Phe-Gly motifs drive fibrillization of TDP-43's prion-like domain droplets

David Pantoja-Uceda[1#], Cristiana Stuani[2#], Douglas V. Laurents[1], Ann E. McDermott[3], Emanuele Buratti[2], Miguel Mompeán[1,3,*]

[1] "Rocasolano" Institute for Physical Chemistry, Spanish National Research Council, Serrano 119, 28006, Madrid, Spain

[2] International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano 99, I-34149, Trieste, Italy

[3] Department of Chemistry, Columbia University, New York, NY 10027, USA

* Correspondence: mmompean@iqfr.csic.es

#These authors contributed equally to this work

Abstract

TDP-43 forms functional and pathological amyloid-like aggregates, with roles in disparate scenarios such as muscle regeneration or neurodegenerative diseases. Using both solution and solid-state NMR, we identify critical regions involved in the monomer to droplet to fibril transitions. This dynamic picture illustrates distinct contributions of Trp-Gly and Phe-Gly motifs: whereas the first mediate droplet assembly but remain flexible, the Phe-Gly repeats form key elements of the fibril core.

Transactive response DNA-binding Protein of 43 kDa (TDP-43) is an RNA-binding protein that forms aberrant aggregates associated to disease contexts, including frontotemporal dementia, amyotrophic lateral sclerosis (ALS)[1], and "LATE" dementia, which had been misdiagnosed as Alzheimer's disease[2]. TDP-43 also forms functional amyloids[3], and participates in at least five different granules[4,5], which highlight the need for structural studies to understand the interplay between these aggregate forms. Structurally, TDP-43 contains a well-folded N-terminal[6], and two RRM domains[7], and a low-complexity, prion-like domain (PrLD) at its C-terminus, which seems pivotal for the formation of diverse amyloid and droplet forms. The PrLD features a hydrophobic helix followed by a Q/N-rich region, surrounded by two Aromatic+Gly-rich segments. Consensus exists that the hydrophobic helix (residues 320-340) in the PrLD[8], with moderate assistance of Trp and other aromatic residues[9], is essential for protein droplet formation via liquid-liquid phase separation (LLPS). Using polypeptide constructs that excluded the two Aromatic+Gly-rich regions, but contained the critical helical (residues 320-340) and Q/N-rich (341-360) segments, cryo-EM structures of polymorphic fibrils were obtained and showed a common "dagger shape" fold[10]. A most recent study used solid-state NMR (SSNMR) to structurally characterize fibrils corresponding to this same TDP-43 (311-360) region and, in contrast to the cryo-EM study, SSNMR revealed the presence of five short β-strands in fibrils from TDP-43(311-360)[11]. Interestingly, these fibrils stem from an intermediate droplet stage, which seems to be true for the entire PrLD of TDP-43[12], as observed for hnRNPA1[13] and other proteins[14]. However, we do not know if the structures formed by the TDP-43(311-360) constructs recapitulate those adopted by the entire PrLD that contains two regions rich in Aromatic-Gly motifs flanking the 311-360 core (*i.e.* at 267-309 & 361-414). These two regions weakly interact during droplet assembly[9] and, interestingly, such symmetric distributions of aromatic residues have just been proposed to encode reversibility of LLPS over amyloid aggregation, for other analogous proteins without a hydrophobic helix like that found in TDP-43[15]. We also lack a well-defined set of conditions that reflects the intrinsic aggregation properties of TDP-43. For instance, ALS has a high incidence in athletes, and intense metabolic activity lowers pH, which induces stress granule formation where TDP-43 is reported to form harmful ALS-relevant aggregates[5]. This, together an uncovered, potentially relevant role of aromatic interactions within droplets that cannot be detected in model fibrils, prompted us to structurally characterize LLPS and amyloid formation at pH 4 by the complete PrLD of TDP-43 using liquid state (LS) and solid state (SS) NMR spectroscopies.

