 doi:10.20944/preprints202209.0442.v1

Article

Different intermolecular interactions drive nonpathogenic liquid-liquid phase separation and potentially pathogenic fibril formation by TDP-43

Yu-teng Zeng[1], Lu-lu Bi[1], Xiao-feng Zhuo[3], Ling-yun Yang[2], Bo Sun[1], Jun-xia Lu[1*]

[1] School of Life Science and Technology, ShanghaiTech University, Shanghai, 201210, China
[2] iHuman Institute, ShanghaiTech University, Shanghai, 201210, China
[3] Shenzhen People's Hospital (The Second Clinical Medical College, Jinan University; The First Affiliated Hospital, Southern University of Science and Technology), Shenzhen, 518020, Guangdong, China
* Jun-xia Lu. lujx@shanghaitech.edu.cn

Abstract: Liquid-liquid phase separation (LLPS) of proteins has been found ubiquitously in eukaryotic cells, critical in the controlling of many biological processes through forming a temporary condensed phase with different bimolecular components. TDP-43 is recruited to stress granules in cells and is the main component of TDP-43 granules and proteinaceous amyloid inclusions in patients with amyotrophic lateral sclerosis (ALS). TDP-43 low complexity domain (LCD) is able to demix in solution forming the protein condensed droplets. The molecular interactions regulating its LLPS were investigated at the protein fusion equilibrium stage, where the droplets stopped growing. We found the molecules in the droplet were still liquid-like but with enhanced intermolecular helix-helix interaction in the LCD. The protein would start to aggregate after about 200 minutes of lag time and aggregate slower than at the condition when the protein does not phase separate or the molecules have a reduced intermolecular helical interaction. A structural transition intermediate towards protein aggregation was also discovered involving a decrease of the intermolecular helix-helix interaction and a reduction in the helicity. Therefore, LLPS and the intermolecular helical interaction could help maintain the stability of TDP-43 LCD.

Keywords: TDP-43; Liquid-liquid phase separation; Solution-state NMR

1. Introduction

Eukaryotic cells contain micro-scale compartments that are formed by specific proteins and nucleic acid through condensation. These compartments do not have bound membranes and are referred as membraneless organelles, regulating diverse processes in cells [1, 2]. Examples include stress granules formed when translation initiation is impaired in response to cellular stresses [3], P granules of C. elegans embryos [4], and Cajal bodies [5], etc. The dynamic equilibrium between condensation and dissolution of these membraneless organelles are mainly controlled by the multivalent but weak interactions of constituting proteins and other biomolecules within the organelles [1]. Proteins with intrinsic disordered low complexity domain (LCD) are often the major components forming these membraneless organelles [6, 7]. In vitro, these proteins with multivalency could demix from solution and form liquid droplets through a protein liquid-liquid phase separation (LLPS) process, providing us a simplified system to understand the physical and chemical forces driving the protein condensation.

TDP-43 (transactive response (TAR) element DNA-binding protein of 43 kDa) is a nuclear ribonucleoprotein. It participates in many processes of RNA regulations and is able to autoregulate its own expression [8]. It is recruited to stress granules in cytoplasm which contains other proteins with LCD such as GTPase activating protein1 (SH3 domain) (G3BP1), T cell intracellular antigen-1 (TIA-1) etc. [3]. Although TDP-43 is not an obligatory stress granule component, it modulates stress granule formation and disassembly [9].

 doi:10.20944/preprints202209.0442.v1

The full-length TDP-43 protein contains 414 residues, with a nuclear localization signal (NLS) and two RNA-recognition motifs at the N-terminal region but a low complexity domain (LCD) at the C-terminal region (267-414) [10]. The first RNA-recognition domain and the LCD are responsible for recruiting TDP-43 into stress granules [8]. In solution, TDP-43 LCD domain [11, 12] is able to form liquid droplets by phase separation. TDP-43 LCD is glycine-rich and prion-like [8]. It can aggregate into amyloids in common solutions and in cells [13]. And the TDP-43 aggregation is considered as the hallmark for amyotrophic lateral sclerosis (ALS) [9, 14]. The protein inclusions are also sometimes found in patients with frontotemporal dementia (FTD) and other neurodegenerative diseases [8]. And those diseases related to the formation of the pathological TDP-43 granules and the abnormal TDP-43 aggregation are considered as TDP-43 proteinopathies. Recent reports indicate that acute recruitment to RNA-containing stress granules maintains TDP-43 solubility [15, 16]. But with prolonged stress, protein aggregation may also be found forming from the stress granules [15].

Most ALS-associated mutations of TDP-43 are located at LCD, indicating the importance of this domain [17]. The mechanisms determining TDP-43 LCD LLPS and the interplay between LLPS and aggregation have been extensively investigated but questions remain. For example, the intrinsic disordered LCDs involved in LLPS are usually composed of polar amino acids punctuated by aromatic residues [18]. The TDP-43 LCD sequence is also like that (table S1a, b, c). But the TDP-43 LCD contains one secondary structure element of two short α-helices connected by a bend (residue 320-343) [13]. TDP-43 LCD sequence contains a higher amount of hydrophobic residues (table S1a), with the helices mostly hydrophobic (table S1c). The intermolecular helix-helix interactions driven by the hydrophobic interactions have been shown as the major forces for TDP-43 phase separation. But this segment also belongs to the amyloidogenic core domain. The droplets TDP-43 LCD forms tend to be relatively small, about 1 μm in diameter in certain conditions (in our hand and also observed by other groups [6, 12]), different from other phase separation proteins [6]. It is also not easy for us to find the condition to observe the ongoing liquid droplet fusion events by differential interference contrast (DIC) microscope. These indicate that the protein fusion equilibrium occurs quickly after the protein is dissolved in solution when the liquid droplet is very small.

So what determines TDP-43 LCD fusion equilibrium? Besides the helical intermolecular interactions, other interactions have also been found important for TDP-43 LCD LLPS. The two flanking regions on both sides of the helices are the intrinsic disordered region (IDR), referred to as IDR1 (276-320) and IDR2 (343-414). IDR2 contains a Gln/Asn-rich (QN) (344-360) motif immediately after the helices (320-343), with the sidechain amide bonds of Q, N residues able to participate in the H-bonding promoting the protein assembly or the amyloid formation [19-21]. Both IDR1 and IDR2 regions contain X-G/S, G/S-X sequence motifs (X represents the aromatic residues Phe, Trp, Tyr). The sidechain H-bonding provided by hydroxyl group of serine, the π-π stacking of aromatic rings or ionic charge-π interactions are all believed to contribute to the intermolecular interactions in protein phase separation [18, 22, 23]. For example, TDP-43 contains three Trp residues. All three Trp residues were found playing roles in regulating the protein LLPS by mutation studies, with the most important Trp located in the helical region [12]. TDP-43 LCD also contains several charged residues including 5 Args (table S1) with a PI (isoelectric point) value of 10.78, and pH of the solution controls its phase separation ability. A decrease of pH to 4-5 could diminish its phase separation completely while increasing pH of solution to 6 and above would cause protein droplet formation [24, 25]. Increasing NaCl concentration also increases the protein LLPS by providing the electrostatic shielding effects [11, 25]. Both the pH and salt effect indicate that the electrostatic repulsion inhibits its phase separation [26]. Therefore, there must exist an intricate balance between different interactions, which fine tune the status of TDP-43 in solution. However, the relative contributions of these interactions to TDP-43 LCD LLPS were not well understood.

In this research, we investigated the special role of the helical region (320-343) and compared it to the rest of the sequence to understand how the balance of the contributions

 doi:10.20944/preprints202209.0442.v1

from these two parts of sequence controlled the LLPS and what happened when the TDP-43 LCD stopped fusion. A structural transition intermediate towards aggregation was also discovered involving a decrease of the intermolecular helix-helix interaction and a reduction in the helicity. Using TDP-43 LCD as a model, our work provided a better understanding of the multivalence controlling the intermolecular interactions and protein phase separation.

2. Results

2.1. TDP-16 phase separation is very sensitive to the solution conditions

2.1.1. TDP-43 LCD droplets fusion events can only be observed in a narrow concentration range.

The liquid-liquid phase separation (LLPS) of TDP-43 LCD (residue 267-414) was first investigated in pH 6.0 10 mM phosphate buffer (PB). A 100 μM protein solution was prepared from dry protein powder. And the droplets were about or less than 1 μm in diameter as shown in figure 1a. The turbidity of the solution was determined using OD600 about 0.47 (figure S1a). A high resolution image was also taken using TEM, where the stained image showed many black spheres with a diameter 1 μm or less (figure S1b). Interestingly, we were not able to detect a very clear droplet fusion event (a short movie in VideoS1), indicating the active fusion has already finished before we can observe. The droplets were able to stick to each other, but remained as separate particles. A controlled fusion of droplets was therefore carried out using optical tweezers to check the fusion ability of the droplets (snapshots shown in figure 1a bottom panel and the video shown in VideoS2). In the video, two droplets were forced to touch each other for as long as 260s (Force profiles shown in figure S1c), but failed to fuse into one droplet. The results confirmed that the droplets formed at this condition had a low capability to fuse at this stage. We referred to these droplets as mature droplets.

