

Point mutations in transactive response DNA-binding protein 43 (TDP-43)'s N-terminal domain compromise its stability, dimerization and functions

Miguel Mompeán<sup>1§</sup>, Valentina Romano<sup>2</sup>, David Pantoja-Uceda<sup>1</sup>, Cristiana Stuani<sup>2</sup>, Francisco E. Baralle<sup>2</sup>, Emanuele Buratti<sup>2,\*</sup>, Douglas V. Laurents<sup>1,\*</sup>

From the:

<sup>1</sup>Instituto de Química Física "Rocasolano", Consejo Superior de Investigaciones Científicas, Serrano 119, E-28006, Madrid, Spain

<sup>2</sup>International Centre for Genetic Engineering and Biotechnology (ICGEB), Padriciano 99, I-34149, Trieste, Italy

<sup>§</sup>Current address: Facultad de Ciencias y Tecnologías Químicas, Universidad de Castilla la Mancha, Avd. C. J. Cela 10, E-13071, Ciudad Real, Spain

Running Title: Stably Folded N-terminal Domain Required for TDP-43 Function

\*To whom Correspondence may be addressed:

[buratti@icgeb.org](mailto:buratti@icgeb.org), [dlaurents@iqfr.csic.es](mailto:dlaurents@iqfr.csic.es)

telephones: +39-040-3757316 (EB) +34 91-561-9400 (DVL)

fax: +39-040-226555 (EB) +34 91-564-2431 (DVL)

**Keywords:** Amyotrophic Lateral Sclerosis (ALS), Microscopic Imaging, Nuclear Magnetic Resonance (NMR), Protein Aggregation, Site-Directed Mutagenesis, Structure-Function, Subcellular fractionation, Transactive DNA-binding protein 43 kDa (TDP-43) (TARDBP)

## ABSTRACT

Transactive response DNA-binding protein 43 (TDP-43) performs multiple tasks in mRNA processing, transport and translational regulation, but it also forms aggregates implicated in amyotrophic lateral sclerosis (ALS). TDP-43's N-terminal domain (NTD) is important for these activities and dysfunctions; however, there is an open debate about whether or not it adopts a specifically folded, stable structure. Here, we studied NTD mutations designed to destabilize its structure utilizing NMR and fluorescence spectroscopies, analytical ultracentrifugation, splicing assays and

in cell microscopy. The substitutions Val 31>Arg and Thr 32>Arg abolished TDP 43 activity in splicing and aggregation processes and even the rather mild Leu 28>Ala mutation severely destabilized the NTD, drastically reducing TDP-43's *in vitro* splicing activity and inducing aberrant localization and aggregation in cells. These findings strongly support the idea that a stably folded NTD is essential for correct TDP-43 function. The stably folded NTD also promotes dimerization which is pertinent to the protein's activities and pathological aggregation, and we present an atomic-level structural model for the TDP-43 dimer based on NMR data. Leu 27 is evolutionarily well conserved even though

**it is exposed in the monomeric NTD. We found here that Leu 27 is buried in the dimer and that the Leu 27>Ala mutation promotes monomerization. In conclusion, our study sheds light on the structural and biological properties of the TDP-43 NTD, indicating that the NTD must be stably folded for TDP-43's physiological functions, and has implications for understanding the mechanisms promoting pathological aggregation of this protein.**

-----

Transactive DNA-Binding Protein 43 kDa (TDP-43) is an essential human protein that is vital to pre- mRNA (1) and microRNA processing (2). Key to these activities are TDP-43's two well-folded and stable RRM domains (spanning residues 106-177 and 192-259, respectively). TDP-43 contains a nuclear localization sequence (NLS, residues 80-102) and is mostly nuclear; however, its nuclear export sequence (NES, residues 238-250) permits it to transport mRNAs to the cytoplasm, and even to synapses as part of neuronal granules (3). TDP-43 also regulates translation by participating in stress granules (4).

Aberrant aggregate forms of TDP-43 are tightly linked to Amyotrophic Lateral Sclerosis (ALS) as inclusions composed of polyubiquitinated, hyperphosphorylated and truncated TDP-43 (5,6) have been reported in 95% of ALS patient motoneurons (7). TDP-43 aggregates are also observed in 60% of Frontotemporal Lobar Degeneration (FTLD: a form of dementia whose symptoms overlap with ALS) patient neurons and the observation of aggregates composed of TDP-43 plus A $\beta$  or polyglutamine suggests that TDP-43 may contribute to other neurodegenerative diseases such as Alzheimer's or Huntington's disease (8,9). Almost all the pathologically linked mutations are localized in the C-terminal region of TDP-43 (CTR, residues 270-414). In contrast to the RRM domains, the CTR is intrinsically disordered and consists of four segments: two of which (residues 267-320 and 367-414) are

rich in G or S, aromatic, G or S motifs reminiscent of G or S, Y, G or S motifs in the RNA-binding protein Fused in Sarcoma / Translocated in Sarcoma (FUS/TLS) which drive the formation of a hydrogel or liquid phase. A third segment (residues 320-340) is a hydrophobic and tends to adopt helical conformations (10-13) and also drives the formation of a distinct, non aqueous liquid phase such as those present in stress granules and neuronal granules (13) (14).

The fourth CTR segment of TDP-43, composed of residues 341 - 367, is "Q/N-rich" as it contains a high proportion of Gln and Asn residues. Whereas variants lacking the Q/N-rich segment do not aggregate, variants containing twelve copies of it recapitulate in cells most of the pathological characteristics seen in ALS (15). In 2015, we advanced an amyloid-like conformer for this segment based on a variety of biochemical, spectroscopic and computational data (16). In early 2016, the Q/N-rich segment, or part of it, was found to be intact in TDP-43 aggregates from *ex-vivo* brain tissue whereas all other segments of the CTR are heavily phosphorylated, deamidated and oxidized (17). Since the exceptionally strong hydrogen-bonding networks of Q/N-rich amyloids (18) could impede such chemical modifications, we have recently interpreted these data as supporting our amyloid-like model of the Q/N-rich segment (19).

In 2012, the N-terminal domain of TDP-43 (NTD) was predicted to adopt a stable fold and was found to drive the formation of large oligomers (20). The NTD is required for TDP-43's physiological functions and pathological aggregation (21-23), but it has been less studied due to its unique sequence, which thwarts structural prediction based on homology modeling, and strong tendency to aggregate. In 2014, an important advance in its structural characterization was reported by Song and coworkers (24), who discovered that NTD constructs with a C-terminal His tag are soluble in low pH, very low ionic strength solution conditions. This construct adopted a minor population of folded conformers; nevertheless, they were able to advance a

medium resolution model for the tertiary fold of the NTD based on a small number of NMR NOE constraints and Rosetta *ab initio* structure prediction.

In 2016, we reported that the NTD is stably folded ( $T_M = 45 - 50^\circ\text{C}$ ) in the context of short (residues 1-77, TDP-43<sub>1-77</sub>) or long (residues 1-102, TDP-43<sub>1-102</sub>) constructs with N-terminal His tags, which permitted the elucidation of the shorter construct's 3D structure to high resolution using NMR methods (25), [PDB=2N4P]. In this study, we observed that the longer construct TDP-43<sub>1-102</sub>, consisting of the NTD plus the NLS, which is rich in cationic residues, is more soluble. The NTD contains two highly conserved consecutive Leu residues at positions 27 and 28. We also found that L27's hydrophobic side chain is mostly solvent exposed, which is unusual for a nonpolar residue, and that L28 is mostly buried and makes important contacts linking different elements of secondary structure. Here, the roles of L27 in promoting dimerization and of L28 in stabilizing the tertiary structure are tested by substituting these residues by Ala. We also study another variant, V31R/T32R designed to introduce two charges into the nonpolar core of the NTD. Since placing even one charge in the hydrophobic core strongly destabilizes proteins (26), this variant is expected to unfold the NTD.

Despite these advances, there is a continuing debate regarding whether the NTD is stably folded and whether this domain needs to be folded for TDP-43 to be active in cells (27) (28)(29). Previously, TDP-43 has been shown to exist as a monomer / dimer *in vivo* (30) and *in vitro* (31). The analysis of truncation mutants (30) and, size exclusion chromatography and a low resolution SAXS envelope (31) indicate the NTD is the domain chiefly responsible for this dimerization. In cells, TDP-43's concentration is exquisitely controlled (32) (33), which further suggests that the monomer / dimer equilibrium could have important functional consequences. However, the conformation of the dimer is currently unknown. Further studies in this area are therefore required, if we consider that TDP-43

affects the maturation and transport of thousands of mRNAs and that changes in TDP-43 concentration, due to aggregation or gene knockdown, strongly alter protein expression (34).

One objective of this manuscript is to test if stably folded NTD is required for TDP-43's activity by characterizing the biological functions of mutations; namely Val 31>Arg, Thr 32>Arg and Leu 28>Ala, designed to disrupt the protein's tertiary structure. The latter variant was chosen for high resolution studies of its structure, stability and dynamics. The second objective is to characterize the conformation of the NTD dimer, on the basis of multidimensional heteronuclear NMR spectroscopy and corroborated by studies of a variant, Leu 27>Ala, which is designed to disrupt the dimer interface without perturbing the structure of the monomer.

## RESULTS

*Characterization of the solution conformation of TDP-43<sub>1-102</sub>.* Taking advantage of the superior solubility of the TDP-43<sub>1-102</sub> construct, we recorded a series of NOESY NMR spectra (2D NOESY and HSQC-NOESY). These spectra yielded hundreds of new NOEs (over 1800 compared to the 1058 peaks used to calculate the structure of TDP-43<sub>1-77</sub>, (25)) and allowed us to determine the NTD's structure to higher resolution (**Figure 1A, Sup. Table 1**). The NMR assignments and final refined structures have been deposited in the Biological Magnetic Resonance Data Bank (BMRB access code: 34081) and the RCSB Protein Data Bank (access code: 5MRG), respectively. Overall, the conformation is very similar to that of the TDP-43<sub>1-77</sub> construct studied previous, except some differences are observed for N76 and Y77 at the end of the last  $\beta$ -strand, and for C39 and G40, which contact those residues. These minor differences could be due to the end effects, particularly the influence of P78 on the conformation of N76 and Y77.

<sup>1</sup>H-<sup>15</sup>N HSQC spectra recorded on <sup>15</sup>N

labeled TDP-43<sub>1-102</sub> which had been transferred into 100% D<sub>2</sub>O buffer, revealed signals belonging to amide groups protected from H/D exchange. This experiment corroborated published results (25) and led to the identification of three new protected residues (S20, Q34 and Y73) (**Figure 1B**). S20 appears to donate an H-bond to E3 in a minority of our family of 20 NMR structures and Q34 and Y73 are H-bonded in the  $\alpha$ -helix and last  $\beta$ -strand, respectively. In our previous *FEBS J* paper, H/D exchange was monitored by 1D <sup>1</sup>H NMR spectra and the protected peaks were identified on the basis of their <sup>1</sup>HN chemical shift and based on a 2D <sup>1</sup>H-<sup>1</sup>H NOESY spectrum that was recorded part way through the exchange. In the present paper, H/D exchange has been followed by 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectra. In the previous experiment, the closeness of the Ser 30 <sup>1</sup>HN signal to those of Thr 25, Ala 38, Val 72, Val 75 and Asn 76, the proximity of Gln 34's <sup>1</sup>HN signal to aromatic side chain <sup>1</sup>H resonances and the closeness of Tyr 73's <sup>1</sup>HN signal to those of Ile 5 and Val 7 made it difficult to identify these slow exchanging peaks. They could be unambiguously identified in the <sup>1</sup>H-<sup>15</sup>N HSQC spectra thanks to the superior separation of resonances in the <sup>15</sup>N dimension.

*The mutation Leu 28>Ala strongly destabilizes TDP-43.* As a further test of the structural integrity of the NTD, the effect of a mutation, **Leu 28>Ala**, designed to disrupt NTD tertiary contacts and hydrophobic core packing, was studied. The <sup>1</sup>H-<sup>15</sup>N HSQC NMR spectrum of this variant, called L28A, shown in **Figure 2A**, revealed a large increase in the number and intensity of peaks clustered in the spectral region (<sup>15</sup>N: 114-126 ppm, <sup>1</sup>H: 8.5-7.9 ppm) typical of denatured protein. For peaks outside this region, analysis of a 3D <sup>1</sup>H-<sup>15</sup>N HSQC-NOESY spectrum allowed the corroboration of the assignment of peaks in the <sup>1</sup>H-<sup>15</sup>N HSQC spectrum which had not shifted relative to their position in wild type TDP-43<sub>1-102</sub>, and the assignment of those which had. The

largest chemical shift perturbations are localized at the mutated site, the second loop, the  $\alpha$ -helix and the turn connecting it to the third  $\beta$ -strand, and the  $\beta$ -hairpin formed by  $\beta$ -strands 4 and 5 (**Figure 2B**). This pattern of perturbations is consistent with the contacts formed by L28's nonpolar side chain in the folded NTD (**Figure S1**).

