TDP-43 α-helical structure tunes liquid-liquid phase separation and function

Alexander E. Conicella[1,2,#], Gregory L. Dignon[3,#], Gül H. Zerze[3], Hermann Broder Schmidt[4], Alexandra M. D'Ordine[2], Young C. Kim[5], Rajat Rohatgi[4,6], Yuna M. Ayala[7], Jeetain Mittal[3,*], and Nicolas L. Fawzi[1,*]

Abstract

Liquid-liquid phase separation (LLPS) is involved in the formation of membraneless organelles (MLOs) associated with RNA processing. Present in several MLOs, TDP-43 undergoes LLPS and is linked to the pathogenesis of amyotrophic lateral sclerosis (ALS). While some disease variants of TDP-43 disrupt self-interaction and function, here we show that designed single mutations can enhance TDP-43 assembly and function via modulating helical structure. Using molecular simulation and NMR spectroscopy, we observe large structural changes in a dimeric TDP-43. Two conserved glycine residues (G335 and G338) are potent inhibitors of helical extension and helix-helix interaction, which are removed in part by variants including the ALS-associated G335D. Substitution to helix-enhancing alanine at either of these positions dramatically enhances phase separation *in vitro* and decreases fluidity of phase separated TDP-43 reporter compartments in cells. Furthermore, G335A increases TDP-43 splicing function in a mini-gene assay. Therefore, TDP-43 helical region serves as a short but uniquely tunable module that shows promise as for controlling assembly and function in cellular and synthetic biology applications of LLPS.

[#]Co-first author
[1]Department of Molecular Pharmacology, Physiology, and Biotechnology, Brown University, Providence, RI
[2]Graduate Program in Molecular Biology, Cell Biology, and Biochemistry, Providence, RI
[3]Department of Chemical and Biomolecular Engineering, Lehigh University, Bethlehem, PA
[4]Department of Biochemistry, Stanford University School of Medicine, Stanford, CA
[5]Center for Materials Physics and Technology, Naval Research Laboratory, Washington, DC 20375
[6]Department of Medicine, Stanford University School of Medicine, Stanford, CA
[7]Edward Doisy Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, MO
[*]Corresponding Authors

Introduction

RNA-binding proteins harboring low-complexity sequences are major constituents of membraneless organelles, particularly ribonucleoprotein (RNP) granules[1-3]. These granules can display liquid-like properties[4], consistent with the view that they assemble through liquid-liquid phase separation (LLPS)[5]. The molecular interactions and physical forces by which RNP granule proteins assemble into liquid-like compartments has been of great interest[6-10]. Multivalent interactions based on several interaction motifs can give rise to LLPS, including globular domains (often in modular chains) binding linear protein or nucleic acid motifs, as well as weak multivalent interactions between polar, aromatic, and charged-patterned disordered protein segments[11,12]. The diversity of interaction types mediating LLPS could in principle allow for different responses to stimuli, as well as variations in permeability, dynamics, and persistence, thus allowing comparable diversity in their respective biological functions[13-15].

TDP-43, a well-known RNP granule component[16-19], is a heterogeneous ribonucleoprotein (hnRNP) family member with genetic and histopathological links to amyotrophic lateral sclerosis (ALS)[20], frontotemporal dementia (FTD)[21], and Alzheimer's disease (AD)[22]. The tandem RRM domains recognize UG-rich sequences found near splice sites[23], contributing to TDP-43's role in regulating splicing of transcripts in mammalian cells[24,25]. Yet, the splicing function of TDP-43 also depends on the TDP-43 globular N-terminal domain[26-30] and its predominantly intrinsically disordered C-terminal domain (CTD, residues 267-414)[31]. One may wonder what is the role of these domains? TDP-43 CTD, rich in polar and aromatic residues, is prone to self-assembly and confers TDP-43 the ability to undergo LLPS both *in vitro*[32] and in cells[33]. Though LLPS requires the polar CTD, LLPS and function of TDP-43 is further enhanced by intermolecular interactions mediated by the folded N-terminal domain and by an evolutionarily conserved partially α-helical subregion of the CTD spanning residues 320-343. This region is required for LLPS of TDP-43 reporter constructs[33] and important for splicing in a well-established mini gene assay[31]. The importance of the conserved region to TDP-43 function is underscored by the presence of 7 ALS-associated variants in this short stretch[34].

Previously we showed that the conserved TDP-43 CTD region mediates intermolecular α-helical contacts enhancing TDP-43 phase separation[32]. In particular, a transiently helical region spanning 321-330[35] is stabilized and extends into the adjacent 331-343 region upon self-association[32]. Importantly, we found that several ALS-associated missense mutations in this region (including Q331K and M337V) disrupt helix-helix interaction, impairing phase separation[32]. Intriguingly, the adjacent 331-343 region that becomes partially helical upon assembly contains two nearby conserved glycine positions, G335 and G338. Because inclusion of designed helical regions have been recently demonstrated to enhance protein phase separation[36], we suspected that these "helix-discouraging" residues may serve as important sites regulating TDP-43 assembly. Therefore we set out to test if TDP-43 helical structure and helix-helix interaction can be tuned to enhance phase separation and increase TDP-43 function by α-helix enhancing mutations at conserved glycine positions.

In this work, we use experiments and simulations to characterize the influence of these two nearby glycine positions on TDP-43 CTD, structure, self-association, phase separation *in vitro* and in cells, and splicing function. We also test the effect of the G335D ALS-associated mutation occurring at one of these sites to understand a potential mechanism leading to disease[37]. Here we pair an extensive experimental and computational structural and biophysical characterization of TDP-43 CTD including a number of mutants targeted at these two glycine residues with an evaluation of the role of these positions on TDP-43 LLPS and splicing function in cells. The combination of high-resolution NMR spectroscopic and computational structural data, phase separation approaches *in vitro* and in cells, and functional data provide an unprecedented and comprehensive atomically detailed picture of the tunability of phase separation, connecting fine control of the sequence, to structure, to interactions and finally to both phase separation and function.

Results

Helical structure of TDP-43 CTD enhanced by dimerization

Our previous work showed that the conserved region (CR) in TDP-43 CTD (residues 320-343) encompasses a region from 321-330 that is partially helical and that helicity is enhanced in 321-330 and spreads to 331-343 upon TDP-43 self-interaction[32]. Yet, the precise extent of increased α-helicity and structural change in the assembled state remains poorly understood, despite the importance of the CR in TDP-43 splicing function[31] and phase separation[32,33]. Because we could not trap the assembled state, we previously relied on experiments sensitive to the small population of assembled state (chemical shift deviation vs. concentration, and relaxation dispersion NMR) present in equilibrium with the monomeric state. Signs of enhanced helicity were present, but direct characterization of the size of, and extent of helicity in, the assembled form could not be determined[32]. Here, to directly visualize the structural changes along the self-assembly pathway, we first used all-atom explicit solvent simulations. Using advanced sampling techniques (see Methods), we determined the equilibrium ensemble of configurations of two copies of a computationally-tractable segment of TDP-43 CTD, TDP-43$_{310-350}$, containing the entire helical region that we previously characterized and validated[32] placed together in a simulation box. First, we found that overall *α*-helical content for the entire two-chain ensemble (i.e. including configurations with and without intermolecular contacts) is greater (up to 60% helical from residues 321-331, Figure S1) than that of the single chain ensemble (up to 50% helical) from our previous work. Next, to search for stable configurations populated in the two chain ensemble, we computed a free energy surface quantifying the relative energetic favorability of different configurations binned along two order parameters selected to represent intermolecular assembly: the number of hydrogen bonds and number of hydrophobic contacts between the two protein chains (Figure 1A). Two distinct minima (ΔG values <10 kJ/mol relative to most populated reference bin), which we term strongly and weakly bound states, are observed. Averaging over only the configurations corresponding to the strongly bound free energy well (i.e. configurations in favorable intermolecular contact), we see a larger increase in *α*-helical content (Figure 1B) and stabilization (~25% population) of a continuous helix spanning from residues 320-334, abruptly ending at residue G335 (Figure 1C).

The simulations suggest heterogeneity in the bound ensemble. Many different residue pairs in the region spanning 320-343 form intermolecular contacts (Figure S1A). The strongly bound ensemble also comprises heterogeneous binding orientations including parallel and antiparallel helix-helix interactions as well as interactions involving disordered conformations of the 320-343 region (Figure S1B,C). Though two chains may not be sufficient to explore the full ensemble of bound configurations (i.e TDP-43 may form trimeric or higher order helical assemblies), these data demonstrate how self-interaction can promote helical stabilization and are consistent with our previous NMR paramagnetic relaxation enhancement data which suggested that TDP-43 CTD 320-343 region visits multiple bound orientations[32]. In summary, stabilization of helical elements upon binding suggests a self-assembly mechanism involving both "conformational selection" and "induced fit" as proposed previously for other helix-based assemblies[38].

Our previous work demonstrated that even in the absence of phase separation, the conserved region of TDP-43 CTD self-assembles mediated by helix-helix contacts into a multimer state invisible to standard NMR approaches[32]. To experimentally trap a dimer state, we instead created a novel approach, mimicking the two chain simulations by forming obligate CTD dimers *in vitro* by covalently crosslinking two CTDs far (>40 residues) from the conserved 320-343 region (Figure 1C). We created two distinct disulfide cross-linked dimers, oxidized by copper (II) phenanthroline, at single cysteines engineered at residues 273 or 387 (Figure 1C). After oxidation, these CTD variants appear at twice the apparent molecular weight compared to reduced control samples in non-reducing SDS-PAGE (Figure 1D). Consistent with stabilization of a structured assembly, fingerprint two dimensional (^{1}H-^{15}N TROSY) spectra of both S273C dimers (10 μM, in monomer units) and S387C dimers (20 μM) reveal shifted, very weak resonances that can be assigned to the 320-343 region by extrapolation of resonances of the WT TDP-43 CTD as a function of increasing concentration (i.e. increasing population of an assembled state) (Figure 1E). This observation strongly suggests that the cross-link does not bias the helix-helix contacts into a particular geometry (i.e. parallel assembly due to symmetric cross-link site), possible due to the long (40 residue) region between the cross-link site and the helical region. Quantification of ^{15}N chemical shift differences between dimeric (oxidized) and monomeric (reduced) CTD revealed upfield chemical shifts from residues 315 to 343, primarily in 321-335, as well as at the site of the crosslink (Figure 1F). Importantly, both cysteine variants display the same pattern of upfield shifts in the 321-340 region (RMSD ~ 0.07 ppm), indicating that the chemical environment and hence the dimeric states sampled by the 321-340 region are independent of crosslink position. Furthermore, the magnitude of the ^{15}N chemical shift differences observed here are similar to those we derived for the assembled state of TDP-43 CTD based on experimental CPMG relaxation dispersion NMR experiments[32]. Interestingly, these obligate dimer samples were not stable at higher concentrations, precluding more extensive NMR characterization. Taken together, the computer simulations and experiments show signatures of extensive enhanced helicity upon dimerization focused in the region from residue 321 to 335.

Mutations at glycine positions enhance CTD *α*-helical structure in the 330-343 region

Given that TDP-43 conserved region helicity seems to be disrupted in the region of G335 and G338, we wondered if we replaced G335 or G338 with other residues, would we increase helicity? To test the hypothesis that substitution of residues that are less helix-discouraging would enhance TDP-43 CTD helicity, we designed several TDP-43 CTD variants at G335 and G338 positions to observe the effect of these amino acids on TDP-43 CTD helicity. First, we conducted single chain all-atom molecular dynamics simulations of the same 41-residue fragment, TDP-43$_{310-350}$, with a set of single point mutations, changing G335 or G338 from glycine to the helix-promoting residue alanine (G335A and G338A). Mutating glycine to alanine (G335A and G338A) results in a significant increase of *α*-helical content in the simulations, particularly in the vicinity of the mutation sites (Figure 2A). The increase in helical content is highest within residues 331-343 (Figure 2B, right and Figure S2A,B), while helical content in residues in the preceding partially helical region 321-330 does not increase (Figure 2B, left). G335D, an ALS-associated variant, has similarly enhanced helicity from 331-343, as do engineered variants G335N (similar to G335D but with the uncharged amide analogous residue) and G335S, suggesting that the enhancements in helicity caused by G335A are due to the removal of the unique backbone conformational flexibility of glycine, not specific side-chain mediated contacts. Quantifying the helical enhancement, the maximum length of contiguous *α*-helical conformations[39] increases for G335A (i.e. helices spanning 320-337) and the population of longer helical conformations increases (Figure 2C). For G338A, long helices are populated in the region including position 338 that are not observed for WT.

