

Dissecting the mechanism of *NOP56* GGCCUG repeat-associated non-AUG translation using cell-free translation systems

Received for publication, December 6, 2024, and in revised form, February 18, 2025 Published, Papers in Press, February 25, 2025,

<https://doi.org/10.1016/j.jbc.2025.108360>

Mayuka Hasumi^{1,†}, Hayato Ito^{1,†}, Kodai Machida², Tatsuya Niwa^{1,3}, Tomoya Taminato⁴, Yoshitaka Nagai⁴, Hiroaki Imataka², and Hideki Taguchi^{1,3,*}

From the ¹School of Life Science and Technology, Institute of Science Tokyo, Yokohama, Japan; ²Department of Applied Chemistry, Graduate School of Engineering, University of Hyogo, Himeji, Hyogo, Japan; ³Cell Biology Center, Institute of Integrated Research, Institute of Science Tokyo, Yokohama, Japan; ⁴Department of Neurology, Kindai University Faculty of Medicine, Osaka-Sayama, Japan

Reviewed by members of the JBC Editorial Board. Edited by Karin Musier-Forsyth

The repeat expansion in the human genome contributes to neurodegenerative disorders such as spinocerebellar ataxia (SCA) and amyotrophic lateral sclerosis. Transcripts with repeat expansions undergo noncanonical translation called repeat-associated non-AUG (RAN) translation. The *NOP56* gene, implicated in SCA36, contains a GGCCTG repeat in its first intron. In tissues of patients with SCA36, poly (Gly-Pro) and poly (Pro-Arg) peptides, likely produced through *NOP56* RAN translation in (*NOP56*-RAN), have been detected. However, the detailed mechanism underlying *NOP56*-RAN remains unclear. To address this, we used cell-free translation systems to investigate the mechanism of *NOP56*-RAN and identified the following features. (i) Translation occurs in all reading frames of the sense strand of *NOP56* intron 1. (ii) Translation is initiated in a 5' cap-dependent manner from near-cognate start codons upstream of the GGCCUG repeat in each frame. (iii) Longer GGCCUG repeats enhance *NOP56*-RAN. (iv) A frameshift occurs within the GGCCUG repeat. These findings provide insights into the similarities between *NOP56*-RAN and other types of RAN translation.

Spinocerebellar ataxia type36 (SCA36) is a neurodegenerative disease caused by a hexanucleotide repeat expansion (1, 2). Patients have been reported in specific regions, including Spain, Japan, and China (3, 4). The primary clinical features include hearing loss and tongue atrophy, along with motor neuron loss and muscle atrophy (1, 2, 5, 6). The GGCCTG hexanucleotide repeat expansion in intron 1 of the *NOP56* gene, whose exon encodes a 60S ribosomal biogenesis protein, is associated with SCA36 (Fig. 1A). Healthy individuals have 5 to 14 GGCCTG repeats, whereas patients with SCA36 typically exhibit expansions ranging from 650 to 2500 repeats. Additionally, it has been reported that as few as 25 repeats can cause the disease (2, 5, 7). The GGCCUG expansion impairs splicing of the *NOP56* intron1, leading to the production of

intron-retained transcripts (8). Indeed, transcripts with aberrant GGCCUG repeats have been detected in patient cells and are thought to be cytotoxic by forming RNA foci (1, 9–12), a feature commonly observed in other repeat-associated neurodegenerative diseases (13, 14). In addition to this RNA-mediated toxicity, a noncanonical process known as repeat-associated non-AUG (RAN) translation contributes to SCA36 pathogenesis (8, 10).

RAN translation, first identified in spinocerebellar ataxia type 8 (SCA8) and myotonic dystrophy type 1 (DM1) (15), is implicated in various neurodegenerative diseases associated with repeat expansions (14, 16). While the molecular mechanisms of RAN translation vary among diseases, a common feature is the involvement of mRNA secondary structures, such as G-quadruplex (G4) and hairpin structures, in the GC-rich repeat expansion region (17). The initiation of C9-RAN (*C9orf72*, GGGGCC) and CGG-RAN (*FMR1*, CGG) is 5'-cap dependent (18, 19), starting from near-cognate AUG codons upstream of the repeat sequences in each reading frame, facilitated by scanning mechanisms involving eIF4A and eIF4F (18–23). Similarly, SCA3-RAN (*ATXN3*, CAG) initiates from non-cognate codons (e.g., CUU, ACU) upstream of the repeat sequences (24). In addition to cap-dependent translation, C9-RAN also employs a cap-independent mechanism resembling an internal ribosomal entry site (IRES), as shown using bicistronic reporter assays (25, 26). Ribosomal frameshifting during elongation on repeat sequences has also been observed, such as GA-to-GP frameshifts in C9-RAN (19, 27, 28) and R-to-G frameshifts in CGG-RAN (29)). These frameshifts result in chimeric peptides with distinct subcellular localization and toxicity compared to non-frameshift peptides (8, 29). Furthermore, the integrated stress response (ISR) enhances the translation efficiency of C9-RAN, CGG-RAN, and SCA3-RAN (20, 24).

Regarding SCA36, the GGCCTG repeat in the *NOP56* gene has been identified in patient cells and mice (8). Possible RAN translation products of this repeat include poly Gly-Leu (GL, from GGC-CUG, 0 frame), poly Ala-Trp (AW, from GCC-UGG, +1 frame), Pro-Gly (PG from

[†] These authors contributed equally to this work.

* For correspondence: Hideki Taguchi, taguchi@bio.titech.ac.jp.

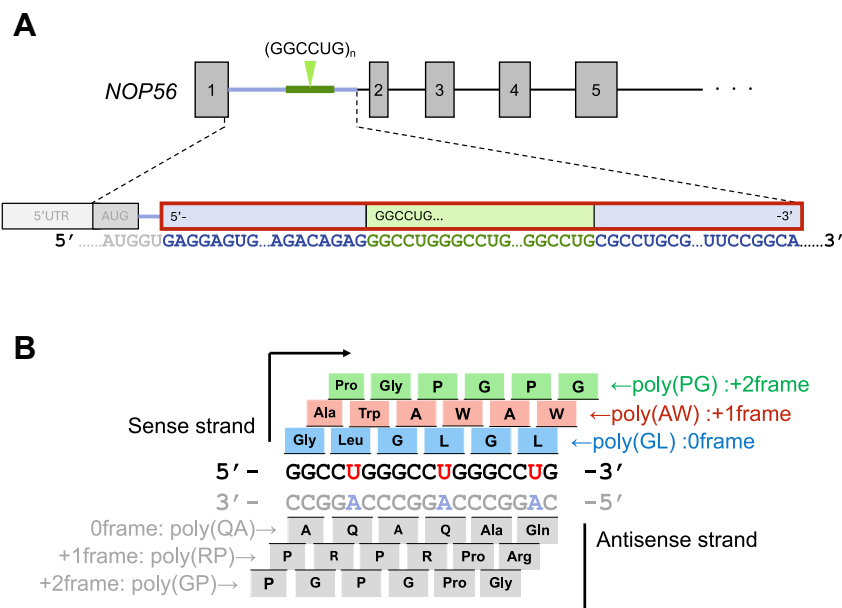


Figure 1. NOP56 gene structure and the GGCCUG-derived putative dipeptide repeats. A, NOP56 gene structure. The GGCCUG repeat in intron 1 of the NOP56 gene and its flanking regions are shown in detail. The red box highlights the region used in this study. B, schematic of putative dipeptide repeats derived from the NOP56 GGCCUG repeat.

CCU-GGG, +2 frame) dipeptide repeats (DPRs) on the sense strand and poly Arg-Pro (RP), poly Glu-Ala (QA), and poly Pro-Gly (GP) DPRs on the antisense strand (Fig. 1B). Poly GP and poly PR products translated in NOP56 GGCCUG repeat sequence have been detected in cells of patients with SCA36 (10). McEachin *et al.* (8) reported that the AUG start codon in exon1 of NOP56 predominantly generates poly PG DPR. In the absence of the exon1 AUG start codon, poly GL, poly AW, and poly GP are also translated from the GGCCUG repeat (10). However, the detailed mechanisms of non-AUG translation initiation, including the specific initiation sites, remain unclear.

Here, we investigated the initiation and elongation mechanisms of the NOP56-GGCCUG RAN translation (NOP56-RAN) using the repeat without an AUG start codon in several cell-free translation systems, based on our recent study of C9-RAN (27). Reporter assays with Western blotting and nano luciferase (Nluc) revealed several features of NOP56-RAN, including the identification of initiation sites, 5' cap-dependency, and the influence of longer repeats. These findings provide new insights into RAN translation, potentially contributing to therapeutic strategies for SCA36.

Results

NOP56-RAN reporter series using nano-luciferase

To investigate the detailed mechanism of NOP56-RAN, we developed a reporter series using only intron1 of the NOP56 gene, which lacks an AUG codon in any reading frame (Figs 1A, 2A). This reporter system fused nano-luciferase (Nluc) and a tandem-HA tag downstream of the GGCCUG repeat. To examine all three possible reading frames, we

inserted one or two nucleotides between the repeat and Nluc (0 nt: GL (0), +1 nt: AW (+1), +2 nt: GP (+2)) (Fig. 2A).