Droplets formed at other protein concentrations were further tested. A higher concentration protein solution (200 μM, figure S2a, VideoS3a) didn't show clear droplet fusion, either. And a lower concentration protein solution (20 μM, figure S2b, VideoS4) only displayed some fusion events. Interestingly, the OD600 readings were low (~0.2) for both 20 μM and 200 μM concentrations (figure S2c), indicating the solution turbidity was not positively correlated to the protein concentration. The turbidity was influenced by a combination effect of particle numbers and sizes in the solution. At 200 μM concentration, the solution looked much clearer than that at 100 μM, and became more turbid once diluted to 100 μM concentration. A video showing the dilution process was displayed in VideoS3b. It took 1-2 min for the buffer to diffuse and a change of equilibrium to take effect, showing more and bigger droplet formation upon dilution. The fusion event was best observed at 40 μM concentration (figure 1b, VideoS5). The protein liquid droplets were also bigger at this condition, about 2-4 μm in diameter.

2.1.2 RNA enhanced TDP-43 LCD droplet fusion

The phase separation property was then investigated in the same PB condition with additional RNA (a final concentration of 20ng/μl). Bigger droplets were formed with the diameter ranging from 1 to 3 μm (figure 1c). TEM also displayed many black spheres (figure S1b). Figure 1c lower panel and Video S6a showed the fusion events. Video S6b showed the process of adding yeast RNA to protein solution in real time. It took 3 min for the RNA to show its effect and a change of equilibrium was displayed by the size of the droplet growing overtime. Therefore, adding RNA to TDP-43 LCD solution would increase the size of the droplets and significantly alter the protein LLPS. The OD600 was about 0.51 (figure S1a), a little bigger than the reading without RNA. The fusion events were also observed for 40 μM TDP-43 LCD solution in the presence of RNA (figure 1d, video S7). Bigger droplets were also observed (1-5 μM in diameter).

 doi:10.20944/preprints202209.0442.v1

2.2. The intermolecular interactions of TDP-43 LCD monitored by solution NMR

2.2.1 TDP-43 LCD in mature droplets is still mobile and liquid-like

The droplets formed by TDP-43 LCD had very little ability to fuse at the concentration of 100 μM in PB quickly after dissolution. ^{1}H-^{15}N HSQC spectrum of freshly prepared TDP-43 LCD (100 μM) in PB was therefore obtained to investigate the molecular status of this condition. The protein will slowly aggregate at this solution condition. The protein aggregation was monitored using THT fluorescence. Figure S3a indicated that the THT fluorescence remained at a low level for about 200 minutes before it raised up. THT tended to bind to protein aggregates with amyloid properties, giving out fluorescence. Therefore, the result suggested a lag time of at least 200 minutes before the protein started to aggregate. The spectrum was taken and finished within the lagging period (figure 2a). The spectra chemical shift assignment was based on the reported work with BMRB accession code: 26823, which was carried out on 20 μM TDP-43 (residue 267-414) in pH 6.1, 10% D2O 20 mM MES buffer [11]. Table S2 compared four reported chemical shifts including BMRB 26823, indicating the chemical shifts were slightly different depending on the sample conditions. Our results matched with BMRB 26823 the best. Table S3 included all the chemical shifts from ^{1}H-^{15}N HSQC obtained for different conditions in our work.

The peak intensity of TDP-43 LCD in PB buffer was shown in figure 2b, displaying a significant signal decrease of the helical region at residue 320-343 (in gray) compared to the rest of the sequence. This was consistent with previous research reports that the helix was involved in the intermolecular interactions, and the dynamic equilibrium caused the decrease of the NMR signal intensity in this region [11, 23]. Besides the helical region, we also observed signal intensity decrease at several places in IDR1 and IDR2 labeled in yellow in figure 2b, indicating those regions were also involved in the intermolecular interactions. These regions were mostly co-localized with the aromatic residues with X-G/S, G/S-X (X represented the aromatic residues Phe, Trp, Tyr, shown in green in the sequence). The observation was also reported by Pantoja-Uceda et. al. on TDP-43 LCD at pH 4.0 [23]. The regions with decreased intensity were generally matched with their reports but differences existed, since the solution condition was very different. The intermolecular interactions provided by these regions could probably be simplified using the so-called "stickers-and-spacers" model [18]. The regions were also overlapped with some of the peptides named as LARKS (low-complexity aromatic-rich kinked segments) by Eisenberg et.al [27].

It was reported the MES buffer in pH 5.5 disfavored the phase separation [24], therefore, the protein molecules were supposed to be mostly in a low level of aggregation or liquid-like state. 70 μM of TDP-43 LCD in pH 5.5 20 mM MES buffer was prepared. A DIC microscopic image was displayed in figure 3a (a video was shown in video S8), showing very few protein droplets. OD600 reading was close to the blank solution (figure 3b). Our observation confirmed a low level of protein LLPS in the MES buffer. In order to investigate whether the molecules in the mature droplet in PB were in liquid-like state, the ^{1}H-^{15}N HSQC spectrum of the TDP-43 LCD (70μM) in pH 5.5 MES buffer was obtained for comparison. The chemical shift assignment was again based on the report with BMRB accession code: 26823, which was on a 20μM protein MES buffer with a higher pH (6.1). The ^{15}N chemical shift difference between these two conditions were minor within ±0.05ppm (figure S3b). Figure S3c showed the original peak intensities for TDP-43 LCD in MES, displaying the similar intensity variations at those regions shown for PB buffer condition. The signal intensity ratio between the two conditions was displayed in figure 3c. The residue peak intensity in the MES buffer was around 0.6-0.7 of that in the PB buffer for most residues except a few showing higher intensity for MES. The intensity level was consistent with a lower concentration of protein in MES (70μM in MES vs 100μM in PB buffer) for the NMR studies, indicating no aggregation of protein in PB. The residues at 320-343 representing the helical region of TDP-43 LCD showed a consistent higher intensity ratio than the other region in figure 3c, which supported the conclusion that the helical intermolecular interactions were reduced in MES. Other higher intensity peaks in MES were labeled in figure 3c and underlined in the sequence in figure 2b, which scattered

 doi:10.20944/preprints202209.0442.v1

throughout the sequence. The high intensity indicated the molecular interactions involving those residues were significantly decreased in MES. The 15N chemical shift changes were shown in figure S3d. A short region at the N-terminus and residue 320-343 displayed bigger negative shifts (up to -0.15ppm) when the buffer was changed from MES to PB. A negative change of amide ^{15}N chemical shift suggested an increased helical structure for the helical region in PB condition [11]. Both amide peak intensity and ^{15}N chemical shift changes, therefore, supported the helix structure at 320-343 was involved in stronger intermolecular interactions in PB. But the protein was still liquid-like without much decrease in the NMR peak intensity for most of the residues in the mature droplets.

2.2.2 The helical intermolecular interaction was reduced during TDP-43 LCD fibrillation in PB

TDP-43 LCD started aggregation after about 200 mins in PB (figure S3a). The aggregation process was monitored using ^{1}H-^{15}N HSQC to see the changes at the residue level. Figure 2c (up) showed the changes in the spectra intensity with time at 100 μM protein concentration in PB buffer, comparing the spectra collected at the 1st hour (1st spectrum, figure 2a), 6th hour, 22nd hour and 53rd hour, respectively. Figure 2d panel displayed the peak intensity ratio, compared all three spectra to the 1st one, showing a consistent decrease in the intensity for most of the residues during the time. However, different changes were again observed around the helical region (residue 312-343, with the additional segment 312-319). At the 6th hour and 22nd hour, the peak intensity at residue 312-343 increased instead of decreased. This indicated the dynamic exchange caused by the helical intermolecular interactions were decreased at the beginning stages of the protein aggregation. The rest of the sequence (IDR1, IDR2), however, had a consistent decrease in the intensity about 20-30%, supporting the protein in slow aggregation. After that, at the 53rd hour, the last spectrum showed a significant decrease of about 90% for all the residues. The ^{15}N chemical shift changes were compared between the 22rd hour and the 1st hour spectra, shown in figure 2c (down) with bigger positive changes (up to 0.13 ppm) mainly at the helical region. This supported a decrease in helicity during the aggregation. Therefore, a decrease of the intermolecular interaction mediated by the helices would be in the pathway for TDP-43 LCD fibrillation. The fibrillation initiated at the IDR1, IDR2 regions since the signal of these regions decreased first.

The protein aggregation with 70 μM concentration in pH 5.5 MES was also studied, showing faster aggregation with only about 150 mins' lag time indicated by the THT fluorescence (figure 3d). The changes of ^{1}H-^{15}N HSQC spectra intensity were also faster. In 22 hours, the intensity for all residues decreased about 50% (figure 3e). Since the helical intermolecular interaction was weaker at this condition, the faster aggregation supported our conclusion that a decrease of the intermolecular interaction mediated by the helices facilitated the protein fibril aggregation. LLPS could be a protective state of the protein and reduce the protein aggregation to some degree.

2.3 TDP-43 LCD aggregation was enhanced by an introduction of 150 mM urea to the solution

In order to further perturb the intermolecular interactions of protein molecules, 150 mM urea was added to TDP-43 LCD solution in PB. ^{1}H-^{15}N HSQC experiment was carried out to investigate how urea changed the protein at the molecular level (figure 4a). This sample showed smaller and less homogeneous spheres observed by DIC microscopy (figure 4b). OD600 reading was about 0.34 (figure S1a), lower than that in PB. We did not observe very clear droplet fusion events (Video S9). Urea is a protein denaturing agent. It would probably change the protein structures and decrease the intermolecular helix-helix interactions of TDP-43 LCD, reducing the protein fusion ability. However, the protein displayed much faster fibril growth as shown in figure S3a, with about 2 hours' lagging time (120 mins).