Two <sup>1</sup>H-<sup>15</sup>N HSQC peaks, for the folded and denatured states, were observed for the side chain HNe of W68 (**Figure 2A**). Based on NMR peak integration, 19 % of L28A TDP-43<sub>1-102</sub> is folded at pH 3.9, 25 °C. Upon lowering the temperature, the folded population, as gauged by this group, increases to 52% at 15 °C and 67% at 5 °C (**Figure S1**). These values are in line with the results of fluorescence-monitored thermal denaturation described below. The lack of signal broadening and the observation of discrete native and denatured peaks for the side chain HNe group of Trp 68 are evidence that the folding/unfolding equilibrium of the L28A variant can be approximated as a two-state process, which is slow on the NMR timescale and that the population of folding intermediates is low.

To get more insight into the L28A variant's conformational stability, its thermal denaturation was followed by fluorescence spectroscopy. Even at low temperature, L28A is not completely folded as its wavelength of maximum emission is 333 nm compared to 329 nm for wt TDP-43<sub>1-102</sub>. The apparent unfolding midpoint temperature is about 15 °C (**Figure 2C**). Although this value is imprecise considering the inability to fit the pre-transition baseline, this apparent T<sub>M</sub> is approximately 30 °C lower than that of the WT TDP-43<sub>1-102</sub> construct, indicating that this variant is strongly destabilized. As will be shown below, this severe destabilization strongly affects TDP-43's subcellular localization, in cell aggregation and its ability to regulate mRNA splicing. By analytical ultracentrifugation, this variant sediments as a monomer; no dimer was observed which is consistent with the lower concentration of folded protein (**Figure 2D**).



NMR relaxation analysis shows increased dynamics in the L28A variant. The heteronuclear  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE of WT TDP-43<sub>1-77</sub> and TDP-43<sub>1-102</sub> constructs and in the TDP-43<sub>1-102</sub> L27A and TDP-43<sub>1-102</sub> L28A variants provides insight into dynamics on the ps-ns timescale (**Figure 3A**). The low NOE ratios show that there are significant dynamics in both the N-terminal His-tag and the C-terminal NLS; this is consistent with the lack of stable conformations in these segments. In contrast, the folded portion (residues-3-77) shows high NOE ratios which approach the values (0.85) expected for static behavior in the elements of secondary structure. Somewhat lower ratio values are observed in the loops, especially that connecting  $\beta$  strands 3 and 4. Similar ps-ns dynamic behavior was seen for the short (1-77) or long (1-102) WT constructs as well as the variant L27A, which was designed to disrupt dimer formation (*vide infra*). By contrast, the L28A variant's  $\{^1\text{H}\}$ - $^{15}\text{N}$  NOE ratio values are considerably lower than WT NTD or the L27A variant, indicating higher dynamics on the fast ps-ns timescale.

WT TDP-43<sub>1-102</sub> dimerizes *in vitro*. As a first step, we studied the oligomerization behavior of WT TDP-43<sub>1-102</sub>, by analytical centrifugation, and found evidence for dimer formation at higher concentrations (**Figure 2D**); namely, a peak with a sedimentation constant of 1.9 S corresponding to dimer is observed at higher concentrations along with a larger peak at 0.8 - 1.1 S which corresponds to monomer. Upon dilution, the peak corresponding to dimer disappears. Similar results were obtained on an independent sample prepared in 100% D<sub>2</sub>O (*vide infra*). Based on these data, an estimated  $K_D$  of 1 mM for the dimer dissociation could be calculated under these experimental conditions: pH 4.0, 25°C. Additional thermal denaturation experiments on WT TDP-43<sub>1-102</sub> monitored by fluorescence spectroscopy showed a mean increase in  $T_M$  of 2.2 °C at 500  $\mu\text{M}$  relative to 77  $\mu\text{M}$  (*data not shown*). This increased  $T_M$  is consistent with a dimer dissociation constant of 0.94 mM.

Previous studies of TDP-43 *in vitro* and in cells at near neutral pH and physiological salt concentrations reported that Cys residues can form disulfide bonds under oxidizing conditions which promote TDP-43 oligomerization (35,36). Cys residues in the RRM domains were found to be mainly responsible, but the participation of the two Cys of the NTD was not ruled out. Here, MALDI-TOF mass spectra revealed a major peak which is consistent with the sequence of  $^{15}\text{N}$ -TDP-43<sub>1-102</sub> and a generally high incorporation (> 90%) of  $^{15}\text{N}$  (**Figure S2**). Only trace peaks whose mass corresponds to a dimer were detected (**Figure S2**). The MALDI-TOF procedure dissociates non-covalently linked dimers, but does not separate dimers linked by disulfide bonds. This point is relevant considering that the NTD contains two rather exposed Cys residues (25) (**Figure S3C**). Since a dimer peak was not detected, these results, as well as  $^{13}\text{C}\beta$  chemical shift values characteristic of reduced Cys, rule out dimerization through the formation of intermolecular disulfide bonds under the conditions studied; namely pH 4.0, 25 °C. MALDI-TOF mass spectra recorded on the L27A and L28A variants also revealed mass peaks expected for these variants, a high level of  $^{15}\text{N}$  incorporation and no evidence for dimers linked by covalent bonds (*data not shown*).

Next, NOESY spectra on a freshly prepared sample dissolved in buffer containing one mM deuterated acetic acid and 100% D<sub>2</sub>O were recorded on concentrated (340  $\mu\text{M}$ ) and diluted (100  $\mu\text{M}$ ) samples. By these spectra, we observed that some signals, for example those of L27, were considerably perturbed (**Figure S3 A**). This strongly suggests that this residue is directly involved in the dimerization interface. The relatively small number of changes is likely due to the fact that these spectra were recorded in D<sub>2</sub>O and most surface HN have exchanged and become invisible. Analytical centrifugation was performed on this sample and revealed a significant population of dimer (*data not shown*). The small number of changes also suggests the interface involves a relatively low number of residues, which is in

agreement with the weak association constant and the relatively low population of dimer.

The NTD surface, shown **Figure S3B**, features several exposed hydrophobic residues. Of them, M1, V57, I60, M63 and V74 are discontinuous and are not conserved by evolution, whereas I16, P19, L27, A33 and P36 are clustered near each other, and are invariable from *H. sapiens* to *Drosophila* (37) and are strictly conserved in *Hydra*, save a conservative substitution of Leu 27 by Met (38) (**Figure S3C**). Exposed and evolutionarily conserved hydrophobic residues commonly form binding surfaces (39,40).

As an additional test for dimer formation, the translational and transversal relaxation rates for the following  $^{15}\text{N}$ -labeled samples: 523  $\mu\text{M}$  WT TDP-43<sub>1-102</sub>, 145  $\mu\text{M}$  WT TDP-43<sub>1-102</sub>, 200  $\mu\text{M}$  L27A TDP-43<sub>1-102</sub>, 132  $\mu\text{M}$  L28A TDP-43<sub>1-102</sub>, at 25 °C in 1 mM acetic acid, pH 3.9, and compared to values previously determined for the TDP-43<sub>1-77</sub> construct at 290  $\mu\text{M}$  (25) (**Figure 3**). These results, show that the transversal relaxation rates,  $R_2$ , of well-folded residues of the concentrated WT TDP-43<sub>1-102</sub> sample are significantly higher than their counterparts in the more dilute TDP-43<sub>1-102</sub> sample, the shorter TDP-43<sub>1-77</sub> construct, and the two variants TDP-43<sub>1-102</sub> L27A and L28A. The translational relaxation rate,  $R_1$ , values of the 523  $\mu\text{M}$  WT TDP-43<sub>1-102</sub> sample are also slightly lower. These results, which can also be represented as the  $R_2/R_1$  ratio, are clear evidence for dimer formation in the 523 microM TDP-43<sub>1-102</sub> sample. From the mean  $R_1$  and  $R_2$  rates, values of the correlation time,  $T_c$ , which measures the tumbling time in solution and is related to the size, were calculated (**Sup. Table 2**) These  $T_c$  values are not directly comparable because of the influence of disordered segments on the experimental  $T_c$  values, but not the calculated  $T_c$  values. Nevertheless, all the samples'  $T_c$  values are in reasonable agreement with the values calculated based on the monomer structure (**Table S2**) and with  $T_c$  values for proteins of this size, except the 523  $\mu\text{M}$  WT TDP-43<sub>1-102</sub> sample. The latter's  $T_c$  value is

consistent with a mixed sample content of monomer and dimer.

*Atomistic model for the solution structure of the NTD dimer:* A sample containing 50% unlabeled TDP-43<sub>1-102</sub> and 50%  $^{13}\text{C}$ ,  $^{15}\text{N}$ -TDP-43<sub>1-102</sub> labeled (total concentration = 330  $\mu\text{M}$ ) was studied to obtain atomic level information on the TDP-43 dimer's structure. In this sample, half the dimers will contain one labeled subunit and one unlabeled subunit which affords the exclusive detection of pure intermolecular proton-proton contacts using a  $^{13}\text{C}$ -edited /  $^{12}\text{C}$ -filtered 2D NOESY experiment, in which only cross-peaks between the unlabeled and labeled monomers within the same dimer are visible. Following this strategy, we could detect 17 intermolecular cross-peaks (**Table S3, Figure S3 D**).

Whereas this limited number of unambiguous NOEs is insufficient to enable the determination of a high-resolution dimer structure, it does reveal many structural features of the interface. Several of these NOEs arise from inter-monomer contacts between L27 side chains, as well as between P19 (at the end of  $\beta$ -strand 2) with E58 (in the loop that connects  $\beta$ -strands 4 and 5). Utilizing a limited number of conformational restraints based just on these NOEs (**Table S3, Figure S3**), the structure of a minimal dimer interface could be calculated (**Figure 4A**). In this structure, the burial of the hydrophobic side chain of L27 (**Figure 4A**), which is exposed in the monomer (*vide supra* & **Figure 1A, S1 & S3**) would provide a favorable free energy change to drive dimerization.

The remaining NOE signals arise from residues 30 - 32 in the  $\alpha$ -helix and P36, which lies right at the end of this element of secondary structure. Employing all possible NOE-derived conformational restraints, a larger dimer interface, which features helix - helix "knobs into grooves" packing could be determined (**Figure 4B**). Whereas this structure represents our proposal for the TDP-43 NTD dimer in solution, it is important to point out that on the basis of this structure, one may expect to see 6 additional NOE crosspeaks between  $^1\text{H}$  in the

$\alpha$ -helices; nevertheless, these signals were unobserved. Dynamic behavior on the micro- to milli-second timescale, which might arise from interconversion between the conformers comprising the minimal and large dimer interfaces, or low signal intensity due to the low population of the dimer could account for why these signals are not detected.

*Leu 27 is a key residue for TDP-43 NTD dimerization.* To further validate the structural model for the NTD dimer, a variant containing the Leu 27 to Ala mutation was prepared and studied. The  $^1\text{H}$ - $^{15}\text{N}$  HSQC of the TDP-43<sub>1-102</sub> L27A variant shows features of the natively folded domain (**Figure 2A**). A 3D  $^{15}\text{N}$ -NOESY-HSQC spectrum was also recorded and its analysis permitted the verification of assignments in the 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum, the confirmation of essential side chain assignments and the corroboration that crucial NOEs defining the tertiary structure were still present in this variant with respect to the WT (**Figure S4**). A plot of the differences in the  $^1\text{H}$ - $^{15}\text{N}$  backbone chemical shifts in the L27A mutant versus the WT construct reveals minor variations except for the residues neighboring the mutation site and residues 21, 22 and 58 (**Figure 2B**). The latter differences may reflect the lack of a populated dimer. In addition, we monitored the thermal denaturation of this variant using fluorescence spectroscopy, and found that the WT TDP-43<sub>1-102</sub> L27A variant ( $T_M = 45^\circ\text{C}$ ) is almost as stable as the WT TDP-43<sub>1-102</sub> WT ( $T_M = 49^\circ\text{C}$ ) at pH 3.9, as illustrated in **Figure 2C**. Analytical centrifugation revealed that the TDP-43<sub>1-102</sub> L27A variant is completely monomeric at concentrations where WT TDP-43<sub>1-102</sub> exists partly as a dimer (**Figure 2D**).

We also analyzed the effect of substituting both L27 and L28 by alanine residues, in the context of the L27A-L28A double mutant. This variant appears to be even less stable than L28A as its  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum is typical of a completely unfolded protein (**Figure S5**). The destabilizing effect caused by the mutations increases from L27A to L28A to L27A-L28A can be appreciated in 7

**Figure S6** where the upfield region of the  $^1\text{H}$  NMR spectrum near 0 ppm shows intense peaks for methyls retaining the native fold in the hydrophobic core of L27A, weaker and shifted signals in L28A, and absent of these peaks in the double mutant.