With strong computational support for enhanced helicity by G335 and G338 point variants, we next characterized the secondary structure of TDP-43 CTD variants experimentally by comparing NMR chemical shifts of WT and mutant TDP-43 CTD variants. Positive values of the difference of the residual ^{13}Cα and ^{13}Cβ chemical shifts compared to a random coil reference database, $\Delta\delta C^{\alpha} - \Delta\delta C^{\beta}$, (Figure 2D) indicate population of α-helical structure. Relative to WT, all of the G335 and G338 point mutants show positive increases in secondary shifts, $\Delta\Delta\delta$ (Figure 2D and Figure S2C) consistent with the enhanced helical content predicted by simulation. As suggested by the simulations, the pattern of local *α*-helical enhancement is distinct for G335A and G338A; enhancements are observed from 330-339 for G335A and from 335-343 for G338A (Figure 2D). Importantly, the values of $\Delta\delta C^{\alpha} - \Delta\delta C^{\beta}$ near the site of mutation nearly reach the level seen in the 321-330 region, which is approximately 50% helical [32]. To compare the helical enhancement between all variants, we computed the average secondary shift across residues 331-343, ($\overline{\Delta\delta C^{\alpha} - \Delta\delta C^{\beta}}$) (Figure 2E). Average secondary shift values for residues 331-343 of WT and mutant CTD are strongly correlated (r = -0.972) with the predicted change in free energy of helix stabilization relative to glycine[40] (Figure S2D). To further probe for enhanced helical

structure in TDP-43 CTD, we measured NMR relaxation parameters sensitive to reorientational motion at each position, which are slower for structured conformations than for disordered conformations. Sensitive to motions on the <1 ns timescale, heteronuclear nuclear Overhauser effect values, {^{1}H}-^{15}N NOE, for G335A and G338A are slightly higher than WT in the region from 330-343 (Figure 2F), consistent with slower motions due to higher population of structured, slow moving conformations. Similarly, the value of the transverse relaxation rate constant, ^{15}N R_2, is higher for G335A and G338A, consistent with slower motions. Because ^{15}N R_2 is also enhanced by transient helix-helix contacts in TDP-43 (Figure S2E, see below) we also measured the cross-correlated relaxation rate, η_{xy}, which is independent of these exchange effects. Like the {^{1}H}-^{15}N NOE, η_{xy} is enhanced from 331-343 in G335A and G338A, consistent with slower motion in the 331-343 region. Differences in ^{15}N spin relaxation parameters for WT, G335A, G338A are negligible across the remainder of the CTD, suggesting that the variants do not change overall folding of the monomeric CTD (e.g. they do not promote the formation of new tertiary structure such as intramolecular helix bundling). Taken together, these data strongly suggest that single point variants at conserved glycine positions in TDP-43 331-343 enhance TDP-43 CTD helicity with a magnitude commensurate with their predicted ability to stabilize α-helix structure.

G335 and G338 mutations enhance intermolecular helix-helix contacts and higher-order assembly

In our previous work, we showed that ALS-associated variants Q331K, M337V, A321G, and A321V disrupt intermolecular helix-helix assembly of TDP-43 CTD[32]. Here, we wondered if TDP-43 CTD variants that enhance helicity in 331-343 region increase TDP-43 CTD assembly. For this purpose, we used the same approach, measuring the concentration-dependent perturbations of NMR resonances in fingerprint (^{1}H-^{15}N HSQC) spectra for G335 and G338 variants at concentrations ranging from 10 μM to 90 μM at conditions where phase separation does not occur (i.e. 0 mM NaCl). We then calculated chemical shift perturbation ($\Delta\delta$) at each concentration relative to a monomeric reference for each variant (i.e. the concentration below which we do not detect significant chemical shift differences = 20 μM for all variants except for G338A = 10 μM; Figure 3A). As shown previously[32], WT CTD displays upfield ^{15}N $\Delta\delta$ in the 321-340 region above 20 μM, indicative of local intermolecular interactions. G335A and G338A show ^{15}N $\Delta\delta$ at the same residue positions as WT but are significantly enhanced at 30 μM compared to WT. Above 30 μM, many of the resonances in G335A and G338A are broadened beyond detection (Figure 3A), consistent with enhanced intermolecular interactions, compared to WT which shows resonance loss due to broadening only above 90 μM.

To quantify the effects of G335 and G338 variants on intermolecular self-assembly in the 321-343 region, we used linear regression to normalize the relative magnitude of ^{15}N $\Delta\delta$ for WT and mutant CTD variants at each concentration to the ^{15}N $\Delta\delta$ observed for WT for 45 μM - 20 μM, which we termed CSP$_{norm}$ (Figure 3B). CSP$_{norm}$ is higher for all G335 and G338 variants relative to WT, suggesting that

these mutations enhance intermolecular interactions in the 321-343 region. G338A displays the largest increase in CSP_{norm} values followed by G335A and G335D, and weaker enhancements caused by G335S and G335N. As we previously observed for WT[32], CSP_{norm} values for all variants increase approximately linearly as a function of increasing CTD concentration, suggesting that the population of assembled state is not saturated at these concentrations. Importantly, the enhancement of helix-helix assembly due to each variant follows the same order as the enhancement in helicity of the monomeric CTD; the only exception is G335D whose lower net charge likely favors assembly in these low salt conditions[41]. Together these observations suggest that higher helicity of TDP-43 CTD monomer enhances helix-helix assembly by increasing the effective affinity.

High concentration samples that contain significant populations of the assembled form, however, show significant line broadening in the region of the helix for all variants, precluding our ability to probe saturation of binding and to determine the dissociation constant for assembly. Instead, we took advantage of the remaining resonances far from the assembly site and still visible at high concentration to probe the average hydrodynamic size as a function of concentration. To this end, we measured translational diffusion coefficients for WT CTD at sample concentrations ranging from 10 μM to 600 μM. Intensity ratios for all concentrations decay as a single exponential as a function of squared gradient strength (Figure S3), suggesting that the translational diffusion of the CTD at each concentration can be best captured by a single diffusion coefficient (D), reflecting the population-weighted average diffusion rates for monomeric and oligomeric states consistent with dynamic, reversible assembly[42]. At low CTD concentrations (50 μM and below) the apparent diffusion coefficient for WT CTD remains largely unchanged (8.8± 0.1 · 10^{-11} m^{2}s^{-1}) (Figure 3C), suggesting that this value reflects the diffusion coefficient of monomeric CTD. This diffusion coefficient corresponds to an approximate radius of gyration of 2.8 nm in dilute aqueous solution, close to that of the similar-length low complexity (LC) domain of hnRNPA2[43] and significantly more extended than the globular reference protein lysozyme (2.03 nm) (Figure S3A), which is of almost identical molecular weight to TDP-43 CTD. At concentrations above 50 μM, however, the apparent diffusion coefficient of the CTD sample decreases, consistent with CTD assembly into higher order species. These diffusion data alone, however, do not uniquely determine the oligomerization state. To provide a reference for the diffusion of a TDP-43 dimer, we also measured the translational diffusion coefficients of the monomeric and covalent dimeric forms of the S273C and S387C CTD variants at low (5 μM) concentrations (Figure 3C, Figure S3B). Diffusion coefficients are similar for both cross-linked dimers. The diffusion coefficient of both reduced monomeric CTD (8.7 ± 0.1 · 10^{-11} m^{2}s^{-1}) is effectively the same as the wild-type whereas the covalent dimer is smaller (6.4 ± 0.1 · 10^{-11} m^{2}s^{-1}). As concentrations exceed 300 μM, the change in diffusion coefficient of wild-type CTD decreases, suggesting saturation of assembly. The diffusion coefficient at high concentrations is also smaller than that of the covalently linked dimer, suggesting that oligomers larger than dimers are present. For the G335A variant, oligomerization occurs at much lower concentration and the apparent diffusion coefficient also plateaus at high concentration, implying

saturation of a similarly-sized state. Hence, G335A self-associates at lower concentrations, implying tighter binding. To demonstrate that higher order structures, not increased nonspecific contacts or increased viscosity due to increased protein concentration, give rise to the slower diffusion coefficients observed, we tracked the diffusion of the helix-disrupting A326P variant[32] that does not assemble (Figure 3C). This variant shows only modestly slower diffusion even at high concentration, and the profile lacks a sigmoidal shape, suggesting that the sigmoidal shape for WT and G335A arises due to assembly in the concentration range of the transition. Taken together, these data suggest TDP-43 CTD assembles into states larger than dimers and that such an assembly happens at lower concentrations for the interaction-enhancing G335A variant.

G335 and G338 variants alter *in vitro* LLPS

Given the relationship between TDP-43 CTD intermolecular self-assembly and LLPS, we next tested if G335 and G338 variants that increase interaction also enhance phase separation. Although the conserved region mediates helix-helix contacts at all conditions, TDP-43 CTD does not readily phase separate in the absence of salt, which appears to induce or enhance interaction between the polar-residue regions of TDP-43 as well as other prion-like domain proteins[32,43,44]. To monitor the effect of G335 and G338 variants on phase separation, we mapped the low concentration arm of the phase diagrams as a function of increasing NaCl using a recently developed approach[7,43,45]. In our case, LLPS was initiated by brief incubation of 20 μM CTD with a given concentration of NaCl, followed by centrifugation to sediment the micron-sized droplets making up the phase-separated state (i.e. the protein dense phase). We then measured the concentration of protein in the supernatant (i.e the protein dispersed phase) for each condition, representative of the saturation concentration of TDP-43 CTD (i.e. the concentration of CTD above which LLPS will occur at these conditions). The amount of TDP-43 CTD remaining in the supernatant decreases for WT CTD as a function of NaCl concentration, suggesting CTD is undergoing LLPS (Figure 4A) with "salting-out" behavior, consistent with our previous work on TDP-43 CTD[32] as well as FUS LC[44] and hnRNPA2 LC[43]. Both the ALS-associated variant Q331K as well as the LLPS-disrupting control variant M337P display higher remaining protein in the supernatant than WT for any given salt concentration, consistent with our previous findings that these mutations discourage CTD assembly and phase separation[32]. Conversely, G335 and G338 variants all show lower concentration of protein remaining in the supernatant compared to WT, suggesting that these single point variants are significantly more prone to phase separate. Micrographs of the mutant CTD after phase separation show no morphological differences of these variants compared to WT and no evidence of increased protein aggregation (Figure 4B) suggesting that changes in protein remaining in the supernatant faithfully report on TDP-43 CTD LLPS and not on a transition to aggregation in these conditions. Of our engineered CTD variants, G335A and G338A show the greatest enhancement of LLPS, with only about 5 μM remaining in the supernatant at 150 mM NaCl (i.e. physiological ionic strength) conditions, a nearly 3x decrease compared to WT (~15 μM). The G335D, G335N, and G335S variants displayed more modest enhancements, with approximately 10

μM remaining. Thus, single substitutions can dramatically enhance phase separation via enhancement of helix-helix contacts. Importantly, these saturation concentrations for TDP-43 CTD variants show a strong correlation with changes in CTD secondary structure (Figure 4C) more so than with amino acid hydropathy (Figure S4A,B), suggesting that α-helical structure stabilized by the amino acid at that position, and not hydrophobicity changes, accounts for the large changes in LLPS. Interestingly, only G335D shows a different rank-order in this saturation concentration as a function of increasing salt concentration – at low salt concentrations (<150 mM NaCl), the saturation concentration for G335D is lower (i.e. greater propensity to phase separate) than that of WT, G335N, and G335S, while at the highest salt concentrations tested (300 mM to 500 mM NaCl), it is approximately the same as WT and is higher (less prone to phase separate) than G335N and G335S. This change as a function of salt may arise because G335D decreases the net charge of TDP-43 CTD, leading to less long range repulsion at low salt conditions lacking salt to screen these effects[41].

To further elucidate the potential contributions of hydrophobicity vs. helicity to TDP-43 phase separation, we conducted coarse-grained simulations of TDP-43 CTD coexistence following the simulation model and framework from our previous work[46-48]. Because this coarse-grained model treats protein chains as flexible polymers and secondary structure is not directly captured, this allows us to decouple the mutational change in hydrophobicity from the changes in helicity, which are inseparable in experiment and traditional all-atom molecular simulation. Even in the absence of helicity in the conserved region, self-interaction at the 310-340 region is slightly favored as seen by elevated contact probability in coarse-grained simulations of the condensed phase (Figure S4C). Changing only the amino acid sequence in the coarse-grained model (e.g. G335A) and hence the favorability of interactions, there is no appreciable change in the phase diagram generated from directly simulating phase coexistence (Figure 4D). We then tested the effects of the α-helical configuration by using rigid body constraints to effectively force the subregion identified by 2-chain all atom simulation (residues 320-334; Figure 1B) into an α-helical configuration (Figure 4E). When we compare the results from fixed-helix coexistence simulations (Figure 4F) and those using fully flexible chains, we see that the phase diagram shifts upward to higher temperatures (Figure 4E) indicating a greater phase separation propensity for the fixed-helix chains. Upon introduction of α-helices, the saturation concentration, indicated by the left arm of the coexistence curve, decreases by roughly a factor of 2 across a broad range of temperatures tested (Figure 4E, inset). These results provide direct evidence that the presence of short α-helical structure can promote self-interaction and hence enhance phase separation of TDP-43 CTD, more so than just the increase of sidechain interaction strength.

G335 and G338 variants affect in-cell LLPS fluidity and splicing function

To elucidate how changes in conserved region helicity impact TDP-43 phase behavior and function in a cellular context, we tested the G335 and G338 point mutations in the context of established cell-based reporter assays for TDP-43 phase dynamics and splicing function. We first introduced G335 and G338 mutations into the TDP-43$_{RRM-GFP}$ construct, previously developed to determine the effect of

single point mutations on TDP-43 phase fluidity in cells by fluorescence recovery after photobleaching (FRAP) of the single, large nuclear TDP-43 reporter droplet phase formed in each cell nucleus, into which partitions the great majority of the expressed TDP-43$_{RRM-GFP}$[33]. We measured fluorescence signal recovery for WT and mutant droplets over time (Figure 5A-D). For comparison, we included only droplets of similar size and signal levels in our analysis. Previously tested CTD variants that disrupt helix-helix interaction show increased fluidity due to decreased intermoleuclar interaction, while removal of the conserved segment prevented phase separation[33]. Here, we now show that CTD mutants G335N, G335S, and G335A instead display decreased droplet fluidity relative to WT (Figure 5B,D). G335A shows the greatest reduction of fluidity, while G335N and G335S show intermediate droplet fluidity. Importantly, the rank order of the in cell fluidity matches that of the *in vitro* C$_{sat}$ (see Figure 4) (except for G335D, see below), though some observations were unexpected, highlighting the importance of the cellular context. In particular, FRAP profiles were variable among G335S droplets for reasons not clear at this point (Figure S5A) and the magnitude of the difference in fluidity between G335S and G335N is not predicted from the *in vitro* TDP-43 CTD LLPS C$_{sat}$. Regarding the disease-associated variant G335D, the in cell FRAP recovery time of G335D droplets is not significantly different from that of WT (Figure 5 B,D), similar to the *in vitro* LLPS at high salt. Taken together, these findings show that increasing CR helix propensity by mutating G335 and G338 not only lowers C$_{sat}$ for phase separation *in vitro*, but also reduces TDP-43 fluidity in a tunable manner correlated with the properties observed *in vitro*. Our data furthermore suggest that C$_{sat}$ and droplet material properties can be precisely tuned by single amino acid changes in the helix-helix interacting region.