The LSNMR ^{13}CO, ^{13}Cα, ^{13}Cβ, ^{15}N and ^{1}HN correlations were obtained using a non-conventional strategy that affords more robust and complete assignments with a minimal set of experiments (see *Online Methods*). In particular, this approach proves useful for low complexity stretches and aggregation-prone samples, and allowed us to assign all residues under the conditions of this study (BMRB entry 50154), even some that unexpectedly did not match reported assignments[8]. The ^{1}H-^{15}N HSQC spectrum of freshly dissolved TDP-43 PrLD shows peaks of approximately uniform intensity, except for residues 321-340 (**Fig. 1a**), for which the ^{13}Cα conformational chemical shifts evidence the formation of α-helical conformations (**Fig. 1b**). Over the course of several hours, the signal intensity of these peaks decreases. This signal loss is interpreted as the formation of high oligomer species with fast transverse relaxation times, as advanced by Fawzi and co-workers in their seminal study[8], under very distinct conditions. Other signals corresponding to aromatic residues and their immediate neighbors show reductions in signal intensity, yet to a minor extent (**Fig. 1b**). When the experiment was repeated at a higher TDP-43 PrLD concentrations, the pattern of signal loss was similar but the rate of the decrease was quicker (**Fig. 1b**). On average, more signal is lost in the C-terminal

side of the PrLD (residues 360-414) than the N-terminal side (residues 267-320) (**Fig. 1b**). This highlights the importance of the former region, which was excluded from previous solid state Zhuo *et al.*[11] and cryo-EM structural studies Cao *et al.*[10]. Importantly, these results are in excellent agreement with droplet formation at nearly physiological pH and salt conditions, indicating that this helix-driven mechanism of droplet formation may be an intrinsic property of the PrLD manifesting over a broad range of conditions, including at pH values below those normally considered physiological but relevant for metabolic stress conditions and lysosome interiors[16].

We observed that, over time, the NMR samples became aggregated and presented extensive droplet formation as well as amyloid-like fibrils (**Fig. 1d** and **Fig.2a**), in agreement with fibrillization coupled to phase separation as reported for both a TDP-43(311-360) polypeptide[11], and the full PrLD (residues 267-414)[12]. Based on the observation that aromatic residues are involved in droplet formation to a lesser extent that the helix, and challenging the recent model of uniform distribution of aromatic residues as a driver for LLPS, but not amyloid formation[15], we explored whether the symmetrically distributed regions containing Aromatic-Gly motifs would form part of the fibril core by taking advantage of their unique structural signature in the NMR spectra. Accordingly, we analyzed the aggregated material that became invisible to LSNMR by means of SSNMR (**Figure 2b**). Although full structure determination by SSNMR requires multiple samples with different labeling schemes[17], CPMAS-based experiments are sensitive to molecular mobility, allowing us to scrutinize the presence of aromatic residues immobilized within the fibril core. **Figure 2b** shows a ^{13}C-^{13}C spectrum recorded under appropriate conditions to ensure detection of sequential residues (*i.e.*, j and j+1) that remain static and rigid. No Trp residues appear in the characteristic region corresponding to aromatic ^{13}C nuclei. This result is curious because Trp residues are key to LLPS[9], and crucial for amyloid formation in the context of the shorter 311-360 TDP-43's PrLD polypeptide segment[10,11]. One potential explanation could be that Trp residues are exposed in the fibril and the aromatic rings are on average dynamic. In contrast, many peaks corresponding to Phe residues are observed (**Figure 2b**). The PrLD of TDP-43 contains seven Phe-Gly / Gly-Phe motifs (**Fig. 1b, bottom**), with only one such motif present in the 311-360 region studied by the cryo-EM or SSNMR studies cited above[10,11]. In our experiments, we identify contacts that correspond to several Phe-Gly and Ser-Gly residues (**Fig. 2b**), which unambiguously indicate that these motifs are immobilized within a fibrillar core which contains β-rich segments, as evinced by distinctive cross-peaks in the Cα/Cβ serine region (**Fig. 2b**). This is reminiscent of the structural model of FUS fibrils, in which the amyloid core is formed by Ser/Gly/Aromatic-rich motifs[18]. On this basis, we propose a molecular picture in which the helix-helix interactions that drive droplet formation facilitate the interaction of aromatic residues within and between PrLD molecules. As these interactions increase and droplet-driven fibrillization begins, the Phe-Gly and Ser-Gly motifs become an essential part of a large fibrillar core. Importantly, since the mechanism for droplet formation presented here at pH 4 is akin to that reported at nearly-physiological conditions, as anticipated by the low content of titratable residues, the same fibrillar structures involving Phe-Gly and Ser-Gly motifs will be also broadly present, at least in the context of this droplet-driven mechanism. In this regard, Shenoy *et al.*[19] have recently reported fibril formation by the PrLD of TDP-43 under different conditions at pH 7.5, and found that residues G380-I383 and A391-G396 are part of the amyloid core. As illustrated in **Fig. 1b**, those residues flank Phe-Gly motifs within the C-terminal part of TDP-43's PrLD.