Translation of the reporter series containing 69 to 71 GGCCUG repeats, we performed *in vitro* translation in HeLa lysate and rabbit reticulocyte lysate (RRL) with mRNA containing a 5'-cap structure. Western blotting detected translation products from all three frames, although the GP products were faint in HeLa lysate (Fig. 2, B and C). Consistent with previous reports (8, 10), products from the GL (0) and AW (+1) frames exhibited higher molecular weights than their predicted sizes (~40 kDa for GL, ~48 kDa for AW) (Fig. 2, B and C). The presence of these larger products became more pronounced as the repeat length increased (Fig. S1A, B), suggesting that longer GL and AW DPRs are aggregation-prone, likely due to their higher hydrophobicity. Notably, overall translation efficiency was higher in RRL than in HeLa lysate (Fig. S1C).

For the luciferase activity of this reporter system, all frames showed 100- to 1000-fold higher activity compared to the control reporter containing GGG-Nluc (Fig. 2, D and E). The highest translation activity was observed in the AW frame, reaching approximately 30% (RRL) or 80% (HeLa lysate) of the activity of AUG + Nluc, a positive control, indicating robust translation even without a canonical AUG start codon. Moreover, we inserted a self-cleaving T2A peptide (30) between DPR and Nluc to investigate whether the aggregation of AW and GL affects Nluc activity (Fig. S2A). As a result, the self-cleavage of DPR enhanced by 1.5-fold Nluc activity in both AW and GL (Fig. S2, B–C). However, this enhancement was insufficient to account for the differences in translation efficiency between AW, GL, and GP (Fig. 2, D and E).

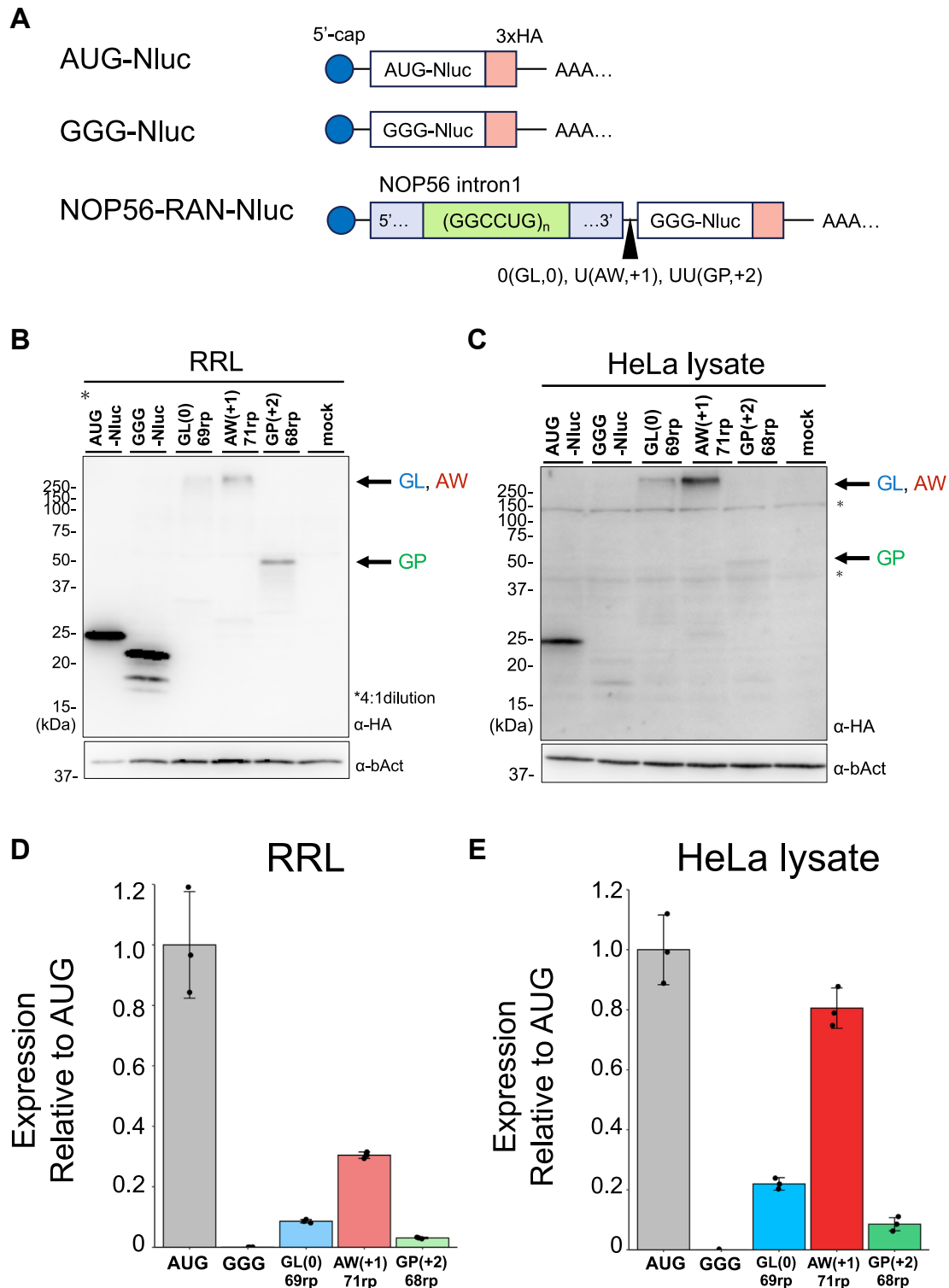


Figure 2. NOP56-RAN translation in cell-free translation systems using nano-luciferase reporter. A, schematic of nano-luciferase (Nluc)-3 × HA reporters. Nluc was used as a reporter enzyme to quantify translation efficiency. B and C, anti-HA Western blot of translation products in each cell-free translation system using the Nluc reporters. B, Rabbit reticulocyte lysate (RRL), (C) HeLa lysate. Asterisks (*) indicate non-specific bands. Predicted molecular sizes: AUG-Nluc: 25 kDa, GL (0): 41 kDa, AW (+1): 46 kDa, and GP (+2): 39 kDa. Anti-β-actin (bAct) was used as a loading control. D and E, expression of NOP56-Nluc reporters normalized to GGG-Nluc expressed in cell-free translation systems. D, RRL and (E) HeLa lysate. Error bars represent standard deviations (±SD) from three independent experiments.

NOP56-RAN utilizes upstream AUG-like codons as initiation sites in each frame

We next examined the 5' m⁷GpppG-cap dependency in NOP56-RAN and found that translation with a 5' ApppG-cap, a cap analog, in HeLa lysate almost abolished (Fig. S3A). This 5'-cap dependency was further confirmed by inhibited translation in the presence of free m⁷GpppG compound (Fig. S3B), which competitively inhibits eIF4E binding.

Based on the 5'-cap dependency, NOP56-RAN is likely initiated from specific sites upstream of the GGCCUG repeat, similar to findings in SCA3-RAN experiments (24). Translation of a series of reporters with various truncations of the 5' end of NOP56 intron1 showed an overall decrease in translation efficiency (Fig. S3, C–F). Based on these results, we generated reporter mRNAs with mutations at the predicted near-cognate start codons, replacing them with AAA (non-initiation) codons, and tested them in conjugation with (GGCCUG)₁₆ repeats (Fig. 3A). The results for each frame were as follows.

- i) GL (0)-frame: In the GL (0)-frame, mutations in either GUG (mut 0–1) or CUG (mut 0–2) reduced luciferase activity by approximately 60%, and by ~80% when both mutations were introduced (HeLa lysate; Fig. 3B, RRL; Fig. S4A), suggesting that either GUG or CUG serves as the initiation codon. Additionally, Western blotting detected different molecular sizes of translation products in each mutant (Fig. S4D), supporting the use of two initiation sites.
- ii) AW (+1)-frame: All three possible mutations in the near-cognate start codons decreased luciferase activity by 40 to 60% (HeLa lysate; Figs. 3C, RRL; S4B, E). Combining mutations in the two near-cognate start codons showed that the double mutation of ACG (mut 1–2) and GUG (mut 1–3) nearly eliminated luciferase activity (Figs. 3C, S4B, E), suggesting that ACG and GUG are critical for AW (+1)-frame initiation.
To obtain further evidence, we performed LC-MS/MS analysis of the AW (+1)-frame initiation site, as previously applied in other RAN translation studies (21, 22, 24). We translated NOP56-RAN reporters with 16 or 71 GGCCUG repeats in RRL, immunoprecipitated the translation products with HA beads, digested them with a trypsin/LysC mix, and analyzed the peptides by LC-MS/MS. We identified two unique peptides from the N-terminal region of the NOP56-RAN AW (+1)-frame, VGVSAACVR and MGVSAACVR (Fig. 3E, Fig. S5, A–C). These peptides were predicted to initiate from the ACG codon, with and without the removal of the N-terminal methionine, strongly suggesting that the AW (+1)-frame is primarily initiated from the ACG codon (mut 1–2).
- iii) GP (+2)-frame: Of the two possible near-cognate start codons, mutation in ACG (mut 2–2) significantly decreased luciferase activity (HeLa lysate; Figs. 3D, RRL; S4, C, F), indicating that the +2 frame initiates from the ACG codon.

Taken together, NOP56-RAN uses specific near-cognate codons upstream of the GGCCUG repeat in each reading frame as initiation sites.

Next, we investigated whether putative initiation codons affect the translation of other frames in HeLa lysate. Mutating the putative initiation codons for the GL (0)-frames enhanced the translation efficiency of the AW (+1)-frame (Fig. S6), suggesting that translation in the AW (+1)-frame is suppressed by translation in the GL (0)-frame.

