bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Hetero-oligomerization of TDP-43 carboxy-terminal fragments with cellular proteins contributes to proteotoxicity

Akira Kitamura[1*], Ai Fujimoto[1], Rei Kawashima[1], Yidan Lyu[1], Kanami Moriya[1], Ayumi Kurata[1], Kazuho Takahashi[1], Reneé Brielmann[2], Laura C. Bott[2], Richard I. Morimoto[2], Masataka Kinjo[1*]

1 Laboratory of Molecular Cell Dynamics, Faculty of Advanced Life Science, Hokkaido University, Sapporo, Japan.
2 Department of Molecular and Cell Biology and Biochemistry, Rice Institute for Biomedical Research, Northwestern University, Evanston, IL, USA.
* To whom correspondence should be addressed

Abstract

Carboxy terminal fragments (CTFs) of TDP-43 contain an intrinsically disordered region (IDR) and form cytoplasmic condensates containing amyloid fibrils. Such condensates are toxic and associated with pathogenicity in several neurodegenerative disorders, including amyotrophic lateral sclerosis and frontotemporal lobar degeneration. However, the molecular details of how the domain of TDP-43 CTFs leads to condensation and cytotoxicity remain elusive. Here, we show that truncated RNA/DNA-recognition motif (RRM) at the N-terminus of TDP-43 CTFs is assembled and leads to the structural transition of the IDR, whereas the IDR itself of TDP-43 CTFs is difficult to assemble even if they are proximate intermolecularly. Hetero-oligomers of TDP-43 CTFs that have recruited other proteins that are essential for proteostasis into low mobile condensates are more toxic than homo-oligomers inside highly mobile condensates, implicating loss-of-function of the endogenous proteins by such oligomers, not necessarily the condensates, is associated with cytotoxicity. Furthermore, such toxicity of TDP-43 CTFs was cell-nonautonomously affected in the nematodes. We speculate that the misfolding and oligomeric characteristics of the truncated RRM at the N-terminus of TDP-43 CTFs define their condensation properties and toxicity by implanting and transmitting structures with toxic properties of the truncated RRM into the IDR.

Keywords: Condensates, TDP-43, amyloid, aggregation, proteostasis

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Introduction

Protein aggregation is a principal phenomenon in which misfolded proteins and intrinsically disordered proteins (IDPs) assemble under certain conditions. Membraneless intracellular assemblies in subcellular compartments of cells and tissues are called protein aggregates or condensates, which are formed by liquid-liquid or liquid-solid phase separation, respectively, and have the ability to concentrate biopolymers (Alberti & Hyman, 2021). Although various IDPs contain intrinsically disordered regions (IDRs), IDRs are known to play functional roles that diverge from those of folded proteins/domains and revolve around their characteristics to act as hubs for protein-protein interactions because IDR conformations are dynamically and rapidly interconverting between free and bound states (Dunker *et al*, 2008; Wu *et al*, 2017). Furthermore, aggregation propensities are not involved in IDRs but domains of protein except IDRs (Uemura *et al*, 2018). However, protein misfolding is known to occur in various situations co- and post-translationally as a result of cell state transitions during stress (Hartl, 2017; Labbadia & Morimoto, 2015). Misfolded proteins are recognized by various cellular mechanisms for protein quality control and can be either refolded by molecular chaperones, or targeted for degradation by the ubiquitin-proteasome system or autophagy-lysosomal pathways to maintain protein homeostasis (proteostasis) in cells (Hartl, 2017; Klaips *et al*, 2018; Labbadia & Morimoto, 2015). If protein quality control fails, misfolded proteins accumulate as aggregates or condensates *in vivo* and *in vitro*. Therefore, it is thought that protein misfolding and conformational transition of IDRs are independent physicochemical properties. Biomolecular condensates including protein and other biomolecules such as RNA are defined as membraneless organelles, which have specialized functions in cells; also, it is then summarized as a new name for cellular inclusion bodies, aggregates, membraneless organelles, and puncta (Banani *et al*, 2017). Condensates also form as a physiological process, for example during cell stress, or in development, even though they are not diseases; however, they are also frequently found in the cytoplasm and nucleus in neurons, muscles, etc., when in neurodegenerative diseases and a certain myopathy (Ahmed *et al*, 2016; Ross & Poirier, 2004). Since one of the cardinal features of neurodegenerative diseases is known to be the presence of protein toxicity (also called proteotoxicity), it is suspected that the toxicity of aggregation-prone proteins that form condensates is associated with neurodegenerative diseases such as amyotrophic lateral sclerosis (ALS), frontotemporal lobar degeneration (FTLD), Alzheimer′s disease, and Parkinson′s disease (Chung *et al*, 2018; Taylor *et al*, 2002). Such disease-associated condensates are deposited inside and outside the cell, and mature and deposited protein aggregates of such disease-associated proteins (e.g., α-synuclein, Aβ,

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

polyglutamine proteins, etc.), called amyloids, are initially thought to be a proteotoxic cause; however, in recent years, their soluble aggregates, the smaller of which are usually called oligomers and protofibrils, have been reported to be highly proteotoxic species (Benilova *et al*, 2012; Kitamura *et al*, 2006; Prots *et al*, 2018). Thus, how oligomers are formed and whether they are toxic are important questions.

More than 97% of both sporadic and familial ALS, transactive response element (TAR) DNA/RNA-binding protein 43 kDa (TDP-43)-positive condensates in the cytoplasm of the affected neurons are featured (Ling *et al*, 2013). TDP-43 condensates are also the hallmark of some forms of frontotemporal dementia (FTD). Such FTDs are classified as FTLD with TDP-43 pathology (FTLD-TDP) (Bang *et al*, 2015). FTLD-TDP subtypes (typically, type A–D) are likely to be distinguished by the type of insoluble proteins in addition to TDP-43 (Laferriere *et al*, 2019). In ALS and FTLD pathology, not only full-length TDP-43 but also TDP-43 CTFs are included in the condensates in the brain, but are rarely detected in the spinal cord (Neumann *et al*, 2006; Smethurst *et al*, 2016). Therefore, it has been proposed that TDP-43 CTFs may not be a prerequisite for neurodegeneration (Berning & Walker, 2019; Lee *et al*, 2011); however, TDP-43 fragmentation and its CTFs are a potential therapeutic target because ectopic expression of TDP-43 CTFs leads to various dysfunctions such as impaired autophagy and proteasome function, reducing functional dimerization/oligomerization of TDP-43, reduced motor function, and cell death, in cellular models and also in animal models (Caccamo *et al*, 2015; Chhangani *et al*, 2021; Fazal *et al*, 2021; Kitamura *et al*, 2016; Wang *et al*, 2015). Consequently, the proteotoxic mechanism of TDP-43 CTFs remains elusive.

Human TDP-43 (hTDP-43) carries an N-terminal ubiquitin-like domain, two RNA/DNA-recognition motifs (RRM1 and RRM2; hTDP-43 106–176 and 191–262, respectively), and a C-terminal glycine-rich region (GRR; hTDP-43 274–414) that is known to be a prion-like domain (Bolognesi *et al*, 2019), including a glycine-rich (hTDP-43 274–310), a hydrophobic patch (318–343), the Q/N-rich (344–360), and glycine/serine-rich domains (hTDP-43 368–414), in addition to several amyloidogenic core regions (Guo *et al*, 2011; Jiang *et al*, 2013). The GRR is also known to be an IDR, which regulates interactions with proteins in other RNA-binding proteins (e.g., hnRNP A1/A2 and FUS/TLS) as well as the self-interaction of TDP-43 via liquid-liquid phase transition (Carey & Guo, 2022; Da Cruz & Cleveland, 2011). hTDP-43 contains typical DEVD-like motifs in regions 86–89 and 216–219 in the primary sequence, which can be cleaved by caspase 3 (DETD and DVMD sites, respectively). hTDP-43 CTFs, 90–414 (35 kDa) and 220–414 (25 kDa), are called TDP35 and TDP25, respectively (Zhang *et al*,

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

2009). Caspase 3 is not the case that produces TDP35 and TDP25, but calpain, alternative splicing, and protein clearance pathways are also responsible for producing such CTFs (Berning & Walker, 2019; Xiao *et al*, 2015; Yang *et al*, 2014). The assembly properties and characteristics of TDP25 condensates have been described (Berning & Walker, 2019). The TDP-43 condensates in the cytoplasm and the nucleus are associated with RNA granules (Loganathan *et al*, 2019); however, the TDP25 condensates did not contain RNA (Kitamura *et al.*, 2016). Furthermore, the structures of the pathological filaments of the GRR were identified (Arseni *et al*, 2022; Li *et al*, 2021); however, it remains unknown how TDP25 forms condensates in the cytoplasm. In the yeast model, although the cytoplasmic condensates increased according to how many regions of the RRM2 of hTDP-43 were included, the GRR itself did not form condensates despite that region 311–320 is known to be the most critical for the amyloidogenic fibrils (Johnson *et al*, 2008; Saini & Chauhan, 2011). Although intact RRM2 is made up of two α-helixes (α1–α2) and five β-strands (β1–β5), β3 and β5 were reported to form non-amyloidogenic fibrils *in vitro* (Wang *et al*, 2013). These reports implicate that the N-terminal conformational change in the truncated RRM2 (tRRM2) of TDP25 could lead to aggregation; however, it is still speculative.

To discover the pathogenic characteristics of TDP-43 CTFs, many model organisms have been established using mice, rats, fruit flies, and worms in addition to tissue culture cells (Berning & Walker, 2019; Chhangani *et al.*, 2021). The nematode worm *Caenorhabditis elegans* has been instrumental to uncover tissue and cell dysfunction and dysregulation of longevity caused by misfolded proteins during especially aging because the lifespan is relatively short compared to mice and rats (Ben-Zvi *et al*, 2009; Nillegoda *et al*, 2015; Vecchi *et al*, 2020). Pan-neuronal expression worm lines of yellow fluorescent protein (YFP)-tagged TDP25 showed significant movement defects associated with protein condensates in the neuronal cell bodies (Zhang *et al*, 2011). On the other hand, to demonstrate proteotoxicity of misfolded proteins associated with human neurodegenerative disorders such as Aβ, expanded polyglutamine proteins, and superoxide dismutase 1 (SOD1), their expression in the body-wall muscle has been widely used (Lee *et al*, 2017; Morley *et al*, 2002; Nussbaum-Krammer & Morimoto, 2014; Sinnige *et al*, 2021). Although studying the cell-nonautonomous effects of such misfolded proteins in neurons (Brignull *et al*, 2006; Nussbaum-Krammer & Morimoto, 2014), there are few reports to demonstrate whether the proteotoxicity of misfolded proteins in the body-wall muscles affects the lifespan (Lee *et al.*, 2017). We reasoned that a new model that expresses TDP25 could be informative for elucidating proteotoxicity during ageing by misfolded protein oligomers and their condensates formed through a

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

liquid-liquid phase transition. We introduced highly toxic TDP25 and further truncated TDP25, which is a highly oligomerized but low toxic variant in murine cells, into *C. elegans* body-wall muscle and analyzed the oligomerization status and protein mobility of TDP-43 variants in the condensates, as well as animal motility and lifespan.

Taken together, using murine neuroblastoma Neuro2a cells and *C. elegans*, we demonstrate the role of misfolding of tRRM2 and GRR as IDR of TDP-43 CTF during their condensation containing amyloidogenic aggregation by analyzing the efficiency of cytoplasmic condensation, molecular fluidity in condensates, and oligomeric states in lysate, of various TDP-43 CTFs. Furthermore, we investigated the proximity properties of the N- or C-terminal side of TDP-43 CTFs using a photo-induced oligomerization tag.

Results:

Trunctaed RRM region at the N-terminus of TDP25 specifies the condensation efficiency.

Although TDP25 (amino acid 220–414; hereafter called C220) is aggregation-prone (Kitamura *et al.*, 2016), it remains elusive how C220 aggregation are deposited in the cells. We first confirmed aggregation propensity of TDP25, C263 (amino acid 263–414), and C274 (amino acid 274–414) tagged with a monomeric green fluorescent protein (GFP) at their N-terminus in murine neuroblastoma Neuro2a cells by observing distinct brighter structures than those are distributed in the cytoplasm (hereafter called condensates) using fluorescence microscopy. C220/TDP25 frequently formed cytoplasmic condensates, but GRR did not (Supplemental Figure S1). Therefore, we hypothesize that the sequence of truncated RRM2 on the N-terminal side of the GRR specifies the aggregation and condensation propensity of the TDP-43 CTFs so that the condensate formation of GFP-tagged TDP43 CTF variants removing 13, 23, and 33 amino acid residues from the N-terminal side of C220 (C233, C243, and C253, respectively; shown in Figure 1A) in Neuro2a cells was analyzed using fluorescence microscopy and the solubility assay followed by western blotting of their cell lysates. C233, C243, and C253 formed condensates in the cytoplasm (Figure 1B). However, the proportion of cells containing C233 condensates was extremely low, while that of C243 and C253 was higher than C233 and C263 but lower than C220 (Figure 1C). C233, C243, and C253 were observed in the 1% SDS-insoluble pellet fraction, but their abundance in the pellet fraction was lower than that of C220 (Figure 1D). For C233, more than half was recovered in the soluble fraction. Almost C263 was recovered in the soluble fraction. It suggests that the sequence on the N-terminal side of GRR is responsible for the cytoplasmic condensation of TDP-43 CTFs. The proportion of dead cells expressing

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

TDP-43 CTFs showed a positively correlative relation with that of cells containing condensates (Figure 1, C & E), suggesting that the condensates of TDP-43 CTFs are responsible for cytotoxicity. Cells expressing C233 and C263 did not show an increase in cell death compared to the expression of GFP monomers as a negative control (Figure 1E), suggesting that C233 and C263 are not cytotoxic.

To compare the mobility and fluidity of C220 and C233 in the cytoplasmic condensates, we performed fluorescence recovery after photobleaching (FRAP) in the TDP-43 CTF-expressing Neuro2a cells (Figure 1F). Because the movement of fluorescent molecules during the photobleaching time reduces the apparent photobleaching efficiency, the normalized relative fluorescence intensities (nRFI) that were corrected by the photobleaching efficiency of immobile GFP were analyzed. The nRFI of C233 just after photobleaching (at 0 s in Figure. 1F) were higher than those of C220 (Figure 1G). The increase in nRFI of C233 just after the photobleaching suggests that a portion of C233 (7.6%; Figure 1G) could be more rapidly associated and dissociated in the condensates than C220. However, there was no difference in the shape of the recovery curve, except for the nRFI value at 0 s between C220 and C233 (Figure 1F), indicating that the fluidity of C220 and C233 in the condensates may be the same.

Next, the oligomerization states of such GFP-tagged TDP-43 CTFs in the 1% Triton X-100-soluble fraction of the cell lysates were analyzed using fluorescence correlation spectroscopy (FCS), which can determine the homo-oligomerization of soluble and diffuse fluorescent species from the fluorescence fluctuation that occurred by passing fluorescent molecules through the confocal detection volume (Kitamura *et al.*, 2006; Kitamura *et al.*, 2016). Photon bursts (also called spikes) in TDP-43 (FL), C220, C233, C243, and C253 were observed during photon count rate records, while not in C263, C274, and GFP (Figure 2A). Counts per particle (CPP), which indicates fluorescence brightness per a single particle and homo-oligomeric states, of C233 was significantly higher compared to other CTFs, FL, and GFP monomers (Figure 2B), suggesting that C233 may be efficiently homo-oligomerized. In contrast, CPPs of other TDP-43 CTFs and FL were not dramatically higher than that of GFP monomers as a control (Figure 2B), suggesting that they may be not so efficiently homo-oligomerized. Next, diffusion time (DT), which corresponds to the passing speed of fluorescent molecules in the detection volume and indicates their apparent molecular weight change, of TDP-43 and its CTFs increased compared to the GFP monomer control (Figure 2C), indicating that TDP-43 CTFs form significantly higher molecular weight oligomers than the GFP monomers in the soluble fraction. The diffusion time of C253, C263, and C274 was slightly slower compared to other CTFs and showed a tendency to be higher than that of GFP monomers. Since the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

apparent molecular weight is assumed by the cube of the diffusion time ratio to the GFP monomer (27 kDa; see Eq. 2), that of C220, C233, and C243 was approximately 3.1 to 3.5 MDa, and that of C253, C263, and C274 was 0.8 to 1.7 MDa. Therefore, C220, C233, and C243 form larger soluble oligomers among FL and TDP-43 CTFs; however, C220 preferentially form hetero-oligomers containing many endogenous molecules, while C233 tends to form relatively homogenous oligomers. Furthermore, these data did not rule out the existence of hetero-oligomers of C233 containing other cellular proteins, because the apparent molecular weight of C233 estimated from the diffusion time was much larger than that of the tetramer, while the CPP of C233 is only about four-fold higher than that of GFP monomers. Therefore, the conformation of the N-terminal region outside of GRR in TDP-43 CTFs may define the structural characteristics for the formation of oligomers with endogenous proteins.

The 220–262 region of TDP-43 is a nucleation domain and incorporates its C-terminal glycine-rich region during its condensation.

To demonstrate whether the N-terminal region defines the condensation propensity of C220, we prepared 220–262 amino acids tagged with GFP, a C-terminal portion of RRM2, of TDP-43 (F220) (Figure 3A). F220 condensates were observed in the cytoplasm in Neuro2a cells (Figure 3B, *left*), while intact RRM2 condensates were not (Figure 3B, *middle*), revealing that F220 itself tends to condensate. However, the population of cells containing cytoplasmic condensates of F220 was lower than that of C220 (Figure 3C). To demonstrate whether the position of F220 relative to GRR could be involved in the condensation propensity, we constructed a mutant in which F220 was fused to the C-terminal of GRR (SW). When expressed in Neuro2a cells, we found that this construct formed cytoplasmic condensates (Figure 3B, *right*); however, the population of cells containing these condensates was not different from that of F220 (Figure 3C). A portion of F220 and SW was recovered in the SDS-insoluble fraction (Fraction P), while the RRM2 and GFP monomers alone were not (Figure 3D). Consequently, F220 is an aggregation-prone peptide, and the presence of GRR on the C-terminal side of the F220 peptide promotes its aggregation propensity. Next, to determine whether the deposition of TDP-43 CTFs in the cytoplasmic condensates is involved in the formation of amyloid fibrils in them, the condensates were stained with an amytracker dye, an amyloid-binding fluorescent probe. Amytracker stained the cytoplasmic condensates of C220, C233, C274, and F220 (Figure 3E and Supplemental Figure S2); however, in a portion of the F220 condensates, the fluorescence intensity of amytracker was low (Figure 3E, *middle lines*), or the condensates stained with amytracker and not with it coexisted (Figure 3E, *bottom*

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

lines). Accordingly, F220 is prone to condensation but could have various types of structures (i.e., amyloids and non-amyloids). However, 1,6-hexanediol, a commonly used inhibitor of weak hydrophobic interactions, was unable to dissolve the cytoplasmic condensates of C220, F220, and C274 (Supplemental Figure S3), suggesting that, in the condensates, these CTFs accumulated tightly with intermolecular entanglements, regardless of amyloid fibril formation. Furthermore, FRAP analysis showed that the recovery proportion of F220 in the cytoplasmic condensates was dramatically higher than that of C220 (Figure 3F). This is because the nRFI of F220 just after photobleaching (at 0 s) was significantly higher than that of C220 (18.5% difference), but there were no significant differences in recovery speed 20 seconds and later between F220 and C220, suggesting a low fluidity of them in the condensates. These results suggest that, although a portion of F220 in the condensates is strongly bound and deposited as in C220, the other portion of F220 can move rapidly between the cytoplasm. Therefore, we come to have the hypothesis that TDP-43 GRR/IDR may play a key role in determining the character of deposited as the cytoplasmic condensates.

