Phase separation behavior of TDP-43 governs its protein interactome and regulation of alternative splicing

Yelyzaveta Zadorozhna[1,2#], Federico Uliana[1#]*, Emanuele Zippo[1,2], Anke Busch[3], Niklas Kretschmer[1], Simone Mosna[1,2], Yongwon Suk[1,2], Jiaxuan Chen[3], Carla Schmidt[4], Lukas Stelzl[1,3], Dorothee Dormann[1,3]*

1. Faculty of Biology, Institute of Molecular Physiology, Johannes Gutenberg University (JGU) Mainz, Germany

2. International PhD Programme (IPP) of IMB Mainz, Germany

3. Institute of Molecular Biology (IMB) Mainz, Germany

4. Faculty of Chemistry, Biochemistry, Johannes Gutenberg University (JGU) Mainz, Germany

# Equal contribution

* co-corresponding authors

Summary

TDP-43 is a nuclear RNA-binding protein regulating numerous steps in RNA metabolism, including alternative splicing. It is a major pathological hallmark of several neurodegenerative diseases, where it forms cytoplasmic aggregates in affected brain regions. TDP-43 can undergo phase separation (PS) and this condensation behavior may be linked to aggregate formation. Whether and how PS governs TDP-43's RNA regulatory functions remains poorly understood.
Here we utilized rationally designed mutations in the TDP-43 low complexity domain to tune TDP-43 PS, yielding a panel of TDP-43 variants with reduced propensity to form condensates ("PS-deficient"), and a panel of TDP-43 variants that form more irreversible, undynamic condensates ("solid-like") *in vitro* and in cells. Affinity proteomics coupled to mass spectrometry identified a set of interactors whose association with TDP-43 is PS-dependent. This includes multiple splicing regulators and the RNA helicase UPF1, which show increased interactions with solid-like TDP-43 variants. Consistently, we identified PS-dependent alternative splicing events that translate into measurable changes in RNA and protein abundance. Our results highlight that TDP-43 PS regulates RNA and protein homeostasis both directly, by altering a subset of TDP-43-dependent alternative splicing events, and indirectly, by changing interactions with other RNA regulatory factors.

Introduction

TDP-43 is an RNA-binding protein known to form aggregates in several neurodegenerative disorders, most notably amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD)[1,2]. However, TDP-43 pathology has also been identified in up to 60% of Alzheimer's disease cases[3,4], and a more recently defined dementia called Limbic-predominant Age-related TDP-43 Encephalopathy (LATE)[5]. TDP-43 is normally found predominantly in the cell nucleus, however in neurons and glial cells in affected brain regions TDP-43 gets lost from the nucleus and forms cytoplasmic aggregates, resulting in a nuclear loss-of-function[1,6,7]. Nuclear loss-of-function is further promoted by cytosolic TDP-43 aggregates, which trap more and more TDP-43 and deplete it from its physiological nuclear localization[8,9].

In healthy cells, nuclear TDP-43 is known to regulate many aspects of RNA metabolism, including alternative splicing, transcription, biogenesis of various micro-RNAs and long non-coding RNAs and alternative polyadenylation, while the smaller cytoplasmic pool of the protein is involved in the regulation of translation and mRNA transport in neurons[10]. Disease-associated pathological aggregation of TDP-43 and the resulting nuclear loss causes a dysregulation of these RNA regulatory processes, e.g. altered splicing of neuron-essential genes[11–14], alterations in translation[15] and changes in polyadenylation and mRNA stability[16,17,17–19]). Additionally, TDP-43 inclusions were found to co-aggregate with other cellular proteins, e.g. nuclear transport machinery components[20,21], RNA processing factors, e.g. splicing and translation factors[22–24], and UPF1, an RNA helicase involved in non-sense mediated decay (NMD)[16], as well as protein quality control components[25,26] and stress granule-associated proteins[27–29]. Some of these proteins also interact with TDP-43 under physiological conditions[30–33], suggesting that TDP-43 may sequester some of its physiological interactors into pathological aggregates, possibly causing additional proteome perturbations and functional deficits.

How TDP-43 aggregates form is still not well understood. It has been proposed that a first step could be coalescence of TDP-43 into physiological or stress-induced biomolecular condensates, such as stress granules[34–36], followed by intra-condensate demixing and TDP-43 aggregation[37]. *In vitro*, purified TDP-43 undergoes phase separation (PS) at physiological concentrations[38], and in cells it is found in numerous nuclear and cytoplasmic condensates, including paraspeckles and Cajal bodies[39,40] and neuronal transport granules[41]. TDP-43 PS is strongly mediated by its intrinsically disordered C-terminal low complexity domain (LCD), in which in particular aromatic residues, such as tryptophans and phenylalanines, but also aliphatic residues and an alpha-helix in the conserved region (CR)[42–46] contribute to a multivalent interaction network and drive PS. Additionally, oligomerization via the folded N-terminal domain is necessary for condensation of the full-length protein[38,40,47].

To what extent PS of TDP-43 is necessary for its regulatory roles in the cell, e.g. its diverse RNA processing functions, is not well understood. Generally, PS appears to be relevant for several DNA/RNA regulatory processes, including DNA damage repair[48,49], transcription[50–57], splicing [58], and translation[59]. For TDP-43, distinct PS phenotypes have been shown to dictate its RNA-binding repertoire and affect its roles in polyadenylation and autoregulation[60,61], to increase global translation[62], and to affect RNP granule transport for local translation in neurons[41,63]. Although these studies provided first evidence of the importance of TDP-43's condensation behavior for cellular function, a system-wide view of the consequences of altered TDP-43 PS is still missing.

Here we utilized a panel of TDP-43 variants with rational amino acid substitutions in the C-terminal LCD, to bidirectionally tune TDP-43 PS through very different types of mutations. Using *in vitro* assays with purified proteins we show that our designed TDP-43

mutants on the one hand include variants that have reduced condensation propensity combined with more dynamic condensate properties (“PS-deficient”), and on the other hand include variants that form amorphous condensates with lower dynamicity and reduced reversibility (“solid-like”). Utilizing a panel of cell lines with inducible expression of these different TDP-43 PS variants in the absence of endogenous TDP-43, we demonstrate that these PS properties are preserved in cells. Using immunoprecipitation and co-condensation assays in cellular lysate coupled to mass spectrometry, we find that solid-like TDP-43 variants exhibit enhanced interactions with several proteins involved in RNA-binding and splicing regulation. One such example is UPF1, for which our data indicate that sequestration into solid-like TDP-43 condensates compromises the protein’s functionality. Using mRNA sequencing, we identify a subset of PS-dependent splicing changes and link them to alterations in RNA and protein levels. Altogether, our data suggest that TDP-43 phase separation regulates RNA processing and thus the proteome both directly and indirectly. In the direct model, TDP-43 condensation directly affects splicing of some of its target genes, leading to alterations in mRNA and protein levels. In the indirect model, TDP-43 condensation alters the interaction with other RNA processing proteins and potentially modulates their functionality, thereby causing broad “secondary effects” on cellular mRNA and protein levels.

Materials and Methods

Cloning of constructs for bacterial expression

pJ4M TDP-43-TEV-MBP-His_6 backbone[38] (Addgene Plasmid #104480) was used to introduce mutations into the WT TDP-43 sequence for bacterial expression of recombinant proteins. Mutations were introduced using restriction cloning of synthetic DNA gBlocks (IDT) containing desired mutations. pJ4M TDP-43-TEV-MBP-His_6 12D/12A and G376V variants have been described earlier[64,65]. The mutations introduced into the distinct TDP-43 variants are listed in Table S1.

Cloning of constructs for the generation of mammalian HeLa T-REx Flp-In cell lines

pcDNA5-FRT-TO backbone (Invitrogen) expressing N-terminally myc-tagged TDP-43 WT[65] was used to introduce PS-altering mutations. The TDP-43 sequence within this backbone is resistant to a TDP-43-specific siRNA (see “siRNA treatments”). Mutations spanning multiple residues (3W, ΔCR, 5R) were introduced as synthetic DNA gBlocks (IDT) via BsrGI and XhoI restriction sites. The G376V point mutation was introduced using site-directed mutagenesis using Q5 polymerase and primers containing the desired mutation. pcDNA pcDNA5-FRT-myc-TDP-43-12D and myc-TDP-43-12A variants were generated earlier[65].

pcDNA5-FRT-TO constructs containing EGFP-tagged TDP-43 WT, 12D, 12A, ΔCR and 5R were generated using Gibson cloning. pcDNA5-FRT-TO-EGFP-G376V was generated using site-directed mutagenesis. 3W mutations were first introduced into the pEGFP-C1-hTDP-43 vector[66], and afterwards EGFP-3W insert was PCR-amplified and inserted into the pcDNA5-FRT-TO backbone.

Expression and purification of TDP-MBP-His_6 recombinant proteins

Recombinant TDP-43-MBP-His_6 WT and its variants were purified according to Wang et al.[67] with minor modifications. Briefly, proteins were expressed in BL21-DE3 Rosetta2 *E. coli*. Protein expression was induced using 0.5 mM IPTG and performed overnight at 16°C. Bacterial cell pellets were lysed in 50 mM Tris pH 8.0, 1M NaCl, 10 mM imidazole, 10% glycerol, 4 mM β-mercaptoethanol, supplemented with aprotinin, leupeptin hemisulfate and pepstatin A at 1 μg/μl each, as well as RNAse A and lysozyme (100 μg/ml each). Lysed material was sonicated and clarified by centrifugation. Protein purification was performed using Ni-NTA agarose beads (Qiagen) and elution was done using the lysis buffer containing 300 mM imidazole. Next, oligomeric and cleavage product species were removed by size exclusion chromatography on the Hiload 16/600 Superdex 200 pg column (GE Healthcare) in 50 mM Tris pH 8.0, 300 mM NaCl, 5 % glycerol, 2 mM TCEP. Fractions of interest were concentrated and stored in the SEC buffer.

Fluorescent labeling of TDP-43-MBP-His_6

To enable fluorescent detection of TDP-43 condensates *in vitro*, TDP-43-MBP-His_6 proteins were labeled at cysteine residues using Alexa Fluor 488 C5 maleimide (Thermo Fisher) at a low labeling efficiency (2-4%) to avoid potential effects of the dye on TDP-43 condensation ability. Protein was incubated with the dye at 100 : 7 (protein : dye) molar ratio for 2 hours at RT. Excess dye was removed by 4 sequential washes using the Amicon Ultra protein concentration columns (30 kDa MWCO, Merck Millipore). For assays where several TDP-43 variants were compared side by side, a maximum of 1% difference in labeling efficiency between protein variants was allowed.

Phase separation assays with purified proteins

Prior to all experiments, proteins were centrifuged at 20 000 g for 10 minutes at 4°C to remove any preformed aggregates/insoluble protein. Phase separation assays were performed in phase separation buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT). TDP-43 condensation was induced by incubating TDP-43 MBP-His_6 (WT or variants) at the indicated concentration with TEV protease (4 µM final concentration) and reactions were incubated for the indicated amount of time (15-60 min) at RT. Phase separation reactions for brightfield microscopy and FRAP experiments were set up in µ-Slide 18-Well-Flat IbidiTreat microscope chambers (Ibidi 81826), for fluorescence microscopy in 384-well plates pre-coated with 10% Pluronics F-127; for all other assays in Eppendorf tubes.

Bright field microscopy: Imaging of unlabeled TDP-43 condensates

Imaging of condensates formed by unlabeled TDP-43 was performed on a Zeiss Axio Observer 7 microscope in the phase contrast mode using a Plan-Apochromat 63x/1.4 Oil Ph3 objective, an Axiocam 705 monocamera and the Zeiss Zen software.

Confocal microscopy: Imaging of fluorescently labeled TDP-43 condensates

Prior to the phase separation reaction setup, 384-well plates were coated with 10% Pluronics F-127 for 1h at RT, and washed with water. Phase separation reactions were performed as described above and imaged using a Revvity Opera Phenix high content confocal microscope (IMB Mainz Microscopy Core Facility) equipped with the Revvity Harmony analysis software. Imaging was performed using the 63x water immersion objective, NA 1.15, binning 1. Alexa 488 imaging channel was used (excitation at 488 nm, emission range: 500-550 nm, time set to 60 ms, power set to 80%).

Condensate reversibility assay

Phase separation reactions were set up in test tubes using 5 µM TDP-43-MBP-His_6 as described in “Phase separation assay with purified proteins”. After the addition of TEV protease, reactions were incubated at RT for 15 or 60 minutes. After the incubation, the “1:1” reactions were immediately centrifuged at 20 000 g for 15 min, 4 °C. Dilution reactions (“1:5” and “1:10”) were diluted with phase separation buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT) 1:5 and 1:10 respectively, incubated for 15 min at RT to enable condensate dissolution, and centrifuged at 20 000 g for 15 min, 4 °C. Proteins from supernatant and pellet fractions were subsequently separated using SDS-PAGE, followed by Western blotting with a rabbit anti-TDP43 antibody diluted 1:1000 (Proteintech, 80001-RR). Western blot signal was quantified as % soluble, where 100 % indicates the sum of pellet and supernatant signal from undiluted 1:1 samples, and the signal coming from the 1:1, 1:5 and 1:10 supernatant samples is represented as a percentage of that value.

C_{sat} (saturation concentration) estimation

For each experiment phase separation reactions were set up in triplicates using 1 µM TDP-43-MBP-His_6 as described in “Phase separation assay with purified proteins”. After the addition of TEV protease, reactions were incubated at RT for 60 minutes and subsequently centrifuged at 20 000 g for 15 min, 4 °C. Obtained pellet and supernatant fractions were analyzed using SDS-PAGE, followed by SYPRO Ruby protein staining. Briefly, gels were fixed in 50% EtOH/10% acetic acid for 30min at RT. Afterwards, SYPRO Ruby solution was applied overnight. Finally, gels were destained in 10% EtOH/10% acetic acid for 30min at RT

and visualized. To obtain C_{sat} values, pellet and supernatant signals were quantified. Assuming the sum of pellet and supernatant signal is 1 μM TDP-43, a corresponding concentration value was assigned for the signal in supernatant alone, indicating an approximate saturation concentration. Each experimental replicate was conducted using three independent phase separation reactions.

Stable cell line generation and cell culture

Parental HeLa T-REx Flp-In host cell line (gift of Christian Behrends, LMU Munich) was grown in Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10% Tet-System Approved fetal bovine serum (Tet-FBS, Thermo Fisher Scientific, A4736401), 10 μg/ml gentamycin and 100 μg/mL zeocin in a humidified incubator at 37°C and 5% CO_2. To generate HeLa T-REx Flp-In cell lines with inducible expression of a single copy of myc- or EGFP-tagged TDP-43 (WT or variants), HeLa T-REx Flp-In host cells were transfected with a pcDNA5-FRT-TO vector containing the TDP-43 variant of interest and a pOG44 Flp recombinase vector (Invitrogen) at a 1:10 ratio (pcDNA5 : pOG44) using Lipofectamine 2000 according to manufacturer's instructions. After transfection, cells were placed into selection medium (DMEM supplemented with 10% Tet-FBS, 10 μg/ml gentamycin, 150μg/ml hygromycin B and 10 μg/ml blasticidin). Once an antibiotic-resistant stable population was established, cells were FACS-sorted using a Bigfoot cell sorter (Invitrogen, IMB Mainz Flow Cytometry core facility) using the 100 μm nozzle at 30 psi based on forward and side scatter profiles and seeded individually for clonal growth. HeLa T-REx Flp-in EGFP-TDP-43 cell lines were induced with 1 μg/ml of doxycycline prior to FACS, and were sorted based on EGFP fluorescence intensity. Individual clones were screened for homogeneity by Opera high-throughput imaging based on EGFP signal or using an anti-myc immunofluorescence staining (mouse 9E10 antibody, IMB Protein Production core facility). Total levels of each expressed TDP-43 variant were also confirmed by Western blotting using a mouse anti-TDP-43 antibody (diluted 1:2000, Proteintech 60019-2-Ig). The amount of TDP-43 expressed was further adjusted by titrating the concentration of doxycycline (see below).

Established cell lines were grown in DMEM supplemented with 10% Tet-approved FBS, 10 μg/ml gentamycin, 150 μg/ml hygromycin B and 10 μg/ml blasticidin. The expression of TDP-43 variants was induced by the addition of doxycycline for 24 hours. The following concentrations were used for individual myc-TDP-43-expressing cell lines: 0.002 μg/ml (WT, 3W, ΔCR), 0.008 μg/ml (5R), 0.03 μg/ml (12A), 0.5 μg/ml (G376V) and 10 μg/ml (12D). For EGFP-TDP-43-expressing cell lines, the following doxycycline concentrations were used: 0.05 μg/ml (WT), 0.1 μg/ml (ΔCR, G376V), 0.5 μg/ml (12D, 5R, 12A), 1 μg/ml (3W).

The parental HeLa T-REx Flp-In host cell line was used to generate cell lysates for the co-condensation experiments.

siRNA treatments

Cells were transfected with a non-targeting siRNA pool (ON-TARGETplus non-targetting control pool, Dharmacon) or a custom TDP-43-specific siRNA (sequence: 5'-GGCUCAAGCAUGGAUUCUAAGUCUU-3', Biomers) at a final concentration of 20 nM using Lipofectamine RNAiMax (Invitrogen) according to manufacturer's instructions. Cells were analyzed 72 hours post-transfection.

Heat stress treatments

To induce heat shock, cells were placed in a humidified incubator at 43°C and 5% CO_2 for 1 hour. Afterwards, cells were undergoing recovery in normal growth conditions (37 °C, 5% CO_2) for 3 hours.

Immunofluorescence staining

Cells were fixed using 3.7% formaldehyde in PBS for 8 minutes and permeabilized with 0.5% Triton in PBS for 5 min. Blocking to prevent unspecific binding of antibodies was done for 30 minutes at RT using blocking buffer (5% donkey serum in PBS with 0.1% Tween-20). Cells were incubated with primary antibodies diluted in the blocking buffer for 1.5 hours (anti-myc mouse 9E10 antibody diluted 1:100, IMB Protein Production core facility), followed by 3 washes using PBS with 0.1% Tween-20. Secondary antibody (anti-mouse conjugated to Alexa Fluor 488, Invitrogen, A21202) was applied in blocking buffer for 1 hour at RT, followed by another 3 washes in PBS with 0.1% Tween-20. Nuclei were stained with 0.5 µg/ml DAPI in PBS (Sigma-Aldrich, D9542) for 5 minutes.

Preparation of cell lysates for Western blotting and full proteome analysis

Cells were washed with ice-cold PBS and collected using a cell scraper. Collected material was centrifuged at 500 x g for 5 minutes at 4 °C, and the pellets were lysed with ice cold RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) for 15 minutes on ice. Cells were then centrifuged at 15 000 x g for 15 minutes at 4 °C, and the supernatant containing cellular proteins was used to quantify protein levels using BCA assay.

SDS-PAGE and Western blotting

For SDS-PAGE, lysates were supplemented with 4x Laemmli buffer (250 mM Tris pH 6,8, 40% glycerol, 8% SDS, 0.1% bromophenol blue, 4% β-mercaptoethanol) and boiled for 5 minutes at 95°C. Proteins were separated on a 10% SDS-PAGE gel, followed by a wet transfer onto a 0.2 µm nitrocellulose membrane. Blocking was done in 5% milk in TBST for 30 minutes, followed by an overnight incubation with the indicated primary antibody. Afterwards, the membrane was washed 3 times with TBST, and incubated with a species-specific secondary antibody conjugated to an IRDye 800CW or 680RD (Li-COR) for 1 hour at RT. Membrane visualization was done using the Odyssey M system (Li-COR). Quantification of the Western blot signal was done using the Image Studio Lite (LI-COR). Antibodies used for Western blotting: rabbit anti-TDP-43 (Proteintech, 80001-RR), mouse anti-TDP-43 (Proteintech, 60019-2-Ig), rabbit anti-GAPDH (Proteintech 60004-1-Ig), rabbit anti-UPF1 (Sigma-Aldrich, HPA019587).

Confocal fluorescence microscopy: Imaging of immunofluorescence-stained cells

To assess TDP-43 nuclear foci counts, cells were seeded onto 96-well plates and after immunofluorescence staining were imaged using the Revvity Opera Phenix high content confocal microscope (IMB Mainz microscopy core facility). Imaging was performed using the 63x water immersion objective, NA 1.15, binning 2. The following settings were applied for each imaging channel: DAPI (excitation at 405 nm, emission range: 435-480nm, time set to 100 ms, power set to 50%), GFP (excitation at 488 nm, emission: 500-550 nm, time set to 200 ms, power set to 100%).

Quantification of nuclear TDP-43 foci was performed using the Revvity Harmony analysis software. Cell nuclei were segmented based on the nuclear myc-TDP-43 signal to ensure all

cells expressing enough of the protein of interest are included. Cells on image borders were omitted. Spots were identified using the spots segmentation and further filtered using the linear classifier function to ensure segmentation of foci that specifically represent TDP-43. For cells that had undergone heat shock (1h HS + 3h recovery), this parameter was adjusted to detect larger, brighter foci formed by TDP-43.

Fluorescence recovery after photobleaching (FRAP)

All FRAP experiments were performed using the STELLARIS 8 FALCON (Leica Microsystems) confocal microscope equipped with the LAS X imaging platform. The images were obtained using a HC PL APO CS2 63x/1.4 oil objective for all experiments.

FRAP of in vitro-formed TDP-43 condensates

For FRAP experiments, TDP-43 condensates were formed from fluorescently labelled TDP-43-MBP-His_6 protein as described in "Phase separation assay with purified proteins". PS-deficient TDP-43 variants (3W, ΔCR, 12D) were analyzed at 40 μM protein concentration, and solid-like TDP-43 variants (5R, 12A, G376V) were analyzed at 15 μM protein concentration. Bleaching experiments were performed between 15 and 30 minutes after TEV cleavage for solid-like variants, and between 15 and 45 minutes after TEV cleavage for PS-deficient variants.

For pre-bleach and post-bleach imaging, 476 nm excitation line from White Light Laser and HyD S3 detector in counting mode (emission band 500 nm - 739 nm) were used. Half-bleaches were done within the selected region of interest in "zoom in" mode using an OPSL488 laser at 488 nm excitation and laser intensity of 25. Five pre-bleach frames were acquired, followed by photobleaching, then 50 post-bleach frames in "minimize mode", and 30 additional frames at 2-second intervals. The pixel size for all images corresponds to 60.18 nm (scan format 512 x 512 px, zoom 6.0).

T-REx Flp-In FRAP of EGFP-TDP-43-containing nuclear foci in HeLa cells

Cells were grown on 8-well chambered slides (Ibidi, 80826) and subjected to TDP-43 knockdown using siRNA (72h total), doxycycline induction (24h total) and, where appropriate, heat stress treatment as described above. FRAP experiments were conducted on live cells at 37°C and 5% CO_2.

For pre-bleach and post-bleach imaging, 476 nm excitation line from White Light Laser and HyD S3 detector in counting mode (emission band 500 nm - 550 nm) were used. Photobleaching was done using the 488 nm excitation line from White Light Laser at intensity of 100. For non-stressed cells, regions of interest were photobleached using the "zoom in" mode, while for heat-shocked cells a point bleach was used. Five pre-bleach frames were acquired, followed by photobleaching, then 40 post-bleach frames in "minimize mode", and 10 additional frames at 2-second intervals. The pixel size for all images corresponds to 72.2 nm (scan format 512 x 512 px, zoom 5.0).

Analysis of FRAP data

All images were prepared using Fiji distribution of ImageJ[68]. Mean fluorescence intensity was quantified in three regions of interest (Roi1 – bleached region, roi2 – whole condensate/cell nucleus, roi3 – background). Further data normalization (double) and quantification were done using the easyFRAP software[69]. Obtained curves were graphed using GraphPad Prism

9, and significance was determined based on area under the curve (AUC) values for each curve obtained.

Molecular dynamics simulations.

Molecular dynamics (MD) simulations were performed using the CALVADOS3 coarse-grained model[70]. In this framework, each residue was represented by a single bead, connected by harmonic bonds, and solvent effects, including ionic strength (fixed at 150mM), were treated implicitly through effective pairwise interactions. In particular, van der Waals interactions were modeled using the Ashbaugh-Hatch potential (a modified Lennard-Jones with attractive part scaled down based on the hydropathy of each specific amino acid) and the salt-screened electrostatics through a Debye-Hückel potential. Within folded regions, beads were positioned at the center of mass of each residue rather than at the Cα atom, and the structures were conserved using elastic networks, as described by Cao et al.[70]. Disordered regions, instead, were modeled as flexible chains.

Protein structures were obtained from AlphaFold2 predictions downloaded in January 2025[71]. Folded domains were defined based on domain annotations from the TED (The Encyclopedia of Domains) database (downloaded August 2026)[72].

All simulations were performed using Langevin dynamics with an integration time step of 0.01 ps and carried out using the python package HOOMD-blue version 3.8.184[73] together with GSD version 2.8.1 for trajectory storage and analysis.

Simulation codes and the HOOMD-blue plugin for the Ashbaugh-Hatch pair potential are available on the following GitHub page (https://github.com/ezippo/ hoomd3_phosphorylation, https://github.com/ezippo/ashbaugh_plugin)[74].

Calculation of contacts between TDP-43 and interaction pairs

For the calculation, 20 random proteins from the list of TDP-43 interactors and 20 proteins from the dataset that were not identified as interactors (controls) were selected. In total the simulations have been performed on 40 interaction pairs.

For each interaction pair, we run a 1×10^8 integration steps (1 μs) long simulation with one TDP-43 chain and one interaction pair in a cubic box of 40nm side length at 250K. Intermolecular contacts between TDP-43 and its pairs were computed by pairwise Euclidean distances between residues accounting for periodic boundary conditions. For each residue pair, a contact was defined when the inter-residue distance was smaller than the sum of the CALVADOS Lennard-Jones σ parameters of the two interacting amino acids.

The calculation of contacts between TDP-43 and interaction pairs has been performed with two strategies. A) Calculating the ratio of frames where contacts between proteins have been identified over frames without contacts, statistical significance was assessed using a Wilcoxon rank-sum test. B) Contact counts were normalized by the length of the interaction partner protein, the length of TDP-43 (414 residues), and the total number of analyzed frames (2,000). Statistical analysis was performed on normalized contact frequencies for both control and interaction datasets. Normality of the distributions was assessed prior to hypothesis testing, and differences between groups were evaluated using an unpaired two-sided t-test (N = 20 per group). Residue-level interaction enrichment along the TDP-43 sequence was quantified by counting the number of contact events observed at each of the 414 positions. A null distribution was generated by randomly shuffling the observed contact

counts across positions 1'000 times. Enrichment for each residue was quantified as the $\log_2$ ratio of observed contact frequency and the median of the null distribution. Empirical *p*values were computed as the fraction of shuffled counts greater than or equal to the observed count.

Simulation of the interaction between one chain of UPF1 and TDP-43 condensates

The condensates in slab geometry were prepared by placing 100 chains of TDP-43 variants in a large cubic simulation box, which was shrunk to a side length of 35 nm and then expanded in the z-direction to 150 nm. Periodic boundary conditions were applied in all directions. Initial chain positions and orientations were randomly assigned.

To simulate the interaction of UPF1 chains with TDP-43 condensates, one chain of UPF1 was placed outside a previously prepared TDP-43 in-silico condensate. The system was simulated at a temperature of 275 K. Three independent replicas were run for each variant for 1×10^8 integration steps, corresponding to 1 µs of simulated time (2'000 frames). Also in this case, the first 1×10^7 steps (0.1 µs) of each simulation were discarded from the analysis to account for thermalization. Contacts were calculated as detailed in ("Calculation of contacts between TDP-43 and interaction pairs"). Density profiles were computed from the combined replicas by counting the number of residues per z-bin, normalized by the total number of residues of the species.

Affinity Purification coupled with Mass Spectrometry (AP-MS)

Antibody crosslinking to protein G Sepharose for immunoprecipitation

Prior to immunoprecipitation experiments, crosslinking of the mouse anti-myc antibody (clone 9E10, IMB Protein Production facility) to protein G Sepharose was performed according to Gersten & Marcgalonis[75].

Beads were equilibrated in binding buffer (20 mM HEPES pH 7.5, 150 mM NaCl) and incubated with the anti-myc antibody for 1 hour at RT in a 2:1 antibody:beads ratio. Beads were extensively washed with borate buffer pH 9.0, and crosslinked using 20 mM dimethyl pimelimidate (DMP) for 30 minutes at RT. The reaction was quenched using 0.2 M ethanolamine, and the beads were washed and stored in IP buffer (50 mM Tris pH 8.0, 150 mM NaCl).

Affinity purification of myc-tagged TDP-43 and MS preparation

HeLa T-REx Flp-In cells expressing a single copy of N-terminally myc-tagged TDP-43 WT or variants were subjected to TDP-43 knockdown (KD) and 48 hours later the expression of myc-TDP-43 variants was induced with doxycycline for 24h as described in "siRNA treatments" and "Stable cell line generation and cell culture".

For each affinity purification experiment, two 15-cm tissue culture dishes were harvested and processed for lysis and affinity purification of the bait protein as previously described by Uliana et al.[76]. As a control, non-induced HeLa Flp-in myc-TDP-43 WT cells with a TDP-43 knockdown were processed in parallel. The experiment was performed with three independent biological replicates. Cell pellets were collected, frozen, and subsequently lysed in 1500 µl of lysis buffer (0.5% NP-40, 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 2.5 mM $MgCl_2$, supplemented with protease inhibitors for 20 minutes on ice. Lysates were sonicated using the Bioruptor Pico (Diagenode) for 1 cycle of 45 sec "on" and 45 seconds "off". Lysates were then treated with Benzonase (50 U/mL) for 20 min at 8 °C. Insoluble material was removed by centrifugation at 20'000 × g for 20 min at 4°C. The cleared supernatant was

incubated overnight at 4°C with 50 µl of beads conjugated to 100 µg of an anti-myc antibody. Beads were then extensively washed with lysis buffer with and without detergent (total wash volume 5 ml). Beads with bound proteins were transferred to a 10 kDa MWCO centrifugal filter unit (Vivacon 500, Sartorius) and processed according to the FASP protocol[77]. Beads in solution were centrifuged at 8'000 g until dry and the proteins bound were denatured and reduced in 8 M urea and 5 mM TCEP prepared in 50 mM ammonium bicarbonate for 30 min and alkylated with 10 mM iodoacetamide for 30 min in the dark. Samples were washed three times with 25 mM ammonium bicarbonate and digested with 1 µg sequencing-grade trypsin (Promega) for 16 h at 37 °C. Proteolysis was quenched by the addition of 5% formic acid and peptides were recovered by centrifugation, purified through C18 StageTip columns according to the manufacturer's instructions (Affinisep), dried in a vacuum concentrator, and resuspended in 20 µL of 0.1% formic acid and 2% acetonitrile prior to mass spectrometry analysis.