NOP56-RAN is enhanced by longer GGCCUG repeats

Studies on RAN translation have shown that longer RNA repeat expansions, such as the GGGGCC repeats in *C9orf72*, enhance translation efficiency (15, 18, 27, 31–33). To investigate this repeat length dependency in NOP56-RAN, we created a reporter series with different GGCCUG repeat lengths (0–71 repeats) (Fig. 4A). In HeLa lysate, translation in all frames was more efficient in the mRNA with 3 repeats than with no GGCCUG repeat (=0 repeat) (Fig. 4, B–D). Notably, in the GL (0) frame, the translation efficiency with 3 repeats was more than 5-fold higher (Fig. 4B). Further repeat expansion generally increased translation efficiency in all frames, although no increase was observed in the GP (+2) frame with more than 16 repeats (Fig. 4, B–D).

The AW (+1)-to-GP (+2) frameshift occurs within NOP56 intron1

Next, we investigated the potential for a frameshift in NOP56-RAN. We focused on a frameshift from the AW (+1)-frame to other frames, as translation from the +1 frame was most efficient in our assay system. We engineered a dual-tagged reporter series, in which the AUG start codon was inserted in-frame in the AW (+1)-frame, upstream of the NOP56 intron1 (Fig. 5A). For detection, we introduced a FLAG tag in the N-terminal region and Nluc-HA downstream of the repeat sequence, adding one or two nucleotides to distinguish frameshift products (Fig. 5A).

Western blotting after translation using RRL showed that the construct for the AW-to-GP frameshift (+1 to +2) produced faint ~50 kDa protein bands in both FLAG and HA detection (Fig. 5B). The ~50 kDa protein corresponds to the approximate calculated size (~40 kDa) for the AW-to-GP frameshifted product. Additionally, no large band (>250 kDa) was observed in the anti-HA detection for the GP (+2) construct, which would indicate a non-frameshifted AW-frame product, as seen in Figure 2. These results suggest an AW-to-GP frameshift under the tested conditions. The luciferase activities for the AW-to-GP and AW-to-GL frameshift constructs relative to the AW-to-AW construct in the 71 repeat constructs were 4.0% and 1.1%, respectively (Fig. 5C), indicating that the AW-to-GP frameshift predominates. The frameshift efficiencies in the 71 repeat series were higher than those in the 3-repeat series (AW-to-GP: 2.5%, AW-to-GL: 0.8%) (Fig. 5C). In addition to the AW-to-GP frameshift, we also investigated the possibility of a GL-to-GP frameshift. The

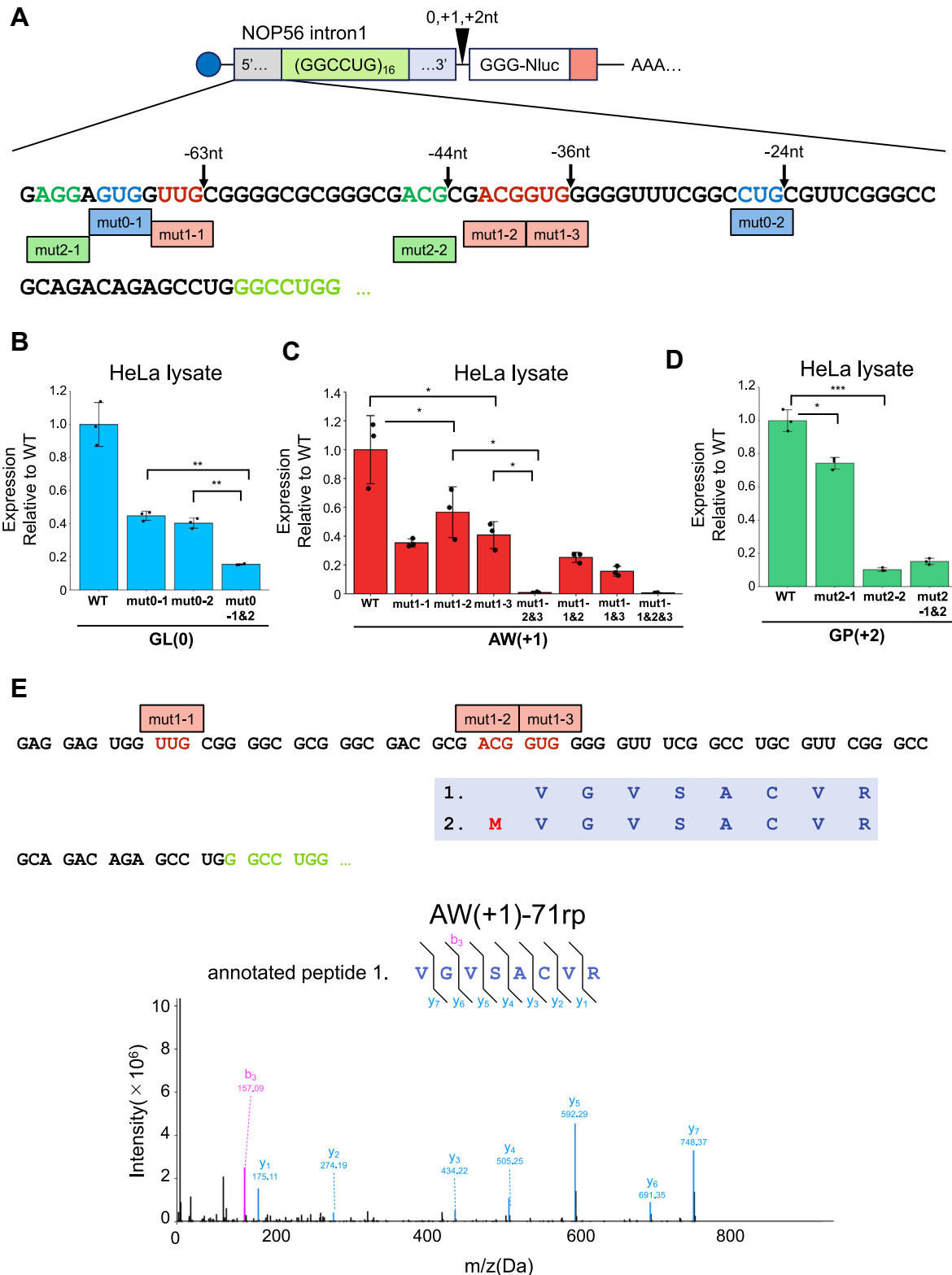


Figure 3. Identification of initiation sites in each reading frame. A, mutations to identify NOP56-RAN start codons. The GGCCUG repeat length is 16 repeats. For example, “mut0-1” refers to the first candidate start codon in the 0 frame. (B–D) Expression of the reporters with mutation relative to the wild-type (WT) reporter expressed in HeLa lysate. B, GL (0), (C) AW (+1) and (D) GP (+2). Error bars represent \pm SD from three independent experiments. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$, two tailed Student’s t test. E, LC-MS/MS analysis of the N-terminal region of the HA-immunoprecipitated and trypsin-digested poly AW-71 repeat protein expressed in RRL. LC-MS/MS analysis identified two peptides, 1-VGVSAACVR, 2-MGVSAACVR. The MS spectrum for peptide 1 is shown, with related MS spectra provided in Fig. S5.