Next, we tested whether C220 and F220 could incorporate TDP-43 GRR into the cytoplasmic condensates by co-expressing both C220 or F220 tagged with GFP (G-C220 or G-F220) and C274 tagged with a red fluorescent protein (RFP) (R-C274) in Neuro2a cells. Confocal microscopy of these cells revealed that condensates of RFP monomers and R-C274 were not observed in cells expressing GFP monomers as negative controls (Figure 4A, ***a*** and ***b***). R-C274 was colocalized with G-C220 in the cytoplasmic condensates, but the RFP monomers were not (Figure 4A, ***c*** and ***d***). In particular, G-F220 was colocalized with R-C274, but bot with RFP monomers (Figure 4A, ***e*** and ***f***). Cells containing condensates of GFP monomers were not detected regardless of R-C274 expression (lanes 1 and 2 in Figure 4B). Cells containing G-C220 condensates were increased by co-expression of R-C274 compared to RFP monomers as a control (lanes 3 and 4 in Figure 4B), suggesting that C274 promotes condensate formation of C220. Meanwhile, cells containing G-F220 condensates did not increase much with the coexpression of R-C274 (lanes 5 and 6 in Figure 4B); however, R-C274-positive condensates were observed only when G-C220 or G-F220 was co-expressed (lanes 4 and 6 in Figure 4C). Although the cell population containing R-C274-positive condensates in G-C220-expressing cells was the same as that in G-F220-expressing cells (lanes 4 and 6 in Figure 4D); therefore, the efficiency of GRR incorporation of F220 is similar to that of C220. As a result, the 220–262 amino acid region of TDP-43 may first assemble and then condensate with involving the GRRs, regardless of whether both the 220–262 region and GRR are included in a single polypeptide.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The hydrophobic region in the intrinsically disordered region of TDP-43 CTF drives protein condensation

Why is the proportion of cells containing C220 and F220 condensates dramatically increased by C274 co-expression, as shown in Figure 4B? The hydrophobic interaction is known to be important for the stabilization of the liquid condensates of TDP-43 (Krainer *et al*, 2021), and the hydropathy plot confirmed that hTDP-43 311–340 is a highly hydrophobic region (HR; Supplemental Figure S4). Nuclear magnetic resonance (NMR) analysis of hTDP-43 311–360 peptide in solution determined that α-helical structure (amino acids 321–340) is included in this HR despite being in the IDR (Zhang *et al*, 2012); however, 282–360 amino acid region including the helical structure is converted to cross-β amyloid filaments in the pathological structure from ALS with FTLD (Arseni *et al.*, 2022). The α-helical structure in GRR is proposed to be involved in homotypic interactions of TDP-43 CTFs in the condensates with tuning liquid-liquid phase transisition (Conicella *et al*, 2020; Conicella *et al*, 2016). Therefore, we determine whether the efficiency of condensate formation is characterized by the HR in GRR using a variant of G-C220 that lacks HR (C220Δ). Furthermore, because TDP-1, a TDP-43 orthologue of *C. elegans*, has a short C-terminal region compared to TDP-43 GRR and does not contain obvious HR (Zhang *et al.*, 2012), we used a chimeric protein that F220 was fused to the C-terminal region of TDP-1 (F220c) (Figure 5A). Although C220Δ and F220c formed the cytoplasmic condensates in Neuro2a cells (Figure 5B), the proportion of cells containing condensates of C220Δ and F220c was significantly lower than that of C220, and it was not different from that of F220 (Figure 5C). The population of dead cells expressing C220Δ and F220c was lower than that of C220, and it was not different from that of F220 (Figure 5D). There was a correlative relationship between cytoplasmic condensate formation and dead cell proportion. Next, to determine the solubility of such CTFs, the amount of the CTFs in SDS-soluble (S) and -insoluble (P) fractions of cell lysates was analyzed using SDS-PAGE followed by western blotting (Figure 5E). A portion of C220Δ was recovered in the P fraction as similar to F220, suggesting that the condensates of C220Δ in the cytoplasm may tightly assemble as aggregates but the assembly state is similar to F220. On the other hand, almost all F220c was recovered in the S fraction, confirming that the C-terminal region of TDP-1 of *C. elegans* does not have a role as HR, and suggesting that HR including α-helical structure in TDP-43 may be involved in the structural transition into tightly assembled aggregates.

To determine whether HR in the GRR is required for the molecular assembly of TDP-43 CTFs in the condensates, HR-deleted C274 (C274Δ) was prepared and analyzed

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

condensation properties (Figure 6A). Observation of the cytoplasmic condensation in Neuro2a cells using confocal fluorescence microscopy showed that, although colocalization of G-C220 and R-C274 in the cytoplasmic condensates was confirmed as shown in Figure 4, that of G-C220 and RFP-tagged C274Δ (R-C274Δ) was not (Figure 6B, ***a*** and ***b***). On the contrary, R-C274 was co-localized with GFP-tagged C220Δ (G-C220Δ) in the condensates, while RFP monomers as a negative control were not (Figure 6B, ***c*** and ***d***). The proportion of cells containing G-C220 condensates was slightly decreased when the co-expressing partners were changed from R-C274 to R-C274Δ (lanes 4 and 5 in Figure 6C), but still higher than that when RFP monomers were expressed (lanes 2 and 5, Figure 6C). However, R-C274Δ was not incorporated into the G-C220 condensates (lane 5 in Figures 6, D and E). This is because, since C274Δ does not contain HR but contains F220, this region may promote the C220 condensation through the intact GRR of the counter polypeptide. The SDS solubility assay showed that, although R-C274 was observed in the P fraction of the cell lysates when G-C220 was co-expressed, R-C274Δ in the P fraction was dramatically decreased and dim (lanes 6 and 8 in Figure 6F).

Next, we investigate whether C274 including HR could be incorporated into the condensates of C220 lacking HR (C220Δ). The cells containing G-C220Δ condensates were increased by co-expression of R-C274 compared to the expression of RFP monomers as a negative control (lanes 3 and 6 in Figure 6C), and this increase was more notable than the increase of the cells containing condensates of G-F220 by co-expression of R-C274 (lanes 5 and 6 in Figure 4B). Co-localization between G-C220Δ and R-C274 in the condensates was observed despite the low efficiency of condensate formation and their co-localization compared to co-expression with HR-including C220 (G-C220) (lanes 4 and 6 in Figure 6, D and E). The SDS solubility of these TDP-43 CTFs in the cell lysates showed that R-C274 was observed in the P fraction when G-C220Δ was co-expressed (lane 12 in Figure 6F). The amount of G-C220Δ was increased by co-expression of R-C274 compared to RFP monomers (lanes 9–12 in Figure 6F). Therefore, HR in GRR is required not only for the intra but also intermolecular assembly of TDP-43 CTFs in the condensates after the aggregation of its N-terminal side.

N- or C-terminal proximity analysis of TDP-43 CTFs in live Neuro2a cells using a blue light-induced oligomerization system.

To investigate the promotion of oligomerization of TDP-43 CTFs when the N-terminal side becomes proximal, a blue-light-induced homo-oligomerization tag engineered from a plant cryptochrome variant, CRY2olig (CRY), was used (Taslimi *et al*, 2014). CRY was

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

tagged to C220, C220Δ, C263, and C263Δ at their N-terminal end. To visualize the oligomerization in live cells, RFP was also tagged at the C-terminal end (CRY-C220-R, CRY-C220Δ-R, CRY-C263-R, and CRY-C263Δ-R, respectively), and expressed in Neuro2a cells. As a control, all cells expressing RFP-tagged CRY (CRY-R) formed foci in the cytoplasm and nucleoplasm by blue-light irradiation (Fig. 7A, ***a***; lane 1 in Fig. 7B). CRY-C220-R foci were not observed in all cells after the light irradiation; however, a low population of cells expressing HR-deleted C220 (CRY-C220Δ-R) formed foci after the light irradiation (Fig. 7A, ***b*** and ***c***; lanes 2 and 3 in Fig. 7B). On the contrary, almost all cells (>90%) expressing CRY-C263-R or CRY-C263Δ-R significantly formed foci after the light irradiation (Fig. 7A, ***d*** and ***e***; lanes 4 and 5 in Fig. 7B). These results suggest that the N-terminal side of TDP-43 GRR may be easy to close regardless of HR (Fig. 7C); however, C220 may have a steric hindrance with difficulties in intermolecular proximity at its N-terminus.

Next, to investigate the promotion of oligomerization of TDP-43 CTFs when the C-terminal side becomes proximal, CRY-R was tagged to C220, C220Δ, C263, and C263Δ at their C-terminal end (C220-CRY-R, C220Δ-CRY-R, C263-CRY-R, and C263Δ-CRY-R, respectively). As a negative control, a variant tagged with GFP on the N-terminal side of CRY-R (G-CRY-R) was used. G-CRY-R formed foci in all cells similar to CRY-R after the light irradiation (Fig. 7A, ***f***; lane 6 in Fig. 7B). Similarly to the N-terminal induction of oligomerization in C220, C220-CRY-R foci were not observed entirely after the light irradiation; however, a low population of cells expressing HR-deleted C220 (C220Δ-CRY-R) formed foci after the light irradiation (Fig. 7A, ***g*** and ***h***; lanes 7 and 8 in Fig. 7B). Unlike CRY-C263-R or CRY-C263Δ-R, the foci-forming population of the cells expressing C263-CRY-R and C263Δ-CRY-R was dramatically low; however, C263Δ-CRY-R showed significantly increased foci-forming cells compared to C263-CRY-R (Fig. 7A, ***i*** and ***j***; lanes 9 and 10 in Fig. 7B). These results suggest that the C-terminal side of TDP-43 GRR, unlike its N-terminal side, may have a steric hindrance with difficulty in intermolecular proximity (Fig. 7D). Furthermore, HR deletion could weaken the steric hindrance of the C-terminal side of GRR, suggesting that HR and α-helical structure in it may help to maintain a relatively rigid conformation even in the IDR.

Condensation properties of GFP-tagged C220 and C233 in nematodes

To demonstrate proteotoxicity of TDP-43 CTFs, we established a new *C. elegans* model expressing G-C220 or -C233 in body-wall muscle. Unlike Neuro2a cells, G-C233 condensates were observed in all fluorescence-positive worms as same as G-C220 on both 4 and 8 days after synchronization by bleaching (hereafter, days 1 and 5 of adulthood,

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

respectively), while GFP monomers did not form any condensates (Figure 8A). The age-dependent change in expressing amounts of GFP monomers, G-C220, and G-C233 was analyzed using western blotting (Figure 8B). The expression levels of GFP monomers did not change from day 1 to 5 of adulthood. However, the amount of G-C220 increased, whereas that of G-C233 did not, from day 1 to 5 of adulthood. Furthermore, high molecular weight species of G-C220 and G-C233 were emerged in day 5, and it was more prominent in G-C220 than G-C233, suggesting that tightly assembled oligomers of G-C220 and G-C233 may be accumulated in the worms as ageing.

The fluidity of G-C220 and G-C233 in the condensates in live worms was compared using FRAP. On both days 1 and 5 of adulthood, the maximum recovery of nRFI of GFP monomers as a freely mobile control was ~100% (Figure 9A). The maximum recovery of nRFI of G-C220 was less than 20%, while that of C233 was more than 75% fluorescence recovery within 270 s (Figure 9A), suggesting that the fluidity of C233 in the condensates was high but not C220. The nRFI just after photobleaching showed that a portion of G-C233 was highly liquid-like and rapidly moving in the condensates compared to G-C220 on days 1 and 5 of adulthood (Supplemental Figure 5). Next, to determine the oligomeric states of G-C220 and G-C233 in the worms, CPP and diffusion time using FCS of them were analyzed in the soluble fraction of worm lysates (Figure 9, B and C). The CPP of G-C220 on day 1 was similar to that of GFP monomers but dramatically increased on day 5, suggesting that G-C220 may become to form homo-oligomers on day 5 of adulthood. On the contrary, the CPP of G-C233 was significantly high on both days 1 and 5 compared to GFP monomers and G-C220 on day 1, and its large variance on day 1 became smaller on day 5, suggesting that G-C233 may form homo-oligomers on both days 1 and 5, but from day 1 to 5, the stoichiometry of C233 may converge to more uniform one. Diffusion times of G-C220 and G-C233 were significantly increased compared to GFP monomers on both days 1 and 5. Although the diffusion time of G-C233 was not different from day 1 to 5, that of G-C220 became dramatically longer. Consequently, G-C220 had a higher percentage of hetero-oligomers with endogenous other proteins than homo-oligomers on day 1 of adulthood; however, G-C220 on day 5 and G-C233 on days 1 and 5 of adulthood contained homo-oligomers significantly, although hetero-oligomers with endogenous proteins could be co-existed because diffusion times of G-C220 and -C233 are dramatically higher than the CPP ratio for GFP monomers as observed in the lysate of Neuro2a cells.

Comparison of proteotoxicity between C220 and C233 in the nematodes

To investigate the proteotoxic effects of G-C220 and G-C233 in the worms, locomotion

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

behavior and lifespan of worms were analyzed (Figure 10, A and B, respectively). The bending activity of the worms expressing G-C220 was low on day 1 compared to those expressing GFP monomers as a control, while that of G-C233 was slightly low but not prominent (Figure 10A). From day 1 to 5, the bending activity of worms expressing GFP monomers was increased, indicating the progression of locomotion behavior by their body growth as aging; however, such a progression was not observed in worms expressing G-C220 and G-C233. Moreover, on day 5, there is a contrast in the bending activities between G-C220 and G-C233 because that of G-C220 decreased compared to day 1 but G-C233 did not. Accordingly, both G-C220 and G-C233 affect locomotion behavior; however, G-C220 affects it more significantly.

The lifespan of worms expressing G-C220 and G-C233 (Median life expectancy: 8.4 and 9.0 days after synchronization by bleaching, respectively) was then significantly shortened compared to worms expressing GFP monomers as a control (Median life expectancy: 12.5 days after the synchronization) (Figure 10B). Although the mortality of the worms expressing G-C220 and G-C233 were similar during young to middle adults (from day 4 to 13 after the synchronization), that of G-C220 significantly decreased during old adults (after day 13 after the synchronization). Accordingly, C220 and C233 expression showed significant proteotoxic effects resulting in a cell-nonautonomous manner in the worms, but that of C233 was slightly mild compared to C220.

Discussion

TDP-43 CTFs are known to form condensates in the cell (Berning & Walker, 2019; Chhangani *et al.*, 2021). Here, we showed that the 220–262 region of tRRM2 on the N-terminal side of the GRR has an important role in the formation of the condensates. The intact RRMs of TDP-43 can be folded stably because the solution and crystal structures of recombinant RRM1, RRM2, and a tandem of them expressed in *E.coli* are determined (Afroz *et al*, 2017; Kuo *et al*, 2014; Lukavsky *et al*, 2013). A 35 kDa TDP-43 CTF containing two intact RRMs (RRM1 and RRM2) and GRR, TDP35 (amino acids 90–414), showed a low propensity for the cytoplasmic condensation and was not toxic because the cell death propensity of TDP35 was as same as that of TDP-43 (Kitamura *et al.*, 2016), in addition to, unlike TDP25/C220, no colocalization between the TDP35 condensates and ubiquitin, suggesting that it would not include misfolded one (Kitamura *et al.*, 2016). Consequently, a stably folded domain at the N-terminal side of the GRR could stabilize the conformation of the CTFs and then be less likely to form condensates containing misfolded ones. On the other hand, peptides in a portion of tRRM2 (amino acids 208–

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

265 and 234–273) are reported to be less stable and form amyloid fibrils *in vitro* (Shimonaka *et al*, 2016; Wang *et al.*, 2013). The cytoplasmic condensates of F220, the tRRM2 region of TDP25/C220, did not necessarily consist only of amyloid fibrils (Figure 3). Although RRM2 contains two α-helixes (α1–2) and five β-strands (β1–5), TDP25/C220 lacks α1, β1, and β2. After the lack of α1, β1, and β2 by the cleavage of RRM2, it is proposed that the remaining β-strands (β3, β4, and β5) in the RRM2 are abnormally exposed because β1 was sandwiched between β3 and β5 in the intact RRM2 (Wang *et al.*, 2013); then, β3 and β5 may be destabilized. As a result, it is thought that exposed β3 and β5 are thought to be lead to misfolding of tRRM2 and may provide an opportunity to form intermolecular assembly (e.g., oligomerization) (Figure 11A).

Amyloid fibril structures of the β-strand-transferred GRR of TDP-43 were identified in patients with ALS and FLTD patients and as recombinant ones using cryo-EM (Arseni *et al.*, 2022; Li *et al.*, 2021). How do such misfolded protein oligomers of TDP-43 CTFs form amyloid fibrils? We showed that the cytoplasmic condensates of GRR (C263 and C274) were quite rarely formed but those of C274 contained amyloid fibrils (Supplemental Figures S1 and S2). The GRR (C274) was incorporated into the condensation of F220 and also into that of C220Δ (Figures 4 & 6). Since C220Δ shares the region of F220 despite lacking the HR, the condensates were formed similarly to F220 (Figure 5). Optogenetic induction of intermolecular assembly using CRY2olig showed that the N-terminal end of GRR could be closed intermolecularly, while that of TDP25/C220 could not. On the other hand, the N-terminus of GRR (C263) allows it to be proximate, suggesting that, before the conformational changes of GRR affected by such as tRRM2, GRR structure may be flexible and is likely to accept to interact with other molecules depending on its conformation. Since the C-terminus of GRR is difficult to be proximate (Figure 7), the molecular assembly of GRR on the N-terminal side can lead to its conformational change, and then the amyloidogenic cross-β sheet transition can progress. tRRM2 may work not only as a driving force for intermolecular assembly of TDP-43 CTFs but also as an opportunity to change the structural transition of GRR, resulting in condensation including amyloid structures with progressive propagation involving the other chains of TDP-43 CTFs having non-amyloid conformations (Figure 11B). Furthermore, there may be some limiting conditions that predispose to the transformation of a disordered GRR conformation into amyloid fibrils in the neurodegenerative disorders: cleavage and unusual translation of TDP-43 to generate its CTFs that should be misfolded, accumulation of misfolded proteins.

Regarding the role of HR and α-helix in the GRR, since the intermolecular assembly of C263 from the C-terminal side was occurred more when HR was lacking

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

(Figure 7), HR may have a role in maintaining the IDR conformation to inhibit intermolecular proximity; conversely, the helix would undergo a conformational change with a rate-determining step, rather than immediately transitioning to amyloid structures just after the intermolecular proximity. The chimera protein of F220 and the C-terminus of TDP-1 of *C. elegans* showed the same tendency for the formation of the cytoplasmic condensate formation as F220 (Figure 5), suggesting that the TDP-43 GRR may be acquired during the evolution process to higher organisms and may play a role in the formation of protein complex through various interactions with other proteins (e.g., hnRNP A1/A2 and FUS/TLS) as well as the self-interaction of TDP-43 (Da Cruz & Cleveland, 2011).