MS acquisition

LC-MS/MS analysis was performed on an Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher) coupled to Dionex Ultimate 3000 RSLCnano System liquid chromatography system. Peptides were loaded on a commercial trap column (µ-Precolumn C18 PepMap100, C18, 300 µm I.D., 5 µm particle size) and separated using a reverse phase column Thermo PepMap100 C18 (50 cm length, 75 µm inner diameter, 35 µm particle size) across a gradient from 8 to 25% in 120 minutes (buffer A: 0.1% (v/v) formic acid; buffer B: 0.1% (v/v) formic acid, 95% (v/v) acetonitrile). The DDA data acquisition mode was set to perform one MS1 scan followed by a maximum of 20 scans for the top 20 most intense peptides with MS1 scans (R=70'000 at 400 m/z, AGC=3e6 and maximum IT=64 ms), HCD fragmentation (NCE=28%), isolation windows (1.4 m/z) and MS2 scans (R=17'500 at 400 m/z, AGC=1e5 and maximum IT=55ms). A minimum AGC target of 2'000 and dynamic exclusion of 20s was applied and charge states lower than two and higher than seven were rejected for the isolation.

Data analysis

Acquired spectra were processed using MaxQuant (version 2.4.13.0) with the integrated Andromeda search engine[78] and searched against the human UniProt reference proteome (downloaded on 27 September 2024; 20'427 entries), supplemented with a database of common contaminants. Database searches were performed with full tryptic specificity, allowing up to two missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification, while oxidation of methionine and protein N-terminal acetylation were included as variable modifications. The "match between runs" feature was enabled with a retention time window of ±0.4 min. The precursor and fragment mass tolerances were set to 20 ppm. Peptide-spectrum match (PSM) and protein FDR were controlled at 1%. Protein abundances were determined using the MaxLFQ algorithm, and only proteins identified with at least two peptides, including at least one unique peptide, were considered for quantification. In the dataset 2'560 proteins were identified, filtered to 2'189 (proteins must be identified in triplicate in at least one condition). Protein intensities were $\log_2$-transformed and median-normalized across samples. Missing values (~15%) were imputed by random sampling from a normal distribution generated for each protein based on the observed intensity distribution across samples. For each protein, the imputation distribution was centered at the minimum observed $\log_2$ intensity minus 3 and assigned a standard deviation equal to the standard deviation of the observed values. Statistical analysis was performed using analysis of

variance (ANOVA) comparing all affinity purifications to the control condition. TDP-43 interactors were defined as proteins that showed significant enrichment in at least one purification with a $\log_2$FC>1 and *p*value (BH adjusted) < 0.05 (226 proteins, excluding TDP-43 itself). The abundance of TDP-43 interactors was normalized for the abundance of the bait and compared to the abundance of TDP-43 WT. For cluster definition, the average abundance across replicates was calculated for each interactor and z-score–transformed. Interactors were subsequently clustered based on the similarity of their abundance profiles using Pearson correlation and the interactors were displayed in the network (Cytoscape v. 5.0.253) based on the interaction retrieved from protein-protein interactions repository (BioGRID v. 5.0.253 considering only physical interactions). GO characterization was performed using EnrichR, only terms comprising at least 4 genes for CC and MF and 9 for BP were considered. To reduce redundancy, GO terms passing a *p*value threshold of 0.05 were subjected to semantic similarity reduction using the reduceSimMatrix approach, which identifies representative parent terms based on a similarity threshold of 0.5. The entire dataset, including raw data, generated Tables and scripts used for the data analysis are available in the PRIDE repository (PRIDE: PXD074566)

Identification of High Confidence Interactor and generation of protein-protein interactions network

A curated analysis of reported interactors was performed. This includes previously published TDP-43 interactomics datasets with appropriate controls[20,25,31–33,79–81], integrated with protein-protein interactions annotated in BioGRID (v. 5.0.253 downloaded Jan 2027, including only the following interactions type "physical association", "direct interaction" ,"colocalization")[82] and manually curated interactions from the literature[83]. Redundancy between datasets was manually removed so that each interaction was reported only once. From the resulting list (2’048 interactors), high-confidence interactors (HCI) were defined as those identified in more than three independent experiments (230 HCI).

The TDP-43 protein–protein interaction network was generated from correlations among interactor profiles. Network clusters were characterized by Gene Ontology cellular component (CC) terms, and proteins were assigned to modules based on annotated interactions from BioGRID (v. 5.0.253) using a guilt-by-association approach and Gene Ontology CC terms.

Characterization of TDP-43 interactome features

Functional association of proteins with TDP-43. Protein-protein interaction data were obtained by STRING (v.12, “9606.protein.links.v12.0.onlyAB.txt”)[84] and interaction confidence score were converted into edge weight for network analysis. Edge costs were defined as the negative natural logarithm of the confidence score, allowing shortest-path computation to favor high-confidence interactions. For each protein (18’841) the shortest path to TDP-43 was calculated and a score is computed as the product of the confidence values of all edges in the path.

The enrichment for specific features within the TDP-43 interactome were assessed by comparing interactors to a random sampled control group matching the size of the TDP-43 interactors defined as binding dependently on the biophysical properties of the bait (N =195). Continuous protein features like AlphaFold2 pLDDT confidence scores (downloaded January 2025)[71], isoelectric point (pI), hydrophobicity (GRAVY index, Kyte Doolittle scale), and predicted secondary structure content, were calculated for each protein and compared

between groups by analyzing their distributions using an unpaired non-parametric Wilcoxon test. Isoelectric point, hydrophobicity index and prediction of secondary structure were computed using Peptides and DECIPHER R Package. Enrichment of functional annotations was evaluated using hypergeometric tests. RG, RGG and extended RGG consensus motifs (G(0–3)X(0–1)RG(1–2)X(0–5)G(0–2)X(0–1)RG(1–2))[85] were searched across protein sequences from the human proteome to identify motif-containing proteins. Annotation of phase-separation–associated proteins was obtained from PhaSepDB (version 2.1, LLPS-related entries only; downloaded December 2022)[86] and from CD-CODE version 2 (including only proteins experimentally observed in biomolecular condensates in at least three independent studies; downloaded January 2025)[87].

Co-condensation experiment

Sample preparation

HeLa total standardized cell lysates were prepared freshly for each experiment. Parental HeLa T-REx Flp-In host cells were lysed in a mild buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.5% NP-40, 2.5 mM $MgCl_2$, supplemented with a protease inhibitor cocktail (Sigma-Aldrich, PIC0004), phosphatase inhibitors (PhosSTOP, Roche, 04906837001) and Benzonase (50 U/mL)) for 30 minutes at 10°C. Obtained lysate was then centrifuged for 20 min at 20 000 g and protein concentration was measured by BCA. Phase separation reactions were set up by the addition of MBP-TDP-43-His_6 variants (10 µM final concentration), HeLa cell lysate (2 mg/ml final concentration) and TEV protease (5 µM final concentration) in the following buffer: 50 mM Tris pH 7.5; NaCl 150 mM, 2 mM DTT. Reactions were incubated for 45 minutes at RT.

Microscopy analysis

Phase separation reactions were set up as described above using fluorescently labelled TDP-43-MBP-His_6 and imaged in 384-well plates as described in “Confocal microscopy: imaging of fluorescently labeled TDP-43 condensates”. Quantification was performed using the Revvity Harmony analysis software. Individual condensates were segmented using the spots segmentation, and condensates on the image borders were omitted.

Immunoblotting analysis

For Western blotting and subsequent detection of TDP-43 and UPF1 levels, samples were centrifuged to separate pellet and supernatant. Proteins were subsequently separated using SDS-PAGE, followed by Western blotting using mouse anti-TDP-43 (diluted 1:2000, Proteintech, 60019-2-Ig) and rabbit anti-UPF1 (diluted 1:1000, Sigma-Aldrich, HPA019587) antibodies. For TDP-43 signal detection, equal amounts of pellet and supernatant were loaded and quantified. For UPF1 signal detection, pellet was overrepresented 5 times compared to supernatant to increase the detection sensitivity.

Sample preparation for MS analysis

Lysate collected before the addition of TEV protease and supernatant (SN) collected after the addition of TEV protease, 45 min incubation and centrifugation, was subjected to sample processing. For quality control, 0.05 µg of BSA was spiked into each sample as an internal standard. Samples were processed using the SP3 protocol[88]. Proteins were reduced (5 mM DTT), alkylated (15 mM iodoacetamide in the dark), and quenched (5 mM DTT). Proteins

were digested overnight with trypsin at 37 °C. After acidification with formic acid, peptides were purified using C18 StageTips for solid-phase extraction.

MS acquisition

LC-MS/MS analysis was performed on an Orbitrap Astral mass spectrometer (Thermo Fisher) coupled to an Vanquish Neo liquid chromatography system (Thermo Fisher). Peptides were separated using a 45-cm reverse phase column (inner diameter: 75 µm; ReproSil-Pur 120 C18-AQ 1.9-µm silica particles, Dr. Maisch GmbH) across a linear gradient from 5 to 40% in 31 minutes (buffer A: 0.1% (v/v) formic acid; buffer B: 0.1% (v/v) formic acid, 80% (v/v) acetonitrile). The mass spectrometer was operated in data-independent acquisition (DIA) with the following parameters: one full FTMS scan in Orbitrap (350–1,050 m/z) at 120,'000 resolution, 20 ms injection time, and AGC target (%)= 300, followed by 175 fixed windows of 4Da from 350 to 1050 m/z, 5ms injection time, and AGC target (%)= 500 (Astral detector). Precursor ions were fragmented with HCD, Normalized Collision Energy 26%, isolation window 4m/z.

Data analysis MS data

Acquired DIA spectra were searched using FragPipe computational platform version 23.0 [89], Philosopher version 5.1.1[90], DIA-NN version 1.8.2[91]. Peptide identification from MS/MS spectra was performed with MSFragger search engine against the Human Proteome reference dataset extended with contaminant list (http://www.uniprot.org/, downloaded on 07.2025, 20'453 proteins) and reverse decoy sequences. Mass tolerance was set to 10 ppm for precursor and for fragment ions. The search parameters were set to include fully tryptic peptides, allowing for up to one missed cleavage-sites, carbamidomethylation on cysteine residues as static modification and up to three variable modification including methionine oxidation and N-terminal acetylation. Peptide and protein inference was FDR controlled at 1% by PeptideProphet and ProteinProphet respectively. The generated library included entries for 159'647 precursors and '7487 protein groups. For the DIA analysis, extraction of quantitative data was performed with DIA-NN version 1.8.2 beta 8 with a maximum FDR of 1% and quantification strategy was set to "Robust LC". Precursor abundance values were extracted from the file "report.tsv" and only proteotypic precursors with intensities > 0 and a PEP < 0.01 were $\log_2$ transformed and median normalized for downstream analysis. For each protein quantified with at least two precursors in every condition, precursor-level filtering was applied. Within each condition, precursor intensities were centered on the median, and precursor variability was quantified as the mean squared error (MSE) across all conditions. Precursors were ranked by their average MSE, and the 50% most variable precursors were excluded, while ensuring that at least two precursors were retained per protein. For the quantification of TDP-43 precursors that included residues that are modified in the variant analyzed were discarded. For each condition, protein abundance was calculated from the top five most intense precursors. With the stringent filtering criteria applied to the dataset, no imputation of missing values was performed. Differential analysis for the abundance of 5'677 proteins was performed with limma package[92]. A linear model without intercept was fitted to protein abundance values across 32 samples representing eight experimental conditions (n = 4 independent replicates per condition). Moderated t-tests were applied using empirical Bayes variance shrinkage, and contrasts were defined to compare each TDP-43 variant against the WT TDP-43 +TEV condition, *p*values were adjusted according to Benjamini-Hochberg method. The entire dataset, including raw data,

generated Tables and scripts used for the data analysis are available in the PRIDE repository (PRIDE: PXD076450)

Dose-response co-condensation experiment (co-coprecipitation profiling)

Sample preparation for MS analysis

For the dose response depletion profiling, increasing concentrations of WT and 5R TDP-43 (0,1,2,5,10 µM) were spiked on a standardized lysate (2 mg/ml) and 10 µM of TEV protease was added to the lysate to remove MBP solubility tag. After 30 minutes at RT the lysates were subjected to centrifugation and 10 µl (~20 µg) supernatant was subjected to proteomics preparation. For quality control, 0.05 µg of BSA was spiked into each sample as an internal standard and samples were processed using the SP3 protocol[88] as described for the co-condensation experiment.

MS acquisition

LC-MS/MS analysis was performed on an Orbitrap Astral mass spectrometer (Thermo Fisher) coupled to an Vanquish Neo liquid chromatography system (Thermo Fisher). Peptides were separated using a 45-cm reverse phase column (inner diameter: 75 µm; ReproSil-Pur 120 C18-AQ 1.9-µm silica particles, Dr. Maisch GmbH) across a linear gradient from 5 to 40% in 31 minutes (buffer A: 0.1% (v/v) formic acid; buffer B: 0.1% (v/v) formic acid, 80% (v/v) acetonitrile). The mass spectrometer was operated in data-independent acquisition (DIA) with the following parameters: one full FTMS scan in Orbitrap (350–1,050 m/z) at 120,'000 resolution, 20 ms injection time, and AGC target (%)= 300, followed by 175 fixed windows of 4 Da from 350 to 1050 m/z, 5 ms injection time, and AGC target (%)= 500 (Astral detector). Precursor ions were fragmented with HCD, Normalized Collision Energy 26%, isolation window 4m/z.

Data analysis MS data

Acquired DIA spectra were searched using FragPipe computational platform version 23.0[89], Philosopher version 5.1.1[90], DIA-NN version 1.8.2[91]. Peptide identification from MS/MS spectra was performed with MSFragger search engine against the Human Proteome reference dataset extended with contaminant list (http://www.uniprot.org/, downloaded on 06.2025, 20'453 proteins) and reverse decoy sequences. Mass tolerance was set to 10 ppm for precursor and for fragment ions. The search parameters were set to include fully tryptic peptides, allowing for up to one missed cleavage-sites, carbamidomethylation on cysteine residues as static modification and up to three variable modification including methionine oxidation and N-terminal acetylation. Peptide and protein inference was FDR controlled at 1% by PeptideProphet and ProteinProphet respectively. The generated library included entries for 175'366 precursors and 7'946 protein groups. For the DIA analysis, extraction of quantitative data was performed with DIA-NN version 1.8.2 beta 8 with a maximum FDR of 1% and quantification strategy was set to "Robust LC". Precursor abundance values were extracted from the file "report.tsv" and only proteotypic precursors with intensities > 0 and a PEP < 0.01 were $\log_2$ transformed and median normalized for downstream analysis. For each protein quantified with at least two precursors in every condition, precursor-level filtering was applied to reduce technical variability. Within each condition, precursor intensities were centered on the median, and precursor variability was quantified as the mean squared error (MSE) across all conditions. Precursors were ranked by their average MSE, and the 50% most variable precursors were excluded, while ensuring that at least two

precursors were retained per protein. For each condition, protein abundance was calculated from the top five most intense precursors. With the stringent filtering criteria applied to the dataset, no imputation of missing values was performed. For all quantified proteins (n = 5'969), dose–response curve fitting was performed at the protein level using the protti R package[93]. A four-parameter logistic model was applied to protein abundance values as a function of TDP-43 spike-in concentration, with concentrations log-transformed prior to model fitting. Proteins co-precipitating with TDP-43 5R were identified by applying a series of filters to dose–response curve parameters. For each protein, values were retained only if the Hill coefficient was >0.7, the ANOVA *p*value was <0.1, and the correlation coefficient was >0.7. The area under the curve (AUC) was then compared between wild-type (WT) and TDP-43 5R samples. 403 proteins were classified as co-precipitating with TDP-43 5R based on the following 2 criteria: i) passed the filtering criteria in 5R but not WT, ii) had an AUC ratio between 5R and WT < 1. The entire dataset, including raw data, generated Tables and scripts used for the data analysis are available in the PRIDE repository (PRIDE: PXD076467).

Characterization of proteins co-precipitating with TDP-43 5R

The characterization of features of co-precipitating proteins (functional association to TDP-43 and enrichment for proteins associated to phase separation according PhaSepDB and CD-CODE) was performed as described in the paragraph "Characterization of TDP-43 interactome features".

Analysis of molecular grammar

The sequence analysis was performed using localCIDER[94] following the methodology described by Han et al.[95]. A total of 128 sequence-based parameters were computed for each protein, including 44 global parameters, 14 local parameters (for which maximum, mean, median, and standard deviation were calculated), and 7 patch-based parameters (basic, acidic, charged, polar, aliphatic, aromatic, and proline). For local parameters, a sliding window of 5 residues was used, while patches were defined as contiguous runs of three or more identical or chemically similar residues. Runs separated by fewer than two residues were merged, and only patches with at least 5 residues were considered. Sequence parameters were calculated across the human proteome as a reference, and values were z-score normalized relative to the human distribution. For proteins co-precipitating with the TDP-43 5R variant or identified as PS-dependent interactors, an average z-score was computed and compared to the reference proteome. Positive Z-scores indicate enrichment, whereas negative values indicate depletion. Statistical significance of enrichment or depletion was assessed using Welch's two-sample t-test. Features significantly identified in at least one dataset (TDP-43 PS-dependent interactome or coprecipitating with 5R)(pval adj < 0.01) and with a coherent enrichment or depletion coherent enrichment between the two conditions are shown in the heatmap.

RNA-Sequencing

RNA extraction

Prior to RNA sequencing, the different stable HeLa T-REx Flp-In myc-TDP-43 cell lines were grown on 10 cm plates and subjected to siRNA treatment (TDP-43-specific or non-targeting control) for 72 hours. Doxycycline induction was performed 18 hours after the siRNA transfection to minimize potential secondary effects caused by long-term TDP-43 depletion.

Cells were washed with ice-cold PBS, collected using cell scrapers and centrifuged at 500 x g for 5 minutes. Total RNA was isolated from cell pellets using the PureLink RNA Mini-Kit (Invitrogen) according to manufacturer's instructions, including the DNAse treatment step.

Library preparation and RNA sequencing

NGS library prep was performed with Illumina's Stranded mRNA Prep Ligation Kit following Stranded mRNA Prep Ligation ReferenceGuide (April 2021) (Document 1000000124518 v02). Libraries were prepared with a starting amount of 1000 ng and amplified in 8 PCR cycles. Two post PCR purification steps were performed to exclude residual primer and adapter dimers. Libraries were profiled in a D5000 Screen Tape on a tape Station 4200 (Agilent technologies) and quantified using the Qubit 1x dsDNA HS Assay Kit, in a QubitFlex Fluorometer (Invitrogen by Thermo Fisher Scientific). All 63 samples were pooled in equimolar ratio and sequenced on 2 NextSeq2000 P4 (100 cycles) FCs,SR for 1x 116 cycles plus 2x 10 cycles for the dual index read and 1 dark cycle upfront R1.

RNA sequencing data analysis

Basic quality controls were done using FastQC (v0.12.1) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Prior to mapping, possibly remaining adapter sequences were trimmed using Cutadapt (v4.4)[96] including a NextSeq-specific quality trimming at the 3' end of reads (--nextseq-trim=20). A minimal overlap of 5 nt between read and adapter was required and only reads with a length of at least 30 nt after trimming (--minimum-length 30) were kept for further analysis.

Reads were mapped using STAR (v2.7.10b)[97] allowing up to 4% of the bases to be mismatched (--outFilterMismatchNoverReadLmax 0.04 --outFilterMismatchNmax 999) and with a splice junction overhang of 115 nt (--sjdbOverhang 115). Genome assembly and annotation of GENCODE[98] release 43 were used during mapping. Secondary hits were removed using Samtools (v1.17)[99].

Uniquely mapped exonic reads per gene were counted using featureCounts of the Subread tool suite (v2.0.0)[100] with parameters --donotsort -t exon -s2. For each mutant, genes differentially expressed between conditions were detected using the DESeq2 R package (v1.42.0)[101].

Alternative splicing analysis

Alternative splicing events were detected using rMATS (v4.1.2)[102]. For each mutant, different conditions were compared pairwise. When running rMATS, the detection of novel splice sites was enabled (--novelSS), clipped alignment were allowed to be used (--allow-clipping) and parameters for reverse-stranded single-end reads were set (--libType fr-firststrand -t single). The read length was set to 116 (--readLength 116), but variable read lengths were allowed (--variable-read-length). The same annotation as during mapping (GENCODE release 43) was used when running rMATS.

Differential analysis of skipped exons was conducted for rescue-versus-KD and KD-versus-NTC conditions. Events were filtered based on the following criteria: average junction counts (IJC and SJC) ≥ 10, false discovery rate (FDR) < 0.05, and inclusion level difference (ΔPSI) > 0.1. Genomic coordinates of significant splicing events were extracted to retrieve sequences (human reference genome, hg38) spanning 500 nucleotides flanking the alternative exons.

Motif enrichment analysis was performed using AME (Analysis of Motif Enrichment) [103] from the MEME suite, utilizing 6-mer motifs or motifs from the CISBP-RNA database.

TDP-43-dependent alternative splicing events were defined as those meeting the following criteria: average junction counts (IJC and SJC) ≥ 10 and significant changes (FDR < 0.05) in the WT condition. Specifically, events were classified as either: (i) "Rescue of exon inclusion" (185 events), where exon inclusion decreased upon knockdown and increased upon rescue, or (ii) "Rescue of exon skipping" (240 events), where exon skipping increased upon knockdown and decreased upon rescue. For TDP-43-dependent events, the ΔPSI observed in the rescue-versus-KD condition for WT was subtracted from the corresponding ΔPSI for each variant. PS-dependent events were defined as those with overall variability exceeding the average variability within solid-like and PS-deficient groups (total: 234 events).

Sashimi plots were generated using the “rmats2sashimi” pipeline which creates sashimi plot visualizations of the rMATS outputs. BAM files and events of interest identified earlier using rMATS were used for the visualization. The pipeline for Sashimi visualization is available under: https://github.com/Xinglab/rmats2sashimiplot.

Differential gene expression analysis

Differential gene expression analysis was performed using precomputed outputs from DESeq2 for pairwise comparisons between Rescue and KD conditions. Differential expression datasets from all TDP-43 variants were combined, retaining genes with at least two significant changes per condition group (solid-like and PS-deficient), resulting in a total of 454 genes. After z-score normalization, PS-dependent genes were defined as those with overall variability exceeding the average variability within solid-like and PS-deficient groups (total: 114 genes).

Gene Ontology (GO) enrichment analysis was conducted using the clusterProfiler package, focusing on Biological Process (BP) terms with q-values < 0.01. Redundant GO terms were reduced using the rrvgo package, clustering terms based on semantic similarity (threshold = 0.7). The median z-score of genes associated with each GO term was used to represent changes.

To compare gene expression between solid-like and PS-deficient conditions, raw read counts were normalized using DESeq2. Genes with low expression (minimum count ≤ 10 across samples) were excluded prior to normalization. Normalized counts were $\log_2$-transformed for comparisons between: (i) solid-like-versus-PS-deficient and (ii) KD-versus-rescue conditions. Median expression values were calculated for each gene per condition, and differences between groups were computed.

Gene set enrichment analysis was performed using clusterProfiler and enrichment plots packages. Ranked gene lists were generated based on $\log_2$ fold changes (solid-like-versus-PS-deficient). UPF1-dependent gene sets were compiled from published datasets[104–106], including genes identified in at least two of the three studies conducted in HeLa or HEK293 cell lines (310 genes in total). Genes identified in *Imamachi et al*. (Table S4), *Fritz et al.* (Table EV2_Supplementary; siUPF1total.vs.NT.log2.fold.change > 1 & FDR < 0.05), and *Colombo et al.* (Table S2; UPF1_FDR < 0.01) were considered for this analysis.

Full proteome characterization

Sample preparation

HeLa T-REx Flp-In cells myc-TDP-43 lines (WT or variants) were subjected to TDP-43 KD knockdown (KD) for 73 hours. Doxycycline induction was performed 18 hours after the siRNA transfection.

For each condition one 15-cm tissue culture dish was harvested and lysed as described in "Preparation of cell lysates for Western blotting and full proteome analysis", and the supernatant containing cellular proteins was used to quantify protein levels using BCA assay. 20 µg of the total proteome were processed using the SP3 protocol[88]. Proteins were reduced (5 mM DTT), alkylated (15 mM iodoacetamide in the dark), and subsequently quenched (5 mM DTT). Proteins were digested overnight with trypsin at 37 °C. After acidification with formic acid, peptides were purified using C18 StageTips for solid-phase extraction.

MS acquisition

LC-MS/MS analysis was performed on an Orbitrap Astral mass spectrometer (Thermo Fisher) coupled to an Vanquish Neo liquid chromatography system (Thermo Fisher). Peptides were separated using a 45-cm reverse phase column (inner diameter: 75 µm; ReproSil-Pur 120 C18-AQ 1.9-µm silica particles, Dr. Maisch GmbH) across a linear gradient from 5 to 40% in 31 minutes (buffer A: 0.1% (v/v) formic acid; buffer B: 0.1% (v/v) formic acid, 80% (v/v) acetonitrile). The mass spectrometer was operated in data-independent acquisition (DIA) with the following parameters: one full FTMS scan in Orbitrap (350–1,050 m/z) at 120,'000 resolution, 20 ms injection time, and AGC target (%)= 300, followed by 175 fixed windows of 4Da from 350 to 1050 m/z, 5ms injection time, and AGC target (%)= 500 (Astral detector). Precursor ions were fragmented with HCD, Normalized Collision Energy 26%, isolation window 4m/z.

Data analysis

Acquired DIA spectra were searched using FragPipe computational platform version 23.1[85], MSFragger version 4.3, Philosopher version 5.1.2[90], DIA-NN version 1.8.2[91]. Peptide identification from MS/MS spectra was performed with MSFragger search engine against the Human Proteome reference dataset extended with contaminant list (http://www.uniprot.org/, downloaded on 12.2025, 20'453 proteins) and reverse decoy sequences. Mass tolerance was set to 10 ppm for precursor and for fragment ions. The search parameters were set to include fully tryptic peptides, allowing for up to one missed cleavage-sites, carbamidomethylation on cysteine residues as static modification and up to three variable modification including methionine oxidation and N-terminal acetylation. Peptide and protein inference was FDR controlled at 1% by PeptideProphet and ProteinProphet respectively. The generated library included entries for 257'251 precursors and 8'981 protein groups. For the DIA analysis, extraction of quantitative data was performed with DIA-NN version 1.8.2 beta 8 with a maximum FDR of 1% and quantification strategy was set to "Robust LC".

Protein group abundance were extracted from the file "report.tsv" and only protein groups with intensities > 0 and a PEP < 0.01 were tested for different normalization approaches[107]. Proteotypic precursors with abundance > 0 were $\log_2$ transformed and median normalized. For each condition, the abundance of 8'694 proteins was calculated from the top three most intense precursors. In case of missing values (less then 1% of the entire dataset), missing intensities were replaced by random values sampled from a normal distribution centered below the minimum observed value (min − 1) with a standard deviation matching the observed data. Differential analysis for the abundance of 8'694 proteins was performed with limma package[92]. A linear model without intercept was fitted to protein abundance values across 48 samples representing 12 experimental conditions (12D,3W,ΔCR,5R,12A,G376V plus/minus dox, n = 4 independent replicates per condition). Moderated t-tests were applied using empirical Bayes variance shrinkage, and contrasts were defined to compare each

variant rescue-versus-KD condition, *p*values were adjusted according to Benjamini-Hochberg method.

Proteins with at least 2 significant changes per condition (PS deficient and solid like) ($log_2FC > 0.5$ & *p*value adjusted < 0.05) (total 302 proteins) were retained for further analysis. A consistency score was defined as the ratio between the mean within-group standard deviation (solid-like and PS-deficient). Proteins with consistency score lower than the total standard deviation across all conditions were considered as PS-dependent proteins (in total 239 proteins). Finally, proteins were stratified based on the direction of regulation of the WT condition into upregulated (log_2FC rescue-versus-KD > 0) and downregulated (log_2FC rescue-versus-KD < 0). GO enrichment was performed using clusterProfiler (BH-adjusted q < 0.01). Redundant terms were reduced using semantic similarity (Wang method) via GOSemSim/rrvgo (similarity > 0.85), retaining representative terms based on significance. The entire dataset, including raw data, generated Tables and scripts used for the data analysis are available in the PRIDE repository (PRIDE: PXD076485)

Network analysis and multi-omics integration

Proteomic datasets of PS-dependent upregulated and downregulated proteins were integrated with transcriptomic and alternative splicing data. For the network analysis, four datasets were utilized: i) PS-dependent interactors of TDP-43 (195 proteins), ii) 234 PS-dependent alternative exons, iii) 397 PS-dependent genes, and iv) 239 PS-dependent proteins.

Direct targets were defined as proteins whose corresponding genes exhibited PS-dependent alternative splicing events. A total of 19 direct targets were identified, representing approximately 8% of the PS-dependent proteins.

Indirect targets were classified into two categories:

i) Cascade-based indirect targets, defined as proteins functionally linked with direct targets.
ii) Interaction-based indirect targets, defined as proteins functionally linked with PS-dependent interactors.

Functional links were established by integrating RNA-protein interaction data from the RNAInter database[108] with functional links from the STRING database[84].

RNA-protein interaction data were sourced from RNAInter and filtered to include only human interactions supported by experimental evidence (either strong or weak, excluding predicted interactions) with interaction scores greater than 0.25. Interactions were further refined to include only those involving mRNAs and proteins present in the dataset. For UPF1, functional links were supplemented with target genes identified in at least two of three published datasets[104–106].

Functional links from the STRING database (v12.0) were included based on a combined score threshold of >0.7.

A multi-layered network was constructed by integrating the following components: (i) PS-dependent alternative splicing events, (ii) PS-dependent RNA expression changes, (iii) PS-dependent protein abundance changes, (iv) PS-dependent interactors, and the functional links described above.

Analysis of PS-dependent protein regulation revealed that among the 131 proteins upregulated in the wild-type (WT) condition, 18 direct PS-dependent events and 64 indirect PS-dependent events mediated by PS-dependent interactors were identified. Conversely, for

the 107 proteins downregulated in the rescue-versus-KD condition, 1 direct PS-dependent event and 69 indirect PS-dependent events mediated by PS-dependent interactors were observed.

