ATP regulates TDP-43 pathogenesis by specifically binding to an inhibitory component of a delicate network controlling LLPS

Dang Mei, Liangzhong Lim, Jian Kang and Jianxing Song*

Department of Biological Sciences, Faculty of Science; National University of Singapore; 10 Kent Ridge Crescent, Singapore 119260;

Short title: ATP regulates TDP-43 pathogenesis

*** Corresponding author; Email:** dbssjx@nus.edu.sg.

Key words: Amyotrophic lateral sclerosis (ALS); TDP-43; Liquid-liquid phase separation (LLPS); Prion-like domain (PLD); Adenosine triphosphate (ATP); NMR spectroscopy.

ABSTRACT

Very recently, liquid-liquid phase separation (LLPS) of cytoplasmic TDP-43 independent of forming stress granule (SG) has been decoded to initiate the neuron death for ALS. Mysteriously neurons maintain ATP concentrations of ~3 mM, but whether ATP modulates TDP-43 LLPS remains completely unknown. Here we characterized the effects of ATP on LLPS of TDP-43 PLD and its six mutants by DIC and NMR. For the first time, the results revealed: 1) ATP does induce and subsequently dissolve LLPS of TDP-43 PLD. 2) ATP achieves the modulation all by specifically binding Arg mostly saturated at 1:100. 3) LLPS of TDP-43 PLD and its exaggeration into aggregation are controlled by a delicate network composed of both attractive and inhibitory interactions, thus rationalizing the susceptibility of TDP-43 PLD to various ALS-causing mutations. Our studies together establish that ATP specifically binds Arg in intrinsically-disordered domains even not RGG-/R-rich, implying that ATP might be a universal regulator for most, if not all, R-containing intrinsically-disordered domains by altering their physicochemical features, conformations, dynamics, LLPS and assembly. Under physiological conditions, TDP-43 even in neuronal cytoplasm is highly bound with ATP and thus inhibited for its toxic LLPS, highlighting a central role of ATP in TDP-43 pathogenesis and aging.

Introduction

Aggregation of TAR DNA-binding protein 43 (TDP-43) in the cytoplasm of motor neurons is a common pathological hallmark of most cases of amyotrophic lateral sclerosis (ALS), which has also been found in other major neurodegenerative diseases, including Alzheimer's (AD), Parkinson's (PD), frontotemporal dementia (FTD), and Huntington's (HD) diseases (1-3). 414-residue TDP-43 consists of the folded N-domain (4-5) and two RRM domains (6), as well as the C-terminal low-complexity (LC) domain over the residues 267-414 which is intrinsically disordered with the amino acid composition similar to those of yeast prion proteins driving the formation of infectious conformations, thus designated as the prion-like domain (PLD) (7,8). Strikingly, despite being intrinsically-disordered, TDP-43 PLD hosts almost all ALS-causing mutations and drives liquid-liquid phase separation (LLPS) of TDP-43, which is functionally needed for forming stress granules (SGs) composed of protein and RNA components but could also exaggerate into aggregation or amyloid fibrils (4-17).

Very recently, a neuronal cell death mechanism has been identified which is initiated by LLPS of cytoplasmic TDP-43 independent of forming conventional SGs (2). Briefly, upon accumulation in the cytoplasm, TDP-43 undergoes LLPS to form long-lived liquid droplets without classical stress granule (SG) markers TIA1 and G3BP1. Furthermore, cytoplasmic TDP-43 droplets act to recruit importin-α and Nup62 as well as induce mislocalization of RanGap1, Ran, and Nup107, consequently resulting in inhibition of nucleocytoplasmic transport, clearance of nuclear TDP-43, and finally cell death. This indicates that TDP-43 can undergo the toxic LLPS to initiate ALS independent of forming SGs.

Mysteriously, all living cells maintain ATP concentrations of 2-12 mM, which are much higher than those required for its previously-known functions at micromolar concentrations (18-21). Only recently, it was decoded that ATP at concentrations > 5 mM hydrotropically dissolves LLPS of RNA-binding proteins including TDP-43, as well as

prevents/inhibits protein aggregation/fibrillation (19). We further found that by specific binding, ATP also mediates LLPS of the intrinsically-disordered RGG-rich domain (20) and fibrillation of the folded RRM domain (21) of FUS. Despite having the lowest ATP concentrations among human cell types, neurons still maintain ATP concentrations of ~3 mM. So two key questions of both physiological and pathological significance arise: 1) can ATP modulate LLPS of TDP-43 PLD which is not RGG- and even not R-rich? 2) If yes, what is the underlying mechanism?

So far, due to the challenge in characterizing LLPS of TDP-43 PLD which results from dynamic and multivalent interactions of the intrinsically-disordered LC domain, even the driving forces for LLPS are still under hot debate despite extensive studies (11-14). In particular, it remains completely unexplored for the effect and mechanism of ATP on LLPS of TDP-43 PLD. Here by DIC microscopy and NMR spectroscopy, we aimed to address the two questions by characterizing the effect of ATP on LLPS as well as its molecular interactions with the wild-type (WT) and six TDP-43 PLD mutants. For the first time, we found that ATP is indeed capable of inducing at low concentrations but dissolving LLPS of TDP-43 PLD at high concentrations. Detailed NMR analysis revealed that ATP achieves the modulation all by specifically binding Arg residues. Further studies on six mutated/deleted mutants decode for the first time that LLPS of TDP-43 PLD is in fact controlled by a delicate interaction network which is constituted not only by the attractive interactions as previously identified (11-14), but also by the inhibitory interaction associated with Arg residues, which were shown here to maintain the reversibility of LLPS so as to prevent its further exaggeration into aggregation. Our unambiguous NMR results here establish that in general ATP can specifically bind Arg residues in the disordered domains which do not need to be RGG- or even R-rich like the FUS RGG-rich domain (20). Furthermore, the finding that TDP-43 LLPS is governed by a delicate interaction network rationalizes its susceptibility to the perturbation by various ALS-causing

mutations to exaggerate into the pathological consequence. The results together support the emerging notion that ATP with concentrations > mM has a large category of previously-unknown functions operating with diverse mechanisms which are fundamentally critical for cell physiology, pathology and aging.

Results

1. ATP modulates LLPS of TDP-43 PLD in a two-stage manner.

TDP-43 has a long C-terminal PLD from residues 267-414 containing six basic residues: Arg268, Arg272, Arg275, Arg293 and Arg361, as well as Lys408 (Fig. 1A). Previous NMR studies from other and our groups indicate that TDP-43 PLD is intrinsically disordered but contains a unique central hydrophobic region over residues 311-343 adopting a partially-folded helical conformation (11,12,16,22). In the present study, TDP-43 PLD over Asn265-Met414 was used as we previously characterized (16,23). By measuring turbidity and visualizing the droplet formation with DIC microscopy, we have first extensively assessed the capacity of TDP-43 PLD to phase separate at different protein concentrations, buffer pH and salt concentrations. We found that at protein concentrations of 15 μM or lower, no LLPS or aggregation was detected for TDP-43 PLD without ATP in 10 mM sodium phosphate buffer from pH 4 to 8 at 25 C, consistent with the previous studies (8-10). On the other hand, at concentrations > 50 μM, it could phase separate without ATP but would exaggerate into aggregation after 1 hr.