To date, several groups have presented studies of intrinsically disordered protein domains undergoing LLPS, wherein the microdroplets formed eventually evolve into amyloid[11,13,14]. Until now, it has been generally assumed that the residues responsible for forming the microdroplet are also those that become amyloid[11]. Here, we have combined LSNMR and SSNMR to obtain atomic level information on both processes. In contrast to previous expectations, we have uncovered that certain residues are responsible for LLPS whereas others are key for amyloidogenesis, including symmetrically distributed aromatic residues proposed to enable reversible droplet formation and not fibrillization[15]. This discovery has important implications for our fundamental understanding of these processes. More excitingly, it means that we should be able to develop inhibitors or modulators that target the residues that are key for one process (*i.e.* the Phe residues which promote amyloidogenesis), without affecting the other residues (namely the Trp residues and the hydrophobic helix which promote microdroplet formation). We close by advancing that, whereas a structure determination is underway, our current observations readily afford a novel picture not contemplated within the current state of knowledge based on the cited SSNMR and cryo-EM structural work.

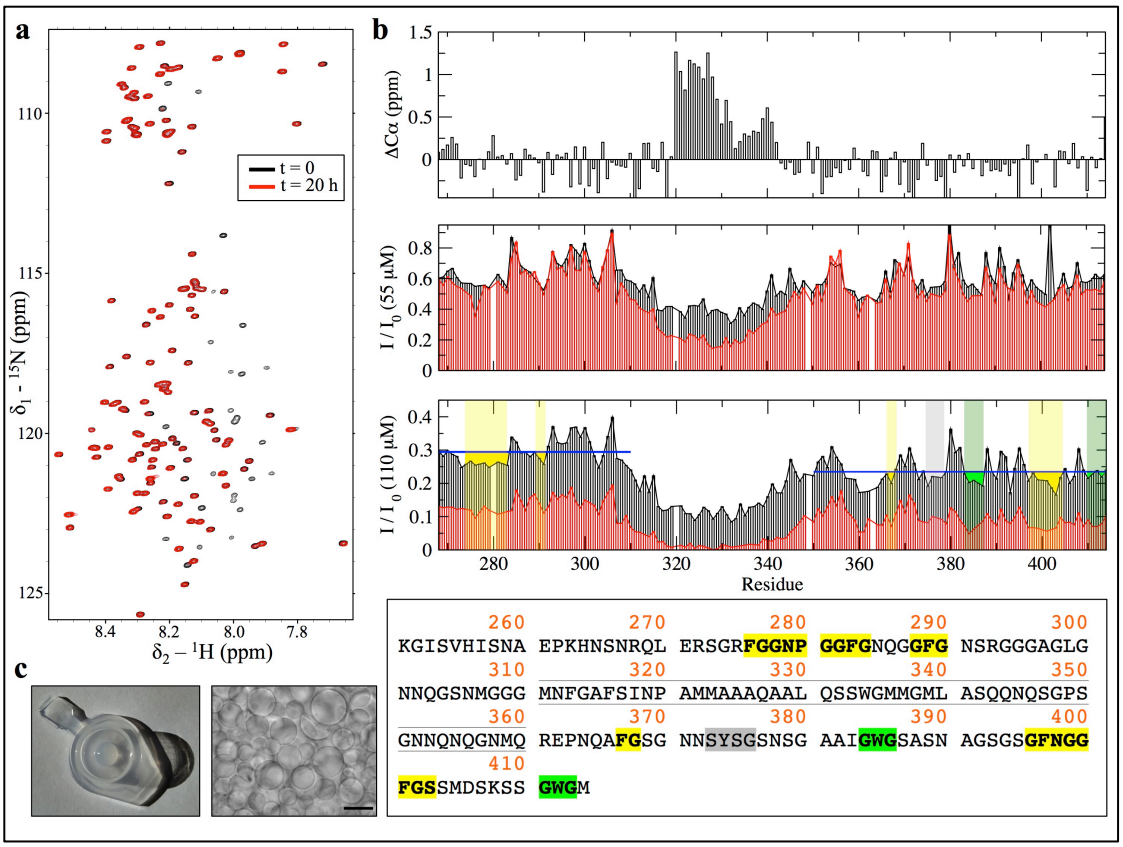

Fig 1 | The PrLD of TDP-43 enables extensive droplet formation at low pH. a, The ^{1}H-^{15}N HSQC spectrum of TDP-43's PrLD at pH 4, 25ºC showing a narrow ^{1}H amide chemical shift range typical of a disordered region. Black peaks correspond to a freshly prepared sample, while red peaks correspond to the same sample after 20 h. **b**, ^{13}Cα conformational chemical shifts reveal a helical region spanning residues 321-340 (top). Plot of the intensity ratios (I/I_0) of aged (20 h, red bars) and fresh samples (black bars) at protein concentration of 55 μM (middle) and 110 μM. Blue lines before and after the helical segment are averaged I/I_0 values, indicating FG (shaded yellow) and WG (shaded green) motifs mapped onto the PrLD sequence. **c**, After the NMR experiments, all samples are aggregated into dense phases and displayed extensive droplet formation (bar scale 100 μm).