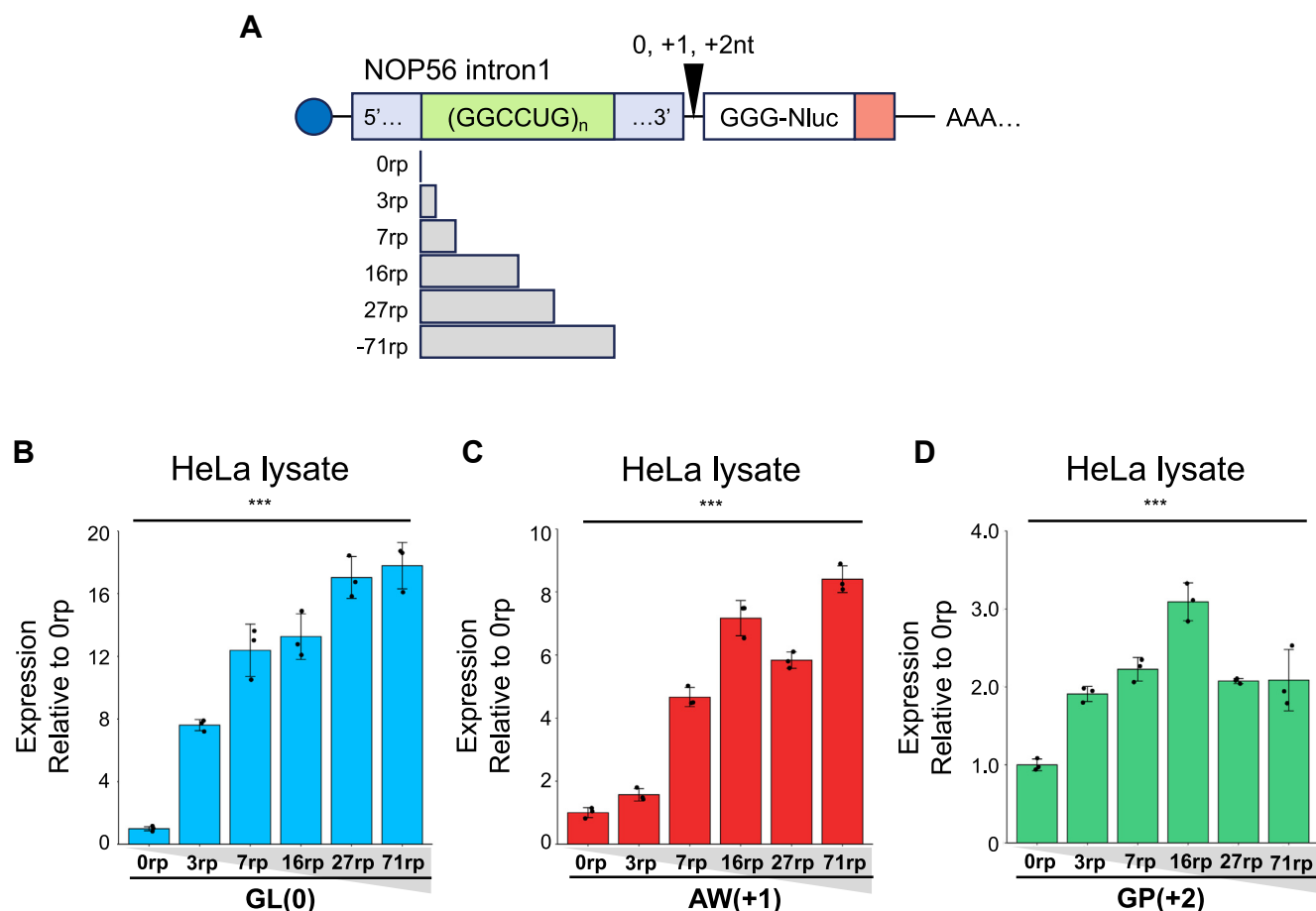


Figure 4. Repeat length dependency in NOP56-RAN. A, schematic of the NOP56-RAN reporter with varying repeat (rp) numbers in each frame. B–D, relative expression levels of NOP56-RAN reporters normalized to the 0 repeat in HeLa lysate. B, GL (0), C, AW (+1), and D, GP (+2). Error bars represent \pm SD from three independent experiments. *** $p < 0.001$, one-way ANOVA with Dunnett's multiple comparison test.

luciferase reporter assay for X-to-GP frameshifts in 71 repeats were 1.2% and 4.0%, respectively. These frameshift efficiencies in 71 repeats were higher than those in 3 repeats (GL-to-GP: 0.6%, AW-to-GP: 2.5%) (Fig. S7, A–C). Although these analyses were performed using RRL due to the large amount of translation products required for the immunoblotting, the luciferase assay showed similar trends when using HeLa lysate (Fig. S7D, E).

Furthermore, we sought to identify the site of the AW-to-GP frameshift. It is known that G or C runs can induce +1 frameshift (19, 29, 34–36), and both C9-RAN and CGG-RAN are associated with +1 frameshifts (19, 29, 34–36). Therefore, we hypothesized that the frameshift might occur at consecutive G or C sequences upstream of the repeat in NOP56-RAN. We identified several G-runs (GGGG, GGG, and GGGGG, Fig. 5D) upstream of the repeat region and mutated them to GAAG (mut A), GAG (mut B), and GAAAG (mut C), respectively (Fig. 5D). When these reporter RNAs were translated in RRL, the frameshift efficiency decreased from 4% to 1% when the GGGGG sequence was mutated (mut C) (Fig. 5E), as confirmed by Western blotting (Fig. S8A). Similarly, the frameshift efficiency decreased from 2% to 1% in HeLa lysate (Fig. S8B). These results suggest that the G-rich region upstream of the repeat contributes to the AW-to-GP frameshift.

NOP56-RAN in human cultured cells

Finally, we investigated the translation of the GGCCUG repeat in HeLa cells to compare the results obtained by cell-free systems with those in cells. The Nluc reporters, including NOP56 intron1 with 71 GGCCUG repeats, were transfected into HeLa cells along with a plasmid carrying firefly luciferase (Fluc) for normalization (Fig. 6A). Western blotting revealed translation products in all frames; both AW and GL, though in low amounts, formed high molecular weight oligomers (Fig. 6B), similar to the cell-free translation conditions. The dual-luciferase reporter assay revealed that the AW (+1) frame was the most prominently translated, with \sim 3 and 10-fold higher translation than the GL (0) and GP (+2) frames, respectively (Figs. 6C, S9A). The trend observed in HeLa cell translation was consistent with the cell-free translation results (Fig. 2E).

We also tested whether ISR stimulates NOP56-RAN translation, as observed in other RAN translations (20, 24, 25). Remarkably, the addition of thapsigargin (TG), which induces ER stress, increased translation of all frames by 1.6 to 2.5-fold (Fig. 6D). This stimulation is likely due to ISR-mediated activation of non-AUG start codons, as TG enhanced the translation of near-cognate start codons, ACG and CUG (Fig. S9C).

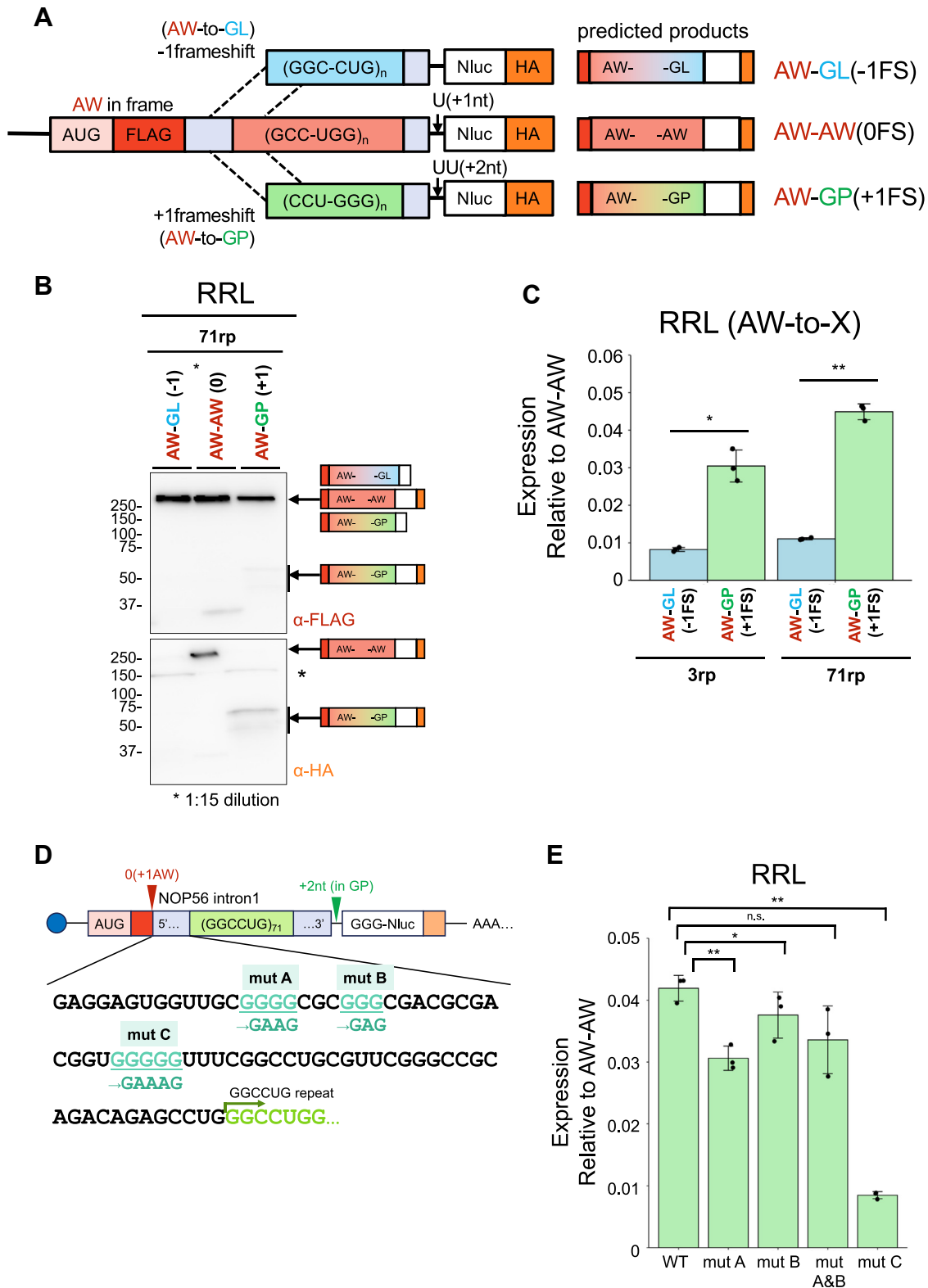


Figure 5. Frameshift analysis in NOP56-RAN. A, Schematic of reporters with an AUG driving expression through an N-terminal FLAG (red) tag in AW (+1), followed by a C-terminal 3 × HA (orange) tag fused to the NOP56 intron 1 with 3 or 71 repeats (rp). B, Anti-FLAG and HA Western blot of the FLAG (AW)-NOP56-Nluc, HA reporter mRNAs expressed in each frame in RRL. Asterisks (*) indicate non-specific bands. Predicted molecular sizes for anti-FLAG: AW-GL and AW-GP (+2) are smaller than 46 kDa (71 rp); AW-AW in frame: 46 kDa or high molecular weight (71 rp). Predicted molecular sizes for anti-HA: AW-GL not detected; AW-AW in frame: 46 kDa; AW-GP: 40 to 47 kDa. C, Luciferase activity of frameshifted products (AW-to-GL, AW-to-GP) expressed in RRL. Relative values are shown when the luciferase activity for in-frame (AW-to-AW) is set to 1. Error bars represent \pm SD from three independent experiments. * p < 0.05; ** p < 0.01, two tailed Student's t test. D, Schematic of reporters with mutations at putative frameshift sites. E, Frameshift efficiency of the AW-to-GP mutations at putative frameshift sites. Relative expression of AW-AW expressed in RRL. Error bars represent \pm SD from three independent experiments. n.s. = not significant; * p < 0.05; ** p < 0.01, two tailed Student's t test.