Nevertheless, upon titrations with ATP, TDP-43 PLD at 15 μM was induced to phase separate, which reached the highest turbidity values (1.62 and 0.94 respectively at pH 7.0 and 5.5) at a molar ratio of 1:100 (PLD:ATP) (Fig. 1B). Further increase of ATP concentrations led to the reduction of turbidity and at 1:1500, the turbidity values were only ~0.05 at both pH. DIC visualization showed that without ATP, no droplets were detected at both pH 7.0 and 5.5. Nevertheless, upon adding ATP, the dynamic droplets were formed at 1:25 with diameter of ~1 μm and at 1:100, many droplets were formed (Fig. 1C). Further increase of ATP concentrations led to reduction of the droplet number and at 1:750 only very few droplets could be detected. At 1:1000, the droplets were completely dissolved. The overall pattern of the

induction and dissolution at pH 7.0 is the same as at pH 5.5 except that the number of the droplets was higher than that at pH 5.5 (Fig. 1C), thus resulting in higher turbidity (Fig. 1B).

2. ATP achieves the modulation of LLPS by specifically binding Arg.

To probe the molecular interactions underlying ATP-induced induction and dissolution of LLPS, we conducted ATP titrations on TDP-43 PLD at 15 μM in 10 mM phosphate buffer at pH 5.5 as monitored by NMR HSQC spectroscopy which is powerful in detecting even very weak perturbations at residue-specific resolution. We selected pH 5.5 for all NMR studies because: 1) several mutants showed severe aggregation above pH 6.0; 2) many HSQC peaks became undetectable at pH 7.0 due to dynamic self-association or/and the enhanced exchange between amide protons and the bulk water (11,12,16).

As seen in Fig. 1D, only a small set of HSQC peaks shifted upon adding ATP even up to 1:1000. Detailed analysis revealed that except for the last residue Met414, the residues with significant shifts are all clustered over N-terminal three Arg residues including Arg268-Gln269-L270-Glu271-Arg272-Ser273 and Arg275-Phe276. Interestingly, the shifting process is largely saturated at the molar ratio of 1:100 (PLD:ATP). We subsequently titrated TDP-43 PLD with AMP under the same conditions. No significant increase in turbidity was observed even with AMP/PLD ratio reaching up to 1:1500 (Fig. S1A). DIC visualization indicated that no droplets or aggregates were detectable at all points of AMP concentrations. We also probed the effect of AMP by NMR and the results showed that AMP induced no significant shifts of HSQC peaks of TDP-43 PLD even up to 1:500 (Fig. S1B). Noticeably, further increase of AMP concentrations triggered the reduction of peak intensity and at 1:1000, HSQC signals were too weak to be detectable. This suggests that despite being undetectable by DIC, AMP induced the dynamic association of TDP-43 PLD which led to the disappearance of HSQC peaks. We also titrated TDP-43 PLD under the same conditions with ADP (Fig. S1 and S2). Upon adding ADP,

the turbidity of the sample increased and at 1:1500, the turbidity reached 1.31, which is even higher than the largest value titrated by ATP (Fig. 1B). Interestingly, however, DIC visualization showed that at 1:200, ADP induced the formation of detectable aggregates which became larger upon increasing ADP concentrations. At 1:1500, very large aggregates > 5 μm were formed which became visible precipitates after 10 min (Fig. S2A). Furthermore, the NMR characterization showed that at 1:25, the HSQC peaks of the residues around the N-terminal three Arg and last residues underwent obvious shifts (Fig. S2B). However, upon further adding ADP, the HSQC signals became too weak to be detected.

We also set to assess whether the inducing effect of ATP is simply due to the screening of the electrostatic interactions by titrating NaCl into TDP-43 PLD. The results showed very minor increase of turbidity (Fig. S3A) as well as no formation of droplets and aggregates as visualized by DIC with NaCl concentrations even up to 150 mM. As monitored by NMR, at 50 mM of NaCl, only HSQC peak of Met414 shifted, while at 150 mM, a set of HSQC peaks shifted (Fig. S3B and S3C), whose pattern, however, is different from that induced by ATP (Fig. 1E). The results thus indicated that at 15 μM and pH 5.5, NaCl has no capacity to induce LLPS of TDP-43 PLD even up to 150 mM. With all results with ATP, ADP, AMP and NaCl together, ATP appears to achieve the two-stage modulation of LLPS of TDP-43 PLD by very specific interaction which needs the presence of the triphosphate chain, more than simple electrostatic interaction (16,20,24).

3. The roles of Arg residues in LLPS and aggregation of TDP-43 PLD.

As NMR titrations showed that the residues clustered over N-terminal three Arg residues are most perturbed, here to determine the roles of five Arg and one Lys residues in the ATP-induced LLPS, we constructed three mutants of TDP-43 PLD by site-directed mutagenesis, namely C2 with Arg268, Arg272 and Arg275 mutated to Ala, C3 with Arg293,

Arg361 and Lys408 mutated to Ala, as well as C4 with all Arg and Lys mutated to Ala (Fig. 1A).

We have extensively assessed the capacity of three mutants to phase separate at different protein concentrations, buffer pH and NaCl concentrations. We found that all three mutants became prone to aggregation at pH above 6. While the proneness of C2 and C3 to aggregation above pH 6.0 might be explained by their pI of ~6.2, it is unexpected for the proneness of C3 with pI of ~3.6 (Fig. 1A), thus implying that factors other than the simple electrostatic interaction also contribute to aggregation of TDP-43 PLD. At pH 5.5, the C2 and C3 samples showed no significant aggregation at 15 μM in 10 mM phosphate buffer. Unexpectedly, the C3 sample at 15 μM started to aggregate even at pH 5.5 30 min after the sample preparation even with aggregates removed by centrifuge. Further, the addition of NaCl even at 50 mM induced visible precipitation for all three samples. Nevertheless, we have successfully acquired HSQC spectra at 15 μM in 10 mM phosphate buffer (pH 5.5) for the freshly prepared samples of three constructs (Fig. S4). All three samples have HSQC spectra with the majority of peaks superimposable to those of WT (Fig. S4), indicating no global alternation of the solution conformations of three mutants.

Subsequently we titrated ATP into the samples of the three mutants at 15 μM in the same buffer. As shown in Fig. 2A, upon adding ATP, the turbidity of the C2 sample increased and reached the highest (0.78) at 1:200 and subsequently reduced. Interestingly, even at 1:1500, the turbidity was still ~0.14, higher than those of WT (~0.05). On the other hand, DIC visualization showed that without ATP, the C2 sample had no droplets or aggregates. With addition of ATP, the droplets were formed at 1:25 but even at 1:200, the droplet number are much less than what was observed on WT (Fig. 2B). The droplets also started to be dissolved with further addition of ATP and at 1:1000, no droplet or aggregate could be observed.

In the absence of ATP, the C3 sample also had no significant turbidity reading. Upon adding ATP, the turbidity increased rapidly and reach the highest (1.73) at 1:100 which is higher than that of WT, and subsequently decreased with further addition of ATP. DIC visualization showed that the C3 sample had no droplet or aggregate without ATP but started to form droplets at 1:25. The number of the droplets reached the highest at 1:100 where many droplets were form. Droplets also became completely dissolved at 1:1000. Noticeably, even for the transparent C4 sample with aggregates removed by centrifuge as checked by DIC, the addition of ATP at 1:10 induced immediate precipitation as monitored by DIC.