The ^{1}H-^{15}N HSQC spectra intensities were summarized in figure 4c (up), comparing the two conditions with or without urea. It showed a slight decrease in the intensity for non-helical regions upon adding 150 mM urea. But it also displayed a slight increase in

 doi:10.20944/preprints202209.0442.v1

the intensity for the helical region, supporting the assumption that the helical intermolecular interaction was perturbed in the presence of urea. The changes of ^{15}N chemical shift were shown in figure 4c (down), indicating positive shifts (up to 0.15 ppm) around the helical region. It suggested a decrease in helicity. Therefore, urea decreased the protein intermolecular interactions and disturbed the helical structure of TDP-43 LCD. Both effects would negatively impact the protein LLPS.

^{1}H-^{15}N HSQC spectrum at the 22nd hour was also collected and the intensity was shown in figure S3e. It showed that the peak intensity in the helical region decreased slightly but the rest of the sequence displayed bigger decreases. The ratios of the intensity changes were shown in figure 4d and compared to that in PB buffer without urea. It showed a bigger decrease in the intensity about 40% with urea and about 30% without urea. This was consistent with the faster fibrillation of TDP-43 with urea shown in figure S3a. During the aggregation up to 22 hours, the helical region showed a little decrease in the intensity in the presence of urea, but an increase in the intensity without urea. As we discussed above, the fibrillation would involve an intermediate step of a partial release of the intermolecular interaction mediated by TDP-43 LCD helices, and a loosening of the helix structure. Both were already introduced upon adding urea at the time the sample was prepared. Therefore, only small changes in the intensity at the helix region were observed after 22 hours in the presence of urea. The fibrillation was enhanced since the solution already went through this intermediate step upon adding urea.

2.4 The disordered region showed less LLPS ability and a slower aggregation

2.4.1 TDP-43 LCD fragments without the helices have a low LLPS ability

TDP-43 LCD is intrinsically disordered except for the short helical region. The α-helical segment mediates the intermolecular interactions and promotes the phase separation. However, our work also showed the rest of the sequence, the IDR, contributed to the intermolecular interactions. In order to compare the contributions of the two parts in determining the protein phase separation properties, we replaced the α-helical segment with two sequences (figure 5a). One was $(EAAAK)_3$, which was a designed helical structure, but did not promote the intermolecular helical-helical interaction [28] (the protein was labeled as TDP-16E). The other one was (GGGGS), which was designed to be a flexible linker to connect the two flanking IDR domains (The protein was labeled as TDP-16G. The wild-type TDP-43 LCD was labeled as TDP-16 in figure 5). The protein solutions were prepared in the PB buffer in 20 μM concentration. As shown in figure 5b, the turbidity of the wild-type protein solution (OD600=0.18) is significantly greater than that of the mutant solutions (OD600=0.04). DIC images (figure 5c, left panel) also confirmed the turbidity reading that only the wild-type solution displayed the protein droplets. The results confirmed the sequence of the α-helical segment has a dominant role in determining the phase separation of TDP-43 LCD, much stronger than the two IDRs.

2.4.2 TDP-43 LCD fragments without the helices still form fibril aggregates

Although TDP-16E and TDP-16G have a low LLPS ability, a long incubation at room temperature would also lead to the protein aggregation. The TEM images of the protein aggregates after 4 days' incubation were shown in figure 5c, middle panel. They also displayed fibril-like images. And the x-ray diffraction of the collected aggregates was shown in figure 5c, right panel. All three displayed two diffraction rings at about 4.7 Å and 10.0 Å, supporting the formation of amyloid fibrils. THT fluorescence was applied to monitor the aggregation process (figure 5d). The TDP-43 LCD had a lag time of about 200 min before the fluorescence started to increase. TDP-16G had a lag time of about 250 min, then a quick increase of THT fluorescence. The absolute intensity of THT fluorescence was not compared since different fibrils would emit fluorescence differently. TDP-16E aggregates were not sensitive to THT binding, without significant fluorescence reading. Therefore, the intrinsic fluorescence of the protein was also monitored to observe the aggregation of TDP-43 LCD and its mutants (figure 5e). During the protein aggregation, the intrinsic fluorescence was decreased. The experiments indicated a faster intensity decrease of wild-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2022 doi:10.20944/preprints202209.0442.v1

type TDP-43 LCD than the other two. The results indicated that the disordered regions of TDP-43 LCD also can form fibrils but much slower without the helical segment.

2.5 RNA enhanced TDP-43 LCD LLPS mainly by mediating the intermolecular interactions with IDR1 and IDR2

In order to understand how RNA interacted with TDP-43 LCD and alters the LLPS, we obtained ^{1}H-^{15}N HSQC spectrum of TDP-43 LCD (100 μM) in the presence of 20ng/μl RNA (figure 6a) and compared it to the protein spectrum in PB buffer. The protein aggregation in the presence of RNA showed the similar lag time (~200 mins) as the PB buffer condition (figure S3a), and the spectrum was taken within the lag time. The signal intensity was displayed in figure 6b for all the residues. It showed a consistent signal decrease for all the residues except the helical region, clearly indicating RNA affected the disordered region mostly. The intensity ratio between with and without RNA was higher close to 1.0 around the helical region, indicating the helical region was less affected by RNA binding (figure 6b right). The changes of ^{15}N chemical shift to the positive direction (up to 0.12 ppm) centered around the helical region, indicated a decrease of the helicity (figure 6b bottom).

In order to further demonstrate the RNA effect on protein LLPS, different amounts of yeast RNA (0-40ng/μl) were added to TDP-16E (100 μM). It showed a consistent increase in OD600 when RNA concentration was increased (figure 6c). DIC microscopy also displayed more and bigger protein droplets formation and clear droplet fusion events (figure 6d, videoS10). Without RNA, TDP-16E showed little LLPS in the PB buffer. It showed the disordered regions IDR1 and IDR2 had a very weak power in inducing LLPS compared to the helical region. However, the addition of RNA brought enhanced interactions with the IDR, which compensated for the lack of the intermolecular interactions provided by the helices.

2.6 The dynamic changes of protein molecules probed by ^{19}F NMR

^{19}F NMR has been used to study protein aggregation for its high sensitivity. It allowed the detection of the aggregation intermediates for Aβ [29] and other amyloids [30]. The TDP-43 LCD has three Trp residues, with one located at the helical region (Trp334) and two at IDR2 (Trp385, Trp412). The three Trp residues play roles in modulating protein LLPS and Trp334Gly mutation reduced the protein LLPS most significantly [12]. The three Trp residues were replaced by 5-fluoro-tryotophan, thus the change of sidechains of the three Trp residues was observed. Four conditions were compared including 40 μM protein in pH 5.5 20mM MES buffer, 40 μM protein in pH 6.0 10mM PB buffer, 40 μM protein with 20ng/μl RNA in pH 6.0 10mM PB buffer, and 100 μM protein in pH 6.0 10mM PB buffer (figure 7). The freshly prepared sample showed only one dominant ^{19}F peak at -125.0 ppm for all conditions, but clearly with difference. Protein in MES showed the narrowest (< 0.1 ppm linewidth at the half height) and highest peak, while the addition of RNA to the solution significantly reduced the intensity of the peak. Both 40 μM and 100 μM protein solutions in PB showed similar spectra, with a similar intensity and a broad shoulder peak at -124.75 ppm. The signal intensity in PB was slightly weaker than that in MES. The results confirmed that in the MES buffer, the three Trp residues were very dynamic and with a similar chemical environment, supporting a low LLPS in the MES buffer. The molecular dynamics were reduced for PB buffer condition, especially for 100 μM protein concentration, causing the weaker signals. The PB buffer condition was LLPS active condition, the shoulder peak for 40 μM protein concentration was mostly from a contribution of Trp334 at the helical region since the helical interaction was the most dominant effect inducing LLPS. The remaining sharp peak at -125.0 ppm was not attenuated significantly for 40 μM condition, but was attenuated significantly for 100 μM protein concentration considering 2.5 times protein concentration. The observation indicated stronger intermolecular interactions involving the helices and the other regions as well. Three Trp residues were least dynamic in the presence of RNA. It supported a stronger interaction between RNA and IDR in promoting LLPS. Although clear difference was observed here

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2022 doi:10.20944/preprints202209.0442.v1

on the freshly prepared samples (40 μM protein concentration) in PB with or without RNA, the DIC images showed active LLPS for both.

^{19}F NMR spectra also changed as a function of time. The peak intensity at -125.0 ppm was attenuated clearly for the two PB buffer conditions without RNA. Additional peak at -119.8 ppm slowly showed up, but remained very weak for all three PB solutions. This peak was likely a peak for the protein oligomer intermediate during the aggregation, but was not investigated here.