Considering the links between TDP-43 assembly and splicing function[26,27], we next sought to investigate whether mutations at G335 and G338 can alter its ability to facilitate exon skipping in the cystic fibrosis transmembrane conductance regulator (CFTR) minigene splicing assay[49]. Previously it was shown that deletion of the conserved region containing the helix in TDP-43 CTD disrupted RNA splicing[31]. Here, we wanted to test if enhanced TDP-43 CTD helicity and helix-helix interaction not only promote phase separation but also alter splicing function. To this end, we performed rescue experiments in TDP-43 knock-down HeLa cells comparing the regulation of CFTR exon 9 splicing by mutant and WT TDP-43. We measured the extent of exon skipping in siRNA treated cells expressing control, siRNA-resistant WT, G335, or G338 mutants (Figure 5E, S5C). The percent of exon exclusion was normalized to WT to calculate the relative splicing regulatory activity for each mutant (Figure 5F, S5D). As a negative control, we included the RNA-binding deficient variant F147L/F149L in our assays. Importantly, G335A variant shows increased activity compared to WT (p value < 0.004), suggesting that a single point substitution can indeed enhance TDP-43 function. Contrary to our expectations and unlike the enhancement observed for G335A, G335D and G338A show slightly decreased splicing activity compared to WT. This difference between G335A and G338A in splicing effect may be due to differential alterations in interactions with binding partners including hnRNPA2 mediated via the CR of TDP-43 CTD that contribute to TDP-43 splicing activity[31] or by non-natural over-stabilization of the

TDP-43 interaction. Indeed, in cellular stress conditions, G338A (and G335A) appear to increase cytoplasmic mislocalization and aggregation of full-length TDP-43 compared to WT (Figure S5B). Thus, even though phase separation and splicing activity can be increased in cells by point mutations in the conserved region, TDP-43 self-interaction, droplet fluidity and splicing function may be precisely balanced by sequence to prevent aggregation and preserve splicing function.

Discussion

Diverse structural elements spanning the spectrum from completely disordered regions to globular domains have been shown to contribute to the interactions that undergird biomolecular LLPS. The role of the conserved region within TDP-43 CTD comes into focus here. We show that the α-helical structure of the region can be increased in direct proportion to the α-helix stabilizing propensity of the amino acid substituted at a glycine position, with the helix-promoting alanine substitution having the greatest effect. These data suggest that the backbone conformational entropy discourages helix formation in the 331-343 region of TDP-43, as substitution of glycine 335 with any of several small amino acids (including G335A, G335S, G335N, and G335D) all enhanced helical structure. This helix stabilization in turn enhances TDP-43 CTD helix-helix interaction and phase separation. These data complement our previous report that variants in the same subsection of the conserved region spanning residues 331-343 that becomes helical only upon helix-helix assembly (M337V ALS-associated mutation and the helix-disrupting M337P engineered variant) both discourage helix-helix assembly and phase separation[32]. Given that the helix-stabilizing variants enhance interactions and phase separation, reducing backbone conformational entropy via substitutions at glycine residues in TDP-43 CTD conserved region appears to stabilize protein interactions at this site.

Importantly, the TDP-43 helical region appears unique among RNA-binding proteins – homologous (i.e. evolutionarily related) hydrophobic-residue-rich sequences are apparently not found in other human proteins and seem out of place amidst the polar (FUS, hnRNPA2, remainder of TDP-43 CTD) or highly charged (ddx4)[3] low complexity sequences known to mediate LLPS. Yet, other phase separating proteins such as TIA-1 contain predominantly disordered domains with conserved helix-inducing residues (especially alanine which is enriched in the TDP-43 conserved region), suggesting that TDP-43 CTD may be one in a class of natural proteins containing helical modules that form upon and contribute to phase separation. Importantly, unlike traditional coiled coil or leucine zipper helix-helix interactions[50], TDP-43 CTD conserved region encodes weak (K_d ~100 µM range) and dynamical helical interactions. This range of affinity is consistent with the multivalent but individually weak contacts formed within membraneless organelles and their dynamic, liquid nature where a tight contact (nM range) with slow off-rate could inhibit responsive reorganization of the functional complexes. Indeed, the apparent interaction strength is commensurate with self-assembly domains already characterized

in the remainder of TDP-43: the polar-rich prion-like majority of the CTD that assembles via individually weak multivalent contacts[32], and TDP-43 N-terminal domain that forms weak (K_d ~100 μM) globular protein chains also enhancing phase separation[26]. Similarly, the contacts formed by TDP-43 CTD conserved region are also heterogeneous[32], unlike amphiphilic α-helices that form ordered, rigid bundles. Indeed, CTD assembly may underlie the ability of this region of TDP-43 to interact with splicing factor partners including hnRNPA2[31,43]. Hence TDP-43 CTD conserved region serves as a dynamic and heterogeneous binding module to enhance TDP-43 phase separation.

The identification of single, specific sites that can be used to substantially alter the assembly and material properties of cellular droplets offers an important opportunity to use TDP-43 CTD as a module to tune phase separation in engineered fusion proteins. The multivalent interactions leading to phase separation lend themselves to modulation[51], and the TDP-43 CTD conserved region is shown here to be a uniquely short and tunable module. Interaction can be finely adjusted by single amino acid position changes to dramatically enhance or, as in our previous work, decrease interaction at this site (Figure 3). Enhanced interactions of this region in turn increases phase separation both *in vitro* (Figure 4) and in cells (Figure 5), leading to enhanced function in cells. Therefore, TDP-43 CTD helical region serves as a small structural element that can be used to perturb or control phase separation. Furthermore, compared with repetitive low complexity sequences (e.g. prion-like regions or arginine rich regions)[11] or globular protein/ linear interaction motif pairs (e.g. SH3:proline-rich)[52], TDP-43 CTD's short (~20 residue) module functions with a fine ability to precisely tune phase separation to the desired degree, using the "knob" of mutations that can precisely increase or decrease interaction. Indeed, just as splicing can be altered by G335A that enhances helix-helix assembly of native TDP-43, variants of the TDP-43 conserved region could be fused to other proteins to regulate their cellular functions within membraneless organelles. Furthermore, just as phase separation of low complexity regions can be controlled by post-translational modification[46], targeting modifications to sites adjacent to or within a TDP-43 helical module could be used to alter phase separation and function responsively. Therefore, a series of TDP-43 CTD variants could make for an important tool for engineering phase separation or, more broadly, for tuning dynamic self-organization.
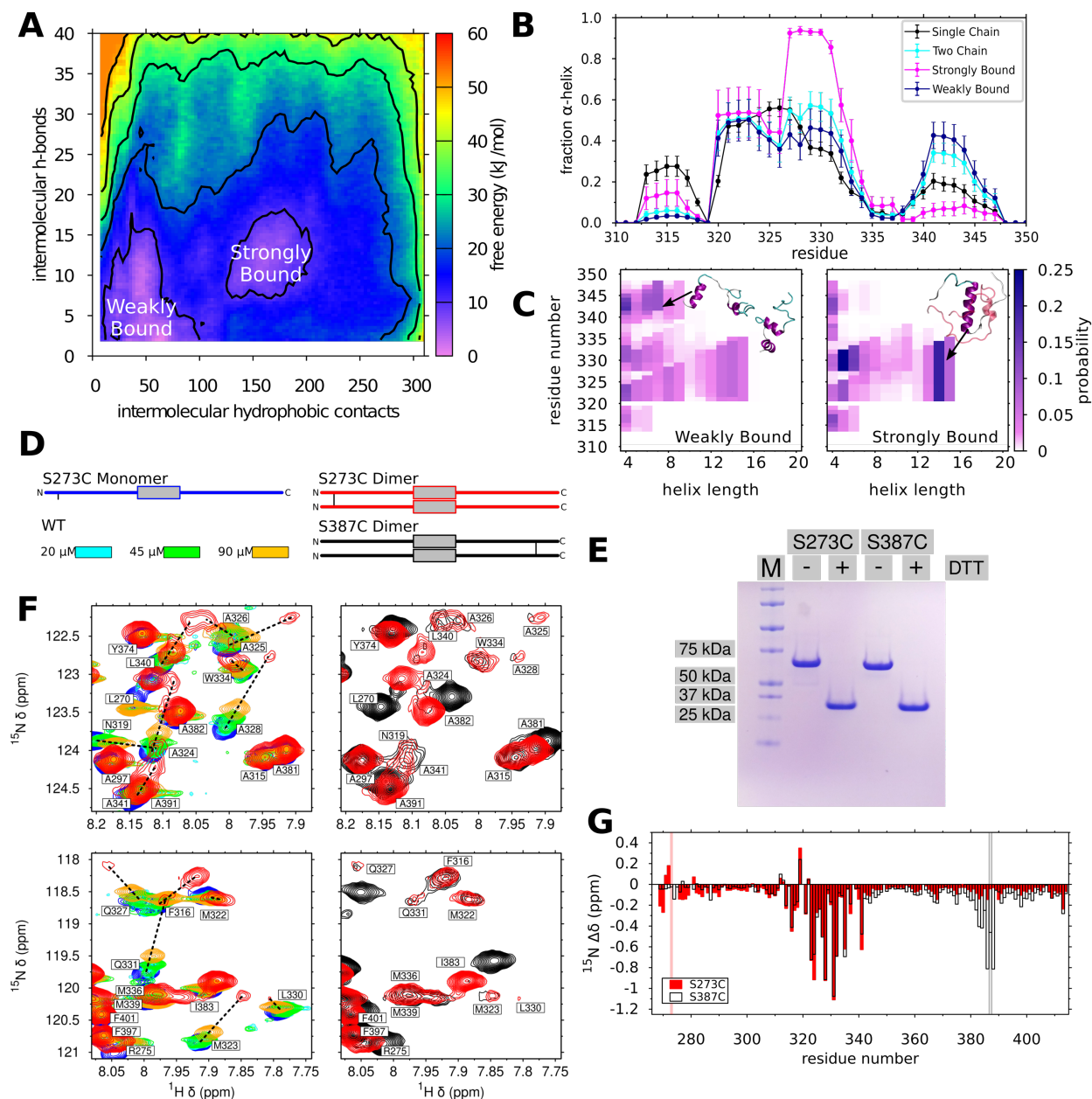
Acknowledgements

Figures

Figure 1: TDP-43 CTD self-associates and forms transient helical structures

A) Free energy landscape of TDP-43$_{310-350}$ intermolecular contacts for all-atom two-chain simulations highlight two minima corresponding to strongly and weakly bound states. B) α-helical content of TDP-43 simulations at each residue. Data are plotted as mean ± s.e.m of n=5 equal divisions of the total data set. C) α-helical secondary structure propensity maps show the contribution of all helical configurations to the overall helical content shown in B. This analysis highlights stabilization of contiguous α-helical structure in the strongly-bound state. D) Schematic of crosslinked CTD variants. E) Non-reducing SDS-PAGE of dimeric S273C and S387C illustrate an apparent reduction of molecular