We conducted the ATP titrations on three constructs under the same conditions as monitored by NMR HSQC (Fig. 2C). Interestingly for C2, at 1:100 only HSQC peak of Met414 showed a large shift. At the higher ratio, the HSQC peaks became weaker and after 1:200 most HSQC peaks became undetectable. Strikingly, for C3, at 1:100 in addition to HSQC peak of Met414, a small set of other HSQC peaks also shifted (Fig. 2C). Detailed analysis revealed that the additionally shifted peaks are all from the residues clustered over the N-terminal three Arg residues, very similar to what was observed on WT titrated by ATP. Intriguingly for C4, the addition of ATP into the freshly-prepared sample at 1:10 triggered the disappearance of most HSQC peaks due to visible precipitation. This set of results suggest that ATP specifically binds Arg residues, which play a key role in maintaining the reversibility of LLPS as well as preventing aggregation. Unexpectedly, even for C2 and C3 whose droplets could be dissolved as well as no detectable aggregates were formed at high ATP concentrations, at molecular level TDP-43 molecules appeared to undergo dynamic association at high ATP concentrations and therefore their HSQC peaks became too broad to be detected.

4. LLPS of TDP-43 PLD is controlled by a delicate interaction network.

To understand the contribution of different regions of TDP-43 PLD to LLPS and aggregation, we further generated C5 with the N-terminal residues 265-276 deleted, C6 with Arg293, Arg361 and Lys408 of C5 mutated to Ala as well as C7 with the hydrophobic region 311-343 deleted (Fig. 1A). We also extensively assessed the capacity of three variants to phase separate and aggregate at different protein concentrations, buffer pH and NaCl concentrations. We found that C5 and C6 samples were also prone to aggregation at pH above 6, again suggesting that the aggregation of TDP-43 PLD is not just governed by electrostatic interaction. On the other hand, addition of NaCl even at 50 mM triggered immediate precipitation for C5 and C6, while no precipitation and turbidity change were observed for C7 even with NaCl up to 150 mM (Fig. S6). Nevertheless, as shown in Fig. S3, at 15 μM in 10 mM sodium phosphate buffer (pH 5.5), the freshly-prepared samples of three Constructs also had HSQC spectra with most peaks highly superimposable to the corresponding ones of WT (Fig. S5), indicating no global changes of their solution conformations.

We then titrated the samples of three Constructs at 15 μM with ATP. As shown in Fig. 3A, for C5, without ATP, the turbidity is <0.1 while the addition of ATP induced increase of turbidity, which reached the highest (0.51) at 1:200. Further addition of ATP led to the decrease of turbidity. By contrast, for C6, the sample without ATP already has a turbidity of 0.19, and addition of ATP to 1:10 led to visible precipitation. On the other hand, for C7, the sample without ATP has a turbidity of 0.023, and addition of ATP even up to 1:1500 led to no significant change of turbidity.

DIC visualization revealed that for C5, without ATP, no droplets were detected while with addition of ATP at 1:25, some droplets were formed. At 1:200, more droplets could be seen but further addition of ATP led to the reduction of the number of droplets, and at 1:1000, all droplets were dissolved. By contrast, for C6, no droplets were formed without ATP, but

addition of ATP only to 1:10 resulted in visible precipitation. Interestingly, for C7, without ATP no droplets were found while the addition of ATP even upon to 1:1500 induced no formation of droplets or aggregates.

Interestingly, when we monitored the ATP titrations by NMR HSQC (Fig. 3C), for C5, only several peaks were significantly shifted which were identified to be from Arg293-Gly294 and Arg361-Glu362. The shifts were largely saturated at 1:200 and the HSQC peak of Lys408 showed no shift even up to 1:1000. On the other hand, due to the strong proneness of C6 to aggregate, we reduced the concentration to 10 μM and rapidly added ATP only at two ratios 1:500 and 1:1000. As shown in Fig. 3C, even at 1:1000, no significant shifts were detected except for Met414. Remarkably, for C7, although no droplets and aggregates were formed with the addition of ATP up to 1:1500, the residues clustered over the N-terminal three Arg residues were also found to have significant shifts which are largely saturated at 1:100 (Fig. 3C, 3D and 4A).

The titrations of NaCl into C7 led to no significant increase of turbidity (Fig. S6A), as well as no formation of droplets or aggregates as visualized by DIC. Further HSQC characterization showed that unlike what were observed on other mutants, the addition of NaCl up to 150 mM induced no disappearance of HSQC peaks but obvious shifts of some HSQC peaks (Fig. S6B). Detailed analysis revealed that the pattern of the shifted residues is very different from that induced by ATP (Fig. 3C). Interestingly, the pattern of the shifted residues of C7 induced by NaCl is more similar to that of WT induced by NaCl (Fig. S6C). The results clearly indicate that the ATP can always specifically bind Arg residues regardless of TDP-43 PLD constructs. Nevertheless, whether to induce LLPS further depends the presence of the unique hydrophobic helical region 311-343, which is shown to be the key driver for LLPS of TDP-43 PLD by the helix dimerization (8).

Discussion

Recently, intrinsically disordered proteins have been demonstrated to act as a key driver for LLPS at the concentrations much lower than those required for the folded proteins such as lysozyme (25). Remarkably, LLPS is now emerging as a common principle for spontaneously forming intracellular membraneless organelles (MLOs) underlying cellular physiology and pathology (7-15,17,25-29). Initially, the physical framework of Flory-Huggins theory developed for the phase separation of polymers (30,31) was adapted for describing protein LLPS. However, with respect to the complexity associated with the intrinsically-disordered proteins, very recently a stickers-and-spacers model has been proposed, in which the groups participating in attractive interactions are defined to be "stickers", while the regions between stickers that do not significantly contribute to attractive interactions are considered to be "spacers" (27). Therefore, to identify types, numbers and strengths of stickers represents a key step to elucidate the general molecular mechanisms of LLPS.

Here, for the first time, we have identified the existence of the inhibitory interaction for LLPS associated with Arg residues, which, however, functions to maintain the reversibility of LLPS and to prevent its exaggeration into aggregation. By site-directed mutagenesis, DIC and NMR visualization, the substitution of Arg residues was shown to result in the loss of the reversibility of LLPS and severe proneness to oligomerization/aggregation of TDP-43 PLD. Previously, the attractive interactions driving LLPS of TDP-43 PLD have been extensively characterized, which mainly include the helix-dimerization between the unique hydrophobic region (11), interactions among aromatic residues and other sequence motifs (12-14,32). Here Arg residues were decoded to contribute the inhibitory interactions to LLPS of TDP-43 PLD. This is very unexpected as previously Arg residues in the intrinsically disordered domains of FUS were shown to be the key sticker for LLPS by establishing π-cation interactions with aromatic residues (20,33-35). Our finding here thus suggests that the roles of Arg in LLPS,

whether to serve as a sticker or inhibitor, are highly context-dependent. In the future, it is of general interest to assess whether other amino acids might also have the context-dependent roles in LLPS.