2.7 Figures

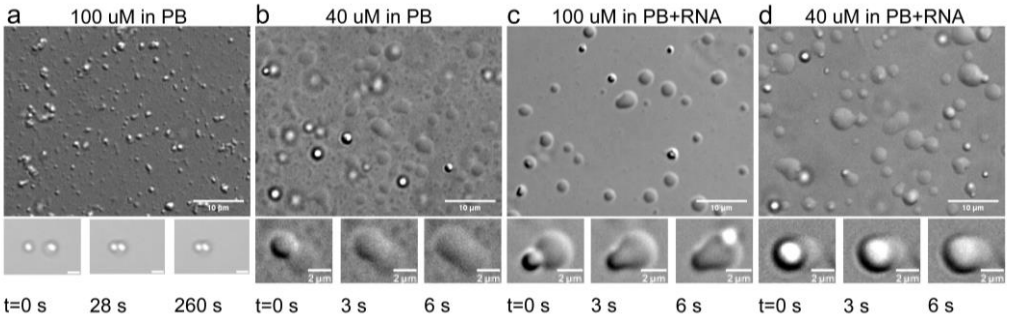

Figure 1. Images of TDP-43 LCD LLPS in four different solution conditions by DIC microscopy. (**a**) 100 μM in PB, 100 μM protein concentration in pH 6.0 10 mM phosphate buffer; (**b**) 40 μM in PB, 40 μM protein concentration in the same PB buffer; (**c**) 100 μM in PB+RNA, 100 μM protein concentration in pH 6.0 10 mM phosphate buffer with yeast RNA (20ng/μL); (**d**) 40 μM in PB+RNA, 40 μM protein concentration in the same PB buffer with RNA. Scale bars are 10 μm. Representative events of TDP-43 LCD droplet fusion were shown at the bottom panel with the scale bar of 2 μm.

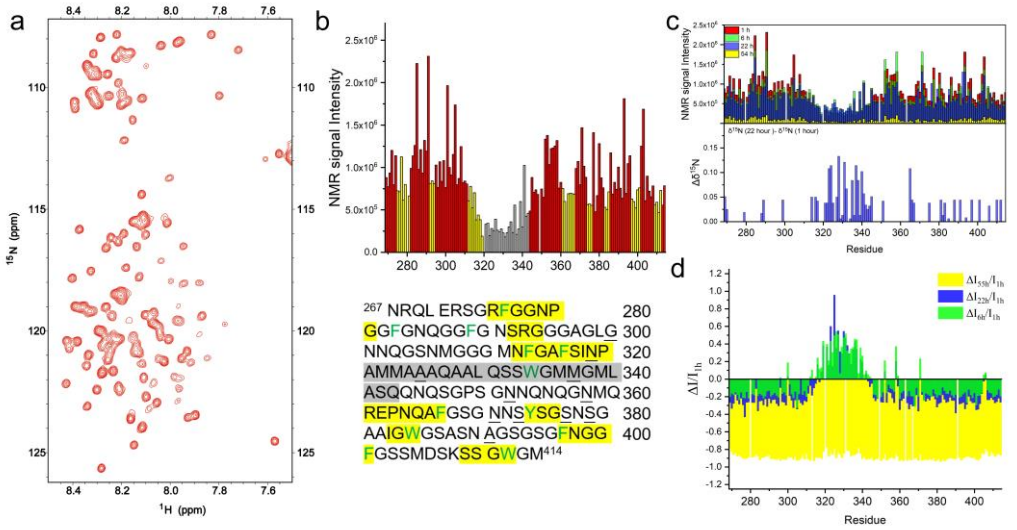

Figure 2. The ^{1}H-^{15}N HSQC spectra of TDP-43 LCD. (a) The ^{1}H-^{15}N HSQC spectra of TDP-43 LCD obtained within 2.5 hours after the sample preparation at a concentration of 100 μM in pH6.0 10mM phosphate buffer and 25℃. (b) NMR peak intensity for each residue obtained from ^{1}H-^{15}N HSQC shown in (a). Helical region of the TDP-43 LCD (in gray) exhibited the lowest signal intensity, while several short segments (in yellow) also exhibited mild signal loss. The amino acid sequence of TDP-43 LCD was shown at the bottom panel with the corresponding highlighted sequences in gray and in yellow. The aromatic amino acids were in green, and the underlined residues were those showing higher intensity in the pH 5.5 MES buffer. (c) (up) Comparison of the NMR signal intensity from ^{1}H-^{15}N HSQC spectra of TDP-43 LCD (100 μM) in pH6.0 10mM phosphate buffer, obtained within 2.5 hours from the sample preparation (labeled as I_{1h} using the approximate starting time, in red) and obtained at ~6 hours (I_{6h}, in green), 22 hours (I_{22h}, in blue) or 53 hours (I_{55h}, in yellow) from the sample preparation. (down)^{15}N chemical shift differences ($\Delta\delta^{15}$N) during TDP-16 fibrillation in PB. $\Delta\delta^{15}$N= δ^{15}N (22nd hour spectrum in PB)- δ^{15}N (1st hour spectrum in PB). (d) NMR signal loss ratio

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2022 doi:10.20944/preprints202209.0442.v1

($\Delta I/I_{1h}$) over time, where ΔI represents the NMR signal intensity after ~6h, 22h or 53h (I_{6h}, I_{22h} or I_{55h}) minus I_{1h}.

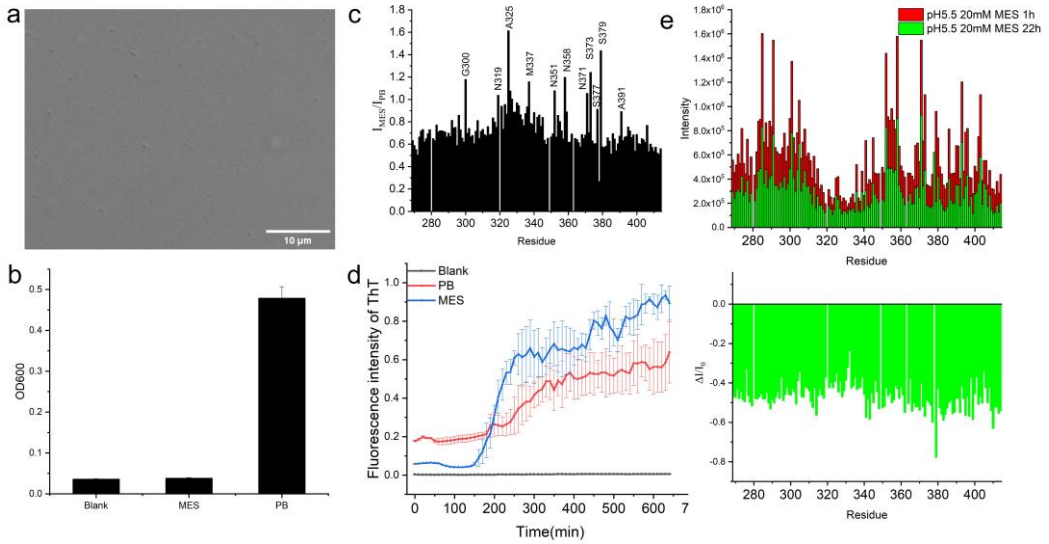

Figure 3. Experiment results on 70 μM TDP-43 LCD in pH 5.5 MES buffer, the LLPS- condition. (**a**) DIC image of TDP-43 LCD in MES buffer. (**b**) OD600 value of 70 μM TDP-43 LCD in pH 5.5 MES buffer (MES) and compare it to the blank MES buffer and 100 μM TDP-43 LCD in pH 6.0 PB buffer (PB). (**c**) Compare the NMR signal intensities of ^{1}H-^{15}N HSQC of TDP-43 LCD between the LLPS+ condition (PB) and the LLPS- condition (MES). (**d**) ThT fluorescence changes of TDP-43 LCD in two different conditions (PB vs MES). (**e**) (up) Comparison of the NMR signal intensity from ^{1}H-^{15}N HSQC spectra of TDP-43 LCD in LLPS- condition in MES (labeled as I_{1h} using the time approximate in the middle of the experiment, in red) and obtained at ~22 hours (I_{22h}, in green) from the sample preparation. (down) NMR signal loss ratio ($\Delta I/I_{1h}$) over time, where ΔI represents the NMR signal intensity after 22h minus I_{1h}.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2022 doi:10.20944/preprints202209.0442.v1

