul Y, Dikstein R 2023) **Translation regulation of specific mRNAs by RPS26 C-terminal RNA-binding tail integrates energy metabolism and AMPK-mTOR signaling** *Nucleic Acids Res* **51**:4415–4428

Hussain T, Llácer JL, Fernández IS, Munoz A, Martin-Marcos P, Savva CG, Lorsch JR, Hinnebusch AG, Ramakrishnan V 2014) **Structural Changes Enable Start Codon Recognition by the Eukaryotic Translation Initiation Complex** *Cell* **159**:597

Jacquemont S, Hagerman RJ, Leehey MA, Hall DA, Levine RA, Brunberg JA, Zhang L, Jardini T, Gane LW, Harris SW, et al 2004) **Penetrance of the Fragile X-Associated Tremor/Ataxia Syndrome in a Premutation Carrier Population** *JAMA* **291**:460–469

Jin P, Duan R, Qurashi A, Qin Y, Tian D, Rosser TC, Liu H, Feng Y, Warren ST 2007) **Pur α Binds to rCGG Repeats and Modulates Repeat-Mediated Neurodegeneration in a Drosophila Model of Fragile X Tremor/Ataxia Syndrome** *Neuron* **55**:556–564

Kearse MG, Green KM, Krans A, Rodriguez CM, Linsalata AE, Goldstrohm AC, Todd PK 2016) **CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins** *Mol Cell* **62**:314–322

Klickstein JA, Mukkavalli S, Raman M 2020) **AggreCount: an unbiased image analysis tool for identifying and quantifying cellular aggregates in a spatially defined manner** *J Biol Chem* **295**:17672–17683

Laakkonen EK, Soliymani R, Karvinen S, Kaprio J, Kujala UM, Baumann M, Sipilä S, Kovanen V, Lalowski M 2017) **Estrogenic regulation of skeletal muscle proteome: a study of premenopausal women and postmenopausal MZ cotwins discordant for hormonal therapy** *Aging Cell* **16**:1276–1287

Li D, Yang J, Huang X, Zhou H, Wang J 2022) **eIF4A2 targets developmental potency and histone H3.3 transcripts for translational control of stem cell pluripotency** *Sci Adv* **8**:478

Li J, Su Y, Chen L, Lin Y, Ru K 2023) **Identification of novel mutations in patients with Diamond-Blackfan anemia and literature review of RPS10 and RPS26 mutations** *Int J Lab Hematol* **45**:766–773

Linsalata AE, He F, Malik AM, Glineburg MR, Green KM, Natla S, Flores BN, Krans A, Archbold HC, Fedak SJ, et al 2019) **DDX 3X and specific initiation factors modulate FMR 1 repeat-associated non-AUG-initiated translation** *EMBO Rep* :20

Liu XM, Yan MQ, Ji SY, Sha QQ, Huang T, Zhao H, Bin Liu H, Fan HY, Chen ZJ 2018) **Loss of oocyte Rps26 in mice arrests oocyte growth and causes premature ovarian failure** *Cell Death Dis* **9**:1–15

Livak KJ, Schmittgen TD 2001) **Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method** *Methods* **25**:402–408

Loomis EW, Sanz LA, Chédin F, Hagerman PJ 2014) **Transcription-Associated R-Loop Formation across the Human FMR1 CGG-Repeat Region** *PLOS Genet* **10**:e1004294

Luan Y, Tang N, Yang J, Liu S, Cheng C, Wang Y, Chen C, Guo YN, Wang H, Zhao W, et al 2022) **Deficiency of ribosomal proteins reshapes the transcriptional and translational landscape in human cells** *Nucleic Acids Res* **50**:6601–6617

Ma L, Herren AW, Espinal G, Randol J, McLaughlin B, Martinez-Cerdeño V, Pessah IN, Hagerman RJ, Hagerman PJ 2019) **Composition of the Intranuclear Inclusions of Fragile X-associated Tremor/Ataxia Syndrome** *Acta Neuropathol Commun* **7**:1–26

Malik I, Kelley CP, Wang ET, Todd PK 2021a) **Molecular mechanisms underlying nucleotide repeat expansion disorders** *Nat Rev Mol Cell Biol* **22**:589–607 https://doi.org/10.1038/s41580-021-00382-6

Malik I, Tseng Y, Wright SE, Zheng K, Ramaiyer P, Green KM, Todd PK 2021b) **SRSF protein kinase 1 modulates RAN translation and suppresses CGG repeat toxicity** *EMBO Mol Med* **13**

Mori K, Weng S-M, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Broeckhoven C Van, et al 2013) **The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS** *Science (80-)* **339**:1335–1338

Perez-Riverol Y, Bai J, Bandla C, García-Seisdedos D, Hewapathirana S, Kamatchinathan S, Kundu DJ, Prakash A, Frericks-Zipper A, Eisenacher M, et al 2022) **The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences** *Nucleic Acids Res* **50**:D543–D552

Piantanida N, La Vecchia M, Sculco M, Talmon M, Palattella G, Kurita R, Nakamura Y, Ronchi AE, Dianzani I, Ellis SR, et al 2022) **Deficiency of ribosomal protein S26, which is mutated in a subset of patients with Diamond Blackfan anemia, impairs erythroid differentiation** *Front Genet* **13**:3573

Pisarev A V., Kolupaeva VG, Yusupov MM, Hellen CUT, Pestova T V 2008) **Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes** *EMBO J* **27**:1609–1621

Plassart L, Shayan R, Montellese C, Rinaldi D, Larburu N, Pichereaux C, Froment C, Lebaron S, O'donohue MF, Kutay U, et al 2021) **The final step of 40s ribosomal subunit maturation is controlled by a dual key lock** *eLife* **10**

Raczynska KD, Ruepp MD, Brzek A, Reber S, Romeo V, Rindlisbacher B, Heller M, Szweykowska-Kulinska Z, Jarmolowski A, Schümperli D 2015) **FUS/TLS contributes to replication-dependent histone gene expression by interaction with U7 snRNPs and histone-specific transcription factors** *Nucleic Acids Res* **43**:9711–9728

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J 2019) **g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update)** *Nucleic Acids Res* **47**:W191–W198

Rosario R, Stewart HL, Choudhury NR, Michlewski G, Charlet-Berguerand N, Anderson RA 2022) **Evidence for a fragile X messenger ribonucleoprotein 1 (FMR1) mRNA gain-of-function toxicity mechanism contributing to the pathogenesis of fragile X-associated premature ovarian insufficiency** *FASEB J* **36**

