Doctoral Dissertation

Short repeat RNA-mediated suppression of the aggregation and cytotoxicity of TDP-43 and its C-terminal fragment TDP25

(TDP-43 及びそのカルボキシ末端断片 TDP25 の凝集およびその細胞毒性の抑制に働く短リピート RNA の研究)

Ai Fujimoto

Laboratory of Cellular and Molecular Sciences
Graduate School of Life Science
Hokkaido University

2025 March

Contents

Abbreviation

cDNA	Complementary deoxyribonucleic acid
D-MEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
FBS	Fetal bovine serum
HCl	hydrogen chloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hnRNP	Heterogeneous nuclear ribonucleoprotein (particle)
HSP	Heat shock protein
KCl	Potassium chloride
LSM	Laser scanning microscope
MALDI-TOF-MS	Matrix-assisted laser desorption ionization-Time of flight -mass spectrometry
$MgCl_2$	Magnesium chloride
miRNA	micro RNA
mRNA	messenger RNA
ncRNA	non-coding RNA
NTP	Nucleoside triphosphate
Proteostasis	protein homeostasis
RNA	Ribonucleic Acid
RNA-seq	RNA-Sequence
RRM	RNA-recognition motif
SDS	Sodium lauryl sulfate

siRNA	small interfering RNA
TAR	Trans activation responsive region
Tris	Tris(hydroxymethyl)aminomethane

Abstract

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that begins with the loss of motor function and eventually leads to respiratory failure, caused by the accumulation of protein aggregates that specifically accumulate in motor neurons, resulting in the degeneration and loss of motor neurons[1]. The accumulation of TDP-43 aggregates and its 25 kDa carboxy-terminal fragment (TDP25; 220–414 amino acids) in motor neurons is thought to be the cause of ALS[2,3]. TDP-43 has two RNA/DNA recognition motifs (RRM1 and RRM2) and a C-terminal intrinsically disordered region (IDR; 274–414 amino acids) that may be the core of amyloid formation[4–6]. This IDR is known to be involved in amyloid formation and contributes to the interaction with other RNA-binding proteins and the regulation of TDP-43 self-aggregation via liquid-liquid phase separation[7,8]. On the other hand, TDP25 is known to be generated by protease cleavage and to have high aggregation properties[9]. TDP25 lacks the N-terminal 29 amino acids of RRM2, and since this deletion region includes the critical amino acids for RNA recognition, it lacks typical RNA binding ability. However, previous research in our laboratory has shown that RNA degradation in cell lysates promotes TDP25 aggregation[10]. From this, it was predicted that some RNA would inhibit the formation of TDP25 aggregates. However, it was unclear what kind of RNA sequence would inhibit the aggregation of TDP25 in this way. Since it is known that the UG repeat RNA recognized by TDP-43 does not interact with TDP25, we had the idea that repeat RNA other than UG repeats might be playing this role[10,11]. We then focused on the GGGGCC repeat sequence observed in the untranslated region of an ALS causative gene *C9orf72*,[12,13] and thought that this might be an RNA sequence that interacts with TDP25 and suppresses its aggregation.

To analyze the molecular interactions between the two proteins, I used fluorescence cross-correlation spectroscopy (FCCS),[14,15] which detects the interaction with a single molecule sensitivity. In cell lysates, it was confirmed that the fluorescently labeled GGGGCC tetra-repeat RNA ($r(G4C2)_4$) interacts with TDP-43,[16] and it was also found that it binds to TDP25. In addition, fluorescently labeled AAAAUU tetra-repeat RNA ($r(A4U2)_4$) also interacted with TDP-43 and TDP25. Next, we performed FCCS analysis of these RNAs using purified TDP-43 and TDP25 from mammalian cells. It was shown that TDP25 interacts directly with RNA without the involvement of other cellular proteins. Next, to investigate the biochemical interaction mechanism between TDP-43/TDP25 and RNA, we performed FCCS analysis of the interaction strength between TDP-43 or TDP25 and RNA were analyzed by FCCS, it was suggested that the RNA interaction of TDP-43 has a higher contribution from electrostatic interactions than from hydrophobic interactions, whereas the RNA interaction mode of TDP25 has a higher contribution from hydrophobic interactions than from electrostatic interactions. Furthermore, to investigate the interaction region between TDP25 and RNA, we used the remaining RRM2 (220–263 amino acids) and IDR, which are defective mutants of TDP25, and found that they no longer interact with $r(G4C2)_4$ and $r(A4U2)_4$, suggesting that the entire TDP25 is necessary for the interaction.

To demonstrate whether these RNAs have a function that inhibits the aggregation of TDP25, we introduced plasmid DNA into mouse neuroblastoma Neuro2a cells to express $r(G4C2)_4$ and $r(A4U2)_4$, the proportion of cells with cytoplasmic aggregates of TDP25 decreased in the cell groups expressing $r(G4C2)_4$ and $r(A4U2)_4$, and cell death caused by TDP25 expression was suppressed by these RNA expression. Next, the directly introduced synthetic $r(G4C2)_4$ and $r(A4U2)_4$ sequences decreased the

proportion of cells with cytoplasmic aggregates of TDP25. Furthermore, transcriptome analysis using next-generation sequencing revealed that there was no induction of known genes that have a protein aggregation inhibitory effect, such as molecular chaperones, in the cell groups expressing $r(G4C2)_4$ and $r(A4U2)_4$. Next, to investigate whether $r(G4C2)_4$ and $r(A4U2)_4$ can inhibit the formation of aggregates of full-length TDP-43 and the cell death caused by its aggregates, we expressed $r(G4C2)_4$ or $r(A4U2)_4$ in human embryonic kidney 293 (HEK293) cells. The proportion of cells with cytoplasmic TDP-43 aggregates decreased, and cell death caused by TDP-43 expression was suppressed by these RNA expressions. Furthermore, as a result of transcriptome analysis using next-generation sequencing, it was confirmed that there was no induction of expression of known genes with a protein aggregation inhibitory effect, as was the case with TDP25 expression.

This study shows that $r(G4C2)_4$ and $r(A4U2)_4$ are RNA sequences that directly bind to TDP25 and TDP-43, suppressing aggregation and cell death. Although proteins with anti-aggregation functions, such as molecular chaperones, have been known, the fact that RNA also has anti-aggregation functions suggests that RNA may act as a molecular chaperone in the cell and contribute to the maintenance of proteostasis.

Keywords

amyotrophic lateral sclerosis, TDP-43, TDP25, protein aggregation, RNA, proteostasis, RNA-Protein interaction, FCCS

Introduction

Amyotrophic lateral sclerosis (ALS) is a refractory neurodegenerative disease characterized by progressive degeneration of upper and lower motor neurons, resulting in a decline in motor function throughout the body. Patients typically succumb to respiratory muscle paralysis within 3 to 5 years after diagnosis[1]. Currently, ALS remains an incurable disease with no treatments or methods available to achieve complete remission. Existing therapies primarily involve palliative measures to slow disease progression[17].

Over 90% of ALS cases are sporadic, with approximately 10% classified as familial ALS. While the exact causes of ALS remain unknown, protein aggregates have been observed in the brains of ALS patients and ALS model mice[18,19]. In 2006, TAR DNA binding protein 43kDa (TDP-43) was identified as a major constituent of these aggregates. Furthermore, it was found that TDP-43 formed in the cytoplasm undergoes phosphorylation and ubiquitination, and its C-terminal fragments (43CTFs) are present in these aggregates.

Subsequent analyses revealed that over 90% of ALS patients have TDP-43-positive aggregates[2,3]. TDP-43 was originally identified as a protein that binds to TAR (trans-acting responsive region), an RNA regulatory sequence element present in the terminal repeat regions of the human immunodeficiency virus type 1 (HIV-1) genome. TDP-43 is a polypeptide consisting of 414 amino acids with a molecular weight of 43kDa and includes two RNA/DNA recognition motifs (RRM1 and RRM2; 106–176 and 191–262 amino acids)[20]. Each of these domains contains two highly conserved sequence motifs essential for nucleic acid binding. At the carboxy terminus, TDP-43 has a glycine-rich region (GRR; 274–414 amino acids), which is known as a prion-like

domain[4]. This GRR includes subdomains enriched in glycine (274–310 amino acids), hydrophobic patches (318–343 amino acids), glutamine/asparagine-rich sequences (344–360 amino acids), glycine/serine-rich sequences (368–414 amino acids), and amyloid-forming cores[5,6]. The GRR is thus referred to as an intrinsically disordered region (IDR) and mediates TDP-43 self-assembly and interactions with other RNA-binding proteins (e.g., hnRNP A2/B1, hnRNP A1, FUS) through liquid-liquid phase separation[7,8]. Notably, most ALS-associated amino acid substitutions in TDP-43 are localized within this IDR[21].

Under normal conditions, TDP-43 predominantly resides in the nucleus but functions as a nuclear-cytoplasmic shuttling protein, maintaining cellular homeostasis by regulating the quality of RNA, including mRNA and small RNA[11,22]. TDP-43 contains canonical DEVD-like motifs at amino acid positions 86–89 and 216–219, which are cleaved by caspase-3, generating the TDP-43 C-terminal fragments TDP35 (90–414 amino acids) and TDP25 (220–414 amino acids)[23]. Other proteases, such as calpain, selective splicing, and protein clearance pathways, are also involved in generating these CTFs[5,24,25]. TDP25 lacks critical amino acids in RRM1 and RRM2 required for RNA binding, making it highly prone to aggregation and capable of inducing cell death[9]. TDP25 aggregates sequester full-length TDP-43, suggesting that both toxic gain-of-function from these aggregates and the loss-of-function of TDP-43 contribute to cytotoxicity [24,26,27]

The mislocalization of TDP-43 contributes to cytotoxicity in ALS through the dysregulation of cellular processes, including gene expression, mRNA maturation, repression of alternative splicing, cryptic exon splicing, and alternative polyadenylation[27,28]. Persistent cytoplasmic mislocalization of TDP-43 results in

aggregate formation and disruption of physiological functions, such as the sequestration of ALS-associated and autophagy-regulating proteins, leading to a decline in proteostasis[29,30]. Consequently, motor neuron degeneration may be associated with TDP-43 mislocalization and the formation of parallel aggregates[26].

Chaperone molecules are known to prevent protein aggregation. Chaperones, a family of proteins, not only assist in protein folding but also recognize and refold partially misfolded proteins to their native state[31]. Certain chaperones can inhibit or disrupt the oligomers that serve as precursors to aggregates[31]. For example, the heat shock protein HSP70 has been shown to suppress TDP25 aggregation[32].

RNA-binding proteins (RBPs), such as TDP-43, interact with RNA and induce the condensation of RBPs through phase separation[33]. RNA degradation induces aggregation of both TDP-43 and TDP25, despite TDP25 lacking major RNA-binding sequences[10,34]. However, RNA binding by TDP-43 antagonizes aggregation and phase transitions[35,36]. Additionally, it has been proposed that certain RNAs, including ribosomal RNA and transfer RNA, possess chaperone-like activities[37–40]. These RNAs regulate the aggregation and condensation of RBPs through phase transitions and facilitate intermolecular assembly based on their physicochemical properties. Nevertheless, the specific RNA molecules that modulate TDP-43 aggregation and contribute to cytoprotection in ALS/FTD remain unclear.

Here, I demonstrate that GGGGCC and AAAAUU hexanucleotide repeat RNAs (r(G4C2) and r(A4U2)) can mitigate proteotoxicity by preventing the aggregation of TDP-43 and TDP25 in a cellular model. The C-terminal domain of TDP-43 is known to recognize the G-quadruplex (Gq) structure in RNA, and Gq DNA protects against protein aggregation[16]. Accordingly, I first focused on whether TDP25 binds to r(G4C2),

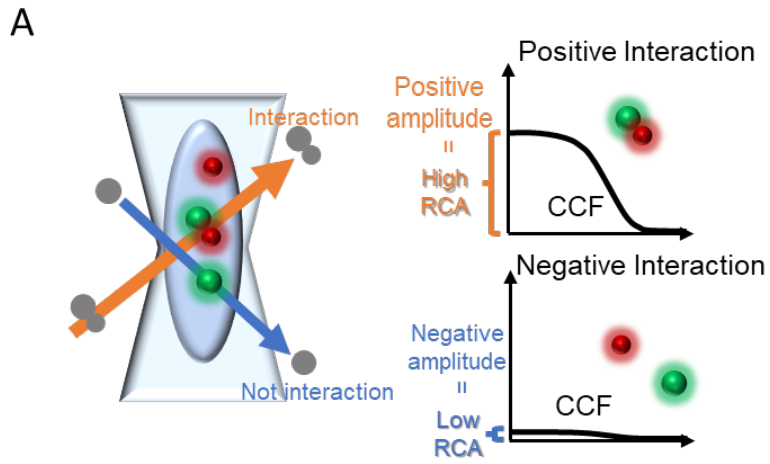
a hexanucleotide repeat RNA folded into a Gq structure and included in the ALS-causative C9orf72 transcript[41]. I then demonstrated that non-ALS-associated short r(G4C2) could have a cytoprotective role by suppressing TDP25 aggregation and mislocalized TDP-43 aggregates. Furthermore, I showed that r(A4U2), generally used as a control sequence for r(G4C2), also binds TDP-43 and TDP25 and suppresses their aggregation while reducing the cytotoxicity caused by their mislocalization and aggregation. These findings suggest that these repeat RNAs may possess chaperone-like activities, preventing protein aggregation and maintaining intracellular proteostasis.

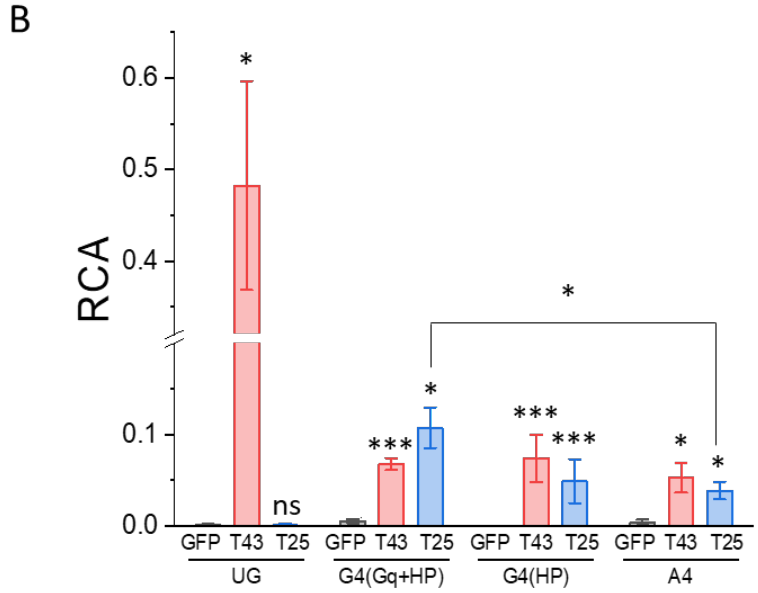
Results

Interaction of GGGGCC and AAAAUU repeat RNA with of TDP-43 and TDP25 in cell lysate.

To analyze the interaction between TDP-43/TDP25 and RNA, I employed a single molecule detection method, fluorescence cross-correlation spectroscopy (FCCS), which can quantitatively determine the strength of interaction between diffuse fluorescent species in solution from the fluorescence fluctuation that occurs by passing fluorescent molecules through the confocal detection volume (Figure 1A) [14,15]. After mixing RNA labeled with a chemical near-infrared fluorescent dye, Alexa Fluor 647 (AF), in lysates of murine neuroblastoma Neuro2a cells expressing green fluorescent protein (GFP)- tagged TDP25 (T25), GFP-tagged TDP-43 (T43), or GFP monomers as the tag, FCCS was performed. As I have previously shown [10], the relative cross-correlation amplitude (RCA) of the AF-labeled 24-mer of uracil guanine repeat RNA (UG), a consensus TDP-43-recognizing sequence, with T43 was high, whereas those with GFP and T25 were not observed (Figure 1B), indicating that the interaction between T25 and UG are none or quite low. The RCA when AF-labeled four repeats of r(G4C2) (G4) in a coexisting state of the hairpin (HP) and Gq structures (Gq + HP) [13,42]was mixed with both T43 and T25 was significantly positive; this suggests that G4 can interact with not only T43 but also T25. The RCA of T43 and T25 with only HP-formed G4 was also positive (Figure 1B), suggesting that the structure of Gq may not be required for the interaction with T43 and T25. Next, to investigate whether T25 also interacts with other RNAs that are not different from G4, I tested AF-labeled four repeats of r(A4U2) (A4), which is often used as a control sequence for r(G4C2)[12,13]. Although I had expected that T25 had not bound to A4, the RCA of T43 and T25 with

A4 was positive but lower than that with G4 in a mixed state of the hairpin (HP) and Gq structures (Figure 1B), suggesting that T43 and T25 can interact with G4 and A4, regardless of their steric structures. Furthermore, as UG did not bind to TDP25, the interaction between TDP25 and G4/A4 may be specific. Next, to determine whether the shorter repeats were adopted for this interaction, two repeats of r(G4C2) (G2) and r(A4U2) (A2) were used. Although the RCAs between T43/T25 and AF-labeled G2 were positive, the RCA between T43/T25 and AF-labeled A2 was not observed (Figure 1C). Therefore, four repeats of r(A4U2) may be required to interact with TDP-43/TDP25. Furthermore, the Gq structure of r(G4C2) is not necessarily required for this interaction.


A
Interaction
Not interaction
Positive Interaction
Positive amplitude = High RCA
CCF
Negative Interaction
Negative amplitude = Low RCA
CCF


B
RCA
0.6
0.5
0.4
0.1
0.0
GFP T43 T25 GFP T43 T25 GFP T43 T25 GFP T43 T25
UG
G4(Gq+HP)
G4(HP)
A4
ns
*
***

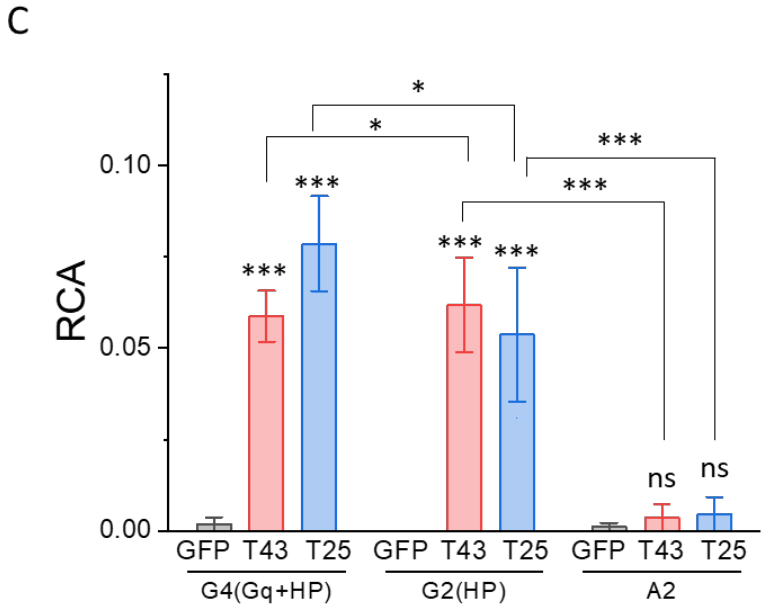

C
RCA
0.10
0.05
0.00
GFP T43 T25 GFP T43 T25 GFP T43 T25
G4(Gq+HP)
G2(HP)
A2
ns
*
***

Figure 1. Interaction of the GGGGCC/AAAAUU repeat RNA with TDP-43 and TDP25 in cell lysates using fluorescence cross-correlation spectroscopy.