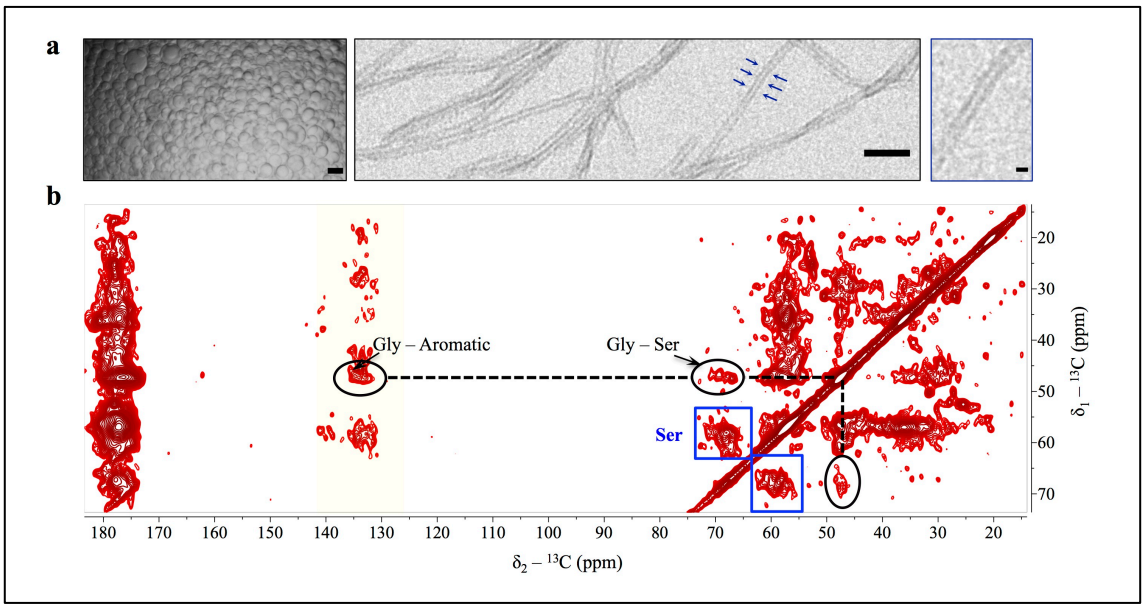

Fig 2 | Fibrils from the PrLD of TDP-43 are stabilized by Phe-Gly motifs. a, Images of droplet (left, scale bar 100 μm) and fibrils (right, scale bar 100 nm and 10 nm for the magnified fibril indicated by blue arrows). **b**, 2D ^{13}C-^{13}C DARR spectrum (250 ms of mixing) of fibrils formed following droplet formation. This experiment detects residues immobilized within the fibril core. Fibrils are rich in β-sheet content, as indicated by the presence of Ser residues showing differences of ^{13}Cβ -^{13}Cα > 8 ppm (blue boxes). Cross-peaks between Ser and Gly (black circles) indicate that Ser-Gly motifs, which are located outside the helical region, are relevant in the fibril form. The lack of signals within 110-130 ppm suggests that Trp aromatic rings may not stabilize the fibril form. In contrast, signals around 130-140 ppm (yellow shaded region) correspond to aromatic ^{13}C from Phe/Tyr residues. Considering that there is only one Tyr in the PrLD sequence, most cross-peaks arise from Phe (^{13}Cδ/^{13}Cε) – Gly (^{13}Cα) (black circles). Importantly, sequential contacts were not observed at shorter mixing times.

Acknowledgments

This work has been funded by Grants CTQ2017-84371-P to D.P.-U. and SAF2016-76678-C2-2-R to D.V.L., from the Spanish MINECO; Grant MCB1412253 from the U. S. National Science Foundation to A.E.M; AriSLA (PathensTDP project) to E.B.; and Grant LCF/BQ/PR19/11700003 from La Caixa Foundation (ID 100010434) to M.M.
NMR experiments were performed in the "Manuel Rico" NMR Laboratory (LMR) of the Spanish National Research Council (CSIC), a node of the Spanish Large-Scale National Facility (ICTS R-LRB).

References

1 Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130-133, doi:10.1126/science.1134108 (2006).

2 Nelson, P. T. *et al.* Limbic-predominant age-related TDP-43 encephalopathy (LATE): consensus working group report. *Brain : A Journal of Neurology* **142**, 1503-1527, doi:10.1093/brain/awz099 (2019).