*Conformational model of the TDP-43<sub>1-414</sub> dimer.* A model of the complete TDP-43 protein, built on the basis of the NTD monomer structure and dimer interface reported here, and the SAXS envelope of (31), shows that the NTD and RRM domains are arranged like beads of a string or rosary and that the disordered C-terminal region extends out away from them (**Figure 4C**). However, considering that the segment connecting the NTD to RRM1 is flexible and that the RRM1 to RRM2 linker is also flexible in the absence of RNA (41), it is quite possible that the folded domains and C-terminal regions can flex and bend to interact with each other, when they are buffeted by other macromolecules or due to attractive interactions with other molecules such as RNA. This structure, and the contacts defining the dimer interface in the N-terminal domain, may well be affected by changes in pH and ionic strength. This is in agreement with this consideration, evidence for interdomain contacts in TDP-43 has been recently reported (28).

*Functional experiments on NTD structural mutants in the L28-F35  $\alpha$ -helix.* To study the biological implications of NTD variants in a cellular context, we engineered a mutation 31V/R-32T/R that was predicted to completely destroy the  $\alpha$ -helix structure and unfold the NTD, because of the well known severely destabilizing effect of introducing charges into the hydrophobic core (26). This mutant was then assayed in a variety of functional assays to test its ability to affect the splicing and aggregation functionality of TDP-43.

First of all, we tested this mutation in an add-back assay (**Figure 5**). This assay which is described in detail by (42) is based on a

minigene system carrying CFTR exon 9, that includes a mutation in a splicing enhancer element within its sequence (C155T) to obtain a 50/50 ratio of exon inclusion/skipping in normal conditions (**Figure 5, lane 1**). This is the optimal condition to see whether a change in TDP-43 structure can result in either a loss-of-function (less exon 9 skipping) or gain-of-function effect (more exon 9 skipping). In this system, when endogenous TDP-43 is removed from the cells by siRNA treatment, the levels of CFTR exon recognition substantially increase to more than 80% (**Figure 5, lane 2**). As expected, splicing inhibition was fully rescued following expression of a si-resistant TDP-43 wild-type protein (**Figure 5, lane 3**) but not when a siRNA-resistant F4L TDP-43 mutant, that cannot bind RNA due to mutations in the RRM1 and RRM2 domains was expressed at similar levels (**Figure 5, lane 4**). The same result was also observed for the 31V/R-32T/R mutant (**Figure 5, lane 5**).

In order to determine whether this change could also be associated with loss-of-function, we took advantage of an aggregation system that we previously set up by adding twelve repetitions of the Q/N prion-like domain of TDP-43 (15). The induction of this protein, when stably transfected in cells, caused the accumulation of aggregates capable of sequestering the endogenous TDP-43 protein and inducing a very well defined loss-of-function phenotype (22,34,43). Therefore, we also inserted the 31V/R-32T/R mutation in these stable cell lines (Flag-TDP-12X-31V/R-32T/R) to see their effect on aggregation and endogenous TDP-43 sequestration (**Figure S7A, S8A**). As shown in **Figure S7B, S8B**, compared with the expression of a wild-type TDP-12x-Q/N, the expression of this protein was not capable of inducing loss-of-function effects in the pre-mRNA splicing of POLDIP3. The reason for this loss is due to the fact that, as shown in co-immunoprecipitation experiments, the Flag-TDP-12X-31V/R-32T/R protein has a very reduced ability to interact with endogenous TDP-43 (**Figure S7C, S8B**). Additional experiments also confirmed that endogenous TDP-43 remains soluble in the

nucleus even in the presence of aggregated Flag-TDP-12X-31V/R-32T/R (*data not shown*).

Having established the importance of folded NTD and its  $\alpha$ -helix in TDP-43 splicing and aggregation properties, it was therefore of interest to see the effects of the L27A and L28A mutants for which we had obtained the structural data described above.

*Functional experiments on NTD structural mutants designed to disrupt quaternary (L27A) and tertiary (L28A) contacts.* First of all, we performed add-back experiments on all the three mutants (L27A-L28A, L27A, and L28A) in HeLa cells (**Figure 6A**). These results show that the double-mutant L27A-L28A cannot recover CFTR exon 9 skipping following its add-back in TDP43-depleted cells (**Figure 6A, lane 5**) just like the 31V/R-32T/R. Furthermore, it almost all localizes in the cytoplasm as observed with immunofluorescence experiments (**Figure 6C**). This is expected, as we have shown that this mutant lacks any native structure in solution, as gauged from the NMR studies described above. More interesting is the L28A variant, which is also unable to recover the exon skipping activity (**Figure 6A, lane 7**) and is also predominantly localized in the cytoplasm. This similar behavior of L28A and the L27-L28A double mutant is evidence that TDP-43 needs to be well, and not partially folded, to perform its biological functions and maintain its predominantly nuclear localization.

In fact, the L27A variant that is shown above to retain the native fold can restore exon skipping just like WT does (**Figure 6A, lane 6**). Interestingly, although this L27A is well localized in the nucleus, it is also quite present in the cytoplasm (**Figure 6C**). We attribute this behavior to this variant's ability to retain the native fold, while being less able than the WT to dimerize.

These results are also reflected when these mutants are transfected in cell lines stably expressing a GFP-TDP-43-12xQ/N-F4L protein (**Figure 7**). First of all, it can be observed that the add-back of L27A can substantially recover the CFTR exon 9 skipping following the



formation of the aggregates (**Figure 7A, lane 6**) whilst this was not observed for the L28A and L27A-L28A mutants (**Figure 7A, lanes 7 and 9**). Moreover, the immunohistochemical localization of L27A in the presence of the aggregates still remains diffuse in the nucleus (**Figure 7B**) as opposed to the L28A and L27A-L28A which colocalize extensively with the aggregates.

## DISCUSSION

Thanks to the superior solubility of the TDP-43<sub>1-102</sub> construct, we have been able to extend the structure of the NTD to higher resolution, corroborate the identity of well protected amide groups in secondary structure, and identify three new protected residues. These results are additional evidence that the NTD is well folded and stable (25) and not poorly folded and marginally stable (24,28).

Our findings are pertinent to the open debate on whether the NTD is stably folded and functionally relevant. In particular, the strong destabilization and structural disruption of the NTD in the context of the L28, L27A/L28A, and V31R/T32R mutations, these variants' dramatically reduced activity in modulating RNA splicing, subcellular mislocalization and strong tendency to aggregate in cells all emphasize the importance of folded NTD for TDP-43's native activity. Based on these results, we conclude that the NTD must be stably folded for TDP-43 to be able to carry out its physiological functions. This conclusion is likely to hold at the level of whole organisms, since the deletion of a short segment of the NTD corresponding to the first  $\beta$ -strand, which we expect would provoke the domain's denaturation, has severe physiological effects in mice (23).

TDP-43 is known to exist as a dimer/monomer equilibrium *in vivo* (30). Under the conditions studied here (25 °C, pH 3.9) the TDP-43 NTD dimerizes through interactions mediated by conserved exposed hydrophobic residues such as L27, P19 and P36 as well as E58. Residues in the alpha helix could contribute additional interactions. The dimerization interface is relatively small and

the association is weak under the conditions examined here. However it could be relevant for TDP-43's physiological activities, particularly in "crowded" conditions, for example when concentrated in sub-cellular compartments such as stress granules or as part of microdroplets (as discussed below).

Considering that the  $K_D$  of the TDP-43 NTD dimer is estimated to be in the range of 1 mM, which is well above the *in vivo* concentration of TDP-43, there is a doubt regarding how much dimer is present in cells. TDP-43 levels in cells are subjected to very fine regulation as the protein binds and regulates the concentration of its own mRNA (32) and also acts to retain its own mRNA in the nucleus to prevent translation (33). Nevertheless, crowding effects from the high *in vivo* concentrations of macromolecules and metabolites act to promote protein oligomerization (44) including TDP-43 dimerization. In addition, *in vivo* post translational modifications, such as the acetylation of K145(45) or phosphorylation (46), are known to increase or decrease, respectively, TDP-43 self association. It is also possible that the monomer association is strengthened at neutral pH and higher ionic strength conditions (20). In addition, gel filtration and SDS-PAGE experiments have also shown that TDP-43 truncation constructs consisting of both RRM domains either linked together or as separate entities also dimerize or tetramerize, respectively, *in vitro* (47). Therefore, TDP-43 dimerization is likely to be aided by interactions among the RRM domains and could be further strengthened when distinct TDP-43 molecules are bound to the same RNA.

It is fascinating that the L27A variant, which is unable to dimerize but can stably fold, is active in modulating splicing. This is logical considering that no TDP-43 dimers were previously detected in the nucleus where splicing occurs (30). In addition to regulating splicing, TDP-43 performs several other functions and dimerization may well be required for some of them. TDP-43 dimers were previously observed in the cytoplasm (30), despite the fact that TDP-43 is more

concentrated in the nucleus. This suggests that some mechanism concentrates TDP-43 in the cytoplasm so as to promote its dimerization. Over the last several years, the idea that cells contain several types of discrete liquid phases known as “microdroplets” such as nucleoli (48), stress granules (49) or neuronal granules which serve to concentrate and organize certain proteins and RNAs has gained acceptance (50-52). TDP-43 has been recently shown to bind to stress granules through the hydrophobic subsegment in the CTR(13,14), and as a working hypothesis, we speculatively propose that the concentration of TDP-43 molecules within these microdroplets could well promote its dimerization and eventually lead to pathological aggregation or altered functionality. Future experiments will be necessary to test this proposal as well as ascertain the physiological and eventually pathological function(s) of the TDP-43 dimer. Nonetheless, our results provide a better understanding of the relationship between TDP-43 N-terminus and its functionality. Most importantly, the level of structural detail achieved by our analyses to characterize the dimerization interface at single amino-acid level may provide may eventually open the way for the development of small molecules/peptides capable of interfering with this process. These effectors could then eventually be used to specifically inhibit TDP-43 aggregation tendencies and enhance its solubility within cells.

## EXPERIMENTAL PROCEDURES

**Protein expression and purification.** Protein expression was essentially performed as previously described in Mompean *et al* (25). Briefly, the sequences coding for TDP-43 N-terminus residues 1-102 and the various mutants were cloned in the BamHI-HindIII of pQE306xHis plasmid. The plasmids were then used to transform M15 bacteria grown on Kana/Amp resistant plates. The bacteria were then grown in M9 media supplemented with D-Glucose (U-13C, 99%, Cambridge Isotope Lab, Inc) at 0.3% v/v and Ammonium-15-Chloride (Sigma) at 0.1% v/v and protein expression was 10

induced overnight with 1 mM IPTG at 30°C. All constructs contained an N-terminal His-tag (MRGSHHHHHHGS) which facilitated the proteins’ purification following the manufacturer’s instructions using Ni-NTA Agarose Qiagen Resin in the presence of Complete PI (Roche). Eluted fractions were then analyzed by 15% SDS-PAGE before staining with Colloidal Coomassie Blue to check for purity. For spectroscopic studies, samples were transferred to buffer containing 85% milliQ water, 15% D<sub>2</sub>O (Cambridge Isotope Labs. Inc) and 1.0 mM acetic acid-d<sub>3</sub> (Aldrich) using PD-10 desalting columns (GE Healthcare) and concentrated utilizing Vivaspinn centrifugal concentrators (Santorius). We used deuterated acetic acid to avoid the <sup>1</sup>H signal from the methyl group of acetic acid.

**Fluorescence spectroscopy.** The thermal denaturation of WT TDP-43<sub>1-102</sub>, and that of the L27A, L28A, L27A–L28A variants, was monitored using a Fluoromax4 spectrofluorimeter (Jobin-Yvon Inc., Edison, NJ, USA). The Xe emission (467 nm) line and the water Raman signal (which is 397 nm when exciting at 350 nm) were used to calibrate the excitation and emission wavelengths, respectively. The experiments were made using fractions from gel filtration or NMR samples (concentrations 50 – 100 μM, save one experiment at 500 μM) and contained the same buffer (1 mM acetic acid, pH 3.9) as the NMR experiments. Emission spectra were recorded over the emission range of 270 - 380 nm, utilizing an excitation wavelength of 280 nm, 2 - 2.5 nm excitation and 2 nm emission slits and a Peltier module to control the temperature, and collecting data at a scan speed of 2 nm·s<sup>-1</sup> over intervals of 1 - 2 °C with a 1 or 2 minute equilibration times, respectively, at the new temperature before recording the spectra.