(A) Illustration of the detection of molecular interactions from the relative cross-correlation amplitude (RCA) using fluorescence cross-correlation spectroscopy. (B) RCA values when AF-labeled RNAs (UG, G4, and A4) that were folded in 100 mM KCl (Gq + HP) or 100 mM LiCl (HP) were mixed in lysates of Neuro2a cells expressing GFP monomers (GFP), GFP TDP-43 (T43), and GFP-TDP25 (T25) with 100 mM KCl (n = 3; bars indicate independent trials and mean ± SE). (C)RCA values when AF-labeled RNAs (G4, G2, and A2) were used. G4 (Gq + HP) was folded in the presence of 100 mM KCl, and then mixed with the cell lysate including 100 mM KCl. G2 (HP) was folded in the presence of 100 mM LiCl, then mixed with the cell lysate including 100 mM KCl (n= 3; mean ± SE). (B & C) Significance: *$p < 0.05$, ***$p < 0.001$, and ns ($p \geq 0.05$).

Direct interaction between G4/A4 and TDP-43/TDP25

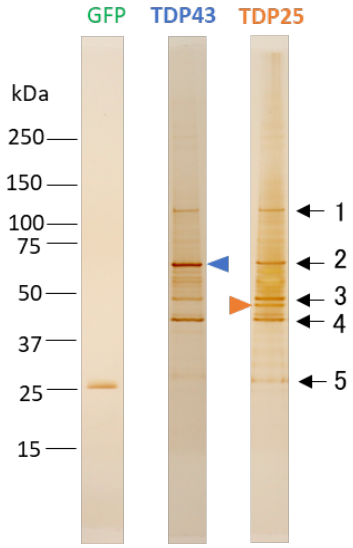
Thus, the interaction analyses between T43/T25 and RNA have been conducted in cell lysates. One of the advantages of FCCS is its ability to measure biomolecular interactions in solution mixtures containing fluorescent molecules. However, it remains unclear whether T25 interacts directly with RNA or if cellular RNA-binding proteins within complexes containing T25 interact with RNA. To confirm the direct interaction with these RNAs, purified GFP, T43, and T25 were prepared using tandem affinity purification with polyhistidine and strep tags. The purified samples were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by silver staining. High-purity GFP was confirmed; however, several cellular proteins were found to be included in the purified T43 and T25 samples (Figure 2a). T43 is a ribonucleoprotein complex (RNP) known to bind RNA sequences capable of forming G-quadruplexes (Gq)[16,43,] Therefore, peptide mass fingerprinting (PMF-MS) was used to identify whether T43 or other RNPs were present among the proteins co-purified with the purified T25. As a result, five main proteins co-purified with T25 were identified: 70 kDa heat shock protein (HSP70), mitochondrial pyruvate carboxylase (mPC), tubulin, and actin. Additionally, GFP-containing fragments were confirmed by Western blotting. Among these, bands corresponding to mPC, tubulin, and actin were also observed in the T43 sample (Figure 3A). However, since HSP70 and mPC have been reported to have RNA-binding properties,[44,45] FCCS was used to confirm that these co-purified cellular proteins did not bind to G4/A4. AF-G4 or AF-A4 was mixed with the lysate of cells expressing GFP- or yellow fluorescent protein (YFP)-tagged HSP70, β-tubulin, α-actin, or mPC, and FCCS analysis was performed under the same conditions as for T43/T25. No significant interactions with any of the proteins were

detected (Figure 3B-E). HSP70 is an abundant molecular chaperone and frequently co-precipitates with highly aggregation-prone proteins such as TDP25. Therefore, the interaction with HSP70 may contribute to maintaining the solubility of TDP25. In other words, it is likely difficult to recover TDP25 in the soluble fraction without HSP70. Furthermore, no endogenous TDP-43 band was observed in the purified T25 sample (Figure 3A), and no interaction between TDP25 and TDP-43 in the cell lysate was detected using FCCS.[10] Thus, the possibility of endogenous TDP-43 being included in the purified T25 sample is low. Additionally, as AF-UG did not bind to T25 in the cell lysate experiments (Figure 1B), it is suggested that G4/A4 directly recognize and bind to TDP25, rather than binding via TDP-43.

As a result, even if other RNPs were included in the purified T25 sample, they are unlikely to contribute to the interaction analysis with G4/A4. Next, the interactions of purified GFP, T43, and T25 with AF-G4/AF-A4 were analyzed using FCCS. When mixed with AF dye alone as a negative control, no positive relative cross-correlation (RCA) was observed for GFP, T43, or T25. However, when mixed with AF-G4 or AF-A4, positive RCA was observed for both T43 and T25, with the RCA being higher for AF-G4 than for AF-A4 (Figure 3F, G).

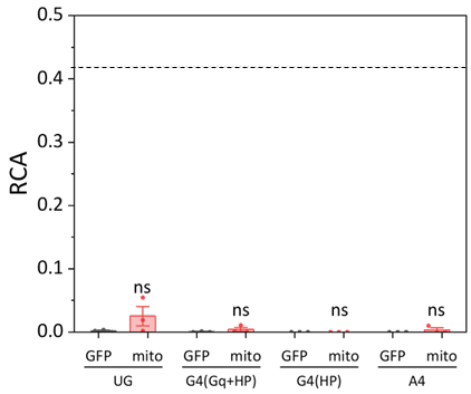
The dissociation constants (Kd) of these interactions calculated from the FCCS analysis were in the submicromolar to micromolar range (Figure 3H). These Kd values are considered reasonable for RNA-protein interactions in solution. Consequently, it was demonstrated that TDP-43 and TDP-25 interact directly with G4 and A4, independent of other RNPs.

A

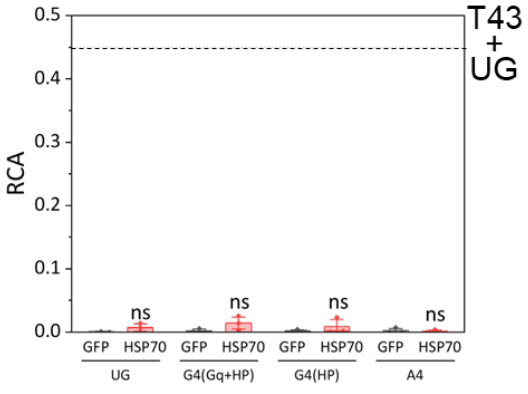


No.	Protein name
1	Mitochondrial pyruvate carboxylase
2	HSAPA8 (Hsp70)
3	α/β-tubulin
4	actin
5	fragment containing GFP

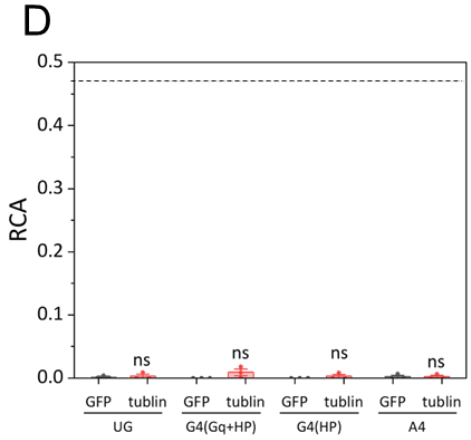
B



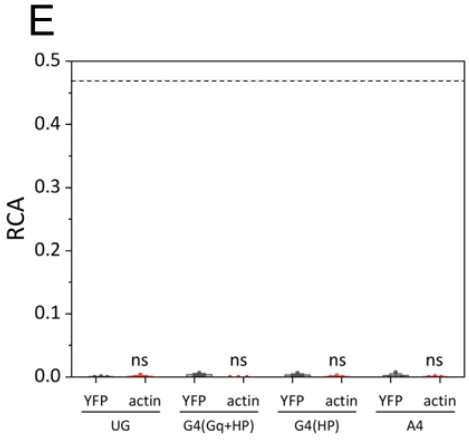
C



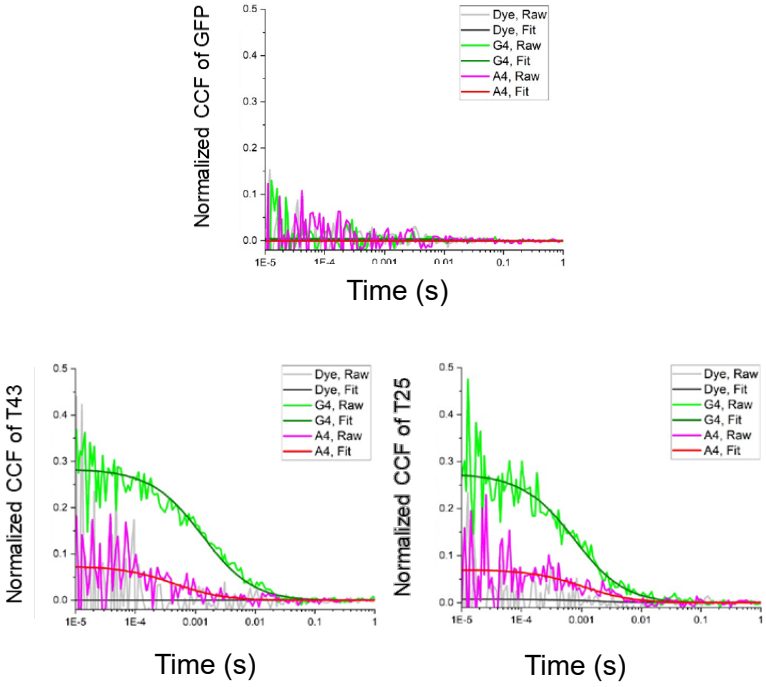
D



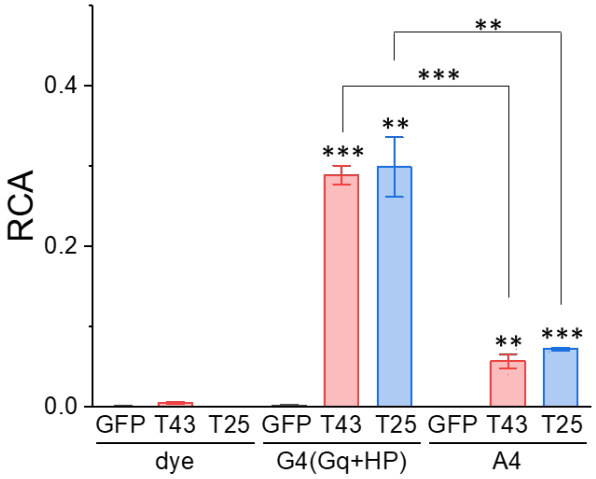
E



F



G



H

RNA	Protein	K_d (μM) (mean±SE ; n=3)
G4(Gq+HP)	T43	0.21 ± 0.02
G4(Gq+HP)	T25	0.16 ± 0.02
A4	T43	2.0 ± 0.7

Figure 2. Direct interaction of the G4/A4 RNA with purified TDP-43 and TDP-25.

SDS-PAGE and silver staining of purified GFP, GFP-TDP-43, and GFP-TDP-25 and attribution table of co-purified proteins. Blue or orange arrows indicate purified target proteins. Numbers indicate co-purified proteins. (B-E) The RCA value of G4/A4-RNA to the proteins that were co-purified with GFP-TDP43 or GFP-TDP25 (mitochondrial pyruvate carboxylase (mito) (B) HSP70, α-tubulin (tubulin) and β-actin (actin) (C), (D), and (E) was examined using fluorescence cross-correlation spectroscopy (FCCS). The relative cross-correlation amplitude (RCA) values when Alexa Fluor 647-labeled RNAs (UG, G4, and A4) that were folded in 100 mM KCl (Gq + HP) or 100 mM LiCl (HP) were mixed in lysates of Neuro2a cells expressing GFP- or YFP-tagged proteins with 100 mM KCl (n = 3; dots and bars indicate independent trials and mean ± SE). GFP or YFP monomers were used as binding negative controls. The dotted line indicates the mean RCA values when UG was mixed with GFP-TDP43 as a positive control. (F) Typical normalized cross-correlation functions (CCFs) when purified proteins were mixed with Alexa Fluor647-labeled RNA. The purified GFP monomers (left), GFP-TDP-43 (middle), and GFP-TDP25 (right) were mixed with Alexa Fluor647 (Dye) or AlexaFluor647-labeled G4 or A4. Raw CCFs (Raw) and fitted curves (Fit) were represented. (G) RCA values when AF-labeled RNAs (G4, and A4) folded in 100 mM KCl were mixed with purified GFP monomers (GFP), GFP-TDP-43 (T43), and GFP-TDP-25 (T25) with 100 mM KCl (n = 3; mean± SE).(H) Dissociation constant (*Kd*) values between T43/T25 (proteins) and AF-labeled G4/A4 (RNA) calculated from (F) (n = 3; mean ± SE). (B, C, D, E, G) Significance: **p < 0.01, ***p < 0.001, and ns (p ≥ 0.05).

Biochemical interaction mechanism between G4/A4 RNA and TDP-43/TDP25

Next, to demonstrate biochemical properties in the interaction between T43/T25 and G4/A4, the interaction strength in the presence of high concentrations of KCl (150, 200, and 250 mM) was analyzed using FCCS. The RCA of neither dual color-fluorescent DNA [46] as a positive control nor the mixture of the dual fluorescent dyes as a negative control changed at the high concentrations of KCl compared to that at 100 mM (Figure S1a), indicating that the FCCS detection system did not change at high KCl concentration. The RCA between T43 and UG dramatically declined with the increase in KCl concentration (cyan line in Figure 2A), suggesting the contribution of electrostatic interaction. The RCA between T25 and UG did not emerge in high KCl concentrations (grey line in Figure 2A). The RCAs between T43 and G4/A4 prominently declined with the increase in KCl concentration (green and purple lines in Figure 2A); however, those between T25 and G4/A4 declined little by little with the increase in KCl concentration (red and yellow lines in Figure 2A). Furthermore, at 150 and 200 mM KCl concentrations, the RCA between T25 and G4/A4 was still positive and higher than that between T43 and G4/A4 (Figure 2A), suggesting that both hydrophobic and electrostatic interactions may contribute significantly to the interactions between TDP25 and G4/A4; however, the electrostatic interaction may relatively highly contribute to that between TDP-43 and G4/A4 compared with TDP25. These trends were similarly observed with the addition of $MgCl_2$ (5, 10, 15 mM), which generates the divalent cation Mg^{2+} (Figure 2B and Figure S1b), and with the addition of heparin (0, 1 ng/μg) (Figure 2C). The divalent cation Mg^{2+} functioned effectively at concentrations 1/10 or lower compared to monovalent cation K^+,

indicating inhibition of charge-based electrostatic interactions. Therefore, the biochemical RNA interaction mechanism, especially the balance of electrostatic and hydrophobic interaction, may differ between TDP-43 and TDP25. On the other hand, the addition of 1,6-hexanediol (HD) (5% and 10%), which interferes with weak hydrophobic interactions, did not alter the RCA between T43 and UG. Similarly, the RCA between T43 and G4/A4 remained unchanged upon the addition of 5% HD. However, the RCA between T25 and G4/A4 significantly decreased when 5% HD was added. This result demonstrated an inverse correlation of RCA values compared with increasing concentrations of KCl and $MgCl_2$ (Figure 2D and Figure S1c). Circular dichroism analysis confirmed that no notable conformational changes in G4/A4 were observed with the increasing concentration of HD (Figure 2E). This suggests that the addition of HD does not affect the tertiary structure of RNA. In conclusion, the biochemical interaction mechanisms with RNA, particularly the balance between electrostatic and hydrophobic interactions, may differ between TDP-43 and TDP25.

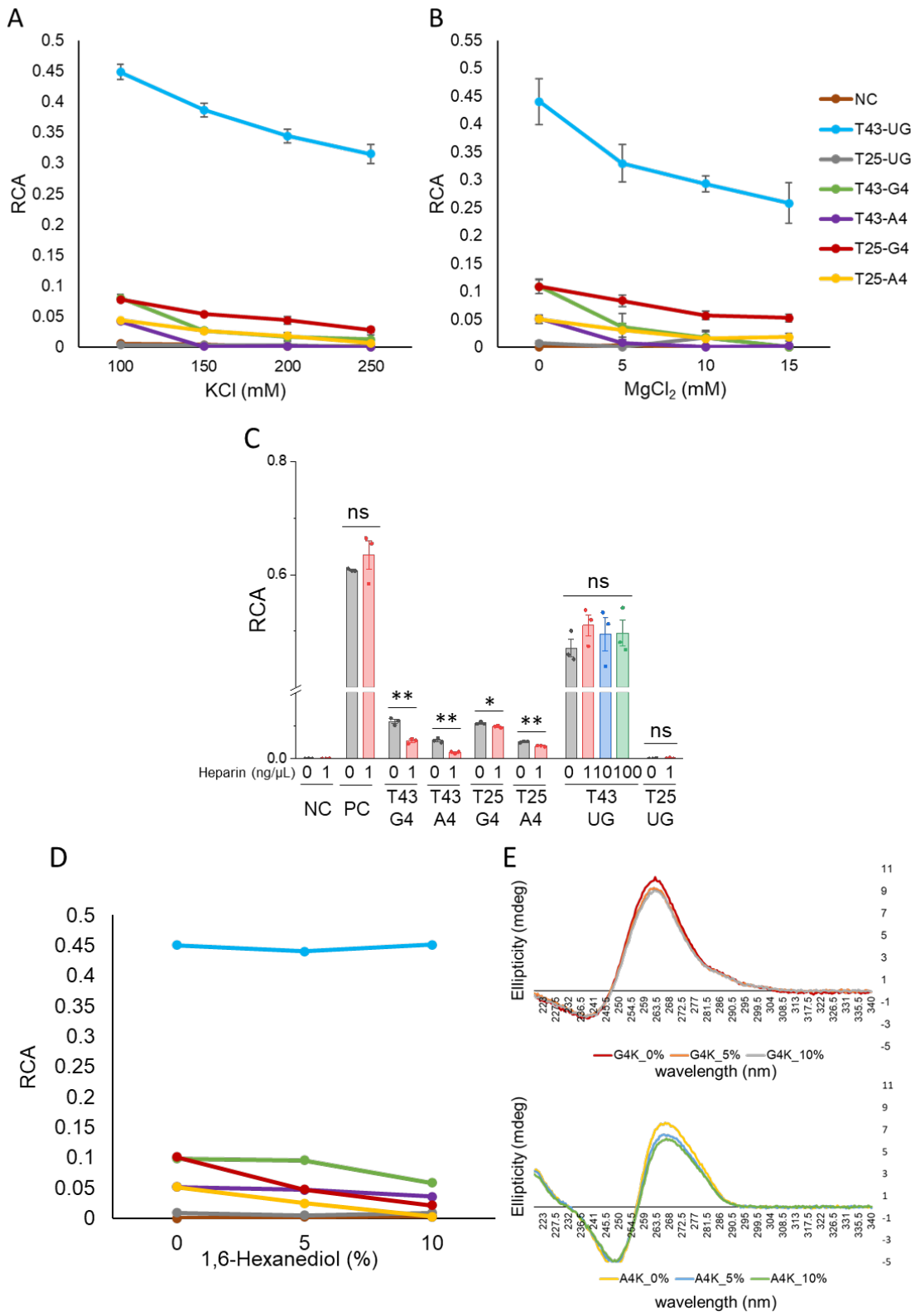

A
RCA
KCl (mM)
B
RCA
MgCl2 (mM)
NC
T43-UG
T25-UG
T43-G4
T43-A4
T25-G4
T25-A4
C
RCA
ns
**
*
Heparin (ng/μL)
NC
PC
T43 G4
T43 A4
T25 G4
T25 A4
T43 UG
T25 UG
D
RCA
1,6-Hexanediol (%)
E
Ellipticity (mdeg)
G4K_0%
G4K_5%
G4K_10%
wavelength (nm)
A4K_0%
A4K_5%
A4K_10%
wavelength (nm)

Figure 3. Biochemical Interaction mechanism of the G4/A4 RNA with TDP-43 and TDP25 in cell lysates using fluorescence cross-correlation spectroscopy.