GGCCUG-RAN translation using cell-free translation systems

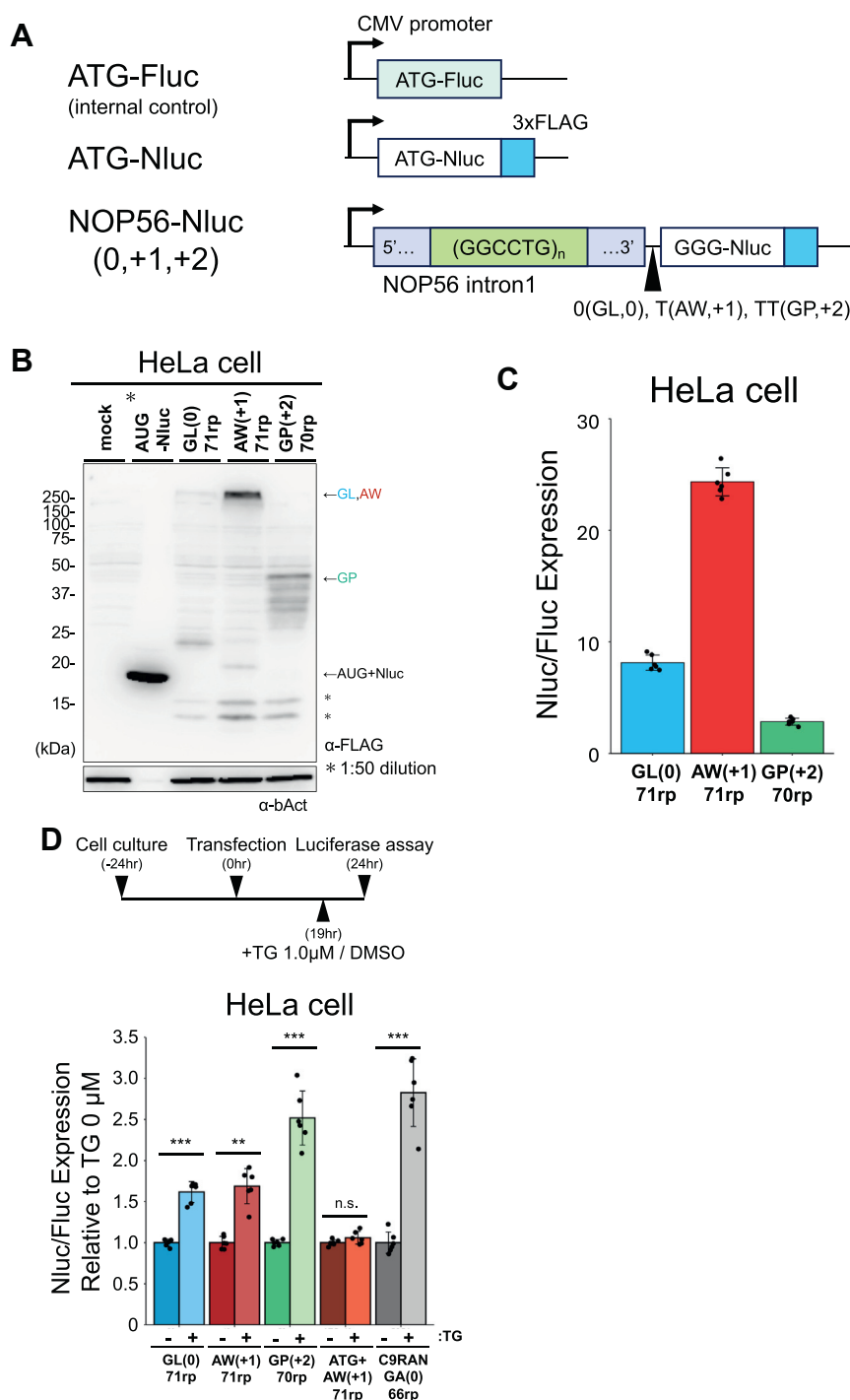


Figure 6. NOP56-RAN in human cultured cells. *A*, schematic of Nluc-3 × FLAG reporters. Nluc was used as a reporter enzyme to quantify translation efficiency. ATG-Fluc was included in all assays for normalization. *B*, anti-FLAG Western blot of the NOP56-Nluc reporter expressed in HeLa cells. Asterisks (*) indicate non-specific bands. Predicted molecular sizes of the products: AUG + Nluc: 21 kDa, GL (0): 38 kDa, AW (+1): 43 kDa, and GP (+2): 36 kDa. *C*, relative expression of the NOP56-Nluc reporters normalized to ATG-Fluc in HeLa cells. Error bars represent standard deviations (\pm SD) from six independent experiments. *D*, *top*, schematic of stress induction by thapsigargin (TG). *Bottom*, relative Nluc/Fluc expression of the NOP56-Nluc reporters normalized to the HeLa cells without TG. Error bars represent standard deviations (\pm SD) from six independent experiments. n.s. = not significant; ** p < 0.01; *** p < 0.001, two tailed Student's t test.

Discussion

Here, we investigated the cell-free translation of *NOP56* GGCCUG repeats by developing a NOP56-RAN reporter series from sequences surrounding the GGCCUG in the *NOP56* intron1. Although previous studies have reported that all frames

can be translated from the GGCCUG repeats (8, 10), we examined the detailed mechanism of NOP56-RAN in the same reporter system using constructs lacking AUG upstream of intron1.

Using this reporter system, we identified putative initiation codons for NOP56-RAN in three frames: CUG and GUG in

the GL (0)-frame, primarily ACG in the AW (+1)-frame, and GP (+2)-frame. In the AW (+1)-frame, MS analysis identified two peptides, VGVSACVR and MGVVSACVR, in the potential N-terminal regions, strongly suggesting that the methionine is incorporated at ACG as the start codon. Similar to other RAN translations, CUG in the GA frame and AGG in the GR frame in C9-RAN, ACG and GUG in the G frame in CGG-RAN, and CUU and ACU in the Q frame in SCA3-RAN have been identified as start codons (21, 22, 24). Taken together, these results suggest that NOP56-RAN utilizes multiple potential initiation sites in all frames upstream of the GGCCUG repeats.

By translating the three frames under the same conditions, the AW (+1)-frame was found to be the most efficient. This may be due to the presence of multiple non-AUG codons in the region before the GGCCUG repeat sequence. In the AW (+1) frame, ACG in (GCG)ACG(G) was identified as the start codon, and the sequence flanking ACG closely resembles a typical Kozak sequence (GCCAUGG) (37). Moreover, it is known that the distance from the initiation codon with a Kozak sequence to bulky secondary structures, such as repeat sequences, is critical for initiating translation (38, 39). In NOP56-RAN, the near-cognate start codon ACG, located closest to the bulky GGCCUG repeat (−36 nt), aligns with the Kozak consensus sequence. This likely accounts for the higher translation efficiency of the AW (+1) frame.

Non-AUG translation, including RAN translation, is known to be enhanced by ISR (20, 25, 40), as observed in NOP56-RAN. One reason for this enhancement is the phosphorylation of eIF2 α , a translation initiation factor (41–43). ISR-induced phosphorylation of eIF2 α is thought to suppress global canonical AUG-start translation and promotes noncanonical translation, such as non-AUG initiation (44, 45). In addition to eIF2 α , ISR also modifies other translation factors, including eIF2B (46, 47) and eIF3d (48). Although the impact of these factors on RAN translation remains unclear, it is possible that the global remodeling of translation landscape by ISR may enhance non-AUG translation by altering the relative abundance of translation-related factors.

As with other RAN translations, NOP56-RAN is overall facilitated by longer repeats. This repeat-length dependence may be due to the bulky secondary structure formed by the repeat RNA (39, 49). Indeed, the GGCCUG repeat in the NOP56 gene is thought to form a G4 structure (9, 12), similar to the GGGGCC repeat in C9-RAN (19, 26, 27).

We observed a +1 frameshift from the AW (+1) to the GP (+2) frame in NOP56-RAN. In general, the mechanisms of −1 and +1 frameshifts differ. The −1 frameshift is primarily involved in viral replication control and occurs in U-rich regions (slippery sequences) just before sequences that form pseudoknot structures (50–53). In contrast, the +1 frameshift, including GA-to-GP and GR-to-GA in C9-RAN, R-to-G in CGG-RAN (19, 27, 29), is thought to be induced by ribosome stalling within repeat RNAs with certain structures, such as G4 (34, 54). The ability of GGCCUG repeats to form G4 RNA structures may contribute to the +1 frameshift.

We observed repeat-length stimulation of the AW-to-GP frameshift, part of which was not abolished by an upstream mutation (Fig. 5, C and E), suggesting that the frameshift occurs within the repeat, as seen in C9-RAN (19, 28). Based on this and previous reports of frameshifts in RAN translation, it can be predicted that frameshifts occur within repeats in guanine-rich regions, albeit as minor events (~4%).

Although this study was primarily conducted in a cell-free system, it provides valuable insights into the mechanism of NOP56-RAN. While the cell-free translation system used in this study employed extract-based systems, we have recently developed a system to investigate RAN translation (C9-RAN) using a reconstituted translation system with only human-derived translation factors (Human PURE system (55, 56)). Using the Human PURE system as an extension of this study, a more detailed mechanism could be explored. For example, reproducing the changes in the proportion of initiating factors, such as eIF2B, and eIF3d, during ISR may help identify potential therapeutic targets. Finally, analyzing the detailed molecular mechanisms of different RAN translations will highlight both similarities and differences, which may help uncover new eukaryotic translation mechanisms unrelated to disease.