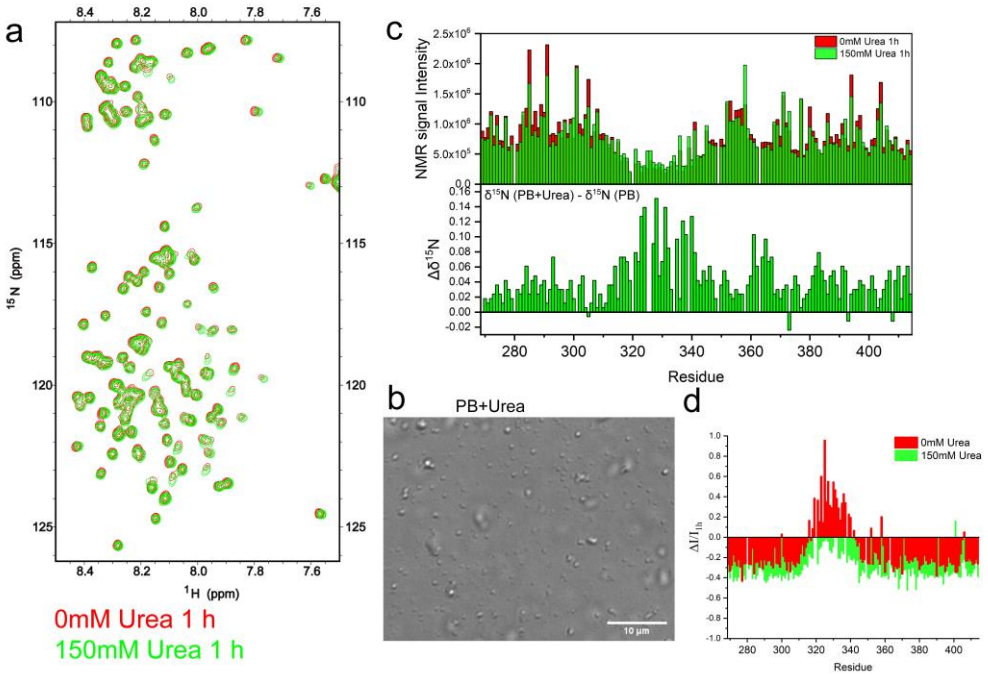

Figure 4. Effect of 150mM urea on TDP-43 LCD liquid-liquid phase separation and protein aggregation. (**a**) The ^{1}H-^{15}N HSQC spectra of TDP-43 LCD obtained within 2 hours from sample preparation at a protein concentration of 100 μM in pH6.0 10mM phosphate buffer with 150mM urea (in green) or without urea (in red). The temperature is 25℃. (**b**) DIC image of 100 μM TDP-43 LCD in PB buffer with 150mM urea. (**c**) (up) Comparison of the NMR signal intensity from ^{1}H-^{15}N HSQC spectra in (a) with (green) or without (red) urea. (down)^{15}N Chemical shift changes ($\Delta\delta^{15}$N) of TDP-43 LCD (100μM) upon adding 150mM urea in pH6.0 10mM phosphate buffer. $\Delta\delta^{15}N = \delta^{15}N$ (PB+Urea) - δ^{15}N (PB). (**d**) NMR signal loss ratio ($\Delta I/I_{1h}$) over time. $\Delta I = I_{22h}$ (NMR signal intensity after ~22hours) - I_{1h}. Samples contained 150mM urea shown in green or without urea shown in red.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2022 doi:10.20944/preprints202209.0442.v1

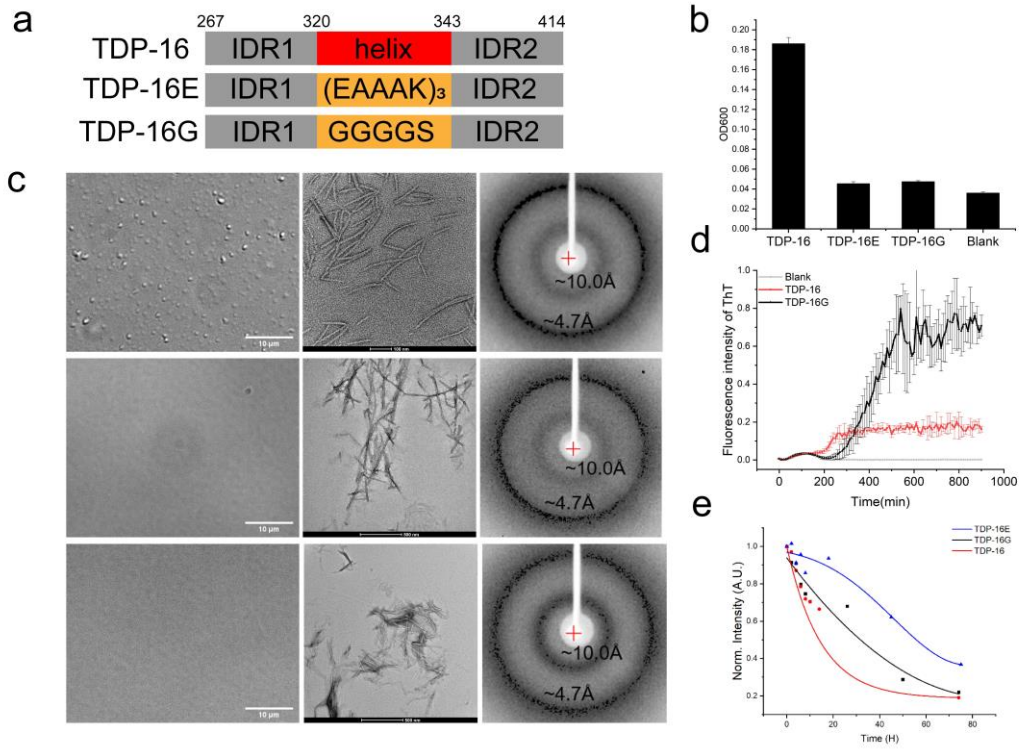

Figure 5. The function of the helical region of TDP-43 LCD in liquid-liquid phase separation and fibrillation. The protein was at a concentration of 20 μM in pH6.0 10mM phosphate buffer for all samples. (**a**) Schematic representation of three domains of TDP-43 LCD (also known as TDP-16) and its mutants TDP-16E and TDP-16G. The helical domain was replaced by $(EAAAK)_3$, a rigid helical sequence, for TDP-16E; The helical domain was replaced by GGGGS, a flexible disordered linker, for TDP-16G. (**b**) Turbidity (OD600 values) of freshly prepared wild-type TDP-43 LCD and its mutants in PB buffer. Error bars represent SD of three replicates. (**c**) Liquid-liquid phase separation of freshly prepared TDP-16, TDP-16E and TDP-16G was imaged by DIC microscopy. Amyloid fibrils formed were imaged by negative-staining TEM and identified by XRD. The fibrils were collected for study after 4 days' incubation of the protein solution. (**d**) Growth rate of TDP-16 and TDP-16G fibrils monitored by ThT fluorescence. Error bars represent SD of three replicates. (**e**) Aggregation rate of TDP-16, TDP-16E and TDP-16G monitored by the protein's intrinsic fluorescence.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2022 doi:10.20944/preprints202209.0442.v1

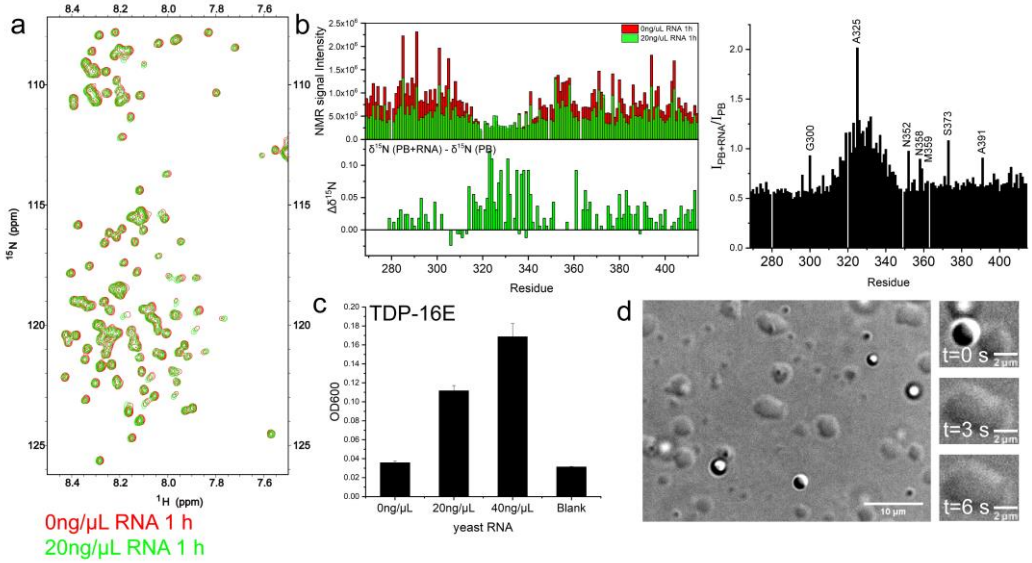

Figure 6. Effect of yeast RNA on TDP-43 LCD liquid-liquid phase separation and protein aggregation. (**a**) The ^{1}H-^{15}N HSQC spectra of 100 μM TDP-43 LCD in pH6.0 10mM phosphate buffer with 20ng/μL yeast RNA (in green) or without yeast RNA (in red), obtained within 2 hours from sample preparation. The temperature is 25℃. (**b**) (left) The NMR signal intensity comparison of the ^{1}H-^{15}N HSQC of TDP-43 LCD (100 μM) in pH6.0 10mM phosphate buffer with (in green) or without 20ng/μL yeast RNA (in red) and ^{15}N Chemical shift changes of TDP-16 (100 μM) upon adding 20ng/μL yeast RNA in pH6.0 10mM phosphate buffer. $\Delta\delta^{15}N = \delta^{15}N$ (PB+RNA) - $\delta^{15}N$ (PB). (Right) The NMR signal intensity ratio shown for all the residues. (**c**) Turbidity (OD600 values) of mutant TDP-16E (100 μM) with different concentrations of yeast RNA. Error bars represent SD of three replicates. (d) Liquid-liquid phase separation of TDP-16E (100 μM) with a concentration of 20ng/μL yeast RNA detected by DIC microscopy. The right panel of d showed a liquid-like droplet fusion in three times steps.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2022 doi:10.20944/preprints202209.0442.v1