(A) RCA values between AF-labeled RNAs (UG, G4, and A4) and GFP-tagged TDP-43/TDP25 in cell lysates at 100, 150, 200, and 250 mM KCl concentrations. G4 (Gq + HP) was folded in the presence of 100 mM KCl, and then mixed with the cell lysate. The dashed line indicates zero. (B) RCA values between AF-labeled RNAs (UG, G4, and A4) and GFP-tagged TDP-43/TDP25 in cell lysates at 0, 5, 10, and 15 mM MgCl2 concentrations. The dashed line indicates zero. (C) RCA of the mixture of GFP-tagged TDP-43/TDP25 and G4, A4, or UG RNA in the presence of represented concentration of heparin. (D) Relative cross-correlation amplitude (RCA) of PC and NC in the presence of 0, 5, and 10% 1,6-hexanediol (HD). The dashed line indicates zero. (E and (F) C ircular dichroism (CD) spectra of G4-RNA (E) and A4-RNA (F) with 100 mM KCl in the presence of 0, 5, and 10% HD. The gray line indicates zero (baseline). A slight decrease in the local maximum ellipticity but no significant changes in the conformation of G4/A4-RNA were observed upon the addition of HD. (n = 3; mean ± SE). (A, B, C, D) Significance: *$p < 0.05$,**$p < 0.01$, and ns ($p \geq 0.05$).

Both truncated RRM2 and IDR are responsible for the interaction between G4/A4 RNA and C-terminal fragments of TDP-43

How does T25 recognize RNA, such as G4 and A4, despite lacking most of its RNA recognition motifs (RRMs)? To elucidate the contribution of the truncated RRM2 and IDR in TDP-43 CTFs to their interaction with G4/A4, two additional truncated variants of TDP25 (F220 and C274 [9,44]; Figure 4A) were analyzed for their interaction with G4/A4 using FCCS in Neuro2a cell lysates. Neither F220 nor C274 interacted with a mixture of Gq and HP, HP-structured G4, or A4 (Figure 4B). This result suggests that both the truncated RRM2 and IDR are necessary for the interaction with G4/A4 and that they do so independently of highly conserved nucleic acid-binding sequences such as the RRM.

Moreover, previous studies have shown that TDP25 exists in an oligomeric and soluble state[9]. Based on this, it is thought that G4/A4 may recognize and bind to specific structures of the protein.

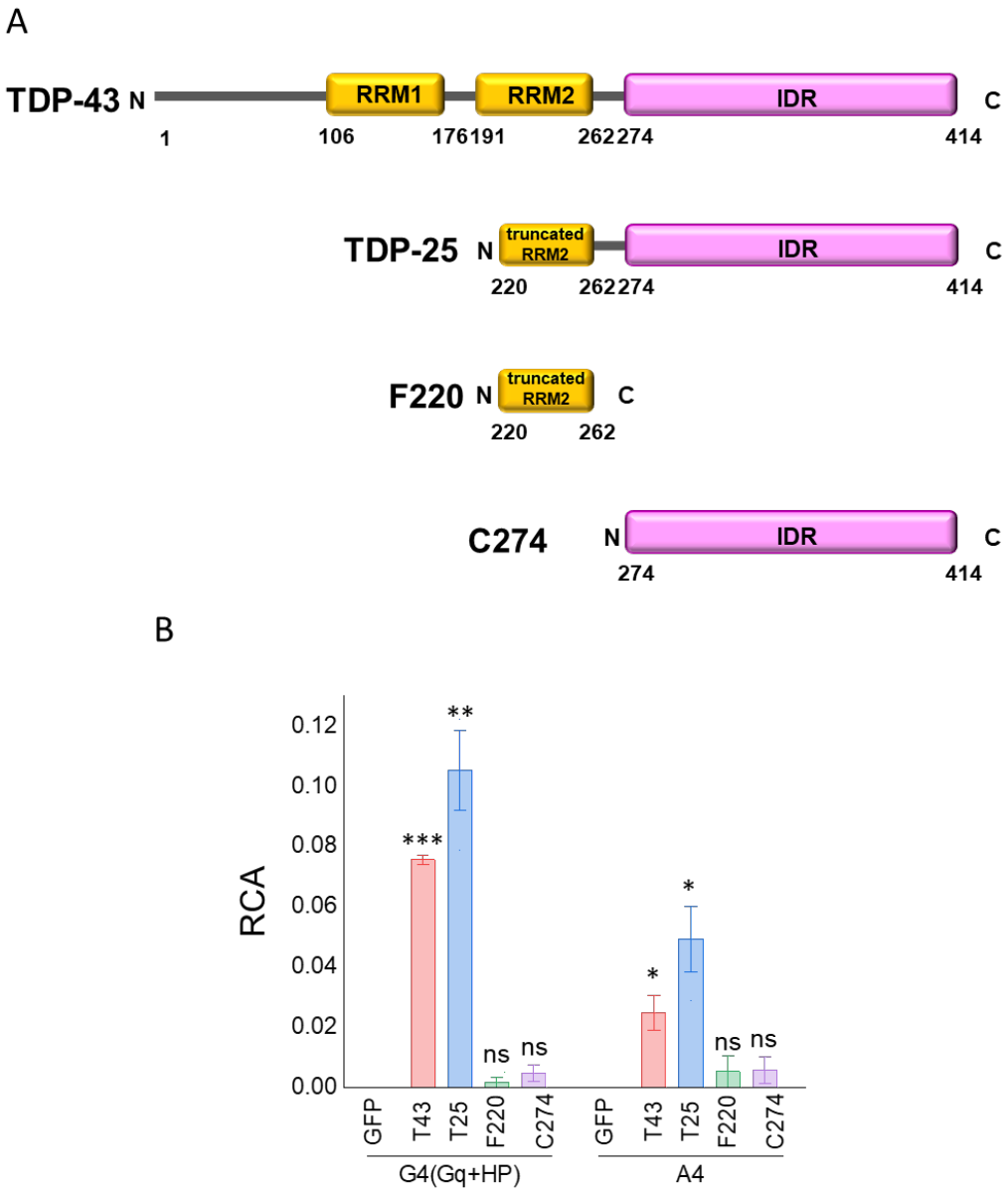


Figure 4. Both truncated RRM2 and IDR are required for interaction with G4/A4 RNA.

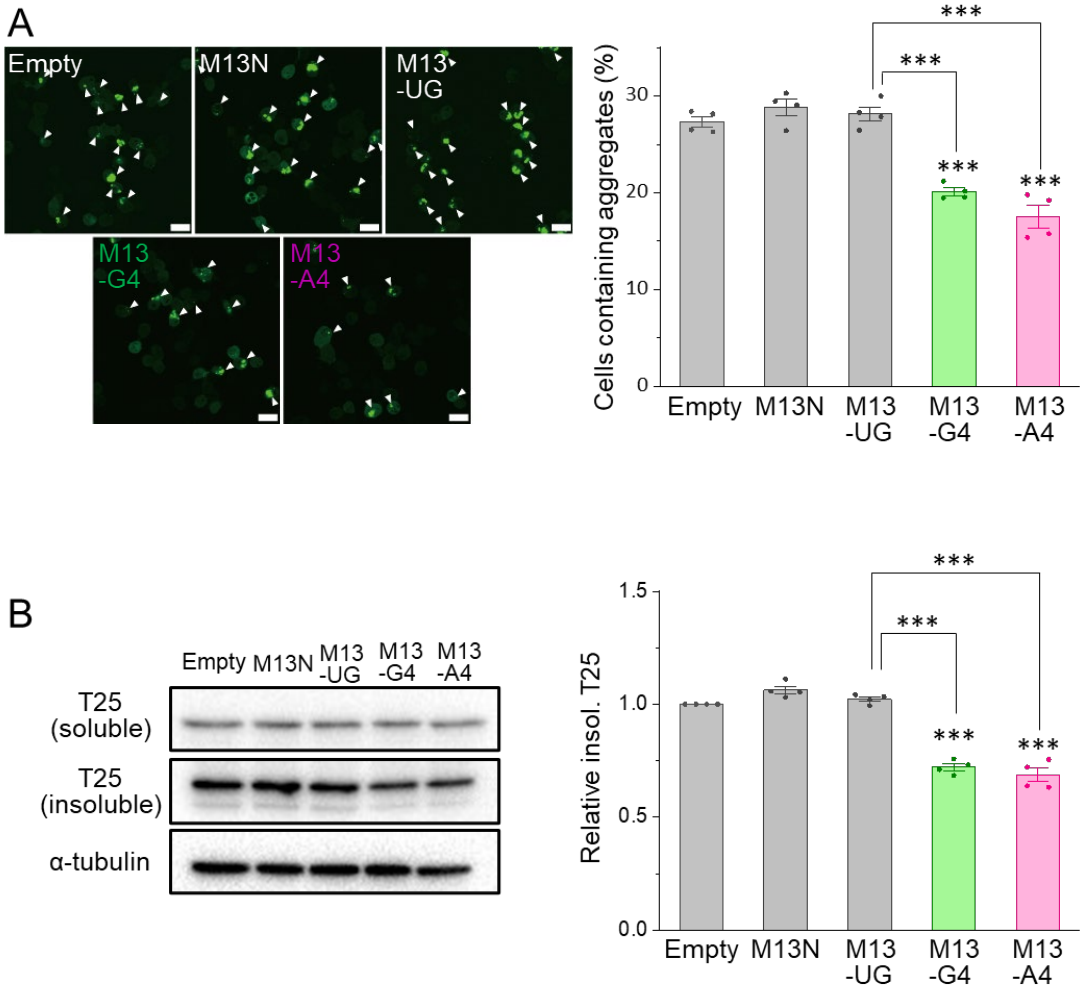
(A) Primary structures of human TDP-43 (1−414 amino acids), TDP-25 (220−414 amino acids), F220 (220−262 amino acids), and C274 (274−262 amino acids). Numbers show the amino acid position. (B) RCA values between AF-labeled RNAs (G4, and A4) that were folded in 100 mM KCl (Gq + HP) and GFP monomers (GFP), GFP-TDP-43 (T43), GFP-TDP-25 (T25), GFP-F220 (220), or GFP-C274 (274) with 100 mM KCl (n = 3; mean ± SE). Significance: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ns ($p \geq 0.05$).

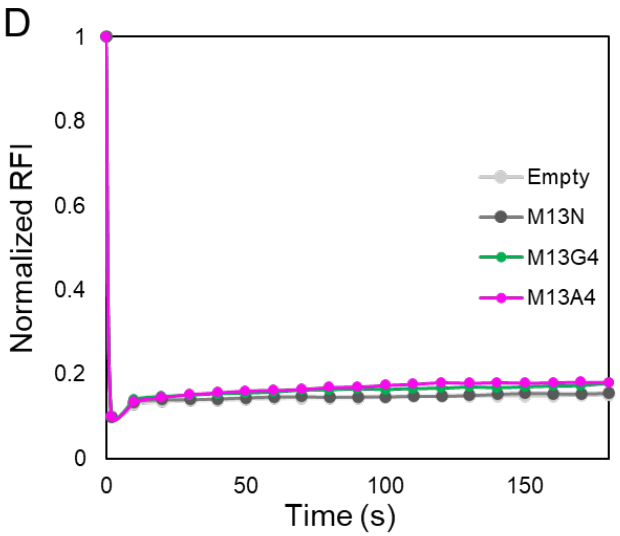
G4/A4 RNA prevents aggregation of TDP25 and ameliorates its cytotoxicity.

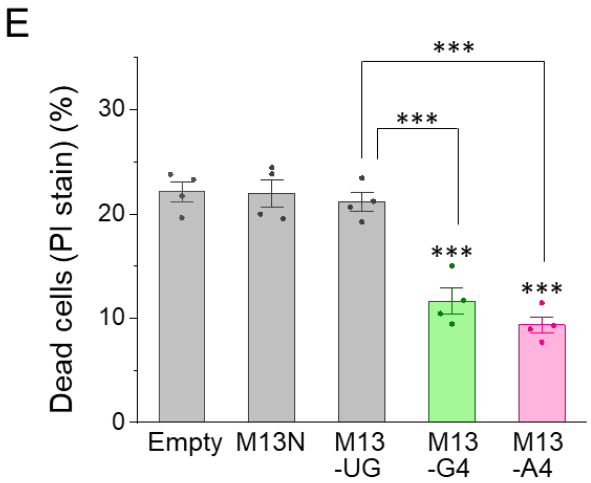
Next, to demonstrate the activity of G4 and A4 that prevent the aggregation of TDP25, G4/A4 tagged with bacteriophage-derived M13 sequences (M13-G4 and M13-A4, respectively) were expressed in Neuro2a cells by transfection of plasmid DNA that transcribes such RNAs under the human H1 promoter. The expression of M13-G4, M13-A4, M13-tagged 24-mer UG repeat RNA (M13-UG), and only M13 tag as a control (M13N) in Neuro2a cells was confirmed by reverse transcription PCR (RT-PCR) (Figure S2a). The population of cells expressing M13-G4 and M13-A4 that harbor T25 aggregates in the cytoplasm was decreased, while that of cells expressing M13N and M13-UG was not (Figure 5A). Although the expression levels of T25 in the soluble fraction of the cell lysates were not changed in cells expressing M13-G4, M13-A4, M13-UG, or M13N compared to empty vector transfection, the levels of T25 in the insoluble fraction significantly decreased in cells expressing M13-G4 and M13-A4 (Figure 5B and Figure S4B). These results suggest that G4/A4 can be an aggregation suppressor for TDP25, similar to a cytoplasmic molecular chaperone HSP70 for TDP25 aggregation[32]. Furthermore, RNA that binds to TDP-43, as in the case of UG repeat RNA, is not a sufficient condition for the prevention of T25 aggregation. Although these G4/A4 are supposed to, the next question is whether the G4/A4 introduced into cells disturbs the intracellular transcriptome, resulting in the indirect suppression of T25 aggregation through the upregulation of stress-induced proteins, such as heat shock proteins. To this end, transcriptomic analysis was performed using RNA-seq when M13-G4, -A4, and -UG were expressed with T25 in Neuro2a cells. No transcripts with a >2- and <0.5-fold change in the significant expression ratio ($p <$

0.05) were observed in both mRNA and non-coding RNA (ncRNA) in the comparison of G4/A4 with UG (Figure 5c). Therefore, it is unlikely that up-regulated chaperone-like genes by expression of G4/A4 inhibit T25 aggregation indirectly; in other words, I suggest that these RNAs directly prevent T25 aggregation. Are the biophysical molecular dynamics of T25 within the aggregates then changed in the expression of G4/A4? To investigate this issue, fluorescence recovery after photobleaching (FRAP) analysis in live Neuro2a cells was performed. T25 was quite immobile and formed amyloids in the aggregates [10,44]; however, the expression of M13-G4/A4 did not alter the immobility of T25 (Figure 5D), suggesting that G4/A4 may act on T25 in the diffusive states in the cytoplasm to prevent aggregation before the formation of deposited aggregates. As a possible reason why G4/A4 cannot act on the deposited aggregates, HSP70 overexpression does not dramatically increase the mobile population of T25 in the aggregates [10,44]. HSP70 is colocalized and dynamically interacts with T25 aggregates[36]; however, RNA is not stained inside the cytoplasmic T25 aggregates 18. This unremarkable difference in T25 mobility in the aggregates, even when overexpressed HSP70 interacts with the aggregates, is consistent with the no significant changes in T25 mobility in the aggregates when M13-G4 and -A4 were expressed, as shown here. To demonstrate whether M13-G4 and -A4 expression ameliorates TDP25 cytotoxicity, I quantified the population of propidium iodide (PI)-stained cells using fluorescence microscopy and DRAQ7-stained cells using fluorescence-activated cell sorting (FACS), indicating dead cells. The population of PI- and DRAQ7-positive and T25-expressing cells was dramatically reduced in M13-G4/A4-expressing cells compared to M13N or M13-UG expression and empty vector transfection as controls (Figure 5E & F), suggesting that G4/A4 plays a role in the

reduction of cytotoxicity in TDP25- expressing cells.


A
Empty
M13N
M13
-UG
M13
-G4
M13
-A4
Cells containing aggregates (%)
0
10
20
30
***
***
***
***
Empty
M13N
M13
-UG
M13
-G4
M13
-A4
B
Empty
M13N
M13
-UG
M13
-G4
M13
-A4
T25
(soluble)
T25
(insoluble)
α-tubulin
Relative insol. T25
0.0
0.5
1.0
1.5
***
***
***
***
Empty
M13N
M13
-UG
M13
-G4
M13
-A4
C
mRNA
ncRNA
<0.5 fold
>2 fold
-log10P
log2FC
A4
G4


D
Normalized RFI
Time (s)
Empty
M13N
M13G4
M13A4


E
Dead cells (PI stain) (%)
***
***
***
***
Empty
M13N
M13 -UG
M13 -G4
M13 -A4

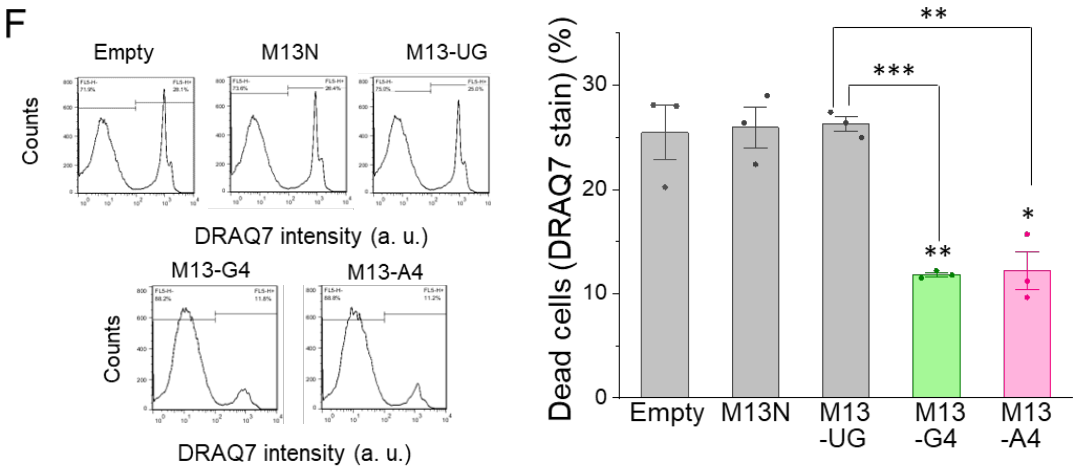

F
Empty
M13N
M13-UG
Counts
DRAQ7 intensity (a. u.)
M13-G4
M13-A4
Counts
DRAQ7 intensity (a. u.)
Dead cells (DRAQ7 stain) (%)
**
***
**
*
Empty
M13N
M13 -UG
M13 -G4
M13 -A4

Figure 5. Reduced aggregates of GFP-TDP-25 in Neuro2a cells expressing G4/A4 RNA.