Consequently, TDP-43 PLD represents the first example whose LLPS is controlled by a delicate interaction network composed of the opposite forces: attractive and inhibitory interactions. As such, the inhibitory interaction might need to be included in the theoretical models for describing LLPS driven by intrinsically-disordered domains. Therefore, in addition to "spacers" and "stickers", it is also of fundamental significance to identify the existence, type and strength of the "inhibitors" in controlling LLPS. Although the existence of the inhibitory interaction may dramatically increase the complexity of the theoretic treatment of LLPS, it rationalizes the observed susceptibility of intrinsically-disordered domains to very minor perturbations such as Met337Val mutation of TDP-43 PLD which are able to significantly alter the properties of LLPS and aggregation.

Mysteriously, ATP ubiquitously exists at high concentrations in all living cells. Therefore, the most important finding in the study is that ATP can modulate LLPS and aggregation of TDP-43 PLD in the two-stage manner: induction at low molar ratios and dissolution at high ratios. Mechanistically, by NMR and site-directed mutagenesis, we unambiguously decoded that ATP achieves the two-stage modulation all by specifically binding to five Arg residues, which only account for 3% of the 150-residue TDP-43 PLD. Surprisingly, the capacity of ATP in modulating LLPS of TDP-43 PLD is not simply resulting from its electrostatic property because under the same condition, NaCl failed to induce LLPS even up to 150 mM with a molar ratio of 1:10000 (PLD:NaCl) despite triggering extensive perturbations onto many residues. It is particularly unexpected to find that only ATP can reversibly modulate LLPS of TDP-43 PLD, while AMP had no capacity and ADP even triggered severe aggregation despite interacting with the same set of residues like ATP as

detected by NMR. This implies that ATP might have some unique capacity that NaCl, AMP and even ADP do not have. Indeed, our previous study showed that ATP highly depends on its unique hydration property associated with the triphosphate chain in modulating LLPS of FUS N-terminal domain (20). Very recently we have found that even without any detectable binding, ATP can still antagonize the crowding-induced destabilization of human lens γS-crystallin most likely by directly targeting the protein hydration shell (36,37). Therefore in the future, despite extremely challenging, it is of both fundamental and biological interest to explore the high-resolution mechanism by which ATP modulates LLPS of TDP-43 PLD with its unique hydration feature.

Previously we were unable to follow the residue-specific perturbations by NMR for the 156-residue intrinsically-disordered RGG-rich domain of FUS due to the extensive disappearance of HSQC peaks at most points of ATP concentrations, as well as the very high content of Arg residues, which contains a total of 25 Arg residues. As a result, we could only indirectly derive the speculation that ATP induces and dissolves LLPS of FUS most likely by targeting Arg residues (20). Here for the first time we were able to unambiguously conclude that ATP induces and dissolves LLPS of TDP-43 PLD by specifically binding Arg residues as supported by the successful assignment of HSQC spectra of TDP-43 PLD at all ATP concentration points in conjecture with site-directed mutagenesis. All results on both TDP-43 and FUS together thus establish that ATP can specifically bind Arg residue of intrinsically-disordered domains.

On the other hand, the exact mechanisms of the ATP-binding to induce LLPS appear to be also context-dependent: for FUS RGG-rich domain with 25 Arg residues, the binding of ATP is sufficient to directly induce LLPS. By contrast, for TDP-43 PLD with only 5 Arg residues, the ATP-binding is insufficient to solely drive LLPS, which thus needs other diving forces, particularly the helix-dimerization (11), which is evidenced by the result that the

deletion of the hydrophobic region abolishes the ATP-induced LLPS although ATP still binds Arg residues. Our current results thus bear the unprecedented implications that in addition to Arg methylation (38,39), ATP is a universal regulator for the functions of most, if not all, intrinsically-disordered R-containing domains whose number is much larger than that of RGG-rich domains in >1,700 human proteins by dynamically binding Arg residues to directly or indirectly modulate their physicochemical properties, conformations, dynamics, LLPS and assembly.

Strikingly, the binding of ATP to Arg residues is largely saturated at the molar ratio of 1:100 for three TDP-43 PLD constructs including WT, PLD with the residues 265-276 deleted and PLD with residues 311-343 deleted (Fig. 4A). In particular, ATP at 1:1000 is sufficient to completely dissolve the droplets. Under physiological conditions, the cytoplasm of neurons has ATP concentrations of ~3 mM but TDP-43 concentrations of only ~1 μM (40), thus with a molar ratio of ~1:3000 (TDP-43:ATP). As such, TDP-43 is largely bound with ATP over its PLD Arg residues as well as RRM1 domain (41), and is mostly inhibited for the toxic LLPS (I of Fig. 4B). Nevertheless, under the pathological or/and aging conditions with significantly increased TDP-43, or/and reduced ATP concentrations, which result in the large reduction of the molar ratios, TDP-43 PLD will undergo LLPS which could even be enhanced by binding with ATP (II of Fig. 4B). Moreover, the formed TDP-43 droplets recruit other proteins such as importin-α and Nup62 (III of Fig. 4B), which also induce mislocalization of RanGap1, Ran, and Nup107 as recently decrypted (2). After a certain period of pathological conditions, dynamic TDP-43 droplets might further exaggerate into irreversible aggregates or/and amyloid fibrils characteristic of ALS (IV of Fig. 4B). Moreover, as previously well documented the ALS-causing mutations in TDP-43 PLD could speed up the transition from dynamic droplets into irreversible aggregates/fibrils. These events lead to inhibition of nucleocytoplasmic transport, clearance of nuclear TDP-43, and ultimately cell death (2).

In summary, TDP-43 PLD represents the first example whose LLPS is controlled by a delicate network composed of both attractive and inhibitory interactions. Arg residues, the key sticker for LLPS of FUS, have been identified here to act as the inhibitor for LLPS of TDP-43 PLD, thus implying that the roles of residues in LLPS is highly context-dependent. Remarkably, ATP specifically targets Arg residues to achieve the two-stage modulation of LLPS of TDP-43 PLD. As ATP, the universal energy currency, exists in all cells with concentrations of 2-12 mM, our study logically establishes that in addition to Arg methylation, ATP is a general and dynamic regulator for the functions of most, if not all, intrinsically-disordered R-containing domains by specifically binding Arg residues to alter their physicochemical properties, conformations, dynamics, LLPS and assembly.

Materials and Methods

Preparation of recombinant WT and mutated TDP-43 PLD proteins.

In the present study, we utilized our previously-cloned DNA construct encoding TDP-43 265-414 (PLD) in a modified vector without tag (16,20,23,41). Six mutants of TDP-43 PLD were further generated by use of the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, United States) as previously described (16).

All seven recombinant TDP-43 PLD proteins were highly expressed in inclusions and their purification followed the previously established protocols by other and our labs in which the RP-HPLC purification was finally used to obtain high-purity proteins with any impurity such as liquid, salt ions and nucleic acids removed (12,16,20,23). Isotope-labeled proteins for NMR studies were prepared by the same procedures except that the bacteria were grown in M9 medium with addition of $(^{15}NH_4)_2SO_4$ for ^{15}N-labeling (16,20,23). The protein concentration was determined by the UV spectroscopic method in the presence of 8 M urea, under which the extinct coefficient at 280 nm of a protein can be calculated by adding up the contribution of Trp, Tyr, and Cys residues (16,20,23,41).