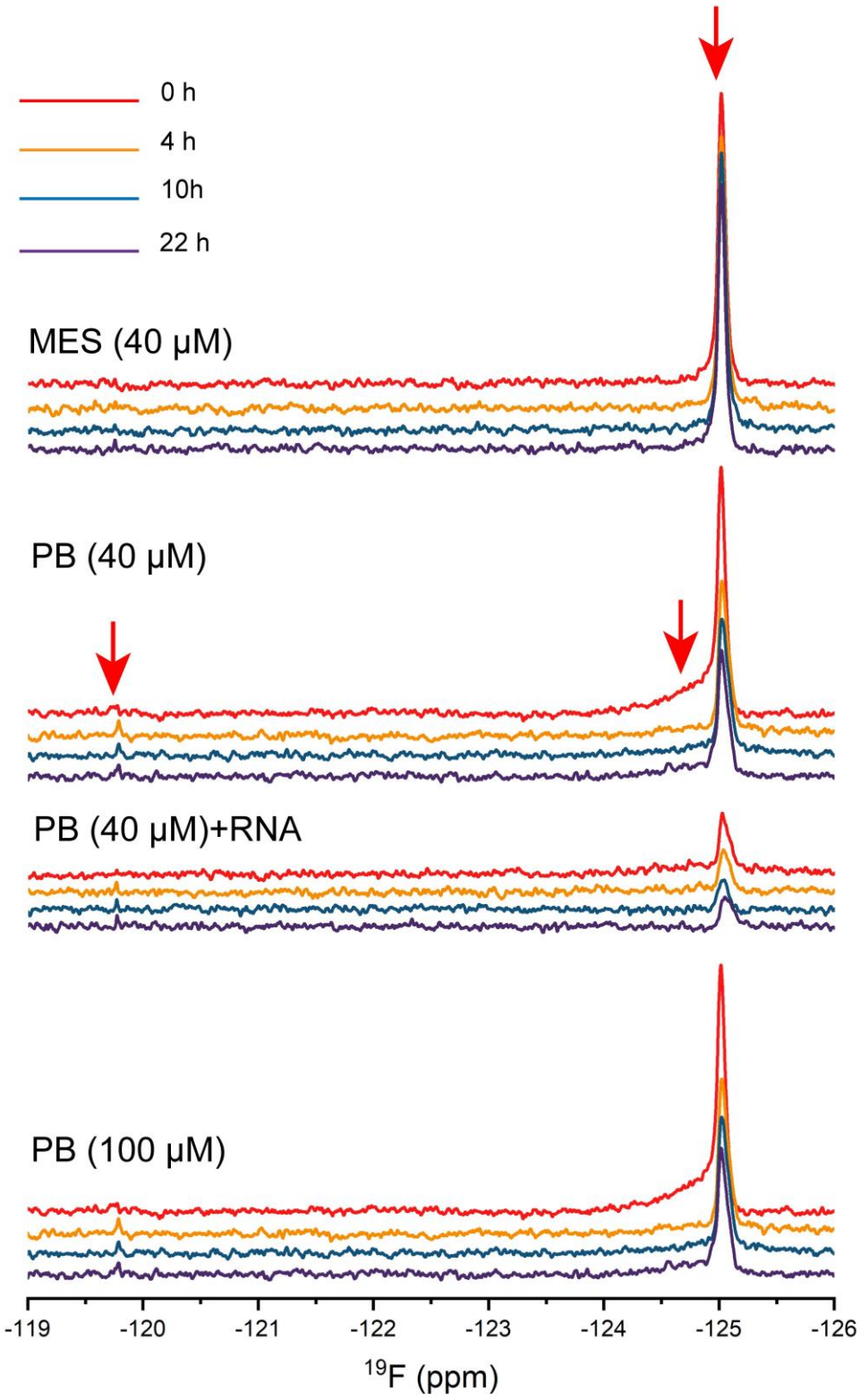

Figure 7. The dynamic changes of protein molecules probed by ^{19}F NMR. Time course spectra of 40 μM protein in pH 5.5 20mM MES buffer, 40 μM protein in pH 6.0 10mM PB buffer, 40 μM protein with 20ng/μL in pH 6.0 10mM PB buffer, and 100 μM protein in pH 6.0 10mM PB buffer were shown. Spectra obtained at ~30 mins (labeled as 0 hour, in red), 4 hours (in orange), 10 hours (in blue) and 22 hours (in purple) from the sample preparation. The temperature is 25℃.

 doi:10.20944/preprints202209.0442.v1

3. Discussion

A fine tune of the different interactions affects the protein LLPS equilibrium and droplet sizes

In this research, we showed ways to manipulate LLPS and the liquid droplet sizes of TDP-43 LCD. By reducing the protein concentrations or adding RNA, the LLPS equilibrium was disturbed. The sizes of the droplet could be increased and the fusion events could be observed by DIC only in a narrow protein concentration range. The active fusion event actually indicates a non-equilibrium situation, where more protein molecules are recruited to the droplet. Protein LLPS requires multivalence and a balance between different intermolecular interactions. To simplify the situation, the interactions can be tentatively put into two groups here for TDP-43 LCD, one is from the helices and the other is from the rest sequence, the IDR sequences. The two groups of interactions could induce the molecules into a type of loosely associated network, connecting the molecules in the helical region and the IDR regions. At 100 μM protein concentration in PB buffer condition, the intermolecular interaction mediated by the helices was probably too strong, but the interactions mediated by the IDR were too weak, not enough to extend the molecular network, therefore, the LLPS stopped at very early stage with small droplet sizes. Decreasing the protein concentration to 40 μM could shift the dynamic equilibrium and reduce the helix-helix intermolecular interactions so that the different intermolecular interactions could be more compatible to each other in the intensity and protein droplet sizes increase. Similarly, adding RNA enhances the interactions mediated by IDR sequences to bring them more compatible in the intensity to the helical interactions and protein droplet sizes increase.

The system we studied here is simplified and the real stress granules or TDP-43 granules in cells contain more variety of molecules including RNA and full-length TDP-43. In this real situation, the TDP-43 LCD helix-helix interaction may not be so dominant if the protein concentration is lower. Some other proteins found in stress granules, such as G3BP1, hnRNPA2 etc contain LCD too and would interact with TDP-43, RNA through a similar interaction mechanism, maintaining the stability. Therefore, the principle gained here should still be applicable in a more complicated system, although more components would have to be taken into consideration.

The molecular status of proteins in the mature droplets

The molecular properties of the mature droplets were also investigated. Since the mature droplets did not fuse in pH 6.0 PB buffer at 100μM protein concentration, the molecules may not still be liquid-like. The 1H-15N HSQC spectra intensity of freshly prepared mature droplets was first compared to the MES condition at 70 μM protein concentration, showing the signal intensity was proportional to the protein concentration for most residues at IDR (figure 3c). Protein was soluble, and showed very low LLPS in the MES condition. Therefore, the protein molecules in the mature droplets were still liquid-like, with similar dynamic properties as the ones in MES for IDR regions. Using THT fluorescence, we found the protein aggregation lag time was 200 min for 100 μM protein concentration in PB (figure S3a), indicating the protein was not severely aggregated within this time frame. Although the mature droplet did not fuse, changes in the solution condition would make the situation change quickly, such as dilution into a buffer or adding RNA. The observations also could indicate that the molecules in the mature droplet were still in an active equilibrium, not in a severe aggregation state. A mature protein droplet should have a certain lifetime before the protein aggregation considering its function in the cell. The membraneless organelle was supposed to have an active function and able to dissolve upon regulation.

Although the ^{1}H-^{15}N HSQC spectra showed the molecules still maintained high dynamic motions in general in the mature droplet, the intermolecular helical interaction was stronger, displayed by lower intensity of the peak and negative ^{15}N chemical shift changes

 doi:10.20944/preprints202209.0442.v1

in PB buffer. We also found the ^{19}F signal intensity from Trp aromatic ring was attenuated (in PB vs in MES) and not proportional to the protein concentrations (40 μM vs. 100 μM in PB), indicating the aromatic residue sidechains were involved in the intermolecular interactions for LLPS. Therefore, the protein molecules in the mature droplet were involved in stronger intermolecular interactions, but still liquid-like and able to change its molecular interactions quickly upon induction.

TDP-43 LCD aggregation intermediate

When the protein aggregates, it usually accompanies a NMR signal loss for all the residues. It was observed for 100 μM protein concentration in PB after 53 hours (figure 2c). It was unexpected for us to observe the signal increase of the helical region during the protein aggregation while the signal decrease of the IDR region was clearly shown (spectra taken at 6th hour and 22nd hour 2c). The signal increase suggested an enhancement of the molecular dynamics of the helical region, probably through a slight release of the helix-helix interaction. Therefore, we observed an intermediate step during TDP-43 LCD aggregation from protein droplets. We then found two other conditions with reduced helix-helix interaction indeed showed faster aggregation with shorter lag time studied using THT fluorescence (70 μM protein concentration in pH5.5 MES and 100 μM protein concentration, 150mM urea in pH6.0 PB). ^{1}H-^{15}N HSQC spectra also showed a similar observation that the two conditions with less helical intermolecular interactions displayed faster decay in the signal intensity. Once the intermolecular interactions brought the molecules together, the protein fibrillation started at the IDR regions, which was followed by the helical region structural conversion.