Rovozzo R, Korza G, Baker MW, Li M, Bhattacharyya A, Barbarese E, Carson JH 2016) **CGG repeats in the 5'UTR of FMR1 RNA regulate translation of other RNAs localized in the same RNA granules** *PLoS One* **11**:1–10

Schmidt EK, Clavarino G, Ceppi M, Pierre P 2009) **SUnSET, a nonradioactive method to monitor protein synthesis** *Nat Methods* **6**:275–277

Schütz S, Fischer U, Altvater M, Nerurkar P, Peña C, Gerber M, Chang Y, Caesar S, Schubert OT, Schlenstedt G, et al 2014) **A RanGTP-independent mechanism allows ribosomal protein nuclear import for ribosome assembly** *eLife* **3**:e03473

Schütz S, Michel E, Damberger FF, Oplová M, Peña C, Leitner A, Aebersold R, Allain FHT, Panse VG 2018) **Molecular basis for disassembly of an importin:ribosomal protein complex by the escortin Tsr2** *Nat Commun* **9**

Sellier C, Buijsen RAM, He F, Natla S, Jung L, Tropel P, Gaucherot A, Jacobs H, Meziane H, Vincent A, et al 2017) **Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome** *Neuron* **93**:331–347

Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A, et al 2013) **Sequestration of DROSHA and DGCR8 by expanded CGG RNA Repeats Alters microRNA processing in fragile X-associated tremor/ataxia syndrome** *Cell Rep* **3**:869–880

Sellier C, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S, Willemsen R, Elliott DJ, et al 2010) **Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients** *EMBO J* **29**:1248–1261

Shelly KE, Candelaria NR, Li Z, Allen EG, Jin P, Nelson DL 2021) **Ectopic expression of CGG-repeats alters ovarian response to gonadotropins and leads to infertility in a murine FMR1 premutation model** *Hum Mol Genet* **30**:923–938

Shi Z, Barna M 2015) **Translating the genome in time and space: specialized ribosomes, RNA regulons, and RNA-binding proteins** *Annu Rev Cell Dev Biol* **31**:31–54

Shi Z, Fujii K, Kovary KM, Genuth NR, Rö St HL, Teruel MN, Correspondence MB 2017) **Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide** *Mol Cell* **67**:83

Szczesny RJ, Kowalska K, Klosowska-Kosicka K, Chlebowski A, Owczarek EP, Warkocki Z, Kulinski TM, Adamska D, Affek K, Jedroszkowiak A, et al 2018) **Versatile approach for functional analysis of human proteins and efficient stable cell line generation using FLP-mediated recombination system** *PLoS One* **13**:e0194887

Tassone F, Iong KP, Tong TH, Lo J, Gane LW, Berry-Kravis E, Nguyen D, Mu LY, Laffin J, Bailey DB, et al 2012) **FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States** *Genome Med* **4**

Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou LP, Mi H 2022) **PANTHER: Making genome-scale phylogenetics accessible to all** *Protein Sci* **31**:8–22

Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen K chun, Scaglione KM, Basrur V, et al 2013) **CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome** *Neuron* **78**:440–455

Tsai BP, Wang X, Huang L, Waterman ML 2011) **Quantitative profiling of in vivo-assembled RNA-protein complexes using a novel integrated proteomic approach** *Mol Cell Proteomics* **10**

Tseng Y-J, Krans A, Malik I, Deng X, Yildirim E, Ovunc S, Tank EMH, Jansen-West K, Kaufhold R, Gomez NB, et al 2024) **Ribosomal quality control factors inhibit repeat-associated non-AUG translation from GC-rich repeats** *Nucleic Acids Res* :gkae137

Tseng Y-J, Sandwith SN, Green KM, Chambers AE, Krans A, Raimer HM, Sharlow ME, Reisinger MA, Richardson AE, Routh ED, et al 2021) **The RNA helicase DHX36/G4R1 modulates C9orf72 GGGGCC hexanucleotide repeat-associated translation** *J Biol Chem* :100914

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J 2016) **The Perseus computational platform for comprehensive analysis of (prote)omics data** *Nat Methods* **13**:731–740

Wiśniewski JR, Zougman A, Nagaraj N, Mann M 2009) **Universal sample preparation method for proteome analysis** *Nat Methods* **6**:359–362

Wright SE, Rodriguez CM, Monroe J, Xing J, Krans A, Flores BN, Barsur V, Ivanova MI, Koutmou KS, Barmada SJ, et al 2022) **CGG repeats trigger translational frameshifts that generate aggregation-prone chimeric proteins** *Nucleic Acids Res* **50**:8674–8689

Yamada SB, Gendron TF, Niccoli T, Genuth NR, Grosely R, Shi Y, Glaria I, Kramer NJ, Nakayama L, Fang S, et al 2019) **RPS25 is required for efficient RAN translation of C9orf72 and other neurodegenerative disease-associated nucleotide repeats** *Nat Neurosci* **22**:1383–1388

Yang N, Chanda S, Marro S, Ng YH, Janas JA, Haag D, Ang CE, Tang Y, Flores Q, Mall M, et al 2017) **Generation of pure GABAergic neurons by transcription factor programming** *Nat Methods* **14**:621–628

Yang YM, Jung Y, Abegg D, Adibekian A, Carroll KS, Karbstein K 2023) **Chaperone-directed ribosome repair after oxidative damage** *Mol Cell* **83**:1527–1537

Yang YM, Karbstein K 2022) **The chaperone Tsr2 regulates Rps26 release and reincorporation from mature ribosomes to enable a reversible, ribosome-mediated response to stress** *Sci Adv* **8**

Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MAC, et al 2011) **Non-ATG-initiated translation directed by microsatellite expansions** *Proc Natl Acad Sci U S A* **108**:260–265

Zu T, Guo S, Bardhi O, Ryskamp DA, Li J, Tusi SK, Engelbrecht A, Klippel K, Chakrabarty P, Nguyen L, et al 2020) **Metformin inhibits RAN translation through PKR pathway and mitigates disease in C9orf72 ALS/FTD mice** *Proc Natl Acad Sci U S A* **117**:18591–18599

Author information

Katarzyna Tutak
Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
ORCID iD: 0000-0002-2801-9956

Izabela Broniarek
Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
ORCID iD: 0000-0001-5318-8228

Andrzej Zielezinski
Department of Computational Biology, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
ORCID iD: 0000-0002-8096-3776

Daria Niewiadomska
Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
ORCID iD: 0000-0001-7100-789X

Tomasz Skrzypczak
Center of Advanced Technology, Adam Mickiewicz University, Poznan, Poland

Anna Baud
Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
ORCID iD: 0000-0003-3710-5722