ATP, ADP and AMP were purchased from SigmaAldrich with the same catalog numbers as previously reported (19-21,41). The protein and ATP samples were all prepared in 10 mM sodium phosphate buffer and $MgCl_2$ was added for stabilization by forming the ATP-Mg complex (19-21,41). The final solution pH values were adjusted with aliquots of very diluted NaOH or HCl.

Differential interference contrast (DIC) microscopy and turbidity measurement

The formation of liquid droplets and aggregates was imaged at 25 ºC on 50 µl of different TDP-43 PLD samples at various protein concentrations in 10 mM sodium phosphate

buffer at different pH values in the absence and in the presence of ATP at molar ratios of 1:0, 1:2.5, 1:10, 1:25, 1:50, 1:100, 1:200, 1:300, 1:400, 1:500, 1:750, 1:1000 and 1:1500 by differential interference contrast (DIC) microscopy (OLYMPUS IX73 Inverted Microscope System with OLYMPUS DP74 Color Camera) as previously described (20,23). NaCl was added into different construct samples to reach 50, 100 and 150 mM respectively. The turbidity values were measured at wavelength of 600 nm for all DIC samples.

NMR characterizations

All NMR experiments were acquired at 25 ºC on an 800 MHz Bruker Avance spectrometer equipped with pulse field gradient units and a shielded cryoprobe as described previously (16,20,21,23,41). To have the enhancing effect of the cryoprobe for NMR signal sensitivity, which is essential for NMR HSQC titration experiments at a very low protein concentration (15 µM), NMR samples had to be prepared in 10 mM sodium phosphate buffer, while pH value was optimized to 5.5 as many HSQC peaks of TDP-43 PLD disappeared at higher pH values due to the enhanced exchange with bulk solvent or/and dynamic aggregation.

For NMR titration studies of the interactions between WT/mutated TDP-43 PLD proteins and ATP, one-dimensional proton and two dimensional ^{1}H-^{15}N NMR HSQC spectra were collected on ^{15}N-labelled samples at a protein concentration of 15 µM in 10 mM sodium phosphate buffer (pH 5.5) at 25 ºC in the absence and in the presence of ATP, ADP and AMP at molar ratios of 1:0, 1:2.5, 1:10, 1:25, 1:50, 1:100, 1:200, 1:300, 1:400, 1:500, 1:750 and 1:1000.

To calculate chemical shift difference (CSD) induced by interactions with ATP and NaCl, HSQC spectra were superimposed and subsequently the shifted peaks were identified, which were further assigned to the corresponding residues of TDP-43 PLD with the NMR resonance assignments previously by our and other groups (8,10,12). The degree of

perturbation was reported by an integrated index calculated by the following formula (20,21,41):

$$((\Delta^{1}H)^{2} + (\Delta^{15}N)^{2}/4)^{1/2}$$

Acknowledgement

This study is supported by Ministry of Education of Singapore (MOE) Tier 1 Grant R-154-000-B92-114 to Jianxing Song.

References

1. Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M., et al. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314, 130–133.

2. Gasset-Rosa F, Lu S, Yu H, Chen C, Melamed Z, Guo L, Shorter J, Da Cruz S, Cleveland DW. Cytoplasmic TDP-43 De-mixing Independent of Stress Granules Drives Inhibition of Nuclear Import, Loss of Nuclear TDP-43, and Cell Death. Neuron. 2019 Apr 17;102(2):339-357.e7.

3. Chiti, F., and Dobson, C.M. (2006). Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. 75, 333–366.

4. H. Qin, L. Lim, Y. Wei, J. Song, TDP-43 N terminus encodes a novel ubiquitin- like fold and its unfolded form in equilibrium that can be shifted by binding to ssDNA, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 18619e18624.

5. M. Mompe, V. Romano, D. Pantoja-Uceda, et al., The TDP-43 N-terminal domain structure at high resolution, FEBS J. 283 (2016) 1242e1260.

6. T. Afroz, E.M. Hock, P. Ernst, et al., Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation, Nat. Commun. 8 (2017) 45.

7. King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA- binding proteins with prion-like domains in neurodegenerative disease. Brain Res. 1462, 61–80.

8. Malinovska, L., Kroschwald, S., and Alberti, S. (2013). Protein disorder, prion propensities, and self-organizing macromolecular collectives. Biochim. Biophys. Acta 1834, 918–931.

9. Boeynaems, S., Alberti, S., Fawzi, N.L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., et al. (2018). Protein phase separation: a new phase in cell biology. Trends Cell Biol. 28, 420–435.

10. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J., Mittag, T., and Taylor, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. Cell 163, 123–133.

11. Conicella, A.E., Zerze, G.H., Mittal, J., and Fawzi, N.L. (2016). ALS mutations disrupt phase separation mediated by a-helical structure in the TDP-43 low- complexity C-terminal domain. Structure 24, 1537–1549.

12. Hao-Ru Li 1 , Wan-Chin Chiang 1 , Po-Chun Chou 1 , Won-Jing Wang 1 , Jie-Rong Huang. TAR DNA-binding protein 43 (TDP-43) liquid–liquid phase separation is mediated by just a few aromatic residues. J. Biol Chem . 2018 Apr 20;293(16):6090-6098.

13. Hermann Broder Schmidt, Ariana Barreau, Rajat Rohatgi. Phase separation-deficient TDP43 remains functional in splicing. Nat Commun . 2019 Oct 25;10(1):4890.

14. Schmidt, H. B., and Rohatgi, R. (2016) In vivo formation of vacuolated multi-phase compartments lacking membranes. Cell Rep. 16, 1228 –1236.

15. Babinchak WM, Haider R, Dumm BK, Sarkar P, Surewicz K, Choi JK, Surewicz WK. The role of liquid–liquid phase separation in aggregation of the TDP-43 low-complexity domain. J Biol Chem . 2019 Apr 19;294(16):6306-6317.

16. 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. 2016; 14: e1002338.

17. Hyman AA, Weber CA, Jülicher F. Liquid-liquid phase separation in biology. Annu Rev Cell Dev Biol. 2014; 30; 39–58. pmid:25288112.

18. Leningher A. Principles of Biochemistry. New York: W. H. Freeman and Company; 2005.

19. Patel A, Malinovska L, Saha S, Wang J, Alberti S, Krishnan Y, et al. ATP as a biological hydrotrope. Science. 2017; 356: 753–756.

20. Kang J, Lim L, Lu Y, Song J. A unified mechanism for LLPS of ALS/FTLD-causing FUS as well as its modulation by ATP and oligonucleic acids. PLoS Biol. 2019 Jun 12;17(6):e3000327.

21. Kang J, Lim L, Song J. ATP binds and inhibits the neurodegeneration-associated fibrillization of the FUS RRM domain. Commun Biol. 2019: 2:223.

22. Jiang LL, Che MX, Zhao J, Zhou CJ, Xie MY, Li HY, He JH, Hu HY. Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. J Biol Chem . 2013 Jul 5;288(27):19614-24.