**Analytical ultracentrifugation.** The samples used for the NMR experiments were subjected to sedimentation velocity experiments at 25 °C, using an Optima XL-A analytical

ultracentrifuge with a UV-VIS variable wavelength detection system. The equilibrium assays were performed on concentrated and samples diluted three to five fold, by centrifugation at 48 krpm and detected using a wavelength of 280 nm. The concentrations of TDP-43<sub>1-102</sub> studied were: WT 0.23 mM and 0.046 mM; L27A 0.20 mM and 0.050 mM; L28A 0.13 mM and 0.032 mM. A second independent WT sample was studied and gave similar results. For the experiments shown in **Figure 2E**, thin cells (pathlength = 3 mm) were used for the concentrated L27A and L28A samples and normal cells (pathlength = 12 mm) were utilized for the dilute samples. This accounts for why the peak absorbance does not decrease despite the substantial dilution. The program SEDFIT (version 15.01b) (53) was used to analyze the sedimentation profiles and obtain the sedimentation coefficients.

*NMR experiments.* All NMR experiments were recorded at 25 °C and pH 3.9, using a Bruker Avance II 800 MHz spectrometer equipped with triple axis resonance cryo-probe with Z-gradients. Based on our reported chemical shift assignments for WT TDP-43<sub>1-77</sub> (25); BMRB 25675, we performed additional standard 3D experiments, including <sup>1</sup>H-<sup>1</sup>H-<sup>15</sup>N HSQC-NOESY, CACB(CO)N(H), HNCO, and HNCA spectra, to ensure that no structural changes occurred, to assign the segment 78–102 and confirm that it is unfolded and does not interact with the folded domain comprising residues 1–77. This assignment process was carried out manually using SPARKY (54), and the <sup>1</sup>H chemical shifts were referenced to sodium 4,4-dimethyl-4-silapentane-1-sulfonate. The <sup>13</sup>C and <sup>15</sup>N chemical shifts were referenced indirectly to <sup>1</sup>H based on these nuclei's gyromagnetic ratios as recommended (55). Interproton nuclear Overhauser effect (NOE) distance restraints were obtained by 2D <sup>1</sup>H-<sup>1</sup>H and <sup>13</sup>C-edited/<sup>12</sup>C-filtered NOESY spectra (120 ms mixing time). The NMR assignments have been deposited in the Biological Magnetic Resonance Data Bank (BMRB access code: 34081).

#### *NMR <sup>15</sup>N relaxation measurements.*

The <sup>15</sup>N relaxation data for WT TDP-43<sub>1-102</sub>, (523 μM and 270 μM), L27A (200 μM) and L28A (130 μM) variants were recorded at 298 K on an 800 MHz spectrometer, with 1024 and 64 complex points along the direct and indirect dimensions and 32 scans. Steady-state <sup>15</sup>N-<sup>1</sup>H heteronuclear NOE measurement (56) were carried out with an overall recycling delay of 10s to ensure the maximal development of NOEs before acquisition and to allow solvent relaxation (57). Heteronuclear NOEs were calculated from the ratio of cross-peak intensities in spectra collected with and without amide proton saturation during the recycle delay. Uncertainties in peak heights were determined from the standard deviation (σ) of the distribution of intensities in the region of the spectra where no signal and only noise was observed.

The characterization of the dynamics and dimer formation by measuring the translational (R<sub>1</sub>), off-resonance rotating frame (R<sub>1ρ</sub>) and transversal (R<sub>2</sub>) relaxation rates using the parameters and procedures previously described for the TDP-43<sub>1-77</sub> construct (25). For the 523 μM WT TDP-43<sub>1-102</sub> sample, four additional spectra were recorded with longer relaxation delays (2.0, 2.2, 2.5 & 3.0 s) in order to accurately fit the R<sub>1</sub> rates, which are somewhat shorter than those of the other samples. The overall correlation time, T<sub>c</sub>, was calculated from the ratio of the mean values of R<sub>1</sub> and R<sub>2</sub> excluding disordered residues. Values of NOE ratios < 0.65, R<sub>2</sub> < the mean R<sub>2</sub> – 1σ and R<sub>1</sub> > the mean R<sub>1</sub> + 1σ were used to define disordered residues (58). In addition to the experimental measurements, the correlation times were also calculated for the well-folded regions of the monomer and the dimer structures using the HydroNMR program (59).

*NMR structure calculations.* The dimeric structure of TDP-43<sub>1-102</sub> was determined in a two-step procedure using CYANA-3.97 (60). First, we resolved the monomeric structure using dihedral angle restraints as determined by TALOS+ based on the chemical shift information, in addition to a

large number NOE cross-peaks obtained from a  $^1\text{H}$ - $^1\text{H}$  2D NOESY spectrum. This data was used as the input for interactive structure calculation; to elaborate, seven cycles of simultaneous automatic assignment of the cross-peaks and torsion angle dynamics. Next, a final simulated annealing calculation was carried out using the distance restraints derived from the consensus assignment achieved in the seventh cycle. The resulting structure is slightly more refined than that of monomeric TDP-43<sub>1-77</sub> that we recently reported, (25), essentially because the samples used in this work yielded superior spectra, as explained in the main text. The final refined structures have been deposited in the RCSB Protein Data Bank (access code: 5MRG), respectively.

In a second step, the  $^{13}\text{C}$ ,  $^{15}\text{N}$ -edited /  $^{12}\text{C}$ ,  $^{14}\text{N}$ -filtered 2D NOESY experiment, which is specifically designed to detect intermolecular contacts, was recorded on a hemi- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled 330  $\mu\text{M}$  TDP-43<sub>1-102</sub> sample. The crosspeaks obtained from this experiment were converted to upper distance limits of 5.5 Å. These data were used as distance restraints to complement the previously obtained dihedral and distance restraints of the monomer. As we noted in our previous study, there is a unique H-bonding pattern that can account for residues showing protection against H/D exchange and used as constraints along with the experimental cross-peaks resulting in a bundle of conformers without violations. Thus, we used all this information together and submitted two copies of TDP-43<sub>1-102</sub> to a second simulated annealing calculation to obtain the final dimeric structure. Finally, the structural model of the NTD dimer and the tandem RRM domain structure solved in 2013 (41) were manually fitted into the previously published SAXS envelope to produce the model for the complete TDP-43 dimer shown in **Figure 4C**.

*Add-back splicing assay and in cell aggregation and subcellular localization with WT & the mutants.* Using specific sets of primers the 31V/R-32T/R, L27A, L28A, and L27A-L28A were inserted in the previously described si-resistant plasmid expressing wild-type TDP-43 (42). Briefly, in order to 12

maximize TDP-43 silencing efficiency, HeLa cells were plated at 30% of confluence (day 0) and two rounds of TDP-43 siRNA transfections were carried out on day 1 and 2, according to the procedure already described (61). On the afternoon of day 2, co-transfection was performed with 1  $\mu\text{g}$  of pFLAG-fusion protein expression vector and 0.5  $\mu\text{g}$  of the CFTR C155T reporter minigene. On day 3, cells were harvested, and total RNA was extracted with EuroGold TRifast (Euroclone, Milan, Italy). Reverse transcription was performed using M-MLV Reverse Transcriptase (Invitrogen, Carlsbad, California, USA), according to the manufacturer's protocol. PCR with DNA Polymerase (New England Biolabs, Ipswich, Massachusetts, USA) using minigene-specific primers was carried out for 35 amplification cycles (95°C for 45s, 54°C for 45s, 72°C for 45s). The expression levels of the added-back TDP-43 proteins were monitored through western blotting, using a commercially available antibody against TDP-43 (Protein Tech, 10782-2-AP). Endogenous tubulin (in-house made mouse monoclonal antibody) and p84 (abcam, Ab487) were used as loading controls. Inclusion levels of exons were quantified using the Qiaexcel platform used to run the PCR products or through ImageJ quantification.

*Stable cell line generation.* HEK293 flip-in cell line (Invitrogen) was grown in DMEM-Glutamax-I (GIBCO) supplemented with 10% fetal bovine serum (GIBCO) and Antibiotic-Antimycotic stabilized suspension (Sigma). The plasmid transfections were carried out using Effectene Transfection reagent (Qiagen) according to manufacturer's instructions. To generate the stable clone, 0.5 g of the Flag-TDP-12X-31V/R-32T/R expressing plasmid were co-transfected together with 0.5 g pOG44 vector that expresses the Flp-recombinase (Invitrogen).

Cells were grown in DMEM-Glutamax-I supplemented with 10% fetal bovine serum, Antibiotic-Antimycotic stabilized suspension (Sigma) and gradually selected using 100g/ml Hygromycin B and 10g/ml of Blastidicin (both purchased from



Invivogen). The induction of Flag-TDP-43-12X-Q/N, Flag-TDP-12X-31V/R-32T/R, and GFP-TDP-43-12XQ/N-F4L proteins was achieved by adding 1 g/ml of tetracycline (Sigma) at the culture medium.

Protein expression was analyzed through western blot experiments using anti-Flag antibody (Sigma, F1804), anti-TDP-43 (Protein Tech, 10782-2-AP) and anti GAPDH (Santa Cruz Biotechnology, sc-25778) as loading control.

*Co-immunoprecipitation assays.* For co-immunoprecipitation assays, HEK293 flip-in stable cell lines expressing Flag-TDP-43-12X-Q/N and Flag-TDP-12X-31V/R-32T/R were induced for 24 hours with 1 µg/ml of tetracycline. Cells were collected in RIPA lysis buffer (50 mM Tris/HCl pH 7.4, 150 mM NaCl, 1% NP-40, 0.1% SDS, 1 mM EDTA pH 8, 1mM PMSF, 0.5% SDC, H<sub>2</sub>O up to the final volume) supplemented with protease inhibitors (Roche, cat. 11836145001) and incubated for 30 minutes at 4°C. After spin down at 500 xg at 4°C, cells were lysed by sonication. The lysates were then incubated with 40 µl of A/G plus Agarose beads (Santa Cruz) to perform a pre-clearing for 1 hour at 4°C. In the meantime, an incubation of 40 µl of A/G plus Agarose beads with 3 µg of anti-Flag antibody (Sigma, F1804) in RIPA buffer, for 2 hours at 4° C was performed. After both incubations, the pre-cleared lysate was incubated with A/G plus Agarose beads/anti-Flag O/N at 4° C. The day after the beads were precipitated and washed with PBS once for 10 minutes at 4°C. The beads were finally resuspended in 50 µl of Resuspension Buffer (50 mM Tris/HCl pH 7.4, 5 mM EDTA, 10 mM DTT, 1% SDS, H<sub>2</sub>O up to the final volume) and 20 µl of SDS 5X loading buffer were added.

Immunoprecipitates were then analyzed by Western blot using anti-flag (Sigma, F1804),

anti-TDP-43 (Protein Tech, 10782-2-AP) or anti-Tubulin (Calbiochem, CP-06) antibodies. A 10% of input from total lysate was loaded on SDS-PAGE and analyzed as control.

*Immunofluorescence microscopy.* For indirect immunofluorescence HEK-Flag-TDP-12X-Q/N, HEK-Flag-TDP-12X-31V/R-32T/R and HEK-GFP-TDP-12X-F4L stable cell lines were induced and processed as previously described (61). The primary antibodies used were anti-Flag (Sigma, F1804) and anti-TDP-43 (Protein Tech, 10782-2-AP). The secondary antibodies anti-mouse-AlexaFluor 594 (cat. A21203), anti-rabbit-AlexaFluor 488 (cat. A21200) and TO-PRO3 (cat. T3605) were all purchased from Life Technologies. Cells were analyzed on a Zeiss LSM 510 Meta confocal microscope.

*Cell lysate fractionation.* To perform cell lysate fractionation in soluble and pellet fractions, 2x10<sup>6</sup> cells (HEK-Flag-TDP-12X-Q/N or HEK-Flag-TDP-12X-31V/R-32T/R) were seeded and induced with 1 µg/ml of tetracycline for 24 hours. Then, cells were collected and lysed with 1ml of RIPA lysis buffer + Protease inhibitor for 30 minutes at 4°C. After a centrifugation at 4000 rpm for 20 minutes, the whole supernatant was furthermore sonicated for 5 minutes to allow a better lysis. Two hundred µg of cell lysate were ultracentrifuged in a clean Beckman polycarbonate thick wall Centrifuge tube (rotor type 70.1Ti) for 1 hour at 25°C at 33000 rpm.

The supernatant was collected and the pellet washed twice with 100 µl of RIPA buffer. Pellet was finally dissolved in Urea buffer (7M urea, 4% CHAPS, 30mM Tris, pH 8.5). To analyze each fraction by Western blot, 10% of input, 10% of soluble fraction (S), and 30% of pellet volume (P), were loaded in a 10% SDS-PAGE.

**Acknowledgements:** We thank Dr. Hanna Yuan, from the Institute of Molecular Biology (Sinica Academy, Taiwan), for sharing the SAXS envelope of the TDP-43 dimer. We are grateful to Prof. Marta Bruix for critically reading the manuscript, and to Dr. Vivian de los Rios and Dr. Juan Roman Luque Ortega (Centro de Investigaciones Biológicas, CSIC) for expert mass spectrometry and analytical ultracentrifugation, respectively. We thank Dr.