However, another question arises here. As reported in the analysis of the pathological amyloid structure of TDP-43 described above, the core structure of the amyloid fibrils of TDP-43 consists of only polypeptide chains of the GRR. Our analysis using FCS suggests that, on average, many TDP25/C220 form heterogeneous oligomers with endogenous proteins in the soluble fraction of the cell lysates, but the existence of homo-oligomers is not denied because highly bright molecules, likely corresponding to highly homogeneous oligomers, were observed during measurements (Figure 2A). The highly homogeneous oligomers would primarily accumulate as amyloid fibrils, resulting in amyloid fibrils-containing condensates, whereas the heterogeneous oligomers including various endogenous proteins would remain in the cytoplasm with the maintenance of the diffusing state. As a result, the heterogeneous oligomers of TDP25/C220 can lead to loss-of-function of various endogenous proteins that are incorporated into those oligomers and then would dysregulate cellular proteostasis. On the other hand, since there are many reports to show the toxicity of amyloid fibrils (Espargaro *et al*, 2016; Fang *et al*, 2014), diffuse homogenous oligomers including amyloid fibrils would also produce cytotoxicity not by loss-of-function but, for example, gain-of-function manner.

C233 formed homogenous oligomers the most efficiently among TDP43 CTFs that were examined here (Figures 2 and 9). Since C233 lacks more than half of β3 strands, β5 may form a metastable structure with β4, which could allow homo-oligomerization of C233, not likely to prefer the hetero-oligomerization with endogenous proteins. Since the condensation efficiency of C253, which only has β5 as tRRM2, was higher than that of C233 and C243, which includes also β4, the β5 probably becomes the conformation unstable. Since β4- and β5-linked peptide (246–258) forms amyloid-like fibrils, while the peptide containing α2 (233–245) did not (Saini & Chauhan, 2011), the region containing α2 can prevent structural changes and misfolding of tRRM2. C233 includes α2, and the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

side chain of Gln 241 in α2 and Asn 259 in β5 forms electrostatic interaction between amine and carbonyl groups in intact RRM2 observed in a determined structure (PDB#3d2w). Since this electrostatic interaction between Gln 241 and Asn 259 exists near the tip of the U-shaped chain formed by α2, β4, and β5 (Supplemental Figure 7), this electrostatic interaction may suppress the dissociation of β4 and β5, likely contributing to the formation of stable homogenous oligomers. Furthermore, such a stable conformation of the C233 oligomers may be a possible reason why C233 formed fewer condensates in Neuro2a cells. Such stable C233 oligomers are diffusive, but less likely to interact with other endogenous proteins, resulting in low toxicity compared to TDP25/C220. Consequently, stabilization of misfolded tRRM2 is important to prevent the production of toxic oligomers; it is thus considered to be a major target for controlling and preventing the aggregation of TDP43 CTFs containing tRRM2. Since a cytoplasmic molecular chaperone, 70 kDa heat shock protein (HSP70) prevents TDP25/C220 condensation through the direct interaction between them (Kitamura *et al*, 2018; Kitamura *et al*, 2017), HSP70 can recognize and stabilize the conformations of tRRM2.

In contrast, a striking difference in worms from the Neuro2a cells is that the condensates of C233 were frequently observed as same as that of TDP25/C220 (Figure 8A), and the fluidity of C233 in the condensates was dramatically higher than TDP25/C220 (Figure 9A). The soluble oligomers of C233 abundantly contain homogeneous ones, which is similar to Neuro2a cells (Figure 9, B & C). Homogeneous oligomers of C233 would dynamically move between the cytoplasm and the condensates. Protein condensates are formed as a result of stress such as heat shock and aging (Jawerth *et al*, 2020; Zhang *et al*, 2022). The biology of condensates is likely various because the condensates formed in response to acute stress such as heat shock are reversible (Alagar Boopathy *et al*, 2022); however, these processes could be perturbed in the presence of disease proteins and/or during ageing (Shin & Brangwynne, 2017). Thus, there would be a mechanism that makes condensates more likely in the body-wall muscle in worms compared to Neuro2a cells. Proteome stability and proteostasis maintenance are widely known to be crucial for the cellular function and lifespan of organisms, and it has been shown to be important for the maintenance of muscle cells and neurons. (Ben-Zvi *et al.*, 2009; Frumkin *et al*, 2014). These findings suggest that misfolded protein oligomers such as C220/TDP25 and C233 and their condensates composed of these oligomers would be toxic and make damage the body-wall muscles. Consequently, we speculate that the cytotoxicity in the body-wall muscles would be involved not in the condensates but the characteristics of the diffuse oligomers (e.g., heterogeneity). Furthermore, the cell-to-cell transmission of neurodegenerative disease-associated protein oligomers, such as TDP-43

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

and ALS-associated and aggregation-prone dipeptide repeat peptides transcribed from the *C9orf72* locus, has been demonstrated (Pirie *et al*, 2021; Westergard *et al*, 2016). In addition to these reports, one may speculate that dynamic oligomers are more easily transmitted among the cells than the static condensates. In the worms we analyzed, the fluorescence molecules were localized only in the BWMs, so that such cell-to-cell transmission of the oligomers would be in circumscribed organs; however, since the condensates were distributed in the body-wall muscles of GFP-positive areas, such transmission may contribute to the prominent form of the condensates especially in C233 in the body-wall muscles in worms compared to Neuro2a.

The worms expressing C220 and C233 in the body-wall muscles showed a shortening of their lifespan (Figure 10). A worm model that expresses polyglutamine-expanded mutant huntingtin, which causes the inherited neurodegenerative disorder Huntington's disease, in the body-wall muscles showed a significantly shorter lifespan (Lee *et al.*, 2017); however, expression of polyglutamine-expanded Ataxin-3 in the body-wall muscles did not affect the lifespan in this model (Christie *et al*, 2014). Although it is still unknown why such differences in the lifespan when the aggregation-prone proteins are expressed in the body-wall muscles, a possible explanation is that the expression level of such proteins would be relevant. Furthermore, the shortening lifespan is involved in the insulin/insulin-like growth factor-1 signaling (IIS) pathway, a major and well-conserved pathway that is involved in the longevity and the proteostasis maintenance in *C. elegans* across eukaryotes (Ben-Zvi *et al.*, 2009; Uno & Nishida, 2016). The key regulator of IIS in *C. elegans* is DAF-16, a FOXO transcription factor homolog (Kenyon *et al*, 1993), The major body-wall muscles health regulator Axin increased longevity through the AMPK signaling pathway and DAF-16 in the IIS (Mallick *et al*, 2020). The short lifespan of the worms expressing C220 and C233 in the body-wall muscless is likely involved in some defections in such signaling pathways maintaining the longevity. The worms pan-neuronally expressing TDP25 showed neurotoxicity (Zhang *et al.*, 2011); however, their lifespan has not been elucidated. Therefore, the worms that express C220 and C233, as we established, should be a useful model that shows the phenotype reflecting the aggregation characteristics and cytotoxicity of TDP-43 CTFs, and to analyze proteostasis dysregulation by misfolded protein and IDPs.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Materials & Methods

Construction of mammalian expression plasmids

To create expression plasmids for Neuro2a cells, DNA fragments of C220, C233, C243, C253, were amplified by PCR using the following primers (Hereafter, underlines of the primer sequences denote restriction enzyme sites for cloning): 5′-GTCAAGCTTCCACCATGGTCTTCATCCCCAAGCCATTCAGG-3′ (forward side for C220), 5′-GCTAAGCTTCGACATTTGCAGATGATCAGAT-3′ (forward side for C233), 5′-GCTAAGCTTCGCTTTGTGGAGAGGACTTGAT-3′ (forward side for C243), 5′-GCTAAGCTTCGATCAGCGTTCATATATCCAA-3′ (forward side for C254), and 5′-CGTGGGATCCCTAGTGATTCATTCCCCAGCCAG-f3′ (reverse side for all of TDP-43 CTFs including glycine-rich region). PCR products were digested and ligated into the HindIII and BamHI sites of pmeGFP-C1 (Kitamura *et al.*, 2016) containing cDNA for a monomeric variant of enhanced GFP that contains the mutation A206K (in this paper, it is named GFP) (pGFP-C220, -C233, -C243, and -C253). The GFP-tagged C263 and C274 expression plasmids were used as same as previously reported (pGFP-C263 and -C274, respectively) (Kitamura *et al.*, 2016). DNA fragments of F220 (220–262) and RRM2 (190–262) were amplified by PCR using the following primers: 5′-GCTAAGCTTCGAGAAAAGTGTTTGTGGGGCGCTG-3′ (forward side of RRM2), the same one for the forward side of C220 (forward side of F220), and 5′-CGTGGATCCTAAGGTTCGGCATTGGATATA-3′ (reverse side of F220 and RRM2). These fragments were inserted into the HindIII and BamHI sites of pmeGFP-C1 as described above (pGFP-F220 and -RRM2). To make a domain-swap mutant (SW), in which F220 was fused to the C-terminal of GRR, two DNA fragments were amplified by PCR using the following primers: 5′-GCTAAGCTTCGGGAAGATTTGGTGGTAATC-3′ (Forward side for GRR of SW), 5′-CAGGTCGACCATTCCCCAGCCAGAAGACTT-3′ (Reverse side for GRR of SW), 5′-CAGGTCGACGTCTTCATCCCCAAGCCATTC-3′ (Forward side for F220 of SW), and the same one for the reverse side of F220 and RRM2 (Reverse side for F220 of SW). These fragments were ligated via the SalI site and inserted into the HindIII and BamHI sites of pmeGFP-C1 as described above (pGFP-SW). To make an expression plasmid for C220Δ (220-414 lacking 311-341 of TDP-43), two DNA fragments were amplified by PCR using the following primers: the same one for the forward side of C220 (forward side of N-terminal fragment of C220Δ), 5′-CCCACCACCCATATTACTAC-3′ (reverse side of N-terminal fragment of C220Δ), 5′-GCCAGCCAGCAGAACCAGTC-3′ (forward side of C-terminal fragment of C220Δ), and the same one for the reverse side of C220 (reverse side of C-terminal fragment of

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

C220Δ). These fragments were phosphorylated by T4 polynucleotide kinase, and then directly inserted into pmeGFP-C1 (pGFP-C220Δ). To make an expression plasmid for C274Δ (274–414 lacking 311–341 of TDP-43), a DNA fragment was amplified by PCR using the following primers: 5′-GCTAAGCTTCGGGAAGATTTGGTGGTAATC-3′ (forward side of C274Δ) and 5′-CGTGGGATCCCTAGTGATTCATTCCCCAGCCAG-3′ (reverse side of C274Δ). The fragment was inserted into the HindIII and BamHI sites of pmeGFP-C1 as described above (pGFP-C274Δ). To prepare the plasmids for expression of TDP-43 CTFs tagged with an red fluorescent protein mCherry (RFP), required fragments for TDP-43 CTFs were obtained by digestion of plasmids described above at the HindIII and BamHI sites, and then inserted into pmCherry-C1 (pRFP-C220, -C274, and -C274Δ).

The synthetic DNA fragment for a chimera encoding amino acids 220–262 of human TDP-43 fused to the C-terminal region of TDP-1 (F44G4.4), a TDP-43 orthologue in *C. elegans* (GTCAAGCTTCGGTCTTCATCCCCAAGCCATTCAGGGCCTTTGCCTTTGTTAC
ATTTGCAGATGATCAGATTGCGCAGTCTCTTTGTGGAGAGGACTTGATCATTA
AAGGAATCAGCGTTCATATATCCAATGCCGAACCTAAGCACAATAGCAATAGA
TCAGTTGGCCCTGATTATGGCCTTCCAGCTGGCTACCGTAACCGCAGAGAAC
GTGATCGACCGGATAGACGACCGATTCAAAATGAAGCACCTCTGCCCATGCC
ATTCGTCCGCCCACCACAAGATTACTCATACCGTCAGCAAAATTCTCCCCTCG
AGAGGAGATACTGGGCACCTGGAGACTCGAGAGGACCAGGATGGCGGGATC
CGTC) was ordered in GeneArt gene synthesis services (Thermo Fisher, Waltham, MA, USA). The fragment was digested and inserted into the HindIII and BamHI sites of pmeGFP-C1 (pGFP-F220c).

To make plasmids for a homo-oligomerization variant of cryptochrome 2 (CRY2olig)-based oligomerization analysis, DNA fragments for N-terminally CRY2olig-tagged TDP-43 CTFs were amplified by PCR from plasmids containing required sequence as described above using the following primers: 5′-GCTCCCGGGCGGTCTTCATCCCCAAGCC-3′ (forward side of C220 and C220Δ), 5′-GCTCCCGGGCGAAGCACAATAGCAATAGAC-3′ (forward side of C263 and C263Δ), and 5′-CGTACCGGTGGGTGATTCATTCCCCAGCCA-3′ (reverse side of C220, C220Δ, C263, and C263Δ). The fragments were digested and inserted into mCherry-N1 at the SmaI and AgeI sites; then a fragment of CRY2olig region that was digested from pCRY2olig-mCherry (#60032; Addgene, Watertown, MA, USA) was inserted at the NheI and SmaI sites (pCRY-C220-, -C220Δ-, -C263-, and -C263Δ-R). DNA fragments for C-terminally CRY2olig-tagged TDP-43 CTFs were amplified by PCR from plasmids

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

containing the required sequence as described above using the following primers: 5′-GCAGCTAGCCACCATGGTCTTCATCCCCAAGCCAT-3′ (forward side of C220 and C220Δ), 5′- GCAGCTAGCCACCATGAAGCACAATAGCAATAGAC-3′ (forward side of C263 and C263Δ), and 5′- CGTCTCGAGCATTCCCCAGCCAGAAGACTTAG-3′ (reverse side of C220, C220Δ, C263, and C263Δ). The fragments were digested and inserted into pCRY2olig-mCherry at the NheI and XhoI sites (pC220-, C220Δ-, C263-, and C263Δ-CRY-R). DNA fragments for C-terminally CRY2olig-tagged meGFP were amplified by PCR from meGFP-C1 using the following primers: 5′-GCAGCTAGCCACCATGGTGAGCAAGGGCGAGGAG-3′ (forward side) and 5′-CGTCTCGAGCTTGTACAGCTCGTCCATGCCG-3′ (reverse side); then inserted into pCRY2olig-mCherry (Addgene) at the NheI and XhoI sites (pGFP-CRY-R).

Cell culture and transfection

Murine neuroblastoma Neuro-2a cells (#CCL-131; ATCC, Manassas, VA, USA) and maintained in DMEM (D5796; Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% FBS (12676029, Thermo Fisher), 100 U/mL penicillin G (Sigma-Aldrich), and 0.1 mg/ml streptomycin (Sigma-Aldrich) as previously described (Kitamura *et al.*, 2016). Cells (2.0×10^5) 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) at 1 day before transfection. The plasmid DNAs (1 μg for expression of TDP-43 CTFs; and the mixture of 0.1 μg pmeGFP-N1 and 0.9 μg pCAGGS for expression of GFP monomers) were transfected using Lipofectamine 2000 (Thermo Fisher). After 1 overnight incubation for transfection, the medium was exchanged for the flesh and maintained prior to microscopic observation. Before the cell lysis, the medium including the transfection reagent was removed, and the cells were washed in PBS at 25°C.

Cell lysis and solubility assay

Neuro2a cells were lysed and fractionated as previously reported (Kitamura *et al.*, 2016). Briefly, cells were lysed in a lysis buffer containing 25 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 1% Noidet P-40, 1% sodium deoxycholate, 0.1% SDS, 0.01 U/μL benzonase, and 1× protease inhibitor cocktail (Sigma-Aldrich). After the centrifugation (20,400 g, 20 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. Samples were applied to a 5-20% polyacrylamide gel (ATTO, Tokyo, Japan) and subjected to electrophoresis in an SDS-containing buffer. The proteins were transferred

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

to a PVDF membrane (Cytiva, Marlborough, MA). The membrane was incubated in 5% skim milk in PBS-T for background blocking for 1 h. After incubation with an anti-GFP antibody (#GF200, Nacalai tesque, Kyoto, Japan), an anti-mCherry antibody (#Z2496N, TaKaRa), or a horseradish peroxidase-conjugated anti-α-tubulin antibody (#HRP-66031, ProteinTech, Rosemont, IL, USA) in CanGet signal immunoreaction enhancer solution 1 (TOYOBO, Osaka, Japan), horseradish peroxidase-conjugated secondary antibodies that are appropriate for primary antibodies (#111-035-144 or #115-035-062; Jackson Immuno Research, West Grove, PA, USA) were inclubated in 5% skim milk in PBS-T. Chemiluminescence signals were acquired using a ChemiDoc MP imager (Bio-Rad, Hercules, CA, USA).

Confocal fluorescence microscopy of Neuro2a cells

Neuro2a cells were observed using an LSM 510 META (Carl Zeiss, Jena, Germany) with a C-Apochromat 40×/1.2NA W Korr. UV-VIS-IR water immersion objective (Carl Zeiss). GFP and mCherry were excited at 488 nm and 543 nm, respectively. Excitation beams were divided through HFT488/543. Fluorescence was divided through a dichroic mirror (NFT545) and corrected through a 505-530 nm band-pass filter (BP505-530) for GFP and a 585 nm long-pass filter (LP585) for mCherry. Pinhole was set to 71 and 82 μm.

Counting of dead cells

Neuro2A cells expressing TDP-43 CTFs tagged with GFP were prepared as described in the other experiments. On 1 day after the transfection, dead cells were stained with a 1.0 μg/mL propidium iodide (PI) solution (Thermo Fisher Scientific) for 5 min. Images of the GFP and PI channels were acquired using an LSM 510 META microscope through a Plan-Neofluar 10×/0.3NA objective at 37°C as previously described (Kitamura *et al.*, 2016). The percentage of dead cells was calculated from the number of PI-positive cells divided by the number of GFP-positive cells.

Amytracker staining

Neuro2a cells expressing TDP-43 CTFs tagged with GFP were fixed in 4% paraformaldehyde buffered with 100 mM Hepes-KOH (pH 7.5). After washing in tris-buffered saline (TBS), cells were stained with a far-red amyloid-specific fluorescent dye (Amytracker 680; Ebba Biotech, Stockholm, Sweden) diluted in PBS (1/500) for 90 min. The excess dye was washed in PBS. Fluorescence images were acquired using an LSM 510 META (Carl Zeiss) and a C-Apochromat 40×/1.2NA W Korr. UV-VIS-IR M27 water immersion objective lens (Carl Zeiss).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Hexanediol treatment of Neuro2a cells

Neuro2A cells expressing TDP-43 CTFs tagged with GFP were prepared as described in the other experiments and grown on a 35 mm glass base dish (IWAKI-AGC Technoglass). On 1 day after the transfection, medium was exchanged for flesh and maintained before microscopic observation; then, the cells were incubated for 1 h. Cells were observed using an LSM 510 META (Carl Zeiss) with a C-Apochromat 40×/1.2NA W Korr UV-VIS-IR M27 water immersion objective. Cells were treated with 6% (w/v) 1,6-Hexanediol (1,6-HD) (Tokyo Chemical Industry, Tokyo, Japan) on the microscope stage at 37°C with 5% CO_2. Confocal imaging of cells positive for the condensates was performed as described in the method for confocal microscopy. Z-stack images were acquired for 7 slices at 2.0 μm intervals. Time-lapse observation of 1,6-HD treated cells was performed until the cells were shrunk or swelled at 5 min. intervals (approx. for 25 min).