(A) Confocal fluorescence images of GFP-TDP-25 (T25) in Neuro2a cells (left). The percentage of cells containing cytoplasmic aggregates of T25 when an empty vector (empty) or M13-tagged RNA was expressed (M13N, M13-UG, M13-G4, and M13-A4) (right) (the number of trials = 4; mean ± SE). (B) Western blotting of Neuro2a cells expressing T25 when M13-tagged RNAs were expressed. Typical blots of T25 in soluble and insoluble fraction, and α-tubulin in the soluble fraction (left). The quantified intensity ratio of T25 in the insoluble fraction compared with that in the empty vector transfection (the number of trials = 4; mean ± SE) (right). The gel loading the insoluble fraction of the cell lysate stained with Coomassie brilliant blue as an internal control was represented in Figure S3-A. (C) Volcano plot of mRNA (left) and ncRNA (right) expression change when M13-G4 or -A4 was expressed in Neuro2a cells compared with M13-UG (the number of trials = 4). Areas of nonsignificant change are shaded in gray (0.5–2-fold change and $p \geq 0.05$). (D) Normalized relative fluorescence intensity (RFI) indicating the mobile fraction of T25 in the cytoplasmic aggregates determined using FRAP (measured cell number = 9; mean ± SE). (E) The population of T25-expressing cells stained with propidium iodide (PI) as a dead cell marker (the number of trials = 4; mean ± SE). (F) Typical intensity diagram (left) and population (right) of T25-expressing cells stained with DRAQ7 as a dead cell marker using FACS (number of trials = 3, mean ± SE). (A, B, E, F)

G4/A4 RNA prevents aggregation of TDP25 and ameliorates its cytotoxicity by direct RNA transfection.

Next, to demonstrate that G4/A4 without any tag directly prevents T25 aggregation, the aggregate formation of T25 in Neuro2a cells introduced with RNA was analyzed. Transfection of synthetic 24-mer UG repeat RNA (UG) and siRNA against the Renilla luciferase mRNA (siRL) [32]as controls pre-introduction of T25-expression vector did not change the population of cells containing T25 aggregates in the cytoplasm compared to mock transfection; however, the direct transfection of synthetic G4/A4 pre-introduction of T25-expression vector significantly decreased the proportion of cells containing T25 aggregates (Figure 6A & B). The levels of T25 in the insoluble fraction decreased in G4-and A4-transfected cells (Figure 6B, C & D). These results confirm the hypothesis that G4/A4 can be used as an aggregation suppressor for T25 by introducing it into cells. Furthermore, the transfection of synthetic G4/A4 post-introduction of T25-expression vector significantly decreased the proportion of cells containing T25 aggregates and the amount of T25 in the insoluble fraction (Figures 6E–H); however, the efficiency of reduction of T25 aggregates during G4/A4 to UG transfection was higher with pre-T25 transfection than with post-T25 transfection (Figures 6). Therefore, G4/A4 may inhibit the aggregation and assembly process before T25 forms aggregates in the cytoplasm that can be viewed by microscopes. These RNAs would maintain the oligomeric states of T25 and/or prevent its intermolecular assembly.

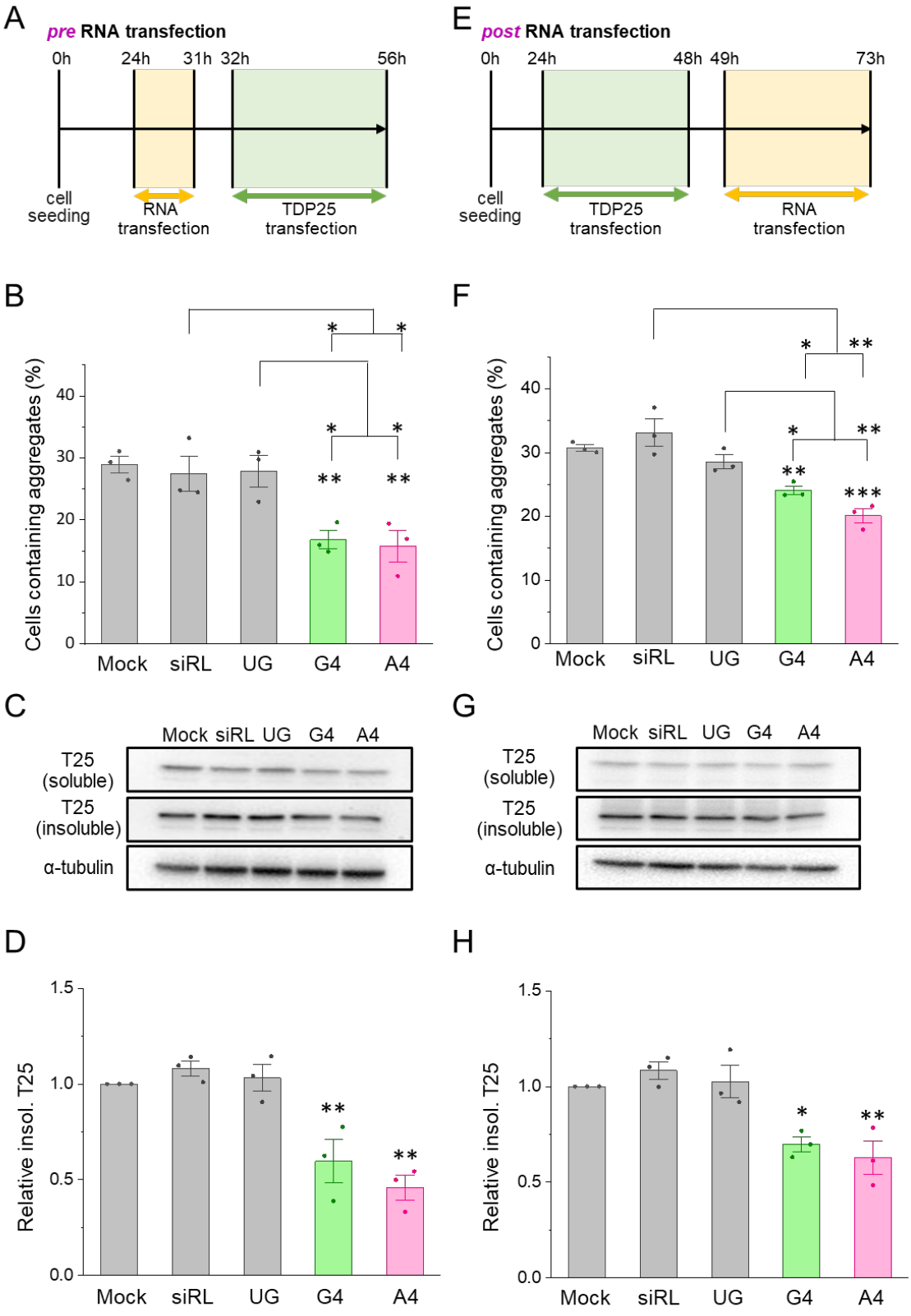

A
pre RNA transfection
0h
24h
31h
32h
56h
cell seeding
RNA transfection
TDP25 transfection
E
post RNA transfection
0h
24h
48h
49h
73h
cell seeding
TDP25 transfection
RNA transfection
B
Cells containing aggregates (%)
Mock
siRL
UG
G4
A4
F
Cells containing aggregates (%)
Mock
siRL
UG
G4
A4
C
Mock siRL UG G4 A4
T25 (soluble)
T25 (insoluble)
α-tubulin
G
Mock siRL UG G4 A4
T25 (soluble)
T25 (insoluble)
α-tubulin
D
Relative insol. T25
Mock
siRL
UG
G4
A4
H
Relative insol. T25
Mock
siRL
UG
G4
A4

Figure 6. Reduced aggregates of GFP-TDP-25 in Neuro2a cells when synthetic RNA was directly introduced.

(A)Scheme of RNA transfection pre-TDP-25-vector introduction. (B) The percentage of cells containing cytoplasmic aggregates of GFP-TDP-25 (T25) when synthetic RNA (Mock, siRL, UG, G4, or A4) was transfected (n = 3; mean ± SE). Significance: *p < 0.05 and **p< 0.01. (C) Western blotting of Neuro2a cells expressing T25 when synthetic RNA (Mock, siRL, UG, G4, or A4) was transfected. (S3-c) SDS-PAGE gel images of the insoluble fractions of cell lysates that are represented in (c), stained with Coomassie brilliant blue R-250. The numbers on the left side of the gel images show the molecular weight of the protein marker standard (Loaded in the lane for 'Marker'). (D) The quantified intensity ratio of T25 in the insoluble fraction compared with that in the empty vector transfection (the number of trials = 3; mean ± SE). (E) Scheme of RNA transfection post-TDP-25-vector introduction. (F) The percentage of cells containing cytoplasmic aggregates of T25 when synthetic RNA (Mock, siRL, UG, G4, or A4) was transfected (n = 3; mean ± SE). Significance: *p< 0.05 and **p< 0.01. h) Western blotting of Neuro2a cells expressing T25 when synthetic RNA (Mock, siRL, UG, G4, or A4) was transfected. (S3-d) SDS-PAGE gel images of the insoluble fraction of cell lysates that are represented in (h), stained with Coomassie brilliant blue R-250. The numbers on the left side of the gel images show the molecular weight of the protein marker standard (Loaded in the lane for 'Marker'). (G)The quantified intensity ratio of T25 in the insoluble fraction compared with that in the empty vector transfection (the number of trials = 3; mean ± SE).

G2/A2 RNA did not prevent aggregation of TDP25.

Next, I examined whether G2 and A2, whose RNA repeats were shortened from four to two, could be used to suppress TDP25 aggregation. Interaction analysis showed that G2 interacted with TDP25, but A2 did not (Figure 1C).

M13-tagged G2/A2 was expressed in Neuro2a cells to the same extent as M13-tagged G4/A4 (Fig. S2b) G2/A2 expression in Neuro2a cells did not alter the cell population harboring T25 aggregates or the level of T25 in the insoluble fraction (Fig. 7A, B) G2 was too short to considered to be unable to form Gq structures efficiently. Perhaps this is why the interaction strength between G2 and T25 was reduced by about 60% compared to G4 (Fig. 1C). Therefore, for r(G4C2) to effectively inhibit TDP25 aggregation in cells, some degree of interaction strength with T25, equivalent to four repetitions of r(G4C2), would be required. Alternatively, the Gq structure may play an important role. It is also important for r(A4U2) to also interact with T25 to effectively inhibit its aggregation, which may require four repeats.

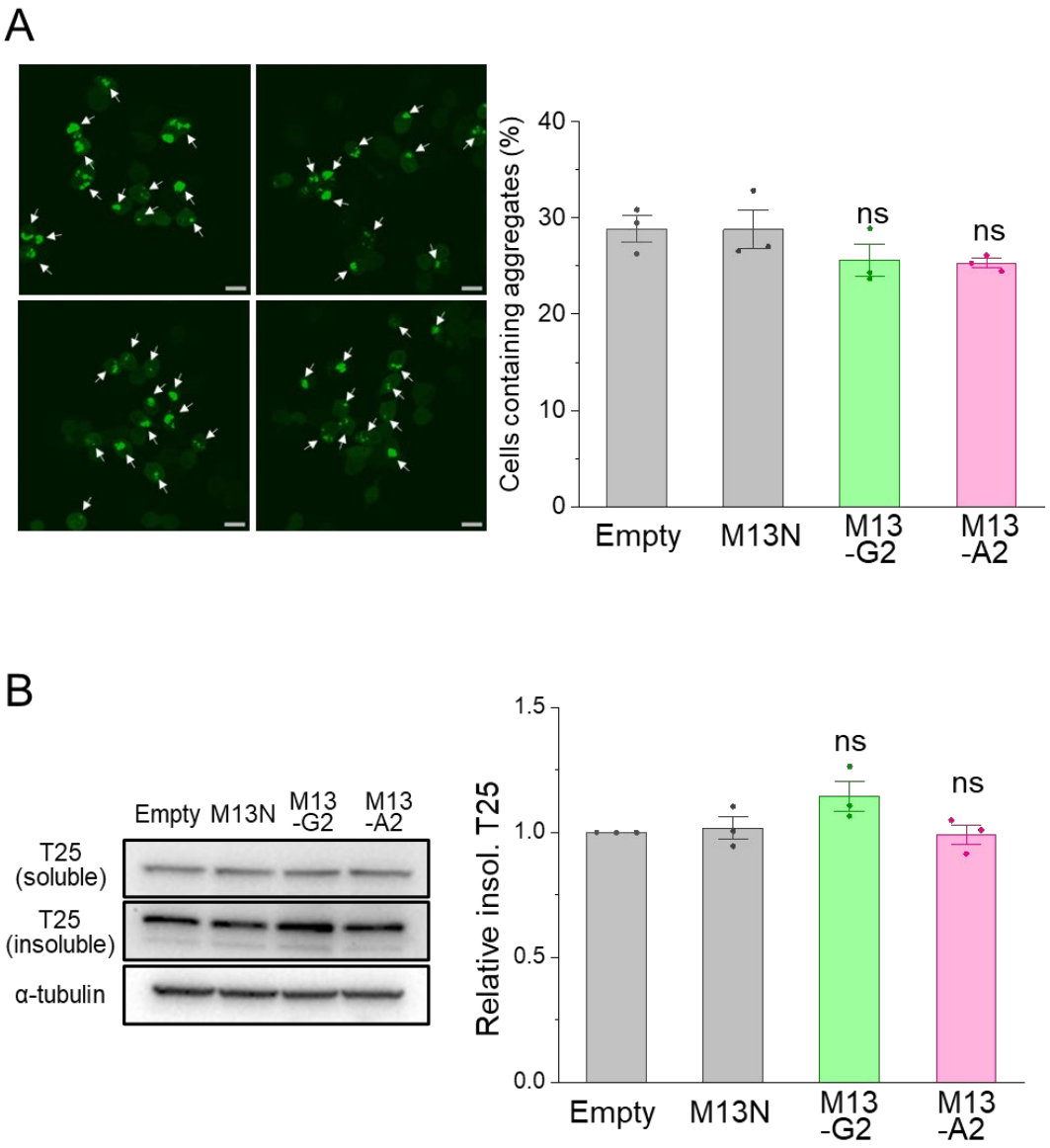


Figure 7. Aggregate formation of GFP-TDP-25 in Neuro2a cells expressing M13-G2/A2.

(A) Confocal fluorescence images of GFP-TDP-25 (T25) in Neuro2a cells (left). The percentage of cells containing cytoplasmic aggregates of T25 when an empty vector (empty) or M13-tagged RNA was expressed (M13N, M13-G2, and M13-A2) (right) (the number of trials = 4; mean ± SE). (B) Western blotting of Neuro2a cells expressing T25 when M13-tagged RNAs were expressed. (A, B) (n = 3; mean ± SE). Significance: ns ($p \geq 0.05$).

G4/A4 RNA prevents aggregation of TDP-43 and ameliorates its cytotoxicity.

Although TDP25 is highly prone to aggregation compared to TDP-43 and its CTFs[24,44,47,48] , the mislocalization of full-length TDP-43 with its aggregation is thought to be more associated with pathogenicity[26,49,50]. However, wild type TDP-43 is not mislocalized and does not form any cytoplasmic aggregates in Neuro2a cells[10]. Thus, I chose a variant of TDP-43 that lacks the ability of nuclear localization and exporting signal sequences and carries amino acid substitutions of C173S/C175S (TDP-43CS) [51]. Furthermore, because GFP-tagged TDP-43CS (T43CS) did not form cytoplasmic aggregates in Neuro2a cells, I used 293 cells, in which T43CS formed cytoplasmic aggregates efficiently. The expression of M13-G4, M13-A4, and M13N in 293 cells was confirmed using RT-PCR (Figure S2c). The population of cells that harbor T43CS aggregates in the cytoplasm was significantly reduced in cells expressing M13-G4 and M13-A4 compared to that of cells expressing M13-UG, M13N, and empty vector transfected cells as controls (Figure 8A). Next, to biochemically detect the reduction in aggregates by expression of M13-G4/A4, these cells were lysed in a buffer containing 0.1% SDS and then fractionated using centrifugation. Almost all T43CS was recovered in the soluble fraction, and no significant change in the T43CS amount both in the soluble and insoluble fraction was observed in M13-G4/A4-expressed cells (Figure 8B and Figure S5B). The aggregates of T43CS may be more loosely assembled than those of T25.