Materials and methods

Plasmids

All plasmids were constructed using standard cloning procedures and Gibson assembly. To maintain the repeat size, we used SURE2 *Escherichia coli* strain for all plasmid preparations. Detailed information on plasmids and primers is summarized in Tables S1 and S2. Briefly, to generate the NOP56-RAN reporter plasmid, the NOP56-intron1 region, lacking an ATG start codon, was PCR-amplified from pUAST-attB-NOP56-GGCCCTG-71 repeats using the primers HI_240 (5'-GAGGAGTGGTTGCGGGGCGCGGGCGACGCGAC-3') and HI_090 (5'-TGCCGGAACCCGTTCCAGGGCAGG-3') with Prime STAR Max DNA Polymerase (Takara). The PCR conditions were 35 cycles of 10 s at 95 °C, 8 s at 58 °C, and 10 s at 72 °C, followed by 5 min at 72 °C. The PCR product was subcloned into the pcDNA5/FRT-T7-3FT vector (27). The size and integrity of the amplified repeat by PCR were verified by gel electrophoresis and sequencing. Various lengths of NOP56 GGCCCTG repeats were randomly generated by PCR. To construct the NOP56-Nluc+3 × HA reporter plasmids (for *in vitro* translation), the Nluc+3 × HA region was amplified from the Nluc+3 × HA sequence (27) and inserted downstream of the pcDNA5_FRT-T7-NOP56 vector. Initiation codon-mutated reporters (Fig. 3A) were generated using PCR and Gibson assembly. For the human-transfected NOP56 reporter plasmids, Nluc-3 × FLAG was amplified and inserted downstream of the pcDNA-FRT-T7-NOP56 vector.

In vitro transcription

Reporter mRNAs were synthesized as previously described (57). Reporter plasmids were linearized with XbaI (Takara), and the digested plasmids were purified using the Wizard SV

GGCCUG-RAN translation using cell-free translation systems

Gel and PCR Purification Kit (Promega). 5' m⁷G-capped mRNAs were synthesized using mMESSAGE mMACHINE T7 Transcription Kit (Thermo Fisher) and subsequently polyadenylated with *E. coli* poly(A) polymerase (NEB). 5' A-capped mRNAs were synthesized using HiScribe T7 High Yield RNA Synthesis Kit (NEB) and G (5') ppp (5') A RNA Cap structure Analog (NEB). The mRNAs were purified by lithium chloride precipitation, and their size and integrity were evaluated through denaturing RNA gel electrophoresis.

In vitro translation in RRL

mRNAs were translated *in vitro* using the Flexi Rabbit Reticulocyte Lysate system (Promega) as previously described (57). Translation reactions were conducted with 3 nM final mRNA and contained 30% RRL, 10 μ M amino acid mix minus methionine, 10 μ M amino acid mix minus leucine, 0.5 mM MgOAc, 100 mM KCl and 0.8 U/ μ l Murine RNase inhibitor (NEB). For 5' cap dependency experiments, uncapped mRNA was used under the same conditions. Translation efficiency was evaluated using luminescence assay and Western blotting as previously described (27). For the luminescence assay, 2 μ l of the sample was diluted in Glo lysis buffer (Promega) and mixed with NanoGlo substrate (diluted 1:50 in NanoGlo buffer) at a 1:1 ratio. After 3 min incubation in the dark with shaking, luminescence was measured using a Varioskan LUX Multimode Microplate Reader (Thermo Fisher Scientific). For Western blotting, 9 μ l of each sample was mixed with 30 μ l of 4 \times SDS sample buffer (240 mM Tris-HCl pH 6.8, 40% (v/v) glycerol, 0.01% (w/v) bromophenol blue, 7% (w/v) SDS, 10% (v/v) of 2-mercaptoethanol) and heated at 70 $^{\circ}$ C for 15 min. Next, 15 μ l of the samples were loaded onto a 13% SDS-polyacrylamide gel for electrophoresis. After separation, proteins were transferred onto PVDF membranes and blocked with 2% (w/v) skim milk in TBS-T and incubated for 1 h with primary antibodies listed in Table S3. Membranes were washed three times with TBS-T for 5 min each and incubated with HRP-conjugated anti-mouse IgG (Sigma-Aldrich). Chemiluminescence signals were detected using an LAS4000 system (Fuji Film) with Immobilon Western Chemiluminescent HRP reagent (Millipore).

In vitro translation in HeLa lysate

In vitro translation using HeLa lysate driven from HeLa cells was performed as previously described (27, 58). Briefly, reactions were conducted with mRNAs at a final concentration of 10 nM and incubated at 32 $^{\circ}$ C for 2 h. In the m⁷GpppG competition assay, m⁷GpppG was added at a final concentration of 250 μ M. Samples were incubated at 32 $^{\circ}$ C for 2 h and then placed on ice to terminate the reaction. The samples were analyzed using the same methods as those applied for *in vitro* translation in RRL.

Cell culture and transfections

HeLa cells were cultured in Dulbecco's Modified Eagle's medium (DMEM; Nacalai Tesque) supplemented with 2 mM L-glutamine, 100 U/ml penicillin, 0.1 mg/ml

streptomycin (Sigma-Aldrich), and 10% (v/v) fetal bovine serum (FBS; Thermo Fisher Scientific). For NOP56-RAN luminescence assays, HeLa cells were seeded in 96-well plates at a density of 2×10^4 cells per well. After 24 h, cells were transfected with 50 ng Fluc control plasmid, 50 ng NOP56-RAN Nluc reporter plasmid, and 0.3 μ l FuGENE HD (Promega) in 20 μ l opti-MEM (Invitrogen) per well. After 24 h, cells were lysed in 50 μ l Glo lysis buffer (Promega). For luminescence assays, 30 μ l of lysate was incubated with 30 μ l ONE Glo EX Fluc substrate in a 1:1 ratio for 3 min in the dark with shaking. Fluc luminescence was measured using a Varioskan LUX Multimode Microplate Reader (Thermo Fisher Scientific). Following Fluc measurement, NanoGlo substrate diluted 1:100 in NanoGlo buffer (Promega) was added to the same plate, and Nluc activity was measured after shaking. For stress induction, HeLa cells were transfected with plasmids for 19 h, followed by treatment with 1 μ M thapsigargin (Thermo Fisher Scientific, T7458) for 5 h. For Western blotting of RAN reporters, HeLa cells were seeded in 12-well plates at a density of 1×10^5 cells per well. After 24 h, cells were transfected with 500 ng NOP56-RAN reporter plasmid using 1.4 μ l FuGENE HD (Promega) and 50 μ l OptiMEM (Invitrogen) per well. After 24 h, cells were harvested and lysed in 100 μ l RIPA buffer (50 mM Tris-HCl pH 6.8, 150 mM NaCl, 0.1% (w/v) SDC, 0.1% (w/v) SDS, 1.0% (w/v) NP-40). Lysates (70 μ l) were mixed with 30 μ l 4 \times SDS sample buffer and heated at 70 $^{\circ}$ C for 15 min. Samples were analyzed by Western blotting using the same procedure as for *in vitro* translation. For eIF2 α and p-eIF2 α detection, after separation, proteins were transferred onto PVDF membranes and blocked with 5% (w/v) bovine serum albumin (BSA) (Sigma-Aldrich) in TBS-T and incubated for 1 h with primary antibodies listed in Table S3. Membranes were washed three times with TBS-T for 5 min each and incubated with HRP-conjugated anti-Rabbit IgG (Sigma-Aldrich). Each experiment was performed in triplicate.

Immunoprecipitation

100 μ l *in vitro* translation reactions were performed using 10 nM mRNA in the RRL (Promega) at 30 $^{\circ}$ C for 90 min, followed by termination on ice. The samples were cleared by centrifugation at 20,000 \times g for 10 min at 4 $^{\circ}$ C, and the supernatants were transferred into new tubes containing 400 μ l TBS (total volume 500 μ l). Each sample was incubated with 10 μ l pre-washed HA agarose beads (Pierce Anti-HA Agarose, 26,181) and rotated at 4 $^{\circ}$ C for 1 h. FLAG beads and HA beads were then washed three times with 500 μ l TBS. For LC-MS/MS analysis, beads were resuspended in 50 μ l Phase Transfer Surfactant (PTS) buffer (12 mM Sodium Deoxycholate (SDC), 12 mM Sodium N-lauroyl sarcosinate (SLS), 100 mM Tris-HCl, pH 9.0) and boiled at 95 $^{\circ}$ C for 5 min. The beads were cleared by column purification, and the flow-through was collected into a new tube. These samples were used for further analysis.

Preparation for LC-MS/MS analysis

After immunoprecipitation of the RRL samples using HA agarose beads and eluted with PTS buffer, LC-MS/MS preparation was performed as previously described (59). First, the samples were reduced by treatment with 10 mM dithiothreitol (DTT) at room temperature for 30 min, then alkylated with 50 mM iodoacetamide in the dark at room temperature for 20 min. The protein mixture was then diluted 5-fold with 50 mM ammonium bicarbonate (ABC). For digestion of the denatured proteins into peptide fragments, 0.25 µg of Trypsin/Lys-C Mix (Promega, V5072) was added and incubated at room temperature for 3 h. An additional 0.5 µg of Trypsin/Lys-C Mix was added and incubated at 37 °C overnight. After digestion, an equal volume of ethyl acetate and 0.5% trifluoroacetic acid (TFA) was added to the samples. The mixture was shaken vigorously for 2 min and centrifuged at 15,700×g for 2 min. The upper ethyl acetate layer was discarded, and the solvent was removed using a centrifugal evaporator. The residual pellet was redissolved in 80 µl of 0.1% TFA and 2% acetonitrile and desalted as follows: the solution was applied to a GL-Tip SDB (GL Sciences), equilibrated with 0.1% TFA and 2% acetonitrile, washed with 0.1% TFA and 2% acetonitrile, and eluted with 0.1% TFA and 80% acetonitrile. The solvent was removed by centrifugal evaporator, and the residual peptides were redissolved in 20 µl of 0.1% TFA and 2% acetonitrile. The solution was centrifuged at 20,000×g for 5 min, and 18 µl of the supernatant was collected into a new tube for LC-MS/MS analysis.