For correspondence: anna.baud@amu.edu.pl

Krzysztof Sobczak
Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University, Poznan, Poland
ORCID iD: 0000-0001-8352-9812

For correspondence: ksobczak@amu.edu.pl

Editors

Reviewing Editor
Ivan Topisirovic
McGill University, Montreal, Canada

Senior Editor
David Ron
University of Cambridge, Cambridge, United Kingdom

Reviewer #2 (Public review):

Summary:

Translation of CGG repeats leads to accumulation of poly G, which is associated with neurological disorders. This is an important paper in which the authors sought out proteins that modulate RAN translation. They determined which proteins in Hela cells were enriched on CGG repeats and affected levels of polyG encoded in the 5'UTR of the FMR1 mRNA. They then showed that siRNA depletion of ribosomal protein RPS26 results in less production of FMR1polyG than in control. Experiments were performed in several cell lines and with

several reporters with differences in repeats and transfection methods to increase confidence that changes were occurring. New data and details of the methods increase confidence that reporter translation but not global translation is diminished by RPS26 knockdown as concluded. The manuscript has been improved by data showing that new proteins are being synthesized in cells following RPS26 knockdown, and that near-cognate start codon usage is diminished in lines when RPS26 is knocked down, but the mechanism by which RPS26 depletion affects translation is still unclear.

Strengths:

- The authors have proteomics data that show enrichment of a set of proteins on FMR1-polyG RNA but not a related RNA.
- Knockdown of RPS26, which was enriched on the FMR1 RNA, led to decreases in cell growth, but surprisingly did not strongly affect global translation, as assessed by puromycin incorporation
- There is some new evidence that near-cognate start codon selection is affected by RPS26 knockdown

Weaknesses:

- The mechanism for RPS26 knockdown affecting translation of the polyG sequences is unclear, whether knockdown is affecting ribosome levels, extra ribosomal RPS26 or ribosome composition is not known.

https://doi.org/10.7554/eLife.98631.2.sa2

Reviewer #3 (Public review):

Tutak et al provide intriguing findings demonstrating that insufficiency of RPS26 and related proteins, such as TSR2 and RPS25, downregulates RAN translation from CGG repeat RNA in fragile X-associated conditions. Using RNA-tagging system and mass spectrometry-based screening, the authors identified RPS26 as a potential regulator of RAN translation. They further confirmed its regulatory effects on RAN translation by siRNA-based knockdown experiments in multiple cellular disease models. Quantitative mass spectrometry analysis revealed that the expression of some ribosomal proteins is sensitive to RPS26 depletion, while approximately 80% of proteins, including FMRP, were not influenced. Given the limited understanding of the roles of ribosomal proteins in RAN translation regulation, this study provides novel insights into this research field. However, certain data do not fully support the authors' critical conclusions.

(1) While the authors substituted the ACG near-cognate initiation codon with other near-cognate codons, such as GTG and CTG, in the luciferase assay (Figure 4F), substitution of the ACG codon with an ATG codon should also be performed. Although they evaluated RPS26 knockdown effect on AUG-dependent FMRP translation in Figure 3C, investigating its effect on AUG-dependent repeat-associated translation (e.g., AUG-CGG-repeat) is necessary to substantiate their claim that ACG codon selection is important for RAN translation downregulation by RPS26 knockdown.

(2) The results of the ASO-based ACG codon-blocking experiment in Figure 4G are difficult to interpret. While RPS knockdown reduces FMRpolyG expression, the effect appears attenuated by the ASO-ACG treatment compared to the control. However, this does not conclusively demonstrate that the regulatory effect is directly due to ACG codon selection during translation initiation for some reasons. For example, ASO-ACG treatment possibly interferes with ribosomal scanning rather than ACG-codon selection, or alters the expression of template CGG repeat RNA. To validate the effect of RPS26 knockdown on ACG codon selection,

experiments using the ACG-to-ATG substituted CGG repeat reporter are recommended, as suggested in comment 1.

(3) The regulatory effects of RPS26 and other molecules on RAN translation have been investigated as effects on the expression levels of FMRpolyG proteins upon knockdown of these molecules in disease model cells expressing CGG repeat sequences (Figures 1C, 1D, 3B, 3C, 3E, 4F, 4G, 5A, 5C, 6A, 6D）. However, FMRpolyG expression levels can be influenced by factors other than RAN translation in these cellular experiments, such as template RNA level, template RNA localization, and FMRpolyG protein degradation. Although the authors evaluated the effect on the expression levels of template CGG repeat RNA, it would be better to confirm the direct effect of these regulators on RAN translation by other experiments. In vitro translation assay that can directly evaluate RAN translation is preferable, but experiments using the ACG-to-ATG substituted CGG repeat reporter, as suggested in comment 1, would also provide valuable insights.

https://doi.org/10.7554/eLife.98631.2.sa1

Author response:

The following is the authors' response to the current reviews.

We thank Reviewers for highlighting the strengths of our work along with suggestions for future directions.

We agree with the Reviewers that RPS26 depletion may impact not only RAN translation initiation and codon selection (as showed in the experiments in Figure 4G), but also other mechanisms, such as speed of PIC scanning, as we stated in the discussion. Although, we did provide the data showing that mRNA of exogenous *FMR1-GFP* does not change upon RPS26 depletion (Figure 3B&C), hence observed effect most likely stems from translation regulation. In addition, an experiment with ASO-ACG treatment (Figure 4G) suggests that near cognate start codon selection or speed of PIC scanning may be a part of the regulation of RAN translation sensitive to RPS26 depletion. In addition, our latest unpublished results (Niewiadomska D. et al., in revision), indicate that FMRpolyG in fusion with GFP is fairly stable, in particular, while derived from long repeats (>90xCGG), suggesting that the protein stability is not at play in RPS26-dependent regulation.

We would like to stress that in order to avoid bias in result interpretation and to mimic the natural situation, the majority of experiments concerning levels of FMRpolyG were performed in cell models with stable expression of ACG-initiated FMRpolyG. Currently, we do not possess a cell model with stable expression of AUG-initiated FMRpolyG, and the experiments based on transient transfection system would not necessarily be comparable to the results obtained in stable expression system. However, we believe that the experiment presented in Figure 2B serves as a good control for overall translation level upon RPS26 depletion indicating that RPS26 insufficiency does not affect global translation and the observed regulation is specific to some mRNAs including the one encoding FMRpolyG frame. We also show that the level of ca. 80% of identified canonical proteins, including FMRP, did not change upon RPS26 silencing (SILAC-MS, Figure 4A). Indeed, we did not explore the ribosome composition upon RPS26 and TSR2 depletion, although, most likely the pool of functional ribosomes in the cell is sufficient enough to support the basal translation level (SUnSET assays, Figure 2B & 5C). However, we cannot exclude possibility that for some mRNAs, including one encoding for FMRpolyG, the observed effect can be partially caused by lowering the number of fully active ribosomes, especially in experiments with transient transfection experiments where transgene expression is hundreds times higher than for average native mRNA.