Results

In vitro analysis of designed TDP-43 variants establishes a panel of "phase separation-deficient" and a panel of "solid-like" TDP-43 variants

To establish whether the phase separation state of TDP-43 influences its functionality, we designed a panel of TDP-43 variants which we predicted to be either condensation-deficient or more condensation-prone. We reasoned that placing selected mutations that change e.g. the charge, hydrophobicity or aromaticity in different positions of the C-terminal LCD would enable us to tune TDP-43's condensation behavior in cells, eventually allowing us to study TDP-43 functionality across different condensation states while avoiding potential mutation-specific effects. Specifically, we designed a panel of six TDP-43 variants, aiming to obtain three TDP-43 variants with reduced condensation properties and three variants with enhanced condensation properties (Fig. 1A). In the 3W variant, three tryptophans (W334G, W385G, W412G) that are essential for LCD phase-separation[44] were replaced with glycines, which tend to have spacer properties in IDRs[67]. In the ΔCR variant, we removed the α-helix-forming conserved region (CR, amino acids 321-340), which is a known driver of TDP-43 phase separation[42,43,60–62,109]. We also utilized a phosphomimetic 12D variant, for which we previously reported a reduced condensation/more liquid-like behavior; and the corresponding 12A variants, which we previously showed to yield more amorphous/aggregate-like condensates[65]. In the 5R variant, we combined five disease-associated mutations (G295R, Q343R, G348R, G357R, G384R), all of which introduce an arginine[110], since arginine-specific interactions were shown to increase the viscosity of TDP-43 condensates[45]. Finally, we introduced the recently reported disease-linked point mutation G376V, which causes TDP-43 to form more amorphous, aggregate-like condensates[64].

To assess the phase separation behavior of the designed variants, we purified bacterially expressed TDP-43 variants with an MBP solubility tag [111], induced its condensation by TEV protease-mediated removal of the MBP tag in a physiological buffer and monitored condensate formation by bright field microscopy (Fig. 1B). In comparison to wild-type (WT) TDP-43, 3W, ΔCR and 12D clearly were condensation-deficient, with visible condensate formation starting only at 10 µM for ΔCR and 12D, and 20 µM for 3W (Fig. 1C). At high protein concentration (20 µM), the 3W and 12D variants formed only few small, round condensates, while ΔCR assembled into much larger condensates that readily fused (Fig. 1C). On the other hand, 5R, 12A and G376V formed small, amorphous or aggregate-like structures, which did not fuse but tended to form agglomerates, instead of round condensates seen for TDP-43 WT (Fig. 1D). Using TDP-43 variants sparsely labeled with a small fluorescent dye, Alexa Fluor 488 (AF488), we recapitulated these condensate phenotypes, while also being able to visualize very small condensates formed by the 3W, ΔCR and 12D variants at 5 µM protein concentration (Fig. S1A), which we were undetectable in assays with unlabeled protein.

As the above described TDP-43 variants led to striking changes in condensate morphology, we proceeded to assess condensate dynamics using fluorescence recovery after photobleaching (FRAP) analysis. To this end, we formed condensates from TDP-43 WT and the PS-deficient 3W, ΔCR and 12D variants sparsely labeled with AF488 at 40 µM protein concentration, performed half bleaches and measured the fluorescence recovery over the time course of ~ 2 min. All three PS-deficient variants showed a faster and higher degree of fluorescence recovery in the bleached region compared to TDP-43 WT (Fig. 1E, G). 3W, ΔCR and 12D condensates readily fused, consistent with their larger size compared to WT condensates. We did the same comparison for TDP-43 WT and the more amorphous 5R, 12A and G376V variants at 15 µM protein concentration. All three variants showed slower fluorescence recovery, with 5R showing the strongest and 12A the most moderate reduction

(Fig. 1F, H). To corroborate the lower dynamicity phenotype of the 5R, 12A and G376V variants, we tested reversibility of condensates formed by these mutants. To this end, 15 or 60 min old condensates were either centrifuged to separate soluble (supernatant) or condensate (pellet) fractions, or diluted 1:5 or 1:10, followed by an additional 15 min incubation and the same centrifugation step (see scheme in Fig. 1I). For TDP-43 WT, increasing levels of protein were detected in the soluble fraction upon dilution, indicating condensate reversibility. In contrast, the 5R, 12A and G376V variants all showed similar amounts of TDP-43 in the soluble fraction even after the dilution steps, indicating lower condensate reversibility or accelerated condensate ageing (Fig. 1J-L, Fig. S1C-D). To address whether any of the three less dynamic, more irreversible TDP-43 variants show a lower saturation concentration (C_{sat}) for condensate formation than TDP-43 WT, we quantified the TDP-43 concentration in the dilute phase by sedimentation analysis followed by SYPRO Ruby staining. This revealed a significantly decreased C_{sat} for 5R and a weakly reduced C_{sat} for G376V compared to WT, while no difference in C_{sat} was seen for the 12A variant (Fig. S1E,F).

Collectively, our *in vitro* characterization of the designed TDP-43 variants (Fig. 1A) demonstrates that by introducing rationally designed mutations in the LCD, we were able to successfully tune TDP-43 phase separation in both directions, creating a panel of phase separation-deficient, more dynamic and liquid-like TDP-43 condensates (3W, ΔCR, 12D), and a panel of amorphous, less dynamic and solid-like TDP-43 condensates (5R, 12A, G376V). For simplicity, we will refer to the first panel (3W, ΔCR, 12D) as “PS-deficient”, and to the second panel (5R, 12A, G376V) as “solid-like” TDP-43 variants.

Rationally designed LCD mutations tune TDP-43 phase separation behavior in cells

As a next step, we wanted to test whether our rationally designed TDP-43 variants tune TDP-43’s phase separation behavior also in cells. As phase separation is a highly concentration-dependent process, we utilized a cellular system where TDP-43 variant levels can be well-controlled. We integrated the different TDP-43 variants into HeLa Flp-In T-REx cells, in which a single copy of a gene-of-interest can be stably integrated and then expressed in a doxycycline-inducible manner. To study the different TDP-43 variants in the absence of endogenous TDP-43, we depleted endogenous TDP-43 using siRNA and then expressed TDP-43 WT or the different TDP-43 variants with an N-terminal myc tag for 24 hours (Fig. 2A). Western blots demonstrated that the TDP-43 knockdown was highly efficient, and by titrating doxycycline amounts, we were able to achieve similar and near-endogenous expression levels of all myc-TDP-43 variants, with the exception of 12D which tended to express at slightly lower levels (Fig. 2B). We reasoned that using a small epitope tag and close to endogenous expression levels should preserve TDP-43’s native phase separation behavior and functionality.

TDP-43 is known to enrich in nuclear foci, such as Cajal bodies, gems and paraspeckles, but also foci of unknown identity and function[40,112]. When we examined TDP-43 localization by high-throughput confocal microscopy in the above-described cellular model system under knockdown of endogenous TDP-43, we observed that the different TDP-43 PS variants showed varying tendencies to localize to such foci. All three PS-deficient variants (3W, ΔCR, 12D) remained more diffuse than TDP-43 WT, showing fewer nuclear TDP-43 foci per cell (Fig. 2C,D). On the other hand, solid-like TDP-43 variants (5R, 12A, G376V) showed similar nuclear foci numbers as TDP-43 WT, with some cells having more TDP-43-rich puncta compared to WT (Fig. 2C-D). Notably, all TDP-43 variants display physiological nuclear

localization, which enables us to compare TDP-43 functionality in the nucleus and excludes mislocalization-driven effects. TDP-43 is also reactive to cellular stress and in response to various types of cellular stress is recruited into large stress-induced nuclear foci[113,114]. To test how TDP-43's recruitment into stress-induced nuclear foci is affected by the PS-altering mutations, we exposed cells to heat stress (1 hour at 43°C) followed by 3-hour recovery at 37°C (Fig. S2A), as TDP-43 WT is prominently recruited into large nuclear foci during this recovery period. Also in this stress paradigm, all three PS-deficient variants (3W, ΔCR, 12D) showed fewer nuclear foci than TDP-43 WT, whereas the solid-like TDP-43 variants (5R, 12A, G376V) showed similar or higher numbers of stress-induced TDP-43 foci (Fig. S2B-C).

Additionally, we sought to determine whether the dynamics of TDP-43 in nuclear foci is altered in cells expressing PS-deficient versus solid-like TDP-43 variants. To this end, we performed fluorescence recovery after photobleaching (FRAP) experiments on HeLa Flp-In T-REx cells expressing EGFP-tagged TDP-43 WT or PS variants at similar levels. To this end, a region in the nucleus containing TDP-43-enriched foci was bleached with a high intensity laser and fluorescence recovery in that region was measured for ~1 minute. All PS-deficient TDP-43 variants showed significantly faster and more efficient fluorescence recovery than TDP-43 WT (Fig. 2E,F), indicating a higher dynamicity of the TDP-43 in cells expressing the 3W, ΔCR, and 12D variants. In contrast, all three solid-like variants showed slower fluorescence recovery than TDP-43 WT, which was significant for the 5R variant (Fig. 2E,F). We also performed FRAP experiments in the above-described heat stress/recovery paradigm (Fig. S2A), by performing point bleaches on the large, stress-induced nuclear TDP-43 foci. Also in these stress-induced nuclear TDP-43 assemblies, all three PS-deficient variants remained significantly more dynamic than TDP-43 WT, whereas a trend towards reduced fluorescent recovery was seen for all three solid-like variants (Fig. S2D-E).

In summary, the alterations in TDP-43 condensation behavior that we observed *in vitro* for our rationally designed LCD variants (Fig. 1A) translates into phenotypic changes in cells, affecting the localization of TDP-43 in physiological and stress-induced nuclear foci as well as the dynamicity of TDP-43 in these nuclear condensates.

The phase separation behavior of TDP-43 governs TDP-43 protein interactions

To evaluate whether the differential phase separation behavior of our rationally designed TDP-43 variants affects TDP-43 protein interactions in the cellular environment, we performed affinity-purification coupled to mass spectrometry (AP-MS). In the experimental setup, TDP-43 WT, three PS-deficient TDP-43 variants (3W, 12D and ΔCR) and three solid-like variants (5R, 12A and G376V) were expressed with a myc-tag in a doxycycline-inducible manner using the above-described HeLa Flp-In T-REx system in absence of endogenous TDP-43 (see Fig. S3A for expression levels of the different TDP-43 variants). After mild cell lysis, TDP-43 interactors were purified using an anti-myc antibody, and the copurified proteins were analyzed by mass spectrometry (Fig. 3A, Table S2). To filter the proteins that specifically interact with TDP-43, we performed a negative-control purification using non-induced TDP-43 WT cells. The number of proteins identified, and their abundance was consistent across all conditions (Fig. S3C,D), and TDP-43 was the most enriched protein purified, with a similar amount being copurified in all samples, except the uninduced control cells (Fig. S3B,E,F). A threshold of $\log_2 FC > 1$ and *p*value adjusted < 0.05 compared to the control purification in at least one condition led to the identification of 226 TDP-43 interactors. Gene ontology (GO) analysis shows that proteins copurified with TDP-43 have a strong enrichment of proteins that bind to RNA and regulate RNA metabolism (mRNA splicing via spliceosome, mRNA transport) and gene expression (Fig. 3B). Cellular components analysis

shows an enrichment for nucleus associated compartments, in agreement with TDP-43's primary localization in the nucleus, as well as membrane bound organelle (e.g. mitochondria) (Fig. 3B). Most of the enriched GO terms were associated with the solid-like TDP-43 variants (in particular 5R); which exhibited a notably high number of interactors involved in RNA-regulatory pathways (Fig. S3G-I). We benchmarked the interactors identified in the generated dataset against a repository of protein-protein interactions (BioGRID v. 5.0.253)[82] and other interactomics datasets acquired for TDP-43 using different approaches (AP-MS, BioID and APEX) in distinct cell lines[20,25,31–33,79–81] (see materials and methods). We defined high-confidence interactors (HCIs) as proteins detected in more than three independent interactomics datasets, to distinguish them from proteins identified less consistently. Our acquired dataset largely recapitulates previous TDP-43 interactome studies, with over half of the proteins having been previously reported as TDP-43 interactors. Among these, 39 proteins were identified as high-confidence interactors (Fig. 3C).

To evaluate whether the interacting proteins were differentially purified based on the phase separation properties of TDP-43 and to ensure that variations in interactor abundance were not due to differences in bait purification yield, the abundance of the identified interactors was normalized for the bait levels, and the resulting levels were compared to the WT condition (Fig. 3D). Unsupervised hierarchical clustering, as well as principal component analysis and cluster correlation of the analyzed purifications indicates two clear and distinct groups based on the phase separation properties of TDP-43 (solid-like and PS-deficient) (Fig. 3D, S4A,B). Notably, the solid-like TDP-43 variants show a massive copurification of interactors compared to the PS-deficient variants (Fig. 3D). Clustering the profile of proteins across all analyzed conditions led to the identification of two clusters: one containing interactors whose abundance is dependent on TDP-43 phase separation behavior (cluster 1, 195 proteins) and another group of interactors, whose abundance is independent of TDP-43 phase separation (cluster 2, 31 proteins) (Fig. S4C). Based on the clustering analysis we generated a network where we overlapped all identified TDP-43 interactors with annotated interactions (BioGRID v. 5.0.253). We then compared this network to 100 randomized control networks sampled from the BioGRID database. TDP-43 interactome network shows more interactions, a higher average number of connections per node, and greater transitivity (the tendency of nodes to form triangles) than the controls (Fig. S4D-F). These findings indicate that the network is highly dense, reflecting extensive interconnectivity among TDP-43 interactors, and highlight the organization into distinct modules (groups of proteins connected by physical and functional associations[115,116] (Fig. 3E). Notably, proteins within the same module share similar clustering profiles, and when they belong to cluster 1 (PS-dependent profile) they display an increased affinity for TDP-43 with solid like properties (see Fig. S4H for the abundance profile of the modules and Fig. S5 for the protein organization within the modules). Among the modules that interact with TDP-43 in a PS-dependent manner (Cluster 1, Fig. 3E), we identified the nuclear pore complex, the spliceosome (Fig. 3F), the nucleoplasm, and various membrane-bound organelles (ER, mitochondria membrane, extracellular vesicle and peroxisome) (Fig. S5). The analysis of network centrality (eigenvector) revealed that proteins forming the spliceosome module are highly linked and occupy central positions within the generated network (Fig. S4G). The spliceosome complex is functionally linked to TDP-43 due to the protein's known roles in splicing regulation[117–119], and indeed we identified the majority of the proteins belonging to this group to be high-confidence interactors of TDP-43.

Taken together, our findings indicates that numerous splicing regulators bind to TDP-43 and that the affinity of these interactions is governed by the phase separation properties of TDP-43.

Phase-separation-dependent TDP-43 interactors are enriched in distinct features

As a next step, we aimed to validate the TDP-43 interactors identified by AP-MS by verifying their functional association with TDP-43, as well as their potential for direct binding to TDP-43. In the first case, the network of proteins generated by the STRING database (v.12)[84] including direct and indirect functional relationships was used to calculate minimal-cost paths distances between identified interactors and TDP-43. Proteins identified as interactors in the AP-MS dataset ranked among the top 20% of TDP-43 functional associated proteins (Fig. S6A).

Secondly, we tested whether the identified interactors are more likely to bind TDP-43 when compared to a control group randomly generated from proteins that did not pass the interactor filtering criteria. Given that TDP-43 is predicted to be a largely disordered protein, we employed MD simulations with the CALVADOS3 one-bead-per-residue coarse-grained model[70] to characterize the dynamic interactions between protein pairs composed of one chain of TDP-43 and another chain constituted by either an identified interactor or a control protein[74,120]. The analysis identifies a significantly greater number of frames predicted to be in a bound state between TDP-43 and its interactors compared to the control group (*p*value = 0.004) (Fig. S6B,C). When normalizing the number of contact events by the total possible contacts (accounting for the length of the proteins involved)[121], TDP-43 interactors exhibited significantly more frequent interactions with TDP-43 than the control proteins (*p*value = $5.4e^{-4}$), suggesting a stronger propensity for these proteins to act as direct interactors (Fig. S6D). The simulation analysis reveals that the contacts mainly occur along the C-terminal LCD of TDP-43 (Fig. S6E).

To identify specific features among the PS-dependent TDP-43 interactors, we compared them to a randomly sampled group of proteins from the human proteome of the same size (N=195) as well as to the PS-independent TDP-43 interactors (N=31). We evaluated several parameters known to influence phase separation, including isoelectric point, hydrophobicity, protein structural content, secondary structure composition and the presence of small linear motifs[67]. No significant differences were observed in the isoelectric point or hydrophobicity of the PS-dependent TDP-43 interactors compared to the control groups (Fig. S6F,G). However, PS-dependent TDP-43 interactors exhibit significantly lower pLLDT confidence score, a good proxy for protein disorder (*p*value=$1.9e^{-4}$, Fig. S6H). Furthermore, the PS-dependent TDP-43 interactors are characterized by a reduced helix content, accompanied by an increased proportion of coiled regions (Fig. S6I).

Short linear motifs are short protein regions characterized by low amino acid complexity and high flexibility, which can promote multivalent interactions and drive phase separation[122,123]. One such motif, the arginine-glycine-rich RGG motif, is well known to promote protein-RNA interactions via electrostatic and π-π interactions and is known to promote protein phase transitions[124,125]. As several TDP-43 interactors are involved in RNA binding (Fig. 3B), we calculated the enrichment of RG, RGG and extended RGG motifs[85] among the PS-dependent TDP-43 interactors and identified a significant enrichment of these motifs (*p*value = $6.9e^{-7}$, Fig. 3G, Fig. S6J,K).

The lower pLLDT score, the shift in secondary structure composition and the enrichment in RG/RGG motifs support the hypothesis that the interactions of this group of proteins with TDP-43 might be governed by phase separation, hence we sought to analyze to which extent the interactors are themselves prone to undergo condensation. Notably, among the list of PS-dependent interactors, 67 and 27 proteins were annotated in CD-CODE[87] and PhaSepDB[86], respectively, two independent repositories cataloging biological molecules that were

experimentally shown or computational predicted to undergo phase separation (Fig. 3H and Fig. S6L).

Collectively, we identified common features among the PS-dependent TDP-43 interactors, e.g. a high degree of disorder, enrichment in RG/RGG motifs and high PS propensity. This suggests that their association with TDP-43 might come from a multivalent network driven by phase separation, where solid-like TDP-43 variants form a stronger network of interactions, thus selectively sequestering/enriching certain interactors more efficiently than TDP-43 WT or PS-deficient TDP-43 variants.

Phase separation properties govern co-condensation of proteins with distinct sequence features in cellular lysates

To investigate whether proteins that preferentially interact with the solid-like TDP-43 variants in AP-MS experiments selectively co-condense with solid-like TDP-43 variants, we modified and further implemented the experimental approach used by Freibaum et al. for reconstituting nuclear and cytoplasmic condensates in total HeLa cell lysates[79]. We spiked recombinant MBP-tagged TDP-43 variants into a standardized HeLa cell lysate (2 mg/ml) and induced TDP-43 condensate formation by addition of TEV protease in the crowded multicomponent environment resembling the cellular milieu. Using fluorescently labelled TDP-43 variants, condensates formed from the recombinant protein in the lysate can be visualized through confocal microscopy; using reactions with unlabeled TDP variants and centrifugation, we can quantify the abundance of TDP-43 and co-depleted proteins in the supernatant by immunoblotting and mass spectrometry (Fig. 4A). The latter approach allows us to decipher which proteins co-partition into TDP-43 condensates with distinct biophysical properties and to what extent.

Microscopy analysis revealed that TDP-43 WT at 10 µM concentration efficiently forms condensates in total HeLa cell lysate (Fig. 4B). In line with our homotypic *in vitro* phase separation experiments (Fig. 1C,D), 3W, ΔCR, and 12D showed a strong deficiency in condensate formation, whereas 5R, 12A, and G376V formed more numerous condensates compared to WT (Fig. 4B,C), with 12A and G376V condensates occupying a larger area (Fig. S7A). In parallel, we quantified the abundance of TDP-43 in the supernatant and pellet fractions by immunoblotting and used mass spectrometry to measure the depletion of proteins from the supernatant. Both approaches showed that the major part (~70%) of solid-like TDP-43 variants were predominantly lost from the supernatant and precipitated into the pellet fraction after lysate centrifugation, while PS-deficient TDP-43 variants remained largely soluble and barely precipitated (Fig. S7B, Fig. 4D).

To identify proteins that co-condense with the different TDP-43 variants, we monitored the depletion of endogenous proteins from the supernatant by mass spectrometry (see Fig. S7C-E for quality control for the proteomics analysis). In total, 371 proteins were depleted ($\log_2$FC < -0.3, *p*value adjusted < 0.1) in at least one TDP-43 variant condition (Fig. S7G, Table S3). Most of the depletion events were observed upon the addition of solid-like TDP-43 variants (in particular 12A and 5R) (Fig. S7F). Among the identified depleted proteins, more than 10% were previously identified as PS-dependent TDP-43 interactors. Although the abundance changes for most endogenous proteins were very modest, proteins identified as PS-dependent interactors displayed a distinct trend: they were more abundant in the supernatant, i.e. more soluble, in presence of the PS-deficient TDP-43 variants compared to solid-like variants (Fig. 4E). This observation indicates that the phase separation properties of TDP-43 play a critical role in modulating protein partitioning into condensates, thereby possibly leading to a “trapping” of certain proteins in more solid-like TDP-43 condensates.

To validate that the observed co-depletion of endogenous proteins is a result of altered phase separation properties of TDP-43, we set up a dose-response experiment, since we reasoned that the co-depletion of proteins should follow a dose-response curve and be dependent on the concentration of added TDP-43. Thus, we set up phase separation reactions with increasing amounts of TDP-43 WT and 5R (due to its prominent solid-like condensation phenotype) and probed protein levels in the supernatant by mass spectrometry (Fig. 5F). In this dataset, we applied stringent filters and selected proteins that either showed a 5R-specific dose-response pattern or had a lower fitted area under the curve in 5R than in WT (see Fig. S8A-C for quality control and methodology). Dose-response depletion profiling retrieved 403 proteins that were specifically trapped by the 5R variant. Using this orthogonal approach, we validated ~11% of the interactors that bound to solid-like TDP-43 variants with higher affinity in our AP-MS experiment (22 proteins; *p*value = $5.18e^{-11}$, hypergeometric test) (Fig. 4G, Table S4). None of the PS-independent interactors were identified, and spliceosome components were the most enriched proteins co-precipitating with 5R condensates (Fig. 4H). The depletion profiles of proteins belonging to this cluster (DDX3X, FUBP3, HNRNPA3, HRNPR, HNRNPUL2), as well as of other nuclear related proteins (NUP160 and SMARCA5) are reported in Fig. S8D-J.

Several lines of evidence suggest that our co-precipitation profiling experiment is a complementary interactomics approach to identify proteins interacting via phase separation. First, although a cell lysate is devoid of native cellular architecture, potentially allowing TDP-43 to interact with proteins it might not encounter under physiological conditions, we identify a functional association between TDP-43 and the proteins that co-precipitate with 5R condensates. The functional association (calculated as the rank position from a minimal cost path to TDP-43) is lower compared to the PS-dependent interactors, but higher compared to a group of random proteins sampled from the identified proteins (Fig. 4I). Second, interactors that co-partition into 5R condensates were enriched in proteins known to undergo phase separation (Fig. 4J and Fig. S8K). Third, PS-dependent interactors and proteins co-precipitating with 5R condensates exhibit shared molecular sequence features, suggesting a common mechanism of interaction. To investigate this, we calculated 128 sequence parameters[94] (see Materials and Methods) to evaluate whether these features were significantly enriched or depleted in the AP-MS or co-precipitation datasets compared to the human proteome (Fig. S9). We focused on parameters that were significantly enriched or depleted in at least one of the two datasets and identified a modest yet significant correlation between them (*p*value = 0.04) (Fig. 5K). This correlation reflects a convergence of sequence features that underlie PS-dependent interactions and co-precipitated proteins. Specifically, these proteins were enriched in acidic and basic amino acids resulting in highly unbalanced charged sequences (high FCR and sigma) typical of strong polyelectrolytes dominated by long range repulsion (Fig. 4L). At the same time, they displayed reduced features that promote order, such as structured interaction propensity, hydrophobicity, proline content, aromaticity (tryptophan), and polar residues (serine, cysteine) (Fig. 4L).

Integration of these orthogonal approaches shows that AP-MS and co-precipitation profiling converge on proteins that are functionally linked to TDP-43, in particular splicing regulators, and are enriched in phase separating proteins with distinct sequence features. Our data suggests that solid-like TDP-43 condensates might trap such proteins, possibly leading to alterations of their function.

Solid-like TDP-43 condensates co-precipitate and trap UPF1

An interesting disease-linked protein that we identified among the TDP-43 PS-dependent interactors (Fig. 5A) and the proteins that co-precipitate with 5R in a dose-dependent manner (Fig. 5B), is UPF1. UPF1 is an RNA-dependent helicase involved in non-sense-mediated decay (NMD), which recently was identified as TDP-43 interactor that becomes dysfunctional in ALS[16,126]. We excluded that UPF1 levels are influenced directly or indirectly by TDP-43, since we observed low variability of UPF1 protein and RNA levels across the analyzed conditions (Fig. S10A,B). Thus, the enrichment of UPF1 with solid-like TDP-43 variants in the AP-MS experiment is unlikely to result from changes in UPF1 expression. To further validate the coprecipitation of UPF1 with solid-like TDP-43 condensates, we performed TDP-43 condensation experiments in the presence of a total HeLa cell lysate, followed by a pellet/supernatant fractionation. We compared TDP-43 variants with distinct phase separation properties and quantified the abundance of UPF1 in the supernatant and pellet fractions by immunoblotting. Pellet quantification confirmed that UPF1 co-condensed more strongly with all solid-like variants compared to the PS-deficient TDP-43 variants (Fig. 5C).

As a second approach, we simulated the trapping effect of UPF1 by TDP-43 variants *in silico*. For this purpose, we used MD simulations and modeled a TDP-43 condensate composed of 100 TDP-43 chains together with a single UPF1 chain or a control protein (albumin) placed outside of the condensate at the beginning of the simulation. This setup allows us to assess whether UPF1 preferentially partitions into the condensate compared to the control protein and whether partitioning depends on the phase separation properties of TDP-43. We quantified the number of contact points between UPF1/albumin and TDP-43 chains, normalizing for protein length to account for the higher contact probability of longer proteins[121] (Fig. S10C shows the distribution of interactions between UPF1 or albumin and the different TDP-43 variants). UPF1 showed stronger interactions (Fig. 5D, S10C) and greater colocalization with TDP-43 density profiles than with albumin (Fig. S10D). These interactions were further enhanced in the presence of solid-like TDP-43 variants, such as 5R and G376V (Fig. 5D, E, S10C), supporting the notion that UPF1 co-partitions into solid-like TDP-43 condensates. Overall, our data suggests that aberrant phase transitions of TDP-43 may trap UPF1 in more solid-like TDP-43 condensates, thus possibly compromising UPF1 function, as recently observed in ALS[16,126].

TDP-43 phase separation properties modulate alternative splicing and RNA levels

Since we identified many splicing regulators among the PS-dependent TDP-43 interactome and the proteins that selectively co-precipitate with solid-like TDP-43, we assessed whether the phase separation propensity of TDP-43 impacts splicing regulation. To this end, we performed mRNA sequencing on RNA extracted from cell lines expressing the different TDP-43 variants. The experimental design included a non-targeting control (NTC), knockdown (KD) of endogenous TDP-43, and rescue by inducible expression of TDP-43 variants in three different clones of each cell line (Fig. 6A, see Fig. S11A for quality control). The rescue restored TDP-43 to slightly higher levels than endogenous conditions, but was comparable across the different TDP-43 variants and clones, as confirmed by RNA levels and immunoblotting (Fig. S11B,C). Due to TDP-43's long-established role in regulating exon skipping and inclusion events[119,127–129], we focused specifically on this group of splicing events. We considered splicing events with significant (FDR< 0.05) changes in percent spliced in (PSI) upon TDP-43 rescue: a positive value (ΔPSI>0.1) indicates exon inclusion, while exon skipping is defined by a negative value (ΔPSI<-0.1) (Table S5). The number of identified events varied across the analyzed TDP-43 variants from ~250 (for 12A) to a maximum of ~1'000 events (for 5R). No clear pattern of the number of splicing events was

associated with the phase separation properties (PS-deficient and solid-like group) (Fig. S11D).

We next examined the sequence in the flanking region (500 nt) of the alternatively spliced exons to identify RNA motifs enriched upon rescue with different TDP-43 variants. Across all conditions, the 6-mers UGUGUG and GUGUGU were significantly enriched, confirming the well-characterized binding preference of TDP-43 for UG-rich sequences, and showing that these motifs are prevalent in the vicinity of the identified alternative splicing events[127] for both the PS-deficient and solid-like variants (Fig. 6B). At first glance, this appears to contrast with an iCLIP study reporting reduced affinity of phase separation-deficient TDP-43 mutants for UG motifs, or when one nucleotide of the UG 6-mer is altered (e.g., C-to-U or G-to-A substitutions "YG containing UG" and "YA-containing UG") (Fig. S11E)[60]. However, this apparent discrepancy might stem from methodological differences. Unlike iCLIP-based assays, which directly capture TDP-43 binding to RNA, our analysis is based on motif enrichment in regions close to alternatively spliced exons.

Enrichment in particular motifs could be a consequence of changes in direct RNA-binding by TDP-43, or changes in interactions of TDP-43 with other splicing regulators. To further test this possibility, we performed analysis of Motif enrichment (AME)[103] to assess the enrichment in RNA motifs known to be bound by RNA-binding proteins (CISBP RNA dataset). In cells expressing solid-like TDP-43 mutants, most prominently 5R, the enrichment of RNA binding motifs was the highest (Fig. 6C), in line with enhanced interactions of RNA-binding proteins with these TDP-43 variants identified by AP-MS (Fig. 3E). We examined the fingerprint of enriched motifs corresponding to previously characterized PS-dependent TDP-43 interactors (Fig. 3E). Among the significantly enriched motifs, we identified motifs associated with HNRNPC, HNRNPH2, HNRNPK, HNRNPL, and SYNCRIP (Fig. 6D), all of which are implicated in splicing regulation. Although for HNRNPC and HNRNPH2 the differences in motifs enrichment are negligible, for the other three (HNRNPK, HNRNPL and SYNCRIP) the motifs are prominently enriched in cells expressing solid-like TDP-43 variants, in particular 5R (Fig. S11F). Together, this analysis suggests that solid-like TDP-43 variants may regulate alternative splicing through the interaction and recruitment of other RNA-binding proteins that then modulate mRNA splicing.