Blue-light induced oligomerization assay in live cells

Plasmids for the expression of CRY2olig-tagged proteins (pCRY2olig-R, pGFP-CRY-R, pCRY-C220-R, pCRY-C220Δ-R, pCRY-C263-R, pCRY-C263Δ-R, pC220-CRY-R, pC220Δ-CRY-R, pC263-CRY-R, and pC263Δ-CRY-R) were transfected into Neuro2a cells at 1 day before the experiments. After incubation for 24 h, the medium was exchanged to a normal growth medium. Cell selection for observation, blue-light irradiation, and imaging was performed using an inverted fluorescence microscope, Axioobserver Z1 (Carl Zeiss) combined with an LSM 510META (Carl Zeiss) and a C-Apochromat 40x/1.2NA W Korr. UV-VIS-IR water immersion objective. Cells expressing CRY2olig-tagged proteins were selected by ocular observation using fluorescence filter sets for mCherry (#49008, Chroma, Bellows Falls, VT, USA) with a mercury lamp. Light irradiation for oligomerization and observation was performed using the LSM. mCherry was excited at 594 nm. The excitation and emission lights were split using a beam splitter (NFT488/594). Fluorescence was collected through a 615 nm long-pass filter (LP615) after through a 545 nm dichroic mirror (NFT545). The pinhole diameter was 150 μm. After 1 frame observation, cells were irradiated at 488 nm (7.04 ± 0.2 μW; mean ± SD; n = 3); pixel dwell time: 0.64 μs; 3 iterations. After 488 nm irradiation, cells were observed at 20 s intervals. Cells forming foci within 2 min were counted. A hypothesis test for the difference in the population proportions was performed using a BellCurve for Excel software (Social Survey Research Information Co., Ltd., Tokyo, Japan).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Construction of worm expression plasmids

For the nematode strain, DNA fragments of meGFP-tagged C220 and C233 were amplified from plasmids pGFP-C220 and -C233, respectively, by PCR using the following primers: 5′-CTCGCTAGCAAAATGGTGAGCAAGGGCGAGGAGC-3′ (forward side) and 5′-GAGGGTACCTTAGTGATTCATTCCCCAGCCAGAAGAC-3′ (Reverse side). A DNA fragment of meGFP monomers was amplified by PCR using the following primers: 5′-CTCGCTAGCAAAATGGTGAGCAAGGGCGAGGAGC-3′ (forward side) and 5′- GAGGGTACCTTACTTGTACAGCTCGTCCATGCCGAG-3′ (reverse side). PCR products were digested and ligated into the NheI and KpnI sites of pPD30.38 containing the *myo-3* promoter and enhancer elements from the *unc-54* myosin heavy-chain locus (pPD30.38::meGFP-C220; pPD30.38::meGFP-C233; pPD30.38::meGFP) (Morley *et al.*, 2002).

C. elegans culture and synchronization

Nematodes were handled using standard methods and cultured at 20°C (Brenner, 1973). For the generation of transgenic animals, 100 ng/μl plasmid DNA encoding meGFP-tagged TDP-43 CTFs (pPD30.38::meGFP-C220, pPD30.38::meGFP-C233, and pPD30.38::meGFP) were linearized with PvuII and mixed with the same volume of 100 ng/μl pBluescript. The mixtures were microinjected into the gonads of adult hermaphrodite N2 animals (Caenorhabditis Genetics Center, Minneapolis, MN, USA). Transgenic animals containing extrachromosal arrays were selected and isolated as strains based on fluorescence in body-wall muscle cells (C220, C233, and GFP strains).

For nematode assays, eggs were recovered by treating young adult animals with alkaline hypochlorite solution containing 0.25 M NaOH and 20% NaOCl (effective chlorine concentration ~ 1%). The eggs were hatched overnight by gently shaking in M9 buffer at 20°C. The next day, the hatched larvae (1 day after synchronization by bleaching) were grown at 20°C on nematode growth medium (NGM) agar plates seeded with OP50 *E. coli* (Caenorhabditis Genetics Center) as a food source. GFP fluorescence-positive nematodes (3 days after synchronization by bleaching) were selected using a stereo-dissection microscope Stemi 305 Trino (Carl Zeiss) with an epi-illumination 470 nm light source (LED470MS-EPI; Optocode Corp., Tokyo, Japan), a 495 nm long passing glass filter (FGL495S; Thorlabs, Newton, NJ, USA), and a 550 nm short passing glass filter (SV0550; Asahi Spectra, Tokyo, Japan), and were transferred to new NGM plates seeded with OP50 using a nematode-picker (platinum wire). Until the following assays, the nematodes were maintained on NGM agar plates seeded with OP50 at 20°C and were transfered to new plates every day.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Confocal fluorescence microscopy of C. elegans

The nematodes were placed on a 1% agarose pad with one drop of 8 mM Levamisole (Tokyo Chemical Industry) and mounted in Halocarbon oil 700 (Sigma-Aldrich) with a No. 1S coverglass (Matsunami, Osaka, Japan). Mounted nematodes were observed using an LSM 510 (Carl Zeiss) with a Plan-Neofluar 10×/0.3NA objective (Carl Zeiss). GFP was excited at 488 nm. The excitation beam was divided through HFT 488. Fluorescence was selected through a 505-550 nm band-pass filter (BP505-550). Bright-field images were acquired using a transmitted light detector for LSM 510 in an inverted microscope Axiovert 100M (Carl Zeiss). The pinhole was set to 140 μm, The X- and Y-scanning sizes were each 1024 pixels, the zoom factor was set to 0.8×, and the pixel dwell time was 3.2 μs.

C. elegans lysis and western blotting

Synchronized nematodes expressing GFP monomers, G-C220, or G-C233 (n = 200) were corrected at 4 or 8 days after the synchronization by bleaching and washed in PBS at 20°C 3 times. After the nematodes were suspended in a lysis buffer containing 20 mM Hepes-KOH (pH 8.0), 420 mM NaCl, 0.2 mM EDTA, 25% glycerol, 0.5 mM DTT, and 1× protease inhibitor cocktail (Sigma-Aldrich), nematodes were frozen at -80°C and thawed for 3 times. After the addition of 1% SDS and 1 M urea, the suspensions were sonicated, then the supernatant was recovered after the centrifugation (20,400 g, 20 min., 25°C). After addition of SDS-PAGE sample buffer, the lysates were denatured at 98°C. The samples were applied to a 12.5% polyacrylamide gel and subjected to electrophoresis in an SDS-containing buffer. The proteins were transferred to a PVDF membrane (Cytiva). The membrane was incubated in 5% skim milk in PBS-T for background blocking for 1 h. After incubation with an anti-GFP antibody (#GF200, Nacalai tesque) or anti-α-tubulin antibody (#sc-32293, Santa Cruz Biotechbology, Dallas, TX, USA) in CanGet signal immunoreaction enhancer solution 1 (TOYOBO), an alkaline-phosphatase-conjugated goat anti-mouse IgG secondary antibody (#SA00002-1, ProteinTech) were incubated in 5% skim milk in PBS-T. Bands were stained in 5-bromo-4-chloro-3-indolyl-phosphate and nitro blue tetrazolium solution (Sigma-Aldrich). The membrane image was acquired using a digital steel camera EOS M6 mark II with a lens EF16-35 mm F4L IS USM (Canon Inc., Tokyo, Japan).

Motility analysis

To analyze the number of body bends, five or fewer nematodes at 4 or 8 days after the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

synchronization by bleaching were transferred with a worm-picker from growing NGM agar plates to bacteria-lacking ones to clean nematodes from bacteria. M9 buffer (1 ml) was gently poured from the edge into the NGM agar plates. After the incubation for 1 min, the bending worms were recorded for 30 s using a digital steel camera EOS kiss digital X3 (Canon Inc.) attached to a stereo-dissection microscope Stemi 305 Trino (Carl Zeiss). The *x*- and *y*-image sizes of the movie were 1920 and 1080 pixels, respectively; and the frame rate was 20 fps. Repeatedly, the bending worms were recorded, and movies for 36 animals were obtained. One body bend corresponds to the movement of the head region that thrashes from one side to the other and back to the starting position. The number of bends was counted by observation of the recorded movie.

Lifespan analysis

Synchronized nematodes (3 days after bleaching; L4 larvae) were transferred onto NGM agar plates seeded with OP50 *E. coli* and maintained at 20°C. The next day, each nematodes at 4 day after the synchronization (Day 1 of adulthood) were transferred to independent NGM agar plates seeded with OP50 after determining the viability of the worms using a stereo dissection microscope Stemi 305 Trino (Carl Zeiss). Worms that did not move when gently prodded and did not show pharyngeal pumping were marked as dead. Two independent trials with 30 worms were conducted. The percentage alive was calculated using all 60 worms of the two independent trials. The generalized Wilcoxon test (Gehan-Breslow method) was performed using a BellCurve for Excel software (Social Survey Research Information Co., Ltd., Tokyo, Japan).

FRAP

Photobleaching experiments were performed on an LSM 510 META system through a C-Apochromat 40×/1.2NA W Korr. UV-VIS-IR M27 objective lens (Carl Zeiss) as previously reported (Kitamura *et al.*, 2016; Kitamura *et al.*, 2017). GFP was excited and photobleached at 488 nm, and fluorescence was separated using a dichroic mirror (NFT545). GFP fluorescence was collected through a band-pass 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 each 512 pixels, and the GFP was photobleached at 488 nm, 100% transmission, for less than 1.92 s. The relative fluorescence intensity was measured using Fiji-ImageJ (Carl Zeiss) and then calculated as reported previously (Kitamura *et al.*, 2017). To obtain the photobleaching efficiency just after the photobleaching (the lowest fluorescence intensity in the photobleached region), cells were fixed in 4% paraformaldehyde (PFA)-containing 100

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

mM Hepes-KOH (pH 7.5) buffer for 19 h, followed by washing in TBS, and photobleaching of GFP in the fixed cells was performed under the same condition (n =10). Normalized relative fluorescence intensities (nRFI) were calculated by subtracting the mean photobleaching efficiency from the RFI. To calculate the maximum recovery proportion, recovery curves associated with the relative fluorescence intensity were fitted using a single component exponential recovery model in Origin 2021 software (OriginLab Corp., Northampton, MA, USA) as previously reported (Kitamura *et al.*, 2017).

The nematodes were placed on a 1% agarose pad with one drop of 8 mM Levamisole (Tokyo Chemical Industry) and mounted in Halocarbon oil 700 (Sigma-Aldrich) with a No. 1S coverglass (Matsunami). Mounted nematodes were observed using an LSM 510 (Carl Zeiss) with a C-Apochromat 40×/1.2NA Korr. water immersion objective (Carl Zeiss). GFP was excited at 488 nm. The excitation beam was divided through HFT 488. Fluorescence was selected through a 505-550 nm band-pass filter (BP505-550). The pinhole was set to 300 μm, and the zoom factor was set to 15×. X- and Y-scanning sizes were 512 and 256 pixels, respectively. GFP was photobleached at 488 nm, 100% transmission, for less than 1.68 s. After the photobleaching, images were acquired at 15 s intervals. The nRFI was obtained by the photobleaching experiment using methanol-fixed nematodes (n = 10) as described for Neuro2a cells.

FCS

Neuro2a cells transiently expressing GFP-tagged TDP-43 CTFs were lysed in a buffer consisting of 50 mM Hepes-KOH (pH 7.5), 150 mM NaCl, 1% Noidet P-40, and 1× protease inhibitor cocktail (Sigma-Aldrich). Supernatants of the lysates were recovered after centrifugation at 20,400 *g* for 10 min. at 4°C, and immediately followed by FCS measurement. For nematodes, in day 4 or 8 after the synchronization, nematodes (n = 200) were collected and shaken in PBS for 1 h at 20°C to remove *E. coli*. Nematodes were suspended in a lysis buffer containing 100 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 1% Triton X-100, 1× protease inhibitor cocktail (Sigma-Aldrich). The freeze-thawed nematodes were sonicated using an ultrasonicator. The supernatants were recovered after centrifugation of the sonicated samples at 20,400 g for 1 min at 4 °C and immediately followed by FCS measurement. FCS was performed using an LSM 510 META + ConfoCor2 (Carl Zeiss) with a C-Apochromat 40×/1.2NA W Korr. UV-VIS-IR water immersion objective lens (Carl Zeiss) as previously reported (Kitamura *et al.*, 2016). Photons were recorded for 300 s. Curve fitting analysis for acquired autocorrelation function was performed using ZEN 2009 software (Carl Zeiss) using a model for one-

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

component 3D diffusion with one-component exponential decay as the following equation (Eq.) 1.

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

where, $G(\tau)$ is an autocorrelation function of the time interval τ; τ_{D} is the diffusion time; N is the average number of particles in the confocal detection volume; T and τ_{T} are the exponential decay fraction and time, respectively; s is the structural parameter determined by measuring the standard fluorescent dye (Alexa Fluor 488) on the same day. Apparent molecular weights were calculated as a Stokes-Einstein relation-modified Eq. 2.

$$Mw_{\mathrm{s}} = \left(\frac{\tau_{\mathrm{D,S}}}{\tau_{\mathrm{D,GFP}}}\right)^3 Mw_{\mathrm{GFP}} \quad \text{Eq. 2}$$

where, Mw_{s} and Mw_{GFP} (= 27 kDa) are the molecular weights of the sample and GFP monomers, respectively; $\tau_{\mathrm{D,\,S}}$ and $\tau_{\mathrm{D,\,GFP}}$ are the diffusion times of the sample and GFP monomers, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Acknowledgments

We thank Rintaro Kawamura for insightful discussions and Sue Fox for technical support. The authors acknowledge Open Facility Division, Global Facility Center, Creative Research Institution, Hokkaido University for use of a chemiluminescence imager. A.K. was supported by a Japan Society for Promotion of Science (JSPS) Grant-in-Aid for Scientific Research on Innovative Areas "Information physics of living matters" (22H04826); by a JSPS Grant-in-Aid for the Promotion of Joint International Research (Fostering Joint International Research) (16KK0156); by a JSPS Grant-in-Aid for Scientific Research (C) (18K06201); by a grant from Hokkaido University Office for Developing Future Research Leaders (L-Station); by a grant from Canon Foundation; by a grant from Hoansha Foundation; by a grant from Hagiwara Foundation of Japan; and by a grant from Nakatani Foundation. M. K. was partially supported by a JSPS Grant-in-Aid for Scientific Research (B) (22H02578).

Author contribution

Akira Kitamura: Conceptualization; Supervision; Funding acquisition; Investigation; Data curation; Visualization; Methodology; Resources; Writing–original draft; Writing–review and editing.

Ai Fujimoto: Investigation; Visualization.

Rei Kawashima: Investigation; Methodology; Visualization.

Yidan Lyu: Investigation.

Kanami Moriya: Investigation; Visualization.

Ayumi Kurata: Investigation.

Kazuho Takahashi: Investigation.

Reneé Brielmann: Investigation; Methodology; Resources.

Laura C. Bott: Methodology; Writing–review and editing.

Richard I. Morimoto: Supervision; Writing–review and editing.

Masataka Kinjo: Supervision; Funding acquisition; Resources; Writing–review and editing.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure legends

Figure 1 Formation of the cytoplasmic condensates of TDP-43 carboxy terminal fragments (CTFs) in Neuro2a cells

(A) Primary structures and abbreviation of full-length TDP-43 (FL) and its carboxy-terminal fragments (CTFs): amino acids 220-414 (C220), 233-414 (C233), 243-414 (C243), 253-414 (C253), 263-414 (C263), and 274-414 (C274). The RRM and GRR denote the RNA/DNA-recognition motif and the glycine-rich region, respectively. (B) Confocal fluorescence and bright field microscopic images of Neuro2a cells expressing GFP monomers (Mo) and GFP-tagged CTFs. White arrows represent the cytoplasmic condensates. Bars = 5 μm. (C) The population of Neuro2a cells carrying cytoplasmic condensates. Bars: mean + 95% CI; Dots: raw values (n = 4). (D) Western blot stained using an anti-GFP and anti-α-tubulin antibody of cell lysates: 1% SDS-soluble (S) and insoluble (P). (E) The population of dead cells expressing TDP-43 CTFs tagged with GFP and GFP monomers (Mo) using a propidium iodide exclusion test. Bars: mean and 95% CI (n = 4). (F) Comparison of time-dependent normalized relative fluorescence recoveries (nRFI) of GFP-tagged C220 and C233 (magenta and green, respectively) in live Neuro2a cells using fluorescence recovery after photobleaching (FRAP). Dots and bars indicate the mean and 95% CI (n = 9). Inset values show mobile fraction from the maximum recovery of nRFI obtained by non-linear curve fitting analysis using a one-component exponential recovery model. (G) Comparison of nRFI just after photobleaching (t= 0) between C220 and C233. The highest and lowest points of the box represent SD, and its bars represent 95% CI. Dots represent raw values. Inset values show the mean. (C, E, and G) *p*-values above the bars and lines: one-way ANOVA with Tukey's test compared to GFP monomers as control and comparison between lines, respectively.

Figure 2 Analysis of soluble oligomers of TDP-43 CTFs using fluorescence correlation spectroscopy (FCS).

(A) Typical photon count rate records during the FCS measurement of GFP monomers (Mo), full-length TDP-43 (FL), and CTFs in Neuro2a cell lysate. The Gray bar on the *y*-axis shows the range of 400 kHz of photon count rate. (B) The plot of normalized counts per particle (CPP), and relative fluorescent brightness per single particle. Bars: mean + 95% CI; Dots: raw values (n = 4). (C) The weighted average of diffusion time was obtained by curve fitting analysis of the autocorrelation function using a two-component diffusion model. Bars: mean and 95% CI; Dots: raw values (n = 4). In *B* and *C*, *p*-values above the bars and lines: one-way ANOVA with Tukey's test compared to GFP monomers

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

as control and comparison between the lines, respectively.

Figure 3 Condensate formation and amyloidogenic properties of the 220–263 region of TDP-43 in Neuro2a cells

(A) Domain structure and abbreviation of TDP-43 CTFs: 220–263 (F220), 191–262 (RRM2), and a mutant in which the position of F220 and C274 was swapped (SW) in addition to C220 and C263. (B) Confocal fluorescence and bright field microscopic images of Neuro2a cells expressing F220, RRM2, and SW tagged with GFP. White arrows represent cytoplasmic condensates. Bars = 5 μm. (C) Plot of normalized counts per particle (CPP), relative fluorescent brightness per a single particle. ND: Not determined. Bars: mean and 95% CI; Dots: raw values (n = 3). (D) Western blot of 1% SDS-soluble (S) and -insoluble (P) cell lysates using an anti-GFP and anti-α-tubulin antibody. Right values (25, 37, and 50 kDa) represent positions of the molecular weight marker. (E) Confocal images Neuro2a cells expressing TDP-43 CTF tagged with GFP stained with a fluorescent tracer for amyloid, Amytracker. Middle: A cell containing less Amytracker-stained condensates of F220. Bottom: A cell containing Amytracker-stained condensates of F220. Inset (bottom): Enlarged condensates enclosed by the white dot lines. White arrows and yellow arrowheads represent the positions of Amytracker-positive and negative condensates in the cytoplasm. The asterisk in the images represents the nucleolus. Bars = 5 μm. In *B* and *D*, *p*-values above the bars and lines: one-way ANOVA with Tukey's test compared to GFP monomers as a control and comparison between the lines, respectively. (F) Comparison of normalized relative fluorescence intensities (nRFI) of C220 and F233 tagged with GFP (magenta and green, respectively) in live Neuro2a cells using fluorescence recovery after photobleaching (FRAP). Dots and bars indicate mean ± 95% CI (n = 6 for C220 and n = 7 for F220). Inset values show estimated mobile fraction from the maximum recovery of nRFI obtained by non-linear curve fitting analysis using a one-component exponential recovery model.