Finally, we agree with the Reviewer that in vitro translation assay would provide the evidence of direct effect of RPS26 on FMRpolyG level, however, we did not manage to overcome technical difficulties in obtaining cellular lysate devoid of RPS26 from vendor companies.

The following is the authors' response to the original reviews.

General Comments

We thank Reviewers for the critical comments and experimental suggestions. We considered most of the advices in the revised version of the manuscript, which allowed for a more balanced interpretation of the results presented, and further supported major statement of the manuscript that insufficiency of the RPS26 and RPS25 plays a role in modulating the efficiency of noncanonical RAN translation from *FMR1* mRNA, which results in the production of toxic polyglycine protein (FMRpolyG). Firstly, performing new experiments, we showed that silencing of the RPS26 and its chaperone protein TSR2, which regulates loading/exchange of RPS26 in maturing small ribosome subunit, did not elicit global translation inhibition. Secondly, we demonstrated that in contrary to RPS26 and RPS25 depletion, silencing the RPS6 protein, a core component of 40S subunit, did not affect FMRpolyG production, further supporting the specific effect of RPS26 and RPS25 on RAN translation regulation of mutant FMR1 mRNA. We also observed that depletion of RPS26, RPS25 and RPS6 had significant negative effect on cells proliferation which is in line with previously published results indicating that insufficiencies of ribosomal proteins negatively affect cell growth. Moreover, we showed that FMRpolyG production is significantly affected by RPS26 depletion while initiated at ACG, but not other near cognate start codons. Importantly, translation of FMRP initiated at canonical AUG codon of the same mRNA upstream the CGGexp was not affected by RPS26 silencing, similarly to vast majority of the human proteome. This implies that RAN translation of *FMR1* mRNA mediated by RPS26 insufficiency is likely to be dependent on start codon selection/fidelity. In essence, we provide a series of evidences indicating that cellular amount of 40S ribosomal proteins RPS26 and RPS25 is important factor of CGGrelated RAN translation regulation. Finally, we also decided to tone down our claims. Now, we state that the RPS26/25/TSR2 insufficiency or depletion, affects RAN translation, rather than composition of 40S ribosomal subunit *per se* influences RAN translation. We have addressed all specific concerns below and made changes to the new version of manuscript.

> ***Public Reviews:***
>
> ***Reviewer #1 (Public Review):***
>
> *Summary:*
>
> *In this manuscript, Tutak et al use a combination of pulldowns, analyzed by mass spectrometry, reporter assays, and fluorescence experiments to decipher the mechanism of protein translation in fragile X-related diseases. The topic is interesting and important.*
>
> *Although a role for Rps26-deficient ribosomes in toxic protein translation is plausible based on already available data, the authors' data are not carefully controlled and thus do not support the conclusions of the paper.*

We sincerely appreciate your rigorous, insightful, and constructive feedback throughout the revision process. We believe your guidance has been instrumental in significantly enhancing the quality of our research. Below, we have addressed your comments pointby-point.

> *Strengths:*

> *The topic is interesting and important.*
>
> *Weaknesses:*
>
> *In particular, there is very little data to support the notion that Rps26-deficient ribosomes are even produced under the circumstances. And no data that indicate that they are involved in the RAN translation. Essential controls (for ribosome numbers) are lacking, no information is presented on the viability of the cells (Rps26 is an essential protein), and the differences in protein levels could well arise from block in protein synthesis, and cell division coupled to differential stability of the proteins.*

We agree that data presented in the first version of the manuscript did not directly address the following processes: ribosome content, global translation rate and cell viability upon RPS26 depletion. Therefore we addressed some of the issues in the revised version of the manuscript. In particular, we showed that RPS26 and TSR2 knock down did not inhibit global translation (new Figure 2B & 4C), hence we concluded that the changes of FMRpolyG level did not arise from general translational shut down. On the other hand, RPS26, RPS25 and RPS6 depletion negatively affected cells proliferation (new Figure 2A,5D,6C), which is in line with a number of previously published researches (e.g. Cheng *et al*, 2019; Havkin-Solomon *et al*, 2023). However, the rate of proliferation abnormalities is limited. We agree that observed effects on RAN translation from mutant FMR1 mRNA may stem from the combination of altered protein synthesis, conditions of the cells but also cis-acting factors of mRNA sequence/structure. In new experiments we showed that single nucleotide substitution of ACG by other near cognate start codons change sensitivity of RAN translation to insufficiency of RPS26 (new Figure 4F). Also the inhibitory effect of antisense oligonucleotide binding to the region of 5'UTR containing ACG initiation codon (ASO_ACG) is different in cells differing in amount of RPS26 (new Figure 4G).

We also agree that our data only partially supports the role of RPS26-defficient ribosomes in RAN translation. Therefore, we have toned down our claims. Now, we state that the RPS26/25/TSR2 insufficiency or depletion affects RAN translation. We also changed the title of the manuscript to: "Insufficiency of 40S ribosomal proteins, RPS26 and RPS25, negatively affects biosynthesis of polyglycine-containing proteins in fragile-X associated conditions" (Previously it was: "Ribosomal composition affects the noncanonical translation and toxicity of polyglycine-containing proteins in fragile X-associated conditions".

> *Specific points:*
>
> *(1) Analysis of the mass spec data in Supplemental Table S3 indicates that for many of the proteins that are differentially enriched in one sample, a single peptide is identified. So the difference is between 1 peptide and 0. I don't understand how one can do a statistical analysis on that, or how it would give out anything of significance. I certainly do not think it is significant. This is exacerbated by the fact that the contaminants in the assay (keratins) are many, many-fold more abundant, and so are proteins that are known to be mitochondrial or nuclear, and therefore likely not actual targets (e.g. MCCC1, PC, NPM1; this includes many proteins "of significance" in Table S1, including Rrp1B, NAF1, Top1, TCEPB, DHX16, etc...).*
>
> *The data in Table S6/Figure 3A suffer from the same problem.*
>
> *I am not convinced that the mass spec data is reliable.*

We thank Reviewer for the comment concerning MS data; however, we believe that it may stem from misunderstanding of the data presented in Table S3 and S6. Both tables represent the output from MaxQuant analysis (so-called ProteinGroup) of MS .raw files, without any

filtering. As stated in the Material&Methods, we applied default parameters suggested by MaxQuant developers to analyze MS data, these include identification of proteins based on at least 1 unique peptide, and thus some of the proteins with only 1 unique peptide are shown in Tables S1 and S3. Reviewer is also right that in this output table common contaminants, such as keratins are included. However, these identifications are denoted as "CON_", and are further filtered out during statistical analysis in Perseus software. During the statistical analysis we first filtered out irrelevant protein groups identifications, such as contaminants, or only identified by site modifications.