3 Vogler, T. O. *et al.* TDP-43 and RNA form amyloid-like myo-granules in regenerating muscle. *Nature* **563**, 508-513, doi:10.1038/s41586-018-0665-2 (2018).

4 Chu, F.-J., *et al*. TDP-43 regulates coupled dendritic mRNA transport-translation processes in cooperation with FMRP and Staufen1. *Cell Reports* **29**, 3118-3133.e6, doi: 10.1016/j.celrep.2019.10.061 (2019).

5 Zhang, P. *et al.* Chronic optogenetic induction of stress granules is cytotoxic and reveals the evolution of ALS-FTD pathology. *eLife* **8**, doi:10.7554/eLife.39578 (2019).

6 Mompean, M. *et al.* The TDP-43 N-terminal domain structure at high resolution. *The FEBS Journal* **283**, 1242-1260, doi:10.1111/febs.13651 (2016).

7 Lukavsky, P. J. *et al.* Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nature Structural & Molecular Biology* **20**, 1443-1449, doi:10.1038/nsmb.2698 (2013).

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

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

10 Cao, Q., Boyer, D. R., Sawaya, M. R., Ge, P. & Eisenberg, D. S. Cryo-EM structures of four polymorphic TDP-43 amyloid cores. *Nature Structural & Molecular Biology* **26**, 619-627, doi:10.1038/s41594-019-0248-4 (2019).

11 Zhuo, X. F. *et al.* Solid-State NMR Reveals the Structural Transformation of the TDP-43 Amyloidogenic Region upon Fibrillation. *Journal of the American Chemical Society*, doi:10.1021/jacs.9b10736 (2020).

12 Babinchak, W. M. *et al.* The role of liquid-liquid phase separation in aggregation of the TDP-43 low-complexity domain. *The Journal of Biological Chemistry* **294**, 6306-6317, doi:10.1074/jbc.RA118.007222 (2019).

13 Gui, X. *et al.* Structural basis for reversible amyloids of hnRNPA1 elucidates their role in stress granule assembly. *Nature Communications* **10**, 2006, doi:10.1038/s41467-019-09902-7 (2019).

14 Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123-133, doi:10.1016/j.cell.2015.09.015 (2015).