Marco Baralle and Dr. Sanja Skaro for preparing the V31R,T32R sample. This work was supported by Grants SAF2013-49179-C2-2-R, SAF2016-76678-C2-2-R and EU JPND AC14/00037 (DVL) and EU JPND RiMOdFTD, Italy, Ministero della Sanita' (EB) and the Thierry Latran Foundation REHNPALS (EB).

**Conflicts of Interest:** The authors declare that they have no conflicts of interest with the contents of this article.

**Author Contributions:** VR, CS, FEB and EB produced TDP-43 N-terminal domain protein and site-directed mutant constructs. MMG, DPU & DVL designed, performed and analyzed the experiments shown in Figures 1, 2, 3 & 4. VR, CS, FEB and EB designed, performed and analyzed the experiments shown in Figures 5, 6 & 7. MMG, FEB, EB & DVL wrote the MS with help from VR, DPU and CS. All authors reviewed the results and approved the final version of the manuscript.

## REFERENCES

1. Buratti, E., and Baralle, F. E. (2012) TDP-43: Gumming up neurons through protein-protein and protein-RNA interactions. *Trends Biochem Sci.* **37**, 237-247
2. Kawahara, Y., and Mieda-Sato, A. (2012) TDP-43 promotes microRNA biogenesis as a component of the Drosha and Dicer complexes. *Proc. Natl. Acad. Sci U.S.A.* **109**, 3347-3352
3. Liu-Yesucevitz, L., Lin, A. Y., Ebata, A., Boon, J. Y., Reid, W., Xu, Y. F., Kobrin, K., Murphy, G. J., Petrucelli, L., and Wolozin, B. (2014) ALS-linked mutations england TDP-43 enriched neuronal RNA granules in the dendritic arbor. *J. Neurosci.* **34**, 4167-4174
4. Colombrita, C., Zennaro, E., Fallini, C., Weber, M., Sommacal, A., Buratti, E., Silani, V., and Ratti, A. (2009) TDP-43 is recruited to stress granules in conditions of oxidative insult. *J. Neurochem.* **111**, 1051-1061
5. Neumann, M., Sampathu, D. M., and Kwong, L. K. (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130-133
6. Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y., and Oda, T. (2006) TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem Biophys Res Com* **351**, 602-611
7. Ling, S.-C., Polymenidou, M., and Cleveland, D. W. (2013) Converging mechanisms in ALS and FTL: disrupted RNA and protein homeostasis. *Neuron* **79**, 416-438
8. Lin, W. L., and Dickson, D. W. (2008) Ultrastructural localization of tDP-43 in filamentous neuronal inclusions in various neurodegenerative diseases. *Acta Neuropathol.* **116**, 205-213.
9. Fuentealba, R. A., Udan, M., Bell, S., Wegorzewska, I., Shao, J., Diamond, M. I., Weihl, C. C., and Baloh, R. H. (2010) Interaction with polyglutamine aggregates reveals a Q/N-rich domain in TDP-43. *J. Biol. Chem.* **285**, 26304-26314
10. Jiang, L. L., Che, M. X., Zhao, J., Zhou, C. J., Xie, M. Y., Li, H. Y., He, J. H., and Hu, H. Y. (2013) Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. *J Biol Chem.* **288**, 19614-19624
11. Jiang, L. L., Zhao, J., Yin, X. F., He, W. T., Yang, H., Che, M. X., and Hu, H. Y. (2016) Two mutations G335D and Q343R within the amyloidogenic core region of TDP-43 influence its aggregation and inclusion formation. *Sci. Rep.* **6**, 23928
12. Lim, L., Wei, Y., Lu, Y., and Song, J. (2016) ALS-causing mutations significantly perturb the self-assembly and interactions with nucleic acid of the intrinsically disordered prion-like domain of TDP-43. *PLoS Biol.* **14**, e1002338
13. Conicella, A. E., Zerze, G. H., Mittal, J., and Fawzi, N. L. (2016) ALS mutations disrupt phase separation mediated by  $\alpha$ -helical structure in the TDP-43 low complexity C-terminal domain. *Structure* **24**, 1-13
14. Schmidt, H. B., and Rohatgi, R. (2016) *In vivo* formation of vacuolated multiphase compartments lacking membranes. *Cell Reports* **16**, 1228-1236
15. Budini, M., Buratti, E., Stuni, C., Guarnaccia, C., Romano, V., De Conti, L., and Baralle, F. E. (2012) Cellular model of TAR DNA-binding protein 43 (TDP-43) aggregation based on its C-terminal Gln/Asn-rich region. *J. Biol. Chem.* **287**, 7512-7525
16. Mompeán, M., Hervás, R., Xu, Y., Tran, T. H., Guarnaccia, C., Buratti, E., Baralle, F., Tong, L., Carrión-Vázquez, M., McDermott, A., and Laurents, D. V. (2015) Structural Evidence of Amyloid Fibril Formation in the Putative Aggregation Domain of TDP-43. *J. Phys. Chem. Lett.* **6**, 2608-2615
17. Kametani, F., Obi, T., Shishido, T., Akatsu, H., Murayama, S., Saito, Y., Yoshida, M., and Hasegawa, M. (2016) Mass spectrometric analysis of accumulated TDP-43

in amyotrophic lateral sclerosis brains. *Scientific Reports* **6**, 23281

18. Mompeán, M., Nogales, A., Ezquerra, T. A., and Laurents, D. V. (2016) Complex system assembly underlies a two-tiered model of highly delocalized electrons. *J. Phys. Chem. Lett.* **7**, 1859-1864
19. Mompeán, M., Baralle, M., Buratti, E., and Laurents, D. V. (2016) An amyloid-like pathological conformation of TDP-43 is stabilized by hypercooperative hydrogen bonds. *Front. Mol. Neurosci.* **9**, 125
20. Chang, C. K., Wu, T. H., Y., W. C., Chiang, M. H., Toh, E. K., Hsu, Y. C., Lin, K. F., Liao, Y. H., Huang, T. H., and Huang, J. J. (2012) The N-terminus of TDP-43 promotes its oligomerization and enhances its DNA binding. *Biochem Biophys Res Com* **425**, 219-224
21. Zhang, Y. J., Caulfield, T., Xu, Y. F., Gendron, T. F., Hubbard, J., Stetler, C., Sasaguri, H., Whitelaw, E. C., Cai, S., Lee, W. C., and Petrucelli, L. (2013) The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Gene.* **22**, 3112-3122
22. Budini, M., Romano, V., Quadri, Z., Buratti, E., and Baralle, F. E. (2015) TDP-43 loss of cellular function through aggregation requires additional structural determinants beyond its C-terminal Q/N prion-like domain. *Hum. Mol. Gene.* **24**, 9-20
23. Sasaguri, H., J., C., Xu, Y. F., Gendron, T. F., Garrett, A., Lee, C. W., Jansen-West, K., Bauer, P. O., Perkerson, E. A., Tong, J., Stetler, c., and Zhang, Y. J. (2016) The extreme N-terminus of TDP-43 mediates the cytoplasmic aggregation of TDP-43 and associated toxicity in vivo. *Brain Res.* **1647**, 57-64.
24. Qin, H., Lim, L. Z., Wei, Y., and Song, J. (2014) 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. USA* **111**, 18619-18624
25. Mompeán, M., Romano, V., Pantoja-Uceda, D., Stuaní, C., Baralle, F. E., Buratti, E., and Laurents, D. V. (2016) The TDP-43 N-terminal domain structure at high resolution. *FEBS J.* **283**, 1242-1260
26. Karp, D. A., Gittis, A. G., Stahley, M. R., Fitch, C. A., Stites, W. E., and García-Moreno, B. (2007) High dielectric constant inside a protein relects structural reorganization coupled to the ionization of an internal Asp. *Biophys. J.* **92**, 2041-2053
27. Chang, C. K., and Huang, T. H. (2016) Untangling the structure of the TDP-43 N-terminal domain. *FEBS J.* **283**, 1239-1241.
28. Wei, Y., Lim, L., Wang, L., and Song, J. (2016) Inter-domain interactions of TDP-43 as decoded by NMR. *Biochem Biophys Res Com* **473**, 614-619
29. Sun, Y., and Chakrabarty, A. (2017) Phase to Phase with TDP-43. *Biochemistry* **56**, 809-823
30. Shiina, Y., Arima, K., Tabunoki, H., and Satoh, J. (2010) TDP-43 dimerizes in human cells in culture. *Cell Mol. Neurobiol.* **30**, 641-652
31. Wang, Y. T., Kuo, P. H., Chiang, C. H., Liang, J. R., Chen, Y. R., Wang, S., Shen, J. C., and Yuan, H. S. (2013) The truncated C-terminal RNA recognition motif of TDP-43 protein plays a key role in forming proteinaceous aggregates. *J Biol Chem.* **288**, 9049-9057
32. Ayala, Y. M., De Conti, L., Avendaño-Vázquez, S. E., Dhir, A., Romano, M., D'Ambrogio, A., Tollervey, J. R., Ule, J., Baralle, M., and Buratti, E. (2011) TDP-43 regulates its mRNA levels through a negative feedback loop. *EMBO J.* **30**, 277-288
33. Avendaño-Vázquez, S. E., Dhir, A., Bembich, S., Buratti, E., Proudfoot, N., and Baralle, F. E. (2012) Autoregulation of TDP-43 mRNA levels involves interplay between transcription, splicing, and alternative polyA site selection. *Genes and Development* **26**, 1679-1684
34. Prpar Mihevc, S. P., Baralle, M., Buratti, E., and Rogeli, B. (2016) TDP-43 aggregation mirrors TDP-43 knockdown, 16affecting the expression levels of a common set



of proteins. *Sci. Rep.* **6**, 33996

35. Cohen, T. J., Hwang, A. W., Unger, T., Trojanowski, J. Q., and Lee, V. M. Y. (2011) Redox signalling directly regulates TDP-43 via cysteine oxidation and disulfide bond cross-linking. *EMBO J.* **31**, 1-12
36. Chang, C. K., Chiang, M. H., Toh, T. K., Chang, C. F., and Huang, T. H. (2013) Molecular mechanism of oxidation-induced TDP-43 RRM1 aggregation and loss of function. *FEBS Lett.* **587**, 575-582
37. Ayala, Y. M., Pantano, S., D'Ambrogio, A., Buratti, E., Brindisi, A., Marchetti, C., Romano, M., and Baralle, F. E. (2005) Human, Drosophila and C. elegans TDP-43: Nucleic acid binding properties and splicing regulatory function. *J. Mol. Biol.* **348**, 575-588.
38. Chapman, J. A., Kirkness, E. F., Simakov, O., and Hampson, S. E. (2010) The dynamic genome of Hydra. *Nature* **464**, 592-596
39. Ma, B., Elkayam, T., Wolfson, H., and Nussinov, R. (2003) Protein - protein interactions: Structurally conserved residues distinguish between binding sites and exposed proteins surfaces. . *Proc. Natl. Acad. Sci U.S.A.* **100**, 5772-5777
40. Guharoy, M., and Chakrabarti, P. (2010) Conserved residue clusters at protein-protein interfaces and their use in binding site identification. *BMC Bioinformatics* **11**, 286
41. Lukavsky, P. J., Daujotyte, D., Tollervey, J. R., Ule, J., Stuani, C., Buratti, E., Baralle, F. E., Damberger, F. F., and Allain, F. H. T. (2013) Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nature Struct. Mol. Biol.* **20**, 1443-1451.
42. D'Ambrogio, A., Buratti, E., Stuani, C., Guarnaccia, C., Romano, M., Ayala, Y. M., and Baralle, F. E. (2009) Functional mapping of the interaction between TDP-43 and hnRNP A2 in vivo. *Nucleic Acids Res* **37**, 4116-4126
43. Romano, V., Quadri, Z., Baralle, F. E., and Buratti, E. (2015) The structural integrity of TDP-43 N-terminus is required for efficient aggregate entrapment and consequent loss of protein function. . *Prion* **9**, 1-9
44. Rivas, G., Fernandez, J. A., and Minton, A. P. (1999) Direct observation of the self-association of dilute proteins in the presence of inert macromolecules at high concentration via tracer sedimentation equilibrium: theory, experiment, and biological significance. *Biochemistry.* **38**, 9379-9388
45. Cohen, T. J., Hwang, A. W., Restrepo, C. R., Yuan, C. X., Trojanowski, J. Q., and Lee, V. M. Y. (2015) An acetylation switch controls TDP-43 function and aggregation propensity. *Nat. Commun.* **6**
46. Li, H. Y., Yeh, P. A., Chiu, H. C., Tang, C. Y., and Tu, B. P. (2011) Hyperphosphorylation as a defense mechanism to reduce TDP-43 aggregation. *PLoS One* **6**, e23075
47. Kuo, P. H., Doudeva, L. G., Wang, Y. T., Shen, C. K. J., and Yuan, H. S. (2009) Structural insights into TDP-43 in nucleic-acid binding and domain interactions. . *Nucleic Acid Res.* **37**, 1799-1808
48. Feric, M., Vaidya, N., Harmon, T. S., Mitrea, D. M., Zhu, L., Richardson, T. M., Kriwacki, R. W., Pappu, R. V., and Brangwynne, C. P. (2016) Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686-1697
49. Patel, A., Lee, H. O., Jawerth, L., Maharama, S., Jahnel, M., Hein, M. Y., Stoyanov, S., Mahamid, J., Saha, S., Frazmann, T. M., Pozniakovski, A., Poser, I., Hyman, A. A., and Alberti, S. (2016) A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066-1077
50. Brangwynne, C. P., Eckmann, C. R., Courson, D. S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F., and Hyman, A. A. (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **324**, 1729-1732
51. Chong, P. A., and Forman-Kay, J. D. 17(2016) Liquid-liquid phase separation in cellular