We have changed the names of Supplementary Table files, giving more detailed description. We hope this will help to avoid misunderstanding for broader public. Secondly, when comparing the data presented in Table S3 and volcano plot presented in Figure 1B, one can notice that indeed the majority of identified proteins are not statistically significant (grey points), thus not selected for further stratification. Lack of significance of these proteins may be partially due to poor MS identification, however, they are not included in the following parts of the manuscript. Further, we selected only eight proteins (out of over 150) for stratification by orthogonal techniques, thus we argue that this step validates the biological relevance of chosen candidate RAN-translation modifiers. One should also keep in mind that pull down samples analyzed by MS often yield lower intensity and identification rates, when comparing to whole cell analysis, as a result of lower protein input or stringent washes used during sample preparation.

Regarding the data presented in Table S6 (SILAC data), we argue that these data are of very good quality. More than 2,000 proteins were identified in a 125min gradient, with over 80% of proteins that were identified with at least 2 unique peptides. Each of three biological replicates was analyzed three times (technical replicates), giving total of 9 high resolution MS runs. Together, we strongly believe that this data is of high confidence.

> *(2) The mass-spec data however claims to identify Rps26 as a factor binding the toxic RNA specifically. The rest of the paper seeks to develop a story of how Rps26-deficient ribosomes play a role in the translation of this RNA. I do not consider that this makes sense.*

Indeed, we identified RPS26 as a protein that co-precipitated with *FMR1* containing expanded CGG repeats (Supplementary Figure 1G) and found that depletion of RPS26 hindered RAN translation of FMRpolyG, suggesting that RPS26 positively affects RAN translation. However, we did not state that RPS26 directly interacts with toxic RNA. In order to confirm the specificity of RAN translation regulation by RPS26 insufficiency, we tested whether depletion of other 40S ribosomal protein, RPS6, affects FMRpolyG synthesis. Our experiments showed that there was no any significant effect on RAN translation efficiency post RPS6 silencing (new Figure 5C). Importantly, we showed that RPS26 depletion did not inhibit global translation (new Figure 2B). In addition, mutagenesis of near-cognate start codon (new Figure 4F) and ASO_ACG treatment (new Figure 4G) provided the evidences that modulation of FMRpolyG biosynthesis by RPS26 level may depend on start codon selection. In essence, our data suggest that RPS26 depletion specifically affects synthesis of FMRpolyG, but not FMRP derived from the same *FMR1* mRNA with CGGexp. However, we do not claim that the observed effect is the consequence of a direct interaction between RPS26 and 5'UTR of FMR1 mRNA. Downregulation of FMRpolyG biosynthesis could be an outcome of the alteration of ribosomal assembly, decrease of efficiency and fidelity of PIC scanning/initiation or impeded elongation or a combination of all these processes. In the manuscript we presented the results of experiments which tested many of these possibilities.

> *(3) Rps26 is an essential gene, I am sure the same is true for DHX15. What happens to cell viability? Protein synthesis? The yeast experiments were carefully carried out under experiments where Rps26 was reduced, not fully depleted to give small growth defects.*

We agree with the Reviewer that RPS26 and DHX15 are essential proteins, similarly to all RNA binding proteins, and caution should be taken during experimental design. To address this, we titrated different concentrations of siRPS26, and found that administration of 5 nM siRPS26, which just partially silenced RPS26, decreased FMRpolyG by around 50% (new Figure 1D). This impact was even greater with 15 nM siRPS26, as we observed around 80% decrease of FMRpolyG.

Havkin-Solomon et al. (2023), showed that proliferation rate is decreased in cells with mutated C-terminus of RPS26, which is required for contacting mRNA. In accordance with this study, we showed that cells with knocked down RPS26 proliferate less efficiently (new Figure 2A), but depletion of RPS26 did not impact the global translation (new Figure 2B). In addition, our SILAC-MS data indicates that ~80% of proteins with determined expression level were not affected by RPS26 insufficiency, and ~20% of the proteins turned out to be sensitive to RPS26 decrease. Although, these data do not take into account the protein stability.

> *(4) Knockdown efficiency for all tested genes must be shown to evaluate knockdown efficiency.*

The current version of the manuscript contains representative western blots with validation of knock-down efficiency (for example in Figure 3B, C, E, Figure 6A) and we included knock-down validations where applicable (Figures 1D, 2B, 4G and 5C).

(5) The data in Figure 1E have just one mock control, but two cell types (control si and Rps26 depletion).

Mock control corresponds to the cells treated with lipofectamine reagent and was included in the study to determine the “background” signal from cells treated with delivery agent and reagents used to measure the apoptosis process. These cells were neither expressing FMRpolyG, nor siRNAs. Luminescence signals were normalized to the values obtained from mock control. We added more details describing this assay in the Figure 1 legend.

> *(6) The authors' data indicate that the effects are not specific to Rps26 but indeed also observed upon Rps25 knockdown. This suggests strongly that the effects are from reduced ribosome content or blocked protein synthesis. Additional controls should deplete a core RP to ascertain this conclusion.*

We agree that observed effects may stem from reduced ribosome content, however, we argue that this is the only possibility and explanation. Previously, it was shown that RPS25 regulates G4C2-related RAN translation, but knock out of RPS25 does not affect global translation (Yamada S, 2019, *Nat. Neuroscience*). Similarly, we showed that KD of RPS26 or TSR2 did not reduce significantly global translation rate (SUnSET assay; new Figure 2B and 5C, respectively).

Moreover, in a new version of manuscript we included a control experiment, where we silenced core ribosomal protein (RPS6) and found that RPS6 depletion did not affect RAN translation from mutant FMR1 mRNA (new Figure 5C), thus strengthening our conclusion about specific RAN translation regulation by the level of RPS26 and RPS25.

Finally, our observation aligns well with current knowledge about how deficiency of different ribosomal proteins alters translation of some classes of mRNAs (Luan Y, 2022, *Nucleic Acids Res*; Cheng Z, 2019, *Mol Cell*). It was shown that depletion of RPS26 affects translation rate of different mRNAs compared to depletion of other proteins of small ribosomal subunit.