23. Wang L, Kang J, Lim L, Wei Y, Song J. TDP-43 NTD can be induced while CTD is significantly enhanced by ssDNA to undergo liquid-liquid phase separation. Biochem Biophys Res Commun. 2018 May 5;499(2):189-195.

24. Song J. Why do proteins aggregate? "Intrinsically insoluble proteins" and "dark mediators" revealed by studies on "insoluble proteins" solubilized in pure water. F1000Research. 2013; 2: 94.

25. Taratuta VG, Holschbach A, Thurston GM, Blankschtein D. Benedek GB. Liquid–liquid phase separation of aqueous lysozyme solutions: effects of pH and salt identity J. Phys. Chem. 1990; 94: 2140– 472.

26. Brangwynne CP, Eckmann CR, Courson DS, Rybarska A, Hoege C, et al. 2009. Germline Pgranules are liquid droplets that localize by controlled dissolution/condensation. Science 324:1729–32

27. Choi J. M., Holehouse A. S., Pappu R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. Annu Rev Biophys . 2020 May 6;49:107-133.

28. Dignon GL, Best RB, Mittal J. Biomolecular Phase Separation: From Molecular Driving Forces to Macroscopic Properties. Annu Rev Phys Chem . 2020 Apr 20;71:53-75.

29. Berry J, Brangwynne CP, Haataja M. Physical principles of intracellular organization via active and passive phase transitions. Rep Prog Phys. 2018; 81: 046601. pmid:29313527.

30.. Flory PJ. 1942. Thermodynamics of high polymer solutions. J. Chem. Phys. 10:51–61

31. Srivastava S. Tirrell MV. Polyelectrolyte Complexation. In: Rice SA, Dinner AR, eds. Advances in Chemical Physics, 161. Hoboken, NJ: John Wiley & Sons; 2016. pp. 499–544.

32. Cao Q, Boyer DR, Sawaya MR, Ge P, Eisenberg DS. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. Nat Struct Mol Biol. 2019 Jul;26(7):619-627.

33. Qamar, S.; Wang, G.; Randle, S. J.; Ruggeri, F. S.; Varela, J. A.; Lin, J. Q.; Phillips, E. C.; Miyashita, A.; Williams, D.; Stro ̈hl, F., et al. FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-π interac- tions. Cell 2018, 173, 720–734.

34. Pak, C. W.; Kosno, M.; Holehouse, A. S.; Padrick, S. B.; Mittal, A.; Ali, R.; Yunus, A. A.; Liu, D. R.; Pappu, R. V.; Rosen, M. K. Sequence determinants of intra- cellular phase separation by complex coacervation of a disordered protein. Molecular cell 2016, 63, 72–85.

35. Wang, J.; Choi, J.-M.; Holehouse, A. S.; Lee, H. O.; Zhang, X.; Jahnel, M.; Maha- rana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D., et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. Cell 2018, 174, 688–699.

36. He Y, Kang J, Song J. ATP antagonizes the crowding-induced destabilization of the human eye-lens protein γS-crystallin. Biochem Biophys Res Commun. 2020 526:1112-1117.

37. He Y, Kang J, Song J. ATP differentially antagonizes the crowding-induced destabilization of human γS-crystallin and its four cataract-causing mutants. Biochem Biophys Res Commun. 2020 In press.

38. Thandapani P, O'Connor TR, Bailey TL, Richard S. Defining the RGG/RG motif. Mol Cell. 2013; 50: 613-623.

39. Guccione E, Richard S. The regulation, functions and clinical relevance of arginine methylation. Nat Rev Mol Cell Biol. 2019 Oct;20(10):642-657.

40. Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovsky A, Poser I, et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. Science. 2018; 360: 918–921.

41. Dang M, Kang J, Lim L, Li Y, Wang L, Song J. ATP is a cryptic binder of TDP-43 RRM domains to enhance stability and inhibit ALS/AD-associated fibrillation. Biochem Biophys Res Commun. 2020 Jan 29;522(1):247-253.