15 Martin, E. W. *et al.* Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694-699, doi:10.1126/science.aaw8653 (2020).

16 Settembre, C., Fraldi, A., Medina, D. L., & Ballabio, A. Signals from the lysosome: a control centre for cellular clearance and energy metabolism. *Nat. Rev. Mol. Cell Biol.* **14**, 283-296, doi:10.1038/nrm3565 (2013).

17 Mompean, M. *et al.* The Structure of the Necrosome RIPK1-RIPK3 Core, a Human Hetero-Amyloid Signaling Complex. *Cell* **173**, 1244-1253 e1210, doi:10.1016/j.cell.2018.03.032 (2018).

18 Murray, D. T., *et al.* Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell* **171**, 615-627, doi:10.1016/j.cell.2017.08.048 (2017).

19 Shenoy, J. *et al.* Structural dissection of amyloid aggregates of TDP-43 and its C-terminal fragments TDP-35 and TDP-16. *The FEBS Journal*, doi:10.1111/febs.15159 (2019).

20 Kjaergaard, M., Brander, S. & Poulsen, F. M. Random coil chemical shift for intrinsically disordered proteins: effects of temperature and pH. *Journal of Biomolecular NMR* **49**, 139-149, doi:10.1007/s10858-011-9472-x (2011).

21 Kjaergaard, M. & Poulsen, F. M. Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *Journal of Biomolecular NMR* **50**, 157-165, doi:10.1007/s10858-011-9508-2 (2011).

22 Schwarzinger, S. *et al.* Sequence-dependent correction of random coil NMR chemical shifts. *Journal of the American Chemical Society* **123**, 2970-2978, doi:10.1021/ja003760i (2001).

23 Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR* **6**, 277-293, doi:10.1007/bf00197809 (1995).

24 Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325-1327, doi:10.1093/bioinformatics/btu830 (2015).

25 Takegoshi, K., Nakamura, S., & Terao, T. (2001). 13C–1H dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **344**, 631–637, doi:10.1016/S0009-2614(01)00791-6 (2001).

Online Methods

Protein expression and purification. The N-terminally hexa-His-tagged PrLD construct corresponding to TDP-43(267-414) was a gift from Prof. Nicolas Fawzi (Addgene plasmid # 98669). Proteins were overexpressed in *E. coli* BL21 Star (DE3) (Life Technologies). Uniform ^{15}N and $^{15}N/^{13}C$ labeling of TDP-43(267-414) was achieved by overexpression in M9 minimal media supplemented with ^{15}N ammonium chloride and ^{13}C glucose as the sole sources of nitrogen and carbon, respectively. *E. coli* BL21 Star (DE3) carrying TDP-43(267-414) were pre-inoculated overnight at 37°C. The day after, 200 ml of M9 minimal media were added to 4 mL of pre-inoculum and the cell growth was monitored using Biophotometer D30 (Eppendorf, Hamburg, Germany) until the growth rate was 0.8 (OD at 600 nm). Then, the protein expression was induced overnight at 30°C by adding 1 mM IPTG.

Cells were collected by centrifugation at 7,000 × g for 1 hour 4°C and the resulting pellet was resuspended in 30 mL lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM Imidazole pH 8.0) supplemented with one tablet EDTA-free protease inhibitor cocktail (Roche Diagnostics) and 1 mM PMSF (SigmaAldrich, St Louis, MO, USA). Cell lysis were performed using EmulsiFlex-C3 (Avestin Europe GmbH, Mannheim, Germany) and Bioruptor UCD-200 (Diagenode, Belgium). Inclusion bodies were recovered by centrifugation of cell lysate at 7,000 x g for 1 hour at 4°C and were resuspended in 20 mL denaturing binding buffer (20 mM Tris-Cl, 500 mM NaCl, 10 mM Imidazole, 1 mM DTT, 8 M urea, pH 8.0). Further centrifugation at 7,000 × g for 1 hour, 4°C was performed in order to pellet any remaining insoluble cell debris.