Figure 4 Co-condensation of C220 and F220 with the glycine-rich region.

(A) Confocal fluorescence and bright field microscopic images of Neuro2a cells expressing GFP-tagged C220 or F220 with the glycine-rich region (GRR; C274) tagged with a red fluorescent protein (RFP). Mo indicates GFP or RFP monomers. White arrows represent C274-positive condensates in the cytoplasm. Bars = 5 μm. (B) The population of Neuro2a cells that carry GFP-positive condensates (cond) in the cytoplasm. (C) The population of Neuro2a cells carrying RFP-positive condensates in the cytoplasm. (D) The ratio of RFP-positive condensates to GFP-positive ones. In *B*, *C*, and *D*, ND: Not

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

determined; N/A: Not applicable; *p*-values above the lines: one-way ANOVA with Tukey's test.

Figure 5 The condensation and cytotoxicity of TDP-43 CTFs that carry hydrophobic region-lacking GRR in Neuro2a cells.

(A) Domain structure and abbreviation of *Homo sapiens* TDP-43, *Caenorhabditis elegans* orthologue of TDP-43 (TDP-1), and its CTFs: amino acid 220-414 (C220), 220-262 (F220), C220 lacking 311-340 region (C220Δ), and a fusion protein of F220 and C-terminal region (C-term) of TDP-1 (F220c). (B) Confocal fluorescence and bright field microscopic images of Neuro2a cells expressing GFP-tagged C220Δ and F220c. Cells containing cytoplasmic condensates (*bottom*) or not containing them (*top*) were represented. (C) The population of Neuro2a cells carrying condensates in the cytoplasm. Mo denotes GFP monomers. Bars = 5 μm. (D) Western blot of 1% SDS-soluble (S) and -insoluble (P) cell lysates using an anti-GFP and anti-α-tubulin antibody. Right values (25, 37, and 50 kDa) represent positions of the molecular weight marker. (E) The proportion of dead cells expressing GFP-tagged C220, F220, C220Δ, and F220c. In *C* and *E*, Bars: mean and 95% CI; Dots: raw values (n = 4); ND: Not determined; *p*-values above the lines: one-way ANOVA with Tukey's test.

Figure 6 The role of the hydrophobic region in the glycine-rich region of TDP-43 to the co-condensation process in Neuro2a cells.

(A) Domain structure and abbreviation of TDP-43 CTFs: amino acids 220–414 (C220), 274–414 (C274), C220 lacking 311–340 region (C220Δ), and C274 lacking 311-340 region (C274Δ). GRR denote a glycine-rich region. (B) Confocal fluorescence and bright field microscopic images of Neuro2a cells co-expressing GFP- and RFP-tagged TDP-43 CTFs. Mo denotes GFP or RFP monomers. White arrows represent cytoplasmic condensates including both GFP- and RFP-tagged CTFs. White arrowheads represent cytoplasmic condensates including only GFP-tagged CTFs. Bars = 5 μm. (C and D) The population of Neuro2a cells carrying GFP-positive and RFP-positive cytoplasmic condensates (cond). (E) Ratio of cells containing RFP-positive condensates in a group of cells containing GFP-positive condensates (C–E) Bars: mean and 95% CI; Dots: raw values (n = 3); ND: Not determined; N/A: Not applicable; *p*-values above the lines: one-way ANOVA with Tukey's test. (F) Western blot of 1% SDS-soluble and -insoluble cell lysates using anti-GFP or anti-RFP antibodies.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 7 Intermolecular assembly of C220 and C263 using a blue light-induced oligomerization tag derived from a cryptochrome variant in Neuro2a cells.
(A) Fluorescence images of Neuro2a cells expressing TDP-43 CTFs: C220, C220Δ, C263, and C263Δ tagged with a cryptochrome-derived oligomerization inducer, CRY2olig (CR), and RFP before and after the 488 nm-light irradiation. The order of the hyphens before and after indicates the N/C-terminal side of the CR tag. White arrowheads represent cells with foci formation after the irradiation. Bars = 5 μm. (B) Population of cells that form foci after the blue-light irradiation obtained from all measured cells. ND: Not determined; *p*-values above the bars and lines: hypothesis test for the difference in the population proportions compared to a tag itself as control and comparison between the lines, respectively; ***$p < 0.001$.

Figure 8 Condensates of TDP-43 C-terminal fragments in nematodes.
(A) Confocal fluorescence and bright field images of nematodes expressing GFP-C220, GFP-C233, and GFP monomers (Mo) in 4 or 8 days after synchronization by bleaching (Day 1 and 5 of adulthood, respectively). Arrowheads indicate the typical position of condensates in body wall muscles. Bars = 0.2 mm. (B) Western blot of nematode lysates in 4 or 8 days after synchronization by bleaching (Day 1 and 5, respectively) using an anti-GFP and anti-α-tubulin antibody. Arrows show the monomer-corresponding bands for GFP-C220, GFP-C233, and GFP monomers. Oligomers indicate high molecular weight species. Asterisk shows a non-specific staining band by anti-GFP antibody.

Figure 9 Biophysical property of the condensates of TDP-43 C-terminal fragments in nematodes.
(A) Fluorescence recovery after photobleaching (FRAP) analysis of GFP monomers as a control and condensates of C220 or C233 tagged with GFP in live nematodes in 4 or 8 days after synchronization by bleaching (Days 1 and 5 of adulthood, respectively) ***Left***: Typical fluorescence images of condensates in nematodes before and after the photobleaching. Arrowheads indicate photobleached positions. Bars = 0.5 μm. Pseudocolor indicates the fluorescence intensities. ***Right***: Plots of the fluorescence recovery curve after the photobleaching of GFP monomers, GFP-C220, and -C233 in live nematodes (mean ± 95% CI; n = 7). The Y-axis represents the mobile fraction (normalized relative fluorescence intensity; nRFI) of C220 and C233 in the condensates and GFP monomers. Student's *t*-test for comparison among GFP monomers, C220, and C233 at the same day: ***$p < 0.001$. Statistical comparison just after the photobleaching ($t = 0$ s) is represented in Supplemental Figure S5. (B and C) Fluorescence correlation

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

spectroscopy (FCS) analysis of GFP-C220, -C233, and GFP monomers (Mo) in nematode lysates in 4 or 8 days after synchronization by bleaching (Day 1 and 5). Counts per particle (CPP) in ***B*** and diffusion time in ***C***. Bars indicate mean and 95% CI of independent 3 independent trials of lysis. *p*-values above the bars and lines: one-way ANOVA with Tukey's test compared to GFP monomers as control and comparison between the lines (###$p < 0.001$ and ***$p < 0.001$), respectively. NS with a cyan color indicates $p > 0.05$.

Figure 10 Bends activity and lifespan of nematodes expressing TDP-43 carboxy terminal fragments in nematodes.

(A) The number of bends per 30 seconds of nematodes expressing GFP-C220, -C233, and GFP monomers in 4 or 8 days after synchronization by bleaching (Day 1 and 5 of adulthood, respectively) in M9 buffer. Bars indicate the mean and 95% CI of 36 nematodes. *p*-values above the bars and lines: one-way ANOVA with Tukey's test compared to GFP monomers as control and comparison between the lines (###$p < 0.001$ and ***$p < 0.001$), respectively. NS with cyan color indicates $p > 0.05$. (B) Survival curve of synchronized nematode expressing GFP-C220, -C233, or GFP monomers (magenta, green, and dark gray, respectively). The *x*-axis shows day after the synchronization by bleaching (i.e., its day 4 corresponds to day 1 of adulthood). *p*-values were obtained from Gehan's generalized Wilcoxon test: ***$p < 0.001$. NS indicates $p > 0.05$.

Figure 11 Model of structural change and aggregate formation of TDP-43 carboxy terminal fragments.

(A) Domain structure of RRM2 and glycine-rich region (GRR) of TDP-43 and a putative model for conformational change of C220. The secondary structure of human RRM2 of TDP-43 is homologically predicted from the solution NMR structure of murine RRM2 (PDB#3D2W): α1 and α2 indicate the first and second α-helix structure in RRM2; β1–5 indicate the five β-strands, respectively. h indicates α-helical structure (322–342) from solution NMR structure of a C-terminal fragment of human TDP-43 (PDB#2N3X). Green β-strands represent potentially amyloidogenic sequences as previously reported. (B) A scheme of intermolecular assembly and aggregate propagation of TDP-43 CTFs that includes truncated RRM2 (tRRM2) and GRR containing hydrophobic region (HR).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Supplemental Figure legends

Supplemental Figure 1: Comparison of cells containing cytoplasmic condensates of C220, C263, and C274 in Neuro2a cells

Dots and bar graphs indicate the population of the cells containing cytoplasmic condensates in Neuro2a cells (mean ± 95% CI; n = 3). *p*-values above the lines: one-way ANOVA with Tukey's test.

Supplemental Figure 2: Amytracker staining of the C233 and C274 condensates in Neuro2a cells

Confocal images of GFP-tagged C233 and C274-expressing Neuro2a cells stained with a fluorescent tracer for amyloid, Amytracker. The asterisk in the images represents the nucleolus. Bars = 5 μm.

Supplemental Figure 3: Confocal fluorescent images of Neuro2a cells expressing C220, C233, and F220 tagged with GFP after the 1,6-Hexanediol treatment.

The confocal fluorescence images of cells containing cytoplasmic condensates were treated with 1,6-hexanediol were represented. Bars = 5 μm.

Supplemental Figure 4: Hydropathy plot of human TDP-43 amino acid sequence

(A) The hydropathy plot was calculated using the Expasy Web server [Kyte J. & Doolittle R.F., ***J. Mol. Biol.***, 157, 105-132 (1982)]. The CTF regions of C220 and C233 represented a cyan and green area, respectively. A highly hydrophobic region in GRR was shown in a yellow area. (B) The hydrophobicity scores of each amino acid are represented (Red to blue: high to low hydrophobicity).

Supplemental Figure 5: Comparison of normalized relative fluorescence intensities (nRFI) just after photobleaching of C220 and C233 tagged with GFP in *C. elegans*.

Box and bars with dots plot of nRFI of GFP monomers, GFP-C220, and GFP-C233 just after photobleaching (t = 0 s in Figure 9A). *p*-values above the lines: one-way ANOVA with Tukey's test (n =7); NS indicates $p > 0.05$.

Supplemental Figure 6: Illustration of the secondary structure of intact RRM2 and GRR of TDP-43 and the cleavage positions of C220, C233, C243, and C253

(A–E) α1 and α2 indicate two α-helixes and β1–5 indicate five β-strands in RRM2 of TDP-43 observed in PDB#3D2W. The h means α-helical structure in GRR of TDP-43

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

observed in PDB#2N3X (Jiang *et al*, 2016). Structure of intact RRM2 and GRR (A), C220/TDP25 and the cleaved fragment of RRM2 (B), C233 (C), C243 (D), and C253 (E).

Supplemental Figure 7: Stereoview of the electrostatic interaction of the amino acid side chain (Q241 and N259) in RRM2

α1 and α2 indicate two α-helixes and β1–5 indicate five β-strands in RRM2 of TDP-43 observed in PDB#3D2W. The balls indicate the atoms of the side chain of Q241 and N259 (red and blue balls represent atoms having positive and negative charges, respectively).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

References:

Afroz T, Hock EM, Ernst P, Foglieni C, Jambeau M, Gilhespy LAB, Laferriere F, Maniecka Z, Pluckthun A, Mittl P *et al* (2017) Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nature communications* 8: 45

Ahmed M, Machado PM, Miller A, Spicer C, Herbelin L, He J, Noel J, Wang Y, McVey AL, Pasnoor M *et al* (2016) Targeting protein homeostasis in sporadic inclusion body myositis. *Sci Transl Med* 8: 331ra341

Alagar Boopathy LR, Jacob-Tomas S, Alecki C, Vera M (2022) Mechanisms tailoring the expression of heat shock proteins to proteostasis challenges. *The Journal of biological chemistry* 298: 101796

Alberti S, Hyman AA (2021) Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nature reviews Molecular cell biology* 22: 196-213

Arseni D, Hasegawa M, Murzin AG, Kametani F, Arai M, Yoshida M, Ryskeldi-Falcon B (2022) Structure of pathological TDP-43 filaments from ALS with FTLD. *Nature* 601: 139-143

Banani SF, Lee HO, Hyman AA, Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nature reviews Molecular cell biology* 18: 285-298

Bang J, Spina S, Miller BL (2015) Frontotemporal dementia. *Lancet* 386: 1672-1682

Ben-Zvi A, Miller EA, Morimoto RI (2009) Collapse of proteostasis represents an early molecular event in Caenorhabditis elegans aging. *Proceedings of the National Academy of Sciences of the United States of America* 106: 14914-14919

Benilova I, Karran E, De Strooper B (2012) The toxic Abeta oligomer and Alzheimer's disease: an emperor in need of clothes. *Nature neuroscience* 15: 349-357

Berning BA, Walker AK (2019) The Pathobiology of TDP-43 C-Terminal Fragments in ALS and FTLD. *Frontiers in neuroscience* 13: 335

Bolognesi B, Faure AJ, Seuma M, Schmiedel JM, Tartaglia GG, Lehner B (2019) The mutational landscape of a prion-like domain. *Nature communications* 10: 4162

Brenner S (1973) The genetics of behaviour. *Br Med Bull* 29: 269-271

Brignull HR, Moore FE, Tang SJ, Morimoto RI (2006) Polyglutamine proteins at the pathogenic threshold display neuron-specific aggregation in a pan-neuronal Caenorhabditis elegans model. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 26: 7597-7606

Caccamo A, Shaw DM, Guarino F, Messina A, Walker AW, Oddo S (2015) Reduced protein turnover mediates functional deficits in transgenic mice expressing the 25 kDa C-terminal fragment of TDP-43. *Human molecular genetics* 24: 4625-4635

Carey JL, Guo L (2022) Liquid-Liquid Phase Separation of TDP-43 and FUS in Physiology and Pathology of Neurodegenerative Diseases. *Front Mol Biosci* 9: 826719

Chhangani D, Martin-Pena A, Rincon-Limas DE (2021) Molecular, functional, and pathological aspects of TDP-43 fragmentation. *iScience* 24: 102459

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Christie NT, Lee AL, Fay HG, Gray AA, Kikis EA (2014) Novel polyglutamine model uncouples proteotoxicity from aging. *PloS one* 9: e96835

Chung CG, Lee H, Lee SB (2018) Mechanisms of protein toxicity in neurodegenerative diseases. *Cellular and molecular life sciences : CMLS* 75: 3159-3180

Conicella AE, Dignon GL, Zerze GH, Schmidt HB, D'Ordine AM, Kim YC, Rohatgi R, Ayala YM, Mittal J, Fawzi NL (2020) TDP-43 alpha-helical structure tunes liquid-liquid phase separation and function. *Proceedings of the National Academy of Sciences of the United States of America* 117: 5883-5894

Conicella AE, Zerze GH, Mittal J, Fawzi NL (2016) ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* 24: 1537-1549

Da Cruz S, Cleveland DW (2011) Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Current opinion in neurobiology* 21: 904-919

Dunker AK, Silman I, Uversky VN, Sussman JL (2008) Function and structure of inherently disordered proteins. *Current opinion in structural biology* 18: 756-764

Espargaro A, Busquets MA, Estelrich J, Sabate R (2016) Key Points Concerning Amyloid Infectivity and Prion-Like Neuronal Invasion. *Frontiers in molecular neuroscience* 9: 29

Fang YS, Tsai KJ, Chang YJ, Kao P, Woods R, Kuo PH, Wu CC, Liao JY, Chou SC, Lin V *et al* (2014) Full-length TDP-43 forms toxic amyloid oligomers that are present in frontotemporal lobar dementia-TDP patients. *Nature communications* 5: 4824

Fazal R, Boeynaems S, Swijsen A, De Decker M, Fumagalli L, Moisse M, Vanneste J, Guo W, Boon R, Vercruysse T *et al* (2021) HDAC6 inhibition restores TDP-43 pathology and axonal transport defects in human motor neurons with TARDBP mutations. *The EMBO journal* 40: e106177

Frumkin A, Dror S, Pokrzywa W, Bar-Lavan Y, Karady I, Hoppe T, Ben-Zvi A (2014) Challenging muscle homeostasis uncovers novel chaperone interactions in Caenorhabditis elegans. *Front Mol Biosci* 1: 21

Guo W, Chen Y, Zhou X, Kar A, Ray P, Chen X, Rao EJ, Yang M, Ye H, Zhu L *et al* (2011) An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nature structural & molecular biology* 18: 822-830

Hartl FU (2017) Protein Misfolding Diseases. *Annual review of biochemistry* 86: 21-26

Jawerth L, Fischer-Friedrich E, Saha S, Wang J, Franzmann T, Zhang X, Sachweh J, Ruer M, Ijavi M, Saha S *et al* (2020) Protein condensates as aging Maxwell fluids. *Science* 370: 1317-1323

Jiang LL, Che MX, Zhao J, Zhou CJ, Xie MY, Li HY, He JH, Hu HY (2013) Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. *The Journal of biological chemistry* 288: 19614-19624

Jiang LL, Zhao J, Yin XF, He WT, Yang H, Che MX, Hu HY (2016) Two mutations G335D and Q343R within the amyloidogenic core region of TDP-43 influence its aggregation and inclusion formation. *Scientific reports* 6: 23928

Johnson BS, McCaffery JM, Lindquist S, Gitler AD (2008) A yeast TDP-43 proteinopathy model:

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. *Proceedings of the National Academy of Sciences of the United States of America* 105: 6439-6444

Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R (1993) A C. elegans mutant that lives twice as long as wild type. *Nature* 366: 461-464

Kitamura A, Iwasaki N, Kinjo M (2018) 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

Kitamura A, Kubota H, Pack CG, Matsumoto G, Hirayama S, Takahashi Y, Kimura H, Kinjo M, Morimoto RI, Nagata K (2006) Cytosolic chaperonin prevents polyglutamine toxicity with altering the aggregation state. *Nature cell biology* 8: 1163-1170

Kitamura A, Nakayama Y, Shibasaki A, Taki A, Yuno S, Takeda K, Yahara M, Tanabe N, Kinjo M (2016) Interaction of RNA with a C-terminal fragment of the amyotrophic lateral sclerosis-associated TDP43 reduces cytotoxicity. *Scientific reports* 6: 19230