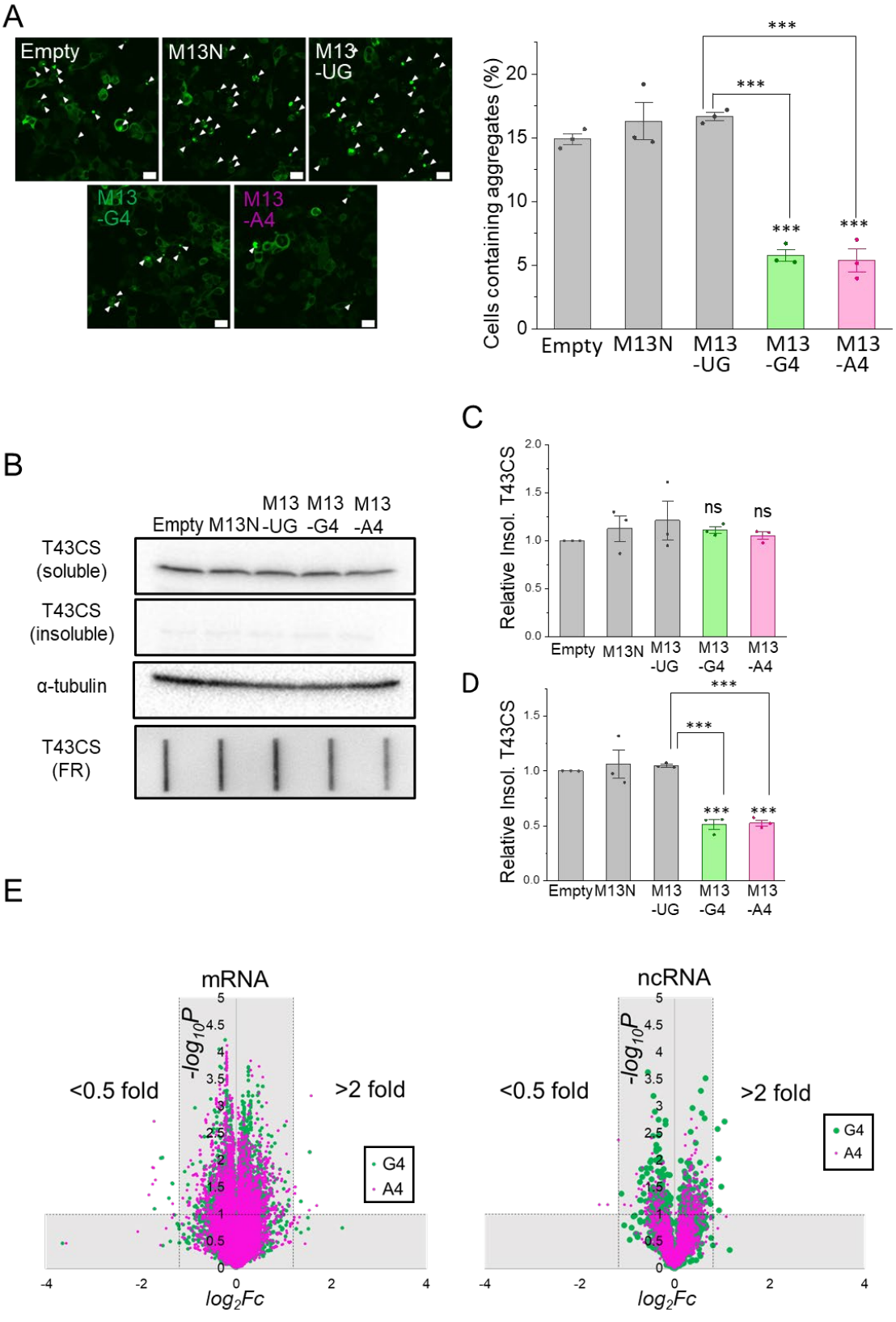
However, the filter retardation assay showed that the amount of T43CS trapped on the membrane was decreased by M13-G4/A4 expression (Figure 8B, C & D), suggesting that T43CS may not exist as complete monomers and may probably exist as oligomers

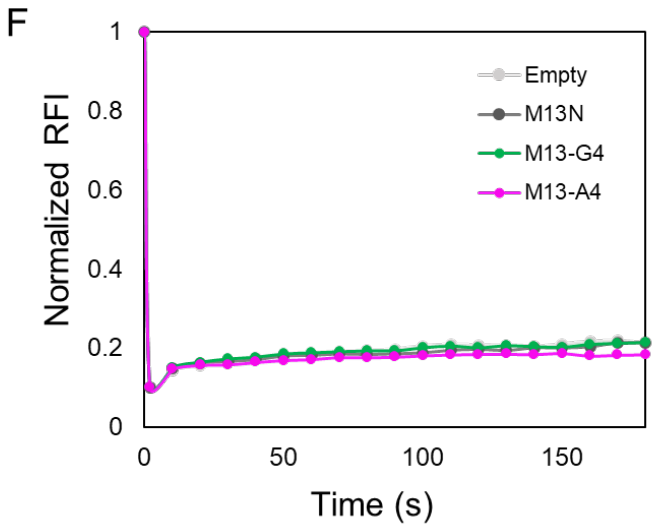
and/or soluble aggregates that were larger than the pore size of the membrane (>0.2 μm); furthermore, they can be solubilized by SDS, and G4/A4 decreases oligomers and/or soluble aggregates of T43CS likely by affecting its oligomerization process. Likewise, for 293 cells, transcriptomic analysis was performed using RNA-seq when G4/A4 were expressed. No transcripts with a >2- and <0.5-fold change in the significant expression ratio ($p < 0.05$) were observed in ncRNA. Although several transcript levels in mRNA were altered in 293 cells by expression of M13-G4/A4 in contrast to Neuro2a cells, none of the transcripts had distinct chaperone functions (Figure 8E and Table S4). Therefore, it is unlikely that up-regulated chaperone-like genes by G4/A4 expression inhibit T43CS aggregation indirectly; in other words, I suggest that these RNAs directly prevent T43CS aggregation as T25.

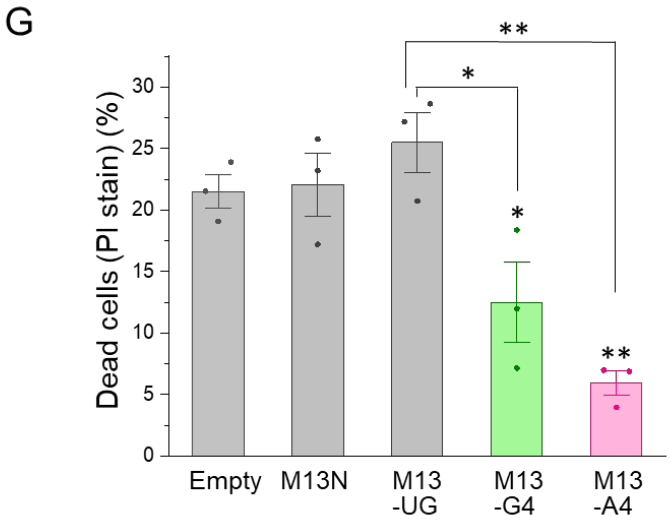
FRAP analysis revealed that T43CS was considerably immobile in the aggregates as T25, and the expression of M13-G4/A4 did not alter the immobility of T43CS (Figure 8F), suggesting that G4/A4-RNA may also act on T43CS in the diffusive states in the cytoplasm and likely not in the aggregates deposited. This is consistent with the findings of the western blot and filter retardation assays. Furthermore, because the cytoplasmic mislocalization of TDP-43 is cytotoxic[26], to demonstrate whether G4/A4-RNA expression ameliorates cytotoxicity, I quantified the population of propidium iodide (PI)-stained cells using fluorescence microscopy and DRAQ7-stained cells using FACS.

The population of PI- and DRAQ7-positive and T43CS-expressing cells was dramatically reduced in M13-G4/A4-expressing cells compared to M13N expression and empty vector transfection as controls (Figure 8G & H), suggesting that G4/A4 plays a role in the reduction of cytotoxicity in T43CS-expressing cells. Therefore,

G4/A4 could allow cells with cytoplasmic mislocalization of TDP-43 to be protected with a decrease of its oligomers/soluble aggregates.


A
Empty
M13N
M13 -UG
M13 -G4
M13 -A4
Cells containing aggregates (%)
***
***
***
***
Empty
M13N
M13 -UG
M13 -G4
M13 -A4
B
Empty
M13N
M13 -UG
M13 -G4
M13 -A4
T43CS (soluble)
T43CS (insoluble)
α-tubulin
T43CS (FR)
C
Relative Insol. T43CS
ns
ns
D
Relative Insol. T43CS
***
***
***
***
E
mRNA
ncRNA
<0.5 fold
>2 fold
-log10P
log2Fc
G4
A4


F
Normalized RFI
Time (s)
Empty
M13N
M13-G4
M13-A4


G
Dead cells (PI stain) (%)
Empty
M13N
M13 -UG
M13 -G4
M13 -A4

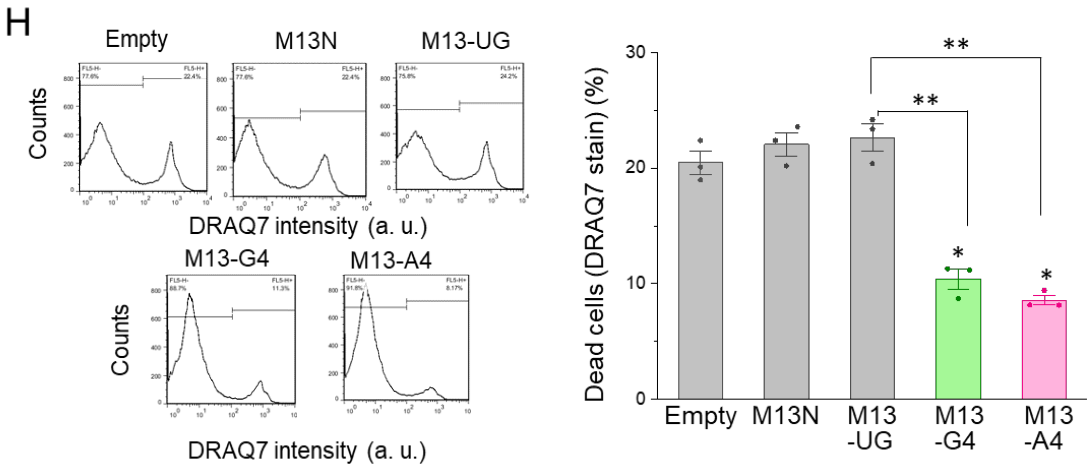

H
Empty
M13N
M13-UG
Counts
DRAQ7 intensity (a. u.)
M13-G4
M13-A4
Dead cells (DRAQ7 stain) (%)

Figure 8. Reduced aggregates of GFP-tagged TDP-43CS in 293 cells expressing G4/A4 RNA.

(A) Confocal fluorescence images of GFP-tagged TDP-43CS (T43CS) in 293 cells. The white arrows indicate the cells containing the aggregates (left). The percentage of cells containing cytoplasmic aggregates of T43CS when empty vector (empty) or M13-tagged RNA was expressed (M13N, M13-UG, M13-G4, and M13-A4) (right) (the number of trials = 3; mean ± SE). (B) Western blotting (top) and filter retardation (FR) (bottom) of 293 cells expressing T43CS when M13-tagged RNAs were expressed. (C) The quantified intensity ratio of T43CS in the insoluble fraction compared with that in the empty vector transfection (the number of trials = 3; mean ± SE). The gel loading the insoluble fraction of the cell lysate stained with Coomassie brilliant blue as an internal control was represented in Figure S3-B. The blots of total T43CS abundance in whole cell lysate are represented in Figure S5. (D) The quantified intensity ratio of the retarded T43CS on the membrane compared with the empty vector transfection (the number of trials = 3; mean ± SE). (E) Volcano plot of mRNA (left) and ncRNA (right) expression change when M13-G4 or M13-A4 was expressed in 293 cells compared with M13-UG (the number of trials = 4). Areas of nonsignificant change are shaded in gray (0.5−2-fold change and $p \geq 0.05$). Transcripts that exhibited significant changes upon expression of M13-G4 or M13-A4 are described in Table S4. (F) Normalized relative fluorescence intensity (RFI) indicating mobile fraction of T43CS in the cytoplasmic aggregates determined using FRAP (measured cell number =7; mean ± SE). (G) The population of T43CS-expressing cells stained with propidium iodide (PI) as a dead cell marker (the number of trials = 3; mean ± SE). (H) Typical intensity diagram (left) and population (right) of T43CS-expressing cells stained with DRAQ7 as a dead cell marker using FACS (number of trials = 3, mean ± SE). (the number of trials = 3; mean ± SE). (A, C, D, G, H) Significance: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Discussion

Interaction of G4/A4 with TDP25 and TDP-43

Interaction analysis revealed that the r(G4C2) sequence and the r(A4U2) sequence interact not only with TDP-43 but also with TDP25. It is known that more than four GGGGCC repeats folds into a mixed structures that include both G-quadruplexes (Gq) and hairpins (HP) in the presence of potassium ions [13,52]. In this study, the interaction of TDP25 and TDP-43 with G4 (four repeats of GGGGCC) was analyzed under conditions of Gq and HP coexistence and HP-only states. The results showed that both TDP25 and TDP-43 interacted with both Gq and HP (Fig. 1B). Even when the repeats were reduced to two (G2), TDP-43 and TDP25 showed interaction (Fig. 1C). This suggests that the interaction of r(G4C2) with TDP-43 and TDP25 depends on the RNA sequence rather than its structure or length. Predictions using RNA secondary structure prediction servers indicate that r(A4U2) does not form thermodynamically stable secondary structures[13,52]. However, TDP-43 and TDP25 interacted with A4 (four repeats of AAAAUU) (Fig. 1B). Reducing the repeat to two did not result in interaction with TDP-43 and TDP25 (Fig. 1C), highlighting that length is essential for interaction with r(A4U2). Taken together, the interaction between TDP-43/TDP25 and G4/A4 is likely based on sequence recognition rather than specific RNA tertiary structures. Despite TDP25 lacking most of the RNA recognition motifs (RRMs) found in TDP-43, it interacted with G4/A4. This implies that G4/A4 might be capable of forming noncanonical RNA interactions, unlike typical RRM-based recognition. When tested with purified proteins, TDP25 and TDP-43 still interacted with G4/A4, whereas other contained proteins, such as pyruvate carboxylase, HSP70, α, β-tubulin, and actin, did not show interaction with G4/A4 (Fig. 2B). This strongly suggests a direct

interaction between G4/A4 and TDP25 or TDP-43 (Fig. 2C). The purified samples, prepared from 1.4×10^7 cells, also contained cytoskeletal elements, likely due to large-scale culture preparation. Furthermore, HSP70 was found only in the purified TDP25, and as previously reported, the aggregation of TDP25 is suppressed by HSP70[32]. This suggests that the presence of HSP70 may help maintain the solubility of TDP25.

Biochemical mechanism of the interaction between TDP25 and TDP-43

The interaction of G4/A4 with TDP25 was significantly reduced upon the addition of 1,6-hexanediol, an inhibitor of hydrophobic interactions, suggesting a substantial hydrophobic contribution. Further analysis revealed that both truncated RRM2 and the IDR domain are required for recognition of G4/A4 (Fig.3). These findings suggest that G4/A4 recognizes the structure of TDP25 and interacts with it. Additionally, previous studies have shown that TDP25 exists in an oligomeric state[9], suggesting that G4/A4 might recognize TDP25 oligomers (Fig.4B).

G4/A4 suppression of TDP25 and TDP-43 aggregation

Expression of G4/A4 in Neuro2a cells through plasmid DNA prevented TDP25 aggregation and reduced cell death (Fig.5). Exogenous G4/A4 introduction showed no impact on the transcriptome of Neuro2a cells, strongly suggesting that G4/A4 directly suppresses TDP25 aggregation (Fig.5C), leading to reduced cell death. FRAP (fluorescence recovery after photobleaching) analysis of TDP25 aggregates showed no fluorescence recovery, indicating that G4/A4 does not function as a protease to degrade aggregates (Fig.5D). Instead, G4/A4 might prevent oligomer aggregation or restore monomers, akin to molecular chaperones[31]. Direct introduction of RNA into cells suppressed TDP25 aggregation regardless of whether it was introduced before or after TDP25 expression, although earlier introduction showed stronger effects (Fig.6). This

indicates that G4/A4 can potentially halt the propagation of TDP25. Similarly, G4/A4 was shown to suppress TDP-43 aggregation and associated cell death in 293 cells (Fig.8). FRAP analysis revealed no fluorescence recovery in TDP-43 aggregates, ruling out the possibility of G4/A4 degrading these aggregates (Fig.8F).

Transcriptome analysis indicated minimal expression changes in known aggregation-suppressing molecules like molecular chaperones, further supporting the direct suppression of aggregation by G4/A4 (Fig.5C, 8E).

RNA length and structure-dependent effects for aggregation suppression

Both r(G4C2) and r(A4U2) exhibited aggregation suppression only at four repeats and not at two repeats, despite interactions with TDP25 and TDP-43 in both conditions. This indicates that the length of r(G4C2) is critical for aggregation suppression (Fig. 5, 6, 7). Since two-repeat r(G4C2) cannot form Gq structures, these results suggest that the Gq structure plays a significant role in aggregation suppression (Fig.1C)[42]. On the other hand, Gq structures are inherently stable and abundant, potentially inducing aberrant phase transitions. Although r(G4C2) was tested at minimal repeats required for Gq structure formation, excessive Gq structures may promote abnormal aggregation, posing risks as therapeutic agents[53–55]. Conversely, r(A4U2), lacking higher-order structures like Gq, may present a safer alternative. Bioinformatics analysis using tools like BLAST identified transcripts containing G4C2/A4U2 repeat sequences, suggesting they might act within cells to prevent TDP-43 or C-terminal fragment (CTF) aggregation.

Conclusions

In conclusion, this is the first demonstration of aggregation suppressing RNA against aggregation-prone TDP-43 and TDP25 in the cell. Both four-repeat r(G4C2) and

r(A4U2) possess anti-aggregation and cytoprotective activities against pathogenic TDP-43 and TDP25.

These RNAs may help maintain proteostasis associated with TDP-43 and its CTFs. Future studies should elucidate the exact stage at which G4/A4 functions during aggregation and identify endogenous RNA sequences with such anti-aggregation properties. Discovering these RNA sequences could establish a new category of molecular biology, with "chaperone RNAs" representing functional molecules with aggregation-suppressive activity.

Significance

Protein aggregation of TDP-43 and its C-terminal fragments are associated with disease progression in neurodegenerative disorders such as ALS. These protein aggregates are so cytotoxic that suppression of aggregate formation enables to reduce the proportion of cell death and degeneration that underlies diseases such as motor neurons. Here, using a binding assay in vitro and cellular expression of short RNAs, four repeats of GGGGCC and AAAAUU RNA sequence were validated as aggregation suppressors for TDP-43 and its 25 kDa C-terminal fragment, which for the first time links anti-aggregation RNAs for TDP-43 and its fragment. Given the use of these short repeat RNAs in anti-aggregation therapy, these studies will have a direct impact on the inhibition of disease progression of ALS.

Methods

Chemicals

HEPES, Tris base, potassium chloride, sodium chloride, lithium chloride, and magnesium chloride were purchased from Fujifilm Wako Chemicals (Osaka, Japan). Heparin sodium salt from porcine intestinal mucosa and 1,6-hexanediol were purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan).

Cell Lines

Murine neuroblastoma Neuro-2a cells (#CCL-131; ATCC, Manassas,VA, USA) and Flp-In T-REx 293 cells (Thermo Fisher, Waltham,MA, USA) were maintained in DMEM (D5796; Sigma-Aldrich, St.Louis, MO, USA) supplemented with 10% FBS (12676029, ThermoFisher), 100 U/mL penicillin G (Fujifilm Wako Chemicals), and 0.1μg/mL streptomycin (Fujifilm Wako Chemicals) as described previously.

Construction of Plasmids

The plasmids for the expression of monomeric enhanced GFP(meGFP)-tagged TDP-25 (pGFP-TDP-25), 220−262 amino acids CTF (F220; truncated RRM2), and 274−414 amino acids CTF(C274) were used as per the protocol established in previousreports.[9,10] The synthetic DNA fragment for a variant of TDP-43that lacks nuclear localization and exporting signal sequences and carrying amino acid substitutions of C173S/C175S (TDP-43CS-mNLS/NES),[51] was ordered via DNA fragment synthesis services(Eurofins Genomics, Tokyo, Japan). The fragment was inserted into the HindIII and BamHI sites of pmeGFP-N1 (pTDP-43CS-GFP). For purification of meGFP-tagged wild type TDP-43, synthetic cDNA fragment for

polyhistidine tag and Strep-tag (hereafter, underlined primer sequences denote restriction enzyme sites for cloning) (5′-TGTACAGTCACCATCATCACCATCACTGGAGCCACCCG-CAGTTCGAAAAATAAGCGGCCGC-3′) was inserted into the BsrGI and NotI sites of wild type TDP-43 fused with meGFP at the C-terminal, as used in our previous study (pTDP-43-GFP-HS).18For purification of GFP-TDP25 and GFP monomers, synthetic cDNA fragment for polyhistidine tag and Strep-tag (5′-GCTAGCCAC-CATGCACCATCATCACCATCACTGGAGCCACCCGCAGTTC-GAAAAACTACCGGT-3′) was inserted into the NheI and AgeI sites of pGFP-TDP-25 and pmeGFP-N1 (pHS-GFP-TDP-25 and pHS-GFP, respectively). The synthetic DNA fragment for M13-taggedRNA (M13N, M13-UG, M13-G4, M13-A4, M13-G2, and M13-A2) (Table S1) was ordered using DNA oligonucleotide synthesis services(Thermo Fisher, Waltham, MA, USA). The fragments were inserted into the BglII and HindIII sites of pSUPERIOR.neo (pSUPn-M13N, M13-UG, pSUPn-M13-G4, pSUPn-M13-A4, pSUPn-M13-G2, andpSUPn-M13-A2). An expression plasmid for GFP-tagged HSP70(pHSP70-GFP) was modified from that for red fluorescent protein-tagged HSP70 in a previous report. Expression plasmids for yellow fluorescent protein (YFP)-tagged β-actin (pEYFP-Actin) and GFP-tagged α-tubulin (pEGFP-Tub) were purchased from Takara Bio(Shiga, Japan). An expression plasmid for GFP-tagged mitochondrial pyruvate carboxylase (pCMV3-PC-GFPSpark; HG15615-ACG) was purchased from Sino Biological Japan (Kanagawa, Japan).