Next, we analyzed alternative splicing events regulated by different TDP-43 variants for rescue-associated changes of exon inclusion ($\Delta PSI > 0.1$) and exon skipping ($\Delta PSI < -0.1$). The overlap of regulated events across conditions was limited, and no distinct group of events could be specifically attributed to the PS properties of different TDP-43 variants (Fig. S11G,H). To address this limitation, we focused on events showing a directional reversal between knockdown (KD) versus non-targeting control (NTC) and rescue versus KD comparisons, meaning that the rescue shifted ΔPSI significantly in the opposite direction of the knockdown effect in WT condition. This approach identified 425 rescue-responsive events, composed of 185 exon inclusion and 240 exon skipping events (Fig. 6E). Motif enrichment analysis of the flanking regions revealed significant enrichment of UGUGUG and GUGUGU motifs (Fig. S12A). Among the rescue-responsive events, we identified several well-characterized TDP-43 splice targets, including *AGRN*[130,131], *ATG4B*[130,132], *MADD*[127,133,134], *PFKP*[130,135], *RANBP1*[130], *BRD8*[133,136] and *POLDIP3*[129,135,137] (see Fig. S12B,C for Sashimi plots of *ATG4B* and *POLDIP3*). Most TDP-43-dependent alternatively spliced exon skipping/inclusion events were located in protein-coding regions, with fewer events occurring in long non-coding RNAs or 5'UTRs (Fig. 6F).

Among 425 rescue-responsive events, we identified 234 PS-dependent alternatively spliced exons. We defined PS-dependent alternatively spliced exons as those whose inclusion levels display lower variability within PS-defined groups (PS-deficient or solid-like) compared to the

variability observed across all experimental conditions; indicating consistent inclusion patterns within each variant class. A summary of the filtering workflow is provided in Fig. S12D. Hierarchical clustering of ΔPSI values (variant versus WT) for events associated with exon inclusion upon rescue revealed two distinct groups, corresponding to the PS properties of the TDP-43 variants (Fig. 6G). In contrast, when analyzing events that exhibit exon skipping upon TDP-43 depletion, we identified fewer PS-dependent events and the data did not cluster based on the phase separation properties of TDP-43 (Fig. S12E). Among the PS-dependent alternatively spliced exons, we identified several previously reported TDP-43 splice targets, including *MADD*[133], *SEMA3F*[119], *BRD8*[133], and *PILRB*[127] (inclusion levels and Sashimi plots for *BRD8*, and *SEMA3F* are shown in Fig. 6H and Fig. S12F,G). Altogether, these findings indicate that TDP-43 phase separation plays a critical role in modulating alternative splicing, either through the recruitment of additional RNA-binding proteins or by influencing the inclusion levels of alternatively spliced exons within TDP-43 splice targets.

Next, we investigated whether TDP-43 PS-dependent alternative splicing affects overall RNA abundance (Table S6). Specifically, the inclusion or exclusion of exons can disrupt the reading frame, potentially introducing premature termination codons that trigger RNA degradation or lead to the incorporation of cryptic exons, which can result in the synthesis of neopeptides[138]. Additionally, alternative splicing may modulate the composition of the 3' untranslated region (3'UTR), or changes to RNA secondary structure, thereby influencing RNA stability or susceptibility to degradation[58]. Differential analysis of RNA levels under WT conditions, comparing KD to NTC and rescue to KD, revealed that genes exhibiting the most significant changes were those previously showing TDP-43-rescue responsive alternative splicing (Fig. 6I). This phenomenon was specifically observed for exons that were included upon TDP-43 knockdownand skipped upon rescue with WT (Fig. S12H), highlighting the role of TDP-43 in the suppression of non-conserved exons[130].

To verify whether RNA abundance changes could be attributed to altered TDP-43 phase separation properties, we identified transcripts that were significantly up- or down-regulated when comparing rescue to TDP-43 KD in at least two variants per group (solid-like and PS-deficient). Subsequently, we filtered transcripts for which the standard deviation of rescue-versus-KD differences within PS-defined groups (PS-deficient or solid-like) was lower than the global standard deviation. This led to the identification of 298 genes, indicating that the phase separation properties of TDP-43 modulate RNA levels of numerous genes, including 11 that were previously found to undergo PS-dependent alternative splicing (Fig. 6J). Most PS-dependent transcripts map to protein-coding genes (Fig. S12I) involved in such pathways as Wnt signaling, ER stress response and ER-associated degradation (ERAD) (Fig. S12J).

However, analysis of RNA levels also indicated possible indirect regulatory effects by proteins interacting with TDP-43. One such example is UPF1, the protein which showed enhanced interaction, co-condensation and co-precipitation with solid-like TDP-43 variants in different, orthogonal assays (Fig. 5). We identified UPF1 RNA targets from published studies in HEK293 and HeLa cells[104–106], focusing on 310 genes reported in at least two independent studies (Fig. S12K). Gene set enrichment analysis (GSEA) confirmed that UPF1 target genes are upregulated upon TDP-43 depletion in our cellular system (Fig. S12L), in line with recently published data[16]. Next, we examined whether these genes exhibit enrichment, depletion, or no significant changes when comparing solid-like TDP-43 variants to PS-deficient variants. GSEA analysis using UPF1 target genes revealed significant enrichment of UPF1 target genes in cells expressing solid-like TDP-43 variants (Fig. 6K). This finding indicates that indeed trapping of UPF1 in solid-like TDP-43 condensates can compromise UPF1 functionality and alter degradation of transcripts via NMD.

Collectively, RNA profiling in cells expressing TDP-43 variants with distinct phase-separation properties reveals pronounced differences in RNA abundance and splicing between PS-deficient and solid-like states, establishing phase separation of TDP-43 as an important modulator of its RNA regulatory function. We observe that for many genes, changes in alternative splicing are tightly coupled to RNA levels, consistent with TDP-43-mediated repression of non-conserved exons. In addition to its direct effects, we found that TDP-43 phase separation likely impacts RNA levels indirectly by modifying interactions with other splicing or RNA regulatory factors, such as UPF1. Our data suggest that solid-like TDP-43 variants may sequester UPF1, thereby disrupting its function and resulting in the upregulation of UPF1 target transcripts.

TDP-43 phase separation regulates the cellular proteome directly and indirectly

After establishing TDP-43 PS-dependent regulation of alternative splicing and RNA-levels, we investigated whether TDP-43 phase separation impacts the regulation of the cellular proteome. To address this, we performed proteome profiling analysis to evaluate the effects of endogenous TDP-43 depletion, followed by re-expression of myc-tagged TDP-43 WT or the different TDP-43 PS variants (see scheme in Fig. S13A). By integrating RNA and proteome datasets, our goal is to define the functional role of TDP-43 phase separation in regulating alternatively spliced exons, RNA and protein levels, as well as to explore the potential role of PS-dependent interactors in regulating protein levels (Fig. 7A, Table S7).

Using a D.I.A. acquisition mode we analyzed the abundance levels of 8’694 proteins (filtering steps for the data analysis are shown in Fig. S13B), with a good coverage of the proteome dynamic range spanning from hundred pM to µM (according to OpenCell[139]) (Fig. S13C, quality control assessments for the proteomics dataset are shown in Fig. S13 D-F). Upon knockdown, TDP-43 protein levels are, on average, 5-fold lower ($\log_2$ scale) compared to the rescue conditions, which displayed relatively consistent levels across the different variants (see Fig. S13G,H for TDP-43 quantification by mass spectrometry and immunoblotting). We first analyzed TDP-43 WT to evaluate the impact of TDP-43 KD and rescue on protein expression levels. Following rescue, 289 proteins were significantly up- or down- regulated. This finding aligns with the crucial role of TDP-43 in RNA metabolism and gene regulation of hundreds of proteins[140], including TDP-43 itself[119,141]. The most pronounced changes in protein levels occurred for genes harboring alternatively spliced exons, and genes showing differential regulation at the protein level also exhibit corresponding changes at the RNA levels (Fig. 7B), in line with observations made by Kozareva et al.[140]. This data indicates a robust correlation between RNA-level regulation and protein-level changes (Fig. S13I), reinforcing the interconnected relationship between the levels of TDP-43, alternative splicing, RNA expression and protein abundance. The integrated analysis of these layers is illustrated in a flow diagram, which connects alternative splicing to RNA and protein expression levels (Fig. 7C). Notably, most of the exon skipping events are linked to elevated RNA and protein levels, underscoring the critical role of TDP-43 in regulating non-conserved exons, which in turn affects protein levels[130]. In contrast, for exon inclusion events, the link between RNA and protein levels is less evident. Overall, 50 out of 172 upregulated proteins are associated with splicing events or RNA changes (Fig. 7C), highlighting the profound influence of TDP-43-dependent RNA regulation on shaping the proteome.

As a next step we sought to investigate how the phase separation properties of TDP-43 influence proteome regulation. We performed a differential analysis comparing protein levels in the rescue-versus-KD conditions across all analyzed TDP-43 PS variants (Fig. S14A). Proteins significantly regulated in at least two variants per group were classified as PS-

dependent if their signal variability within the group (solid-like or PS-deficient) was lower than the overall variability. These PS-dependent proteins were further divided into two groups based on their regulation in WT rescue-versus-KD (131 up- and 107 down-regulated proteins). Among up-regulated proteins, we identified several key biological processes linked to protein quality control (ERAD pathway), protein clearance, cytoskeleton regulation and vesicle trafficking (Fig. S14B), with ERAD components also appearing in the PS-dependent RNA dataset (Fig. S12J). Downregulated proteins were enriched for roles in transcription activation during ER stress and lipid metabolism (Fig. S14C). Figure S15 provides details on the identified PS-dependent proteins, their relative abundance, and their association with PS-dependent alternative splicing and RNA levels. Analysis of relative abundance across the conditions reveals higher protein levels in solid-like TDP-43 variants compared to PS-deficient variants, underscoring the role of TDP-43 phase separation in regulating proteome composition.

Next, we investigated whether changes in PS-dependent protein abundance are linked to alternative splicing and RNA levels. We identified 19 PS-dependent alternative splicing events (as defined in the previous paragraph) occurring in proteins whose levels are PS-dependent. Notably, many of these events (11) involve exon skipping during rescue, which is associated with an increase in protein levels upon rescue. For four of these events, we monitored how PS-dependent alternative splicing influences RNA and protein levels (Fig. 7D). This analysis suggests that, for a specific subset of genes, TDP-43 phase separation directly influences alternative splicing, resulting in altered protein expression. We define these events as directly regulated by phase separation. Figure 7E,F illustrates two examples of opposing trends, where PS-dependent alternative splicing results in distinct protein levels. Approximately 13% of the PS-regulated proteins that are upregulated (18 proteins) and only one downregulated protein can be directly linked to TDP-43-dependent alternative splicing (Fig. 7G).

To address the origin of the remaining PS-dependent protein abundance changes, we followed an approach recently proposed by Kozareva et al.[140] which introduced a *"secondary cascade"* model demonstrating that the mis-regulation of a few genes upon TDP-43 knockdown leads to widespread changes on the proteome level. Only a limited number of functional associations between PS-dependent proteins could be established, however, when considering PS-dependent interactors, which we defined earlier in the interactomics experiment (Fig. 3), a different picture emerged. Several PS-dependent interactors are annotated as RNA binding proteins, harboring RGG motifs and playing a role in gene regulation. We identified that more than 50% of PS-up or down-regulated proteins are functional targets of these PS-dependent interactors (Fig. 7G, Fig. S16A,B reports functional link between PS-dependent interactors and PS-dependent proteins). We define these events as being indirectly regulated by phase separation through TDP-43 interactors. Figures S17, S18 summarize in a network the organization of PS-dependent proteins that are regulated both directly and indirectly by phase separation.

Two additional observations support the proposed indirect regulation by TDP-43 PS-dependent interactors. First, RNA-binding motif analysis around the alternatively spliced exons revealed binding signatures of PS-dependent TDP-43 interactors (e.g., HNRNPK, HNRNPL, and SYNCRIP) (Fig. 6D). Second, transcriptome analysis showed that several genes upregulated in cells expressing solid-like TDP-43 variants are known targets of UPF1 (Fig. 6K), an RNA decay factor identified in our datasets as a PS-dependent interactor (Fig. 5A,B). For some UPF1 targets (e.g. CHID1 and RNF213) we observed a correlation between the PS-dependent UPF1 binding to TDP-43, and increased levels of CHID1 and RNF213 (Fig. 7 H,I). These observations suggest a model in which UPF1 gets trapped in solid-like TDP-43 condensates, impairing its ability to efficiently perform RNA degradation. This may

result in increased abundance of UPF-dependent transcripts and corresponding proteins (Fig. 7J).

Together, our observations indicate that RNA and protein levels are regulated by TDP-43 phase separation properties and reflect a convoluted picture composed of the direct RNA-regulatory activity of TDP-43 via alternative splicing, and the indirect contribution of RNA-binding proteins whose function is modulated by the phase separation properties of TDP-43 (Fig. 7J).

Discussion

Tuning TDP-43 phase separation with diverse LCD mutations allows probing condensation-state-dependent functional changes in cells

While the role of phase separation in TDP-43 aggregation in neurodegenerative disease has been widely investigated[21,37,142,143], the relevance of TDP-43 condensation in supporting the protein's physiological functions in the cell is only beginning to be investigated. As phase separation of TDP-43 might be altered in disease, and targeting its phase transitions is even explored as a therapeutic strategy[20,143–145], it is of great importance to understand how and to what extent TDP-43 condensation plays a role for its functionality in RNA processing and global RNA/protein homeostasis.

Some efforts have already been made to address the functional relevance of TDP-43 condensation: one study deleted or mutated the alpha-helical conserved region (CR) in TDP-43, which drives TDP-43 condensation[42,43,60], and showed that these CR perturbances alter TDP-43's global RNA-binding profile, its autoregulation, and certain TDP-43-mediated 3'-end processing events[60]. A similar approach was adopted in mice, whereby deletion of the CR from mouse TDP-43 was shown to cause neuronal and behavioral impairments and to enhance translation globally via an increased association of TDP-43 with the translation machinery[62]. Other studies perturbed not only the CR, but also other condensation-relevant sequences in TDP-43[45,61,111]. However, those studies only investigated the effects on a small number of alternative splicing events and came to discrepant conclusions[45,111], or confirmed that TDP-43 phase separation is necessary for TDP-43 autoregulation[61].

In our work, we took a mutational approach to create variants of TDP-43 with PS-deficiency or solid-like condensate properties. We made sure to target different regions of the LCD and to utilize very different types of mutations to exclude region/mutation-specific effects. Utilizing this panel of TDP-43 variants, where each condensation state is represented by three completely different mutants, we were able to search for PS-dependent rather than sequence-dependent changes in TDP-43 functionality, making it likely that the effects we see come from changes to the condensation state and not the mutations themselves.

Our *in vitro* characterization of the chosen TDP-43 variants revealed distinct behaviors: PS-deficient phenotypes for the 3W, ΔCR, and 12D variants, and more solid-like properties for the 5R, 12A, and G376V variants. These findings align with previously published observations on the "molecular grammar" of TDP-43 phase separation, which has highlighted an important role of aromatic and aliphatic residues[44–46,109] and the alpha-helical conserved region[42,43,60,146]. Our findings furthermore support the notion that charge plays a key role in regulating TDP-43 phase separation, as negatively charged residues in the C-terminus (12D) reduce condensation[65], while positively charged arginine residues reduce condensate dynamicity[45]. Finally, the highly liquid-like phenotype of 3W (tryptophans substituted by glycines) and the more solid-like behavior of G376V are in line with glycine residues having a liquefying effect, consistent with previous observations in the FET proteins[67] and glycine substitutions in TDP-43's alpha-helical conserved region[43].

Importantly, we observed that the PS behavior observed in homotypic PS assays *in vitro* (e.g. number of condensates, dynamics as measured by FRAP) was largely recapitulated when the same variants were stably expressed in cells in the absence of endogenous TDP-43. This is not always the case for other protein systems, as a previous study reported discrepancies between the *in vitro* PS behavior of the chromosomal passenger complex and its behavior in cells or cytomimetic media[147]. In our system, TDP-43 variants retained their PS behavior both in cells and cellular lysates, supporting the notion that TDP-43 robustly forms condensates under physiological conditions. In healthy cells, where the protein is

predominantly nuclear, TDP-43 enriches in multiple nuclear foci. Based on earlier studies, these might represent Cajal bodies, gems and paraspeckles, but also other assemblies of unknown identity[40,112,114]. To what extent the TDP-43 foci of undetermined identity overlap with transcriptional or splicing condensates[58,148] and are functionally linked to the regulation of transcription and splicing remains to be investigated.

TDP-43 phase separation modulates protein interactions, e.g. with RNA-binding proteins

Interestingly we found that protein interactome rearrangements were associated with distinct phase separation properties: solid-like TDP-43 variants exhibit enhanced interactions with numerous other proteins, particularly RNA-binding proteins (RBPs) and nuclear pore complex components. This finding was validated in experiments with recombinant proteins in cellular lysates. Inspired by previous studies that reconstituted condensates in cellular lysates[79,149], we developed this experimental framework that enables the characterization of biophysical properties through microscopy or the identification of other cellular condensate components by mass spectrometry. A key advantage of this system is that it closely mimics cellular conditions, providing a near-physiological multicomponent context, yet it allows for easy tuning of experimental conditions (e.g. variation of protein concentration, salt, pH, PTMs etc.) as is normally only feasible in *in vitro* experiments.

The interactomics dataset shows that several RBPs involved in splicing regulation interact with TDP-43 in a PS-dependent manner. This observation prompted us to investigate whether the expression of TDP-43 variants with distinct PS properties could lead to global changes in splicing regulation. Indeed, we found significant number of splicing events that depended on TDP-43's ability to phase separate or were altered upon expression of solid-like PS variants. Previous studies largely used specific gene reporters to assess the role of PS in splicing and lack consensus on whether disease-linked or phase separation-altering mutations in TDP-43 affect its regulatory splicing activity[45,65,111]. A possible explanation has been proposed by Hallegger et al.[60], by demonstrating that the effects of TDP-43 condensation are context-specific to particular RNA binding regions (unusually long clusters of motifs of characteristic types and density). While we cannot rule out that there are indeed context-dependent RNA binding preferences among our PS variants, we observed a potential indirect effect of TDP-43 PS on splicing regulation, which we linked to PS-dependent protein interactors. Specifically, our analysis of the regions flanking the alternatively spliced exons shows that exons that are differentially spliced by solid-like TDP-43 are more enriched in RNA-binding motifs associated with PS-dependent RBP interactors, pinpointing that the recruitment of other RBPs into TDP-43 condensates may modulate splicing events, that we further linked to abundance changes at the protein level.

Our interactomics dataset corroborates previously reported TDP-43 interactors[20,31–33,65,79–81] and reveals how the interactome is organized into functionally coherent protein modules[115] (Fig. 3). The TDP-43 interactome is highly promiscuous, ranking among the top 3% of proteins with the most extensive interaction network (BioGRID version 5.0.253[82]). Its composition varies significantly based on subcellular localization and physiological or pathological conditions[25,150]. Key proteostasis-related proteins, such as HSP70, HSPB1, and co-chaperones BAG2 and BAG3, regulate TDP-43 condensation and inclusion formation[151–153]. Additionally, other RBPs, such as SRRM2, NUFIP2, and HNRNPC, modulate TDP-43 localization and splicing activity, hence they have been suggested as potential therapeutic targets for modulating TDP-43 function[126]. Our work demonstrates that the TDP-43 interactome is not solely dictated by cellular context but is also governed by its PS properties, given that all variants generated are functionally active and localized in the nucleus. This suggests that phase separation is a key mechanism through which TDP-43

interacts with specific partners to execute distinct functional activities. Our data suggests that more solid-like states of TDP-43, for example as they arise in disease (e.g. through disease-associated mutations or aberrant post-translational modifications), can trap physiological TDP-43 interactors. Importantly, this process is selective, favoring proteins with intrinsically disordered regions containing charged patches.

Phase separation modulates TDP-43 functional activity via direct and indirect mechanisms

Functional changes within the cell depend on the modulation of the cellular proteome at various levels, including alterations in protein abundance, reshaping of the interactome, and post-translational modifications[154]. In this study, we demonstrate that phase separation of TDP-43 is a critical mechanism that governs the reorganization of the cellular proteome. To uncover the global role of TDP-43 PS, we employed a multi-layered approach[140,155] combining interactomics, RNA-Seq and proteomics data. By integrating these results, we describe how the effect of altered TDP-43 condensation, which might occur in diseases such as ALS or FTD, regulates alternative splicing, RNA and protein levels. Notably and not surprisingly, the effects observed in our study are more subtle and nuanced compared to those observed when studying the consequences of TDP-43 depletion[140]. However, our findings nicely complement the observation of a “cascade effect” reported upon TDP-43 knockdown, with nearly half of the differentially expressed proteins functional linked to PS-dependent interactors[140]. In our study, we find that TDP-43 PS can regulate the splicing and abundance of RNAs, and consequently the proteome, both directly via PS-dependent alternative splicing of TDP-43 target genes, and indirectly via protein interactors that are modulated by TDP-43 PS properties (Fig. 7J).

A notable example of the indirect phase separation-mediated regulation is UPF1. While UPF1 has been reported to be neuroprotective in models of TDP-43 dysfunction[156,157], a more direct pathological link has been identified recently as both proteins were found to interact and co-aggregate in disease[16]. UPF1 functions in non-sense-mediated decay (NMD) and a set of UPF-dependent mRNA targets normally undergo degradation via NMD[104–106]. Our data indicates that solid-like TDP-43 variants interact more strongly with UPF1, and UPF1 co-precipitates with solid-like TDP-43 condensates. Sequestration of UPF1 into undynamic TDP-43 condensates may compromise UPF1 function, in line with the observed increase in the levels of UPF1 RNA targets specifically in cell lines expressing solid-like PS variants. We exemplify this with CHID1 and RNF213, two UPF1 targets whose RNA and protein abundance linearly correlate with the abundance of UPF1 interacting with TDP-43.

Taking together, our study offers evidence that the cellular proteome is directly and indirectly regulated via TDP-43 phase separation, the indirect mechanisms involving global TDP-43 interactome rearrangements caused by altered TDP-43 PS properties.

Limitations of the study

Our PS-dependent protein-protein interactions are based on AP-MS and co-precipitation assays, which both were performed on cell lysates. These approaches, unlike proximity-based labeling approaches that preserve cellular architecture, cannot capture transient interactions in living cells or identify which interactions are specific to the native cellular environment, e.g. nucleus vs cytoplasm.

The integration of large-scale datasets is inherently constrained by differences in sensitivity, variability, definition of thresholds and data quality, which can introduce bias into multi-layer analyses. For instance, changes in protein levels may not be detectable for certain alternative splicing events (e.g., SEMA3F). These challenges are further exacerbated by

additional regulatory mechanisms that TDP-43 is involved in that were not accounted for in the analysis. Specifically, we did not consider PS-dependent regulation of alternative polyadenylation at the 3′UTR, which can influence mRNA stability[60] or PS-dependent regulation of mRNA translation[62]. More in-depth analysis of these individual TDP-43 functions would be needed to define the exact molecular mechanisms by which TDP-43 PS regulates the proteome. Although we made every effort to control the levels of rescued TDP-43 across all our experiments, it was not always possible to achieve identical protein abundance. This variability added complexity to distinguishing the specific effects of phase separation from those related to TDP-43 abundance. Additionally, the datasets acquired are high-dimensional, characterized by a large number of variables but relatively few observations (6 analyzed TDP-43 variants). This imbalance can limit statistical power, make it more challenging to identify consistent patterns, and increase the possibility of correlations that may not accurately represent causal relationships. Finally, our analysis was performed in HeLa cells, which do not reflect the gene expression, RNA splicing and degradation patterns or regulatory dynamics of neuronal cells[158–160]. Consequently, our analysis is limited to studying TDP-43 phase separation effects on gene expression in general, without addressing neuron-specific processes. Future studies should validate these findings in neuronal models to better understand the physiological relevance of TDP-43 phase separation in neuronal gene regulation.

Acknowledgements

The authors thank all members of the Dormann lab (in particular Emre Pekbilir), Julian König and Christian Münch for critical discussion. We acknowledge the IMB Flow Cytometry Core Facility, especially Stefanie Möckel and Stephanie Nick, for their assistance with experiment setup and analysis. Cell sorting was performed on the Invitrogen Bigfoot (funded by the Deutsche Forschungsgemeinschaft (DFG), project number 511658729). Support by the Microscopy, Genomics, Bioinformatics, Proteomics and Protein Production Core Facilities of IMB Mainz is gratefully acknowledged. The authors especially thank the Proteomics Core Facility for assistance with MS experiments and the Microscopy Core Facility for expert advice in imaging setup and analyses. Funding by the DFG supported the Revvity Opera Phenix (project ID 316215830), the Leica Stellaris 8 FALCON (project ID 497669232) and the ORBITRAP ASTRAL (project ID 524805621). Funding by ReAlity (Landesinitiative des Landes Rheinland-Pfalz) supported purchase of the ZEISS Axio Observer wide field microscope.

This work was funded by the European Union (EU) through the ERC CoG TDP-ASSEMBLY (Project 1011258709), the DFG within the Heisenberg Programme (project ID 442698351), the SFB1551 “Polymer Concepts in Cellular Function” (project ID 464588647) and the SPP2191 “Molecular Mechanisms of Functional Phase Separation (project ID 419139133). F.U. acknowledges funding by a Marie Curie Individual Fellowship (Grant agreement number 101207537).

Figure legends

Figure 1: In vitro analysis of designed TDP-43 variants establishes a panel of "phase separation-deficient" and a panel of "solid-like" TDP-43 variants.

1A. TDP-43 variants characterized in this study. Schematic diagram of full-length TDP-43 with its N-terminal domain (NTD), RNA-recognition domains (RRM1, RRM2) and disordered LCD (aa. 261-414) is shown. Variants designed to be phase separation-deficient are highlighted in turquoise, while variants designed to be more phase separation-prone than WT are shown in red. Created in part with BioRender.com.

1B. Scheme illustrating workflow of the *in vitro* condensate microscopy assay. TDP-43-MBP-His_6 is incubated with TEV protease to proteolytically remove the MBP solubility tag; reactions are incubated in phase separation buffer for 45 minutes and resulting condensates are imaged using brightfield microscopy. Created in part with BioRender.com.

1C-D. Representative brightfield microscopy images of condensates formed by the variants predicted to be PS-deficient (C) or PS-prone (D) in comparison to TDP-43 WT at 5, 10 and 20 µM. Imaging was done 45 minutes post TEV addition. Negative controls without TEV protease addition (-TEV) contain TDP-43-MBP-His_6 at 20 µM. Scale bar: 5 µm.

1E-F. Representative confocal fluorescence microscopy images for the FRAP experiments. Condensates from PS-deficient TDP-43 variants and WT as control were formed at 40 µM protein concentration (E), while for the amorphous mutants and corresponding WT control (F), a protein concentration of 15 µM was used. Bleached regions of interest (ROI) are indicated in yellow. Scale bar:2 µm.

1G-H. Quantification of FRAP analysis illustrating a faster and higher fluorescence recovery for the PS-deficient variants compared to WT (G) and a lower fluorescence recovery of the amorphous variants compared to WT (H). An average of 14 condensates were analyzed per condition. Mean values are shown, with standard deviation indicated in shaded regions. Significance was determined by quantifying areas under the curve for individual condensates followed by one-way ANOVA with a Dunnett's multiple comparisons test to WT. P values are as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

1I. Workflow illustrating the condensate reversibility assay. Condensate formation was induced using TEV protease cleavage, followed by a 15 or 60 min incubation and centrifugation step for undiluted (1:1) samples, or dilution (1:5 or 1:10) in phase separation buffer and an additional incubation for 15 minutes. Samples were centrifuged to separate supernatant (soluble TDP-43) and pellet (condensed TDP-43) fractions. Created in part with BioRender.com.

1J. Quantification of condensate reversibility. Dilution was performed at 60 minutes post condensate induction. % soluble values were determined by quantifying the signal coming from the soluble (S) fraction as a percentage of total signal (P+S fractions) for a given mutant. Error bars indicate standard deviation across three independent experiments. Significance was determined by quantifying the signal for 1:10, followed by a one-way ANOVA with a Dunnett's multiple comparisons test to WT. P values are as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

1K-L. Representative Western blot images for TDP-43 WT, or the less dynamic 5R, 12A and G376V variants tested for condensate reversibility after 60 minutes of incubation time. Pellet (P) and supernatant (S) fractions were loaded for undiluted (1:1) samples, while for 1:5 and 1:10 samples only supernatant (S) fractions are shown. Blots were incubated with a TDP-43-specific antibody.

Figure 2. Rationally designed LCD mutations tune nuclear TDP-43 phase separation behavior under physiological conditions.

2A. Scheme illustrating the experimental setup in all cellular experiments (silencing of endogenous TDP-43 knockdown by siRNA for 48h, followed by doxycycline-induced expression of myc-tagged TDP-43 variants for 24h). Created in part with BioRender.com.

2B. TDP-43 Western blot showing myc-TDP-43 expression levels in HeLa T-REx Flp-in cells expressing the indicated myc-TDP-43 variants in the absence of endogenous TDP-43 and after dox-induction. Parental HeLa Flp-in T-REx host cells were used to show endogenous TDP-43 levels as well as its depletion after knockdown (KD TDP-43). GAPDH blot is shown as loading control.

2C. Representative confocal images of HeLa T-REx Flp-in cells expressing the indicated myc-TDP-43 variants stained with an anti-Myc antibody (green). Nuclei were stained with DAPI (blue). Scale bar: 10 µm.

2D. Quantification of nuclear TDP-43-rich foci from confocal high throughput images shown in Fig. 2C. Violin plots show the distributions of foci counts per cell. Median is indicated by a dashed line, 1st and 4th quartiles are shown in thick dotted lines.

2E. Representative FRAP images in HeLa T-REx Flp-in cells expressing the indicated EGFP-tagged TDP-43 variants. Bleached nuclear regions are indicated as yellow dashed squares. Top right quadrants show a zoomed-in image of the bleached region. Scale bar: 5 µm.