Kitamura A, Yuno S, Muto H, Kinjo M (2017) Different aggregation states of a nuclear localization signal-tagged 25-kDa C-terminal fragment of TAR RNA/DNA-binding protein 43 kDa. *Genes to cells : devoted to molecular & cellular mechanisms* 22: 521-534

Klaips CL, Jayaraj GG, Hartl FU (2018) Pathways of cellular proteostasis in aging and disease. *The Journal of cell biology* 217: 51-63

Krainer G, Welsh TJ, Joseph JA, Espinosa JR, Wittmann S, de Csillery E, Sridhar A, Toprakcioglu Z, Gudiskyte G, Czekalska MA *et al* (2021) Reentrant liquid condensate phase of proteins is stabilized by hydrophobic and non-ionic interactions. *Nature communications* 12: 1085

Kuo PH, Chiang CH, Wang YT, Doudeva LG, Yuan HS (2014) The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. *Nucleic acids research* 42: 4712-4722

Labbadia J, Morimoto RI (2015) The biology of proteostasis in aging and disease. *Annual review of biochemistry* 84: 435-464

Laferriere F, Maniecka Z, Perez-Berlanga M, Hruska-Plochan M, Gilhespy L, Hock EM, Wagner U, Afroz T, Boersema PJ, Barmettler G *et al* (2019) TDP-43 extracted from frontotemporal lobar degeneration subject brains displays distinct aggregate assemblies and neurotoxic effects reflecting disease progression rates. *Nature neuroscience* 22: 65-77

Lee AL, Ung HM, Sands LP, Kikis EA (2017) A new Caenorhabditis elegans model of human huntingtin 513 aggregation and toxicity in body wall muscles. *PloS one* 12: e0173644

Lee EB, Lee VM, Trojanowski JQ (2011) Gains or losses: molecular mechanisms of TDP43-mediated neurodegeneration. *Nature reviews Neuroscience* 13: 38-50

Li Q, Babinchak WM, Surewicz WK (2021) Cryo-EM structure of amyloid fibrils formed by the entire low complexity domain of TDP-43. *Nature communications* 12: 1620

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Ling SC, Polymenidou M, Cleveland DW (2013) Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* 79: 416-438

Loganathan S, Lehmkuhl EM, Eck RJ, Zarnescu DC (2019) To Be or Not To Be...Toxic-Is RNA Association With TDP-43 Complexes Deleterious or Protective in Neurodegeneration? *Front Mol Biosci* 6: 154

Lukavsky PJ, Daujotyte D, Tollervey JR, Ule J, Stuani C, Buratti E, Baralle FE, Damberger FF, Allain FH (2013) Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nature structural & molecular biology* 20: 1443-1449

Mallick A, Ranawade A, van den Berg W, Gupta BP (2020) Axin-Mediated Regulation of Lifespan and Muscle Health in C. elegans Requires AMPK-FOXO Signaling. *iScience* 23: 101843

Morley JF, Brignull HR, Weyers JJ, Morimoto RI (2002) The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* 99: 10417-10422

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM *et al* (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* 314: 130-133

Nillegoda NB, Kirstein J, Szlachcic A, Berynskyy M, Stank A, Stengel F, Arnsburg K, Gao X, Scior A, Aebersold R *et al* (2015) Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature* 524: 247-251

Nussbaum-Krammer CI, Morimoto RI (2014) Caenorhabditis elegans as a model system for studying non-cell-autonomous mechanisms in protein-misfolding diseases. *Disease models & mechanisms* 7: 31-39

Pirie E, Oh CK, Zhang X, Han X, Cieplak P, Scott HR, Deal AK, Ghatak S, Martinez FJ, Yeo GW *et al* (2021) S-nitrosylated TDP-43 triggers aggregation, cell-to-cell spread, and neurotoxicity in hiPSCs and in vivo models of ALS/FTD. *Proceedings of the National Academy of Sciences of the United States of America* 118

Prots I, Grosch J, Brazdis RM, Simmnacher K, Veber V, Havlicek S, Hannappel C, Krach F, Krumbiegel M, Schutz O *et al* (2018) alpha-Synuclein oligomers induce early axonal dysfunction in human iPSC-based models of synucleinopathies. *Proceedings of the National Academy of Sciences of the United States of America* 115: 7813-7818

Ross CA, Poirier MA (2004) Protein aggregation and neurodegenerative disease. *Nature medicine* 10 Suppl: S10-17

Saini A, Chauhan VS (2011) Delineation of the core aggregation sequences of TDP-43 C-terminal fragment. *Chembiochem : a European journal of chemical biology* 12: 2495-2501

Shimonaka S, Nonaka T, Suzuki G, Hisanaga S, Hasegawa M (2016) Templated Aggregation of TAR DNA-binding Protein of 43 kDa (TDP-43) by Seeding with TDP-43 Peptide Fibrils. *The Journal of biological chemistry* 291: 8896-8907

Shin Y, Brangwynne CP (2017) Liquid phase condensation in cell physiology and disease. *Science* 357

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Sinnige T, Meisl G, Michaels TCT, Vendruscolo M, Knowles TPJ, Morimoto RI (2021) Kinetic analysis reveals that independent nucleation events determine the progression of polyglutamine aggregation in C. elegans. *Proceedings of the National Academy of Sciences of the United States of America* 118

Smethurst P, Newcombe J, Troakes C, Simone R, Chen YR, Patani R, Sidle K (2016) In vitro prion-like behaviour of TDP-43 in ALS. *Neurobiology of disease* 96: 236-247

Taslimi A, Vrana JD, Chen D, Borinskaya S, Mayer BJ, Kennedy MJ, Tucker CL (2014) An optimized optogenetic clustering tool for probing protein interaction and function. *Nature communications* 5: 4925

Taylor JP, Hardy J, Fischbeck KH (2002) Toxic proteins in neurodegenerative disease. *Science* 296: 1991-1995

Uemura E, Niwa T, Minami S, Takemoto K, Fukuchi S, Machida K, Imataka H, Ueda T, Ota M, Taguchi H (2018) Large-scale aggregation analysis of eukaryotic proteins reveals an involvement of intrinsically disordered regions in protein folding. *Scientific reports* 8: 678

Uno M, Nishida E (2016) Lifespan-regulating genes in C. elegans. *NPJ Aging Mech Dis* 2: 16010

Vecchi G, Sormanni P, Mannini B, Vandelli A, Tartaglia GG, Dobson CM, Hartl FU, Vendruscolo M (2020) Proteome-wide observation of the phenomenon of life on the edge of solubility. *Proceedings of the National Academy of Sciences of the United States of America* 117: 1015-1020

Wang X, Ma M, Teng J, Che X, Zhang W, Feng S, Zhou S, Zhang Y, Wu E, Ding X (2015) Valproate Attenuates 25-kDa C-Terminal Fragment of TDP-43-Induced Neuronal Toxicity via Suppressing Endoplasmic Reticulum Stress and Activating Autophagy. *International journal of biological sciences* 11: 752-761

Wang YT, Kuo PH, Chiang CH, Liang JR, Chen YR, Wang S, Shen JC, Yuan HS (2013) The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates. *The Journal of biological chemistry* 288: 9049-9057

Westergard T, Jensen BK, Wen X, Cai J, Kropf E, Iacovitti L, Pasinelli P, Trotti D (2016) Cell-to-Cell Transmission of Dipeptide Repeat Proteins Linked to C9orf72-ALS/FTD. *Cell reports* 17: 645-652

Wu S, Wang D, Liu J, Feng Y, Weng J, Li Y, Gao X, Liu J, Wang W (2017) The Dynamic Multisite Interactions between Two Intrinsically Disordered Proteins. *Angewandte Chemie* 56: 7515-7519

Xiao S, Sanelli T, Chiang H, Sun Y, Chakrabartty A, Keith J, Rogaeva E, Zinman L, Robertson J (2015) Low molecular weight species of TDP-43 generated by abnormal splicing form inclusions in amyotrophic lateral sclerosis and result in motor neuron death. *Acta neuropathologica* 130: 49-61

Yang Z, Lin F, Robertson CS, Wang KK (2014) Dual vulnerability of TDP-43 to calpain and caspase-3 proteolysis after neurotoxic conditions and traumatic brain injury. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* 34: 1444-1452

Zhang H, Shao S, Zeng Y, Wang X, Qin Y, Ren Q, Xiang S, Wang Y, Xiao J, Sun Y (2022) Reversible phase separation of HSF1 is required for an acute transcriptional response during heat shock. *Nature cell biology*

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Zhang T, Hwang HY, Hao H, Talbot C, Jr., Wang J (2012) Caenorhabditis elegans RNA-processing protein TDP-1 regulates protein homeostasis and life span. *The Journal of biological chemistry* 287: 8371-8382

Zhang T, Mullane PC, Periz G, Wang J (2011) TDP-43 neurotoxicity and protein aggregation modulated by heat shock factor and insulin/IGF-1 signaling. *Human molecular genetics* 20: 1952-1965

Zhang YJ, Xu YF, Cook C, Gendron TF, Roettges P, Link CD, Lin WL, Tong J, Castanedes-Casey M, Ash P *et al* (2009) Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity. *Proceedings of the National Academy of Sciences of the United States of America* 106: 7607-7612

Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

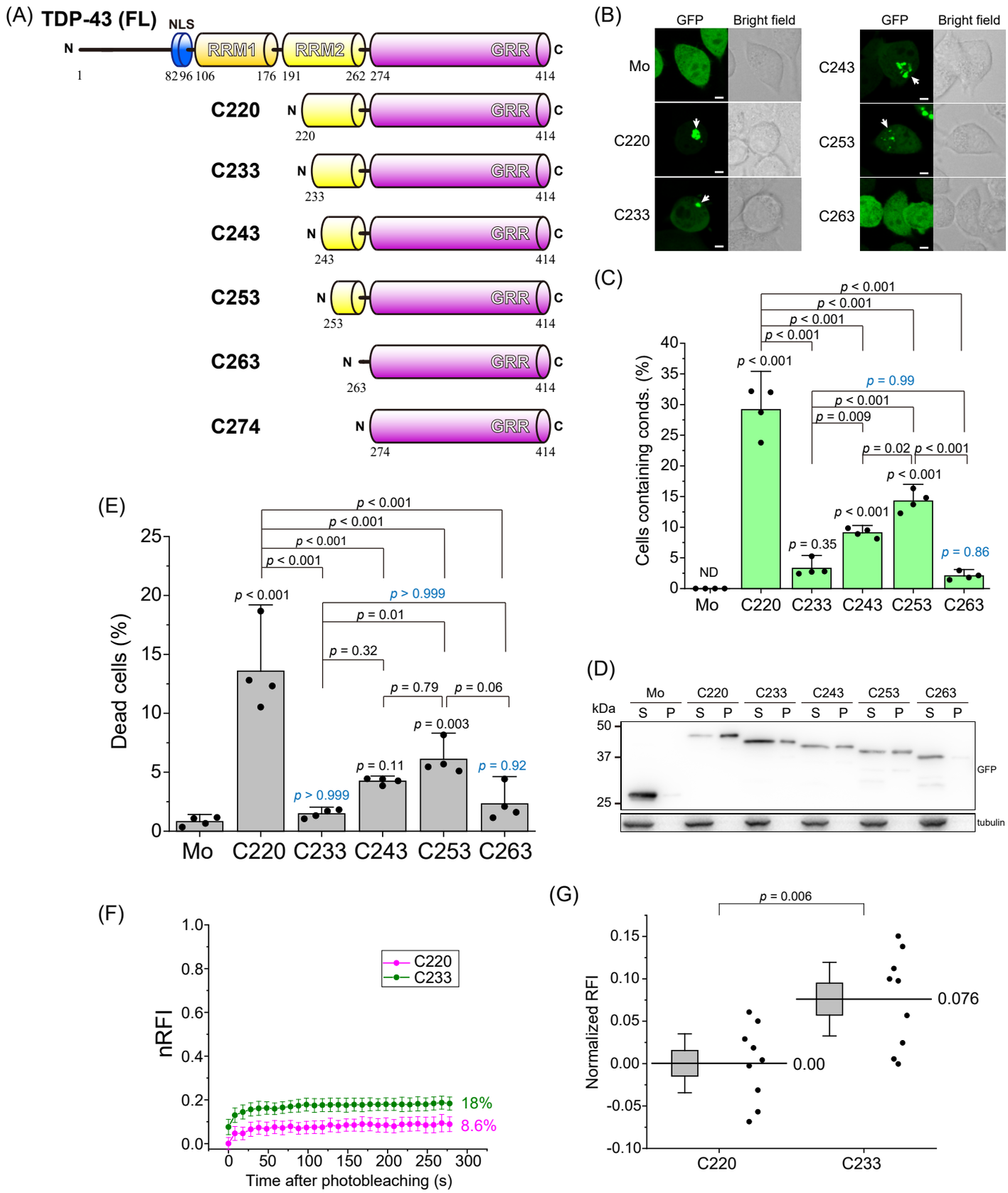
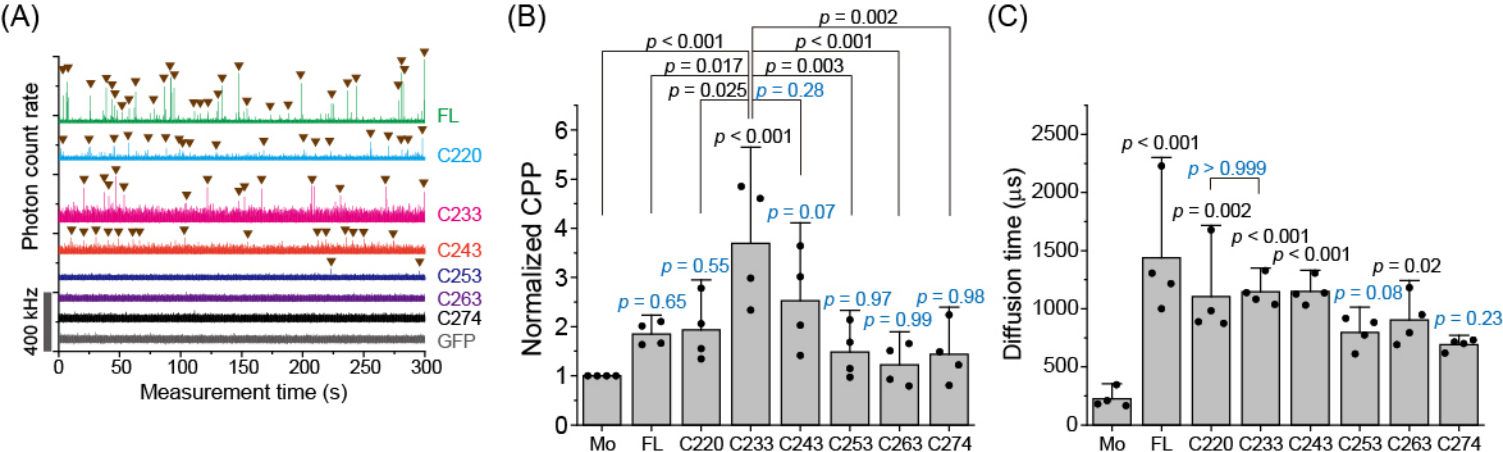
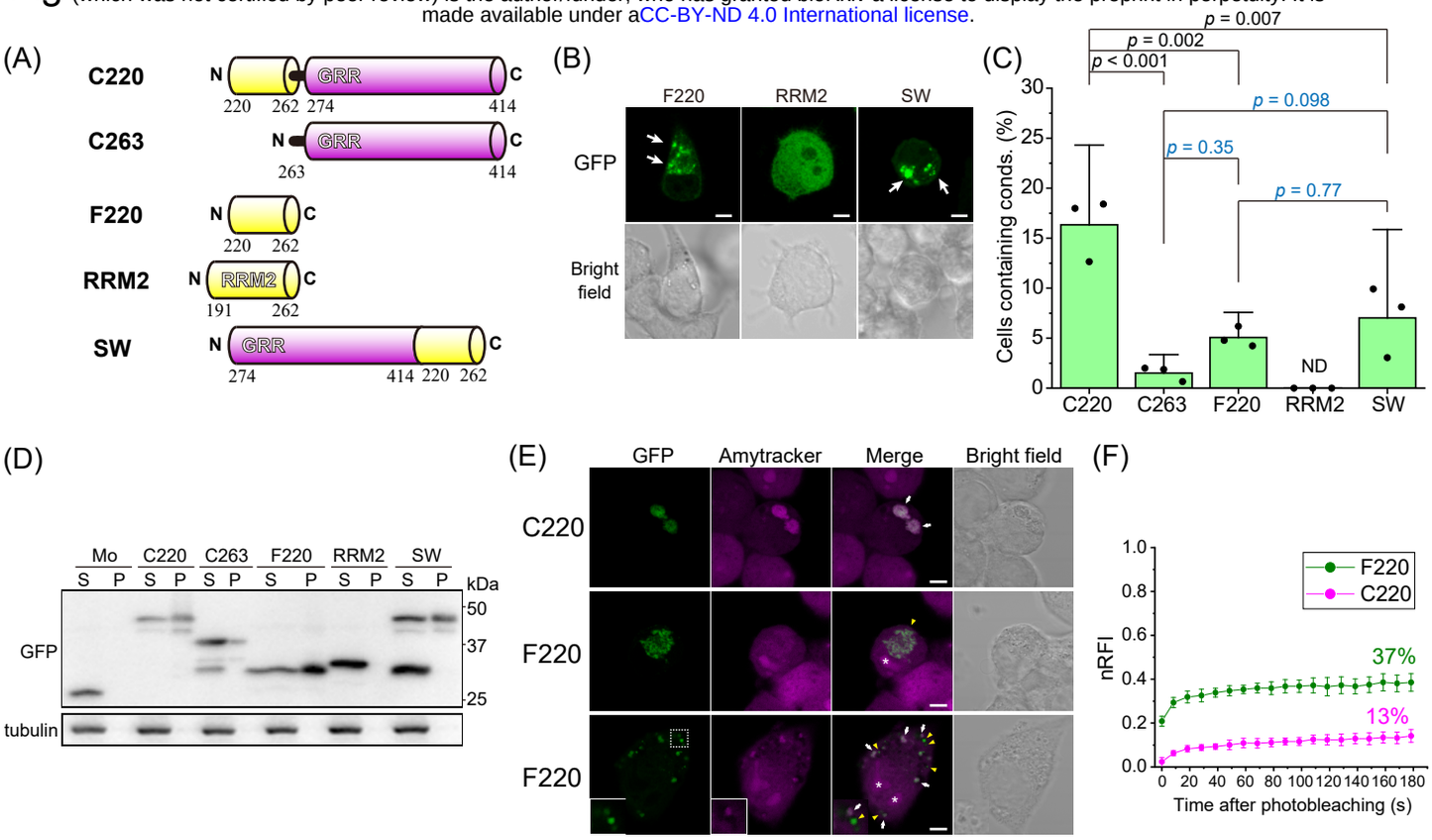


Figure 2


(A)
Photon count rate
FL
C220
C233
C243
C253
C263
C274
GFP
400 kHz
0
50
100
150
200
250
300
Measurement time (s)
(B)
Normalized CPP
0
1
2
3
4
5
6
p < 0.001
p = 0.017
p = 0.025
p = 0.28
p = 0.003
p < 0.001
p = 0.002
p < 0.001
p = 0.07
p = 0.55
p = 0.65
p = 0.97
p = 0.99
p = 0.98
Mo
FL
C220
C233
C243
C253
C263
C274
(C)
Diffusion time (μs)
0
500
1000
1500
2000
2500
p < 0.001
p > 0.999
p = 0.002
p < 0.001
p < 0.001
p = 0.08
p = 0.02
p = 0.23
Mo
FL
C220
C233
C243
C253
C263
C274