Previous NMR studies indicated that for residue 321-330, the α-helical structure only populated about 50% of the conformational ensemble, while for residue 331-343, the helical population was even smaller [11]. Although those values were obtained mostly at 20 μM protein concentration at pH 6.1 MES, where the protein was not in droplets, it indeed showed the helices in TDP-43 LCD were in dynamic exchange in the conformation, not a strong helical structure. Previous research also indicated increased helical structure upon protein LLPS, consistent with our results [15]. Therefore, a proper helix-helix interaction is needed for protein LLPS and at the same time, the intermolecular interaction also helps maintain the stability of the helical structure. Without a good intermolecular interaction, the helical structure can get loose easily and the protein is easy to aggregate as shown in urea condition or our pH 5.5 MES condition. But even in the mature droplet state, for a long time incubation, there are still opportunities for a temporary off of the helix-helix interaction and a slightly loosening of the helicity. That would explain the increase in the HSQC spectra intensity at the helical region during the protein aggregation after long time incubation. This observation was consistent with studies on the cell, showing the recruitment of TDP-43 into granules would protect the protein from fast aggregation [15, 16]. Some of the ALS-associated mutations were also reported to alter the helix-helix interactions or the helical propensity [31]. Our observation again reinforced the importance of the helices for TDP-43 LCD LLPS and aggregation.

The polar, aromatic residue rich sequence of low complexity domain

X-G/S, G/S-X (X represents the aromatic residues) sequences have been found in many proteins with LLPS. Interestingly, the nucleoporins in nuclear pore complexes also contain many phenylalanyl-glycyl (FG)-rich repeats at their selective filter for the random fuzzy interactions with their cargo the transport factors [32]. Similar observation on the NMR signal intensity attenuation was also reported on this FG-rich region of nucleoporins. TDP-43 was originally expressed in nucleus and was transported to the cytoplasm through the nuclear pore complexes. We speculate that the intermolecular interactions between TDP-43 LCD and the nucleoporins may also be present during the transportation of TDP-43 out of the nucleus.

In conclusion, this research studied the protein molecular properties when TDP-43 LCD formed the mature droplets. They were still liquid-like although the intermolecular

 doi:10.20944/preprints202209.0442.v1

interactions were stronger than the lower protein concentration conditions or no LLPS conditions. The protein exit and reentrant equilibrium could be shifted by modifying the solution environment, here adding RNA or dilution of the protein was demonstrated. The protein in the mature droplets would aggregate gradually but the aggregation was slower than some conditions with decreased helix-helix intermolecular interactions. A partial loosening of the helical intermolecular interaction was identified as the aggregation intermediate step. Not all interactions were probed here, and the studies were only carried out at a very simplified system to exclude other influence factors. But the principles could be applied to a more complicated and realistic case. Since a single factor could influence the protein LLPS significantly, the information gained could provide useful guidance to design ligands to fine tune the protein phase behaviors.

4. Materials and Methods

Expression and Purification

The cDNA of human TDP-43 LCD (from residue 267 to 414) was derived from plasmid encoded thioredoxin (Trx)-fused TDP-43 LCD (gift from Prof. Hong-Yu Hu), the cDNA of mutant TDP-16E and TDP-16G were synthesized directly and all cDNA were cloned into Pet32M with the N-terminal hexa-His-tag. Proteins were overexpressed in E. coli BL21(DE3) or Rosetta (DE3). The uniformly labeled peptide was expressed in M9 minimal medium containing 4 g of glucose and 1.5 g of $^{15}NH_4Cl$ per liter. Unlabeled peptide was expressed in LB medium. Cells were grown at 37℃ until the OD600 reading was 0.8. Then, the protein expression was induced overnight at 22℃ by adding 0.5 mM isopropyl β-D-thiogalactoside (IPTG) for TDP-43 LCD, or induced 6h at 37℃ by adding 1 mM IPTG for TDP-16E and TDP-16G.

5-fluoro-tryptpphan (5FW) labeled peptide was obtained by adding 5-fluoroindole (5FI) in M9 minimal medium. In brief, after cells were grown to OD600 about 0.6 in 1 liter M9 minimal medium, cells were centrifuged and the pellets were transferred into 0.5 liter M9 minimal medium containing 2 g of glucose, 0.75 g of NH_4Cl, 30 mg of 5FI, 20 mg of Tyr, 20 mg of Phe and 0.5 g glyphosate (to suppress the pentose phosphate pathway for aromatic amino acid synthesis). Following incubation for 30 minutes, 0.5 mM (IPTG) was added to initiate protein expression.

Cells were collected by centrifugation at 10,000g for 10 min and the cell pellet was resuspended in 100 mL lysis buffer (50mM Tris-HCl, 300mM NaCl, pH 8.0) with 1 mM PMSF, lysed by French press and centrifuged at 28,000g for 30min, the peptide in the inclusion bodies were washed with water then resuspended in 20 mL denaturing binding buffer (50mM Tris-HCl, 300mM NaCl, 8M urea, pH8.0) until most inclusion bodies dissolved at 4℃. Further centrifugation at 28,000g for 10min, 4℃ and the supernatant was purified by Ni-NTA affinity column, while the elution buffer contained 8 M urea and 500 mM imidazole.

For TDP-43 LCD, the protein solution with 8 M urea and 500 mM imidazole was dialyzed in water for 1 day at room temperature. All dialysates were collected and lyophilized. The dried sample was then dissolved in 30% formic acid and subsequently purified by reverse-phase HPLC on a C18 column eluted by a water-acetonitrile solvent system. The HPLC elution containing pure recombinant proteins was lyophilized and stored at -80℃ for further experiments.

For TDP-16E and TDP-16E, the protein is initially stored in 8 M urea and desalted into the phosphate buffer with a 0.5mL ZebaSpin Desalting column (Thermo Scientific) and diluted to 100uM or 20uM for experiments.

Turbidity Measurements

TDP-43 LCD and variants were dissolved in different buffers at 25℃ and incubated for 5 minutes. 100 μL of samples were transferred to a 96-cell plate. Turbidity was measured using a plate reader (Enspire, PerkinElmer) monitoring the absorbance at 600 nm. The tested solution conditions included PB buffer, pH6.0 10mM phosphate buffer;

 doi:10.20944/preprints202209.0442.v1

PB+RNA, pH6.0 10mM phosphate buffer with yeast RNA (20ng/uL, Sigma); PB+Urea, pH6.0 10mM phosphate buffer with 150mM Urea; pH5.5 20mM MES buffer.

Thioflavin-T assays

TDP-43 LCD and variants were dissolved in different buffers containing 20 μM THT at 25℃ and transferred to a 96-cell plate. The fluorescence emission at 480 nm was measured using a plate reader (Enspire, PerkinElmer) with excitation wavelength at 430 nm [13]. 5 seconds of shaking being applied before each reading. Blank was pH 6.0 10mM phosphate buffer only.

Intrinsic Fluorescence Spectroscopy

20 μM of TDP-43 LCD and variants were dissolved in a phosphate buffer and 150 μL of the samples were transferred to a cuvette (Quartz SUPRASIL, Hellma). For fluorescence spectroscopy measurements (FluoroMax-4, HORIBA), the excitation wavelength was set to 295 nm, the emission wavelength range was 310-500 nm. Both slits were 5 nm, and the scanning step was 1 nm [33, 34]. The variation of fluorescence maximum intensity with time indicated the aggregation rate.

Differential interference contrast (DIC) Microscopy

TDP-43 LCD and variants were dissolved in different buffers at 25℃ and incubated for 5 minutes. For all samples, 5 μL of protein solutions were dropped onto the bottom of a glass dish. Then the solution was checked by inverted microscope (Nikon ECLIPSE Ti, Nikon) and imaged by digital camera (ORCA-Flash 4.0, HAMAMATSU) with a 60×1.49 NA oil objective. Blank was pH 6.0 10mM phosphate buffer only.

Optical tweezers

An optical tweezer microscope C-trapTM from LUMICKS (Amsterdam, The Netherlands) with two steerable traps was used to perform controlled fusion of droplets [35]. 100μM of TDP-43 LCD was dissolved in a phosphate buffer and droplets were formed in minutes. These droplets flowed into the chamber just before data acquisition. A 1064 nm laser with low light intensity (< 0.5 W) was applied to minimize heating. One droplet was held in place by a trap, and the other steerable trap was used to capture other droplets and bring them toward the stationary droplet with a velocity of 0.04 μM s−1 until the two droplets' surface touched [36]. Force-extension and image data were taken at 5 Hz. Touch times were determined via analysis from the laser signal and confirmed with video signals.

Negative-staining transmission electron microscope (TEM)

To observe the droplets, 100 μM of TDP-43 LCD was dissolved in different buffers and incubated for 5 minutes. To observe the fibril, 20 μM of TDP-43 LCD and variants were incubated for 4 days before imaging. In total, 5ul sample solution was adsorbed to the glow-discharged TEM grid (Cu, 300 mesh; Beijing Zhongjingkeyi Technology Co., Ltd.) for 45 s. Then the grid was washed using 5 μL of water for 3 s, and finally stained with 5 μL of 2% uranyl acetate for 45 s. TEM images were obtained on a transmission electron microscope (Talos L120C, FEI). The acceleration voltage was 120KeV. Exposure time for each image was 2 s.