2F. Quantification of fluorescence recovery from FRAP experiments shown in Fig. 2E. Regions of interest in nuclei were bleached and fluorescence recovery in these regions was measured for ~1 minute. Between 12-14 cells were analyzed per cell line across three biological replicates, means of all fluorescence recovery curves are plotted, with standard deviation indicated in the shaded region. Significance was determined by quantifying areas under the curve for individual condensates followed by one-way ANOVA with a Dunnett's multiple comparisons test to WT. P values are as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Figure 3: The phase separation behavior of TDP-43 governs TDP-43 protein interactions.

3A. Workflow for the identification of TDP-43 interactors by AP-MS. Phase separation-deficient variants (3W, ΔCR, 12D), solid-like variants (5R, G376V, 12A) or TDP-43 WT were expressed with a myc tag to enable copurification of protein interactors under mild lysis conditions (N=3). Copurified proteins were subjected to MS identification and quantification. To filter specific interactors, the purification was also performed for a mock condition without the expression of the bait. All purifications were done in the absence of endogenous TDP-43 (after siRNA-mediated depletion). Created in part with BioRender.com.

3B. GO enrichment analysis of TDP-43 interactors. GO terms with at least 10 genes associated, a $\log_2$FC enrichment > 1 and with a *p*value < $1e^{-5}$ are displayed. To reduce redundancy, only terms based on a similarity are shown. Circle corresponds to the number of genes per term, and color indicates the GO category (red: Biological Process (BP); green: Cellular Component (CC); blue: Molecular Function (MF)).

3C. Overlap of TDP-43 interactors identified in this dataset with interactors reported from other TDP-43 interactomics experiments. HCIs: High confidence interactors, defined as TDP-43 interactors identified in at least three independent interactomics datasets.

3D. Unsupervised hierarchical clustering of TDP-43 interactor levels identified for solid-like and PS-deficient variants of TDP-43. Values reported in the heatmap show the average (three independent replicates) intensity of interactors as log_2 fold change compared to TDP-43 WT. Proteins with an enhanced interaction intensity are shown in red, while those with weaker interactions are shown in blue.

3E. Protein-protein interaction network of TDP-43. Proteins identified as TDP-43 interactors are clustered based on the abundance profile across all TDP-43 purifications (plot profile in the left). Proteins are organized in modules according to reported protein-protein interactions and their localization. The edges (connection between two nodes) indicate interactions identified in at least two independent experiments (BioGRID v. 5.0.253)[82]. Novel interactors, low and high confidence interactors are shown as nodes with different colors (grey, yellow and red respectively).

3F. Protein-protein interaction network for the spliceosome module. The width of the edge between the nodes is proportional to the number of independent studies describing the interactions. Novel interactors, low and high-confidence interactors are shown with distinct colors (grey, yellow and red respectively).

3G-H. Enrichment analysis for TDP-43 interactors (PS-dependent/PS-independent) and for the control group. Enrichment was evaluated for proteins containing an extended RGG motif G(0-3)X(0,1)RG(1,2)X(0,5)G(0,2)X(0,1)RG(1,2)[85] (G), and for proteins annotated as associated to phase separation from the CD-CODE database (only proteins reported to participate in biomolecular condensates in at least three independent experiments were considered) (H). Reported values correspond to enrichment calculated using a hypergeometric test.

Figure 4: Phase separation-dependent interactors with distinct sequence features co-precipitate with solid-like TDP-43 variants in cellular lysates

4A. Experimental scheme for the co-condensation experiment. Purified MBP-tagged TDP-43 was spiked into standardized HeLa cell lysate (2 mg/ml) and formation of condensates was triggered by TEV protease addition. Formed condensates were examined by confocal fluorescence microscopy or pelleted by centrifugation. The supernatant and pellet fractions were examined by Western blotting (WB), and soluble proteins in the supernatant were quantitatively analyzed by mass spectrometry. Created in part with BioRender.com.

4B. Representative confocal microscopy images of TDP-43 condensates formed at 10 µM protein concentration in the presence of HeLa cell lysate (2 mg/ml) upon TEV cleavage (45 minutes). MBP-TDP-43-His_6 proteins were labeled with Alexa Fluor 488. Maximum intensity projections are shown. Scale bar: 10 µm.

4C. Quantification of the number of condensates per well for the images shown in B. 12 fields of view were quantified per condition. Mean values of 3 experimental replicates are shown. Error bars indicate standard deviation. Significance was determined using one-way ANOVA with a Dunnett's multiple comparisons test to WT.

4D. Abundance of soluble TDP-43 measured in the supernatant fraction (SN) by mass spectrometry. Abundance values were normalized to the corresponding TDP-43 levels in the

absence of TEV (first column). The bar plot represents the mean values, error bars indicate the standard deviation, and individual points correspond to biological replicates (N = 4).

4E. Abundance levels of TDP-43 PS-dependent interactors (N = 43). Only interactors whose abundance changed significantly under at least one condition were considered (*p*value BH adjusted < 0.05). The level of endogenous soluble proteins in the supernatant following centrifugation is reported as z-score, calculated from the mean protein value for each condition (N =4).

4F. Experimental workflow for the dose-response co-precipitation profiling upon titration of TDP-43 WT and 5R. Increasing amounts of purified protein (0,1,2,5,10 µM) were added to total HeLa cell lysate (2 mg/ml) and proteins levels were measured by mass spectrometry before and after TEV protease addition and centrifugation. 403 proteins displaying a dose-dependent depletion from the supernatant following increasing concentration of spiked-in TDP-43 5R variant but not WT were classified as proteins co-precipitating with solid-like mutants (example profile shown below). See material and methods and Fig. S8C for details on the data analysis procedure. Created in part with BioRender.com.

4G. Venn diagram showing the overlap between the proteins identified as PS-dependent interactors by AP-MS (Fig. 3) and proteins co-precipitating specifically with the 5R variant. Dose-response depletion profiling confirmed co-depletion of 22 proteins with the 5R variant, representing ~11% of the 195 proteins previously identified in the AP-MS dataset

4H. Enrichment analysis for TDP-43 modules identified in the AP-MS dataset and proteins co-precipitating with 5R. Reported values correspond to enrichment calculated using a hypergeometric test.

4I. Distribution of the functional association of proteins with TDP-43 was analyzed across three groups: control (403 randomly sampled proteins from the identified set), coprecipitating with 5R (403 proteins), and PS-dependent interactors (195 proteins). For each group, the shortest path to TDP-43 within the STRING network was calculated, and a path score was determined by multiplying the interaction confidence values along the path. Proteins were then ranked based on their respective path scores.

4J. Enrichment analysis for proteins annotated to be associated to phase separation from the CD-CODE database for phase separation dependent interactors, proteins co-precipitating with 5R and random control proteins. Only proteins reported to partition into biomolecular condensates in at least three independent studies were considered. Reported values correspond to enrichment calculated using a hypergeometric test.

4K. Correlation between significant grammar sequence features (N =74, BH adjusted *p*value < 0.01, Welch test) of TDP-43 PS-dependent interactors and proteins co-precipitating with 5R. Sequence parameter values for each protein were calculated using CIDER[94] and expressed as z-scores relative to the mean and standard deviation of the human proteome. Each point represents a grammar sequence feature deviation from the proteome average, with positive and negative values indicating enrichment or depletion, respectively.

4L. Conserved grammar sequence features for proteins identified as PS-dependent interactors and co-precipitating with 5R. A scale from blue (negative values) to red (positive values) displays respectively the depletion and enrichment of properties compared to the mean properties of the human proteome. The significance of the deviation from the human proteome is reported (BH adjusted *p*value, Welch test , ** < 0.01 and * < 0.1). Among the 128 grammar features calculated, this heatmap only reports features that show significant differences compared to the human proteome in at least one condition and that are characterized by a coherent enrichment between the two conditions.

Figure 5: Phase separation behavior of TDP-43 governs interactions with UPF1, leading to sequestration of UPF1 in solid-like TDP-43 condensates.

5A. UPF1 abundance in the AP-MS experiment across the different TDP-43 variants. The level of UPF1 was normalized for the abundance of TDP-43 purified in each condition and the values are normalized to the average of the WT condition. Single points represent biological replicates, and the boxplot reports the distribution of the values with the mean (central line), the interquartile range (1st and 3rd quartile) and whiskers (1.5XIQR, interquartile range).

5B. UPF1 shows a dose-dependent depletion from the supernatant upon titration of 5R, but not WT TDP-43, into Hela cell lysate. Single points represent biological replicates, and the dashed line indicates the fitted dose-response curve.

5C. The abundance of UPF1 was assessed by Western Blot in both the pellet and supernatant fractions. The pellet fraction was loaded at five times the volume of the supernatant fraction. Lower panel: quantification of UPF1 in the pellet fraction was performed using four independent replicates, each derived from distinct batches of cell lysate. The values were normalized to set the abundance of UPF1 in TDP-43 WT sample to 1. The mean values and standard error are presented.

5D. Number of interactions between 1 UPF1/albumin (CTRL) chain and 100 chains of TDP-43 variants, as calculated from MD simulations (sum across 2000 frames, 1 µs). Interaction counts are normalized based on all possible contact pairs. Results are derived from three independent simulation runs.

5E. Representative visualization of *in silico* condensates generated using simulations with the coarse-grained HPS model. In red one chain of UPF1 is shown, while 100 chains of TDP-43 are shown in grey. The simulation was performed by adding one chain of UPF1 externally to a condensate generated from 100 molecules of TDP-43 (Start). The other snapshots show the last frame of the simulations (1 µs) for UPF1 with condensates formed from TDP-43 WT, PS-deficient variants (3W, ΔCR and 12D) and solid-like variants (5R, 12A and G376V).

Figure 6: TDP-43 phase separation properties modulate alternative splicing and RNA levels

6A. RNA sequencing experiment setup to study alternative splicing and gene expression changes in cell lines expressing the different TDP-43 PS variants. Created in part with BioRender.com.

6B. Identification of RNA motifs. The flanking regions (500nt) in the proximity of the alternative exon were analyzed to infer motif enrichment. Bottom, motif enrichment analysis for 4’096 60mers in PS-deficient variants (3W, ΔCR and 12D) and solid-like variants (5R, 12D and G376V). The significance was calculated with Fisher exact test.

6C. Distribution of *p*values (Fisher’s exact test) for RNA motif enrichment (CISBP-RNA for homo sapiens DNA encoded) within the flanking regions of alternatively spliced exons. Only RNA motifs significantly enriched (*p*value < 0.01) are plotted as single points. The boxplot reports the distribution of the values with the median (central line), the interquartile range (1st and 3rd quartile) and whiskers (1.5XIQR, interquartile range).

6D. Significance enrichment (Fisher exact test) for RNA motif of proteins identified as PS-dependent TDP-43 interactors (HNRNPC, HNRNPH2, HNRNPK, HNRNPL and SYNCRIP). The enrichment for each TDP-43 variant is shown as single point. The boxplot reports the distribution of the values with the average (central line), the interquartile range (1st and 3rd quartile) and whiskers (1.5XIQR, interquartile range). The motif enrichment (according to CISBP-RNA for homo sapiens DNA encoded) is shown in the right box. The RNA motifs enrichment for each single TDP-43 variant is shown in Fig. S11F.

6E. TDP-43-dependent alternative exons. A total of 425 events were identified in the WT condition (185 with rescue of exon inclusion and 240 with rescue of exon skipping in the rescue condition). Only events that show statistically significant changes (FDR< 0.05) between knockdown (KD) and NTC ('non targeting control') as well as between WT rescue and KD, and that exhibit opposite inclusion directions between the two conditions, were considered. Two well-characterized alternative splice targets of TDP-43 (*POLDIP3* and *ATG4B*) are shown as examples.

6F. Annotation of alternative exons identified as TDP-43-dependent. CDS = coding sequence.

6G. Unsupervised hierarchical clustering of 240 exon-skipping events following rescue, showing differences in exon inclusion levels for TDP-43 variants compared to WT TDP-43. Values reported in the heatmap show the differences in inclusion levels, with positive values (increase in inclusion) in red and negative values (increase in skipping) in blue. In the annotation bar on the left, the events identified as PS-dependent splicing events are shown (events where the standard deviation within the PS groups is lower than the total standard deviation). The unsupervised hierarchical cluster for 185 exon-inclusion events is displayed in Fig. S12E.

6H. Examples of TDP-43- and PS-dependent alternative exons. Inclusion levels of alternative exons from *SEMA3F*, *BRD8* are shown under NTC, KD and rescue conditions. The levels for each replicate are shown as single point, while the boxplot shows the average (central line) and the interquartile range (1st and 3rd quartile).

6I. Volcano plot showing RNA level changes following TDP-43 knockdown (left) and subsequent rescue with TDP-43 WT (right). The x-axis represents the $\log_2$ fold change ($\log_2$FC), while the y-axis indicates the significance of changes, with *p*values adjusted for multiple comparisons using the Benjamini-Hochberg (BH) correction. Exons included upon TDP-43 knockdown and exons skipped upon rescue with TDP-43 WT are highlighted in orange.

6J. Unsupervised hierarchical clustering of 298 genes following TDP-43 rescue. The upper panel displays 120 upregulated genes, while the lower panel shows 177 downregulated genes, as determined by differential analysis between rescue and knockdown conditions. The heatmap represents z-score differences in gene expression, with red indicating positive deviations and blue indicating negative deviations from the mean expression levels of each gene. The annotation bar on the left highlights identified alternative exons, with TDP-43-dependent exons shown in light green and PS-dependent exons in dark green.

6K. Gene set enrichment analysis (GSEA) comparing RNA levels for solid-like TDP-43 variants (5R, 12A and G376V) and PS-deficient variants (3W, ΔCR and 12D). Gene set: UPF1 target genes in HEK293 and HeLa cells[104–106]. The y-axis displays the running enrichment score (ES) (upper panel), the position of UPF1 target genes (middle panel) and corresponding $\log_2$ fold changes (lower panel); the x axis shows the ranked position of each gene in the set. NES (normalized enriched score) and *p*value adjusted are shown in the plot.

Figure 7: TDP-43 phase separation directly and indirectly regulates the cellular proteome

7A. Workflow for the identification of PS-dependent events across multiple levels of gene regulation, including alternative splicing, RNA abundance, protein levels, and protein–protein interactions. PS-dependent events were filtered independently in each dataset and subsequently integrated through network analysis to identify regulatory links. Created in part with BioRender.com.

7B. Volcano plot showing differences in protein abundance between TDP-43 knockdown (KD) and rescue with TDP-43 WT. The x-axis represents differential protein abundance, while the y-axis shows statistical significance (adjusted *p*value, Benjamini–Hochberg correction). Proteins associated with TDP-43-regulated alternative splicing are highlighted in dark red, while genes identified in the RNA-seq dataset as upregulated or downregulated upon rescue are highlighted in red and cyan. The lower panel shows the density distribution of protein abundance changes for genes that are differently regulated at the RNA level.

7C. Sankey diagram displaying events depending on TDP-43 WT levels. The first level represents alternative splicing changes (exon inclusion and skipping), the second and third level represent RNA and protein changes, respectively, classified as upregulated (red) and downregulated (blue). The color of the links indicates whether an event is associated to upregulated or down regulation in the differential analysis between rescue and KD.

7D. Sankey diagram illustrating regulatory events influenced by the PS properties of TDP-43. The diagram highlights PS-dependent alternative splicing events associated with PS-dependent proteins (direct regulated by PS). Proteins directly regulated by PS are depicted in red (upregulated) and blue (downregulated), with their corresponding links emphasized in pale red and pale blue.

7E-F. Examples demonstrating how the PS properties of TDP-43 directly impact alternative splicing, resulting in coordinated changes in protein abundance (MADD and NCAPG2). In the case of MADD, solid-like TDP-43 variants exhibit reduced exon inclusion, which correlates with increased protein abundance. Conversely, for NCAPG2, reduced exon inclusion is associated with decreased protein abundance.

7G. Summary of the PS-dependent effects on protein abundance. Proteins directly linked to alternative splicing events are shown in red (direct mechanism). Proteins associated with PS-dependent TDP-43 interactors are shown in green (indirect mechanism). Proteins for which no mechanistic link could be established are shown in cyan.

7H-I. Examples demonstrating how the PS properties of TDP-43 indirectly impact protein abundance (CHID and RNF213). TDP-43 solid-like variants trap mRNA decay factor UPF1, possibly reducing its availability for the RNA degradation machinery and leading to increased protein levels of UPF1 target genes (CHID and RNF213).In the proposed examples the abundance of CHID and RNF213 correlates with the abundance of UPF1 that binds to TDP-43 with distinct phase separation properties.

7J. Distinct mechanisms by which TDP-43 phase separation regulates the proteome. Variations in phase separation behavior influence TDP-43 organization, leading to two potential effects: (1) a direct effect, where TDP-43 phase separation modulates alternative splicing of TDP-43 target genes and corresponding protein levels, and (2) an indirect effect, where TDP-43 phase separation changes interactions with RNA regulator factors (e.g. UPF

show as example), which in turn regulate RNA and protein abundance. Created in part with BioRender.com.

Supplementary Figure legends

Fig. S1

S1A. Representative confocal fluorescence microscopy images of condensates formed by TDP-43 variants in comparison to TDP-43 WT at 5, 10 and 20 µM. Imaging was performed 45 minutes after TEV addition. Proteins were fluorescently labeled with Alexa Fluor 488 C5 malemide. Scale bar: 5 µm.

S1B. Quantification of condensate reversibility. Dilution was performed at 15 minutes post condensate induction. % soluble values were determined by quantifying the signal coming from the soluble (S) fraction as a percentage of total signal (P+S fractions) for a given mutant. Error bars indicate standard deviation across three independent experiments. Significance was determined by quantifying the signal for 1:10, followed by a one-way ANOVA with a Dunnett's multiple comparisons test to WT. P values are as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

S1C-D. Representative Western blot images for TDP-43 WT, or the less dynamic 5R, 12A and G376V variants tested for condensate reversibility after 15 minutes of incubation time. Pellet (P) and supernatant (S) fractions were loaded for undiluted (1:1) samples, while for 1:5 and 1:10 samples only supernatant (S) fractions are shown. Blots were incubated with a TDP-43-specific antibody.

S1E. Saturation concentration (C_{sat}) values obtained for condensates formed by TDP-43 WT and solid-like TDP-43 variants. SYPRO Ruby signal from S and P fractions was quantified to obtain a total signal corresponding to 1 µM protein. Afterwards, protein concentration corresponding to the amount of signal in S fraction was quantified and assigned as C_{sat}. Three independent experimental replicates were used for the quantification and each replicate corresponds to a C_{sat} value quantified from three phase separation reactions processed on the same day. Error bars indicate standard deviation. Significance was determined by a one-way ANOVA with a Dunnet's multiple comparisons test to WT. P values are as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

S1F. Representative image showing a SYPRO Ruby-stained gel with TDP-43-containing supernatant (S) and pellet (P) fractions used to estimate C_{sat} values. Three reactions were performed for each C_{sat} estimation replicate and loaded side by side (r1-3).

Fig. S2

S2A. Scheme illustrating the heat shock treatment and recovery procedure (silencing of endogenous TDP-43 knockdown by siRNA for 48h, doxycycline-induced expression of myc-tagged TDP-43 variants for 24h, followed by a 1-hour heat stress treatment at 43°C and a 3-hour recovery period at 37°C). Created in part with BioRender.com.

S2B. Representative confocal images of HeLa T-REx Flp-in cells expressing the indicated myc-TDP-43 variants subjected to heat stress as illustrated in Fig. S2A. myc-TDP-43 variants were stained with an anti-Myc antibody (green). Nuclei were stained with DAPI (blue). Scale bar: 10 µm.

S2C. Quantification of nuclear stress-induced TDP-43 rich foci formed during heat shock recovery from confocal high throughput images shown in Fig. S2B. Violin plots show the distributions of foci counts per cell. Median is indicated by a dashed line, 1st and 4th quartiles are shown in thick dotted lines.

S2D. Representative FRAP images in HeLa T-REx Flp-In cells expressing the indicated EGFP-tagged TDP-43 variants during heat stress recovery. Point bleaches were performed on individual TDP-43-rich condensates, and are indicated with yellow arrows. Top right quadrants show a zoomed-in image of the bleached condensate. Scale bar is 5 µm.

S2E. Quantification of FRAP in HeLa T-REx Flp-In cells during heat stress recovery. Stress-induced TDP-43-rich condensates were bleached using a point bleach and their fluorescence recovery was monitored for ~1 minute. Between 8-13 cells were analyzed per cell line across three biological replicates, means of all fluorescence recovery curves are plotted, with standard deviation indicated in the shaded region. Significance was determined by quantifying areas under the curve for individual condensates followed by one-way ANOVA with a Dunnett's multiple comparisons test to WT. P values are as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Fig. S3

S3A-B) Immunoblot showing expression levels of the bait (TDP-43) across the analyzed variants (A) and its levels across purifications (B). INP: Input (1% of the total amount), FT: Flow-through (1% of the total amount), IP: Immunopurification (10% of the total amount).

S3C-D) Number of identifications and protein abundance levels across analyzed conditions.

S3E) Abundance of the bait (TDP-43) purified in the different conditions.

S3F) Volcano plot showing the proteins enriched in the different purifications for the different TDP-43 variants. High confidence interactors (HCIs) are shown in black, and abundance of the bait (TDP-43) is highlighted.

3G-I) Sankey plot showing the GO enrichment for each purification (CC: Cellular Component (G); MF: Molecular Function (H); BP: Biological Process (I)). For the GO analysis, only terms with at least 4 (CC, MF) and 9 (BP) genes associated and identified as parent term based on similarity are shown. For each gene only the term with the lowest *p*value is associated.

Fig. S4

S4A) Correlation matrix of TDP-43 purifications. Pearson correlation coefficients were computed for all pairs of purifications. Euclidean distances between conditions were calculated from the interactor intensities normalized for TDP-43 levels to guide hierarchical clustering with the Ward.D2 method.

S4B) Principal component analysis of TDP-43 purifications. Eigenvalues of the two principal components (Dim1 and Dim2) visualize the proportion of variance explained. TDP-43 variants are highlighted in different colors, and every dot represents a replicate condition.

S4C) Hierarchical clustering of TDP-43 interactors. TDP-43 interactors abundance was z-score transformed and the Pearson correlation coefficients were calculated across all conditions. The resulting correlation matrix was used to guide hierarchical clustering, applying complete linkage to define clusters of interactors with similar enrichment patterns.

SD-F) Characterization of the network generated and comparison to a distribution of control networks generated by sampling the interactions of 226 random proteins from BioGRID 100 times. The plots report the number of interactions (D), the average node degree (E) and the

transitivity (F), the likelihood that nodes connected to a common node are also connected to each other.

S4G) Organization of the network generated; proteins are ranked according to their eigenvector centrality. Proteins associated to the spliceosome complex (according to QuickGO annotation) are highlighted in red.

S4H) z-score profile of interactors according to their module assignment. Profile of every protein is shown in grey, while the median profile is shown in red.

Fig. S5

The protein-protein interaction network of TDP-43 is organized into modules based on reported interactions and protein localization. Proteins within each module are displayed using a circular layout. Edges between nodes represent interactions identified in at least two independent experiments (BioGRID v. 5.0.253). Nodes are color-coded to distinguish novel interactors, low- and high-confidence interactors, represented by grey, yellow, and red respectively.

Fig. S6

S6A) Functional association of proteins with TDP-43. 18’841 proteins were ranked based on their functional association with TDP-43. For each protein present in the STRING network (v.12), the shortest interaction path to TDP-43 was calculated, and a path score was derived as the product of the interaction confidence values along the path. Lower panel shows the distribution of the functional score for all the identified TDP-43 interactors (Fig. 3E).

S6B) Prediction of interactions between TDP-43 and its interacting proteins compared to the control group. The number of molecular contacts was quantified for 20 control pairs (TDP-43 paired with proteins identified in the AP-MS dataset but not classified as interactors) and 20 protein pairs consisting of TDP-43 and proteins identified as TDP-43 interactors in the AP-MS dataset. Molecular dynamics simulations were conducted for each protein pair using the CALVADOS3 coarse-grained force field. Interactions were evaluated by calculating the probability of a bound state over the one of unbound state. Significance is calculated by by Wilcoxon rank-sum test, N = 20.

S6C) Representative examples of interactions are shown between a control protein (DAB2) and a TDP-43 interactor (NUP188) with TDP-43. The lower panel displays the minimal distance between the two protein pairs for each frame. The upper panel includes a tile plot, showing the fraction of frames in which interactions are identified. On top is a snapshoot of the simulation between DAB2 (red) and TDP-43 (blue) and between NUP188 (red) and TDP-43 (blue).

S6D) Prediction of interactions between TDP-43 and its interacting proteins compared to the control group. The number of molecular contacts was quantified for 20 control pairs (TDP-43 paired with proteins identified in the AP-MS dataset but not classified as interactors) and 20 protein pairs consisting of TDP-43 and proteins identified as TDP-43 interactors in the AP-MS dataset. Molecular dynamics simulations were conducted for each protein pair using the CALVADOS3 coarse-grained force field. Interactions were evaluated by calculating the number of contacts normalized to the total number of possible contact pairs. Significance is calculated by unpaired two-side t-test, N=20.

S6E) Residue-level interaction enrichment along the TDP-43 sequence. The log_2 enrichment of contact frequencies calculated for each residue compared to a null distribution generated by shuffling residue-wise contact counts 1'000 times. The pLLDT confidence score obtained from AlphaFold2 predictions is showed in the lower panel. Annotated domains include the N-terminal domain (NTD, 1-81), the NLS region (82-103), RNA recognition motifs (RRM1 and RRM2, 84-262) and the C-terminal LCD (LCD, 263-414).

S6F-G) Isoelectric point (pI) (D) and median GRAVY index per protein (E) (calculated according to Kyte Doolittle) of the PS-dependent and -independent TDP-43 interactors, and for proteins in the control group. A statistical analysis was performed using an unpaired Wilcoxon test to compare PS-dependent interactors and the control group (N = 195).

S6H) Distribution of median of AlphaFold2 pLLDT confidence score for TDP-43 interactors and for the control group. The median per-protein pLLDT scores were calculated for proteins identified as PS-dependent TDP-43 interactors (N =195), PS-independent interactors (N=31), and for proteins in the control group (N=195). A statistical analysis was performed using an unpaired Wilcoxon test to compare PS-dependent interactors and the control group (N = 195).

S6I) Distribution of the fraction of the predicted secondary structure calculated for proteins identified as TDP-43 interactors: PS-dependent interactors (N =195), and PS-independent interactors (N=31) or the control group (N=195).

S6J,K,L) Enrichment analysis for PS-dependent and -independent TDP-43 interactors, and for proteins in the control group. Enrichment was evaluated for proteins containing at least one 'RG' motifs (J), 'RGG' motifs (K) and proteins annotated as associated to phase separation from PhaSepDB database (v. 2.1) (L). Reported values correspond to the enrichment calculated using a hypergeometric test.

Fig. S7

S7A) Quantification of the TDP-43 condensates area per well for the images shown in Fig. 4B. 12 fields of view were quantified per condition. Mean values of 3 experimental replicates are shown. Error bars indicate standard deviation. Significance was determined using one-way ANOVA with a Dunnett's multiple comparisons test to WT.

S7B) Immunoblot of TDP-43 on supernatant (soluble) and pellet (insoluble) fractions after centrifugation for the indicated TDP-43 variants. The band above 95 kDa represents TDP-43-MBP-His_6, while the band around 43 kDa corresponds to TDP-43 without tag.

S7C) Precursors identified under the acquired conditions and their detection across all samples. The color scale indicates precursor completeness, expressed as the percentage of conditions in which each precursor was identified; red bars denote complete detection across the dataset.

S7D) Abundance of the precursors before normalization and filtering.

S7E) Principal component analysis (PCA) showing sample stratification, with PS-deficient and solid-like conditions separating along the second principal component.

S7F) Overlap between proteins co-condensing (log_2FC <-0.3 and BH adjusted *p*values < 0.1) upon the spike-in of the different TDP-43 variants. The comparison was performed using TDP-43 WT + TEV as reference. Proteins co-condensing in at least 2 conditions with solid-like TDP-43 condensates are highlighted in red.

S7G) Volcano plot showing the differential changes of protein levels upon the spike-in of the indicated TDP-43 variants into the HeLa cell lysate. Proteins showing a significant depletion (log_2FC <-0.3 and BH adjusted *p*values < 0.1) and identified previously as PS-dependent interactors are highlighted in red.

Fig. S8

S8A) Precursors identified under the acquired conditions and their detection across all samples. The color scale indicates precursor completeness, expressed as the percentage of conditions in which each precursor was identified; red bars denote complete detection across the dataset.

S8B) Abundance of the precursors before normalization and filtering.

S8C) Data analysis strategy for dose-response depletion profiling of protein levels upon titrating in TDP-43 WT or 5R. No imputation was applied, and only the 5'969 proteins identified with at least two precursors per condition were included in the analysis. Protein abundance was calculated as the median intensity of the top five precursors, following a filtering procedure designed to remove low-quality measurements, including precursors with poor identification confidence (PEP), low intensity (>0), and low reproducibility (excluding 50% of precursors per protein with the highest median standard error per condition). Dose-response curve fitting was applied to all quantified proteins, and proteins were retained if they exhibited a 5R-specific dose-response or a lower area under the curve (AUC) for the 5R condition compared to WT, yielding a final set of 403 proteins. Filtering criteria for curve fitting included the Hill coefficient (> 0.7), ANOVA *p*value (< 0.1), and correlation (> 0.7).

S8D-J) Abundance profiles for the indicated proteins upon titration of WT and 5R. Single points represent biological replicates, and the dashed line indicates the fitted dose-response curve (according to the criteria described above). The ANOVA *p*value and EC50 calculated for the 5R condition are reported for each profile. Representative profiles are shown for SMARCA5, NUP160, DDX3X, FUBP3, HNRNPA3, HNRNPR, and HNRNPUL2.

S8K) Enrichment analysis for proteins annotated to be associated with phase separation from PhaSepDB database (v. 2.1). Reported values correspond to enrichment calculated using a hypergeometric test.

Fig. S9

Sequence features ("molecular grammar" for proteins identified as co-precipitating with 5R (left) or PS-dependent interactors (right). Only sequence features identified as significant (BH adjusted *p*value, Welch test < 0.1) are shown. A scale from blue (negative values) to red (positive values) displays the depletion and enrichment, respectively, of properties compared to the mean properties of the human proteome. The significance of the deviation from the human proteome is reported (BH adjusted *p*value, Welch test ,*** < 0.001, ** < 0.01 and * < 0.1).