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 3



bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 4

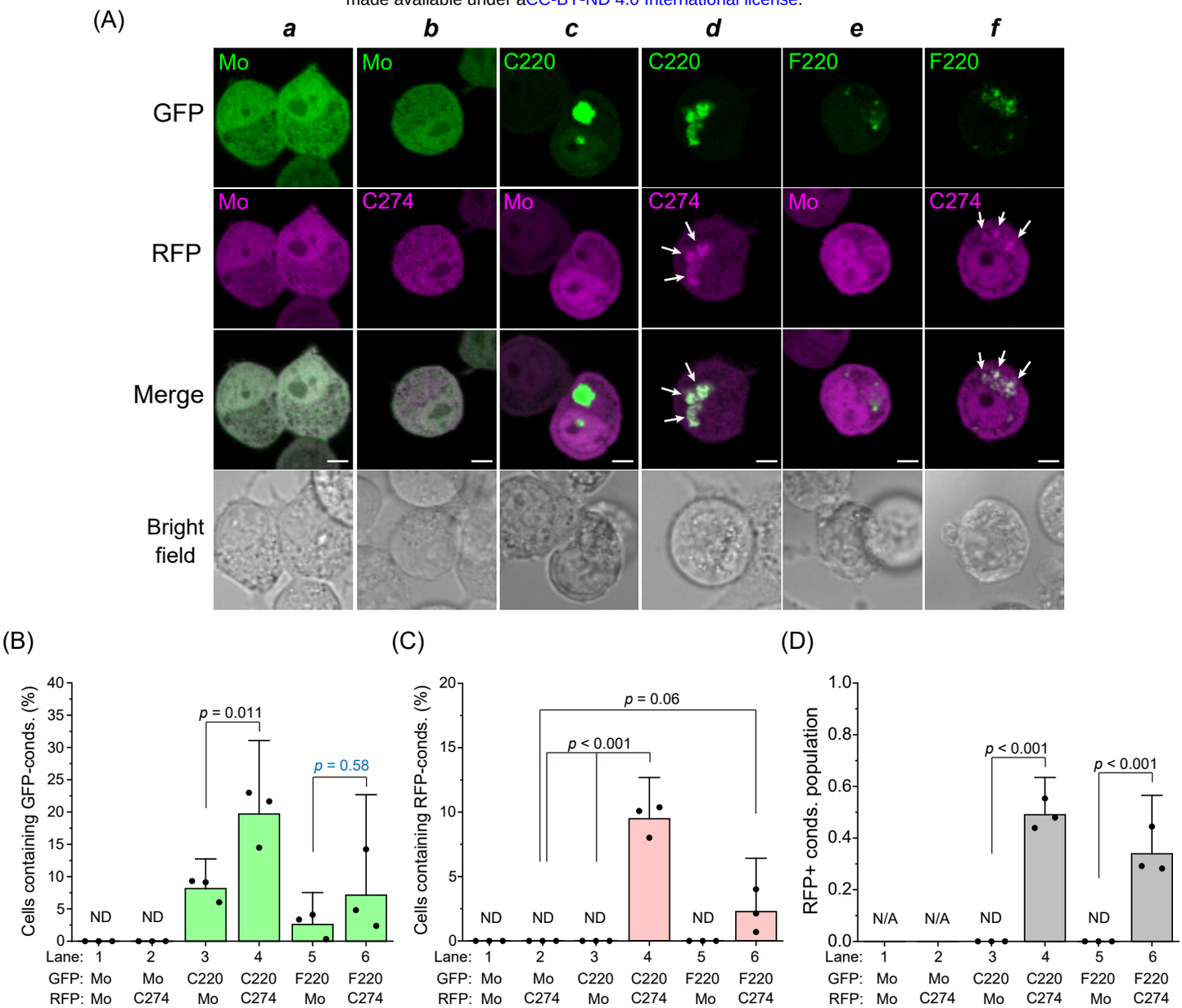
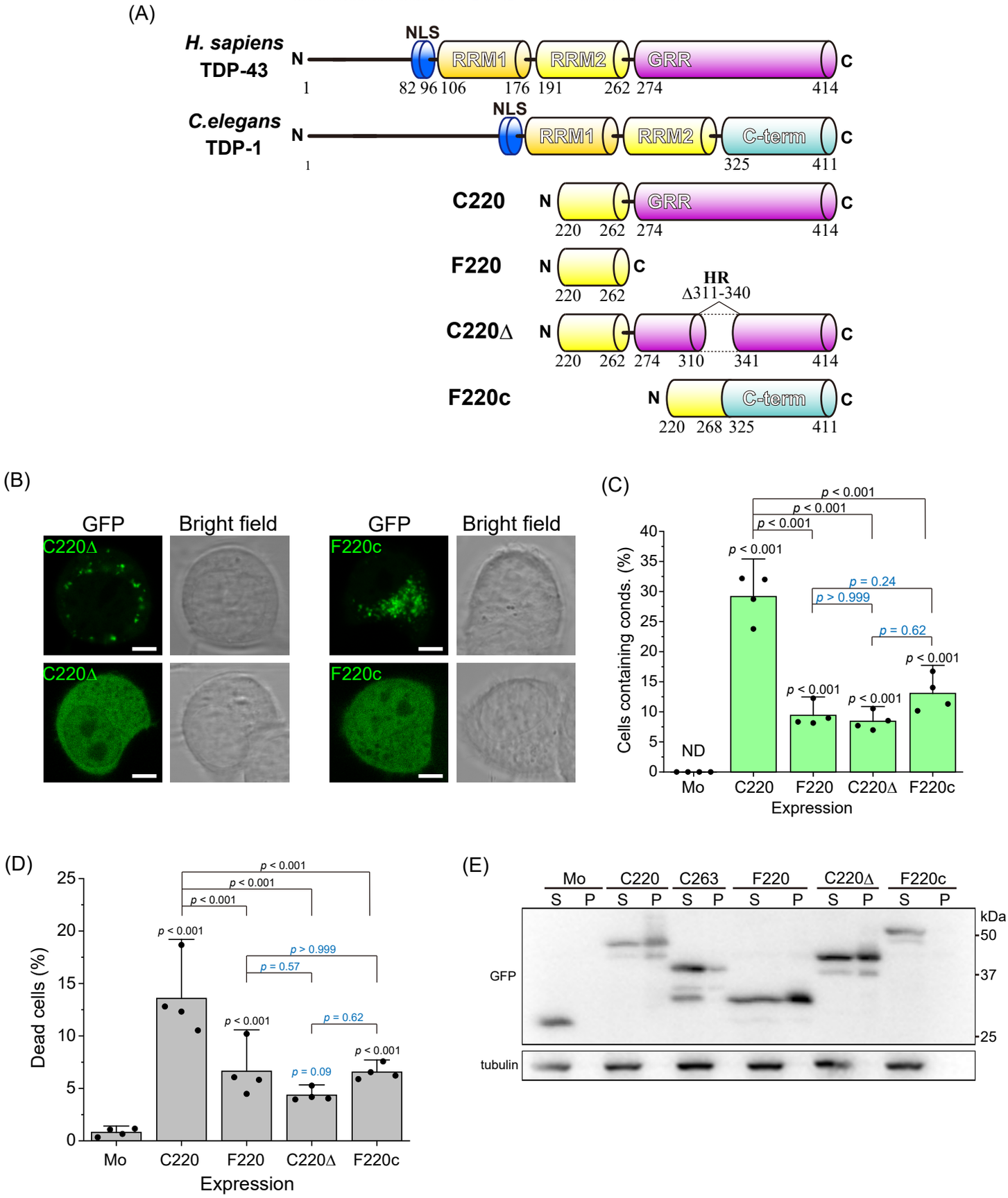


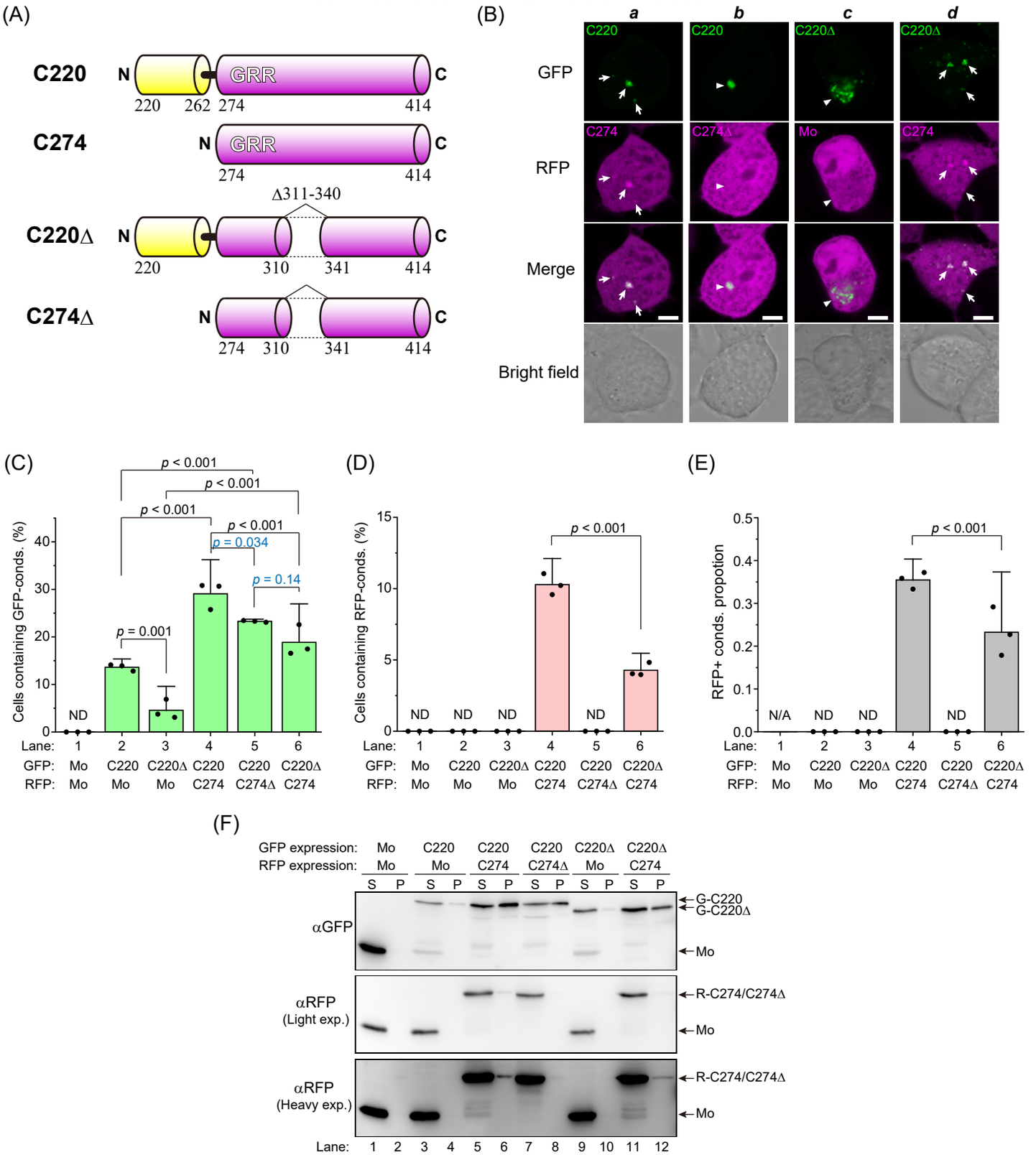
Figure 5

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

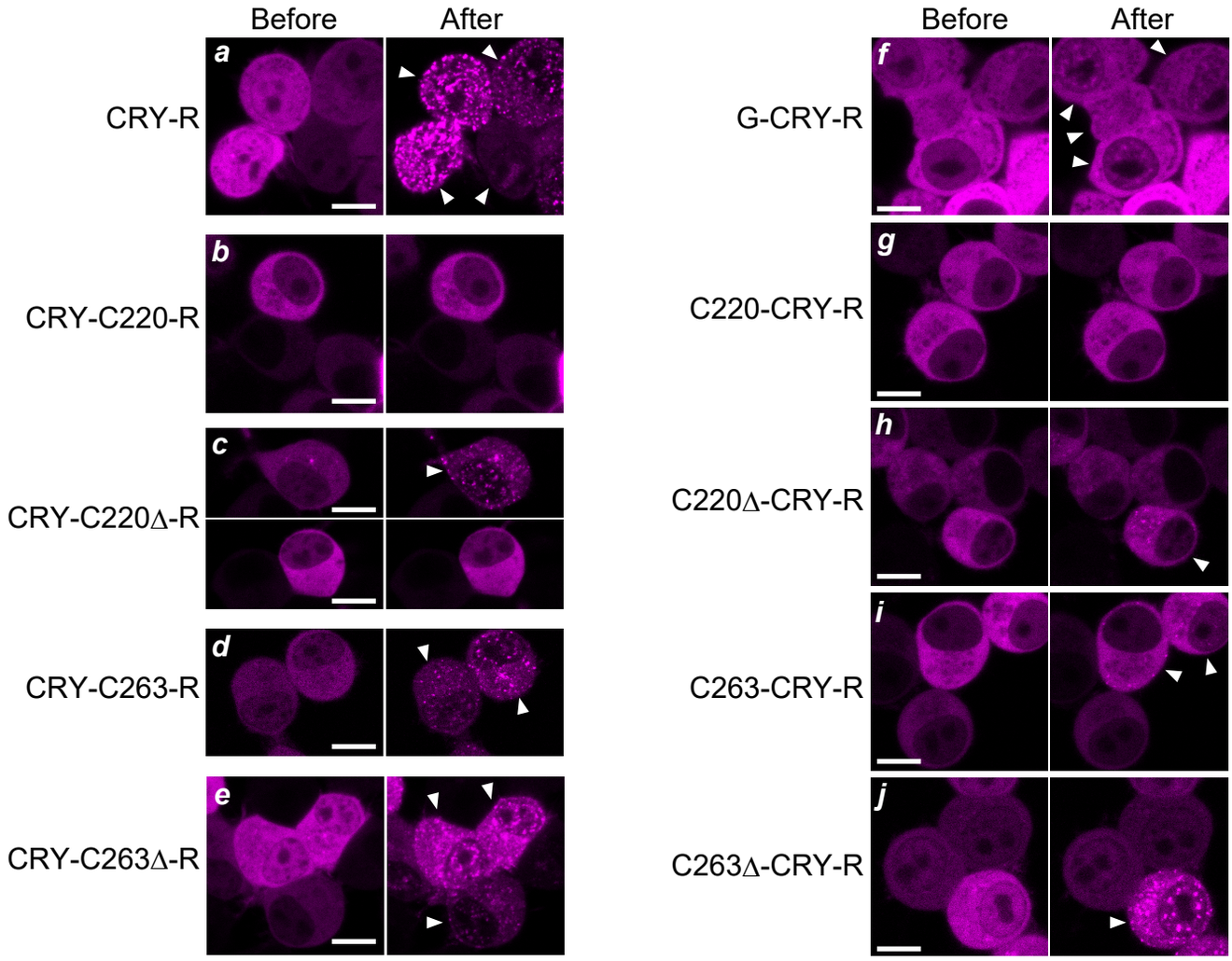
Figure 6



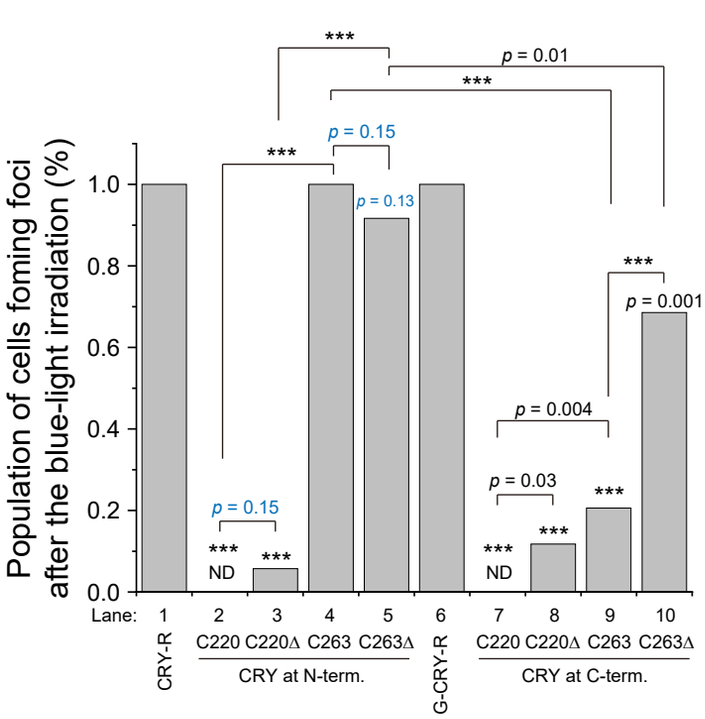
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 7

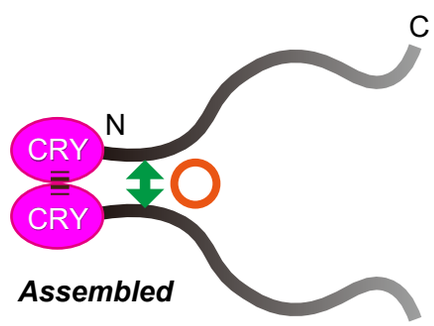
(A)



(B)



(C)



(D)