> *(7) Supplemental Figure S3 demonstrates that the depletion of S26 does not affect the selection of the start codon context. Any other claim must be deleted. All the 5'-UTR logos*

> *are essentially identical, indicating that "picking" happens by abundance (background).*

Supplementary Figure 3D represents results indicating that the mutation in -4 position (from G to A) did not affect the RAN translation regardless of RPS26 presence or depletion. However, this result does not imply that RPS26 does not affect the selection of start codon of sequence- or RNA structure-context. We verified this particular -4 position, as it was suggested previously as important RPS26-sensitive site in yeasts (Ferretti M, 2017, *Nat Struct Mol Biol*). We agree with Reviewer that all 5'UTR logos presented in our paper did not show statistical significance for neither tested position for human mRNAs. On the contrary, we observed that regulation sensitive to RPS26 level depends on the selection of start codon of RAN translation, in particular ACG initiation (new Figure 4F&G). RPS26 depletion affected ACG-initiated but not GTG- or CTG-initiated RAN translation.

In the previous version of the manuscript, we wrote that we did not identify any specific motifs or enrichment within analyzed transcripts in comparison to the background. On the other hand, we found that the GC-content among analyzed transcripts is higher within 5'UTRs and in close proximity to ATG in coding sequences (Figure 4D), what suggests the importance of RNA stable structures in this region. In addition, we showed that mRNAs encoding proteins responding to RPS26 depletion have shorter than average 5'UTRs (new Figure 4E).

> *(8) Mechanism is lacking entirely. There are many ways in which ribosomes could have mRNA-specific effects. The authors tried to find an effect from the Kozak sequence, unsuccessfully (however, they also did not do the experiment correctly, as they failed to recognize that the Kozak sequence differs between yeast, where it is A-rich, and mammalian cells, where it is GGCGCC). Collisions could be another mechanism.*

Indeed, collisions as well as other mechanisms such as skewed start codon fidelity may have an effect on efficiency of FMRpolyG biosynthesis. In the current version of the manuscript, we show that RPS26 amount-sensitive regulation seems to be start codonselection dependent (new Figure 4F&G).

> ***Reviewer #2 (Public Review):***
>
> *Summary:*
>
> *Translation of CGG repeats leads to the accumulation of poly G, which is associated with neurological disorders. This is a valuable paper in which the authors sought out proteins that modulate RAN translation. They determined which proteins in Hela cells bound to CGG repeats and affected levels of polyG encoded in the 5'UTR of the FMR1 mRNA. They then showed that siRNA depletion of ribosomal protein RPS26 results in less production of FMR1polyG than in control. There are data supporting the claim that RPS26 depletion modulates RAN translation in this RNA, although for some results, the Western results are not strong. The data to support increased aggregation by polyG expression upon S26 KD are incomplete.*

We thank the Reviewer for critical comments and suggestions. We sincerely appreciate your rigorous, insightful, and constructive feedback throughout the revision process.

Below each specific point, we addressed the mentioned issues.

> *Strengths:*
>
> *The authors have proteomics data that show the enrichment of a set of proteins on FMR1 RNA but not a related RNA.*

We thank Reviewer for appreciation of provided MS-screening results, which identified proteins enriched on FMR1 RNA with expanded CGG repeats.

> *Weaknesses:*
>
> *- It is insinuated that RPS26 binds the RNA to enhance CGG-containing protein expression. However, RPS26 reduction was also shown previously to affect ribosome levels, and reduced ribosome levels can result in ribosomes translating very different RNA pools.*

In previous version of the manuscript we did not state that RPS26 binds directly to RNA with expanded CGG repeats and we did not show the experiment indicating direct interaction between studied RNA and RPS26. What we showed is that RPS26 was enriched on *FMR1* RNA MS samples, however, we did not verify whether it is direct or indirect interaction. We also tried to test hypothesis that lack of RPS26 in PIC complex may affect efficiency of RAN translation initiation via specific, previously described in yeast Kozak context (Ferretti M, 2017, *Nat Struct Mol Biol*). As we described this hypothesis was negatively validated. However, we showed that other features of 5'UTR sequences (e.g. higher GC-content or shorter leader sequence) are potentially important for translation efficiency in cells with depleted RPS26.

Indeed, RPS26 is involved in 40S maturation steps (Plassart L, 2021, *eLife*) and its insufficiency or mutations or blocking its inclusion to 40S ribosome may result in incomplete 40S maturation, which subsequently might negatively affect translation per se. However, we did not observe global translation inhibition after RPS26 depletion or depletion of TSR2, the chaperon involved in incorporation/exchange RPS26 to small ribosomal subunit (new Figure 2B and 5C). In addition, our SILAC-MS data indicates that majority of studied proteins (including FMRP, the main product of *FMR1* gene) were not affected by RPS26 depletion which can be carefully extrapolated to global translation. In revised manuscript we also showed that relatively low silencing of RPS26 also decreased FMRpolyG production in model cells (new Figure 1D).

We agree that reduced ribosome levels can result in different efficiency of translation of different RNA pools. We enhance this statement in revised manuscript. However, we also showed that the same mRNA containing different near cognate start codons (single/two nucleotide substitution) specific to RAN translation, or targeting this codon with antisense oligonucleotides resulted in altered sensitivity of FMR1 mRNA translation to RPS26 depletion (new Figure 4F).

> *- A significant claim is that RPS26 KD alleviates the effects of FMRpolyG expression, but those data aren't presented well.*

We thank the Reviewer for this comment. In the new version of the manuscript, we have added new microscopic images and improved the explanation of Figure 1E. We have also completed the interpretation of Figure 1F in the main text, figure image as well as figure legend, and we hope that these changes will ameliorate understanding of our data.

> ***Recommendations For The Authors:***
>
> *- A significant claim is that RPS26 KD alleviates the effects of FMR polyG expression, but those data aren't presented well:*
>
> *Figure 1D (supporting data in S2) and 2D - the authors need to show representative images of a control that has aggregation and indicate aggregates being counted on an image. The legend states that there are no aggregates, but the quantification of*

> *aggregates/nucleus is ~1, suggesting there are at least 1 per cell. It is preferred to show at least a representative of what is quantified in the main figure instead of a bar graph.*

The representative images of control and siRPS26-treated cells are now shown in revised version of Figure 1E. Additionally, we completed the Figure legend concerning this part, as well as extended description of the experiment in Materials&Methods section.

> *Figure 1E - it is unclear what luminescence signal is being measured. Is this a dye for an apoptotic marker? More information is needed in the legend.*

This information was added to the legend of modified Figure 1F (previously 1E) as suggested.