Fig. S10

S10A-B) Quantification of RNA (A) and protein (B) levels of UPF1 across the described conditions. TDP-43 was depleted using siRNA (KD), followed by re-expression of myc-tagged WT, PS-deficient or solid-like TDP-43 variants (Rescue). Single points represent biological

replicates and the boxplot reports the distribution of the values with the mean (central line), the interquartile range (1st and 3rd quartile) and whiskers (1.5XIQR, interquartile range).

S10C) Interactions per frame between UPF1 or albumin as control (CTRL) protein and 100 chains of TDP-43 variants, as calculated from molecular dynamics simulations (2'000 frames, 1 µs).

S10D) Density profiles from molecular dynamics simulations showing interactions between TDP-43 variants (100 chains, gray) with UPF1 (one chain, cyan) and albumin (control, one chain, red). Profiles average three replica simulations and include trajectories after 100 ns (total time of simulation 1 µs).

Fig. S11

S11A) Number of total reads for each condition analyzed.

S11B) Number of reads for TDP-43 ($\log_2$ scale). The boxplot reports the distribution of the values with the mean (central line), the interquartile range (1st and 3rd quartile) and whiskers (1.5XIQR, interquartile range).

S11C) Immunoblots of TDP-43 levels for all examined cell clones (3 clones / TDP-43 variant, each in non-tageting control (NTC), knockdown (KD) and Rescue condition). GAPDH is shown as loading control.

S11D) Number of significative (FDR< 0.05) alternative exons identified. Events with exon inclusion (PSI > 0.1) and with exon skipping (PSI < -0.1) are shown respectively in orange and cyan.

S11E) Heatmap showing the motif enrichment of TDP-43 PS-dependent motifs identified and grouped by Hallegger et al. ("YG-containing", "YA-containing", "AA-containing")[60]. Significantly enriched RNA motifs are shown in scale of red.

S11F) Significance enrichment (Fisher exact test) for RNA motif enrichment of proteins identified as PS-dependent TDP-43 interactors (HNRNPC, HNRNPH2, HNRNPK, HNRNPL, and SYNCRIP). The RNA motifs enrichment for each single TDP-43 variant is show as barplot. The motif enrichment (according to CISBP-RNA for homo sapiens DNA encoded) is shown in the right box.

S11G-H) Upset plot showing the overlap of events significantly included (PSI>0.1) (G) or skipped (PSI < 0.1) (H).

Fig. S12

S12A) Motif analysis enrichment for 4'096 6-mers in TDP-43- dependent alternatively spliced exons. The significance was calculated with Fisher exact test.

S12B-C) Sashimi plots showing alternatively spliced exons in *POLDIP3* and *ATG4B*.

S12D) Summary of the data analysis approach for identifying TDP-43- and PS-dependent alternative exons. TDP-43-dependent alternative exons (425 events) were identified based on significantly rescued ΔPSI shifts in the opposite direction of the knockdown effect under WT conditions. Events exhibiting reduced variability within PS-dependent groups (PS-deficient, solid-like) compared to the total variability were classified as PS-dependent alternative exons (234 events).

S12E) Unsupervised hierarchical clustering of 185 exon inclusion events following rescue, showing differences in exon inclusion levels for TDP-43 variants compared to WT TDP-43. Values reported in the heatmap show the differences in inclusion levels, with positive values (increase in inclusion) in red and negative values (increase in skipping) in blue. In the annotation bar on the left, events identified as PS-dependent splicing events are shown.

S12F-G) Sashimi plots showing alternatively spliced exons in *SEMA3F* and, *BRD8*. The plot shows exon coverage and splice junction usage for NTC, KD and rescue in TDP-43 WT TDP-43 cells and the rescue for all TDP-43 variants analyzed.

S12H) Distribution of gene expression changes ($\log_2$ fold change) for genes containing exons that are included (left panel) or skipped (right panel) upon rescue. Each panel compares KD-versus-NTC and rescue-versus-KD conditions. Boxplots show the median (center line), interquartile range (Q1–Q3), and whiskers extending to 1.5 × IQR. Statistical significance was assessed using an unpaired two-sided t-test.

S12I) Annotation of genes identified as PS-dependent. lncRNA= long non-coding RNA

S12J) GO analysis enrichment (BP) for terms significantly enriched (*p*value < 0.01). GO terms are grouped based on semantic similarity, measured by the Jaccard index of their associated genes (threshold = 0.95)

S12K) Venn diagram showing the overlap of UPF1- dependent genes in from HEK293 and HeLa cells based on the dataset generated by Fritz et al., Colombo et al, and Imamachi et al. [104–106]. In bold, the number of the genes identified in at least two datasets and considered as UPF1- dependent genes.

S12L) Gene set enrichment analysis (GSEA) comparing RNA levels for TDP-43 rescue and knockdown condition. Gene set: UPF1 target genes in HEK293 and HeLa cells defined in Fig. S12K. The y axis displays the running enrichment score (ES) (upper panel), the position of UPF1 target genes (middle panel) and corresponding $\log_2$ fold changes (lower panel); the x axis shows the ranked position of each gene in the set. NES (normalized enriched score) and *p*value adjusted are shown in the plot.

Fig. S13

S13A) Experimental workflow for proteome profiling of TDP-43 variants with distinct PS properties. Proteome profiling was performed in TDP-43 KD cells rescued with different TDP-43 variants: WT, solid-like variants (5R, 12A, G376V), and PS-deficient variants (3W, ΔCR, 12D). Differential analysis is performed between rescue and KD conditions. Created in part with BioRender.com.

S13B) Analysis pipeline used to identify proteins whose abundance is associated with the PS properties of TDP-43. After protein identification, protein groups are filtered based on matrix completeness. Median-normalized protein intensities are analyzed using the limma framework to compare rescue versus KD conditions. Proteins showing at least two significant changes within either the PS-deficient or solid-like variant groups are considered for further analysis. Protein abundances are subsequently z-score normalized and PS-dependent proteins are defined as those for which the standard deviation within the PS-deficient and solid-like groups is lower than the global standard deviation across all conditions. This analysis identified 107 PS-dependent downregulated proteins and 131 PS-dependent upregulated proteins.

S13C) Comparison of the dynamic range of the protein abundance (expressed as protein concentration in the cell) in the acquired dataset with the OpenCell dataset (HEK293 cells).

S13 D) Precursors identified under the acquired conditions and their detection across all samples. The color scale indicates precursor completeness, expressed as the percentage of conditions in which each precursor was identified; red bars denote complete detection across the dataset.

S13E) Distribution of precursor intensities across analyzed conditions. Boxplots represent the distribution of precursor abundance values, with the central line indicating the median, boxes representing the interquartile range (first and third quartiles), and whiskers extending to 1.5× the interquartile range.

S13F) Coefficient of variation (CV) of protein abundance measurements across conditions. Violin plots show the distribution of CV values, with the median indicated by a red line.

S13G) Quantification of TDP-43 abundance across all analyzed conditions. Only peptide precursors detected in all conditions were used to estimate TDP-43 abundance.

S13H) Immunoblot showing the level of expression of TDP-43 across all conditions.

S13I) Correlation between RNA and protein abundance changes in TDP-43 WT rescue-versus-KD conditions. All genes and proteins are displayed in light grey, while those significantly regulated at both RNA and protein levels (adjusted *p*value < 0.05) are highlighted in blue. The Spearman correlation coefficient is calculated based on the subset of significantly regulated genes and proteins.

Fig. S14

S14A) Volcano plot showing differences in protein abundance between TDP-43 variants with distinct PS-properties between rescue and KD condition. The x-axis represents differential protein abundance, while the y axis shows statistical significance (adjusted *p*value, Benjamini–Hochberg correction). Proteins identified as PS-dependent upregulated and downregulated are shown in red and blue respectively.

S14B-C. Gene Ontology (GO) enrichment analysis (Biological Process) for terms significantly enriched (*p*value < 0.01) among upregulated (B) and downregulated (C) proteins upon rescue. Parental terms, determined based on semantic similarity, are displayed. Enrichment significance was calculated using a hypergeometric test, and bar colors represent the average protein levels in the differential comparison between solid-like and PS-deficient conditions.

Fig. S15

Summary of PS-dependent proteins and their mechanisms of regulation. The heatmap categorizes proteins as upregulated (left) or downregulated (right) upon rescue with TDP-43 WT.
The following layers of information are displayed from left to right:

- Regulatory mechanism indicates whether regulation is direct or indirect.
- Link to PS-dependent alternative splicing events specifies whether the protein is associated with exon inclusion or skipping.
- Link to PS-dependent RNA abundance changes highlights whether RNA abundance is upregulated or downregulated upon rescue.

- Functional link with TDP-43 PS-dependent interactors report whether the protein is regulated by PS-dependent interactors, this includes interaction between protein and mRNA (RNAInter database), functional link (STRING database) and both.
- Z-scores of protein abundance changes display protein abundance differences between solid-like and PS-deficient conditions (positive z-scores (red) indicate higher abundance in the solid-like condition, while negative z-scores (blue) indicate lower abundance).
- Differential analysis for WT compares protein abundance changes between rescue and knockdown (KD) conditions.

Fig. 16

S16A-B. Functional links between TDP-43 PS-dependent interactors (rows) and indirect targets (columns). The color code indicates the nature of the link: functional (STRING), protein-mRNA interaction (RNAInter), or both.

Fig. S17-18

Network representation of PS-dependent events across different levels of gene regulation for upregulated proteins (Fig. S17) and downregulated proteins (Fig. S18). The vertical axis illustrates the direct mechanism of regulation, where PS properties influence alternative splicing, RNA levels, and protein levels. The horizontal axis represents the indirect mechanism of regulation, in which PS-dependent TDP-43 interactors modulate protein levels. Direct interactions are depicted with solid black lines, while functional associations derived from database curation are shown with dotted lines based on the source of the link (purple=STRING, light violet=RNAInter, and dark violet= STRING and RNAInter). Nodes represent identified genes or proteins, with orange nodes indicating proteins annotated as regulators of gene expression or RNA binding (based on QuickGO annotations).

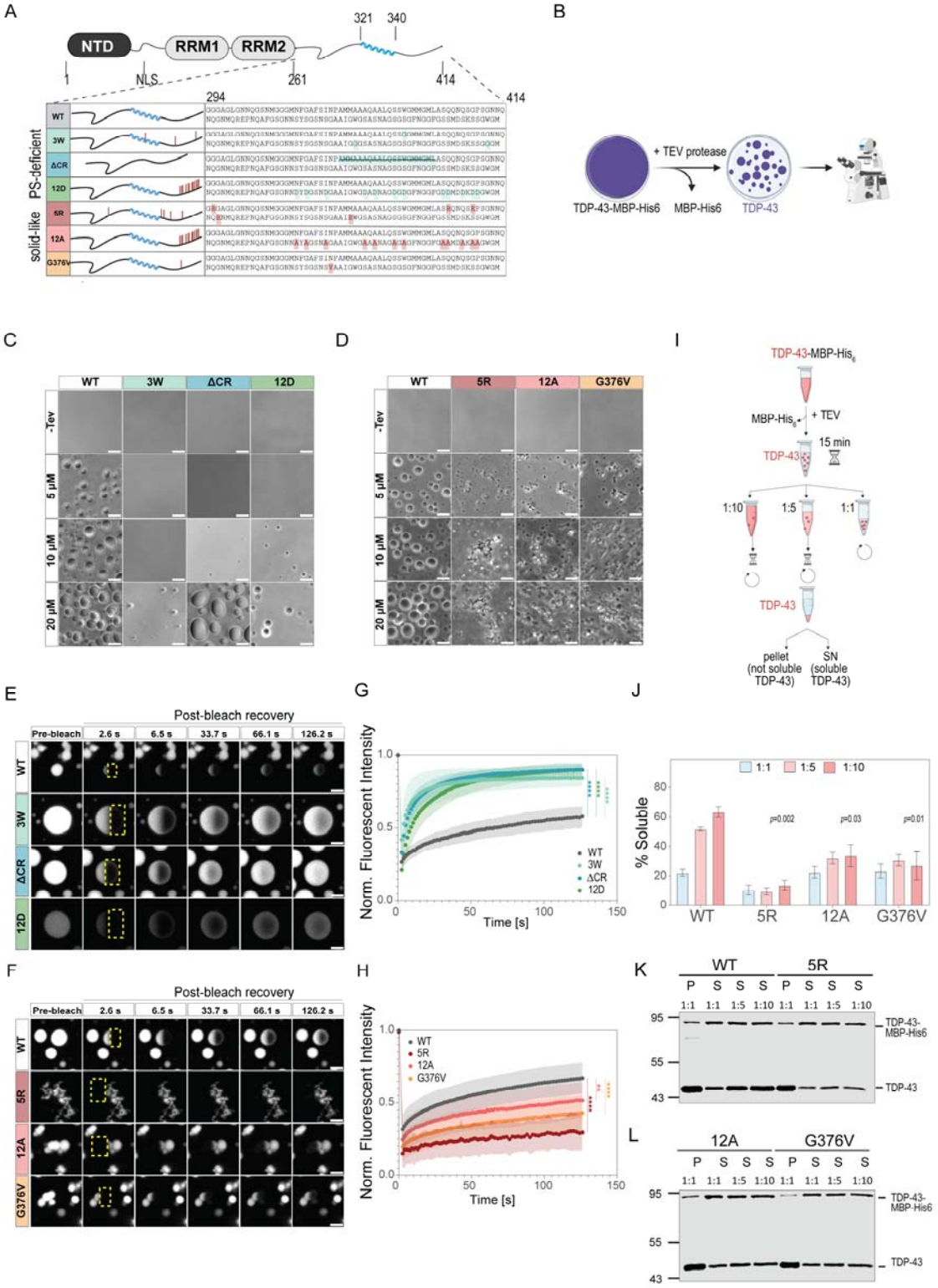
A
NTD
RRM1
RRM2
NLS
1
261
321
340
414
294
PS-deficient
solid-like
WT
3W
ΔCR
12D
5R
12A
G376V
B
+ TEV protease
TDP-43-MBP-His6
MBP-His6
TDP-43
C
WT
3W
ΔCR
12D
-Tev
5 µM
10 µM
20 µM
D
WT
5R
12A
G376V
I
TDP-43-MBP-His6
MBP-His6
+ TEV
15 min
1:10
1:5
1:1
pellet (not soluble TDP-43)
SN (soluble TDP-43)
E
Post-bleach recovery
Pre-bleach
2.6 s
6.5 s
33.7 s
66.1 s
126.2 s
G
Norm. Fluorescent Intensity
Time [s]
J
% Soluble
p=0.002
p=0.03
p=0.01
F
H
K
L
P
S
1:1
1:5
1:10
TDP-43-MBP-His6
TDP-43

Figure 1

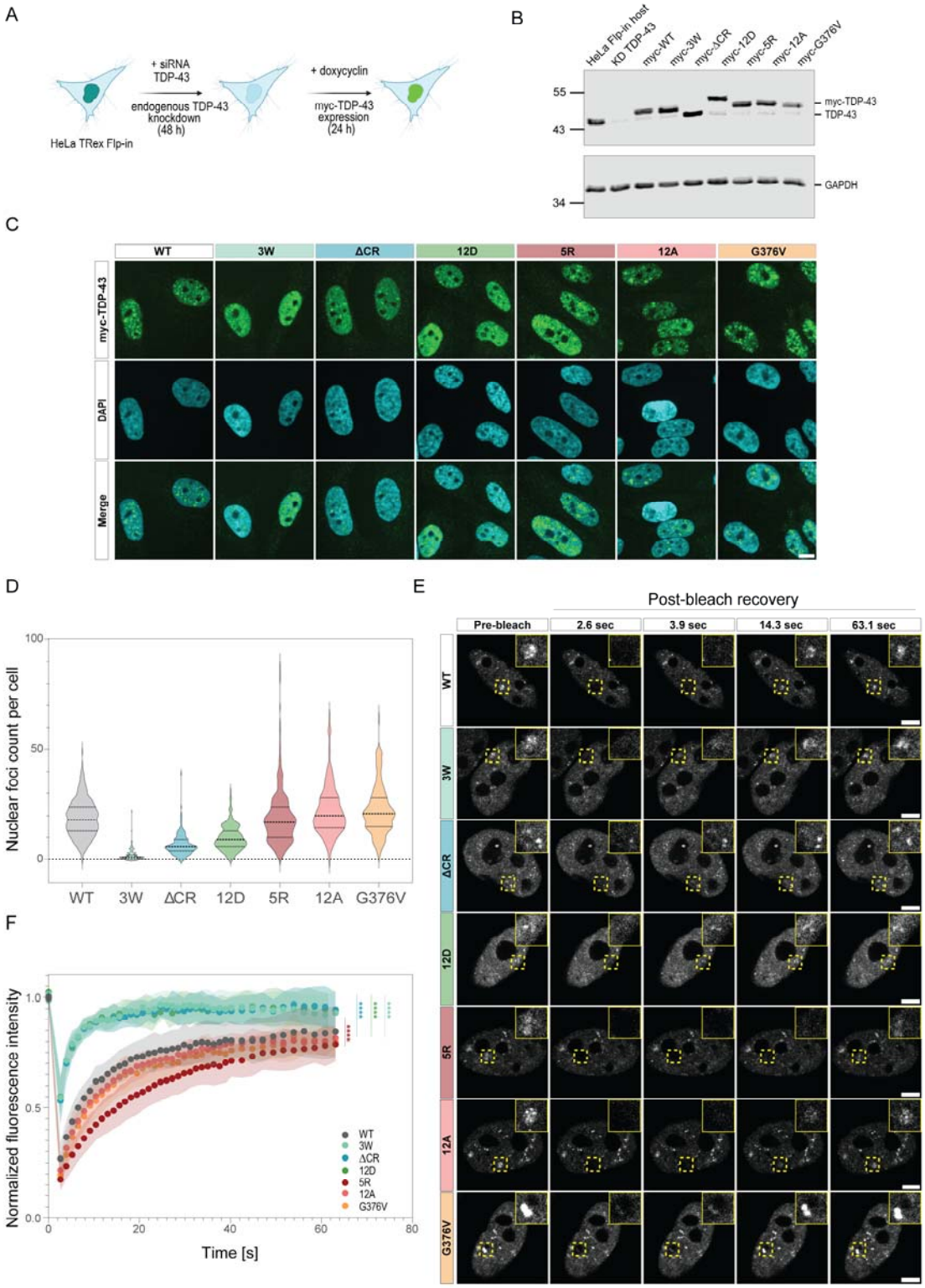
A
+ siRNA
TDP-43
endogenous TDP-43
knockdown
(48 h)
+ doxycyclin
myc-TDP-43
expression
(24 h)
HeLa TRex Flp-in
B
HeLa Flp-in host
KD TDP-43
myc-WT
myc-3W
myc-ΔCR
myc-12D
myc-5R
myc-12A
myc-G376V
55
43
34
myc-TDP-43
TDP-43
GAPDH
C
WT
3W
ΔCR
12D
5R
12A
G376V
myc-TDP-43
DAPI
Merge
D
Nuclear foci count per cell
100
50
0
WT
3W
ΔCR
12D
5R
12A
G376V
E
Post-bleach recovery
Pre-bleach
2.6 sec
3.9 sec
14.3 sec
63.1 sec
F
Normalized fluorescence intensity
1.0
0.5
0.0
0
20
40
60
80
Time [s]
WT
3W
ΔCR
12D
5R
12A
G376V

Figure 2

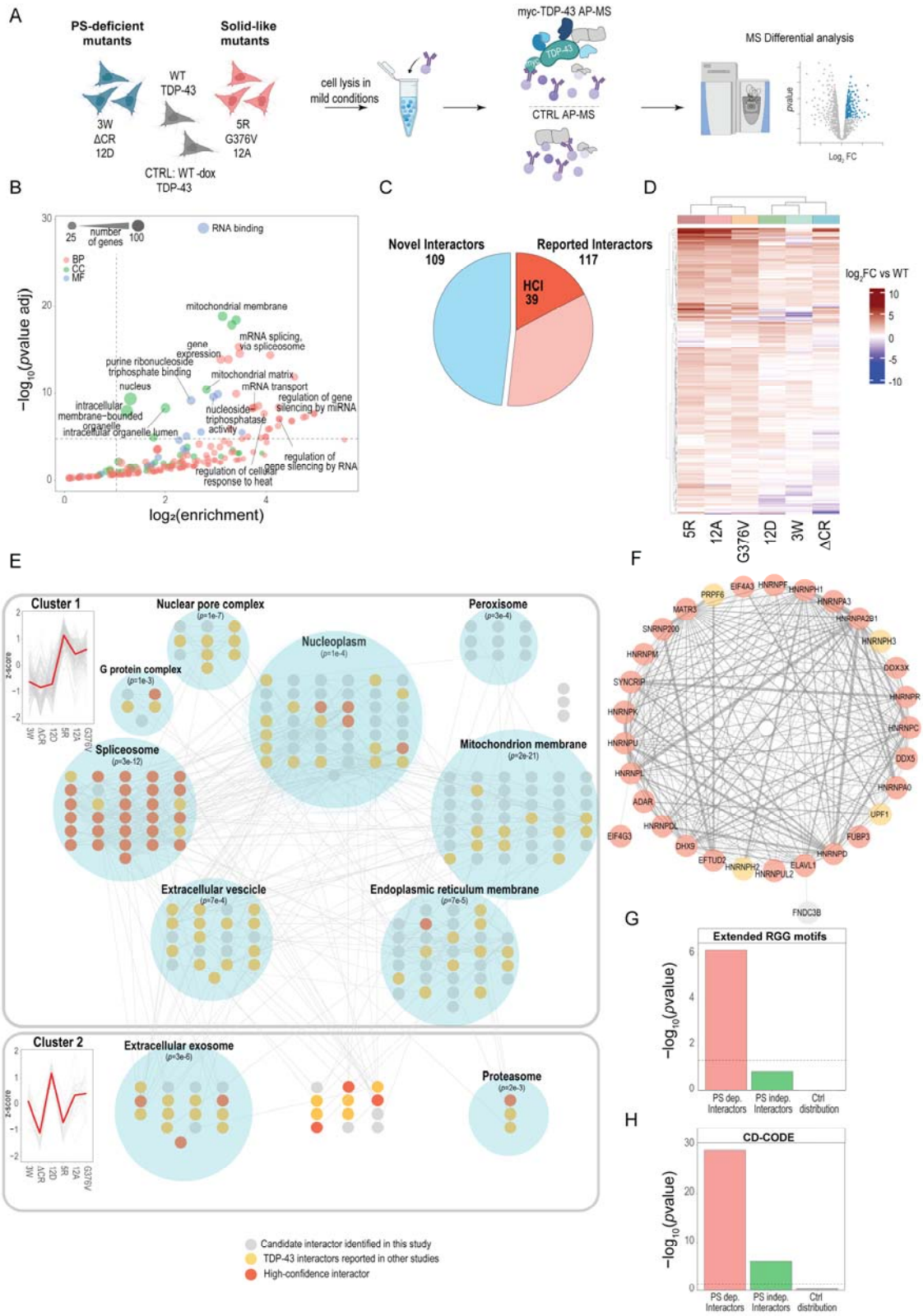
A
PS-deficient mutants
Solid-like mutants
WT TDP-43
3W ΔCR 12D
5R G376V 12A
CTRL: WT-dox TDP-43
cell lysis in mild conditions
myc-TDP-43 AP-MS
CTRL AP-MS
MS Differential analysis
pvalue
Log2 FC
B
RNA binding
number of genes
BP
CC
MF
mitochondrial membrane
mRNA splicing, via spliceosome
gene expression
purine ribonucleoside triphosphate binding
mitochondrial matrix
nucleus
mRNA transport
regulation of gene silencing by miRNA
intracellular membrane-bounded organelle
nucleoside-triphosphatase activity
intracellular organelle lumen
regulation of gene silencing by RNA
regulation of cellular response to heat
−log10(pvalue adj)
log2(enrichment)
C
Novel Interactors 109
Reported Interactors 117
HCI 39
D
log2FC vs WT
5R
12A
G376V
12D
3W
ΔCR
E
Cluster 1
Nuclear pore complex
Peroxisome
Nucleoplasm
G protein complex
Spliceosome
Mitochondrion membrane
Extracellular vescicle
Endoplasmic reticulum membrane
Cluster 2
Extracellular exosome
Proteasome
Candidate interactor identified in this study
TDP-43 interactors reported in other studies
High-confidence interactor
F
G
Extended RGG motifs
−log10(pvalue)
PS dep. Interactors
PS indep. Interactors
Ctrl distribution
H
CD-CODE

Figure 3

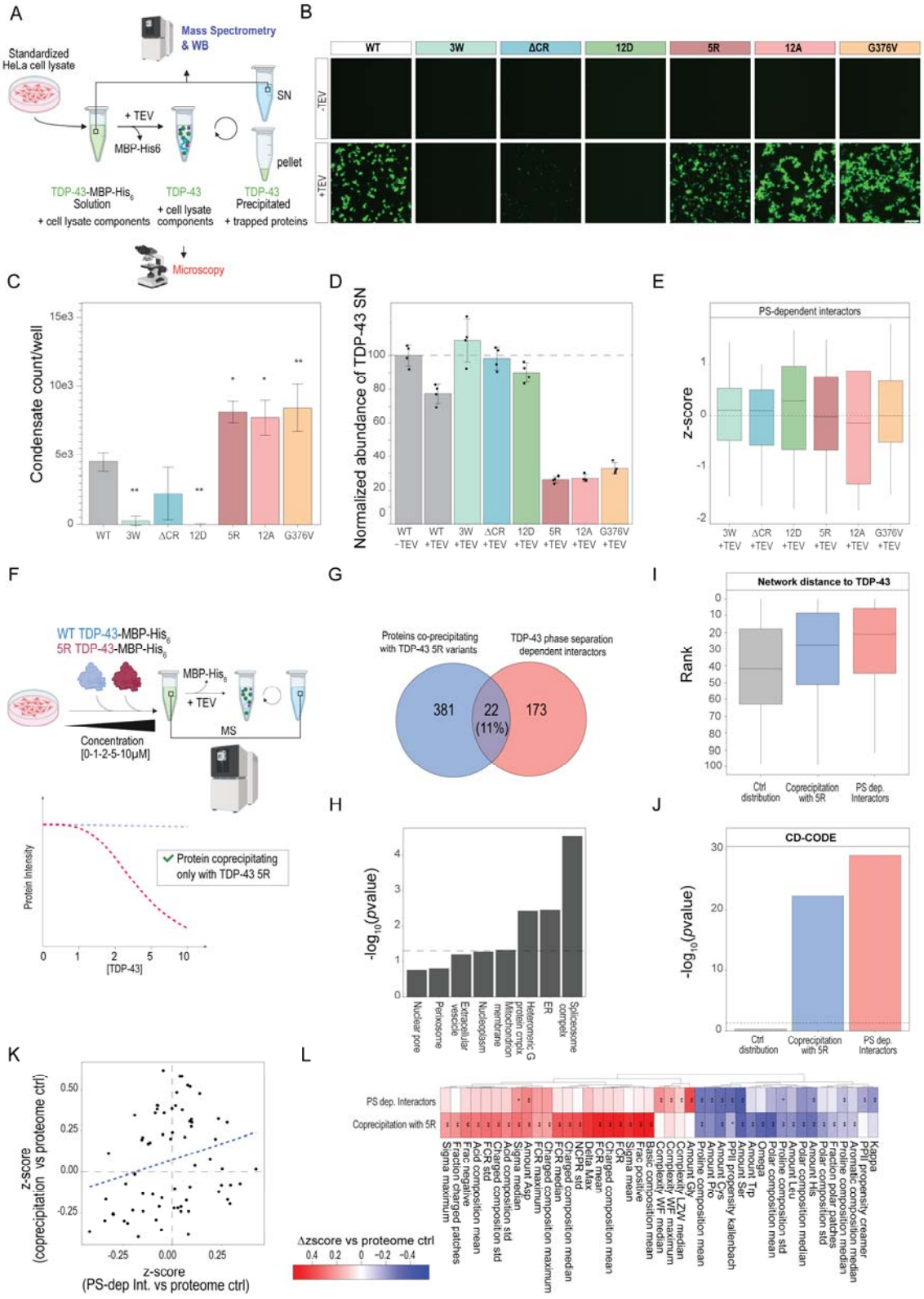
A
Mass Spectrometry & WB
Standardized HeLa cell lysate
+ TEV
MBP-His6
SN
pellet
TDP-43-MBP-His6 Solution + cell lysate components
TDP-43 + cell lysate components
TDP-43 Precipitated + trapped proteins
Microscopy
B
WT
3W
ΔCR
12D
5R
12A
G376V
-TEV
+TEV
C
Condensate count/well
D
Normalized abundance of TDP-43 SN
E
PS-dependent interactors
z-score
F
WT TDP-43-MBP-His6
5R TDP-43-MBP-His6
Concentration [0-1-2-5-10μM]
MS
Protein Intensity
Protein coprecipitating only with TDP-43 5R
[TDP-43]
G
Proteins co-precipitating with TDP-43 5R variants
TDP-43 phase separation dependent interactors
381
22 (11%)
173
H
-log10(pvalue)
Nuclear pore
Peroxisome
Extracellular vesicle
Nucleoplasm
Mitochondrion membrane
Heteromeric G protein cmplx
ER
Spliceosome complex
I
Network distance to TDP-43
Rank
Ctrl distribution
Coprecipitation with 5R
PS dep. Interactors
J
CD-CODE
-log10(pvalue)
K
z-score (coprecipitation vs proteome ctrl)
z-score (PS-dep Int. vs proteome ctrl)
L
PS dep. Interactors
Coprecipitation with 5R
Δzscore vs proteome ctrl