Ni-NTA Agarose beads (Qiagen Inc, Gaithersburg, MD, USA) were used to immunoprecipitate (IP) the target protein. Before the IP reaction, beads were washed three times in 15 mL of denaturing binding buffer and each washing step was interspersed with centrifugation at 500 × g for 5 minutes, RT. Beads were then incubated with cell lysate for 4 hours rotating and successively collected by centrifugation at 500 × g for 5 minutes, RT. Finally, beads were washed three times in 2 ml of denaturing binding buffer and protein was eluted with imidazole gradient buffers, as described below:

(1) 1 ml 20 mM Tris-Cl, 500 mM NaCl, 250 mM imidazole, 1 mM DTT, 8 M urea, with rotation of 2 hours at RT.
(2) 1ml, 20 mM Tris-Cl, 500 mM NaCl, 300mM imidazole, 1 mM DTT, 8 M urea, with rotation overnight at 4°C.
(3) 1ml 20 mM Tris-Cl, 500 mM NaCl, 300mM imidazole, 1 mM DTT, 8 M urea, with rotation of 2 hours at RT.
(4) 1ml 20 mM Tris-Cl, 500 mM NaCl, 500 mM imidazole, 1 mM DTT, 8 M urea, with rotation of 6 hours at RT.
Each elution step was interspersed with centrifugation at 500 ×g for 5 minutes, RT.

The pH of the PrLD constructs was then lowered to 4.0, and applied to a PD-10 gel filtration column (GE Healthcare Ltd, Little Chalfont, UK), which had been previously pre-equilibrated with the LSNMR buffer: 1 mM CD_3COOD in 85/15 H_2O/D_2O. The eluted fractions containing the PrLD construct, as identified by UV absorbance, were then concentrated using an ultrafiltration device (Amicon Corp., Danvers, MA, USA) to concentrations between 50 and 120 micromolar.

LSNMR experiments. Liquid state NMR experiments were collected at pH 4.0 and 298 K on a Bruker Avance II 800 MHz (^{1}H frequency) spectrometer equipped with a cryoprobe and Z-gradients at two protein concentrations of 55 and 110 μM. To obtain NMR assignments, we combined the following set of experiments:

(1) ^{1}H-^{15}N HSQC (4 scans, 12 and 20 ppm as spectral widths for ^{1}H and ^{15}N, respectively, and transmitter frequency offsets of 4.75 and 116.5 ppm for ^{1}H and ^{15}N, respectively).
(2) H(NCOCA)HN (12 scans, 12, 20, and 4.25 ppm as ^{1}H, ^{15}N, and ^{1}H spectral widths, respectively, and transmitter frequency offsets of 4.75, 116.5, and 7.125 ppm for ^{1}H, ^{15}N, and ^{1}H, respectively).
(3) (H)N(COCA)NH (12 scans, 12, 20, and 20 ppm as ^{1}H, ^{15}N, and ^{1}H spectral widths, respectively, and transmitter frequency offsets of 4.75, 116.5, and 116.5 ppm for ^{1}H, ^{15}N, and ^{15}N, respectively).
These experiments provided unambiguous sequential connectivities between spin systems, which were assigned with the following two experiments:
(4) HNCA (4 scans, 12, 20, and 30 ppm as ^{1}H, ^{15}N, and ^{13}C spectral widths, respectively, and transmitter frequency offsets of 4.75, 116.5, and 54 ppm for ^{1}H, ^{15}N, and ^{13}C, respectively)
(5) CBCA(CO)NH (8 scans, 12, 20, and 75 ppm as ^{1}H, ^{15}N, and ^{13}C spectral widths, respectively, and transmitter frequency offsets of 4.75, 116.5, and 44 ppm for ^{1}H, ^{15}N, and ^{13}C, respectively). From these spectra, we obtained ^{13}Cα and ^{13}Cβ that unambiguously afforded the assignment of the PrLD construct at the conditions described in this work. Next, to obtain complete backbone assignments, ^{13}CO and ^{1}Hα nuclei were assigned using the two following spectra:
(6) 3D HNCO (8 scans, 12, 20, and 12 ppm as ^{1}H, ^{15}N, and ^{13}C spectral widths, respectively, and transmitter frequency offsets of 4.75, 116.5, and 174 ppm for ^{1}H, ^{15}N, and ^{13}C, respectively).
(7) HBHA(CO)NH (12 scans, 12, 20, and 6, ppm as ^{1}H, ^{15}N, and ^{1}H spectral widths, respectively, and transmitter frequency offsets of 4.75, 116.5, and 4.75 ppm for ^{1}H, ^{15}N, and ^{1}H, respectively)
Finally, a CC(CO)NH experiment was performed to assign other side chain ^{13}C nuclei distinct from ^{13}Cβ (8 scans, 12, 20, and 75 ppm as ^{1}H, ^{15}N, and ^{13}C spectral widths, respectively, and transmitter frequency offsets of 4.75, 116.5, and 39 ppm for ^{1}H, ^{15}N, and ^{13}C, respectively)