weight by approximately one half upon addition of disulfide-breaking reducing agent (1 mM DTT). F) (Left) Select regions from ^{1}H-^{15}N TROSY spectra of dilsulfide-linked homodimeric TDP-43 CTD S273C (red) show line broadening and large upfield ^{15}N chemical shift differences in 321-343 region compared to monomeric reduced S273C (blue), consistent with formation of structure. Chemical shifts for residues in the dimer state were assigned by extrapolation (dashed black line) from spectra of 20 μM (cyan), 45 μM (green), and 90 μM (orange) WT. (Right) Overlay of ^{1}H-^{15}N TROSY spectra regions for homodimeric S273C and S387C are similar in the central 320-340 region. G) Quantification of ^{15}N $\Delta\delta$ values between homodimeric S273C (red) and S387C (black) and their respective monomers. Gray boxes indicate the positions of S273C and S387C disulfide crosslinks.
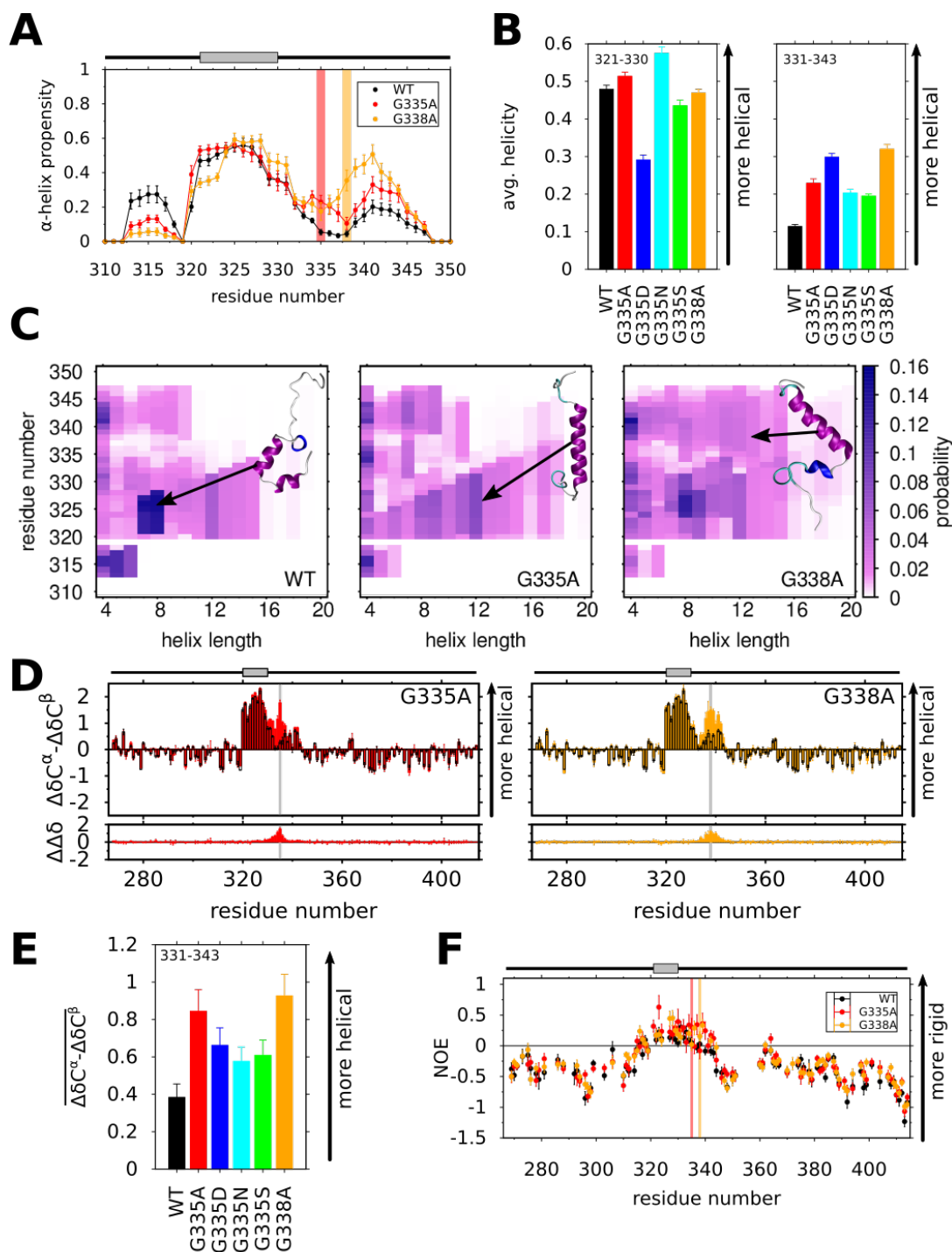
Figure 2: G335 and G338 substitutions enhance in TDP-43 CTD helical region enhance helical stability.

A) Per-residue α-helical propensities for WT, G335A, and G338A for atomistic simulations of TDP-43$_{310-350}$ show higher fraction of α-helical structure near the sites of mutation (highlighted by red, G335A, or orange, G338A, bars). Data are plotted as mean ± s.e.m of n=5 equal divisions of the total

data set. B) Average simulated helicity is mostly unaffected for residues 321-330 (left) away from G335 and G338 mutation sites but enhanced for 331-343 (right) surrounding the mutation site. C) Maps of contiguous α-helical structure location (y- axis) and length (x-axis) for TDP-43$_{310-350}$ show increased probability for longer helix structure in G335A and G338A relative to WT. One example configuration from a subpopulation of the simulation ensemble (indicated by the black arrow) is displayed for each variant. D) Experimental NMR secondary chemical shifts ($\boldsymbol{\Delta\delta C^{\alpha} - \Delta\delta C^{\beta}}$) and the differences in secondary shifts with respect to WT ($\triangle\triangle\delta$) for G335A and G338A show increased α-helical structure near the site of mutation. The secondary shift values for WT are overlaid in black for comparison. E) The average experimental secondary shift ($\overline{\boldsymbol{\Delta\delta C^{\alpha} - \Delta\delta C^{\beta}}}$) for TDP-43 331-343 of WT and mutant CTD highlight increases in local α-helical structure in all variants. Error bars are SEM. F) Higher {^{1}H}-^{15}N heteronuclear nuclear Overhauser effect (NOE) values measured for G335A (red) and G338A (orange) compared to WT (black) near the site of mutation indicate slowed local protein backbone motion, consistent with enhanced helical structure. Mutation positions G335A and G338A highlighted with gray bars. For all above panels: Simulation data are plotted as mean and SEM from 10 equal blocks of equilibrated structural ensemble. Unless otherwise stated, experimental data are plotted as mean and SD estimated from signal to noise ratio.
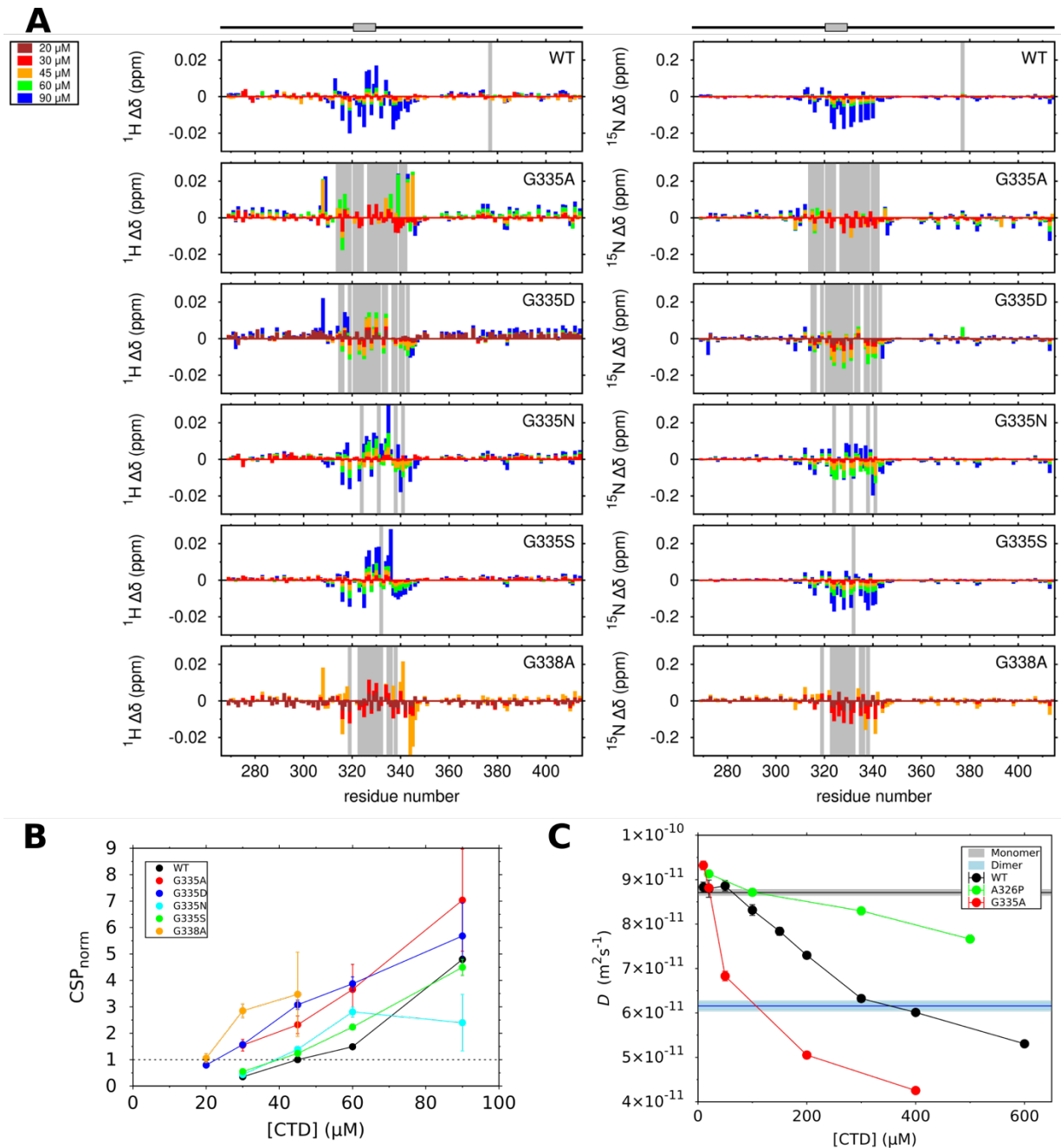
Figure 3: Mutations enhance TDP-43 CTD helix-helix interaction and assembly.

A) Larger ^{15}N chemical shift differences ($\Delta\delta$) in the helical region of TDP-43 CTD for G335A and G338A compared to WT are consistent with enhanced helix-helix interactions. Shifts are reported at concentrations from 20 μM to 90 μM with respect to a monomeric (low concentration) reference (20 μM for WT and G335A; 10 μM for G338A). Gray boxes indicate the loss of detectable peak intensity due to line broadening. B) Normalized chemical shift differences vs concentration are higher for mutant

CTD compared to WT, consistent with increased self-assembly for G335 and G338 variants. Error bars are SD estimated by 100 bootstrapping simulations derived from experimental uncertainty. C) Diffusion coefficients calculated from NMR diffusion data plotted as a function of CTD concentration show that WT (black) and G335A (red) diffuse slower than A326P (green). Average diffusion coefficients for S273C and S387C dimers and S273C and S387C monomers are indicated by blue and black horizontal lines, respectively. Data are plotted as mean ± s.d., indicated for monomer and dimer diffusion coefficients by gray and light blue boxes, respectively.
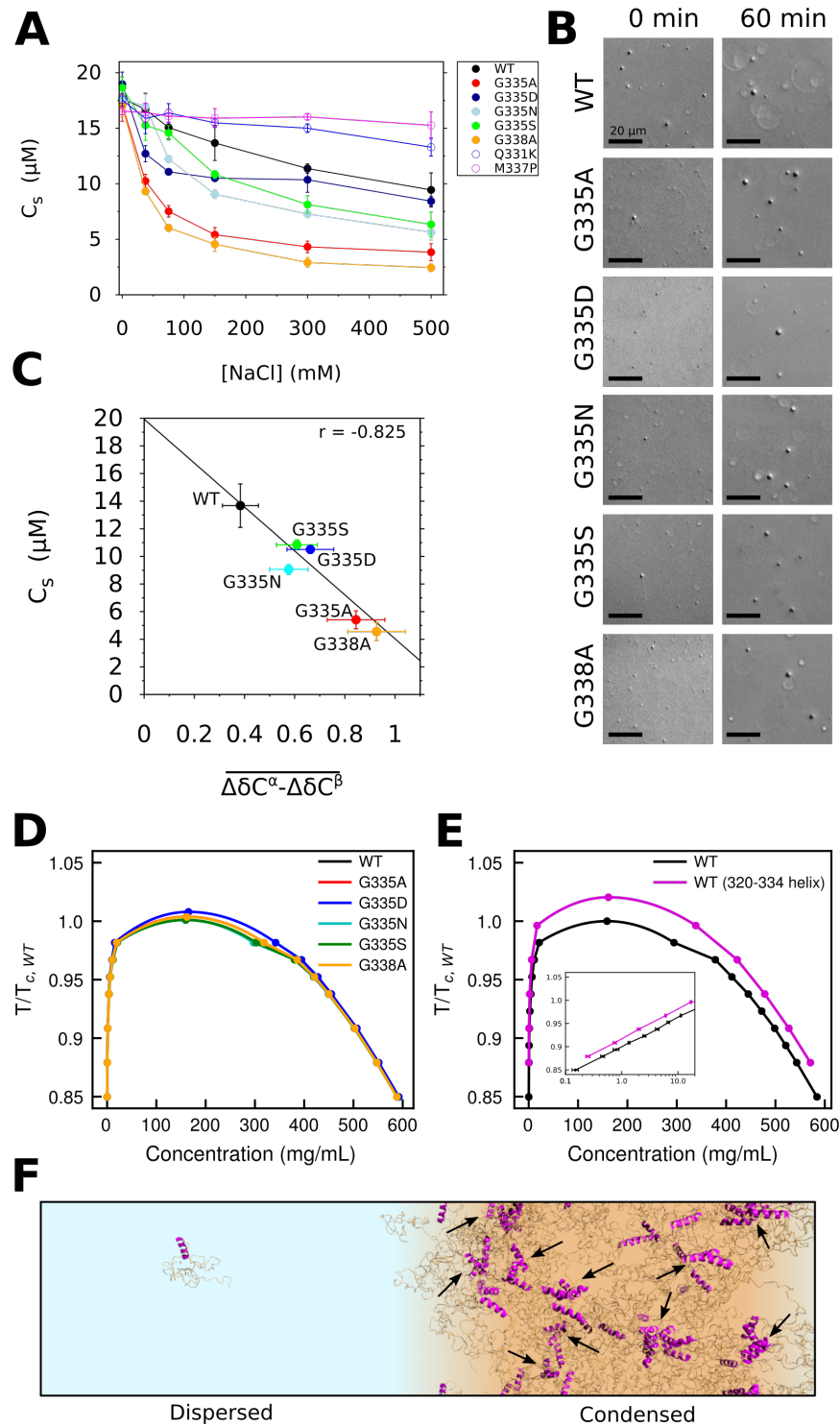
Figure 4: Mutations at G335 and G338 alter *in vitro* LLPS