> *- Some of the Western blots are not very convincing. Better evidence for the changes in bar graphs would improve how convincing the data are:*
>
> *Fig 2B. The western for FMR95G in the first model is not very convincing. The difference by eye for the second siRNA seems to give a larger effect than the first for 95G construct but they appear almost the same on the graph. More supporting information for the quantification is needed.*

We provided better explanation for WB quantification in M&M section in the manuscript. Alos, we provided additional blot demonstrating independent biological replicate of the mentioned experiment in supplementary materials (Supplementary Figure S2E).

> *Figure 4A, the blots for RPS26 and FMR95G are not convincing. They are quite smeary compared to all of the others shown for these proteins in other figures. Could a different replicate be shown?*

We provided additional blot demonstrating the effect on transiently expressed FMRpolyG affected by depletion of TSR2 in COS7 cell line (Supplementary Figure S4A).

> *Figure 5A and 5B blots are not ideal. Could a different replicate be shown? Or show multiple replicates in the supplemental figure?*

We provided additional blots from the same experiment, although data is not statistically significant, most likely due to low quality of normalization factor, which is Vinculin (Supplementary Figure S5A). Nevertheless, the level of FMRpolyG is decreased by ~70% after RPS25 silencing in SH-SY5Y cells.

> *Figure 2C. Please use the same y axes for all four Westerns in B and C. One would like to compare 95 and 15 repeats, but it is difficult when the y axes are different.*

Thank you for this comment. The y axis was adjusted as suggested by the Reviewer.

> *Figure 3D-The text suggests a significant difference between positive and negative responders that is not clear in the figure.*

In the main body of the manuscript we state that: “We did not observe any significant differences in the frequency of individual nucleotide positions in the 20-nucleotide vicinity of the start codon relative to the expected distribution in the BG”, which is in line with the graph showed in Figure 4D (previously 3D).

> ***Reviewer #3 (Public Review):***
>
> *Tutak et al provide interesting data showing that RPS26 and relevant proteins such as TSR2 and RPS25 affect RAN translation from CGG repeat RNA in fragile X-associated conditions. They identified RPS26 as a potential regulator of RAN translation by RNAtagging system and mass spectrometry-based screening for proteins binding to CGG repeat RNA and confirmed its regulatory effects on RAN translation by siRNA-based knockdown experiments in multiple cellular disease models and patient-derived fibroblasts. Quantitative mass spectrometry analysis found that the expressions of some ribosomal proteins are sensitive to RPS26 depletion while approximately 80% of proteins including FMRP were not influenced. Since the roles of ribosomal proteins in RAN translation regulation have not been fully examined, this study provides novel insights into this research field. However, some data presented in this manuscript are limited and preliminary, and their conclusions are not fully supported.*
>
> *(1) While the authors emphasized the importance of ribosomal composition for RAN translation regulation in the title and the article body, the association between RAN translation and ribosomal composition is apparently not evaluated in this work. They found that specific ribosomal proteins (RPS26 and RPS25) can have regulatory effects on RAN translation (Figures 1C, 2B, 2C, 2E, 4A, 5A, and 5B), and that the expression levels of some ribosomal proteins can be changed by RPS26 knockdown (Figure 3B, however, the change of the ribosome compositions involved in the actual translation has not been elucidated). Therefore, their conclusive statement, that is, "ribosome composition affects RAN translation" is not fully supported by the presented data and is misleading.*

We thank the Reviewer for critical comments and suggestions. We agree that the initial title and some statements in the text were misleading and the presented data did not fully support the aforementioned statement regarding ribosomal composition affecting FMRpolyG synthesis. Therefore, in the revised version of the manuscript we included a control experiment indicating that depletion of another core 40S ribosomal protein (RPS6) did not impact the FMRpolyG synthesis (new Figure 5C), which supports our hypothesis that RPS26 and RPS25 are specific CGG-related RAN translation modifiers. To precisely deliver a main message of our work, we changed the title that will indicate the specific effect of RPS26 and RPS25 insufficiency on RAN translation of FMRpolyG. Proposed title: “Insufficiency of 40S ribosomal proteins, RPS26 and RPS25 negatively affects biosynthesis of polyglycine-containing proteins in fragile-X associated conditions”. We also changed all statements regarding “ribosomal composition” in main text of the new version of manuscript.

> *(2) The study provides insufficient data on the mechanisms of how RPS26 regulates RAN translation. Although authors speculate that RPS26 may affect initiation codon fidelity and regulate RAN translation in a CGG repeat sequence-independent manner (Page 9 and Page 11), what they really have shown is just identification of this protein by the screening for proteins binding to CGG repeat RNA (Figure 1A, 1B), and effects of this protein on CGG repeat-RAN translation. It is essential to clarify whether the regulatory effect of RPS26 on RAN translation is dependent on CGG repeat sequence or near-cognate initiation codons like ACG and GUG in the 5' upstream sequence of the repeat. It would be better to validate the effects of RPS26 on translation from control constructs, such as one composed of the 5' upstream sequence of FMR1 with no CGG repeat, and one with an ATG substitution in the 5' upstream sequence of FMR1 instead of near-cognate initiation codons.*

We agree that the data presented in the manuscript implies that insufficiency of RPS26 plays a pivotal role in the regulation of CGG-related RAN translation and in the revised version of the manuscript we included a series of experiments indicating that ACG codon selection

seems to be an important part of RPS26 level-dependent regulation of polyglycine production (new Figure 4F&G; see point 3 below for more details). Importantly, in the luciferase assay showed on Figure 4F we used the AUG-initiated firefly luciferase reporter as normalization control.

Moreover, to verify if FMRpolyG response to RPS26 deficiency depends on the type of reporter used, we repeated many experiments using FMRpolyG fused with different tags. The luciferase-based assays were in line with experiments conducted on constructs with GFP tag (new Figure 1D), thus strengthening our previous data. Moreover, in the series of experiments, we show that FMRP synthesis which is initiated from ATG codon located in *FMR1* exon 1, was not affected by RPS26 depletion (Figure 3E & 4C), even though its translation occurs on the same mRNA as FMRpolyG. This indicates a specific RPS26 regulation of polyglycine frame initiated from ACG near cognate codon.