Figure 4

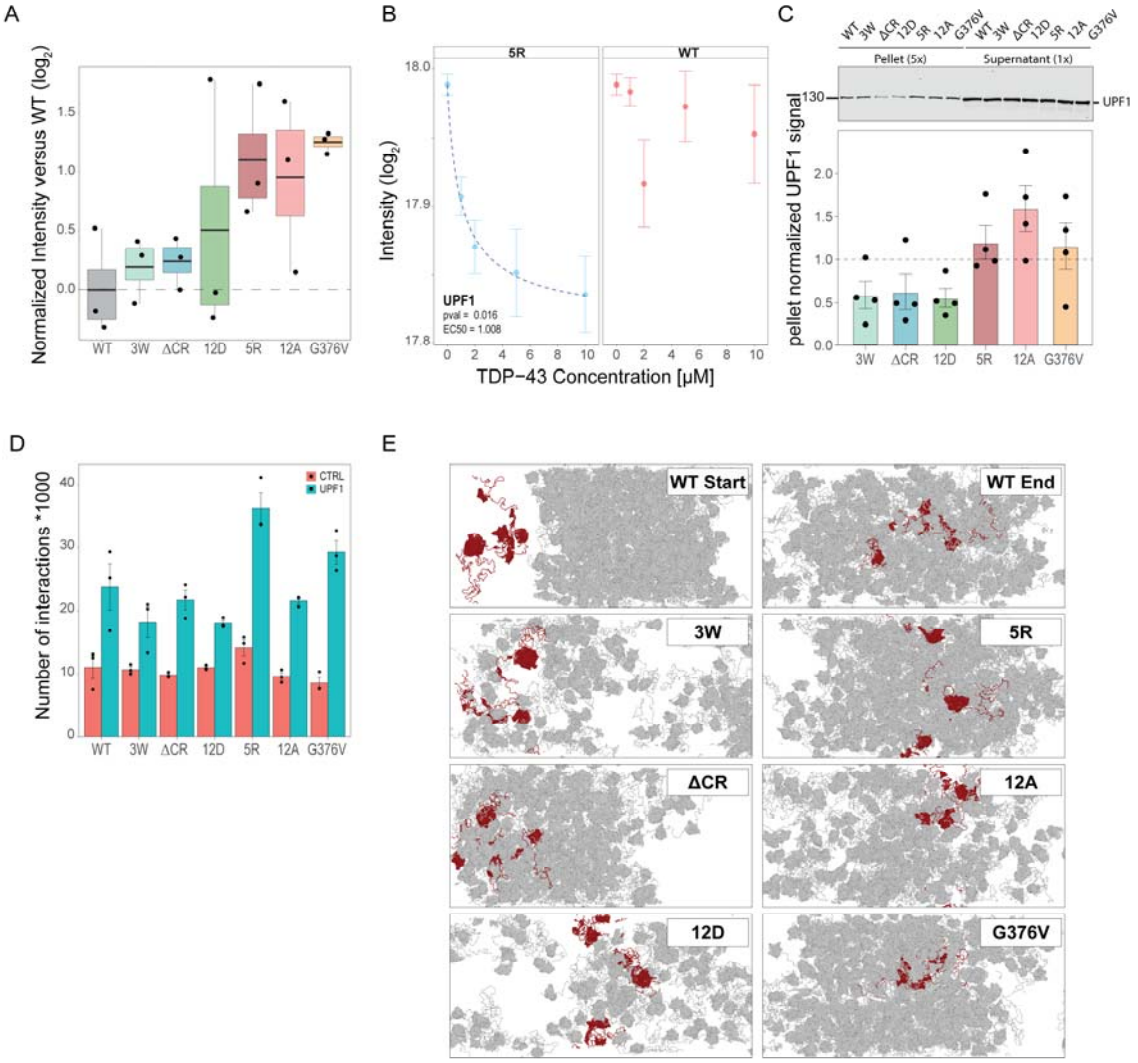
A
Normalized Intensity versus WT (log2)
WT
3W
ΔCR
12D
5R
12A
G376V
B
5R
WT
Intensity (log2)
UPF1
pval = 0.016
EC50 = 1.008
TDP−43 Concentration [μM]
C
Pellet (5x)
Supernatant (1x)
UPF1
pellet normalized UPF1 signal
D
CTRL
UPF1
Number of interactions *1000
E
WT Start
WT End
3W
5R
ΔCR
12A
12D
G376V

Figure 5

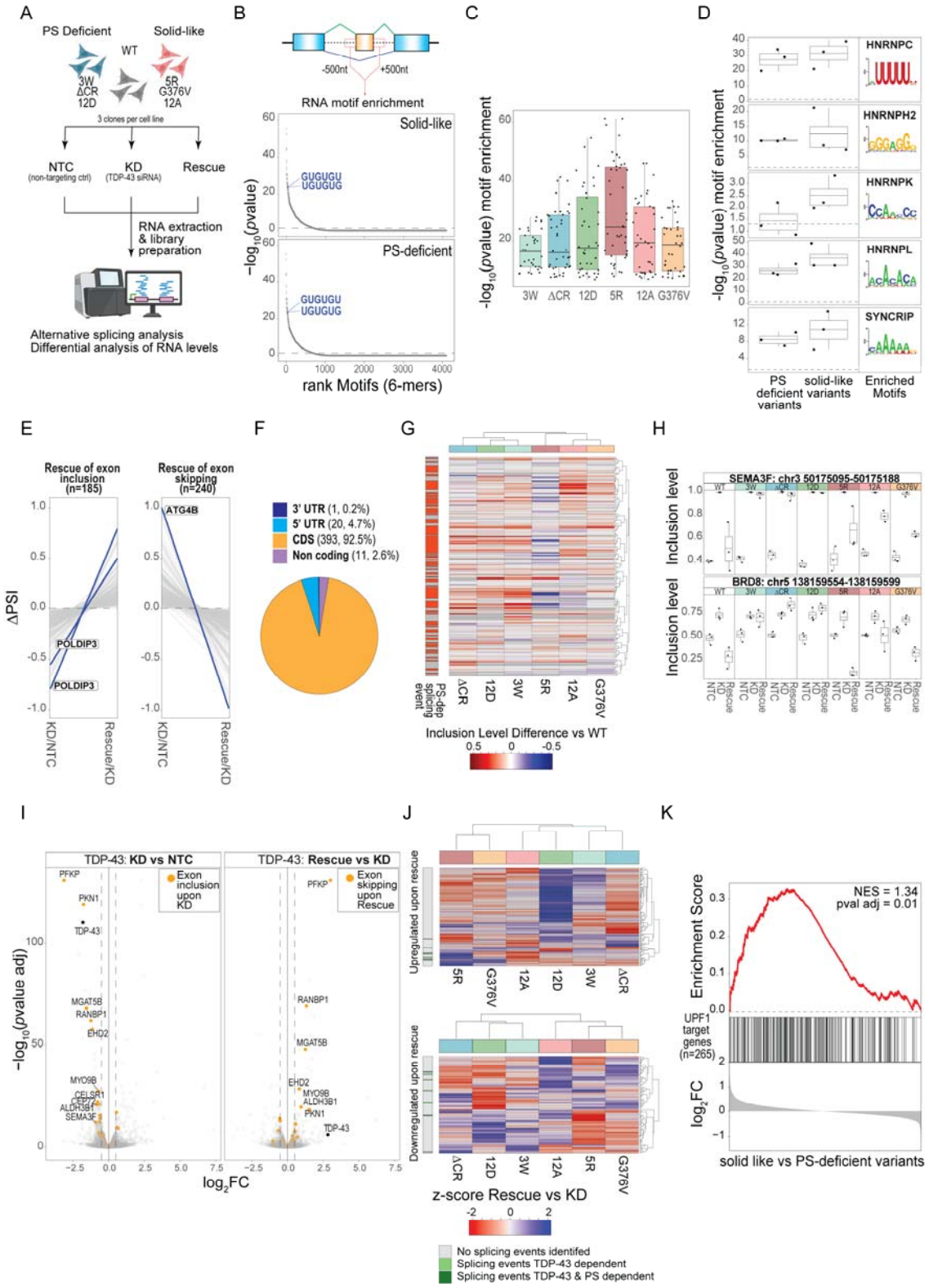
A
PS Deficient
Solid-like
WT
3W
ΔCR
12D
5R
G376V
12A
3 clones per cell line
NTC
(non-targeting ctrl)
KD
(TDP-43 siRNA)
Rescue
RNA extraction
& library
preparation
Alternative splicing analysis
Differential analysis of RNA levels
B
-500nt
+500nt
RNA motif enrichment
Solid-like
PS-deficient
GUGUGU
UGUGUG
−log10(pvalue)
rank Motifs (6-mers)
C
-log10(pvalue) motif enrichment
3W ΔCR 12D 5R 12A G376V
D
HNRNPC
HNRNPH2
HNRNPK
HNRNPL
SYNCRIP
-log10(pvalue) motif enrichment
PS deficient variants
solid-like variants
Enriched Motifs
E
Rescue of exon inclusion (n=185)
Rescue of exon skipping (n=240)
ΔPSI
POLDIP3
ATG4B
KD/NTC
Rescue/KD
F
3' UTR (1, 0.2%)
5' UTR (20, 4.7%)
CDS (393, 92.5%)
Non coding (11, 2.6%)
G
PS-dep splicing event
ΔCR
12D
3W
5R
12A
G376V
Inclusion Level Difference vs WT
H
SEMA3F: chr3 50175095-50175188
BRD8: chr5 138159554-138159599
Inclusion level
WT
3W
ΔCR
12D
5R
12A
G376V
NTC
KD
Rescue
I
TDP-43: KD vs NTC
TDP-43: Rescue vs KD
Exon inclusion upon KD
Exon skipping upon Rescue
PFKP
PKN1
TDP-43
MGAT5B
RANBP1
EHD2
MYO9B
CELSR1
ALDH3B1
SEMA3F
−log10(pvalue adj)
log2FC
J
Upregulated upon rescue
Downregulated upon rescue
5R
G376V
12A
12D
3W
ΔCR
z-score Rescue vs KD
No splicing events identifed
Splicing events TDP-43 dependent
Splicing events TDP-43 & PS dependent
K
NES = 1.34
pval adj = 0.01
Enrichment Score
UPF1 target genes (n=265)
log2FC
solid like vs PS-deficient variants

Figure 6

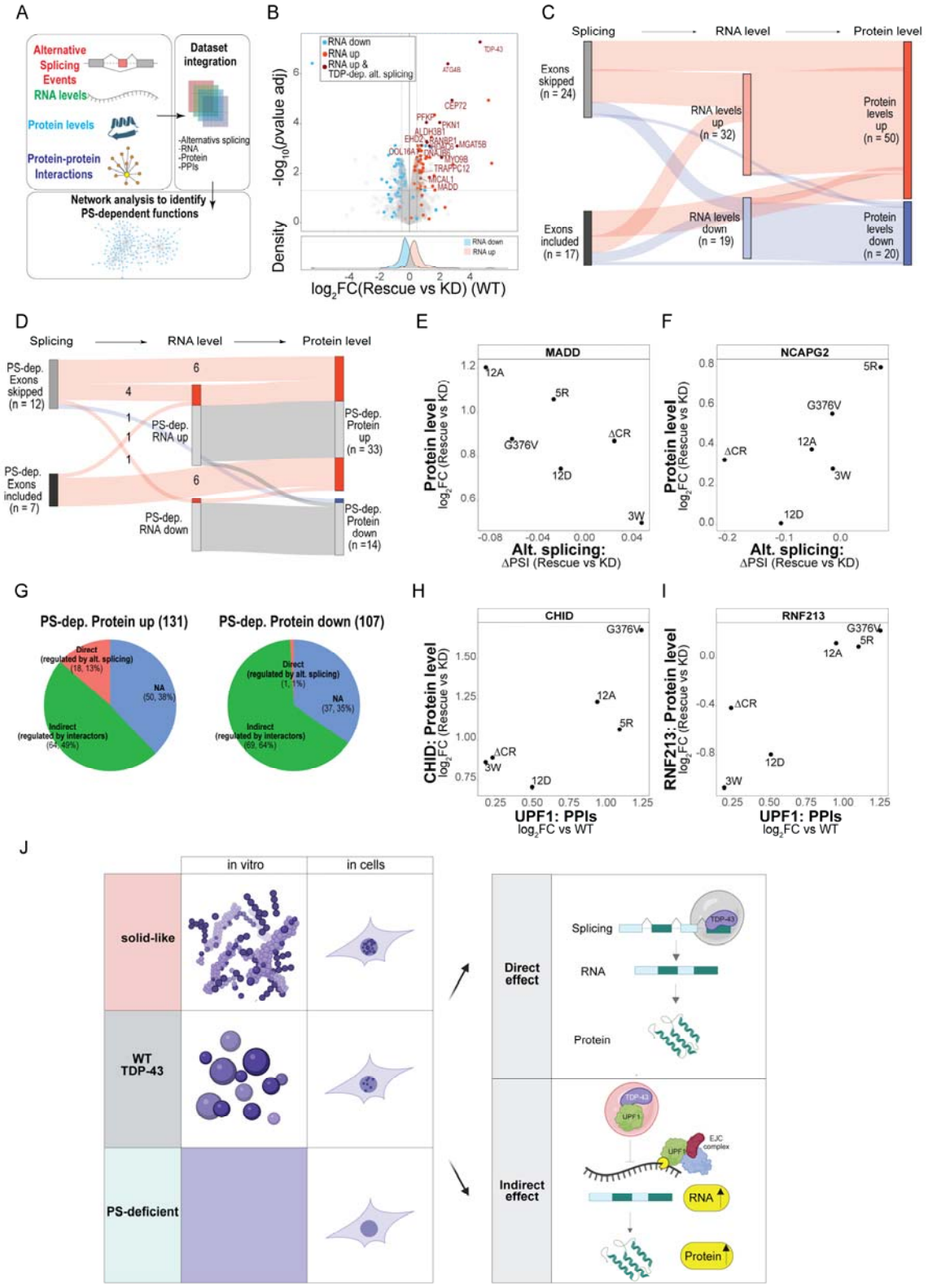
A
Alternative Splicing Events
RNA levels
Protein levels
Protein-protein Interactions
Dataset integration
-Alternativs splicing
-RNA
-Protein
-PPIs
Network analysis to identify PS-dependent functions
B
RNA down
RNA up
RNA up & TDP-dep. alt. splicing
-log10(pvalue adj)
Density
log2FC(Rescue vs KD) (WT)
C
Splicing
RNA level
Protein level
Exons skipped (n = 24)
Exons included (n = 17)
RNA levels up (n = 32)
RNA levels down (n = 19)
Protein levels up (n = 50)
Protein levels down (n = 20)
D
PS-dep. Exons skipped (n = 12)
PS-dep. Exons included (n = 7)
PS-dep. RNA up
PS-dep. RNA down
PS-dep. Protein up (n = 33)
PS-dep. Protein down (n =14)
E
MADD
Protein level log2FC (Rescue vs KD)
Alt. splicing: ΔPSI (Rescue vs KD)
F
NCAPG2
G
PS-dep. Protein up (131)
PS-dep. Protein down (107)
Direct (regulated by alt. splicing)
Indirect (regulated by interactors)
NA
H
CHID
CHID: Protein level log2FC (Rescue vs KD)
UPF1: PPIs log2FC vs WT
I
RNF213
RNF213: Protein level log2FC (Rescue vs KD)
J
in vitro
in cells
solid-like
WT TDP-43
PS-deficient
Direct effect
Indirect effect
Splicing
RNA
Protein
TDP-43
UPF1
EJC complex

Figure 7

Bibliography

1. Ling, S.-C., Polymenidou, M., and Cleveland, D.W. (2013). Converging Mechanisms in ALS and FTD: Disrupted RNA and Protein Homeostasis. Neuron *79*, 416–438. https://doi.org/10.1016/j.neuron.2013.07.033.

2. Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., et al. (2006). Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. Science *314*, 130–133. https://doi.org/10.1126/science.1134108.

3. Amador-Ortiz, C., Lin, W., Ahmed, Z., Personett, D., Davies, P., Duara, R., Graff-Radford, N.R., Hutton, M.L., and Dickson, D.W. (2007). TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. Annals of Neurology *61*, 435–445. https://doi.org/10.1002/ana.21154.

4. Josephs, K.A., Whitwell, J.L., Weigand, S.D., Murray, M.E., Tosakulwong, N., Liesinger, A.M., Petrucelli, L., Senjem, M.L., Knopman, D.S., Boeve, B.F., et al. (2014). TDP-43 is a key player in the clinical features associated with Alzheimer's disease. Acta Neuropathol *127*, 811–824. https://doi.org/10.1007/s00401-014-1269-z.

5. Nelson, P.T., Dickson, D.W., Trojanowski, J.Q., Jack, C.R., Boyle, P.A., Arfanakis, K., Rademakers, R., Alafuzoff, I., Attems, J., Brayne, C., et al. (2019). Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. Brain *142*, 1503–1527. https://doi.org/10.1093/brain/awz099.

6. Mackenzie, I.R., Rademakers, R., and Neumann, M. (2010). TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. The Lancet Neurology *9*, 995–1007. https://doi.org/10.1016/S1474-4422(10)70195-2.

7. Rummens, J., and Da Cruz, S. (2025). RNA-binding proteins in ALS and FTD: from pathogenic mechanisms to therapeutic insights. Mol Neurodegeneration *20*, 64. https://doi.org/10.1186/s13024-025-00851-y.

8. Rummens, J., Khalil, B., Yıldırım, G., Silva, P., Zorzini, V., Peredo, N., Wojno, M., Ramakers, M., Bosch, L.V.D., Damme, P.V., et al. (2025). TDP-43 seeding induces cytoplasmic aggregation heterogeneity and nuclear loss of function of TDP-43. Neuron *0*. https://doi.org/10.1016/j.neuron.2025.03.004.

9. Scialò, C., Zhong, W., Jagannath, S., Wilkins, O., Caredio, D., Hruska-Plochan, M., Lurati, F., Peter, M., Cecco, E.D., Celauro, L., et al. (2025). Seeded aggregation of TDP-43 induces its loss of function and reveals early pathological signatures. Neuron *0*. https://doi.org/10.1016/j.neuron.2025.03.008.

10. Ratti, A., and Buratti, E. (2016). Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins. Journal of Neurochemistry *138*, 95–111. https://doi.org/10.1111/jnc.13625.

11. Klim, J.R., Williams, L.A., Limone, F., Guerra San Juan, I., Davis-Dusenbery, B.N., Mordes, D.A., Burberry, A., Steinbaugh, M.J., Gamage, K.K., Kirchner, R., et al. (2019). ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. Nat Neurosci *22*, 167–179. https://doi.org/10.1038/s41593-018-0300-4.

12. Brown, A.-L., Wilkins, O.G., Keuss, M.J., Hill, S.E., Zanovello, M., Lee, W.C., Bampton, A., Lee, F.C.Y., Masino, L., Qi, Y.A., et al. (2022). TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. Nature *603*, 131–137. https://doi.org/10.1038/s41586-022-04436-3.

13. Ma, X.R., Prudencio, M., Koike, Y., Vatsavayai, S.C., Kim, G., Harbinski, F., Briner, A., Rodriguez, C.M., Guo, C., Akiyama, T., et al. (2022). TDP-43 represses cryptic exon inclusion in the FTD-ALS gene UNC13A. Nature *603*, 124–130. https://doi.org/10.1038/s41586-022-04424-7.

14. Melamed, Z., López-Erauskin, J., Baughn, M.W., Zhang, O., Drenner, K., Sun, Y., Freyermuth, F., McMahon, M.A., Beccari, M.S., Artates, J.W., et al. (2019). Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. Nat Neurosci *22*, 180–190. https://doi.org/10.1038/s41593-018-0293-z.

15. Bjork, R.T., Mortimore, N.P., Loganathan, S., and Zarnescu, D.C. (2022). Dysregulation of Translation in TDP-43 Proteinopathies: Deficits in the RNA Supply Chain and Local Protein Production. Front Neurosci *16*, 840357. https://doi.org/10.3389/fnins.2022.840357.

16. Alessandrini, F., Wright, M., Kurosaki, T., Perloff, O.S., Ngo, M., Kofler, J.K., Donnelly, C.J., Maquat, L.E., and Kiskinis, E. (2026). TDP-43 dysfunction compromises UPF1-dependent mRNA metabolism in ALS. Neuron *114*, 640-660.e10. https://doi.org/10.1016/j.neuron.2025.11.001.

17. Bryce-Smith, S., Brown, A.-L., Chien, M.Z.Y.J., Dattilo, D., Mehta, P.R., Mattedi, F., Barattucci, S., Mikheenko, A., Zanovello, M., Pellegrini, F., et al. (2025). TDP-43 loss induces cryptic polyadenylation in ALS/FTD. Nat Neurosci *28*, 2190–2200. https://doi.org/10.1038/s41593-025-02050-w.

18. Eck, R.J., Valdmanis, P.N., Liachko, N.F., and Kraemer, B.C. (2025). Alternative 3' UTR polyadenylation is disrupted in the rNLS8 mouse model of ALS/FTLD. Mol Brain *18*, 1. https://doi.org/10.1186/s13041-025-01174-1.

19. Zeng, Y., Lovchykova, A., Akiyama, T., Rayner, S.L., Maheswari Jawahar, V., Liu, C., Sianto, O., Guo, C., Calliari, A., Prudencio, M., et al. (2025). TDP-43 nuclear loss in FTD/ALS causes widespread alternative polyadenylation changes. Nat Neurosci *28*, 2180–2189. https://doi.org/10.1038/s41593-025-02049-3.

20. Chou, C.-C., Zhang, Y., Umoh, M.E., Vaughan, S.W., Lorenzini, I., Liu, F., Sayegh, M., Donlin-Asp, P.G., Chen, Y.H., Duong, D.M., et al. (2018). TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. Nat Neurosci *21*, 228–239. https://doi.org/10.1038/s41593-017-0047-3.

21. Gasset-Rosa, F., Lu, S., Yu, H., Chen, C., Melamed, Z., Guo, L., Shorter, J., Da Cruz, S., and Cleveland, D.W. (2019). Cytoplasmic TDP-43 De-mixing Independent of Stress Granules Drives Inhibition of Nuclear Import, Loss of Nuclear TDP-43, and Cell Death. Neuron *102*, 339-357.e7. https://doi.org/10.1016/j.neuron.2019.02.038.

22. Dammer, E.B., Fallini, C., Gozal, Y.M., Duong, D.M., Rossoll, W., Xu, P., Lah, J.J., Levey, A.I., Peng, J., Bassell, G.J., et al. (2012). Coaggregation of RNA-Binding Proteins in a Model of TDP-43 Proteinopathy with Selective RGG Motif Methylation and a Role for RRM1 Ubiquitination. PLoS ONE *7*, e38658. https://doi.org/10.1371/journal.pone.0038658.

23. van der Zee, J., Dillen, L., Baradaran-Heravi, Y., Gossye, H., Koçoğlu, C., Cuyt, I., Dermaut, B., Sieben, A., Baets, J., De Jonghe, P., et al. (2021). Family-based exome sequencing identifies *RBM45* as a possible candidate gene for frontotemporal dementia and amyotrophic lateral sclerosis. Neurobiology of Disease *156*, 105421. https://doi.org/10.1016/j.nbd.2021.105421.

24. Collins, M., Riascos, D., Kovalik, T., An, J., Krupa, K., Krupa, K., Hood, B.L., Conrads, T.P., Renton, A.E., Traynor, B.J., et al. (2012). The RNA-binding motif 45 (RBM45) protein accumulates in inclusion bodies in amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration with TDP-43 inclusions (FTLD-TDP) patients. Acta Neuropathol *124*, 717–732. https://doi.org/10.1007/s00401-012-1045-x.

25. Riemenschneider, H., Guo, Q., Bader, J., Frottin, F., Farny, D., Kleinberger, G., Haass, C., Mann, M., Hartl, F.U., Baumeister, W., et al. (2022). Gel-like inclusions of C-terminal fragments of TDP-43 sequester stalled proteasomes in neurons. EMBO Rep *23*, e53890. https://doi.org/10.15252/embr.202153890.

26. Wang, P., Wander, C.M., Yuan, C.-X., Bereman, M.S., and Cohen, T.J. (2017). Acetylation-induced TDP-43 pathology is suppressed by an HSF1-dependent chaperone program. Nat Commun *8*, 82. https://doi.org/10.1038/s41467-017-00088-4.

27. Volkening, K., Leystra-Lantz, C., Yang, W., Jaffee, H., and Strong, M.J. (2009). Tar DNA binding protein of 43 kDa (TDP-43), 14-3-3 proteins and copper/zinc superoxide dismutase (SOD1) interact to modulate NFL mRNA stability. Implications for altered RNA processing in amyotrophic lateral sclerosis (ALS). Brain Research *1305*, 168–182. https://doi.org/10.1016/j.brainres.2009.09.105.

28. Liu-Yesucevitz, L., Bilgutay, A., Zhang, Y.-J., Vanderwyde, T., Citro, A., Mehta, T., Zaarur, N., McKee, A., Bowser, R., Sherman, M., et al. (2010). Tar DNA Binding Protein-43 (TDP-43) Associates with Stress Granules: Analysis of Cultured Cells and Pathological Brain Tissue. PLOS ONE *5*, e13250. https://doi.org/10.1371/journal.pone.0013250.

29. Bentmann, E., Neumann, M., Tahirovic, S., Rodde, R., Dormann, D., and Haass, C. (2012). Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). J Biol Chem *287*, 23079–23094. https://doi.org/10.1074/jbc.M111.328757.

30. Freibaum, B.D., Chitta, R.K., High, A.A., and Taylor, J.P. (2010). Global Analysis of TDP-43 Interacting Proteins Reveals Strong Association with RNA Splicing and Translation Machinery. J. Proteome Res. *9*, 1104–1120. https://doi.org/10.1021/pr901076y.

31. Ling, S.-C., Albuquerque, C.P., Han, J.S., Lagier-Tourenne, C., Tokunaga, S., Zhou, H., and Cleveland, D.W. (2010). ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. Proceedings of the National Academy of Sciences *107*, 13318–13323. https://doi.org/10.1073/pnas.1008227107.

32. Cheng, F., Chapman, T., Venturato, J., Davidson, J.M., Polido, S.A., Rosa-Fernandes, L., San Gil, R., Suddull, H.J., Zhang, S., Macaslam, C.Y., et al. (2025). Proteomics Analysis of the TDP-43 Interactome in Cellular Models of ALS Pathogenesis. Journal of Neurochemistry *169*, e70079. https://doi.org/10.1111/jnc.70079.

33. Feneberg, E., Gordon, D., Thompson, A.G., Finelli, M.J., Dafinca, R., Candalija, A., Charles, P.D., Mäger, I., Wood, M.J., Fischer, R., et al. (2020). An ALS-linked mutation in TDP-43 disrupts normal protein interactions in the motor neuron response to oxidative stress. Neurobiol Dis *144*, 105050. https://doi.org/10.1016/j.nbd.2020.105050.

34. Dormann, D., and Haass, C. (2011). TDP-43 and FUS: a nuclear affair. Trends in Neurosciences *34*, 339–348. https://doi.org/10.1016/j.tins.2011.05.002.

35. Wolozin, B., and Ivanov, P. (2019). Stress granules and neurodegeneration. Nat Rev Neurosci *20*, 649–666. https://doi.org/10.1038/s41583-019-0222-5.

36. Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013). Stress granules as crucibles of ALS pathogenesis. J

Cell Biol *201*, 361–372. https://doi.org/10.1083/jcb.201302044.

37. Yan, X., Kuster, D., Mohanty, P., Nijssen, J., Pombo-García, K., Garcia Morato, J., Rizuan, A., Franzmann, T.M., Sergeeva, A., Ly, A.M., et al. (2025). Intra-condensate demixing of TDP-43 inside stress granules generates pathological aggregates. Cell *188*, 4123-4140.e18. https://doi.org/10.1016/j.cell.2025.04.039.

38. Wang, A., Conicella, A.E., Schmidt, H.B., Martin, E.W., Rhoads, S.N., Reeb, A.N., Nourse, A., Ramirez Montero, D., Ryan, V.H., Rohatgi, R., et al. (2018). A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. EMBO J *37*, e97452. https://doi.org/10.15252/embj.201797452.

39. Naganuma, T., Nakagawa, S., Tanigawa, A., Sasaki, Y.F., Goshima, N., and Hirose, T. (2012). Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. EMBO J *31*, 4020–4034. https://doi.org/10.1038/emboj.2012.251.

40. Pérez-Berlanga, M., Wiersma, V.I., Zbinden, A., De Vos, L., Wagner, U., Foglieni, C., Mallona, I., Betz, K.M., Cléry, A., Weber, J., et al. (2023). Loss of TDP-43 oligomerization or RNA binding elicits distinct aggregation patterns. The EMBO Journal *42*, e111719. https://doi.org/10.15252/embj.2022111719.

41. Gopal, P.P., Nirschl, J.J., Klinman, E., and Holzbaur, E.L.F. (2017). Amyotrophic lateral sclerosis-linked mutations increase the viscosity of liquid-like TDP-43 RNP granules in neurons. Proceedings of the National Academy of Sciences *114*, E2466–E2475. https://doi.org/10.1073/pnas.1614462114.

42. Conicella, A.E., Zerze, G.H., Mittal, J., and Fawzi, N.L. (2016). ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure *24*, 1537–1549. https://doi.org/10.1016/j.str.2016.07.007.

43. Conicella, A.E., Dignon, G.L., Zerze, G.H., Schmidt, H.B., D'Ordine, A.M., Kim, Y.C., Rohatgi, R., Ayala, Y.M., Mittal, J., and Fawzi, N.L. (2020). TDP-43 α-helical structure tunes liquid–liquid phase separation and function. Proc. Natl. Acad. Sci. U.S.A. *117*, 5883–5894. https://doi.org/10.1073/pnas.1912055117.

44. Li, H.-R., Chiang, W.-C., Chou, P.-C., Wang, W.-J., and Huang, J. (2018). TAR DNA-binding protein 43 (TDP-43) liquid–liquid phase separation is mediated by just a few aromatic residues. Journal of Biological Chemistry *293*, 6090–6098. https://doi.org/10.1074/jbc.AC117.001037.

45. Schmidt, H.B., Barreau, A., and Rohatgi, R. (2019). Phase separation-deficient TDP43 remains functional in splicing. Nat Commun *10*, 4890. https://doi.org/10.1038/s41467-019-12740-2.

46. Mohanty, P., Shenoy, J., Rizuan, A., Mercado-Ortiz, J.F., Fawzi, N.L., and Mittal, J. (2023). A synergy between site-specific and transient interactions drives the phase separation of a disordered, low-complexity domain. Proceedings of the National Academy of Sciences *120*, e2305625120. https://doi.org/10.1073/pnas.2305625120.

47. Afroz, T., Hock, E.-M., Ernst, P., Foglieni, C., Jambeau, M., Gilhespy, L.A.B., Laferriere, F., Maniecka, Z., Plückthun, A., Mittl, P., et al. (2017). Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. Nat Commun *8*, 45. https://doi.org/10.1038/s41467-017-00062-0.

48. Chappidi, N., Quail, T., Doll, S., Vogel, L.T., Aleksandrov, R., Felekyan, S., Kühnemuth, R., Stoynov, S., Seidel, C.A.M., Brugués, J., et al. (2024). PARP1-DNA co-condensation drives DNA repair site assembly to prevent disjunction of broken DNA ends. Cell *0*. https://doi.org/10.1016/j.cell.2024.01.015.