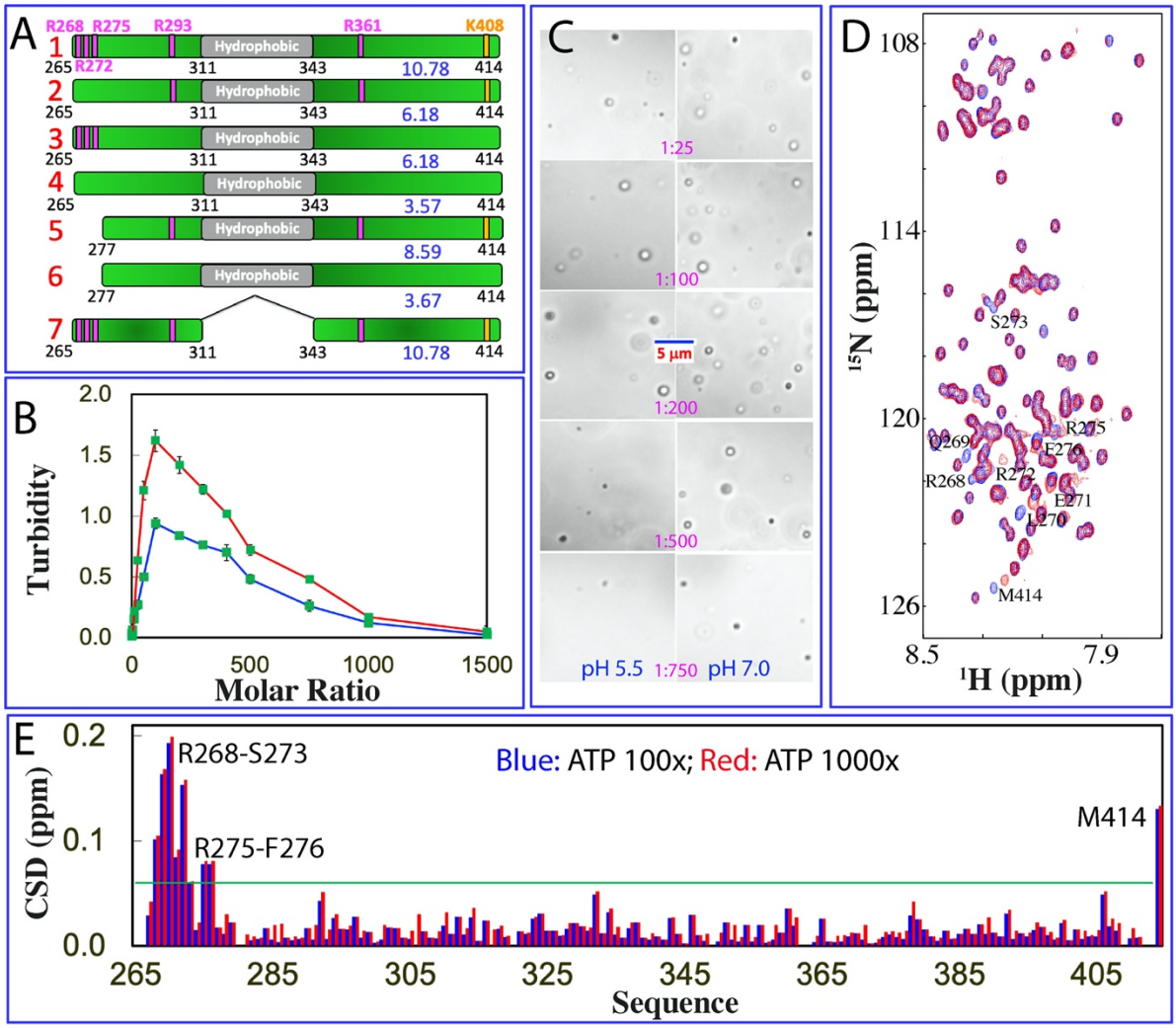

Figure 1. DIC and NMR visualization of interactions of WT TDP-43 PLD with ATP. (A) Schematic representation of seven constructs of TDP-43 PLD studied here, including WT and six mutants. (B) Turbidity curves of WT TDP-43 PLD at 15 µM in the presence of ATP at different molar ratios in 10 mM sodium phosphate buffer at pH 5.5 (green) and 7.0 (red). (C) DIC images of TDP-43 PLD at 15 µM in the presence of ATP at different molar ratios in 10 mM sodium phosphate buffer respective at pH 5.5 and 7.0. (D) ^{1}H-^{15}N NMR HSQC spectra of the ^{15}N-labeled TDP-43 PLD at 15 µM in the absence (blue) and in the presence of ATP at a molar ratio of 1:1000 (red) in 10 mM sodium phosphate buffer (pH 5.5). (E) Chemical shift difference (CSD) of TDP-43 PLD between in the absence and in the presence ATP at 1:100 (blue) and 1:1000 (red). The green line is used to indicate the value: Average + STD at the ATP ratio of 1:1000. The residues with CSD values > Average + STD are defined as significantly perturbed residues.

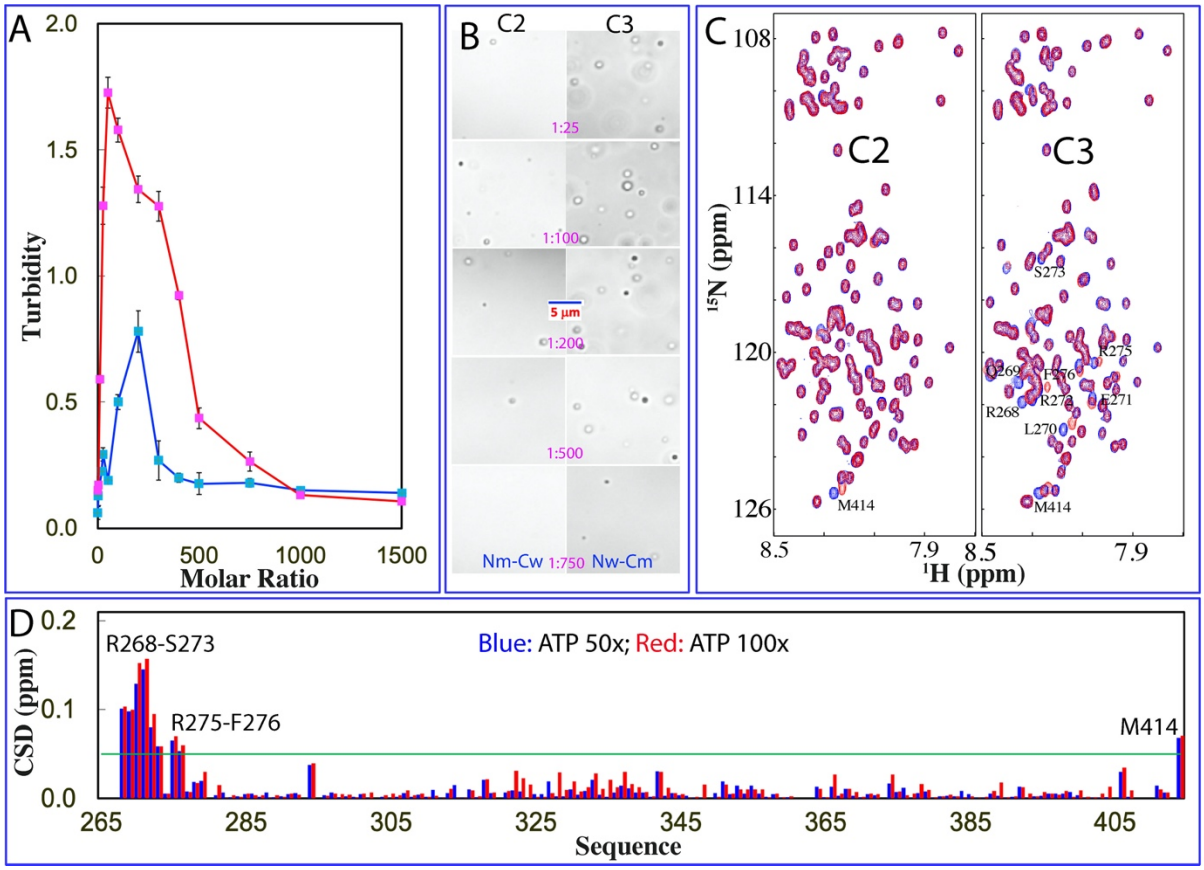

Figure 2. DIC and NMR visualization of interactions of C2- and C3-PLD with ATP. (A) Turbidity curves of TDP-43 C2- (blue) and C3-PLD (red) at 15 µM in the presence of ATP at different molar ratios in 10 mM sodium phosphate buffer at pH 5.5. (B) DIC images of TDP-43 C2- and C3-PLD at 15 µM in the presence of ATP at different molar ratios in 10 mM sodium phosphate buffer respective at pH 5.5. (C) HSQC spectra of the ^{15}N-labeled TDP-43 C2-PLD (I) at 15 µM in the absence (blue) and in the presence of ATP at a molar ratio of 1:50, as well as C3-PLD (II) at 15 µM in the absence (blue) and in the presence of ATP at a molar ratio of 1:100. (D) Chemical shift difference (CSD) of TDP-43 C3-PLD between in the absence and in the presence ATP at 1:50 (blue) and 1:100 (red). The green line is used to indicate the value: Average + STD at 1:1000.