A) Phase diagrams for WT and mutant CTD measured for 0 to 500 mM NaCl. Error bars represent SD

of three replicates. B) The saturation concentration (C_s) of WT and mutant CTD at 150 mM NaCl ("physiological conditions") correlates well with the average change in secondary shift ($\overline{\Delta\delta}$) (Pearson r = -0.825). Error bars represent SD of three replicates for C_s and SEM for $\overline{\Delta\delta C^\alpha - \Delta\delta C^\beta}$. C) DIC micrographs of LLPS for WT TDP-43 CTD and variants G335A, G335D, G335N, G335S, and G338A at 0 min and 60 min after addition of NaCl to 300 mM. Scale bar is 20 μm. D) Coexistence curves of TDP-43 CTD from coarse-grained (CG) slab simulations of WT and G335/G338 mutants illustrate no discernable difference in LLPS behavior as a result of mutation. E) Coexistence curves of TDP-43 from CG slab simulations illustrate reduced thermal miscibility when residues 320-334 are treated as a rigid body (magenta) instead of fully flexible (black). Inset shows the ~2-fold change to left arm of the phase diagram with temperature vs. concentration (on a log scale). F) A snapshot of the TDP-43 CTD CG slab simulation where CTD molecules harboring rigid α-helical substructure (magenta) visit both dispersed and condensed states in phase coexistence. Arrows indicate oligomers formed by helix-helix contacts.
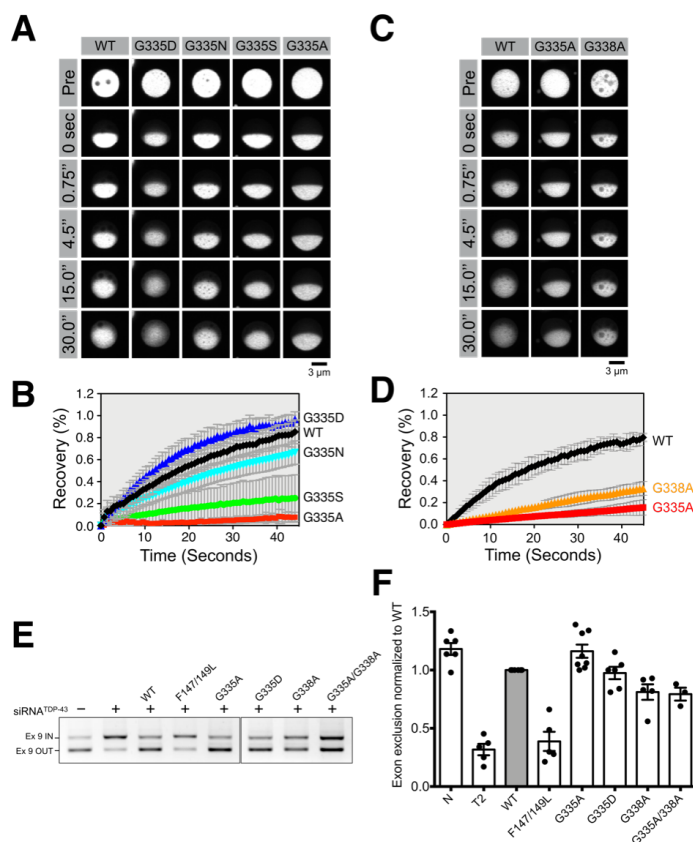

Figure 5: Mutations at G335 and G338 alter in-cell LLPS fluidity and splicing function

A-D) Fluorescence recovery after in cell half-droplet photobleaching and quantification for TDP-43_RRM-GFP reporter construct with (A,B) WT, G335D, G335N, G335S, and G335A or (C,D) WT, G335A, and G338A CTD sequences. TDP-43_RRM-GFP shows differential droplet fluidity for G335 variants relative to

WT. E) Example agarose gel of PCR products to quantify the fraction of exon 9 CFTR exclusion (OUT) expressed from a minigene construct transfected in HeLa cells. Cells were treated with control, TDP-43 siRNA or TDP-43 siRNA plus WT or mutant vectors. F) The splicing activity was calculated as the ratio of exon 9 exclusion relative to WT. Error bars are SEM of 3 or more replicates.

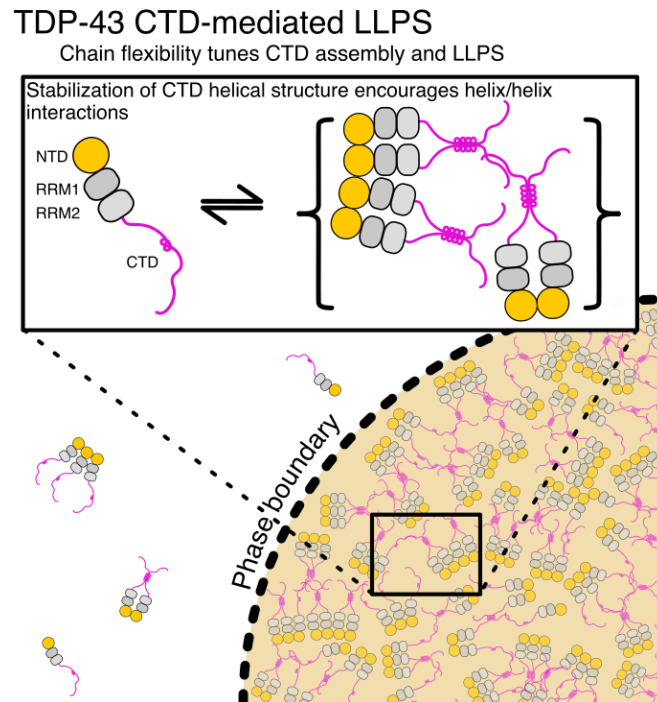**Figure 6: A model for contributions of CTD chain flexibility to TDP-43 LLPS**

The TDP-43 CTD is sufficient for assembly into sparsely populated oligomeric states and LLPS states at low and high salt concentrations, respectively. Engineered or disease-causing mutations that stabilize CTD α-helical structure result in increased helix/helix intermolecular interactions and phase separation.

Methods

Protein Expression and Purification

Point mutations were introduced into the TDP-43$_{267-414}$ (CTD) construct previously described[32] using Quikchange mutagenesis or overlap extension PCR with mutagenic primers[53].

WT and mutant CTD constructs were expressed in *Escherichia coli* BL21 Star DE3 (Invitrogen) cells in either LB media or, where indicated as uniformly labeled ^{15}N or ^{15}N /^{13}C, in M9 minimal media supplemented with ^{15}NH$_4$Cl and/or ^{13}C -glucose as the sole nitrogen and carbon sources, respectively. Recombinant proteins were purified as previously described[32], with the exception that the TEV cleavage buffer was substituted with 20 mM Tris, 500 mM GndHCl, pH 8.0 for more efficient cleavage of the purification tag. Additionally, proteins were exchanged into Storage Buffer (20 mM Tris, 8 M urea, pH 8.0) before flash freezing and storage at -80 $^{\circ}$C. For all experiments, purified proteins were desalted into MES Buffer (20 mM MES, pH 6.1) using 0.5 mL Zeba spin desalting columns (Thermo Scientific) in accordance with manufacturer instructions.

NMR Spectroscopy

All NMR experiments were preformed in 20 mM MES buffer pH 6.1 (pH adjusted with NaOH) supplemented with 10% D$_2$O. Experiments were recorded on Bruker Avance spectrometers operating at either 850 MHz or 500 MHz ^{1}H Larmor frequency with TCI z-gradient cryoprobes at 25° C. Backbone resonance assignments for CTD variants were obtained by standard triple resonance experiments HNCACB and CBCACONH as previously described[32]. Subsequent secondary chemical shifts were obtained from comparison of experimental C$^{\alpha}$ and C$^{\beta}$ chemical shifts with their respective predicted random coil reference chemical shifts[54,55].

^{15}N R_1, ^{15}N R_2, and {^{1}H } ^{15}N heteronuclear NOE experiments were measured for 20 μM WT, 20 μM G335A, and 10 μM - 20 μM G338A ^{15}N-labeled CTD at 500 MHz ^{1}H Larmor frequency using standard pulse sequences (hsqct1etf3gpsitc3d, hsqct2etf3gpsitc3d, hsqc-noef3gpsi). Each ^{15}N R_1 experiment comprises seven interleaved ^{15}N R_1 relaxation delays: 100 ms, 200 ms, 300 ms, 400 ms, 600 ms, 800 ms, and 1000 ms with a recycle delay of 1.5 s. Each ^{15}N R_2 experiment comprises six interleaved ^{15}N R_2 relaxation delays with a B_1 field of 556 Hz: 16.4 ms, 32.8 ms, 65.5 ms, 131.2 ms, 229.6 ms, and 262.4 ms with a recycle delay of 2 s. {^{1}H}-^{15}N heteronuclear NOE experiments comprise interleaved proton saturation sequences and reference sequences without proton saturation with a recycle delay of 5 s. Measurement of η_{xy} was accomplished using the pulse sequence described previously[56], with a relaxation delay (T) of 42 ms and recycle delay of 1.5 s.

Chemical shift perturbations were quantified for either WT and mutant CTD variants by measuring the chemical shifts of ^{1}H-^{15}N crosspeaks of HSQCs acquired at 850 MHz ^{1}H Larmor frequency at concentrations ranging from 10 μM to 90 μM as previously described[32].

Pulsed-field gradient diffusion measurements for WT and mutant CTD at concentrations ranging from

10 µM to 600 µM acquired using the longitudinal eddy current delay with bipolar gradient scheme[57] provided in the Bruker standard library (ledbpgppr2), with 16 gradient strengths: 0.96 G/cm, 5.39 G/cm, 9.92 G/cm, 13.62 G/cm, 16.58 G/cm, 20.10 G/cm, 23.75 G/cm, 27.26 G/cm, 30.83 G/cm, 34.37 G/cm, 37.90 G/cm, 41.38 G/cm, 44.82 G/cm, 48.08 G/cm, 51.33 G/cm, and 54.46 G/cm. The diffusion time was set to 250 ms with encoding/decoding gradient pulse lengths of 1 ms, using a recycle delay of 2 s. Gradient powers were calibrated based on the known diffusion coefficient of trace H_2O in a 99.8 % D_2O sample[58]. In all cases, intensity ratios were calculated by least-squares regression of select ^1H spectral regions for each gradient power against those of a reference gradient power (0.96 G/cm). Standard error of the intensity ratios was estimated with 100 boostrapping simulations.

All NMR spectra were processed with NMRPipe[59] and analyzed with NMRFAM-Sparky[60]. ^{15}N R_1, ^{15}N R_2, {^1H} ^{15}N NOE, and η_{xy} values and their respective standard errors were calculated using in-house scripts.

Estimating R_{ex} for WT and G335A CTD

Chemical exchange contributions to ^{15}N R_2 (R_{ex}) were calculated based on the relationship:

$$R_2 = \kappa\eta_{xy} \tag{1}$$

where κ is a scaling parameter that can be approximated for all residues as the trimmed mean of R_2/η_{xy}[61]. In our case, two values of κ were calculated due to the differential relaxation rates across the CTD resulting from both structured and disordered regions. η_{xy} rates were first separated into one of two equal-width bins, and their respective κ values were obtained by solving equation (1). Site-specific R_{ex} values were then calculated from R_2, κ, and η_{xy}, as:

$$R_{ex} = R_2 - \kappa\eta_{xy} \tag{2}$$

Crosslinking

CTD variants harboring engineered cysteine residues at one of two serine positions (S273C or S387C) were buffer exchanged and diluted to 100 µM in Storage Buffer (see above, which includes 8M urea to maintain solubility). Crosslinking was initiated by addition of copper (II) phenanthroline[62] to a final concentration of 1 mM, followed by incubation at room temperature for at least 1 hour. Reactions were quenched by addition of EDTA to 5 mM, and the resulting mixture was purified by size-exclusion chromatography with a 26/600 Superdex 75 equilibrated in Storage Buffer supplemented with 1 mM EDTA. Fractions corresponding to dimeric CTD were pooled, concentrated up to 750 µM ([dimer]), flash frozen, and stored at -80C. NMR samples were prepared by diluting 12.5 µL of the dimeric stock into 437.5 µL MES Buffer, resulting in a final urea concentration of 200 mM. Samples were then centrifuged for 10 minutes at 12,000 rpm, and the concentration of the supernatant was measured by A_{280}. The samples were then diluted to the final desired protein concentration with 1:35 Storage Buffer:MES Buffer, thus ensuring the urea concentration remains constant. In order to prepare monomeric control samples, the samples were supplemented with 1 mM DTT (Sigma) prior to centrifugation and at all subsequent steps.

Quantifying phase separation *in vitro*

WT and mutant CTD variants were desalted into MES Buffer and diluted to 40 µM. Phase separation was initiated by adding an equal volume of MES Buffer supplemented with either 0, 75, 150, 300, 600, or 1000 mM NaCl, resulting in final salt concentrations of 0, 37.5, 75, 150, 300, and 500 mM NaCl, respectively. Samples were mixed by gentle agitation and promptly centrifuged at 12,000 rpm for 10 min. The concentration of CTD in the supernatant was measured by A_{280} (E_{280} = 17990 $M^{-1}cm^{-1}$) with a Nanodrop 2000c. All experiments were performed in triplicate.

Microscopy

WT and mutant CTD were desalted into MES Buffer and diluted to 40 µM. Phase separation was initiated by addition of an equal volume of MES Buffer supplemented with either 0 mM or 600 mM NaCl, resulting in final salt concentrations of 0 or 300 mM NaCl, respectively. Samples were spotted onto glass coverslips, and differential interference contrast (DIC) micrographs were acquired with a Zeiss Axiovert 200M as previously described[32]. Images were processed with ImageJ.