LC-MS/MS

LC-MS/MS measurements were performed using an Easy-nLC1000 nanoflow liquid chromatography system and a Q-Exactive tandem mass spectrometer equipped with a nano-ESI ion source (Thermo Fisher Scientific). The trap column was a 2 cm × 75 µm capillary column packed with 3 µm C18-silica particles (Nikkyo Technos, Japan). The flow rate was set to 300 nl/min. Separation was conducted using a 10 to 40% linear acetonitrile gradient over 30 min in the presence of 0.1% formic acid. MS/MS data were acquired in data-dependent acquisition (DDA) mode, controlled by the Xcalibur 4.0 program (Thermo Fisher Scientific). The DDA settings were as follows: the resolution was 70,000 for the full MS scan and 17,500 for the MS2 scan; the AGC target was 3.0×10^6 for the full MS scan and 5.0×10^5 for the MS2 scan; the maximum IT was 60 ms for both the full MS and MS2 scans; the full MS scan range was m/z 310 to 1500, and the top 10 signals in each full MS scan were selected for the MS2 scan. DDA measurements were performed twice for each sample as technical replicates.

Protein identification

The MS/MS data were analyzed using the Proteome Discoverer 2.4 software bundled with the Sequest HT search engine (Thermo Fisher Scientific) to obtain the MS/MS spectra and peptide search parameters. MS2 spectra were searched against all *Oryctolagus cuniculus* ORF sequences

obtained from the UniProt database (<https://www.uniprot.org/taxonomy/9986>, downloaded on 20/3/2023) and predicted N-terminal candidate sequences listed in Table S4. The search parameters were as follows: MS1 and MS2 tolerance were set to 10 ppm and 0.2 Da, respectively; carbamidomethylation of cysteines (57.02146 Da) was considered a static modification, while acetylation of the protein N-terminus and oxidation of methionine (15.9949 Da) were dynamic modifications. The Percolator algorithm was used to determine the false discovery rate (FDR), and proteins/peptides with $\leq 1\%$ FDR were retained for further analysis. MS/MS spectra assigned to the peptides of interest were manually examined (Table S4).

Quantification and statistical analysis

All statistical analyses were performed using custom R code. The presented quantitative data represent the mean \pm SD from a minimum of three or six independent experiments. For comparison of NLuc reporter luciferase activity, one-way ANOVA and Student's *t* test were performed to confirm statistical difference between control and experimental groups.

Data availability

Raw data files are available in the Mendeley Data repository (<https://doi.org/10.17632/kc64gxd8r3.1>).

Supporting information—This article contains supporting information.

Acknowledgments—We thank the Center for Integrative Biosciences at Science Tokyo for DNA sequencing and the Cell Biology Center Research Core Facility at Science Tokyo for the Q-Exactive mass spectrometry measurements.

Author contributions—M. H., H. Ito, H. Imataka, K. M., T. T., and Y. N., resources; H. T., M. H., H. Ito., and K. M. writing—review & editing; H. T., M. H., H. Ito., and T. N. writing—original draft; H. T. supervision; H. T. project administration; H. T., and Y. N. funding acquisition; H. T., M. H., and H. Ito., T. N. data curation; H. T., M. H., and H. I. conceptualization. M. H. visualization; M. H., H. Ito., and T. N. methodology; M. H. and H. Ito. investigation.

Funding and additional information—This work was supported by MEXT Grants-in-Aid for Scientific Research (grant numbers JP26116002, JP18H03984, JP21H04763, and JP20H05925 to HT, JP20H05927 to YN), AMED-CREST: under Grant Number JP21gm1410008 (to H. T.), JST, the establishment of university fellowships towards the creation of science technology innovation and JSPS KAKENHI (grant number JPMJFS2112 and 24KJ1067 to H Ito), Uehara Memorial Foundation (to H. T.), Mitsubishi Foundation (to H. T.), and Daiichi Sankyo (to H. T.).

Conflict of interest—The authors declare that they have no conflict of interests with the contents of this article.

Abbreviations—The abbreviations used are: AW, alanine-tryptophan; DPR, dipeptide repeat; Fluc, Firefly Luciferase; GL, glycine-leucine; GP, glycine-proline; NLuc, Nano Luciferase; RAN

translation, Repeat Associated Non-AUG translation; RRL, rabbit reticulocyte lysate; SCA, Spinocerebellar ataxia; TG, Thapsigargin.