signaling systems. *Curr. Op. in Struct. Biol.* **41**, 180-186

52. Uversky, V. N. (2017) Intrinsically disordered proteins in overcrowded milieu: Membranless organelles, phase separation and intrinsic disorder. *Curr. Op. in Struct. Biol.* **44**, 18-30
53. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619
54. Goddard, T. D., and Kneller, D. G. SPARKY. 3 Ed., San Francisco
55. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D., Wright, P. E., and Wüthrich, K. (1988) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *Pure & Appl. Chem.* **70**, 117-142
56. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by  $^{15}\text{N}$  NMR relaxation. *Biochemistry* **33**, 5984–6003
57. Renner, C., Schleicher, M., Moroder, L., and Holak, T. A. (2002) Practical aspects of the 2D  $^{15}\text{N}$ - $^1\text{H}$ -NOE experiment. *J. Biomol.NMR* **23**, 23–33
58. Pawley, N. H., Wang, C., Koide, S., and Nicholson, L. K. (2001) An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of  $^{15}\text{N}$  relaxation parameters. *J. Biomol. NMR* **20**, 149-165
59. Ortega, A., and García de la Torre, J. (2005) Efficient, accurate calculation of rotational diffusion and NMR relaxation of globular proteins from atomic level structures and appropriate hydrodynamic calculations. *J. Am. Chem. Soc.* **127**, 12764-12765
60. Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol.* **278**, 353-378
61. Ayala, Y. M., Zago, P., D'Ambrogio, A., Xu, Y. F., Petrucelli, L., Buratti, E., and Baralle, F. E. (2008) Structural determinants of the cellular localization and shuttling of TDP-43. *J. Cell Sci.* **121**, 3778-3785

## FIGURE LEGENDS

### FIGURE 1. Structure of TDP-43<sub>1-102</sub> confirms that the NTD is a stable, well folded domain.

(A) Solution Structure of TDP-43<sub>1-102</sub> calculated on the basis of 1628 restraints (see Table S1), shown in atomic (upper figure) and ribbon+atomic (lower figure) representations. Residues -11 to 0 (the His tag, sequence MRGSHHHHHHGS) is colored in dark brown, and residues 78-102 (the NLS) in gray. The zoomed view on the right, highlights the minor differences between the new structure (rose gold, PDB:5MRG) and the previous one (green, PDB:2N4P); namely the C-terminal residues N76 and Y77 now establish contacts with C39 and G40 (red labels), which allows elongation of the  $\beta$ -strands (red circles), and a half turn of 3/10 helix spanning residues P46-S48 (purple labels and circle). This last element of structure is partially populated in the previous structure of the TDP-43<sub>1-77</sub> construct (PDB 2N4P) and now appears in the lowest energy structure in the new calculation obtained with the longer construct: TDP-43<sub>1-102</sub>. Gray arrows indicate the close proximity in the tertiary fold of residues P19, L27, and E58 (black labels), involved in the dimerization interface (see text).

(B) 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum of TDP-43<sub>1-102</sub> recorded at pH\* 3.9, 25 °C in 1.0 mM deuterated acetic acid after transfer into 100% D<sub>2</sub>O. pH\* is the pH reading of the pH meter without correction for the deuterium isotope effect. The blue, green and red spectra were recorded after one hour, 1.5 days and 10 days, respectively of exchange. For clarity, the green and red spectra are displaced 0.07 and 0.14 ppm, respectively, to the right along the x-axis. Ser 20, Gln 34 and Tyr 73,

whose slow exchanging HN groups are identified here, are circled.

**FIGURE 2. L28A destabilizes and breaks tertiary contacts whereas L27A disrupts the dimerization interface.**

(A) 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of TDP-43<sub>1-102</sub> recorded at pH\* 3.9, 25 °C in 1.0 mM deuterated acetic acid in 15 % D<sub>2</sub>O: WT = blue, L27A = green; L28A = red. Selected residues are labeled. Black labels refer all three variants. Brown labels refer to very similar WT + L27A peaks.

(B & C) Weighted average  $^1\text{H}$  &  $^{15}\text{N}$  chemical shift differences ( $\Delta\delta$ ) (calculated as the  $\sqrt{(\Delta\delta(^{15}\text{N}/5))^2 + (\Delta\delta(^1\text{H}))^2}$ ) for WT - L28A (panel B) and WT - L27A (panel C). Asterisks mark residues, for example T25, which moved or broadened so much in the variant so as to prevent assignment. The small “o” indicate residues 45, 47, 50, 51, 53 and 54 in the L27A variant where the chemical shift differences with WT are zero. Other positions lacking bars are either Pro (at sequence numbers 15, 19, 36, 46, 64, 78) or do not have data due to ambiguous assignment in the variant. Blue arrows (beta-strands), parallelogram (alpha-helix) and L (for loops) in panel B indicate the approximate position of secondary structural elements.

(D) Thermal denaturation of WT TDP-43<sub>1-77</sub> = cyan, and TDP-43<sub>1-102</sub> WT = blue and its mutants L27A = green and L28A = red, followed by fluorescence spectroscopy as the ratio of the emission intensity at 330 nm (which is maximal for the folded protein) to 355 nm (where the emission of denatured protein peaks). The curves show the fit of a two state denaturation model to the data points, from which the apparent temperature midpoints ( $T_M$ ) values were calculated. These experiments were performed in 1 mM acetic acid, pH 3.9, at low protein concentrations (50 – 100  $\mu\text{M}$ ).

(E) Sedimentation velocity of WT = blue, L27A = green, L28A = red. The darker solid lines correspond to the more concentrated samples and the lighter dashes are the diluted samples. The concentrations studied were WT concentrated 230  $\mu\text{M}$ , dilute 46  $\mu\text{M}$ , L27A concentrated 200  $\mu\text{M}$ , dilute 50  $\mu\text{M}$  and L28A concentrated 130  $\mu\text{M}$  and dilute 32  $\mu\text{M}$ . The minor shifts of the monomer peaks upon dilution are caused by nonideality effects due to low ionic strength. The experimental sensitivity permits detection of as little as 1% dimer, and none was detected for L27A or L28A. These experiments were performed 25 °C on the same samples utilized for the NMR experiments and contained 1.0 mM deuterated acetic acid, pH 3.9 and 15% D<sub>2</sub>O. For clarity, the profiles of the variants are displaced along the y-axis.

**FIGURE 3. TDP-43 NTD Dynamics and Dimerization Characterized by NMR Relaxation Measurements.** All measurements were obtained at 800 MHz ( $^1\text{H}$ ), in 15% D<sub>2</sub>O, 85% H<sub>2</sub>O, in 1.0 mM acetic acid (d<sub>3</sub>), pH 3.9, 25 °C. WT TDP-43<sub>1-102</sub> 523 microM = purple; TDP-43<sub>1-102</sub> 270 microM = blue; TDP-43<sub>1-77</sub> 290 microM = turquoise; L27A TDP-43<sub>1-102</sub> 200 microM = green; L28A TDP-43<sub>1-102</sub> 130 microM = red (A)  $^{15}\text{N}\{^1\text{H}\}$  NOE ratio; (B) Transversal Rates,  $R_2$  ( $\text{s}^{-1}$ ); (C) Translational Relaxation Rates,  $R_1$  ( $\text{s}^{-1}$ ); (D) The  $R_2 / R_1$  ratio.

**FIGURE 4. Atomistic structural model of TDP-43 dimerization via its NTD.**

(A) The minimal dimerization interface as defined by the intersubunit NOEs involving P19, L27 and E58. These residues are circled on the left and shown in a zoomed view on the right. The two subunits are colored silver and rose gold.

(B) Additional NOEs detected between P36-P36 and certain residues located in the  $\alpha$ -helix, suggest a larger dimerization interface which includes dynamic, transient coiled-coil-like packing of the two helices. These regions are circled on the left and represented in a zoomed view on the right.

(C) Structural model of the complete TDP-43 dimer structure, based on the NTD dimer model and high resolution monomer structure of the NTD<sub>19</sub> reported here (silver/ rose gold), the structures

of the RRM domains (blue) (41) and a representative disordered conformer of the C-terminal region (CTR, purple). The N-terminal domains are labeled and circled in red. The RRM domains and the CTR are labeled black.

**FIGURE 5. Functional add-back of TDP-43 31V/R-32T/R mutant.**

(A) Schematic drawing showing the experimental design. WT TDP-43 (*blue, top*) reduces splicing; this action is minimal when TDP-43 is knocked down (*middle*). In the context of knocked-down endogenous WT TDP-43, the ability of TDP-43 mutant proteins to modulate splicing can be tested (*bottom*).

(B) RT-PCR of CFTR exon 9 isoforms upon transient transfection of HeLa cell line added back with si-resistant TDP-43 WT, F4L and 31V/R-32T/R mutants. Quantification of three independent experiments is also shown. Western blot of transfected plasmids using an anti-Flag antibody and loading control using anti-Tubulin antibody is also shown.

**FIGURE 6. Functional add-back and immunolocalization assay of TDP-43 3L27 and L28 mutants.**

(A) RT-PCR of CFTR exon 9 isoforms upon transient transfection of HeLa cell line added back with si-resistant TDP-43 WT and F4L, L27A, L28A, and L27A-L28A mutants. Quantification of three independent experiments is also shown.

(B) Immunofluorescence experiment of HeLa cells transiently transfected with TDP-43 WT and L27A, L28A, and L27A-L28A. All Flag-TDP-43 proteins are shown in red, nuclei are stained with TO-PRO3 in blue, and the merge is shown.

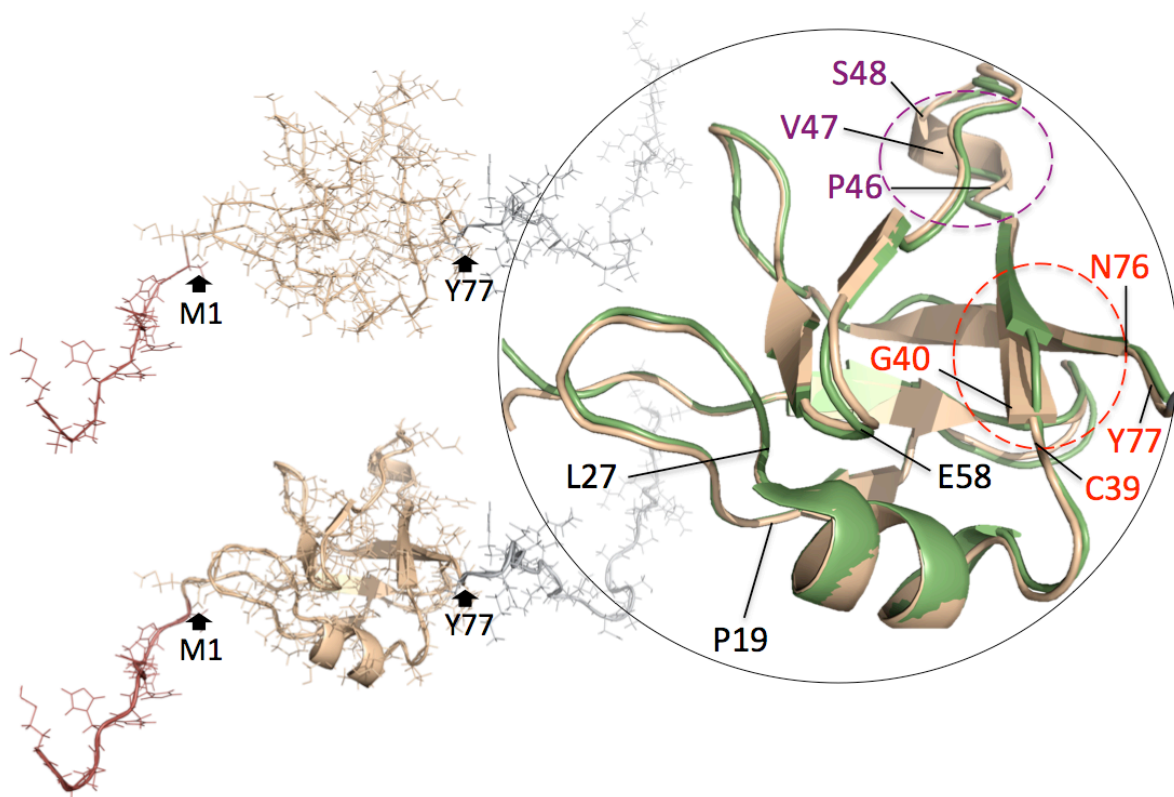
**Figure 7 Functional assay of TDP-43 L27 and L28 mutants in the presence of TDP-43 aggregates.**

(A) RT-PCR of CFTR exon 9 isoforms upon transient transfection of GFP-TDP-12x-F4L stable cell line with TDP-43 WT, TDP-43-F4L, TDP-43-L27A, TDP-43-L28A and TDP-43-L27/28A mutants. Quantification of three independent experiments is also shown.