Translational Diffusion

Translational diffusion coefficients of WT and mutant CTD were calculated by non-linear least squares fit of CH_3 ^1H spectral region intensity ratios to the Stejskal-Tanner equation[63]:

$$\frac{I(g)}{I_0} = e^{-\gamma D \delta^2 g^2 (\Delta - \frac{\delta}{3})}$$ (3)

where γ is the ^1H gyromagnetic ratio, D is the translational diffusion coefficient, δ is the length of the bipolar encoding/decoding gradient element, Δ is the diffusion time, and g is the gradient strength. Standard error in the diffusion coefficients was estimated from the variance-covariance matrix of nonlinear least-squares model parameters.

TDP-43_{RRM-GFP} reporter assays

HEK-293T cells (ATCC catalog number CRL-3216; RRID:CVCL_0063) maintained in DMEM high glucose (GE Healthcare) supplemented with 10% FBS, 2mM L-glutamine (Gibco), 1mM non-essential amino acids (Gibco) and 1% penicillin/streptomycin (Gibco) at 37°C and 5% CO_2 were seeded onto ibidi µ-slides at 30,00 cells/well and transfected with TDP-43_{RRM-GFP} reporter constructs (300 ng) using X-tremeGENE 9 (Roche). After 24 hours, the culture medium was exchanged with L-15 medium (Gibco) supplemented with 10% FBS, and the cells were imaged using a Leica SP8 laser-scanning confocal microscope equipped with a 488-nm laser, 63x glycerol objective (NA 1.3) and a temperature-controlled incubation chamber (Life Imaging Services) set to 37°C.

Half-bleach experiments were recorded and quantified using the FRAP module of the Leica Application Suite X software as previously described[33]. In short, for each recorded time point (t), the fluorescence intensities within the bleached droplet hemisphere were then normalized to the fluorescence intensity of the corresponding unbleached droplet hemisphere. These normalized, time-dependent fluorescence

intensities I_t were then used to calculate the fluorescence recovery (FR) according to the following formula:

$FR(t) = (I_t - I_{t0}) / (I_{before\ bleaching} - I_{t0})$, with t0 being the first time point observed after photobleaching. Two replicate experiments per construct were performed and at least ten individual droplets from different cells were analyzed per experiment. Replicate measurements (mean ± SD) were plotted using GraphPad Prism. To ensure that only similarly-sized droplets of comparable intensities were bleached, the same bleach window and detector settings were used. Droplet diameters and fluorescence intensities were quantified using Mathematica (Wolfram Research) as previously described[33].

Splicing assays

WT and mutant siRNA-resistant constructs were generated by site-directed mutagenesis in N-terminally FLAG-tagged pCMV2[64]. HeLa cells were grown in growth media-Dulbecco's modified Eagle's medium, 4,500 mg/l glucose, l-glutamine, and sodium bicarbonate and supplemented with filtered fetal bovine serum at 10%. Cells were transiently transfected with Lipofectamine 2000 and Oligofectamine in the case of siRNA (Life Technologies) according to manufacturer protocols. RNAi-mediated downregulation of TDP-43 was carried out as previously described[64] and siCONTROL Nontargetting siRNA#1 (Dharmacon) was used as control. TDP-43 downregulation was confirmed by immunoblotting. Splicing assays were performed using the CFTR reporter minigene as previously described[64]. WT or mutant siRNA-resistant vectors were co-transfected in T2-treated cells. To quantify exon skipping we performed standard PCR with primers specific for exon 9 flanking regions following RNA extraction and cDNA synthesis. The percent exon exclusion was calculated using ImageJ analysis.

All-atom simulations

Parallel tempering in the well-tempered ensemble (PT-WTE) simulations[65,66] were conducted on a 41-residue fragment of TDP-43 CTD, TDP-43$_{310-350}$ using the GROMACS-4.6.7 molecular dynamics software package[67]. We use the Amber03ws force field, which has been developed and parameterized for use with disordered proteins and protein folding[68]. For simulations of WT TDP-43$_{310-350}$ dimer, we also apply a well-tempered metadynamics (WT-MetaD) bias[69] on two collective variables, intermolecular hydrophobic contacts and intermolecular hydrogen bonds to aid in sampling bound and unbound states.

Coarse-grained simulations

Coarse-grained simulations were conducted using an amino-acid-resolution model with 20 residue types to capture sequence specificity, having interactions based on relative hydropathies of each amino acid[48,70]. Each system was simulated at a range of temperatures using constant volume and temperature using a Langevin thermostat, and slab geometry, which allows for efficient sampling of coexistence densities. Simulations of phase coexistence were conducted using HOOMD-Blue v2.1.5 software package[71,72]. For simulations of helical TDP-43, residues 320-334 were constrained to a helical configuration using rigid body dynamics.

Analysis of Simulations

Secondary structures were calculated using the DSSP algorithm[73], and secondary structure maps are used to show heterogeneity of helicity within the equilibrium ensemble[39]. From the PT-WTE-MetaD dimer simulation, we generate the free energy profile with respect to the two collective variables[74]. All analysis done on the dimer simulation was reweighted based on the applied bias. Chemical shifts were calculated using SPARTA+[75], and normalized using random-coil shifts from the Poulson webserver[55].

Data Availability

NMR chemical shift assignments in this paper for TDP-43 CTD G335A (27750), G335N (27788), G335S (27789), G338A (27790), and G335D (in process) can be obtained online from the Biological Magnetic Resonance Database (BMRB, http://www.bmrb.wisc.edu/). TDP-43 CTD WT chemical shifts were obtained from the BMRB (26823). Plasmids generated for this project can be found at Addgene.org.

Author Contributions

JM and NLF conceived and designed research. AEC and AMD conducted biophysical experiments. GLD and GHZ performed and analyzed simulations. HBS and YMA conducted in-cell experiments. YCK developed tools to aid in setup and analysis of simulations. RR, YMA, JM, and NLF funded and supervised research. AEC, GLD, JM, and NLF wrote manuscript.

Competing Interests

The authors declare no competing interests.

References

1   Kato, M., Han, T. W., Xie, S., Shi, K., Du, X., Wu, L. C., Mirzaei, H., Goldsmith, E. J., Longgood, J., Pei, J., Grishin, N. V., Frantz, D. E., Schneider, J. W., Chen, S., Li, L., Sawaya, M. R., Eisenberg, D., Tycko, R. & McKnight, S. L. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753-767, (2012).

2   Mitrea, D. M., Cika, J. A., Guy, C. S., Ban, D., Banerjee, P. R., Stanley, C. B., Nourse, A., Deniz, A. A. & Kriwacki, R. W. Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* **5**, (2016).

3   Nott, T. J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowietz, A., Craggs, T. D., Bazett-Jones, D. P., Pawson, T., Forman-Kay, J. D. & Baldwin, A. J. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol Cell* **57**, 936-947, (2015).

4   Brangwynne, C. P., Eckmann, C. R., Courson, D. S., Rybarska, A., Hoege, C., Gharakhani, J., Julicher, F. & Hyman, A. A. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732, (2009).

5   Hyman, A. A., Weber, C. A. & Julicher, F. Liquid-liquid phase separation in biology. *Annu Rev Cell Dev Biol* **30**, 39-58, (2014).

6   Brangwynne, C. P., Tompa, P. & Pappu, R. V. Polymer physics of intracellular phase transitions.

Nat Phys **11**, 899-904, (2015).

7 Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C. C., Eckmann, C. R., Myong, S. & Brangwynne, C. P. The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc Natl Acad Sci U S A* **112**, 7189-7194, (2015).

8 Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T. & Taylor, J. P. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **163**, 123-133, (2015).

9 Patel, A., Lee, Hyun O., Jawerth, L., Maharana, S., Jahnel, M., Hein, Marco Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, Titus M., Pozniakovski, A., Poser, I., Maghelli, N., Royer, Loic A., Weigert, M., Myers, Eugene W., Grill, S., Drechsel, D., Hyman, Anthony A. & Alberti, S. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **162**, 1066-1077, (2015).

10 Murray, D. T., Kato, M., Lin, Y., Thurber, K. R., Hung, I., McKnight, S. L. & Tycko, R. Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell* **171**, 615-627 e616, (2017).

11 Lin, Y., Protter, D. S., Rosen, M. K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell*, (2015).

12 Amaya, J., Ryan, V. H. & Fawzi, N. L. The SH3 domain of Fyn kinase interacts with and induces liquid-liquid phase separation of the low-complexity domain of hnRNPA2. *J Biol Chem* **293**, 19522-19531, (2018).

13 Quiroz, F. G. & Chilkoti, A. Sequence heuristics to encode phase behaviour in intrinsically disordered protein polymers. *Nat Mater* **14**, 1164-1171, (2015).

14 Kroschwald, S., Munder, M. C., Maharana, S., Franzmann, T. M., Richter, D., Ruer, M., Hyman, A. A. & Alberti, S. Different Material States of Pub1 Condensates Define Distinct Modes of Stress Adaptation and Recovery. *Cell Rep* **23**, 3327-3339, (2018).

15 Schuster, B. S., Reed, E. H., Parthasarathy, R., Jahnke, C. N., Caldwell, R. M., Bermudez, J. G., Ramage, H., Good, M. C. & Hammer, D. A. Controllable protein phase separation and modular recruitment to form responsive membraneless organelles. *Nat Commun* **9**, 2985, (2018).

16 Alami, N. H., Smith, R. B., Carrasco, M. A., Williams, L. A., Winborn, C. S., Han, S. S., Kiskinis, E., Winborn, B., Freibaum, B. D., Kanagaraj, A., Clare, A. J., Badders, N. M., Bilican, B., Chaum, E., Chandran, S., Shaw, C. E., Eggan, K. C., Maniatis, T. & Taylor, J. P. Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. *Neuron* **81**, 536-543, (2014).

17 Bentmann, E., Neumann, M., Tahirovic, S., Rodde, R., Dormann, D. & Haass, C. Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA-binding protein of 43 kDa (TDP-43). *The Journal of biological chemistry* **287**, 23079-23094, (2012).

18 Colombrita, C., Zennaro, E., Fallini, C., Weber, M., Sommacal, A., Buratti, E., Silani, V. & Ratti, A. TDP-43 is recruited to stress granules in conditions of oxidative insult. *J Neurochem* **111**, 1051-1061, (2009).

19 Liu-Yesucevitz, L., Bilgutay, A., Zhang, Y. J., Vanderweyde, T., Citro, A., Mehta, T., Zaarur, N., McKee, A., Bowser, R., Sherman, M., Petrucelli, L. & Wolozin, B. Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS One* **5**, e13250, (2010).

20 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., McCluskey, L. F., Miller, B. L., Masliah, E., Mackenzie, I. R., Feldman, H., Feiden, W., Kretzschmar, H. A., Trojanowski, J. Q. & Lee, V. M. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130-133, (2006).

21 Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y. & Oda, T. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Commun* **351**, 602-611, (2006).

22 Josephs, K. A., Murray, M. E., Whitwell, J. L., Tosakulwong, N., Weigand, S. D., Petrucelli, L.,

Liesinger, A. M., Petersen, R. C., Parisi, J. E. & Dickson, D. W. Updated TDP-43 in Alzheimer's disease staging scheme. *Acta Neuropathol* **131**, 571-585, (2016).

23 Lukavsky, P. J., Daujotyte, D., Tollervey, J. R., Ule, J., Stuani, C., Buratti, E., Baralle, F. E., Damberger, F. F. & Allain, F. H. Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat Struct Mol Biol* **20**, 1443-1449, (2013).

24 Tollervey, J. R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., Konig, J., Hortobagyi, T., Nishimura, A. L., Zupunski, V., Patani, R., Chandran, S., Rot, G., Zupan, B., Shaw, C. E. & Ule, J. Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci* **14**, 452-458, (2011).

25 Xiao, S., Sanelli, T., Dib, S., Sheps, D., Findlater, J., Bilbao, J., Keith, J., Zinman, L., Rogaeva, E. & Robertson, J. RNA targets of TDP-43 identified by UV-CLIP are deregulated in ALS. *Mol Cell Neurosci* **47**, 167-180, (2011).

26 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., Shewmaker, F., Naik, M. T., Mittag, T., Ayala, Y. M. & Fawzi, N. L. A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. *EMBO J* **37**, (2018).

27 Afroz, T., Hock, E. M., Ernst, P., Foglieni, C., Jambeau, M., Gilhespy, L. A. B., Laferriere, F., Maniecka, Z., Pluckthun, A., Mittl, P., Paganetti, P., Allain, F. H. T. & Polymenidou, M. Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. *Nat Commun* **8**, 45, (2017).

28 Romano, V., Quadri, Z., Baralle, F. E. & Buratti, E. The structural integrity of TDP-43 N-terminus is required for efficient aggregate entrapment and consequent loss of protein function. *Prion* **9**, 1-9, (2015).

29 Mompean, M., Romano, V., Pantoja-Uceda, D., Stuani, C., Baralle, F. E., Buratti, E. & Laurents, D. V. Point mutations in the N-terminal domain of transactive response DNA-binding protein 43 kDa (TDP-43) compromise its stability, dimerization, and functions. *J Biol Chem* **292**, 11992-12006, (2017).

30 Mompean, M., Romano, V., Pantoja-Uceda, D., Stuani, C., Baralle, F. E., Buratti, E. & Laurents, D. V. The TDP-43 N-Terminal Domain Structure at High Resolution. *FEBS J*, (2016).