References

- Kobayashi, H., Abe, K., Matsuura, T., Ikeda, Y., Hitomi, T., Akechi, Y., *et al.* (2011) Expansion of intronic GGCCCTG hexanucleotide repeat in NOP56 causes SCA36, a type of spinocerebellar ataxia accompanied by motor neuron involvement. *Am. J. Hum. Genet.* **89**, 121–130
- García-Murias, M., Quintáns, B., Arias, M., Seixas, A. I., Cacheiro, P., Tarrio, R., *et al.* (2012) “Costa da Morte” ataxia is spinocerebellar ataxia 36: Clinical and genetic characterization. *Brain* **135**, 1423–1435
- Valera, J. M., Diaz, T., Petty, L. E., Quintáns, B., Yáñez, Z., Boerwinkle, E., *et al.* (2017) Prevalence of spinocerebellar ataxia 36 in a US population. *Neurol. Genet.* **3**, e174
- Lopez, S., and He, F. (2022) Spinocerebellar ataxia 36: from mutations toward therapies. *Front Genet.* **13**, 837690
- Obayashi, M., Stevanin, G., Synofzik, M., Monin, M. L., Duyckaerts, C., Sato, N., *et al.* (2015) Spinocerebellar ataxia type 36 exists in diverse populations and can be caused by a short hexanucleotide GGCCCTG repeat expansion. *J. Neurol. Neurosurg. Psychiatr.* **86**, 986–995
- Ohta, Y., Hayashi, T., Nagai, M., Okamoto, M., Nagotani, S., Nagano, I., *et al.* (2007) Two cases of spinocerebellar ataxia accompanied by involvement of the skeletal motor neuron system and bulbar palsy. *Intern. Med.* **46**, 751–756
- Zou, J., Wang, F., Gong, Z., Wang, R., Chen, S., Zhang, H., *et al.* (2023) A Chinese SCA36 pedigree analysis of NOP56 expansion region based on long-read sequencing. *Front Genet.* **14**, 1110307
- McEachin, Z. T., Gendron, T. F., Raj, N., García-Murias, M., Banerjee, A., Purcell, R. H., *et al.* (2020) Chimeric peptide species contribute to divergent dipeptide repeat pathology in c9ALS/FTD and SCA36. *Neuron* **107**, 292–305
- Hirayanagi, K., Ozaki, H., Tsukagoshi, S., Furuta, N., and Ikeda, Y. (2021) Porphyrins ameliorate spinocerebellar ataxia type 36 GGCCCTG repeat expansion-mediated cytotoxicity. *Neurosci. Res.* **171**, 92–102
- Todd, T. W., McEachin, Z. T., Chew, J., Burch, A. R., Jansen-West, K., Tong, J., *et al.* (2020) Hexanucleotide repeat expansions in c9FTD/ALS and SCA36 confer selective patterns of neurodegeneration in vivo. *Cell Rep.* **31**, 107616
- Matsuzono, K., Imamura, K., Murakami, N., Tsukita, K., Yamamoto, T., Izumi, Y., *et al.* (2017) Antisense oligonucleotides reduce RNA foci in spinocerebellar ataxia 36 patient iPSCs. *Mol. Ther. Nucleic Acids* **8**, 211–219
- Zhang, Y., Roland, C., and Sagui, C. (2018) Structural and dynamical characterization of DNA and RNA quadruplexes obtained from the GGGGCC and GGGCCT hexanucleotide repeats associated with C9FTD/ALS and SCA36 diseases. *ACS Chem. Neurosci.* **9**, 1104–1117
- Cleary, J. D., and Ranum, L. P. (2017) New developments in RAN translation: insights from multiple diseases. *Curr. Opin. Genet. Dev.* **44**, 125–134
- Wright, S. E., and Todd, P. K. (2023) Native functions of short tandem repeats. *Elife* **12**, e84043
- Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., *et al.* (2011) Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A* **108**, 260–265
- Castelli, L. M., Huang, W. P., Lin, Y. H., Chang, K. Y., and Hautbergue, G. M. (2021) Mechanisms of repeat-associated non-AUG translation in neurological microsatellite expansion disorders. *Biochem. Soc. Trans.* **49**, 775–792
- Liu, H., Lu, Y. N., Paul, T., Periz, G., Banco, M. T., Ferré-D’Amaré, A. R., *et al.* (2021) A helicase unwinds hexanucleotide repeat RNA G-quadruplexes and facilitates repeat-associated non-AUG translation. *J. Am. Chem. Soc.* **143**, 7368–7379
- Kearse, M. G., Green, K. M., Krans, A., Rodriguez, C. M., Linsalata, A. E., Goldstrohm, A. C., *et al.* (2016) CGG repeat-associated non-AUG translation utilizes a cap-dependent scanning mechanism of initiation to produce toxic proteins. *Mol. Cell.* **62**, 314–322
- Tabet, R., Schaeffer, L., Freyermuth, F., Jambeau, M., Workman, M., Lee, C. Z., *et al.* (2018) CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nat. Commun.* **9**, 152
- Green, K. M., Glineburg, M. R., Kearse, M. G., Flores, B. N., Linsalata, A. E., Fedak, S. J., *et al.* (2017) RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nat. Commun.* **8**, 2005
- Boivin, M., Pfister, V., Gaucherot, A., Ruffenach, F., Negroni, L., Sellier, C., *et al.* (2020) Reduced autophagy upon C9ORF72 loss synergizes with dipeptide repeat protein toxicity in G4C2 repeat expansion disorders. *EMBO. J.* **39**, e100574
- Zhang, Y., Glineburg, M. R., Basrur, V., Conlon, K., Wright, S. E., Krans, A., *et al.* (2022) Mechanistic convergence across initiation sites for RAN translation in fragile X associated tremor ataxia syndrome. *Hum. Mol. Genet.* **31**, 2317–2332
- Van ’t Spijker, H. M., Stackpole, E. E., Almeida, S., Katsara, O., Liu, B., Shen, K., *et al.* (2022) Ribosome profiling reveals novel regulation of C9ORF72 GGGGCC repeat-containing RNA translation. *RNA* **28**, 123–138
- Jazurek-Ciesiolka, M., Ciesiolka, A., Komur, A. A., Urbanek-Trzeciak, M. O., Krzyzosiak, W. J., and Fiszler, A. (2020) RAN translation of the expanded CAG repeats in the SCA3 disease context. *J. Mol. Biol.* **432**, 166699
- Cheng, W., Wang, S., Mestre, A. A., Fu, C., Makarem, A., Xian, F., *et al.* (2018) C9ORF72 GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2 α phosphorylation. *Nat Commun.* **9**, 15
- Sonobe, Y., Ghadge, G., Masaki, K., Sandoel, A., Fuchs, E., and Roos, R. P. (2018) Translation of dipeptide repeat proteins from the C9ORF72 expanded repeat is associated with cellular stress. *Neurobiol. Dis.* **116**, 155–165
- Ito, H., Machida, K., Hasumi, M., Ueyama, M., Nagai, Y., Imataka, H., *et al.* (2023) Reconstitution of C9orf72 GGGGCC repeat-associated non-AUG translation with purified human translation factors. *Sci. Rep.* **13**, 22826
- Latallo, M. J., Wang, S., Dong, D., Nelson, B., Livingston, N. M., Wu, R., *et al.* (2023) Single-molecule imaging reveals distinct elongation and frameshifting dynamics between frames of expanded RNA repeats in C9ORF72-ALS/FTD. *Nat Commun.* **14**, 5581
- Wright, S. E., Rodriguez, C. M., Monroe, J., Xing, J., Krans, A., Flores, B. N., *et al.* (2022) CGG repeats trigger translational frameshifts that generate aggregation-prone chimeric proteins. *Nucleic. Acids. Res.* **50**, 8674–8689
- Liu, Z., Chen, O., Wall, J. B. J., Zheng, M., Zhou, Y., Wang, L., *et al.* (2017) Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep* **19**, 2193
- Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., *et al.* (2013) RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. U S A* **110**, E4968–E4977
- Sellier, C., Buijsen, R. A. M., He, F., Natla, S., Jung, L., Tropel, P., *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
- Bañez-Coronel, M., Porta, S., Kagerbauer, B., Mateu-Huertas, E., Pantano, L., Ferrer, I., *et al.* (2012) A pathogenic mechanism in huntington’s disease involves small CAG-repeated RNAs with neurotoxic activity. *PLoS Genet.* **8**, e1002481
- Yu, C. H., Teulade-Fichou, M. P., and Olsthoorn, R. C. L. (2014) Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic. Acids. Res.* **42**, 1887–1892
- Atkins, J. F., Gesteland, F. F., Ft Reid, B., and Anderson, C. W. (1979) Normal tRNAs promote ribosomal frameshifting. *Cell* **18**, 1119–1131
- Maehigashi, T., Dunkle, J. A., Miles, S. J., and Dunham, C. M. (2014) Structural insights into +1 frameshifting promoted by expanded or modification-deficient anticodon stem loops. *Proc Natl Acad Sci U S A* **111**, 12740–12745

37. Kozak, M. (1986) Point mutations Define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283–292
38. Kozak, M. (1989) Circumstances and mechanisms of inhibition of translation by secondary structure in eucaryotic mRNAs. *Mol. Cell. Biol.* **9**, 5134–5142
39. Kozak, M. (1990) Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. U S A* **87**, 8301–8305
40. Westergard, T., McAvoy, K., Russell, K., Wen, X., Pang, Y., Morris, B., *et al.* (2019) Repeat-associated non- AUG translation in C9orf72- ALS/FTD is driven by neuronal excitation and stress. *EMBO Mol. Med.* **11**, e9423
41. Kashiwagi, K., Yokoyama, T., Nishimoto, M., Takahashi, M., Sakamoto, A., Yonemochi, M., *et al.* (2019) Structural basis for eIF2B inhibition in integrated stress response. *Science* **364**, 495–499
42. Sidrauski, C., Mcgeachy, A. M., Ingolia, N. T., and Walter, P. (2015) The small molecule ISRIB reverses the effects of eIF2 α phosphorylation on translation and stress granule assembly. *Elife* **4**, e05033
43. Andreev, D. E., Bf O'connor, P., Fahey, C., Kenny, E. M., Terenin, I. M., Dmitriev, S. E., *et al.* (2015) Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *Elife* **4**, e03971
44. Spriggs, K. A., Bushell, M., and Willis, A. E. (2010) Translational regulation of gene expression during conditions of cell stress. *Mol. Cell.* **40**, 228–237
45. Pakos-Zebrucka, K., Koryga, I., Mnich, K., Lujic, M., Samali, A., and Gorman, A. M. (2016) The integrated stress response. *EMBO. Rep.* **17**, 1374–1395
46. Lawrence, R. E., Shoemaker, S. R., Deal, A., Sangwan, S., Anand, A. A., Wang, L., *et al.* (2024) A helical fulcrum in eIF2B coordinates allosteric regulation of stress signaling. *Nat. Chem. Biol.* **20**, 422–431
47. Oliveira, M. M., Lourenco, M. V., Longo, F., Kasica, N. P., Yang, W., Ureta, G., *et al.* (2021) Correction of eIF2-dependent defects in brain protein synthesis, synaptic plasticity, and memory in mouse models of Alzheimer's disease. *Sci Signal* **14**, eabc5429
48. Mukhopadhyay, S., Amodeo, M. E., and Lee, A. S. Y. (2023) eIF3d controls the persistent integrated stress response. *Mol. Cell.* **83**, 3303–3313
49. Kozak, M. (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**, 1–34
50. Meydan, S., Klepacki, D., Karthikeyan, S., Margus, T., Thomas, P., Jones, J. E., *et al.* (2017) Programmed ribosomal frameshifting generates a copper transporter and a copper chaperone from the same gene. *Mol. Cell.* **65**, 207–219
51. Chamorro, M., Parkin, N., and Varmus, H. E. (1992) An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger. *RNA* **89**, 713–717
52. Saffert, P., Adaml, F., Schieweck, R., Atkins, J. F., and Ignatova, Z. (2016) An expanded CAG repeat in huntingtin causes +1 frameshifting. *J. Biol. Chem.* **291**, 18505–18513
53. Wang, X., Xuan, Y., Han, Y., Ding, X., Ye, K., Yang, F., *et al.* (2019) Regulation of HIV-1 gag-pol expression by shiftless, an inhibitor of programmed -1 ribosomal frameshifting. *Cell* **176**, 625–635.e14
54. Giedroc, D. P., and Cornish, P. V. (2009) Frameshifting RNA pseudoknots: structure and mechanism. *Virus. Res.* **139**, 193–208
55. Machida, K., Kanzawa, K., Shigeta, T., Yamamoto, Y., Tsumoto, K., and Imataka, H. (2018) Huntingtin polyglutamine-dependent protein aggregation in reconstituted cells. *ACS Synth. Biol* **7**, 377–383
56. Machida, K., Mikami, S., Masutani, M., Mishima, K., Kobayashi, T., and Imataka, H. (2014) A translation system reconstituted with human factors proves that processing of encephalomyocarditis virus proteins 2A and 2B occurs in the elongation phase of translation without eukaryotic release factors. *J. Biol. Chem.* **289**, 31960–31971
57. Fujino, Y., Ueyama, M., Ishiguro, T., Ozawa, D., Ito, H., Sugiki, T., *et al.* (2023) FUS regulates RAN translation through modulating the G-quadruplex structure of GGGGCC repeat RNA in C9orf72-linked ALS/FTD. *Elife* **12**, RP84338
58. Mikami, S., Kobayashi, T., Masutani, M., Yokoyama, S., and Imataka, H. (2008) A human cell-derived in vitro coupled transcription/translation system optimized for production of recombinant proteins. *Protein. Expr. Purif.* **62**, 190–198
59. Masuda, T., Saito, N., Tomita, M., and Ishihama, Y. (2009) Unbiased quantitation of Escherichia coli membrane proteome using phase transfer surfactants. *Mol. Cell. Proteomics.* **8**, 2770–2777