> *(3) The regulatory effects of RPS26 and other molecules on RAN translation have all been investigated as effects on the expression levels of FMRpolyG-GFP proteins in cellular models expressing CGG repeat sequences Figures 1C, 2B, 2C, 2E, 4A, 5A, and 5B). In these cellular experiments, there are multiple confounding factors affecting the expression levels of FMRpolyG-GFP proteins other than RAN translation, including template RNA expression, template RNA distribution, and FMRpolyG-GFP protein degradation. Although authors evaluated the effect on the expression levels of template CGG repeat RNA, it would be better to confirm the effect of these regulators on RAN translation by other experiments such as in vitro translation assay that can directly evaluate RAN translation.*

We agree that there are multiple factors affecting final levels of FMRpolyG-GFP proteins including aforementioned processes. We evaluated the level of *FMR1* mRNA, which turned out not to be decreased upon RPS26 depletion (Figure 3B&C), therefore, we assumed that what we observed, was the regulation on translation level, especially that RPS26 is a ribosomal protein contacting mRNA in E-site. We believe that direct assays such as in vitro translation may be beneficial, however, depletion of RPS26 from cellular lysate provided by the vendor seems technically challenging, if not completely impossible. Instead, we focused on sequence/structure specific regulation of RAN translation with the emphasis on start-codon initiation selection. It resulted in generating the valuable results pointing out the RPS26 role in start codon fidelity (Figure 4F&G). These new results showed that translation from mRNAs differing just in single or two nucleotide substitution in near cognate start codon (ACG to GUG or ACG to CUG), although results in exactly the same protein, is differently sensitive to RPS26 silencing (new Figure 4F). Similar differences were observed for translation efficiency from the same mRNA targeted or not with antisense oligonucleotide complementary to the region of RAN translation initiation codon (new Figure 4G). These results also suggest that stability of FMRpolyG is not affected in cells with decreased level of RPS26.

> *(4) While the authors state that RPS26 modulated the FMRpolyG-mediated toxicity, they presented limited data on apoptotic markers, not cellular viability (Figure 1E), not fully supporting this conclusion. Since previous work showed that FMRpolyG protein reduces cellular viability (Hoem G, 2019,Front Genet), additional evaluations for cellular viability would strengthen this conclusion.*

We thank the Reviewer for this suggestion. We addressed the apoptotic process in order to determine the effect of RPS26 depletion on RAN translation related toxicity (Figure 1F). In revised version of the manuscript, we also added the evaluation on how cells proliferation was affected by RPS26, RPS25, RPS6 and TSR2 depletion. Our data indicate that TSR2 silencing slightly impacted the cellular fitness (new Figure 5D), whereas insufficiencies of RPS26, RPS25 and RPS6 had a much stronger negative effect on proliferation (new Figure 2A, 5D, 6C), which

is in line with previous data (Cheng Z 2019, *Mol Cell*; Luan Y, 2022, *Nucleic Acids Res)*. The difference in proliferation rate after treatment with siRPS26 makes proper interpretation of cellular viability assessment very difficult.

> ***Recommendations For The Authors:***
>
> *(1) It would be nice to validate the effects of overexpression of RPS26 and other regulators on RAN translation, not limited to knockdown experiments, to support the conclusion.*

We did not performed such experiments because we believed that RPS26 overexpression may have no or marginal effect on translation or RAN translation. It is likely impossible to efficiently incorporate overexpressed RPS26 into 40S subunits, because the concentration of all ribosomal proteins in the cells is very high.

> *(2) It would be better to explain how authors selected 8 proteins for siRNA-based validation (Figure 1C, 1D, S1D) from 32 proteins enriched in CGG repeat RNA in the first screening.*

We selected those candidates based on their functions connected to translation, structured RNA unwinding or mRNA processing. For example, we tested few RNA helicases because of their known function in RAN translation regulation described by other researchers. This explanation was added to the revised version of the manuscript.

> *(3) Original image data showing nuclear FMRpolyG-GFP aggregates should be presented in Figure 1D.*

The representative images of control and siRPS26-treated cells are now shown in modified version of Figure 1E and described with more details in the legend.

> *(4) Image data in Figure 2A and 2D have poor signal/noise ratio and the resolution should be improved. In addition, aggregates should be clearly indicated in Figure 2D in an appropriate manner.*

The stable S-FMR95xG cellular model is characterized by very low expression of RANtranslated FMR95xG, therefore, it is challenging to obtain microscopic images of better quality with higher GFP signal. In the L-99xCGG model expression of transgene is higher. Therefore, we provided new image in the new version of Figure 3D (former 2D). Moreover, we showed aggregates on the image obtained using confocal microscopy (new Supplementary Figure 2D).

> *(5) The detailed information on patient-derived fibroblast (age and sex of the patient, the number of CGG repeats, etc.) in Figure 2F needed to be presented.*

This information was added to the figure legend (Figure 3F; previously 2F) and in the Material and Methods section as suggested.

> *(6) It would be better to normalize RNA expression levels of FMR1 and FMR1-GFP by the housekeeping gene in Figure S2C, like other RT-qPCR experimental data such as Figure 2B.*

Normalization of FMR1-GFP to GAPDH is now shown in modified version of Figure S2C (right graph) as requested by the Reviewer.

> *(7) It would be better to add information on molecular weight on all Western blotting data.*
>
> *(8) Marks corresponding to molecular weight ladder were added to all images.*

Full blots, including protein ladders were deposited in Zenodo repository, under doi: 10.5281/zenodo.13860370

References

Cheng Z, Mugler CF, Keskin A, Hodapp S, Chan LYL, Weis K, Mertins P, Regev A, Jovanovic M & Brar GA (2019) Small and Large Ribosomal Subunit Deficiencies Lead to Distinct Gene Expression Signatures that Reflect Cellular Growth Rate. *Mol Cell* 73: 36-47.e10

Havkin-Solomon T, Fraticelli D, Bahat A, Hayat D, Reuven N, Shaul Y & Dikstein R (2023) Translation regulation of specific mRNAs by RPS26 C-terminal RNA-binding tail integrates energy metabolism and AMPK-mTOR signaling. *Nucleic Acids Res* 51: 4415–4428

Hoem,G., Larsen,K.B., Øvervatn,A., Brech,A., Lamark,T., Sjøttem,E. and Johansen,T. (2019) The FMRpolyGlycine protein mediates aggregate formation and toxicity independent of the CGG mRNA hairpin in a cellular model for FXTAS. Front. Genet., 10, 1–18.

Luan Y, Tang N, Yang J, Liu S, Cheng C, Wang Y, Chen C, Guo YN, Wang H, Zhao W, *et al* (2022) Deficiency of ribosomal proteins reshapes the transcriptional and translational landscape in human cells. *Nucleic Acids Res* 50: 6601–6617

Plassart L, Shayan R, Montellese C, Rinaldi D, Larburu N, Pichereaux C, Froment C, Lebaron S, O'donohue MF, Kutay U, *et al* (2021) The final step of 40s ribosomal subunit maturation is controlled by a dual key lock. *Elife* 10

https://doi.org/10.7554/eLife.98631.2.sa0