49. Spegg, V., and Altmeyer, M. (2021). Biomolecular condensates at sites of DNA damage: More than just a phase,. DNA Repair (Amst) *106*, 103179. https://doi.org/10.1016/j.dnarep.2021.103179.

50. Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. Cell *175*, 1842-1855.e16. https://doi.org/10.1016/j.cell.2018.10.042.

51. Sabari, B.R., Dall'Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science *361*, eaar3958. https://doi.org/10.1126/science.aar3958.

52. Shrinivas, K., Sabari, B.R., Coffey, E.L., Klein, I.A., Boija, A., Zamudio, A.V., Schuijers, J., Hannett, N.M., Sharp, P.A., Young, R.A., et al. (2019). Enhancer Features that Drive Formation of Transcriptional Condensates. Molecular Cell *75*, 549-561.e7. https://doi.org/10.1016/j.molcel.2019.07.009.

53. Wei, M.-T., Chang, Y.-C., Shimobayashi, S.F., Shin, Y., Strom, A.R., and Brangwynne, C.P. (2020). Nucleated transcriptional condensates amplify gene expression. Nat Cell Biol *22*, 1187–1196. https://doi.org/10.1038/s41556-020-00578-6.

54. Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. Science *361*, 412–415. https://doi.org/10.1126/science.aar4199.

55. Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P., et al. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. Nat Struct Mol Biol *25*, 833–840. https://doi.org/10.1038/s41594-018-0112-y.

56. Bremer, A., Lang, W.H., Kempen, R.P., Sweta, K., Taylor, A.B., Borgia, M.B., Ansari, A.Z., and Mittag, T. (2025). Reconciling competing models on the roles of condensates and soluble complexes in transcription factor function. Molecular Cell *85*, 2718-2732.e7. https://doi.org/10.1016/j.molcel.2025.06.008.

57. Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. Cell *169*, 13–23. https://doi.org/10.1016/j.cell.2017.02.007.

58. Giudice, J., and Jiang, H. (2024). Splicing regulation through biomolecular condensates and membraneless organelles. Nat Rev Mol Cell Biol *25*, 683–700. https://doi.org/10.1038/s41580-024-00739-7.

59. Parker, D.M., Winkenbach, L.P., and Osborne Nishimura, E. (2022). It's Just a Phase: Exploring the Relationship Between mRNA, Biomolecular Condensates, and Translational Control. Front. Genet. *13*, 931220. https://doi.org/10.3389/fgene.2022.931220.

60. Hallegger, M., Chakrabarti, A.M., Lee, F.C.Y., Lee, B.L., Amalietti, A.G., Odeh, H.M., Copley, K.E., Rubien, J.D., Portz, B., Kuret, K., et al. (2021). TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. Cell *184*, 4680-4696.e22. https://doi.org/10.1016/j.cell.2021.07.018.

61. Koehler, L.C., Grese, Z.R., Bastos, A.C.S., Mamede, L.D., Heyduk, T., and Ayala, Y.M. (2022). TDP-43 Oligomerization and Phase Separation Properties Are Necessary for Autoregulation. Front Neurosci *16*, 818655. https://doi.org/10.3389/fnins.2022.818655.

62. Gao, J., Wang, L., Ren, X., Dunn, J.R., Peters, A., Miyagi, M., Fujioka, H., Zhao, F., Askwith, C., Liang, J., et al. (2021). Translational regulation in the brain by TDP-43 phase separation. J Cell Biol *220*, e202101019. https://doi.org/10.1083/jcb.202101019.

63. Vishal, S.S., Wijegunawardana, D., Salaikumaran, M.R., and Gopal, P.P. (2022). Sequence Determinants of TDP-43 Ribonucleoprotein Condensate Formation and Axonal Transport in Neurons. Front Cell Dev Biol *10*, 876893. https://doi.org/10.3389/fcell.2022.876893.

64. Zibold, J., Lessard, L.E.R., Picard, F., Da Silva, L.G., Zadorozhna, Y., Streichenberger, N., Belotti, E., Osseni, A., Emerit, A., Errazuriz-Cerda, E., et al. (2024). The new missense G376V-TDP-43 variant induces late-onset distal myopathy but not amyotrophic lateral sclerosis. Brain *147*, 1768–1783. https://doi.org/10.1093/brain/awad410.

65. Gruijs da Silva, L.A., Simonetti, F., Hutten, S., Riemenschneider, H., Sternburg, E.L., Pietrek, L.M., Gebel, J., Dötsch, V., Edbauer, D., Hummer, G., et al. (2022). Disease-linked TDP-43 hyperphosphorylation suppresses TDP-43 condensation and aggregation. EMBO J *41*, e108443. https://doi.org/10.15252/embj.2021108443.

66. Ederle, H., Funk, C., Abou-Ajram, C., Hutten, S., Funk, E.B.E., Kehlenbach, R.H., Bailer, S.M., and Dormann, D. (2018). Nuclear egress of TDP-43 and FUS occurs independently of Exportin-1/CRM1. Sci Rep *8*, 7084. https://doi.org/10.1038/s41598-018-25007-5.

67. Wang, J., Choi, J.-M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D., et al. (2018). A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. Cell *174*, 688-699.e16. https://doi.org/10.1016/j.cell.2018.06.006.

68. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

69. Koulouras, G., Panagopoulos, A., Rapsomaniki, M.A., Giakoumakis, N.N., Taraviras, S., and Lygerou, Z. (2018). EasyFRAP-web: a web-based tool for the analysis of fluorescence recovery after photobleaching data. Nucleic Acids Res *46*, W467–W472. https://doi.org/10.1093/nar/gky508.

70. Cao, F., von Bülow, S., Tesei, G., and Lindorff-Larsen, K. (2024). A coarse-grained model for disordered and multi-domain proteins. Protein Sci *33*, e5172. https://doi.org/10.1002/pro.5172.

71. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

72. Lau, A.M., Bordin, N., Kandathil, S.M., Sillitoe, I., Waman, V.P., Wells, J., Orengo, C.A., and Jones, D.T. (2024). Exploring structural diversity across the protein universe with The Encyclopedia of Domains. Science *386*, eadq4946. https://doi.org/10.1126/science.adq4946.

73. Anderson, J.A., Glaser, J., and Glotzer, S.C. (2020). HOOMD-blue: A Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations. Computational Materials Science *173*, 109363. https://doi.org/10.1016/j.commatsci.2019.109363.

74. Zippo, E., Dormann, D., Speck, T., and Stelzl, L.S. (2025). Molecular simulations of enzymatic phosphorylation of disordered proteins and their condensates. Nat Commun *16*, 4649. https://doi.org/10.1038/s41467-025-59676-4.

75. Gersten, D.M., and Marchalonis, J.J. (1978). A rapid, novel method for the solid-phase derivatization of IgG antibodies for immune-affinity chromatography. J. Immunol. Methods *24*, 305–309. https://doi.org/10.1016/0022-1759(78)90133-3.

76. Uliana, F., Ciuffa, R., Mishra, R., Fossati, A., Frommelt, F., Keller, S., Mehnert, M., Birkeland, E.S., van Drogen, F., Srejic, N., et al. (2023). Phosphorylation-linked complex profiling identifies assemblies required for Hippo signal integration. Molecular Systems Biology *n/a*, e11024. https://doi.org/10.15252/msb.202211024.

77. Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat Methods *6*, 359–362. https://doi.org/10.1038/nmeth.1322.

78. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367–1372. https://doi.org/10.1038/nbt.1511.

79. Freibaum, B.D., Messing, J., Yang, P., Kim, H.J., and Taylor, J.P. (2021). High-fidelity reconstitution of stress granules and nucleoli in mammalian cellular lysate. Journal of Cell Biology *220*, e202009079. https://doi.org/10.1083/jcb.202009079.

80. Schreiber, K.J., Kadijk, E., and Youn, J.-Y. (2024). Exploring Options for Proximity-Dependent Biotinylation Experiments: Comparative Analysis of Labeling Enzymes and Affinity Purification Resins. J. Proteome Res. https://doi.org/10.1021/acs.jproteome.3c00908.

81. Blokhuis, A.M., Koppers, M., Groen, E.J.N., van den Heuvel, D.M.A., Dini Modigliani, S., Anink, J.J., Fumoto, K., van Diggelen, F., Snelting, A., Sodaar, P., et al. (2016). Comparative interactomics analysis of different ALS-associated proteins identifies converging molecular pathways. Acta Neuropathol *132*, 175–196. https://doi.org/10.1007/s00401-016-1575-8.

82. Oughtred, R., Stark, C., Breitkreutz, B.-J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., et al. (2019). The BioGRID interaction database: 2019 update. Nucleic Acids Res. *47*, D529–D541. https://doi.org/10.1093/nar/gky1079.

83. Helmold, B.R., Pauss, K.E., and Ozdinler, P.H. (2023). TDP-43 protein interactome informs about perturbed canonical pathways and may help develop personalized medicine approaches for patients with TDP-43 pathology. Drug Discovery Today *28*, 103769. https://doi.org/10.1016/j.drudis.2023.103769.

84. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A.L., Fang, T., Doncheva, N.T., Pyysalo, S., et al. (2023). The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res *51*, D638–D646. https://doi.org/10.1093/nar/gkac1000.

85. Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG motif. Mol Cell *50*, 613–623. https://doi.org/10.1016/j.molcel.2013.05.021.

86. You, K., Huang, Q., Yu, C., Shen, B., Sevilla, C., Shi, M., Hermjakob, H., Chen, Y., and Li, T. (2020). PhaSepDB: a database of liquid-liquid phase separation related proteins. Nucleic Acids Res *48*, D354–D359. https://doi.org/10.1093/nar/gkz847.

87. Rostam, N., Ghosh, S., Chow, C.F.W., Hadarovich, A., Landerer, C., Ghosh, R., Moon, H., Hersemann, L., Mitrea, D.M., Klein, I.A., et al. (2023). CD-CODE: crowdsourcing condensate database and encyclopedia. Nat Methods *20*, 673–676. https://doi.org/10.1038/s41592-023-01831-0.

88. Hughes, C.S., Moggridge, S., Müller, T., Sorensen, P.H., Morin, G.B., and Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. Nat Protoc *14*, 68–85. https://doi.org/10.1038/s41596-018-0082-x.

89. Kong, A.T., Leprevost, F.V., Avtonomov, D.M., Mellacheruvu, D., and Nesvizhskii, A.I. (2017). MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. Nat Methods *14*, 513–520. https://doi.org/10.1038/nmeth.4256.

90. da Veiga Leprevost, F., Haynes, S.E., Avtonomov, D.M., Chang, H.-Y., Shanmugam, A.K., Mellacheruvu, D., Kong, A.T., and Nesvizhskii, A.I. (2020). Philosopher: a versatile toolkit for shotgun proteomics data analysis. Nat Methods *17*, 869–870. https://doi.org/10.1038/s41592-020-0912-y.

91. Yu, F., Teo, G.C., Kong, A.T., Fröhlich, K., Li, G.X., Demichev, V., and Nesvizhskii, A.I. (2023). Analysis of DIA proteomics data using MSFragger-DIA and FragPipe computational platform. Nat Commun *14*, 4154. https://doi.org/10.1038/s41467-023-39869-5.

92. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res *43*, e47. https://doi.org/10.1093/nar/gkv007.

93. Quast, J.-P., Schuster, D., and Picotti, P. (2022). protti: an R package for comprehensive data analysis of peptide- and protein-centric bottom-up proteomics data. Bioinformatics Advances *2*, vbab041. https://doi.org/10.1093/bioadv/vbab041.

94. Holehouse, A.S., Das, R.K., Ahad, J.N., Richardson, M.O.G., and Pappu, R.V. (2017). CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. Biophysical Journal *112*, 16–21. https://doi.org/10.1016/j.bpj.2016.11.3200.

95. Han, Q., Zhong, Y., Quan, B., Wang, T.-Y., and Chou, T.-F. Decoding Condensate Composition Reveals Tunable Interactomes and Functions of Disordered Proteins.

96. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

97. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

98. Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res *47*, D766–D773. https://doi.org/10.1093/nar/gky955.

99. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. Gigascience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

100. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

101. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

102. Wang, Y., Xie, Z., Kutschera, E., Adams, J.I., Kadash-Edmondson, K.E., and Xing, Y. (2024). rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data. Nat Protoc *19*, 1083–1104. https://doi.org/10.1038/s41596-023-00944-2.

103. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data | BMC Bioinformatics | Springer Nature Link https://link.springer.com/article/10.1186/1471-2105-11-165.

104. Colombo, M., Karousis, E.D., Bourquin, J., Bruggmann, R., and Mühlemann, O. (2017). Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA *23*, 189–201. https://doi.org/10.1261/rna.059055.116.

105. Imamachi, N., Salam, K.A., Suzuki, Y., and Akimitsu, N. (2017). A GC-rich sequence feature in the 3' UTR directs UPF1-dependent mRNA decay in mammalian cells. Genome Res. *27*, 407–418. https://doi.org/10.1101/gr.206060.116.

106. Fritz, S.E., Ranganathan, S., Wang, C.D., and Hogg, J.R. (2022). An alternative UPF1 isoform drives conditional remodeling of nonsense-mediated mRNA decay. EMBO J *41*, EMBJ2021108898. https://doi.org/10.15252/embj.2021108898.

107. Chawade, A., Alexandersson, E., and Levander, F. (2014). Normalyzer: a tool for rapid evaluation of normalization methods for omics data sets. J Proteome Res *13*, 3114–3120. https://doi.org/10.1021/pr401264n.

108. Kang, J., Tang, Q., He, J., Li, L., Yang, N., Yu, S., Wang, M., Zhang, Y., Lin, J., Cui, T., et al. (2022). RNAInter v4.0: RNA interactome repository with redefined confidence scoring system and improved accessibility. Nucleic Acids Res *50*, D326–D332. https://doi.org/10.1093/nar/gkab997.

109. Schmidt, H.B., and Rohatgi, R. (2016). In Vivo Formation of Vacuolated Multi-phase Compartments Lacking Membranes. Cell Reports *16*, 1228–1236. https://doi.org/10.1016/j.celrep.2016.06.088.

110. Buratti, E. (2015). Functional Significance of TDP-43 Mutations in Disease. In Advances in Genetics (Elsevier), pp. 1–53. https://doi.org/10.1016/bs.adgen.2015.07.001.

111. Wang, A., Conicella, A.E., Schmidt, H.B., Martin, E.W., Rhoads, S.N., Reeb, A.N., Nourse, A., Ramirez Montero, D., Ryan, V.H., Rohatgi, R., et al. (2018). A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. The EMBO Journal *37*, e97452. https://doi.org/10.15252/embj.201797452.

112. Tsuiji, H., Iguchi, Y., Furuya, A., Kataoka, A., Hatsuta, H., Atsuta, N., Tanaka, F., Hashizume, Y., Akatsu, H., Murayama, S., et al. (2013). Spliceosome integrity is defective in the motor neuron diseases ALS and SMA. EMBO Mol Med *5*, 221–234. https://doi.org/10.1002/emmm.201202303.

113. Wang, C., Duan, Y., Duan, G., Wang, Q., Zhang, K., Deng, X., Qian, B., Gu, J., Ma, Z., Zhang, S., et al. (2020). Stress Induces Dynamic, Cytotoxicity-Antagonizing TDP-43 Nuclear Bodies via Paraspeckle LncRNA NEAT1-

Mediated Liquid-Liquid Phase Separation. Molecular Cell *79*, 443-458.e7. https://doi.org/10.1016/j.molcel.2020.06.019.

114. Lang, R., Hodgson, R.E., and Shelkovnikova, T.A. (2024). TDP-43 in nuclear condensates: where, how, and why. Biochem Soc Trans *52*, 1809–1825. https://doi.org/10.1042/BST20231447.

115. Gstaiger, M., and Aebersold, R. (2013). Genotype–phenotype relationships in light of a modular protein interaction landscape. Mol. BioSyst. *9*, 1064–1067. https://doi.org/10.1039/C3MB25583B.

116. Hartwell, L.H., Hopfield, J.J., Leibler, S., and Murray, A.W. (1999). From molecular to modular cell biology. Nature *402*, C47–C52. https://doi.org/10.1038/35011540.

117. Tziortzouda, P., Van Den Bosch, L., and Hirth, F. (2021). Triad of TDP43 control in neurodegeneration: autoregulation, localization and aggregation. Nat Rev Neurosci *22*, 197–208. https://doi.org/10.1038/s41583-021-00431-1.

118. Buratti, E., and Baralle, F.E. (2010). The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation. RNA Biol *7*, 420–429. https://doi.org/10.4161/rna.7.4.12205.

119. Polymenidou, M., Lagier-Tourenne, C., Hutt, K.R., Huelga, S.C., Moran, J., Liang, T.Y., Ling, S.-C., Sun, E., Wancewicz, E., Mazur, C., et al. (2011). Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nat Neurosci *14*, 459–468. https://doi.org/10.1038/nn.2779.

120. Changiarath, A., Flores-Solis, D., Michels, J.J., Rodriguez, R.H., Hanson, S.M., Schmid, F., Zweckstetter, M., Padeken, J., and Stelzl, L.S. (2024). Promoter and Gene-Body RNA-Polymerase II co-exist in partial demixed condensates. Preprint at bioRxiv, https://doi.org/10.1101/2024.03.16.585180 https://doi.org/10.1101/2024.03.16.585180.

121. Rekhi, S., Garcia, C.G., Barai, M., Rizuan, A., Schuster, B.S., Kiick, K.L., and Mittal, J. (2024). Expanding the molecular language of protein liquid-liquid phase separation. Nat Chem *16*, 1113–1124. https://doi.org/10.1038/s41557-024-01489-x.

122. Davey, N.E., Van Roey, K., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H., and Gibson, T.J. (2012). Attributes of short linear motifs. Mol Biosyst *8*, 268–281. https://doi.org/10.1039/c1mb05231d.

123. Pawson, T., and Nash, P. (2003). Assembly of cell regulatory systems through protein interaction domains. Science *300*, 445–452. https://doi.org/10.1126/science.1083653.

124. Chowdhury, M.N., and Jin, H. (2023). The RGG motif proteins: Interactions, functions, and regulations. Wiley Interdiscip Rev RNA *14*, e1748. https://doi.org/10.1002/wrna.1748.

125. Chong, P.A., Vernon, R.M., and Forman-Kay, J.D. (2018). RGG/RG Motif Regions in RNA Binding and Phase Separation. J Mol Biol *430*, 4650–4665. https://doi.org/10.1016/j.jmb.2018.06.014.

126. Xie, L., Zhu, Y., Hurtle, B.T., Wright, M., Robinson, J.L., Mauna, J.C., Brown, E.E., Ngo, M., Bergmann, C.A., Xu, J., et al. (2025). Context-dependent Interactors Regulate TDP-43 Dysfunction in ALS/FTLD. Preprint at bioRxiv, https://doi.org/10.1101/2025.04.07.646890 https://doi.org/10.1101/2025.04.07.646890.

127. Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Župunski, V., et al. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat Neurosci *14*, 452–458. https://doi.org/10.1038/nn.2778.

128. Buratti, E., Dörk, T., Zuccato, E., Pagani, F., Romano, M., and Baralle, F.E. (2001). Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. The EMBO Journal *20*, 1774–1784. https://doi.org/10.1093/emboj/20.7.1774.

129. Fiesel, F.C., Weber, S.S., Supper, J., Zell, A., and Kahle, P.J. (2012). TDP-43 regulates global translational yield by splicing of exon junction complex component SKAR. Nucleic Acids Research *40*, 2668–2682. https://doi.org/10.1093/nar/gkr1082.

130. Ling, J.P., Pletnikova, O., Troncoso, J.C., and Wong, P.C. (2015). TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. Science *349*, 650–655. https://doi.org/10.1126/science.aab0983.

131. Irwin, K.E., Jasin, P., Braunstein, K.E., Sinha, I.R., Garret, M.A., Bowden, K.D., Chang, K., Troncoso, J.C., Moghekar, A., Oh, E.S., et al. (2024). A fluid biomarker reveals loss of TDP-43 splicing repression in presymptomatic ALS–FTD. Nat Med *30*, 382–393. https://doi.org/10.1038/s41591-023-02788-5.

132. Torres, P., Rico-Rios, S., Ceron-Codorniu, M., Santacreu-Vilaseca, M., Seoane-Miraz, D., Jad, Y., Ayala, V., Mariño, G., Beltran, M., Miralles, M.P., et al. (2024). TDP-43 regulates LC3ylation in neural tissue through ATG4B cryptic splicing inhibition. Acta Neuropathol *148*, 45. https://doi.org/10.1007/s00401-024-02780-4.

133. De Conti, L., Akinyi, M.V., Mendoza-Maldonado, R., Romano, M., Baralle, M., and Buratti, E. (2015). TDP-43 affects splicing profiles and isoform production of genes involved in the apoptotic and mitotic cellular pathways.

Nucleic Acids Res *43*, 8990–9005. https://doi.org/10.1093/nar/gkv814.

134. Zhang, Y., Zhou, L., and Miller, C.A. (1998). A splicing variant of a death domain protein that is regulated by a mitogen-activated kinase is a substrate for c-Jun N-terminal kinase in the human central nervous system. Proc Natl Acad Sci U S A *95*, 2586–2591. https://doi.org/10.1073/pnas.95.5.2586.

135. Roczniak-Ferguson, A., and Ferguson, S.M. (2019). Pleiotropic requirements for human TDP-43 in the regulation of cell and organelle homeostasis. Life Science Alliance *2*. https://doi.org/10.26508/lsa.201900358.

136. Cortese, A., Laurà, M., Casali, C., Nishino, I., Hayashi, Y.K., Magri, S., Taroni, F., Stuani, C., Saveri, P., Moggio, M., et al. (2018). Altered TDP-43-dependent splicing in HSPB8-related distal hereditary motor neuropathy and myofibrillar myopathy. European Journal of Neurology *25*, 154–163. https://doi.org/10.1111/ene.13478.

137. Shiga, A., Ishihara, T., Miyashita, A., Kuwabara, M., Kato, T., Watanabe, N., Yamahira, A., Kondo, C., Yokoseki, A., Takahashi, M., et al. (2012). Alteration of POLDIP3 Splicing Associated with Loss of Function of TDP-43 in Tissues Affected with ALS. PLOS ONE *7*, e43120. https://doi.org/10.1371/journal.pone.0043120.

138. Seddighi, S., Qi, Y.A., Brown, A.-L., Wilkins, O.G., Bereda, C., Belair, C., Zhang, Y.-J., Prudencio, M., Keuss, M.J., Khandeshi, A., et al. (2024). Mis-spliced transcripts generate de novo proteins in TDP-43-related ALS/FTD. Science Translational Medicine *0*, eadg7162. https://doi.org/10.1126/scitranslmed.adg7162.

139. Cho, N.H., Cheveralls, K.C., Brunner, A.-D., Kim, K., Michaelis, A.C., Raghavan, P., Kobayashi, H., Savy, L., Li, J.Y., Canaj, H., et al. (2022). OpenCell: Endogenous tagging for the cartography of human cellular organization. Science *375*, eabi6983. https://doi.org/10.1126/science.abi6983.

140. Kozareva, V., Liu, Z., Blake, K., Qi, Y.A., Rollinson, S., Seddighi, S., Alsaidi, M., Tsitkov, S., Prudencio, M., Petrucelli, L., et al. Integrative multi-omic analysis links TDP-43-driven splicing defects to cascading proteomic disruption of ALS/FTD pathways.

141. Ayala, Y.M., De Conti, L., Avendaño-Vázquez, S.E., Dhir, A., Romano, M., D'Ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E., et al. (2011). TDP-43 regulates its mRNA levels through a negative feedback loop. EMBO J *30*, 277–288. https://doi.org/10.1038/emboj.2010.310.

142. Combe, P., Subecz, C., Le Goff, G., Plamont, M.A., Bohl, D., and Gueroui, Z. Concentration-dependent cytoplasmic phase separation of TDP-43 drives aggregation and proteinopathy. The FEBS Journal *n/a*. https://doi.org/10.1111/febs.70429.

143. Mann, J.R., Gleixner, A.M., Mauna, J.C., Gomes, E., DeChellis-Marks, M.R., Needham, P.G., Copley, K.E., Hurtle, B., Portz, B., Pyles, N.J., et al. (2019). RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. Neuron *102*, 321-338.e8. https://doi.org/10.1016/j.neuron.2019.01.048.

144. Darling, A.L., and Shorter, J. (2021). Combating deleterious phase transitions in neurodegenerative disease. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1868*, 118984. https://doi.org/10.1016/j.bbamcr.2021.118984.

145. Guo, L., Mann, J.R., Mauna, J.C., Copley, K.E., Wang, H., Rubien, J.D., Bergmann, C.A., Carey, J.L., Merjane, J., Ngo, M., et al. (2026). Defining RNA oligonucleotides that reverse deleterious phase transitions of RNA-binding proteins with prion-like domains. Mol Cell *86*, 114-134.e10. https://doi.org/10.1016/j.molcel.2025.12.009.

146. Rizuan, A., Shenoy, J., Mohanty, P., dos Passos, P.M., Mercado Ortiz, J.F., Bai, L., Viswanathan, R., Zaborowksy, J., Wang, S.-H., Johnson, V., et al. (2025). Structural details of helix-mediated multimerization of the conserved region of TDP-43 C-terminal domain. Nat Commun *16*, 10528. https://doi.org/10.1038/s41467-025-65546-w.

147. Hedtfeld, M., Dammers, A., Koerner, C., and Musacchio, A. (2024). A validation strategy to assess the role of phase separation as a determinant of macromolecular localization. Molecular Cell *84*, 1783-1801.e7. https://doi.org/10.1016/j.molcel.2024.03.022.

148. Pei, G., Lyons, H., Li, P., and Sabari, B.R. (2025). Transcription regulation by biomolecular condensates. Nat Rev Mol Cell Biol *26*, 213–236. https://doi.org/10.1038/s41580-024-00789-x.

149. Lyons, H., Veettil, R.T., Pradhan, P., Fornero, C., De La Cruz, N., Ito, K., Eppert, M., Roeder, R.G., and Sabari, B.R. (2023). Functional partitioning of transcriptional regulators by patterned charge blocks. Cell *186*, 327-345.e28. https://doi.org/10.1016/j.cell.2022.12.013.

150. Zhong, W., Scialò, C., Gatta, B., Häfliger, M., Leu, N., Lurati, F., Peter, M., Ramesh, N., Roschitzki, B., Jagannath, S., et al. (2025). Lysosomal escape and TMEM106B fibrillar core determine TDP-43 seeding outcomes. Preprint at bioRxiv, https://doi.org/10.64898/2025.12.19.695531 https://doi.org/10.64898/2025.12.19.695531.

151. Lu, S., Hu, J., Arogundade, O.A., Goginashvili, A., Vazquez-Sanchez, S., Diedrich, J.K., Gu, J., Blum, J., Oung, S., Ye, Q., et al. (2022). Heat-shock chaperone HSPB1

regulates cytoplasmic TDP-43 phase separation and liquid-to-gel transition. Nat Cell Biol *24*, 1378–1393. https://doi.org/10.1038/s41556-022-00988-8.

152. Yu, H., Lu, S., Gasior, K., Singh, D., Vazquez-Sanchez, S., Tapia, O., Toprani, D., Beccari, M.S., Yates, J.R., Da Cruz, S., et al. (2021). HSP70 chaperones RNA-free TDP-43 into anisotropic intranuclear liquid spherical shells. Science *371*, eabb4309. https://doi.org/10.1126/science.abb4309.

153. Lu, S., Zhang, S., Oung, S., Diedrich, J.K., Han, P., Arnold-Garcia, O., Ohkubo, T., Arogundade, O.A., Vazquez-Sanchez, S., Zhang, K., et al. (2025). TDP-43 skein-like inclusions are formed by BAG3- and HSP70-guided co-aggregation with actin-binding proteins. Nat Cell Biol, 1–13. https://doi.org/10.1038/s41556-025-01789-5.

154. Bludau, I., and Aebersold, R. (2020). Proteomic and interactomic insights into the molecular basis of cell functional diversity. Nat Rev Mol Cell Biol *21*, 327–340. https://doi.org/10.1038/s41580-020-0231-2.

155. Ciuffa, R., Uliana, F., Mannion, J., Mehnert, M., Tenev, T., Marulli, C., Satanowski, A., Keller, L.M.L., Rodilla Ramírez, P.N., Ori, A., et al. (2022). Novel biochemical, structural, and systems insights into inflammatory signaling revealed by contextual interaction proteomics. Proceedings of the National Academy of Sciences *119*, e2117175119. https://doi.org/10.1073/pnas.2117175119.

156. Barmada, S.J., Ju, S., Arjun, A., Batarse, A., Archbold, H.C., Peisach, D., Li, X., Zhang, Y., Tank, E.M.H., Qiu, H., et al. (2015). Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. Proceedings of the National Academy of Sciences *112*, 7821–7826. https://doi.org/10.1073/pnas.1509744112.

157. Gomez, N., Hsieh, C., Li, X., Dykstra, M., Waksmacki, J., Altheim, C., Bechar, Y., Klim, J., Zaepfel, B., Rothstein, J., et al. (2024). Counter-regulation of RNA stability by UPF1 and TDP43. Preprint at bioRxiv, https://doi.org/10.1101/2024.01.31.578310 https://doi.org/10.1101/2024.01.31.578310.

158. Olivieri, J.E., Dehghannasiri, R., Wang, P.L., Jang, S., de Morree, A., Tan, S.Y., Ming, J., Ruohao Wu, A., Tabula Sapiens Consortium, Quake, S.R., et al. (2021). RNA splicing programs define tissue compartments and cell types at single-cell resolution. eLife *10*, e70692. https://doi.org/10.7554/eLife.70692.

159. Huang, L., Lou, C.-H., Chan, W., Shum, E.Y., Shao, A., Stone, E., Karam, R., Song, H.-W., and Wilkinson, M.F. (2011). RNA Homeostasis Governed by Cell Type-Specific and Branched Feedback Loops Acting on NMD. Molecular Cell *43*, 950–961. https://doi.org/10.1016/j.molcel.2011.06.031.

160. Laman Trip, D.S., van Oostrum, M., Memon, D., Frommelt, F., Baptista, D., Panneerselvam, K., Bradley, G., Licata, L., Hermjakob, H., Orchard, S., et al. (2025). A tissue-specific atlas of protein–protein associations enables prioritization of candidate disease genes. Nat Biotechnol, 1–14. https://doi.org/10.1038/s41587-025-02659-z.