Proton chemical shifts were directly referenced using DSS on a PrLD sample prepared for this purpose, and ^{13}C and ^{15}N chemical shifts chemical shifts were referenced indirectly. Conformational chemical shifts, $\Delta(^{13}C\alpha)$, where calculated as $\delta^{13}C\alpha(exp) - \delta^{13}C\alpha(ref)$, with $\delta^{13}C\alpha(exp)$ being our measured chemical shifts and $\delta^{13}C\alpha(ref)$ reference values obtained using the sequence of the TDP-43 PrLD construct and the data compiled by Poulsen, Dyson, and co-workers, at 25 ºC and pH 4, as implemented in the chemical shift calculator http://spin.niddk.nih.gov/bax/nmrserver/Poulsen_rc_CS[20-22].

All spectra were processed using either NMRPipe[23] or Topspin 4.0.8 (Bruker Biospin, Germany), and peak-picking and spectral assignment was conducted using NMRFAM-Sparky[24]. The NMR chemical shifts are deposited in the Biological Magnetic Resonance Bank (BMRB) under accession code 50154.

SSNMR experiments. Solid-state NMR experiments were recorded on the aggregated material that formed following droplet formation during our LSNMR studies, which was amplified by seeding soluble protein prepared as described above. As a reference, from the most concentrated LSNMR samples (~110 micromolar), 500 microliters contained approximately 1 mg of material assuming that 100% of the protein converted into fibrils. After three weeks of incubation, less than 5 mg of fibrillar material was packed on a rotor and analyzed on a Bruker 17.6 T spectrometer (750 MHz ^{1}H frequency) using an HCN 1.3 mm MAS probe. A 2D ^{13}C-^{13}C Dipolar Assisted Rotational Resonance (DARR) spectrum[25] was recorded with a mixing time of 250 ms, spinning the rotor containing the sample at a MAS rate of 20 kHz. This mixing time seemed appropriate to observe cross-peaks corresponding to sequential (j to j+1) residues, to detect correlations between Gly and Phe, Trp or Tyr, based on the characteristic chemical shifts of the Gly ^{13}Cα and ^{13}Caro chemical shift values distinctive for each type of aromatic residue. A contact time of 2 ms was used for cross-polarization transfer. The spectrum was recorded with 64 scans, using a spectral width of 220.9 ppm in the direct and indirect dimensions, with acquisition times of 20 and 2.5 ms, respectively, and setting the transmitter frequency offset to 100 ppm. The DARR spectrum was processed using Topspin 4.0.8, with chemical shifts referenced to DSS.

Visualization of amyloid fibrils and protein droplets. Two types of samples were directly visualized by confocal and transmission electron microscopies, to check for the presence of protein droplets and amyloid fibrils, respectively. One type of sample was directly collected after our LSNMR experiments **(Figure 1C**), whereas the second type of samples consisted of a larger fraction resulting from collecting up to 8 mg of material prepared under the conditions used for LSNMR. Liquid droplets were directly visualized by spotting the samples onto glass coverslips using a Leica TCS SP2 inverted confocal microscope. For transmission electron microscopy, samples were directly adsorbed onto carbon-coated 300-mesh copper grids, and negatively stained by incubation with 2% uranyl-acetate. Amyloid fibrils were visualized on a JEOL JEM-1011 electron microscope equipped with a TVIPS TemCam CMOS, with images acquired at a magnification of 30,000x and an accelerating voltage of 1,000 kV.