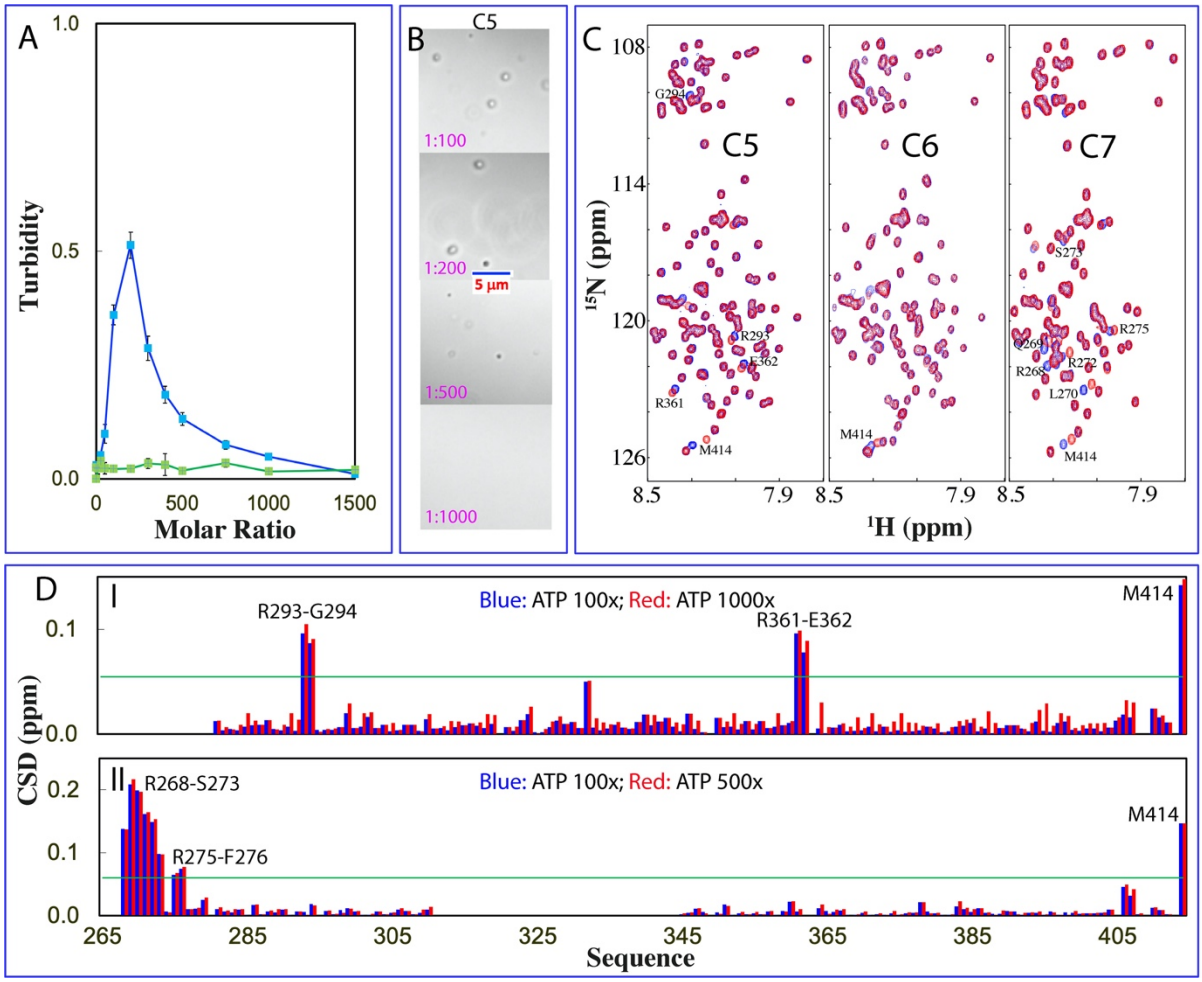

Figure 3. DIC and NMR visualization of interactions of C5- and C7-PLD with ATP. (A) Turbidity curves of TDP-43 C5- (blue) and C7-PLD (green) at 15 µM in the presence of ATP at different molar ratios in 10 mM sodium phosphate buffer at pH 5.5. (B) DIC images of TDP-43 C5-PLD at 15 µM in the presence of ATP at different molar ratios. (C) HSQC spectra of the ^{15}N-labeled TDP-43 C5- (I) and C7-PLD (III) at 15 µM, as well as C6-PLD (II) at 10 µM in the absence (blue) and in the presence of ATP at a molar ratio of 1:1000. (D) Chemical shift difference (CSD) of TDP-43 C5- (I) and C7-PLD (II) between in the absence and in the presence ATP at 1:100 (blue) and 1:1000 (red). The green line is used to indicate the value: average and STD values at 1:1000.

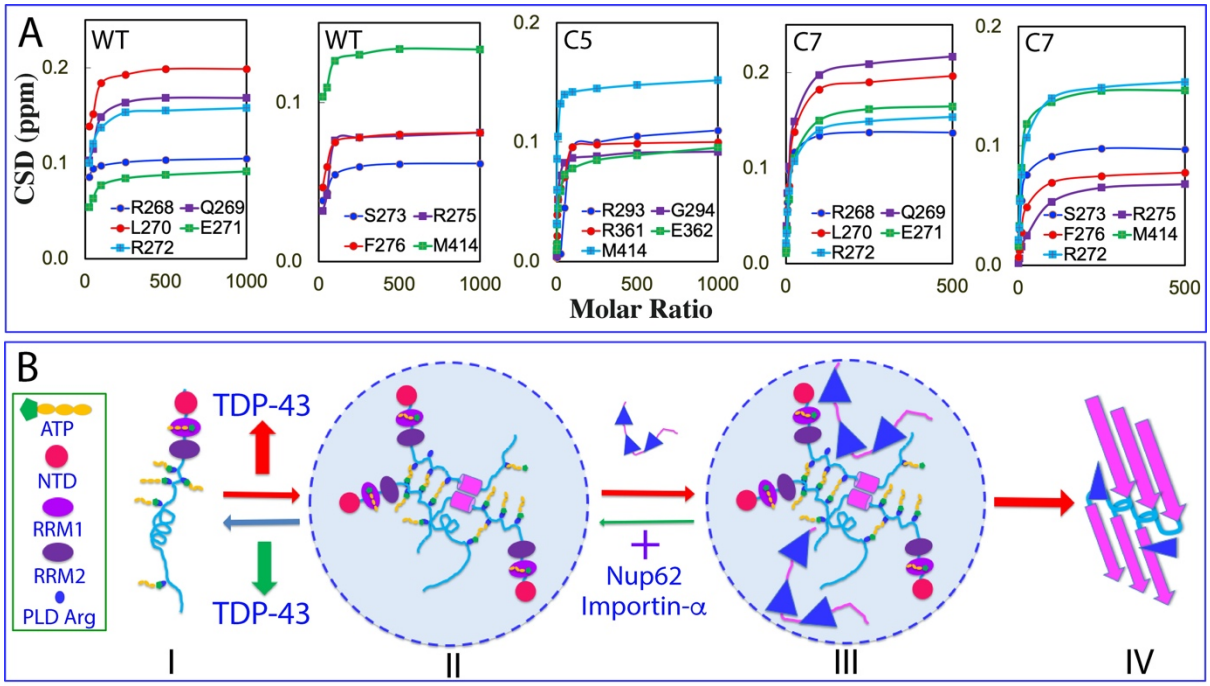

Figure 4. Speculative model for cytoplasmic TDP-43 as modulated by ATP
(A) Concentration-dependent CSD for significantly-shifted residues of WT TDP-43 PLD, C5- and C7-PLD. (B) (I) Cytoplasmic TDP-43 in neurons is not phase separated due to being highly bound with ATP under the physiological conditions in which TDP-43 has low concentrations of ~1 μM while ATP has concentrations of ~3 mM. (II) Cytoplasmic TDP-43 reversibly phase separates into dynamic droplets which may be enhanced by ATP under pathological conditions with a accumulation of cytoplasmic TDP-43, or/and with a significant reduction of ATP concentrations. (III) Cytoplasmic TDP-43 droplets are capable of recruiting other proteins such as importer-α and Nup62. (IV) With long incubation time, TDP-43 droplets may undergo the irreversible transition from the dynamic droplets into aggregation or/and amyloid fibrillation, which ultimately leads to the death of neurons.