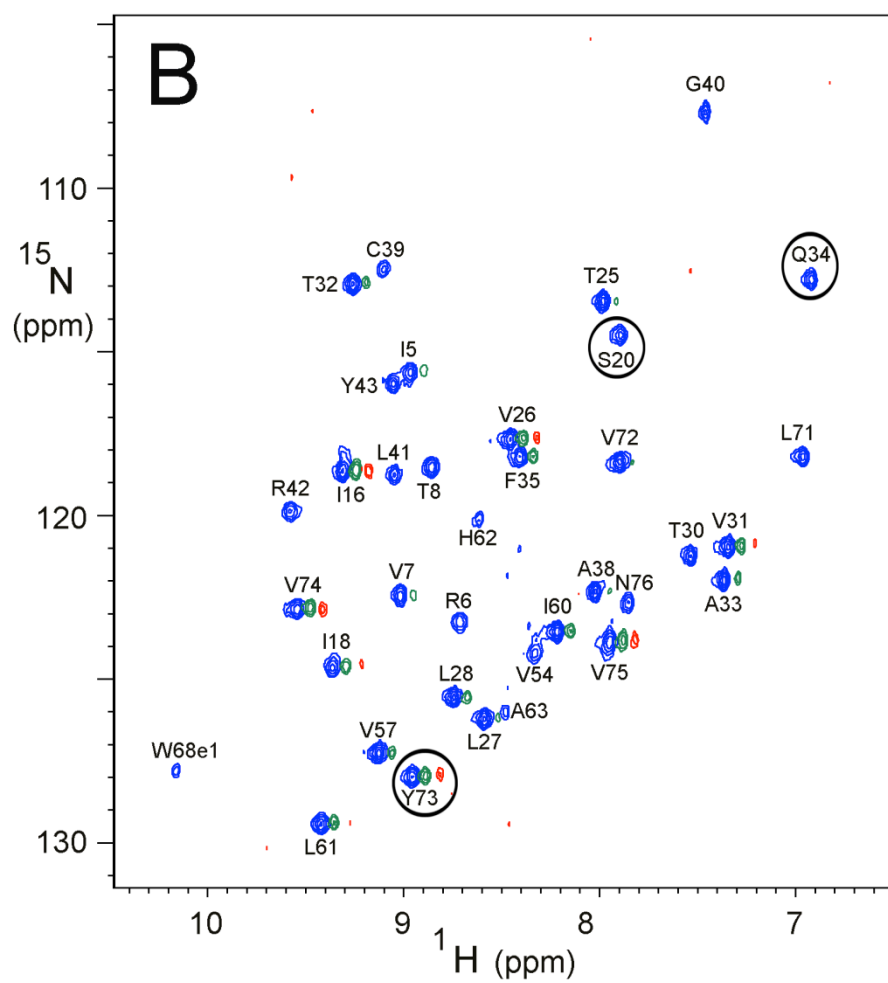
(B) Immunofluorescence experiment of GFP-TDP-12x-F4L stable cell line transiently transfected with TDP-43 WT, TDP-43-F4L, TDP-43-L27A, TDP-43-L28A and TDP-43-L27/28A mutants. A merge of GFP (green) /Anti-Flag (red) /TO-PRO3 (blue) is shown.



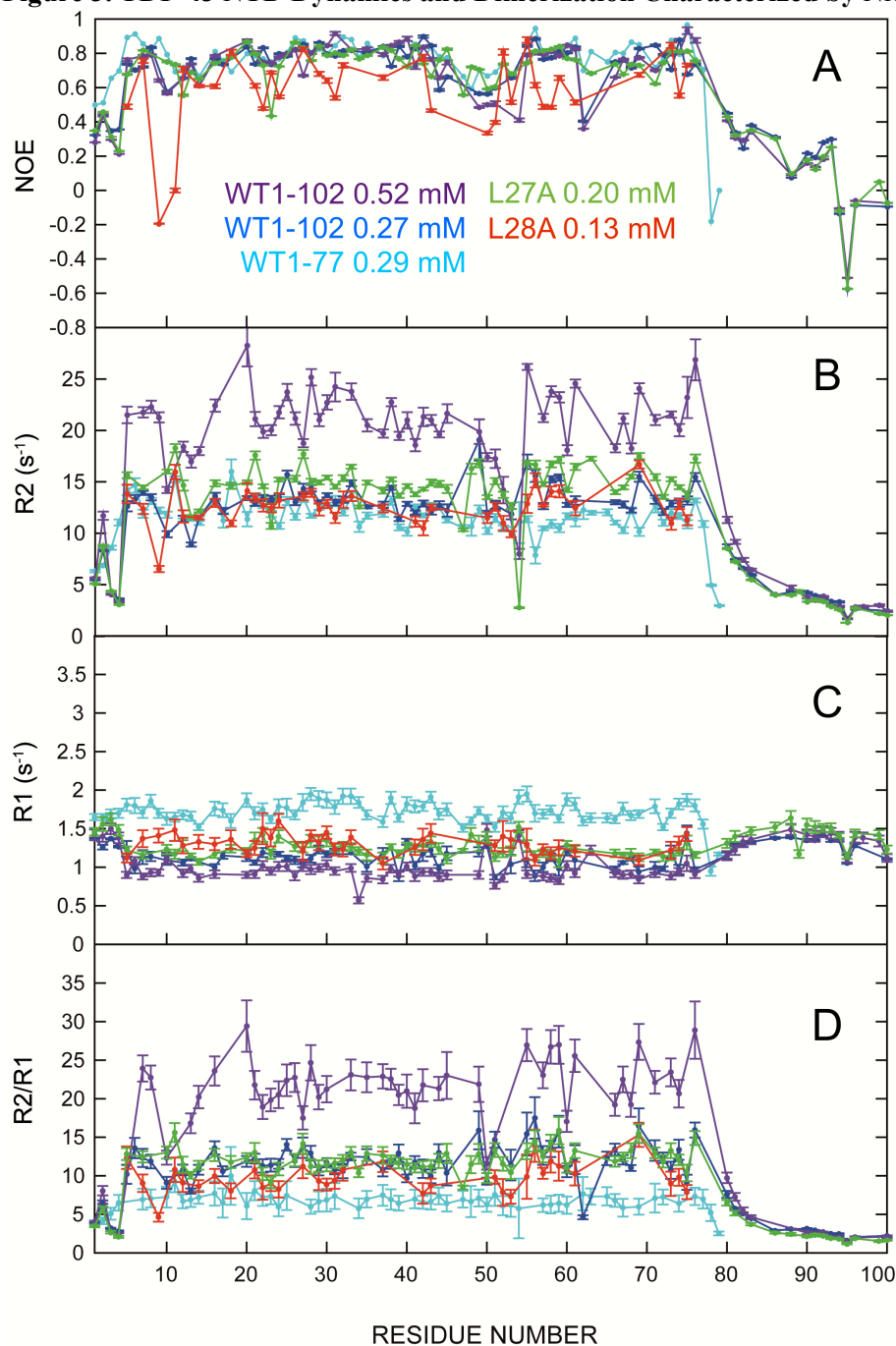
**Figure 1. (A) Structure of TDP-43<sub>1-102</sub> (monomer)**