RNA Synthesis

All synthetic RNAs were ordered using RNA oligonucleotide synthesis services

(AJINOMOTO BioPharma, Osaka, Japan). The consequences used in this study are presented in Table S2. For fluorescence assay, the 5′ ends of the RNAs were labeled with AlexaFluor 647, a fluorescent dye, followed by purified by high performance liquid chromatography using oligonucleotide synthesis and purification services (AJINOMOTO BioPharma). For introduc-tion into the cells, synthetic RNAs with no modification (AJINOMOTO BioPharma) were used. The sequence of siRNA against renilla luciferase mRNA (siRL) was obtained from a reference.

Circular Dichroism (CD)

CD spectra of synthetic RNAs in an HD-containing buffer were recorded in a 1 mm path length cuvette at 20 °C using a spectropolarimeter (J-820, Jasco, Tokyo, Japan) equipped with a Peltier temperature controller in a 220−340 nm measurement range,0.5 nm pitch wavelength, 1 nm bandwidth, 100 nm/min scannings peed. Spectral contribution from the buffer was subtracted from all CD spectra.

Cell Culture and Transfection

Neuro2a (2.0×10^5) or Flp-In T-REx 293 cells (4.0×10^5) were spread in a 35 mm glass-bottom dish for microscopic experiments (3910−035; IWAKI-AGC Technoglass, Shizuoka, Japan) or a 35 mm plastic dish for cell lysis (150318; Thermo Fisher) 1 day before transfection. Before spreading Flp-In T-REx 293 cells, the dishes were coated with 0.01 mg/mL poly-L-lysine. The mixture of pGFP-TDP-25 (1 μg) or pTDP-43CS-GFP (1 μg) and pSUPERIOPR.neo (1 μg), pSUPn-M13N (0.25 μg mixed with 0.75 μg of the empty vector to normalized the RNA expression level), pSUPn-M13-G4 (1 μg), pSUPn-M13-A4 (1 μg), pSUPn-M13-G2 (1 μg), or pSUPn-M13-A2 (1 μg) were

transfected using 5.0 μL of Lipofectamine 2000 (Thermo Fisher). After overnight incubation for transfection, the medium was replaced. Neuro2a and Flp-In T-REx 293 cells were cultured for 24 and 48 h after transfection, respectively, before microscopic analysis or cell lysis. For protein purification, Neuro2a cells (3.0×10^6) were spread in a 150 mm plastic dish (TR4003; NIPPON Genetics, Tokyo, Japan) 1 day before transfection. A mixture of 14 μg of salmon sperm DNA (BioDynamics Laboratory Inc., Tokyo, Japan) and 16 μg of expression plasmid DNA (pTDP-43-GFP-HS, pHS-GFP-TDP-25, or pHS-GFP) was diluted in a 150 mM NaCl solution, followed by the addition of 77.5 μg polyethyleneimine (PEI) (Fujifilm Wako Chemicals) in a 10 mM Tris-HCl buffer (pH 8.0). The mixture was added to the dishes after 15 min of incubation. After 24 h of incubation, the cells were lysed for protein purification. For purification, 30, 30, and 2 dishes were prepared for TDP-43, TDP25, and GFP monomers, respectively. For sequential transfection of RNA and plasmids, Neuro2a cells (2.0×10^5) were spread in a 35 mm glass-bottom dish (IWAKI-AGC Technoglass) 1 day before the first transfection. For RNA transfection pre-T25 vector introduction, after the cells were transfected with RNA (25 pmol) using 5.0 μL of Lipofectamine RNAi MAX (Thermo Fisher), they were incubated for 7 h. After the medium exchange, the cells were transfected with pGFP-TDP-25 (1 μg) using 2.5 μL of Lipofectamine 2000 (Thermo Fisher). The cells were cultured for 24 h before microscopic analysis. For RNA transfection post-T25 vector introduction, after the cells were transfected with pGFP-TDP-25 (1 μg) using 2.5 μl of Lipofectamine 2000 (Thermo Fisher), the cells were incubated for 25 h. After the medium exchange, the cells were transfected with RNA (25 pmol) using 5.0 μL of Lipofectamine RNAi MAX (Thermo Fisher). The cells were cultured for 24 h before the microscopic analysis.

Protein Purification

GFP-tagged protein-expressing Neuro2a cells were collected in tubes after trypsinization. Lysis buffer containing 50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 1% Noidet P-40 (Fujifilm Wako Chemicals), and 1× protease inhibitor cocktail (P8340, Sigma-Aldrich) was added to a tube placed on ice, and the supernatant was recovered aftercentrifugation at 20,400 × g for 10 min at 4 °C. Ni-NTA agarose beads (25 μL slurry for GFP monomers and 75 μL slurry for TDP-43 and TDP25; Fujifilm Wako Chemicals) were added to the supernatant, followed by rotation and incubation for 60 min. The beads were washed three times in a buffer containing 50 mM HEPESKOH (pH 7.5), 150 mM NaCl, 20 mM imidazole, and 0.2% Noidet P-40. After the proteins were eluted in a buffer containing 50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 250 mM imidazole, and 0.2% Noidet P-40, the protein solutions were applied to a spin column containing a Strep-Tactin XT Superflow suspension (Zymo Research, Irvine, CA, USA). After the column was washed in a buffer containing 50 mM HEPES-KOH (pH 7.5), 150 mM NaCl, and 0.2% Noidet P40, the proteins were eluted in a buffer containing 50 mM HEPESKOH (pH 7.5), 150 mM NaCl, 0.2% Noidet P-40, and 50 mM biotin (Fujifilm Wako Chemicals). The proteins were concentrated in a buffer containing 20 mM HEPES-KOH (pH 7.5) and 0.2% Noidet P40 using a dialysis spin column (Amicon Ultra-10; Merck, Darmstadt, Germany). The concentrations of purified fluorescent proteins were measured by using fluorescence correlation spectroscopy (FCS). The purified proteins were immediately used for the FCCS measurements. For mass spectroscopic analysis, the purified proteins were separated on a 5−20% gradient gel (ATTO, Tokyo, Japan) using SDS-PAGE, followed by silver staining using silver stain reagents (423413; Cosmo Bio, Tokyo, Japan). The bands of

interest were cut and analyzed using MALDI-TOF-MS, followed by peptide mass fingerprinting, which was performed using an outsourced service (Genomine, Pohang, South Korea).

Fluorescence Cross-Correlation Spectroscopy

Before FCCS measurement, 10 μM Alexa Fluor 647-labeled RNAs were heat-denatured and cooled (7 °C/min) in a buffer containing 20 mM Tris-HCl (pH 8.0) and 100 mM KCl for the mixture of Gq and HP or LiCl for the only HP as described previously.31 The tertiary structure of these RNA was confirmed using circular dichroism and nondenatured electrophoresis.31 Neuro2a cells were lysed in a buffer containing 20 mM HEPES-KOH (pH 7.5), 1% Noidet P-40, and a 1 x protease inhibitor cocktail (Sigma-Aldrich). The supernatants of the lysates were recovered after centrifugation at 20,400 × g for 10 min at 4 °C. The supernatant was mixed with 100 nM RNA in KCl, $MgCl_2$, heparin, or HD-containing buffer and incubated for 5 min before the measurement. Double-strand 500 bp DNA as a positive control (PCDNA) for FCCS was amplified by PCR with ATTO488- and ATTO647N-labeled primers and lambda DNA as a PCR template and was purified using MicroSpin S-400 HR columns (GE Healthcare, Chicago, IL, USA) as previously reported.32 The mixture of ATTO488- and ATTO647N-labeled primers (ATTO488-λDNA-R and ATTO647N-λDNA-F, respectively) was used as a negative control (NC-DNA) for FCCS. The DNA sequences as PC-DNA, ATTO488-λDNA-R, and ATTO647N-λDNA-F are represented in Table S3. FCCS was performed using LSM 510 META + ConfoCor3 (Carl Zeiss) with a C-Apochromat 40 × /1.2NA W. UV−VIS-IR water immersion objective lens (Carl Zeiss) as previously reported.18 The photons were recorded over 100 s. Curve

fitting analysis for the acquired auto- and cross-correlation function was performed using ZEN software (Carl Zeiss) with a model for two-component 3D diffusion involving a one-component exponential blinking fraction in the fitting for autocorrelation function, as shown in the following eq 1, and with a model for two-component 3D diffusion without a blinking fraction (T = 0 in eq 1) in the fitting for cross-correlation function.

$$G(\tau) = 1 + \left[1 + \frac{T}{1-T}\exp\left(-\frac{\tau}{\tau_T}\right)\right]\frac{1}{N}\sum_i^2 F_i \left[\left(1 + \frac{\tau}{\tau_{\mathrm{D},i}}\right)^{-1}\left(1 + \frac{\tau}{s^2\tau_{\mathrm{D},i}}\right)^{-\frac{1}{2}}\right] \quad \text{Eq. 1}$$

where $G(\tau)$ represents an auto- or cross-correlation function of the time interval τ; $\tau_{\mathrm{D},i}$ represents the diffusion time (i = 1 or 2); F_i represents the fraction of the component (i = 1 or 2); N represents the average number of particles in the confocal detection volume; T and τ_{T} represent the exponential blinking fraction and its meantime, respectively (these are zero for cross-correlation function); s represent the structural parameter determined by analyzing the standard fluorescent dyes (ATTO 488 and Cy5) on the same day. The relative cross-correlation amplitude (RCA) was calculated using Eq. 2.

$$\mathrm{RCA} = \frac{N_{\mathrm{C}}}{N_{\mathrm{G}}} \quad \text{Eq. 2}$$

where N_{C} and N_{G} represent the number of interacting and total green molecules, respectively. Dissociation constants (K_{d}) were calculated as reported previously[56].

RT-PCR

The total RNA of cells expressing M13-tagged RNA was isolated using an RNeasy Mini Kit (QIAGEN, Venlo, Netherlands) according to the manufacturer's instructions. The RNA was eluted in diethyl pyrocarbonate-treated ultrapure water. To detect M13-tagged RNAs, total RNA (1.9 μg) was used to synthesize first-strand cDNA using the Mir-X miRNA First-Strand Synthesis Kit (TaKaRa) according to the manufacturer's instructions. PCR was performed using a 0.25 U PrimeSTAR HS DNA polymerase (TaKaRa) with 0.2 mM dNTP mixture, 2 pmol of M13-forward primers (5'-gtaaaacgacggccagt-3'), 2 pmol of M13-reverse primers (5'-gagcggataacaatttcacacagg-3'), and template cDNA (corresponding to 15 ng of total RNA). Amplification was performed using a PC320 thermal cycler (Astec-Bio, Fukuoka, Japan). The products were separated on a 5–20% polyacrylamide gel (ATTO) in Tris-glycine buffer. The gel was stained with SYBR Gold (Thermo Fisher), and gel images were acquired using a Limited-STAGE II gel imager with blue LED light (AMZ System Science, Osaka, Japan).

Confocal microscopy and cell counting

GFP-expressing cells were observed using an LSM 510 META microscope (Carl Zeiss, Jena, Germany) with a Plan-Apochromat 10×/0.45NA objective (Carl Zeiss). GFP was excited at a wavelength of 488 nm. The excitation beam was divided using an HFT405/488. Fluorescence was corrected via a 505-570 nm band-pass filter (BP505-570) and then acquired using a photo-multiplier tube. The pinhole was set to 72 mm. Transmission bright-field images were acquired simultaneously with GFP fluorescence images. Aggregate-positive cells were counted using the Fiji-ImageJ software. The

population of cells that harbored cytoplasmic aggregates was normalized to the number of GFP-positive cells. Multiple fluorescent images were acquired at randomly selected locations and then GFP-positive and aggregate-positive cells were counted using the all acquired images. Aggregate-positive cells were counted when the brightness of the structures was higher than that in the cytoplasm. The total number of GFP-positive cells was counted ranging from 200 to 400 per trial. Statistical analysis of three independent transfections was performed (Student's t-test), and the results are plotted in the graph (trial number = 3; mean ± SE). Propidium iodide (PI) staining for determining dead cells

was performed as previously reported 18. The population of PI-positive cells was normalized to the total number of GFP-positive cells. The total number of cells that were counted ranged from 392 to 735 for TDP25 and from 761 to 1327 for TDP-43 per trial from the randomly acquired multiple fields of fluorescent images. Statistical analysis of three independent trials was performed (Student's t-test), and the results are plotted in the graph (trial number = 3; mean ± SE).

Fluorescence recovery after photobleaching

Photobleaching experiments were performed with an LSM 510 META system using a C-FRAP objective lens (Carl Zeiss), as previously reported 18, 35-36. GFP was excited and photobleached at 488 nm. The GFP fluorescence was collected through a bandpass filter (BP505-570). The pinhole size was set to 200 μm, the zoom factor was set to 5×, and the interval time for image acquisition was set to 10 s. The X- and Y-scanning sizes were 512 pixels, and the GFP was photobleached at 488 nm with 100% transmission for less than 1.92 s. Relative

fluorescence intensities (RFIs) were measured using the Fiji-ImageJ software and calculated as reported previously 18, 35-36. Normalized RFIs were calculated using the mean RFI of GFP-TDP25 and TDP-43CS-GFP in the aggregates with an empty vector transfection just after photobleaching.

Cell lysis, solubility assay, and filter retardation assay

Cells expressing GFP-tagged TDP25 and TDP-43CS were lysed and fractionated as reported previously 18. Briefly, cells were lysed in a lysis buffer containing 25 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 1% Noidet P-40, 0.1% SDS, 0.25 U/mL of benzonase (Sigma-Aldrich), and 1× protease inhibitor cocktail (Sigma-Aldrich). After centrifugation (20,400 × *g*, 10 min, 4°C), the supernatants were recovered, and the pellets were solubilized in PBS containing 1 M urea. SDS-PAGE sample buffer-mixed lysates were denatured at 98°C for 5 min. The samples were applied to a 12.5% polyacrylamide gel (ATTO) and subjected to electrophoresis in an SDS-containing buffer. Proteins were transferred onto polyvinylidene difluoride membranes (Cytiva, Marlborough, MA, USA).

The membrane was blocked with PBS-T buffer containing 5% skim milk for 1 h. Horseradish peroxidase-conjugated anti-GFP (#598-7, MBL, Tokyo, Japan) or anti-α-tubulin antibodies (#HRP-66031, Proteintech, Rosemont, IL, USA) were reacted in CanGet Signal Immunoreaction Enhancer Solution 1 (TOYOBO, Osaka, Japan). To detect endogenous HSP70, an anti-HSP70 antibody (#10995-1-AP, Proteintech) was reacted in the blocking buffer, and then reacted with a horseradish peroxidase-conjugated anti-rabbit IgG antibody (#111-035-144, Jackson ImmunoResearch, West Grove, PA, USA). Chemiluminescent signals were acquired using a ChemiDoc MP

imager (Bio-Rad, Hercules, CA, USA). As an internal control for the insoluble fraction, the insoluble fractions were applied to a 12.5% polyacrylamide gel (ATTO) and subjected to electrophoresis in an SDS-containing buffer, followed by staining with Coomassie brilliant blue R-250 (Fujifilm Wako Chemicals). Images of the stained gels were obtained using a Limited-STAGE II gel imager with transparency white light (AMZ System Science). For filter retardation assay, the cells were lysed in a PBS buffer containing 1% SDS and 0.25 U/mL of benzonase (Sigma-Aldrich). The lysates were applied to a cellulose acetate membrane with a pore size of 0.2 μm (Advantec Toyo, Tokyo, Japan) using a Bio-Dot SF blotter (Bio-Rad). The membrane was blocked in PBS-T buffer containing 5% skim milk for 1 h. The GFP on the membrane was detected by western blotting, as described above. All the intensities were measured using Fiji-ImageJ. After subtracting the values in the background region where there are no bands, the ratio value of the intensity of TDP-43/TDP25 in the insoluble fraction or trapped on the filter to the control was calculated.

FACS analysis

RNA-transfected Neuro2a and 293 cells were prepared as same as for cell lysis and confocal microscopy. Cells were trypsinized and suspended in a medium containing FBS. After centrifugation (1,000 × *g*, 3 min, 25°C), the cells were suspended in Opti-MEM I (Thermo Fisher). The cell concentration was adjusted to 6.0 × 105 cells/ml after cell counting. Subsequently, 3 uM DRAQ7 (BioStatus, Leicestershire, UK) was added for dead cell staining. Immediately before FACS analysis, the cells were passed through a cell strainer (#352235, Corning Incorporated, Corning, NY, USA), and fluorescence of DRAQ7-positive cells was detected using a 668-722 nm band pass

filter with 638 nm on a JSAN desktop cell sorter (Bay Bioscience, Brookline, MA, USA). The percentage of dead cells was determined based on higher fluorescence intensity.

RNA-seq and data processing

The total RNA of the cells expressing M13-tagged RNA was recovered in the same manner as that of RT-PCR. The total RNA concentration was measured using a Quantus Fluorometer with a QuantiFluor RNA system (Promega). The quality of total RNA was examined using a fragment analyzer system with an Agilent HS RNA kit (Agilent Technologies). Libraries were prepared using an MGIEasy RNA Directional Library Prep Set (MGI Tech Co., Ltd.) according to the manufacturer's instructions. The concentration of the libraries was measured using a Qubit 3.0 Fluorometer with a dsDNA HS assay kit (Thermo Fisher). The quality of the libraries was analyzed using a fragment analyzer with a dsDNA 915 reagent kit (Agilent Technologies). Circular DNA was created using the libraries and an MGIEasy Circularization kit (MGI Tech Co., Ltd.), according to the manufacturer's protocol. DNA Nano Ball (DNB) was prepared using a DNBSEQG400RS high-throughput sequencing kit (MGI Tech Co., Ltd.) according to the manufacturer's protocol. DNBs were sequenced using a DEBSEQ-G400 sequencer (MGI Tech Co., Ltd.). After removing the adaptor sequence using Cutdapt ver. 1.9.1 software, Pair reads that were less than 40 bases and a quality score of less than 20 were removed using a sickle ver. 1.33 software. Mapping was performed by referring to a murine genome reference (GRCm39) using the hisat2 ver. 2.2.1 software. The number of reads was counted using the featureCounts ver. 2.0.0 software. Reads per kilobase per million (RPKM) and transcripts per million (TPM)

values were calculated. After removing transcripts with low read counts (< 19), TPM values derived from total RNA obtained four times, independently, were averaged. The mean fold change and *p*-values were calculated for cells expressing M13-G4/A4 in comparison with cells expressing M13-UG.

Volcano plots were prepared using Microsoft Excel.