31 D'Ambrogio, A., Buratti, E., Stuani, C., Guarnaccia, C., Romano, M., Ayala, Y. M. & Baralle, F. E. Functional mapping of the interaction between TDP-43 and hnRNP A2 in vivo. *Nucleic Acids Res* **37**, 4116-4126, (2009).

32 Conicella, A. E., Zerze, G. H., Mittal, J. & Fawzi, N. L. ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **24**, 1537-1549, (2016).

33 Schmidt, H. B. & Rohatgi, R. In Vivo Formation of Vacuolated Multi-phase Compartments Lacking Membranes. *Cell Rep* **16**, 1228-1236, (2016).

34 Buratti, E. TDP-43 post-translational modifications in health and disease. *Expert Opin Ther Targets* **22**, 279-293, (2018).

35 Lim, L., Wei, Y., Lu, Y. & Song, J. ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43. *PLoS Biol* **14**, e1002338, (2016).

36 Roberts, S., Harmon, T. S., Schaal, J. L., Miao, V., Li, K. J., Hunt, A., Wen, Y., Oas, T. G., Collier, J. H., Pappu, R. V. & Chilkoti, A. Injectable tissue integrating networks from recombinant polypeptides with tunable order. *Nat Mater* **17**, 1154-1163, (2018).

37 Corrado, L., Ratti, A., Gellera, C., Buratti, E., Castellotti, B., Carlomagno, Y., Ticozzi, N., Mazzini, L., Testa, L., Taroni, F., Baralle, F. E., Silani, V. & D'Alfonso, S. High frequency of TARDBP gene mutations in Italian patients with amyotrophic lateral sclerosis. *Hum Mutat* **30**, 688-694, (2009).

38 Mittal, J., Yoo, T. H., Georgiou, G. & Truskett, T. M. Structural ensemble of an intrinsically disordered polypeptide. *J Phys Chem B* **117**, 118-124, (2013).

39 Iglesias, J., Sanchez-Martínez, M. & Crehuet, R. SS-map. *Intrinsically Disordered Proteins* **1**,

e25323, (2013).

40 Bryson, J. W., Betz, S. F., Lu, H. S., Suich, D. J., Zhou, H. X., O'Neil, K. T. & DeGrado, W. F. Protein design: a hierarchic approach. *Science* **270**, 935-941, (1995).

41 Mompean, M., Chakrabartty, A., Buratti, E. & Laurents, D. V. Electrostatic Repulsion Governs TDP-43 C-terminal Domain Aggregation. *PLoS Biol* **14**, e1002447, (2016).

42 Bodner, C. R., Dobson, C. M. & Bax, A. Multiple tight phospholipid-binding modes of alpha-synuclein revealed by solution NMR spectroscopy. *Journal of molecular biology* **390**, 775-790, (2009).

43 Ryan, V. H., Dignon, G. L., Zerze, G. H., Chabata, C. V., Silva, R., Conicella, A. E., Amaya, J., Burke, K. A., Mittal, J. & Fawzi, N. L. Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol Cell* **69**, 465-479 e467, (2018).

44 Burke, K. A., Janke, A. M., Rhine, C. L. & Fawzi, N. L. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol Cell*, (2015).

45 Mackenzie, I. R., Nicholson, A. M., Sarkar, M., Messing, J., Purice, M. D., Pottier, C., Annu, K., Baker, M., Perkerson, R. B., Kurti, A., Matchett, B. J., Mittag, T., Temirov, J., Hsiung, G. R., Krieger, C., Murray, M. E., Kato, M., Fryer, J. D., Petrucelli, L., Zinman, L., Weintraub, S., Mesulam, M., Keith, J., Zivkovic, S. A., Hirsch-Reinshagen, V., Roos, R. P., Zuchner, S., Graff-Radford, N. R., Petersen, R. C., Caselli, R. J., Wszolek, Z. K., Finger, E., Lippa, C., Lacomis, D., Stewart, H., Dickson, D. W., Kim, H. J., Rogaeva, E., Bigio, E., Boylan, K. B., Taylor, J. P. & Rademakers, R. TIA1 Mutations in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia Promote Phase Separation and Alter Stress Granule Dynamics. *Neuron* **95**, 808-816 e809, (2017).

46 Monahan, Z., Ryan, V. H., Janke, A. M., Burke, K. A., Rhoads, S. N., Zerze, G. H., O'Meally, R., Dignon, G. L., Conicella, A. E., Zheng, W., Best, R. B., Cole, R. N., Mittal, J., Shewmaker, F. & Fawzi, N. L. Phosphorylation of the FUS low-complexity domain disrupts phase separation, aggregation, and toxicity. *EMBO J* **36**, 2951-2967, (2017).

47 Dignon, G. L., Zheng, W., Best, R. B., Kim, Y. C. & Mittal, J. Relation between single-molecule properties and phase behavior of intrinsically disordered proteins. *Proc Natl Acad Sci U S A* **115**, 9929-9934, (2018).

48 Dignon, G. L., Zheng, W., Kim, Y. C., Best, R. B. & Mittal, J. Sequence determinants of protein phase behavior from a coarse-grained model. *PLoS Comput Biol* **14**, e1005941, (2018).

49 Buratti, E., Dork, T., Zuccato, E., Pagani, F., Romano, M. & Baralle, F. E. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J* **20**, 1774-1784, (2001).

50 Landschulz, W. H., Johnson, P. F. & McKnight, S. L. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759-1764, (1988).

51 Shin, Y., Berry, J., Pannucci, N., Haataja, M. P., Toettcher, J. E. & Brangwynne, C. P. Spatiotemporal Control of Intracellular Phase Transitions Using Light-Activated optoDroplets. *Cell* **168**, 159-171 e114, (2017).

52 Li, P., Banjade, S., Cheng, H. C., Kim, S., Chen, B., Guo, L., Llaguno, M., Hollingsworth, J. V., King, D. S., Banani, S. F., Russo, P. S., Jiang, Q. X., Nixon, B. T. & Rosen, M. K. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336-340, (2012).

53 Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59, (1989).

54 Kjaergaard, M., Brander, S. & Poulsen, F. M. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *J Biomol NMR* **49**, 139-149, (2011).

55 Kjaergaard, M. & Poulsen, F. M. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *J Biomol NMR* **50**, 157-165, (2011).

56 Kroenke, C. D., Loria, J. P., Lee, L. K., Rance, M. & Palmer, A. G. Longitudinal and transverse H-1-N-15 dipolar N-15 chemical shift anisotropy relaxation interference: Unambiguous determination of

rotational diffusion tensors and chemical exchange effects in biological macromolecules. *Journal of the American Chemical Society* **120**, 7905-7915, (1998).

57 Wu, D. H., Chen, A. D. & Johnson, C. S. An Improved Diffusion-Ordered Spectroscopy Experiment Incorporating Bipolar-Gradient Pulses. *J Magn Reson Ser A* **115**, 260-264, (1995).

58 Holz, M. & Weingartner, H. Calibration in Accurate Spin-Echo Self-Diffusion Measurements Using H-1 and Less-Common Nuclei. *Journal of Magnetic Resonance* **92**, 115-125, (1991).

59 Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-293, (1995).

60 Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327, (2015).

61 Wang, C., Grey, M. J. & Palmer, A. G., 3rd. CPMG sequences with enhanced sensitivity to chemical exchange. *J Biomol NMR* **21**, 361-366, (2001).

62 McIntosh, P. R. & Freedman, R. B. Characteristics of a copper-dependent cross-linking reaction between two forms of cytochrome P-450 in rabbit-liver microsomal membranes. *Biochem J* **187**, 227-237, (1980).

63 Stejskal, E. O. & Tanner, J. E. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *Journal of Chemical Physics* **42**, 288-+, (1965).

64 Li, W., Reeb, A. N., Lin, B., Subramanian, P., Fey, E. E., Knoverek, C. R., French, R. L., Bigio, E. H. & Ayala, Y. M. Heat Shock-induced Phosphorylation of TAR DNA-binding Protein 43 (TDP-43) by MAPK/ERK Kinase Regulates TDP-43 Function. *J Biol Chem* **292**, 5089-5100, (2017).

65 Bussi, G., Gervasio, F. L., Laio, A. & Parrinello, M. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J Am Chem Soc* **128**, 13435-13441, (2006).

66 Deighan, M., Bonomi, M. & Pfaendtner, J. Efficient Simulation of Explicitly Solvated Proteins in the Well-Tempered Ensemble. *J Chem Theory Comput* **8**, 2189-2192, (2012).

67 Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* **4**, 435-447, (2008).

68 Best, R. B., Zheng, W. & Mittal, J. Balanced Protein-Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. *J Chem Theory Comput* **10**, 5113-5124, (2014).

69 Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Phys Rev Lett* **100**, 020603, (2008).

70 Kim, Y. C. & Hummer, G. Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding. *J Mol Biol* **375**, 1416-1433, (2008).

71 Anderson, J. A., Lorenz, C. D. & Travesset, A. General purpose molecular dynamics simulations fully implemented on graphics processing units. *J Comput Phys* **227**, 5342-5359, (2008).

72 Glaser, J., Nguyen, T. D., Anderson, J. A., Lui, P., Spiga, F., Millan, J. A., Morse, D. C. & Glotzer, S. C. Strong scaling of general-purpose molecular dynamics simulations on GPUs. *Comput Phys Commun* **192**, 97-107, (2015).

73 Kabsch, W. & Sander, C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637, (1983).

74 Bonomi, M., Barducci, A. & Parrinello, M. Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics. *J Comput Chem* **30**, 1615-1621, (2009).

75 Shen, Y. & Bax, A. SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. *J Biomol NMR* **48**, 13-22, (2010).

76 Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132, (1982).

SI Figures
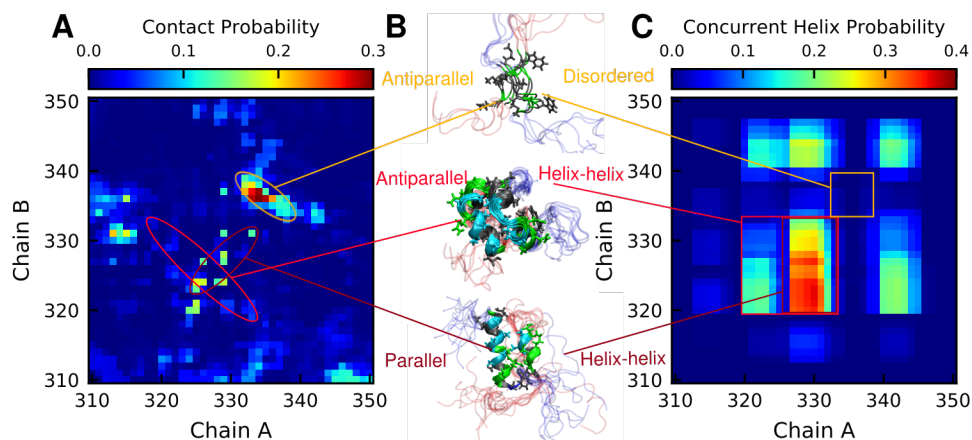
**SI Figure 1: Self-association of TDP-43 CTD vs helical content**

A) Intermolecular contact probability map of two chain simulation of TDP-43$_{310-350}$ calculated by distance between α-carbon atoms of each residue. Cutoff value for contacts is set to 8 Å. B) Clusters of bound configurations highlighting regions that are in contact, and helical with the exception of the disordered contact on the top snapshot. C) Per-residue probability of concurrent helix formation between the two chains, showing which regions of the protein are commonly helical at the same time.
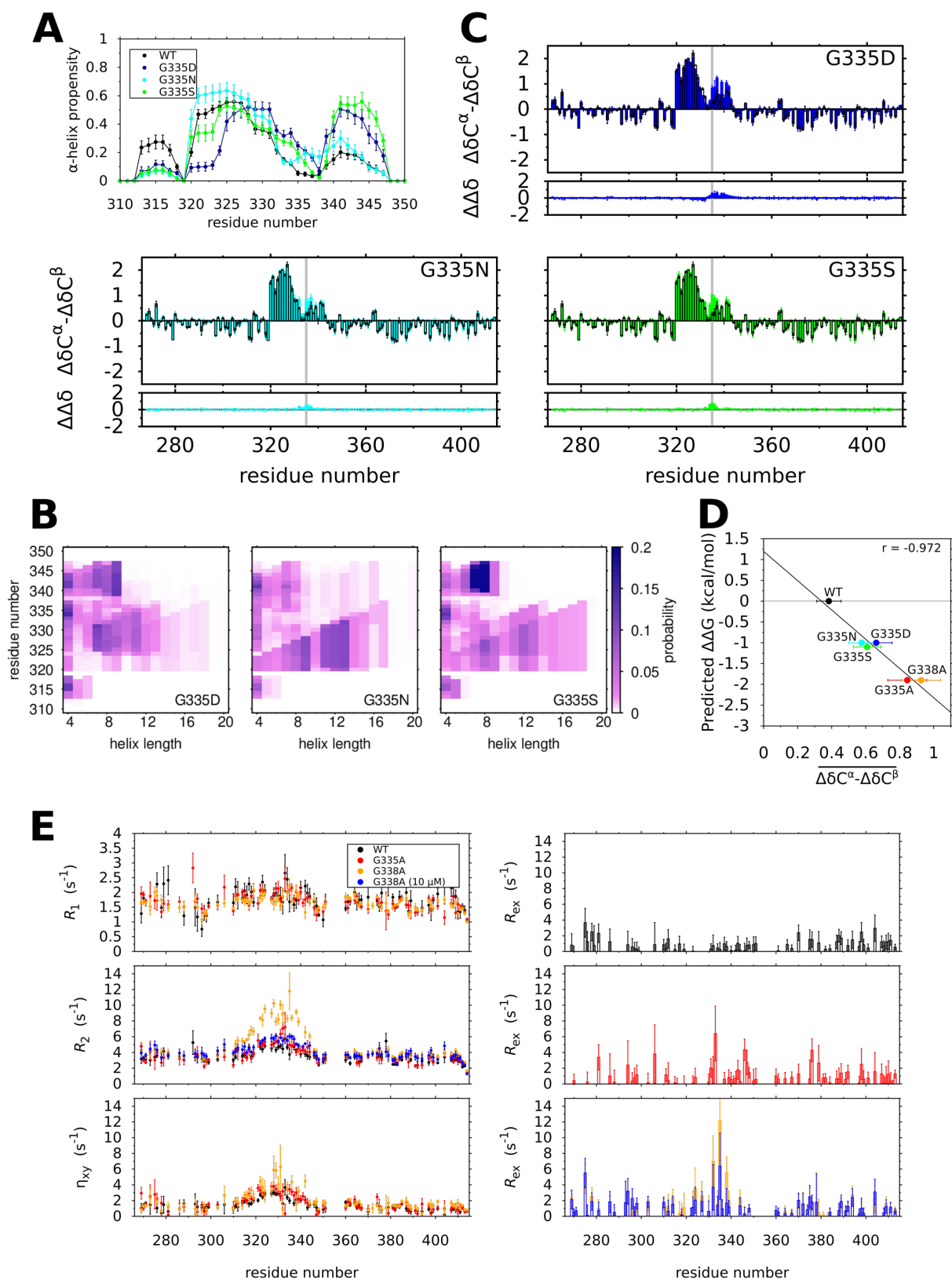
SI Figure 2: Mutations at G335 and G338 enhance local α-helical structure