X-ray diffraction (XRD)

20 μM of TDP-43 LCD and variants were incubated for 4 days first. The solution mixtures were then centrifuged at 50,000 rpm for 2 h (Optima Max-TL, BECKMAN COULTER) and the pellets were collected. The precipitation was applied to a Single Crystal X-ray Diffraction instrument (Bruker D8 VENTURE, Bruker) for the measurement and the light source was Cu Kα radiation at 1.54184Å wavelength [37].

 doi:10.20944/preprints202209.0442.v1

Solution-state NMR

The samples were dissolved in 90% H_20/10% D_2O pH6.0 phosphate buffer with or without RNA/Urea or pH5.5 MES buffer. All 2D ^{1}H-^{15}N HSQC NMR experiments were recorded on Bruker 800M Hz AVANCE III spectrometer at 298K. The spectrum was first taken within 2 hours from the sample preparation and more spectra were taken again after ~6, 22 or 53 hours. All spectra were collected with the following parameters: 128* and 2048* complex pairs in the indirect ^{15}N and direct ^{1}H dimensions, 32 scans, 13.9 and 28 ppm as spectral widths for ^{1}H and ^{15}N. Each experimental time was approximately 1 hour 52 minutes. The ^{1}H-^{15}N HSQC peak assignments were based on a published chemical shift list deposited in BMRB under the accession code 26823. A summary of ^{1}H,^{15}N chemical shifts from four different publications were listed in table S2 to show the similarity and difference between the different samples (BMRB code:26823, 50154, 26728,26816).

Proton chemical shifts were directly referenced using DSS on a TDP-43 LCD sample prepared for this purpose, and ^{15}N chemical shifts were referenced indirectly. All spectra were processed using either Sparky or Topspin 4.1.3. All chemical shifts gained from ^{1}H-^{15}N HSQC spectra with various conditions prepared in this work were reported in table S3.

^{19}F NMR Spectroscopy

The samples were dissolved in 90% H_20/10% D_2O pH6.0 phosphate buffer with or without RNA or pH5.5 MES buffer. All NMR spectra were recorded on a Bruker AVANCE-600 MHz spectrometer (Bruker Biospin, Billerica, MA) at 298K. The spectrum was first taken within ~30 mins from the sample preparation and more spectra were taken again after ~4, 10 or 22 hours. All spectra were collected with the following parameters: 40960 complex points and 1024 scans. Each experimental time was approximately 30 minutes. All samples contained TFA as an internal reference, which was set at -75.6 ppm. Line broadening of 10 Hz was used to process the final spectra. Origin 2018 and MestReNova were used for plotting the data.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: title; Table S1: title; Video S1: title.

Author Contributions: The research was designed by J. -X L, B. S. and Y.Z.; The samples were prepared by Y.Z. and X.Z.; The NMR Spectroscopy were performed by L.Y.; The optical tweezers were performed by L.B.; The data analysis was performed by J. -X L. and Y.Z.; The original manuscript was written by J. -X L.; and revised by Y.Z., B.S. and L.Y.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by grants from the National Key Research and Development Program of China (2017YFA0504804 to J.-X.L.), and the Natural Science Foundation of China (No. 32171185 and No. 31770790 to J.-X.L.).

Acknowledgments: We would like to thank the Molecular Imaging Core Facility (MICF) at School of Life Science and Technology, ShanghaiTech University for DIC studies, the Biomolecular NMR Facility at ihuman Institute and the School of Life Science and Technology, ShanghaiTech University for NMR studies. TEM images were taken at the Bio-Electron Microscopy Facility of ShanghaiTech University. The authors would also like to thank the support from the Analytical Instrumentation Center (contract no. SPST-AIC 10112914), the School of Physical Science and Technology at ShanghaiTech University. We thank Dr. Dongsheng Liu for the advises in ^{19}F labeling of protein.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Banani, S.F., et al., Biomolecular condensates: organizers of cellular biochemistry. Nat Rev Mol Cell Biol, 2017. 18(5): p. 285-298.
2. Wang, B., et al., Liquid-liquid phase separation in human health and diseases. Signal Transduct Target Ther, 2021. 6(1): p. 290.
3. Buchan, J.R. and R. Parker, Eukaryotic stress granules: the ins and outs of translation. Molecular Cell, 2009. 36(6): p. 932-941.
4. Brangwynne, C.P., et al., Germline P granules are liquid droplets that localize by controlled dissolution/condensation. Science, 2009. 324(5935): p. 1729-32.
5. Machyna, M., P. Heyn, and K.M. Neugebauer, Cajal bodies: where form meets function. Wiley Interdiscip Rev RNA, 2013. 4(1): p. 17-34.
6. Molliex, A., et al., Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell, 2015. 163(1): p. 123-33.
7. Antifeeva, I.A., et al., Liquid-liquid phase separation as an organizing principle of intracellular space: overview of the evolution of the cell compartmentalization concept. Cell Mol Life Sci, 2022. 79(5): p. 251.
8. Tziortzouda, P., L. Van Den Bosch, and F. Hirth, Triad of TDP43 control in neurodegeneration: autoregulation, localization and aggregation. Nat Rev Neurosci, 2021. 22(4): p. 197-208.
9. Dewey, C.M., et al., TDP-43 aggregation in neurodegeneration: are stress granules the key? Brain Res, 2012. 1462: p. 16-25.
10. Francois-Moutal, L., et al., Structural Insights Into TDP-43 and Effects of Post-translational Modifications. Front Mol Neurosci, 2019. 12: p. 301.
11. Conicella, A.E., et al., ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure, 2016. 24(9): p. 1537-49.
12. Li, H.R., et al., The physical forces mediating self-association and phase-separation in the C-terminal domain of TDP-43. Biochim. Biophys. Acta, 2017. 1866(2): p. S1570963917302303.
13. Jiang, L.L., et al., Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. J. Biol. Chem., 2013. 288(27): p. 19614-19624.
14. Sun, Y. and A. Chakrabartty, Phase to Phase with TDP-43. Biochemistry, 2017. 56(6): p. 809-823.
15. Gasset-Rosa, F., et al., Cytoplasmic TDP-43 De-mixing Independent of Stress Granules Drives Inhibition of Nuclear Import, Loss of Nuclear TDP-43, and Cell Death. Neuron, 2019. 102(2): p. 339-357 e7.
16. Mann, J.R., et al., RNA Binding Antagonizes Neurotoxic Phase Transitions of TDP-43. Neuron, 2019. 102(2): p. 321-338 e8.
17. Buratti, E., Functional Significance of TDP-43 Mutations in Disease. Adv Genet, 2015. 91: p. 1-53.
18. Martin, E.W., et al., Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. Science, 2020. 367(6478): p. 694-699.
19. Reijns, M.A., et al., A role for Q/N-rich aggregation-prone regions in P-body localization. J Cell Sci, 2008. 121(Pt 15): p. 2463-72.
20. Fan, H.C., et al., Polyglutamine (PolyQ) diseases: genetics to treatments. Cell Transplant, 2014. 23(4-5): p. 441-58.
21. Mompean, M., et al., Structural Evidence of Amyloid Fibril Formation in the Putative Aggregation Domain of TDP-43. J Phys Chem Lett, 2015. 6(13): p. 2608-15.
22. Murray, D.T. and R. Tycko, Side Chain Hydrogen-Bonding Interactions within Amyloid-like Fibrils Formed by the Low-Complexity Domain of FUS: Evidence from Solid State Nuclear Magnetic Resonance Spectroscopy. Biochemistry, 2020. 59(4): p. 364-378.
23. Pantoja-Uceda, D., et al., Phe-Gly motifs drive fibrillization of TDP-43's prion-like domain condensates. PLoS Biol, 2021. 19(4): p. e3001198.
24. Pakravan, D., et al., Liquid-Liquid Phase Separation Enhances TDP-43 LCD Aggregation but Delays Seeded Aggregation. Biomolecules, 2021. 11(4).
25. Lim, L., et al., ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43. Plos Biology, 2016. 14(1): p. e1002338.
26. Mompean, M., et al., Electrostatic Repulsion Governs TDP-43 C-terminal Domain Aggregation. PLoS Biol, 2016. 14(4): p. e1002447.
27. Guenther, E.L., et al., Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. Nat Struct Mol Biol, 2018. 25(6): p. 463-471.
28. Chen, X., J.L. Zaro, and W.C. Shen, Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev, 2013. 65(10): p. 1357-69.
29. Suzuki, Y., et al., Resolution of oligomeric species during the aggregation of Abeta1-40 using (19)F NMR. Biochemistry, 2013. 52(11): p. 1903-12.
30. Larda, S.T., et al., Dynamic equilibria between monomeric and oligomeric misfolded states of the mammalian prion protein measured by 19F NMR. J Am Chem Soc, 2013. 135(28): p. 10533-41.
31. Conicella, A.E., et al., TDP-43 alpha-helical structure tunes liquid-liquid phase separation and function. Proc Natl Acad Sci U S A, 2020. 117(11): p. 5883-5894.
32. Hough, L.E., et al., The molecular mechanism of nuclear transport revealed by atomic-scale measurements. Elife, 2015. 4.

 doi:10.20944/preprints202209.0442.v1

33. Zhuo, X.F., et al., Solid-State NMR Reveals the Structural Transformation of the TDP-43 Amyloidogenic Region upon Fibrillation. J Am Chem Soc, 2020. 142(7): p. 3412-3421.
34. Ladokhin, A.S. Fluorescence Spectroscopy in Peptide and Protein Analysis. 2006.
35. Patel, A., et al., A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. Cell, 2015. 162(5): p. 1066-77.
36. Gui, X., et al., Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly. Nat Commun, 2019. 10(1): p. 2006.
37. Zhang, J., et al., Hydroxyapatite Formation Coexists with Amyloid-like Self-Assembly of Human Amelogenin. Int J Mol Sci, 2020. 21(8).