Quantification and Statistical Analysis

All the data were presented as mean ± standard error from at least three independent trials. Student's t-tests were performed using Microsoft Excel.

Supplemental

Table S1. cDNA sequence for plasmid DNA

Name	cDNA sequence (5' to 3')
TDP-43CS-mNLS/mNES	AAGCTTATGTCTGAATATATTCGGGTAACCGAAGATGAGAACGATGAGCCCATTGAAATACCATCGGAAGACGATGGGACGGTGCTGCTCTC CACGGTTACAGCCCAGTTTCCAGGGGCGTGTGGGCTTCGCTACAGGAATCCAGTGTCTCAGTGTATGAGAGGTGTCCGGCTGGTAGAAGGAA TTCTGCATGCCCCAGATGCTGGCTGGGGAAATCTGGTGTATGTTGTCAACTATCCAAAAGATAACGCCGCTGCCATGGATGAGACAGATGCT TCATCAGCAGTGGCCGTGGCCGCTGCAGTCCAGAAAACATCCGATTTAATAGTGTTGGGTCTCCCATGGAAAACAACCGAACAGGACCTGAA AGAGTATTTTAGTACCTTTGGAGAAGTTCTTATGGTGCAGGTCAAGAAAGATCTTAAGACTGGTCATTCAAAGGGGTTTGGCTTTGTTCGTT TTACGGAATATGAAACACAAGTGAAAGTAATGTCACAGCGACATATGATAGATGGACGATGGTCCGACTCCAAACTTCCTAATTCTAAGCAA AGCCAAGATGAGCCTTTGAGAAGCAGAAAAGTGTTTGTGGGGCGCTGTACAGAGGACATGACTGAGGATGAGCTGCGGGAGTTCTTCTCTCA GTACGGGGATGTGATGGATGTCTTCATCCCCAAGCCATTCAGGGCCTTTGCCTTTGTTACATTTGCAGATGATCAGGCCGCGCAGTCTGCCT GTGGAGAGGACGCCGCTGCCAAAGGAATCAGCGTTCATATATCCAATGCCGAACCTAAGCACAATAGCAATAGACAGTTAGAAAGAAGTGGA AGATTTGGTGGTAATCCAGGTGGCTTTGGGAATCAGGGTGGATTTGGTAATAGCAGAGGGGGTGGAGCTGGTTTGGGAAACAATCAAGGTAG TAATATGGGTGGTGGGATGAACTTTGGTGCGTTCAGCATTAATCCAGCCATGATGGCTGCCGCCCAGGCAGCACTACAGAGCAGTTGGGGTA TGATGGGCATGTTAGCCAGCCAGCAGAACCAGTCAGGCCCATCGGGTAATAACCAAAACCAAGGCAACATGCAGAGGGAGCCAAACCAGGCC TTCGGTTCTGGAAATAACTCTTATAGTGGCTCTAATTCTGGTGCAGCAATTGGTTGGGGATCAGCATCCAATGCAGGGTCGGGCAGTGGTTT TAATGGAGGCTTTGGCTCAAGCATGGATTCTAAGTCTTCTGGCTGGGGAATGAATCACGCGGATCC
M13N	GGATCCCCGTAAAACGACGGCCAGTGAGCTCACCTGTGTGAAATTGTTATCCGCTCTTTTTAAGCTT
M13-UG	GGATCCCCGTAAAACGACGGCCAGTGGGGCCGGGGCCGGGGCCGGGGCCGAGCTCACCTGTGTGAAATTGTTATCCGCTCTTTTTAAGCTT
M13-G4	GGATCCCCGTAAAACGACGGCCAGTTGTGTGTGTGTGTGTGTGTGTGGAGCTCACCTGTGTGAAATTGTTATCCGCTCTTTTTAAGCTT
M13-A4	GGATCCCCGTAAAACGACGGCCAGTAAAATTAAAATTAAAATTAAAATTGAGCTCACCTGTGTGAAATTGTTATCCGCTCTTTTTAAGCTT
M13-G2	GGATCCCCGTAAAACGACGGCCAGTGGGGCCGGGGCCGAGCTCACCTGTGTGAAATTGTTATCCGCTCTTTTTAAGCTT
M13-A2	GGATCCCCGTAAAACGACGGCCAGTAAAATTAAAATTGAGCTCACCTGTGTGAAATTGTTATCCGCTCTTTTTAAGCTT

Table S2. Synthetic RNA sequence

Name	Synthetic RNA sequence (5' to 3')
UG	UGUGUGUGUGUGUGUGUGUGUG
G4	GGGGCCGGGGCCGGGGCCGGGGCC
A4	AAAAUUAAAAUUAAAAUUAAAAUU
G2	GGGGCCGGGGCC
A2	AAAAUUAAAAUU
siRL sense	GCCAAGAAGUUUCCUAAUATT
siRL antisense	UAUUAGGAAACUUCUUGGCTT

Table S3. DNA sequences for FCCS controls.

Name	DNA sequence (5' to 3')
PC-DNA	GATGAGTTCGTGTCCGTACAACTGGCGTAATCATGGCCCTTCGGGGCCATTGTTTCTCTGT GGAGGAGTCCATGACGAAAGATGAACTGATTGCCCGTCTCCGCTCGCTGGGTGAACAACTG AACCGTGATGTCAGCCTGACGGGGACGAAAGAAGAACTGGCGCTCCGTGTGGCAGAGCTGA AAGAGGAGCTTGATGACACGGATGAAACTGCCGGTCAGGACACCCCTCTCAGCCGGGAAAA TGTGCTGACCGGACATGAAAATGAGGTGGGATCAGCGCAGCCGGATACCGTGATTCTGGAT ACGTCTGAACTGGTCACGGTCGTGGCACTGGTGAAGCTGCATACTGATGCACTTCACGCCA CGCGGGATGAACCTGTGGCATTTGTGCTGCCGGGAACGGCGTTTCGTGTCTCTGCCGGTGT GGCAGCCGAAATGACAGAGCGCGGCCTGGCCAGAATGCAATAACGGGAGGCGCTGTGGCTG ATTTCGATAACC
ATTO647N-λDNA-F	GATGAGTTCGTGTCCGTACAACT
ATTO488-λDNA-R	GGTTATCGAAATCAGCCACAGCG

Table S4.

Transcripts that exhibited changes upon expression of G4- or A4-RNA in 293 cells

mRNA

	Fold change	p value	G4/A4
DENND11	3.0	0.0006	A4
MAGED4B	2.9	0.0071	G4
NTAN1	2.4	0.0316	A4
ZNF251	2.2	0.0345	A4
PARN	2.1	0.0019	A4
TRPV2	2.1	0.0077	A4
SLC27A1	2.0	0.0023	G4
SKIV2L	2.0	0.0340	G4
NFASC	0.8	0.0359	A4
COL24A1	0.5	0.0133	A4
RETSAT	0.5	0.0283	G4
FAM120B	0.5	0.0163	G4
CHCHD10	0.4	0.0274	G4
PLEKHM1	0.3	0.0091	G4
ASMT	0.3	0.0019	A4
DHRS4	0.3	0.0282	A4
RAB11FIP3	0.3	0.0437	A4

ncRNA

	Fold change	p value	G4/A4
LOC100506472	2.1	0.0019	G4
UXT-AS1	0.5	0.0413	G4
FAM106A	0.4	0.0042	A4

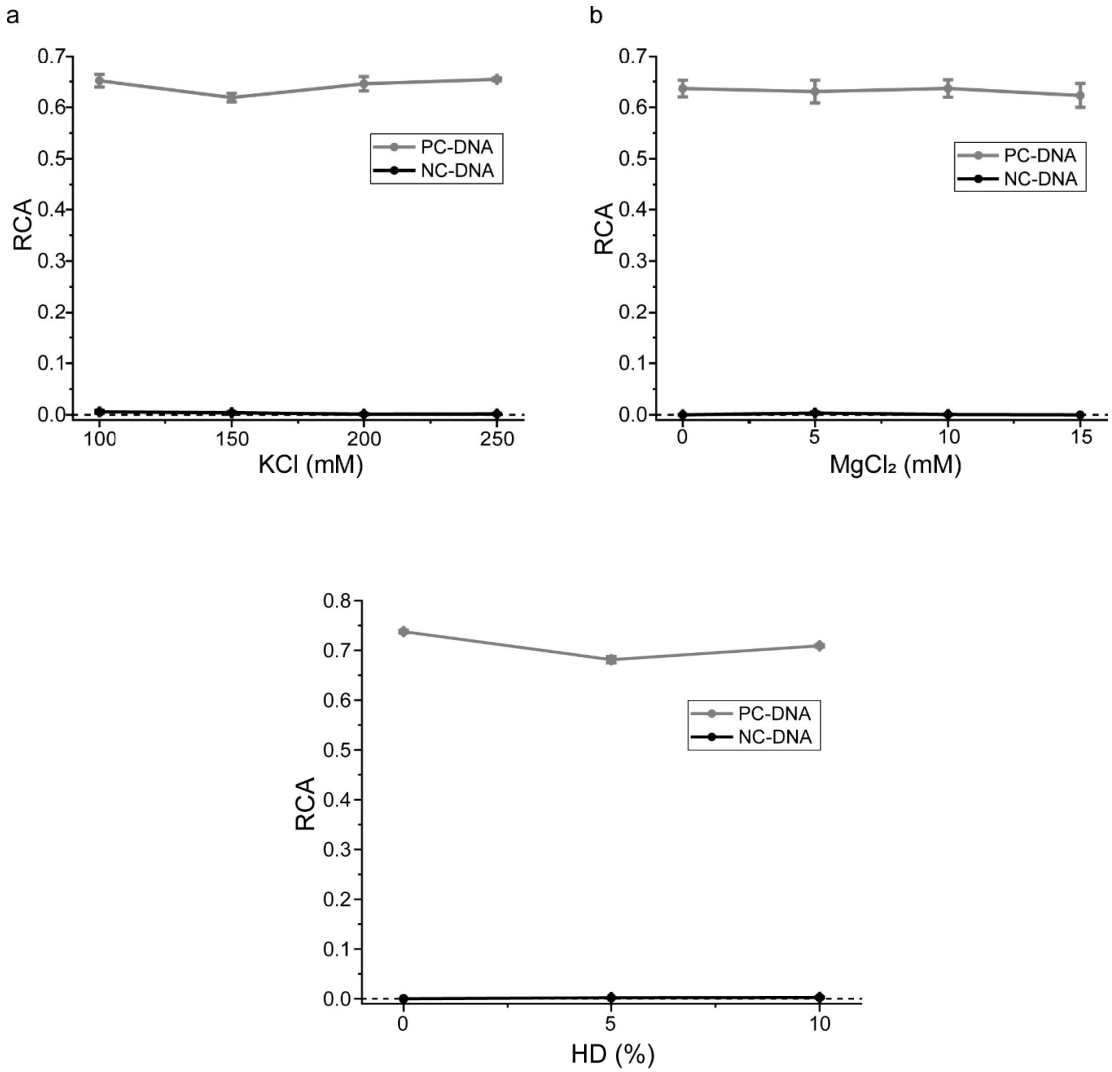


Figure S1. Interaction strength in the presence of electrostatic interaction inhibitors.

(a) Relative cross-correlation amplitude (RCA) of PC and NC in the presence of 100–250 mM KCl. The dashed line indicates zero. (b) RCA of PC and NC in the presence of 0–15 mM MgCl2. The dashed line indicates zero. (c) Relative cross-correlation amplitude (RCA) of PC and NC in the presence of 0, 5, and 10% 1,6-hexanediol (HD). The dashed line indicates zero.

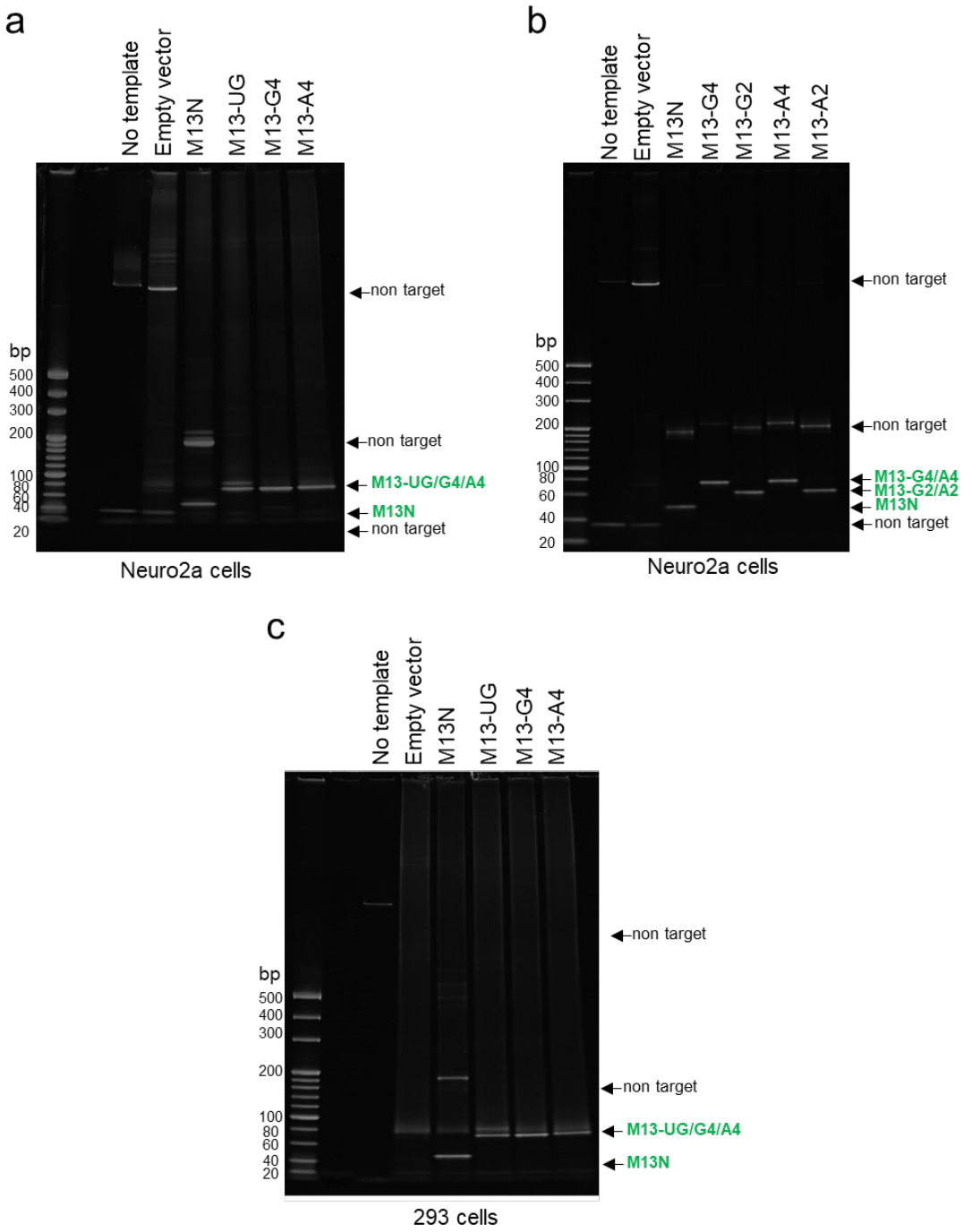
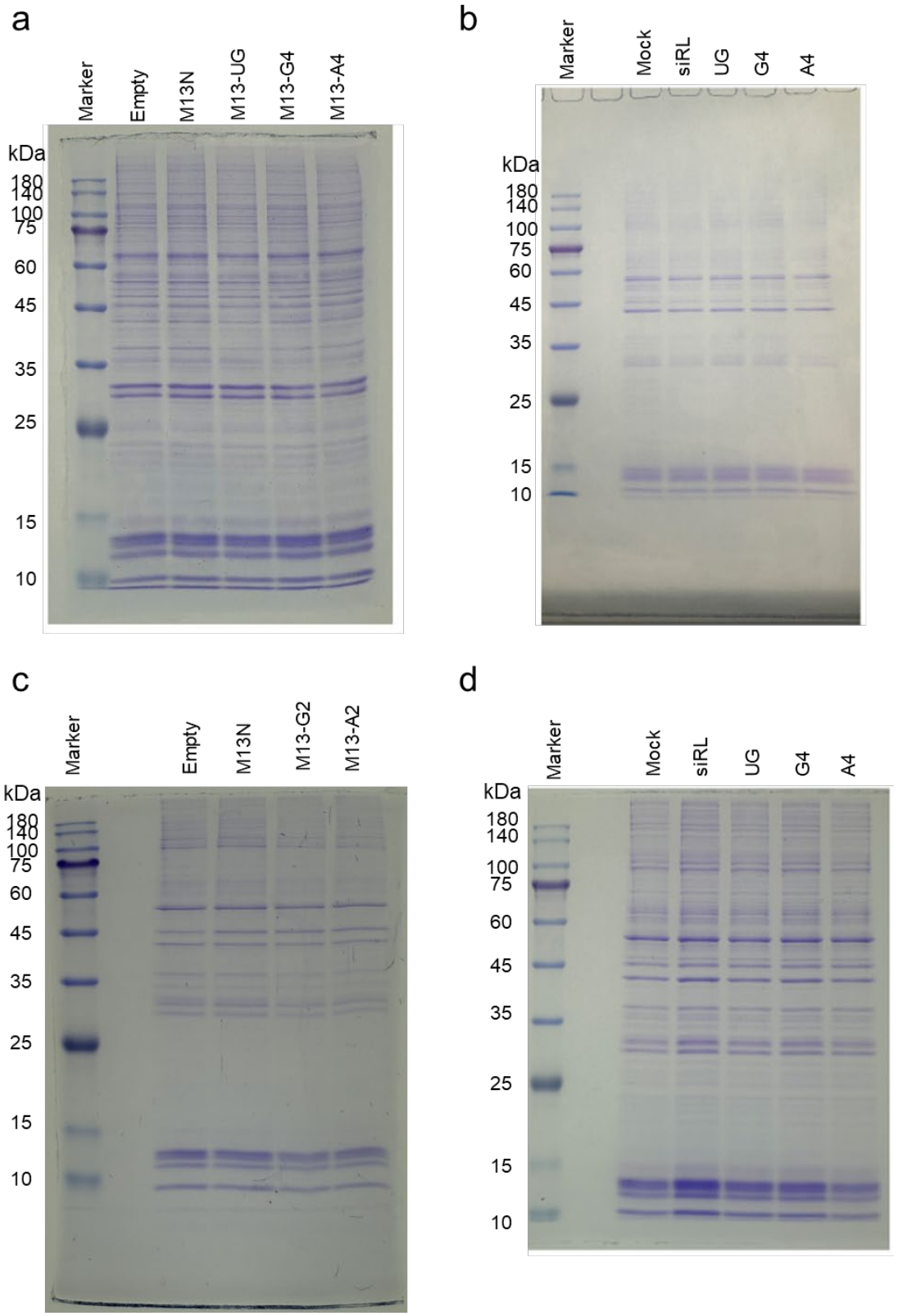


Figure S2. Confirmation of the expression levels of M13-tagged RNA using RT-PCR.