A) α-helical propensities for WT, G335D, G335N, and G335S CTD$_{310-350}$ determined by all-atom MD

simulation. An unexpected dip for G335D far from the site of mutation (residues 320-324) may be due to a change in net charge of the peptide. Error bars are SEM. B) Secondary chemical shifts ($\Delta\delta C^{\alpha}$ − $\Delta\delta C^{\beta}$) and the differences in secondary shifts with respect to WT ($\Delta\Delta\delta$) for G335D, G335N, and G335S CTD. Error bars are SD. C) Predicted change in free energy $\Delta\Delta G$ for monomeric α-helix formation versus the average change in secondary shift ($\overline{\Delta\delta C^{\alpha} − \Delta\delta C^{\beta}}$) for residues 331-343 of G335 and G338 variants. D) ss-maps for G335D, G335N and G335S TDP-43$_{310-350}$ E) ^{15}N spin-relaxation parameters R_1, R_2, and η_{xy}, as well as calculated R_{ex} values are plotted for WT, G335A, and G338A at 10-20 µM sample concentration. Error bars are SD.
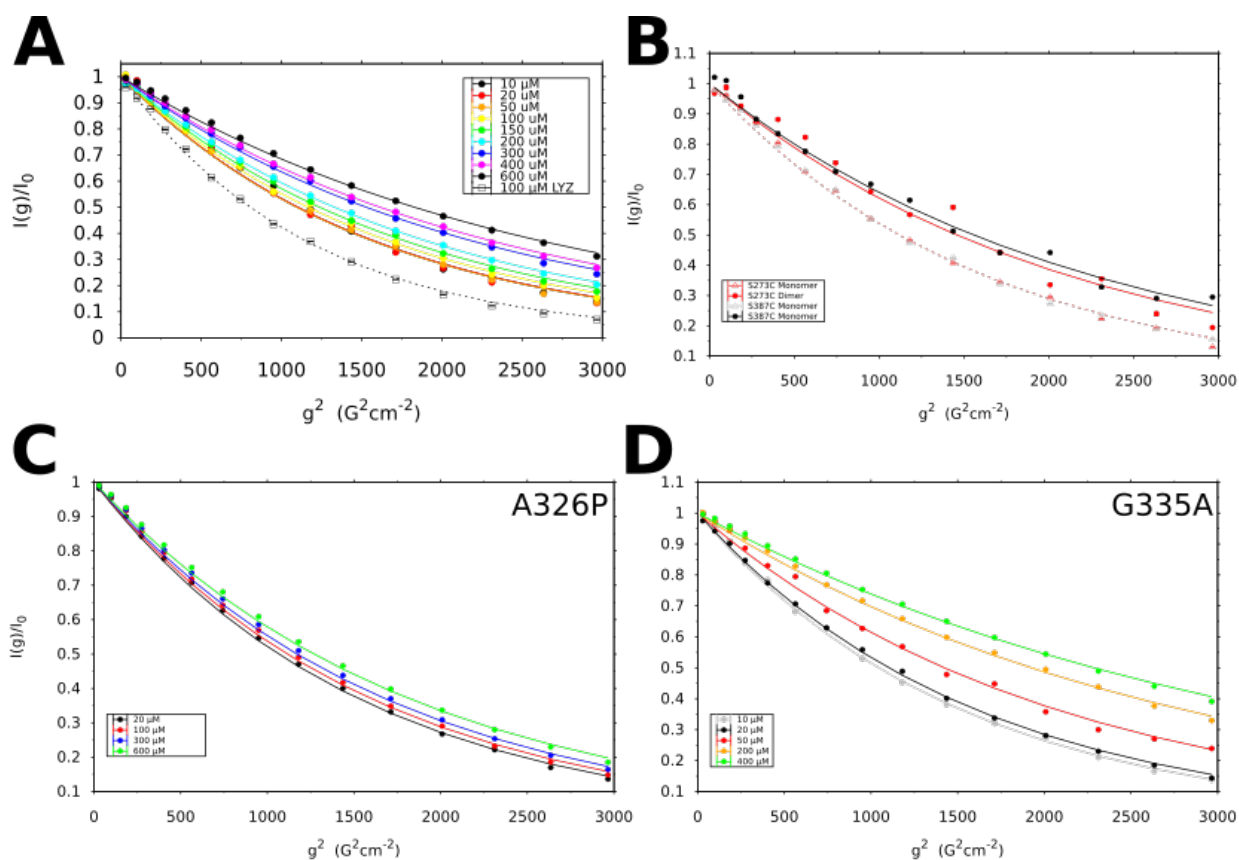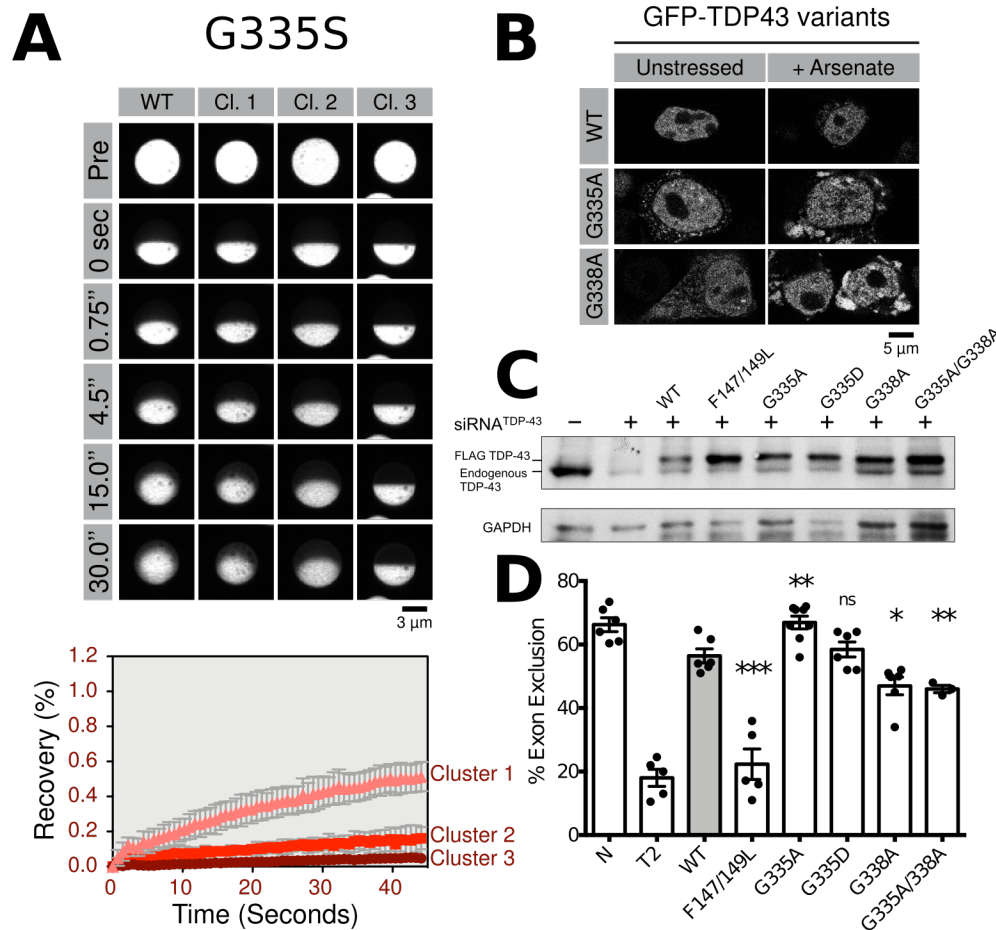

SI Figure 3: Mutations at G335 and G338 enhance intermolecular helix-helix contacts and higher-order assembly

A) Diffusion NMR displayed as normalized resonance intensity ($I(g)/I_0$) versus squared gradient strength (g^2) show that diffusion rate changes rapidly for WT CTD from 150 µM to 300 µM. Lines are best fit to the Stejskal-Tanner equation. Control sample of 100 µM lysozyme (LYZ) shown as black dashed line. Error bars are SD from 100 bootstrapping simulations. Intensity ratios of methyl ^{1}H resonances for monomeric/dimeric S273C and S387C B), A326P C), and G335A D) plotted as a function of squared gradient strength (g^2). Error bars are SD from 100 bootstrapping simulations. The solid and dashed lines represent the best-fit nonlinear-least squares solution of the Stejskal-Tanner equation.
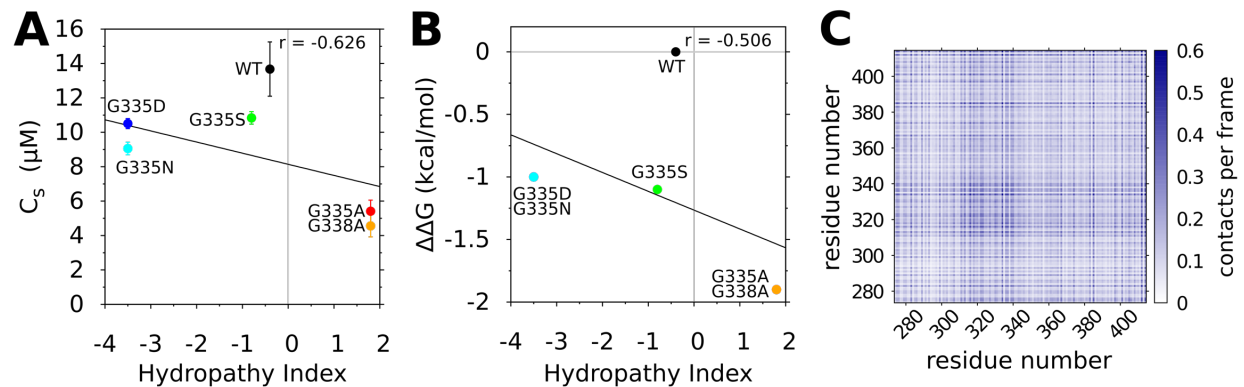
SI Figure 4: Mutations at G335 and G338 alter *in vitro* LLPS

(A) Correlation plot of C_s values at 150 mM NaCl versus hydropathy index based on the scales described by Kyte and Doolittle[76] shows poor correlation, suggesting that changes in residue hydrophobicity of G335 and G338 variants do not explain changes in phase separation. (B) Correlation plot of the change in free energy formation for a given amino acid to adopt monomeric α-helical structure with respect to glycine (ΔΔG) versus hydropathy index. C) Intermolecular contact map for coarse-grained slab simulation of TDP-43 CTD highlights residues 314-340 as having the highest propensity of intermolecular interaction in the equilibrated simulation ensemble.

SI Figure 5: Mutations at G335 and G338 alter in-cell LLPS and splicing function

A) Fluorescence recovery after in cell half-droplet photobleaching for G335S TDP-43$_{RRM-GFP}$ illustrates heterogeneity of G335S droplet fluidity. Observations of TDP-43 reporter droplet fluidities from individual cells were manually grouped by FRAP recovery rates into clusters to demonstrate the heterogeneity that gives rise to large uncertainties for G335S data analyzed without grouping (seen for G335S in Figure 5B). Scale bar is 3 μm. B) Fluorescence microscopy of GFP-TDP-43 WT, G335A, and G338A highlight the formation of cytosolic puncta in mutant variants after arsenate-induced cellular stress. Scale bar is 5 μm. C) Western blot of G335 and G338 variants used in exogenous CFTR exon 9 minigene assay shows protein expression levels between WT and mutant variants that are comparable when compared to GAPDH loading control variations. WT and mutant constructs were expressed in TDP-43 siRNA-treated HeLa cells. D) Percent CFTR exon 9 skipping measured from control and siRNA TDP-43 (T2) treated HeLa, expressing siRNA-resistant WT and mutant constructs. Non-specific siRNA was used as control (N). Statistical significance and SEM calculated for n≥3, using unpaired, two-tailed, Welch-corrected t-test (P *** P = 0.0008, ** P = 0.004, and *P = 0.02).