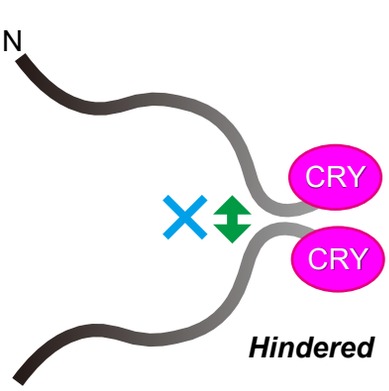
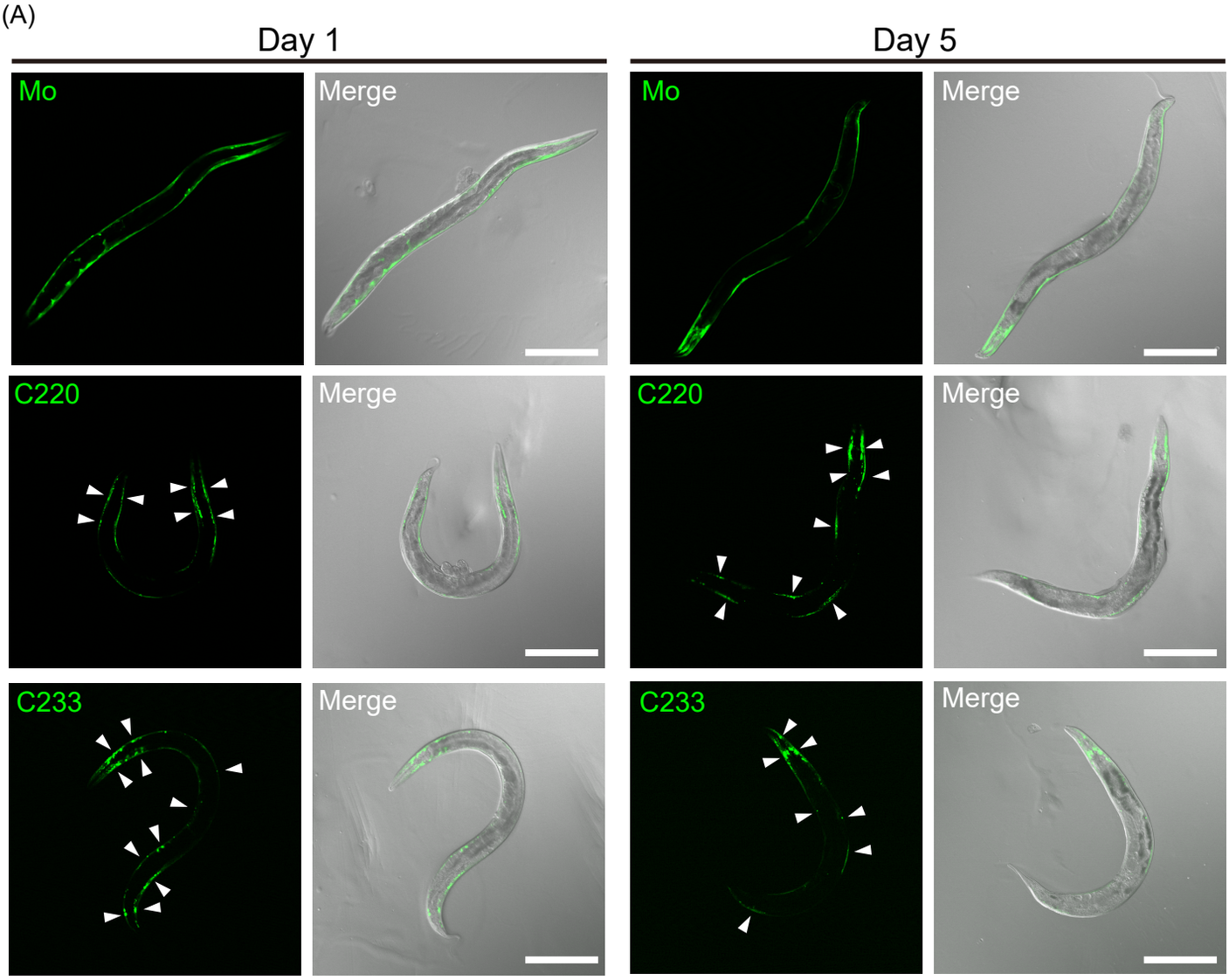


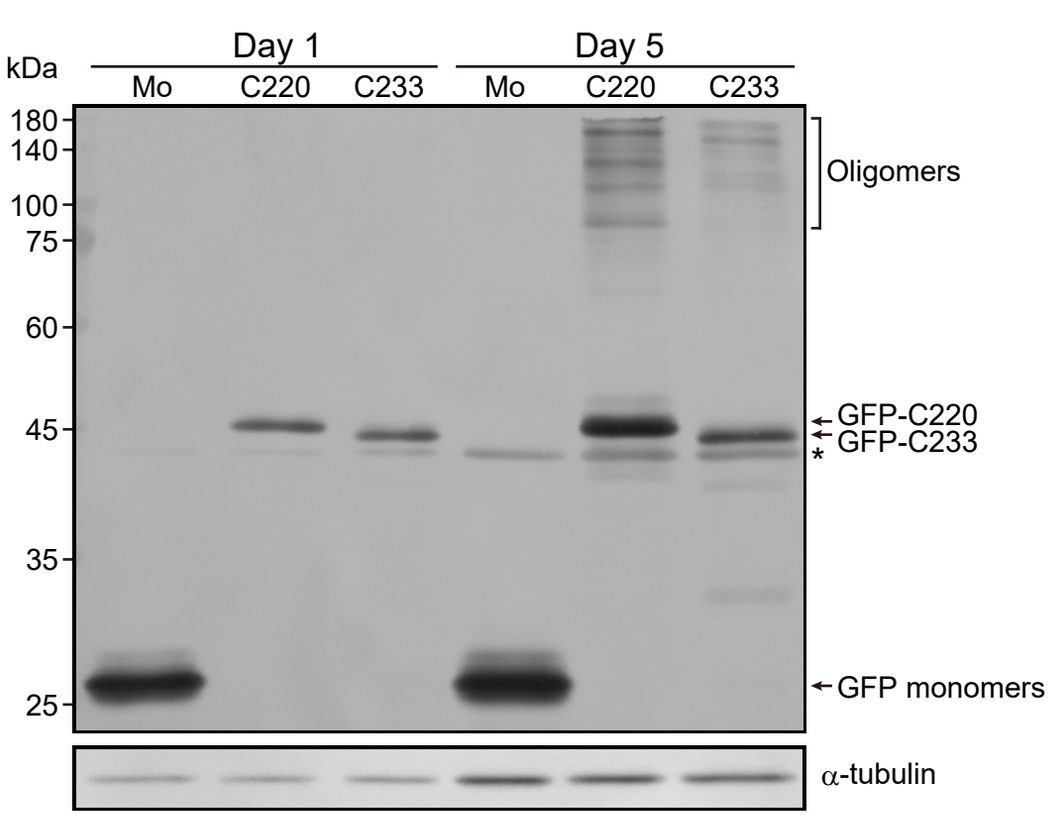
Figure 8

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

(A)



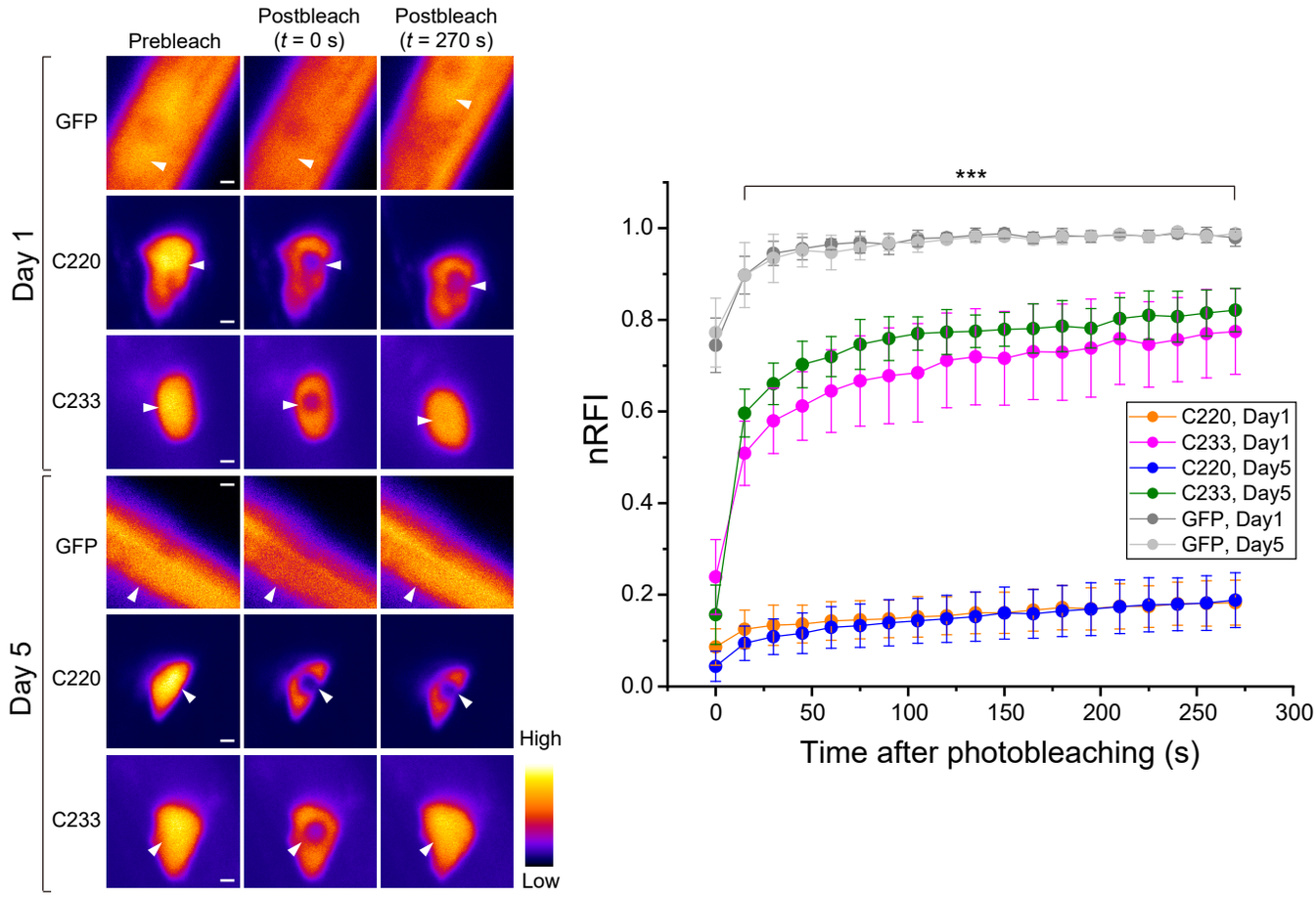
(B)



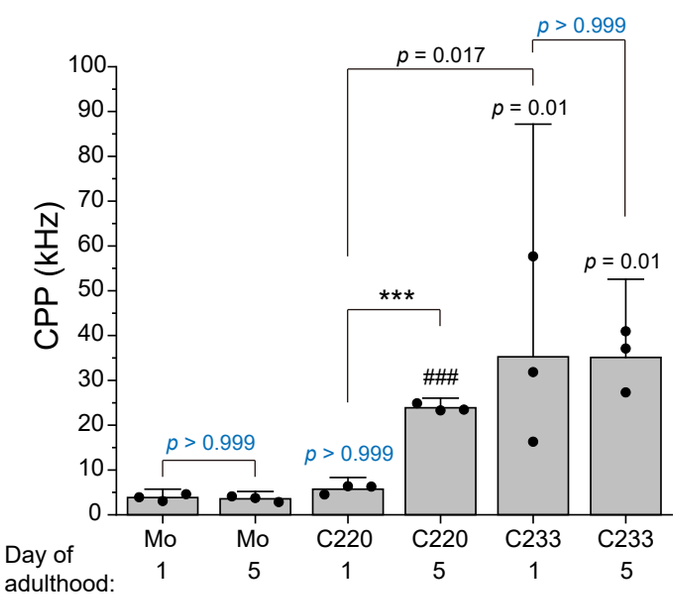
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Figure 9

(A)



(B)



(C)

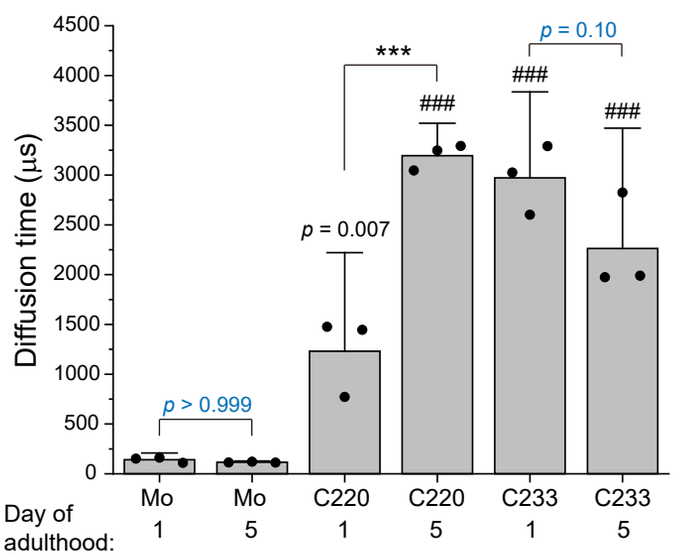


Figure 10

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

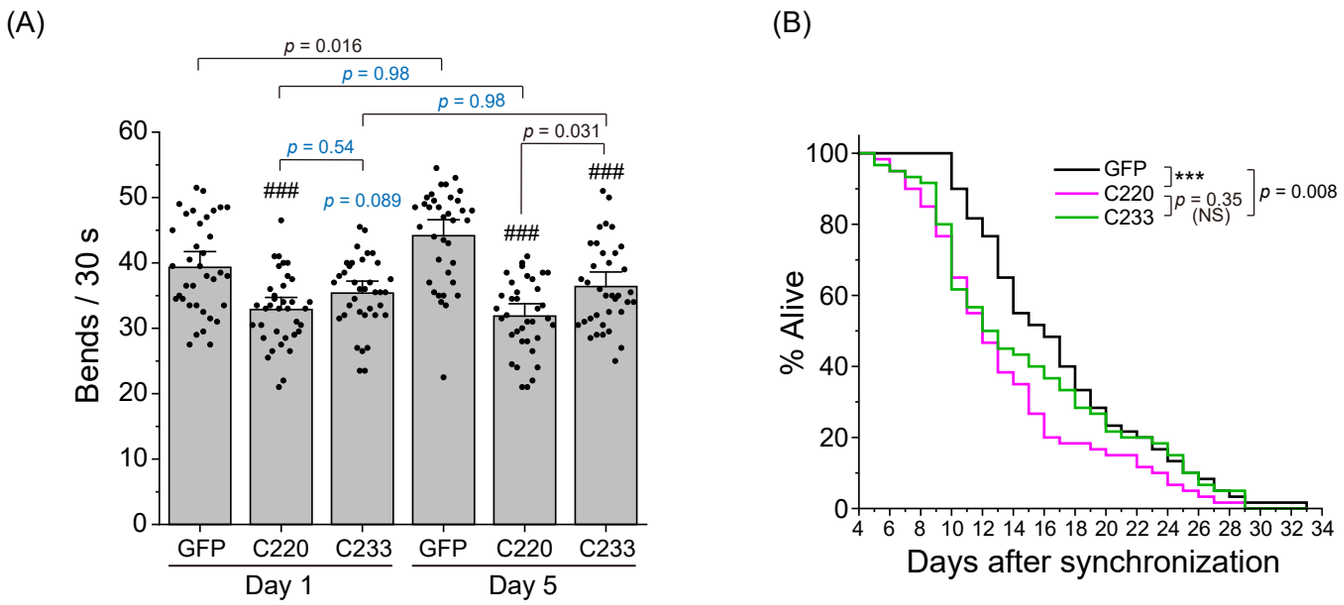
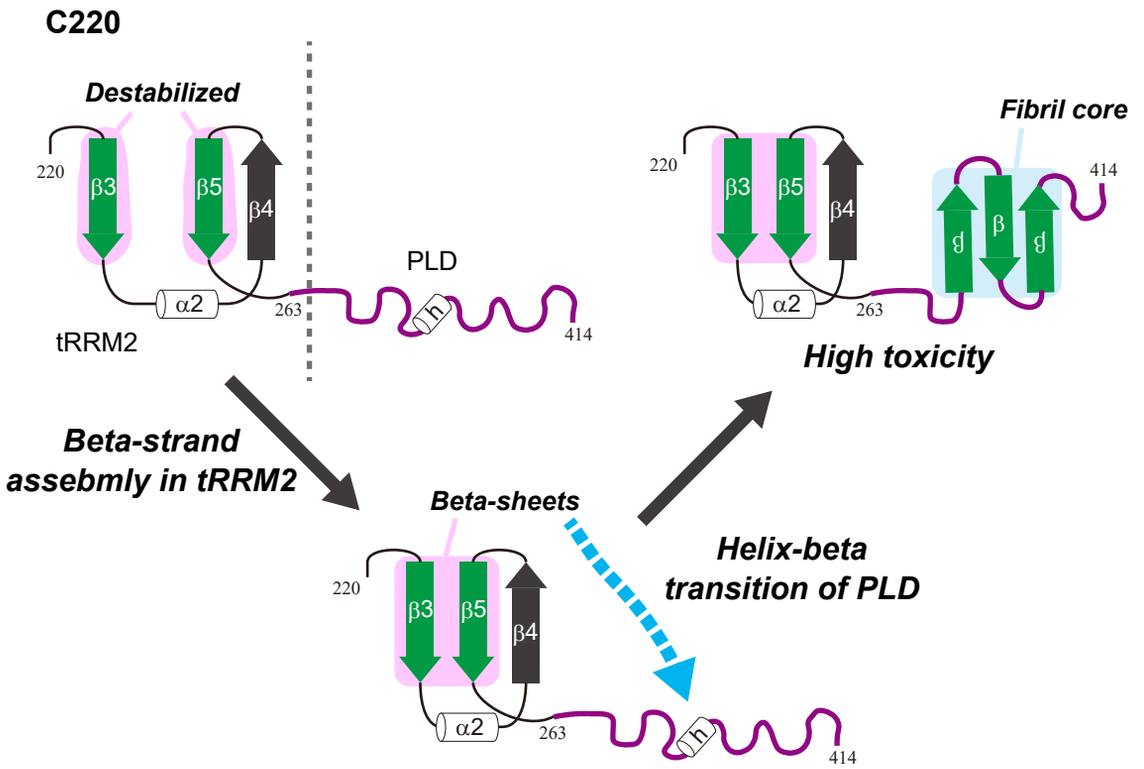
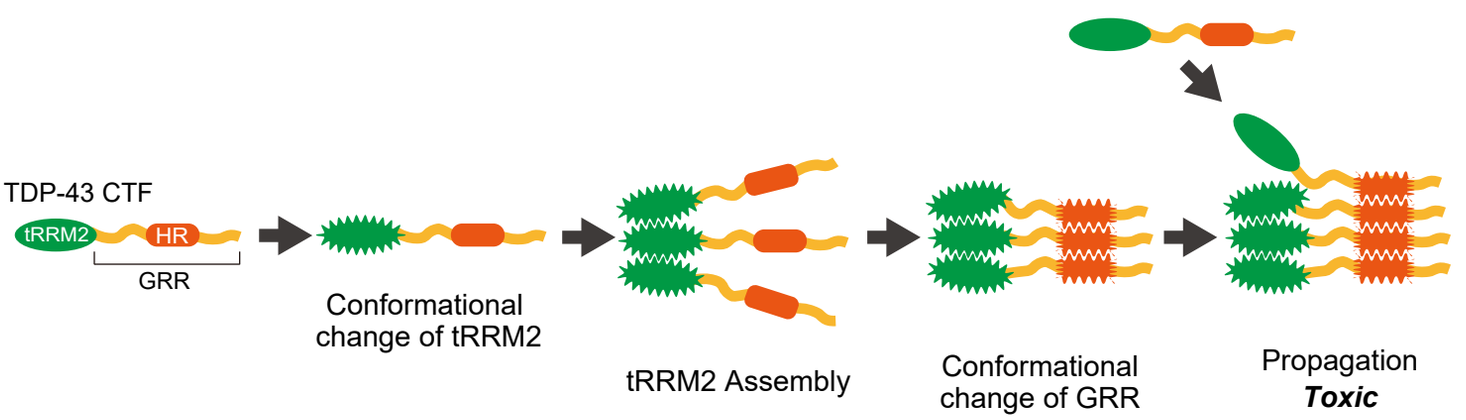


Figure 11

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.22.493003; this version posted May 23, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

(A)



(B)