(a, b & c) The amplified DNA was analyzed on the gel derived from Neuro2a (a, b) and Flp-In T-Rex 293 cells (b). The green letters indicate the bands of interest (M13N, M13-UG, M13-G4, M13-G2, M13-A4 and –A2). The values shown on the left of the gel images indicate the size of the molecular weight markers for double-stranded DNA.


a
Marker
Empty
M13N
M13-UG
M13-G4
M13-A4
kDa
180
140
100
75
60
45
35
25
15
10
b
Marker
Mock
siRL
UG
G4
A4
kDa
180
140
100
75
60
45
35
25
15
10
c
Marker
Empty
M13N
M13-G2
M13-A2
kDa
180
140
100
75
60
45
35
25
15
10
d
Marker
Mock
siRL
UG
G4
A4
kDa
180
140
100
75
60
45
35
25
15
10

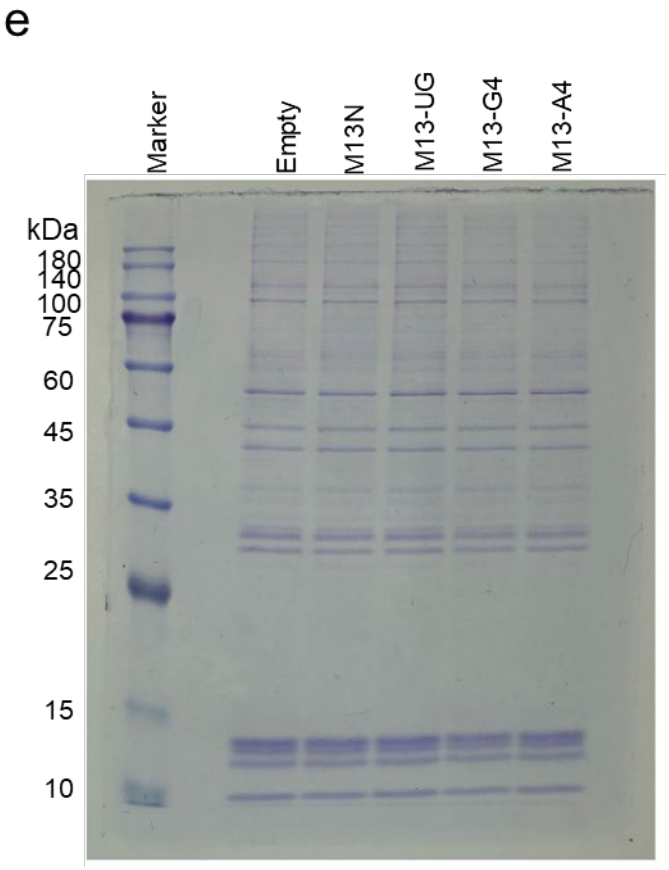


Figure S3. Loading control of insoluble fraction of cell lysate.

(a, b, c, d, e) SDS-PAGE gel images of insoluble fraction of cell lysates are represented in Fig. 5b (a), Fig. 6C (b), Fig. 6G (c), Fig. 7B (d) and 8B (b), stained with Coomassie brilliant blue R-250. The numbers on the left side of the gel images show the molecular weight of the protein marker standard (Loaded in the lane for 'Marker'). The insoluble fraction of Neuro2a cells expressing GFP-tagged TDP25 (a) when an empty vector (Empty) or M13-tagged RNA was expressed (M13N, M13-UG, M13-G4, and M13-A4). (b) when an empty vector (Empty) or M13-tagged RNA was expressed (M13N, M13-G2, and M13-A2) (c, d) when synthetic RNA (Mock, siRL, UG, G4, or A4) was transfected (e) Flp-In T-Rex 293 cells expressing GFP-tagged TDP-43when an empty vector (Empty) or M13-tagged RNA was expressed (M13N, M13-UG, M13-G4, and M13-A4).

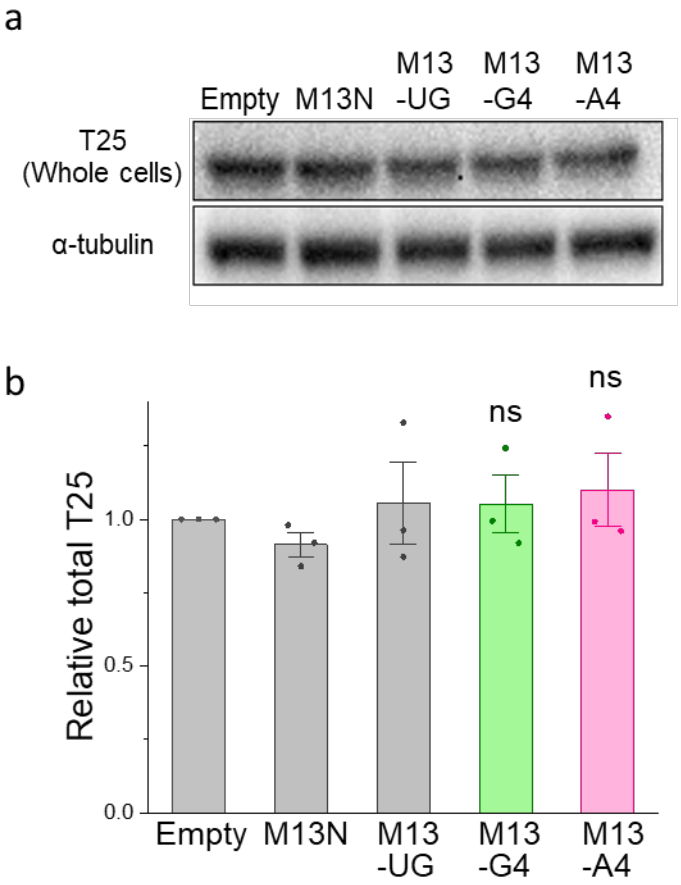


Figure S4. Total GFP-TDP-25 abundance in cell lysate when expressing M13-G4/A4.

(a) Western blotting of T25 and α-tubulin as a loading control in the whole cell lysates when expressing T25 (top). (b) The quantified intensity ratio of total T25 in the whole cell lysates compared with that in the empty vector transfection (the number of trials = 3; mean ± SE)(bottom). Significance: ns ($p \geq 0.05$).

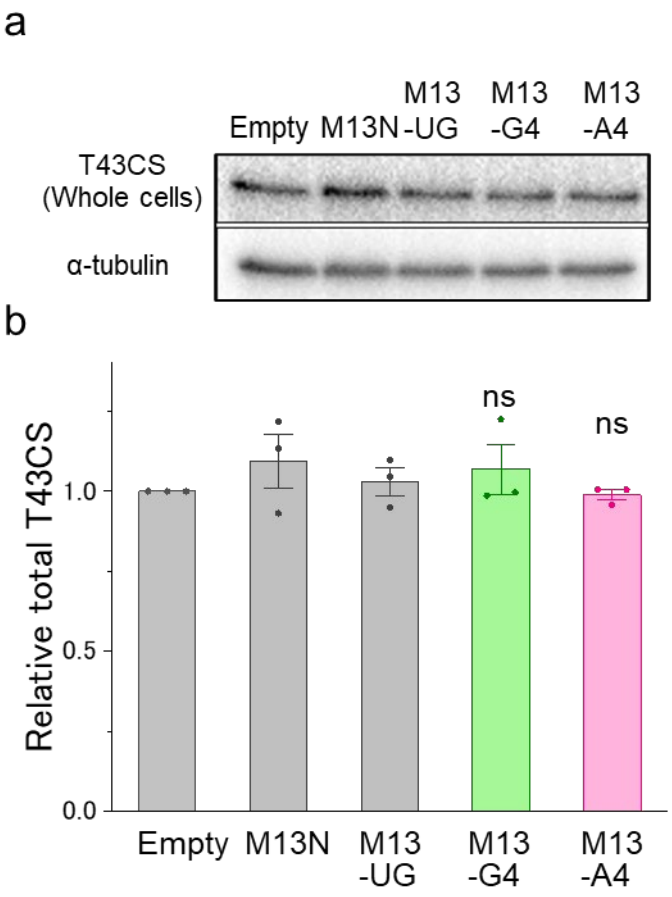


Figure S5. Total GFP-TDP-43CS abundance in cell lysate when expressing M13-G4/A4.

(a) Western blotting of T43CS and α-tubulin as a loading control in the whole cell lysates when expressing T43CS (top). (b)The quantified intensity ratio of total T43CS in the whole cell lysates compared with that in the empty vector transfection (the number of trials = 3; mean ± SE) (bottom). Significance: ns ($p \geq 0.05$).

Acknowledgments

This research was conducted under the supervision of Associate Professor Akira Kitamura of the Laboratory of Cellular and Molecular Sciences (formerly the Laboratory of Molecular Cell Dynamics), Hokkaido University. We are grateful to Associate Professor Akira Kitamura for all of his support, from basic to applied experimental skills and thinking.

Professor Emeritus Masataka Kinjo discussed in depth the data analysis of FCS and FCCS. We also thank him for providing us with abundant experimental equipment and environment.

Asuka Murata, an experimental assistant, not only conducted this research, but also gave me a great deal of mental support.

I am grateful to my senior colleague, Dr Sho Oasa (Karolinska Institutet), who carefully taught me the basic theory of FCS. Mr. Ryosuke Fukushima taught me basic microscope handling and statistical analysis.

I am deeply thankful to Professor Kenji Monde, Professor Toyoyuki Ose, and Professor Ryota Uehara for their unwavering support and insightful guidance, which have been invaluable to my work.

D. student, this research was enriched by the support of JST SPRING, Grant JPMJSP2119.

I would also like to express my gratitude to my seniors, classmates, and juniors in the Laboratory of Cellular and Molecular Sciences (formerly Laboratory of Molecular Cell Dynamics) for their cooperation and support. Finally, I would like to thank my parents, family, and friends for their financial and emotional support.

Reference

1. Hardiman, O., van den Berg, L. H. & Kiernan, M. C. Clinical diagnosis and management of amyotrophic lateral sclerosis. *Nat Rev Neurol* **7**, 639–649 (2011).

2. Neumann, M. *et al.* Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science (1979)* **314**, 130–133 (2006).

3. Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun* **351**, 602–611 (2006).

4. Bolognesi, B. *et al.* The mutational landscape of a prion-like domain. *Nat Commun* **10**, 4162 (2019).

5. Guo, W. *et al.* An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nat Struct Mol Biol* **18**, 822–830 (2011).

6. Jiang, L.-L. *et al.* Structural Transformation of the Amyloidogenic Core Region of TDP-43 Protein Initiates Its Aggregation and Cytoplasmic Inclusion. *Journal of Biological Chemistry* **288**, 19614–19624 (2013).

7. Da Cruz, S. & Cleveland, D. W. Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Curr Opin Neurobiol* **21**, 904–919 (2011).

8. Carey, J. L. & Guo, L. Liquid-Liquid Phase Separation of TDP-43 and FUS in Physiology and Pathology of Neurodegenerative Diseases. *Front Mol Biosci* **9**, (2022).

9. Kitamura, A. *et al.* Hetero-oligomerization of TDP-43 carboxy-terminal fragments with cellular proteins contributes to proteotoxicity. *Commun Biol* **7**, 743 (2024).

10. Kitamura, A. *et al.* Interaction of RNA with a C-terminal fragment of the amyotrophic lateral sclerosis-associated TDP43 reduces cytotoxicity. *Sci Rep* **6**, 19230 (2016).

11. Lukavsky, P. J. *et al.* Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat Struct Mol Biol* **20**, 1443–1449 (2013).

12. Cooper-Knock, J. *et al.* Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* **137**, 2040–2051 (2014).

13. Haeusler, A. R. *et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195–200 (2014).

14. Kitamura, A., Nagata, K. & Kinjo, M. Conformational Analysis of Misfolded Protein Aggregation by FRET and Live-Cell Imaging Techniques. *Int J Mol Sci* **16**, 6076–6092 (2015).

15. Bacia, K., Kim, S. A. & Schwille, P. Fluorescence cross-correlation spectroscopy in living cells. *Nat Methods* **3**, 83–89 (2006).

16. Ishiguro, A., Kimura, N., Watanabe, Y., Watanabe, S. & Ishihama, A. TDP-43 binds and transports G-quadruplex-containing mRNAs into neurites for local translation. *Genes to Cells* **21**, 466–481 (2016).

17. Antao, V. C. & Horton, D. K. The National Amyotrophic Lateral Sclerosis (ALS) Registry. *J Environ Health* **75**, 28–30 (2012).

18. Bruijn, L. I., Miller, T. M. & Cleveland, D. W. UNRAVELING THE MECHANISMS INVOLVED IN MOTOR NEURON DEGENERATION IN ALS. *Annu Rev Neurosci* **27**, 723–749 (2004).

19. Boillée, S., Vande Velde, C. & Cleveland, D. W. ALS: A Disease of Motor Neurons and Their Nonneuronal Neighbors. *Neuron* **52**, 39–59 (2006).

20. Ou, S. H., Wu, F., Harrich, D., García-Martínez, L. F. & Gaynor, R. B. Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol* **69**, 3584–3596 (1995).

21. Lagier-Tourenne, C. & Cleveland, D. W. Rethinking ALS: The FUS about TDP-43. *Cell* **136**, 1001–1004 (2009).

22. Mackness, B. C., Tran, M. T., McClain, S. P., Matthews, C. R. & Zitzewitz, J. A. Folding of the RNA Recognition Motif (RRM) Domains of the Amyotrophic Lateral Sclerosis (ALS)-linked Protein TDP-43 Reveals an Intermediate State. *Journal of Biological Chemistry* **289**, 8264–8276 (2014).

23. Zhang, Y.-J. *et al.* Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proceedings of the National Academy of Sciences* **106**, 7607–7612 (2009).

24. Berning, B. A. & Walker, A. K. The Pathobiology of TDP-43 C-Terminal Fragments in ALS and FTLD. *Front Neurosci* **13**, (2019).

25. Xiao, S. *et al.* Low molecular weight species of TDP-43 generated by abnormal splicing form inclusions in amyotrophic lateral sclerosis and result in motor neuron death. *Acta Neuropathol* **130**, 49–61 (2015).

26. Suk, T. R. & Rousseaux, M. W. C. The role of TDP-43 mislocalization in amyotrophic lateral sclerosis. *Mol Neurodegener* **15**, 45 (2020).

27. Polymenidou, M. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci* **14**, 459–468 (2011).

28. Buratti, E. Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. *Frontiers in Bioscience* **13**, 867 (2008).

29. Hipp, M. S., Kasturi, P. & Hartl, F. U. The proteostasis network and its decline in ageing. *Nat Rev Mol Cell Biol* **20**, 421–435 (2019).

30. Labbadia, J. & Morimoto, R. I. The Biology of Proteostasis in Aging and Disease. *Annu Rev Biochem* **84**, 435–464 (2015).

31. Hartl, F. U. & Hayer-Hartl, M. Molecular Chaperones in the Cytosol: from Nascent Chain to Folded Protein. *Science (1979)* **295**, 1852–1858 (2002).

32. Kitamura, A., Iwasaki, N. & Kinjo, M. Molecular chaperone HSP70 prevents formation of inclusion bodies of the 25-kDa C-terminal fragment of TDP-43 by preventing aggregate accumulation. *Cell Stress Chaperones* **23**, 1177–1183 (2018).

33. Tauber, D., Tauber, G. & Parker, R. Mechanisms and Regulation of RNA Condensation in RNP Granule Formation. *Trends Biochem Sci* **45**, 764–778 (2020).

34. Aarum, J. *et al.* Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates. *EMBO Rep* **21**, (2020).

35. Rengifo-Gonzalez, J. C. *et al.* The cooperative binding of TDP-43 to GU-rich RNA repeats antagonizes TDP-43 aggregation. *Elife* **10**, (2021).

36. Mann, J. R. *et al.* RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. *Neuron* **102**, 321-338.e8 (2019).

37. Son, A., Huizar Cabral, V., Huang, Z., Litberg, T. J. & Horowitz, S. G-quadruplexes rescuing protein folding. *Proceedings of the National Academy of Sciences* **120**, (2023).

38. Kaiser, C. M., Goldman, D. H., Chodera, J. D., Tinoco, I. & Bustamante, C. The Ribosome Modulates Nascent Protein Folding. *Science (1979)* **334**, 1723–1727 (2011).

39. Horowitz, S. & Bardwell, J. C. A. RNAs as chaperones. *RNA Biol* **13**, 1228–1231 (2016).

40. Choi, S. Il *et al.* Protein Solubility and Folding Enhancement by Interaction with RNA. *PLoS One* **3**, e2677 (2008).

41. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72**, 245–256 (2011).

42. Kitamura, A. *et al.* Trans-cis isomerization kinetics of cyanine dyes reports on the folding states of exogeneous RNA G-quadruplexes in live cells. *Nucleic Acids Res* **51**, (2023).

43. Gitler, A. D. & Shorter, J. RNA-binding proteins with prion-like domains in ALS and FTLD-U. *Prion* **5**, 179–187 (2011).

44. Kerr, A. G. *et al.* The long noncoding RNA ADIPINT regulates human adipocyte metabolism via pyruvate carboxylase. *Nat Commun* **13**, 2958 (2022).

45. Kishor, A. *et al.* Hsp70's RNA-binding and mRNA-stabilizing activities are independent of its protein chaperone functions. *Journal of Biological Chemistry* **292**, 14122–14133 (2017).

46. Sasaki, A., Yamamoto, J., Jin, T. & Kinjo, M. Raster image cross-correlation analysis for spatiotemporal visualization of intracellular degradation activities against exogenous DNAs. *Sci Rep* **5**, 14428 (2015).

47. Chhangani, D., Martín-Peña, A. & Rincon-Limas, D. E. Molecular, functional, and pathological aspects of TDP-43 fragmentation. *iScience* **24**, 102459 (2021).

48. Calabrese, J. M. & Sharp, P. A. Characterization of the short RNAs bound by the P19 suppressor of RNA silencing in mouse embryonic stem cells. *RNA* **12**, 2092–2102 (2006).

49. Kitamura, A. *et al.* Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. *Nat Cell Biol* **8**, 1163–1169 (2006).

50. Fang, Y.-S. *et al.* Full-length TDP-43 forms toxic amyloid oligomers that are present in frontotemporal lobar dementia-TDP patients. *Nat Commun* **5**, 4824 (2014).

51. Shodai, A. *et al.* Aberrant Assembly of RNA Recognition Motif 1 Links to Pathogenic Conversion of TAR DNA-binding Protein of 43 kDa (TDP-43). *Journal of Biological Chemistry* **288**, 14886–14905 (2013).

52. Carey, J. L. & Guo, L. Liquid-Liquid Phase Separation of TDP-43 and FUS in Physiology and Pathology of Neurodegenerative Diseases. *Front Mol Biosci* **9**, (2022).

53. Oldani, E. G. *et al.* Manipulating TDP43 Aggregation via RNA G-quadruplexes. Preprint at https://doi.org/10.1101/2024.04.30.591873 (2024).

54. Cammas, A. & Millevoi, S. RNA G-quadruplexes: emerging mechanisms in disease. *Nucleic Acids Res* gkw1280 (2016) doi:10.1093/nar/gkw1280.

55. Fay, M. M., Anderson, P. J. & Ivanov, P. ALS/FTD-Associated C9ORF72 Repeat RNA Promotes Phase Transitions In Vitro and in Cells. *Cell Rep* **21**, 3573–3584 (2017).

56. Oyama, R. *et al.* Protein–protein interaction analysis by C-terminally specific fluorescence labeling and fluorescence cross-correlation spectroscopy. *Nucleic Acids Res* **34**, e102–e102 (2006).