**Figure 1. (B)** HN groups protected against hydrogen/deuterium exchange

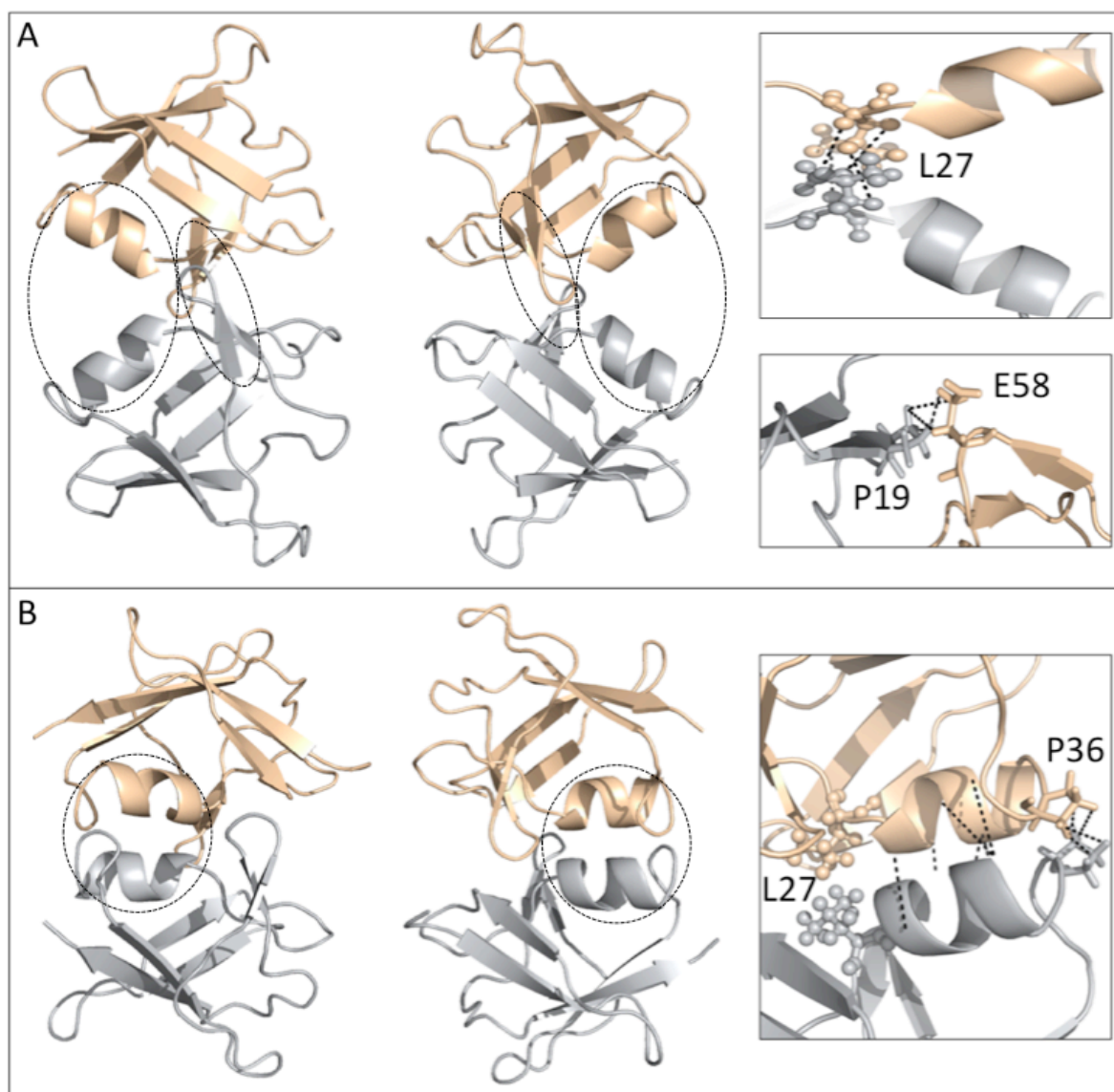




**Figure 3. TDP-43 NTD Dynamics and Dimerization Characterized by NMR**

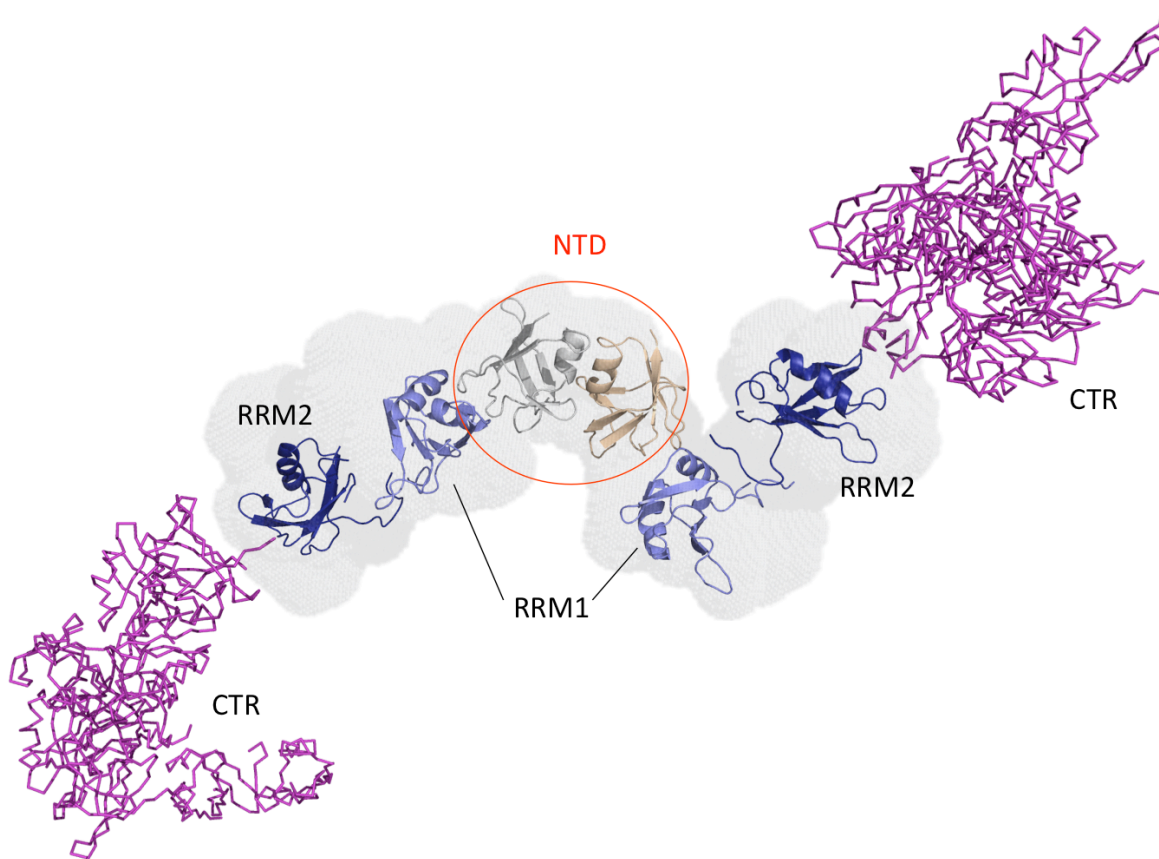
**Relaxation Measurements.** All measurements were obtained at 800 MHz (<sup>1</sup>H), in 15% D<sub>2</sub>O, 85% H<sub>2</sub>O, in 1.0 mM acetic acid (d<sub>3</sub>), pH 3.9, 25 °C. WT TDP-43<sub>1-102</sub> 523 microM = purple; TDP-43<sub>1-102</sub> 145 microM = blue; TDP-43<sub>1-77</sub> 290 microM = turquoise; L27A TDP-43<sub>1-102</sub> 200 microM = green; L28A TDP-43<sub>1-102</sub> 130 microM = red

**Figure 4.** Atomistic model of TDP-43 dimerization via its NTD.

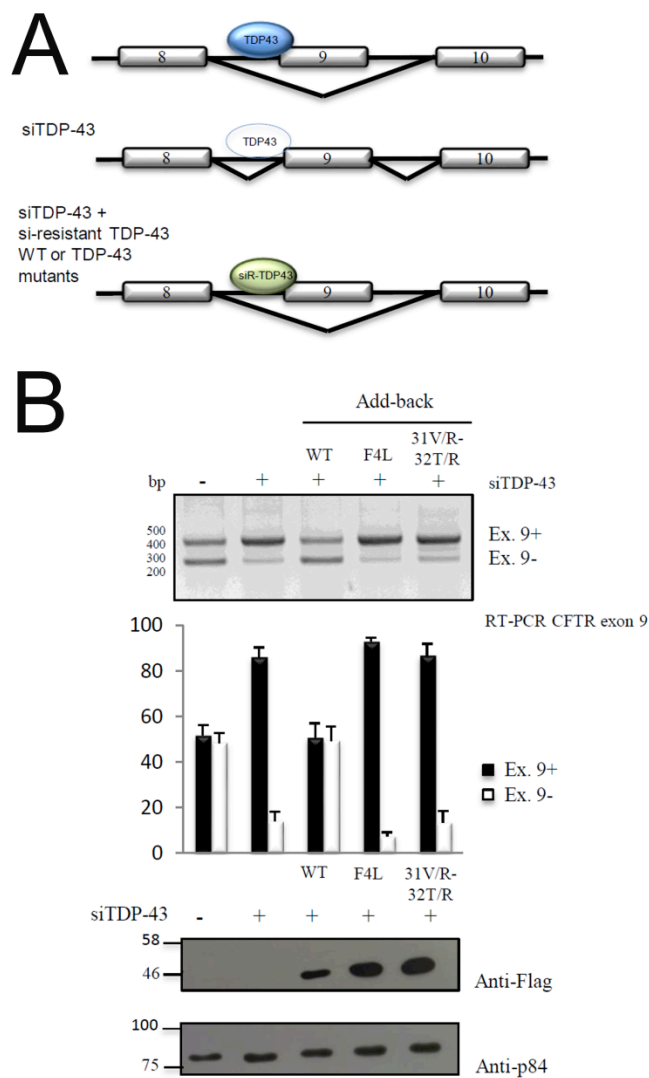




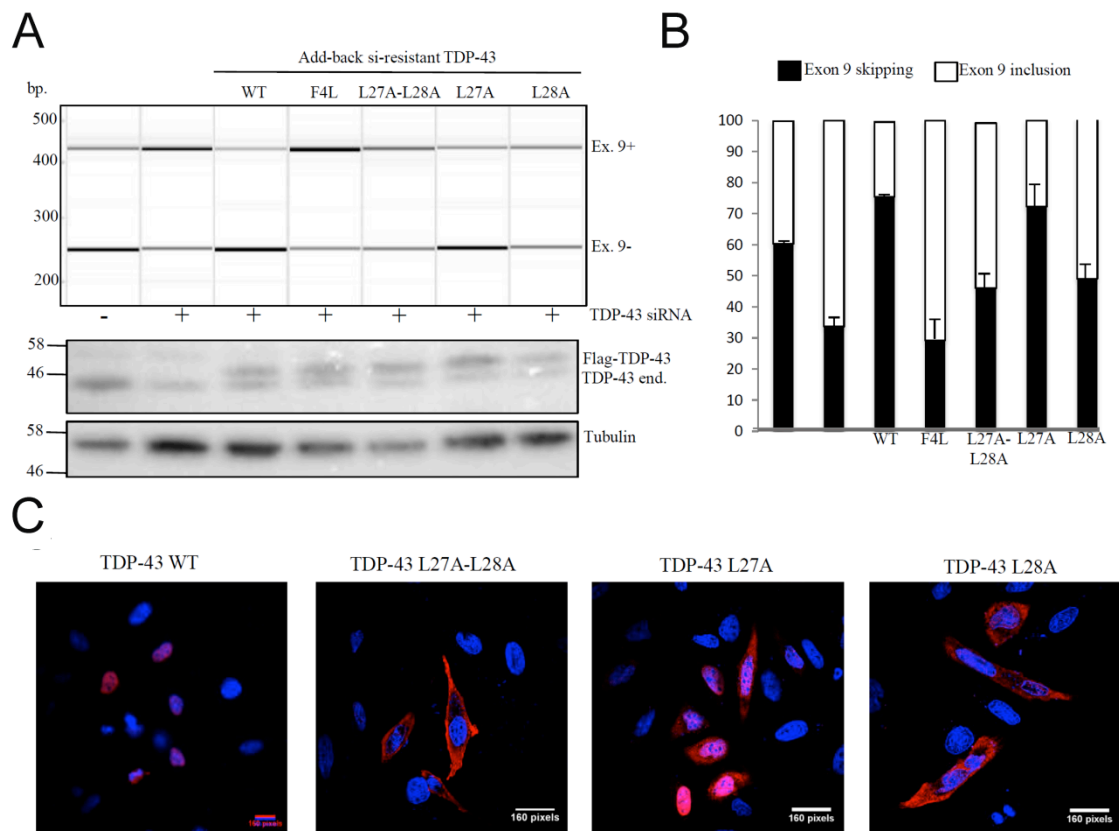
**Figure 4C.** Structural Model of Dimeric TDP-43 (residues 1-414)



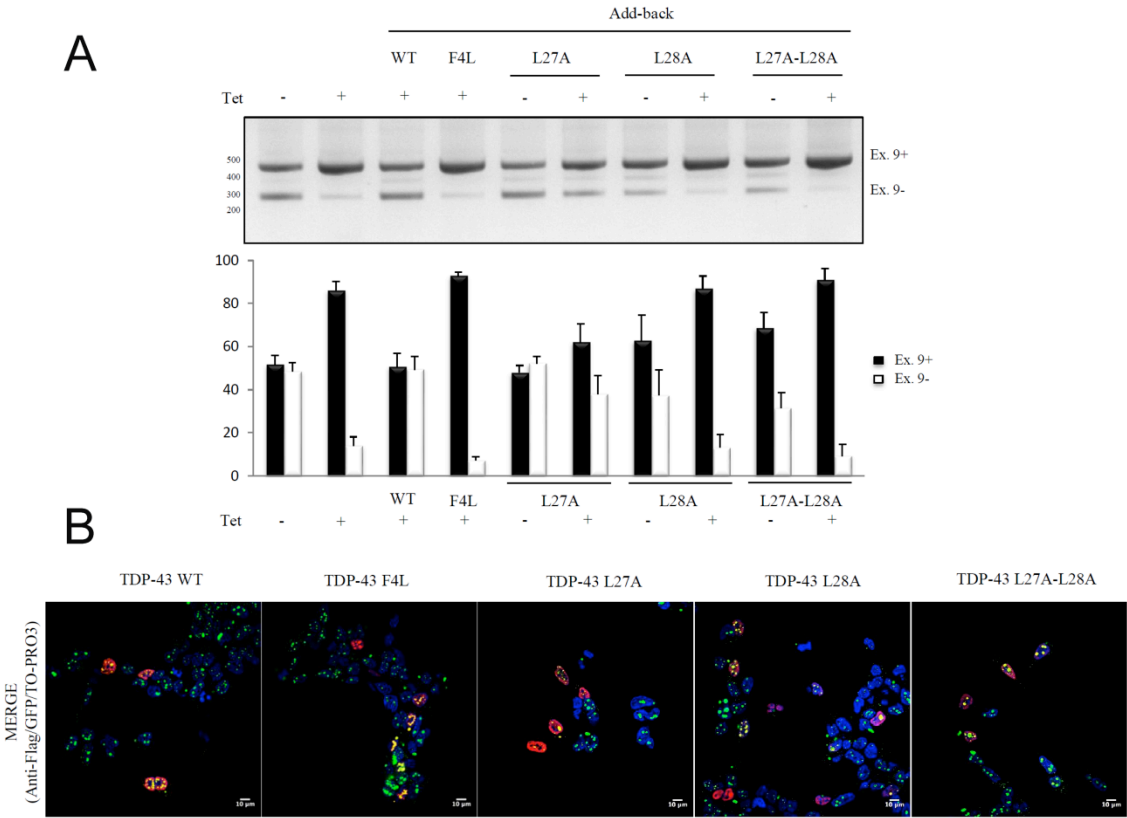
**Figure 5. Functional add-back of TDP-43 31V/R-32T/R mutant.**



**Figure 6.** Functional add-back and immunolocalization assay of TDP-43 L27A and L28A and L27A+L28A variants.



**Figure 7 Functional assay of TDP-43 L27 and L28 mutants in the presence of TDP-43 aggregates.**



**Point mutations in transactive response DNA-binding protein 43 (TDP-43)'s  
N-terminal domain compromise its stability, dimerization and functions**  
Miguel Mompeán, Valentina Romano, David Pantoja-Uceda, Cristiana Stuaní, Francisco E.  
Baralle, Emanuele Buratti and Douglas V. Laurents

*J. Biol. Chem.* published online May 31, 2017

---

Access the most updated version of this article at doi: [10.1074/jbc.M117.775965](https://doi.org/10.1074/jbc.M117.775965)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

Supplemental material:

<http://www.jbc.org/content/suppl/2017/05/31/M117.775965.DC1>

This article cites 0 references, 0 of which can be accessed free at

<http://www.jbc.org/content/early/2017/05/31/jbc.M117.775965.full.html#ref-list-1>