


ARTICLE

DOI: 10.1038/s41467-017-00062-0

OPEN

# Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation

Tariq Afroz<sup>1</sup>, Eva-Maria Hock<sup>1</sup>, Patrick Ernst<sup>2</sup>, Chiara Foglieni<sup>3</sup>, Melanie Jambeau<sup>1</sup>, Larissa A.B. Gilhesp<sup>1</sup>, Florent Laferriere<sup>1</sup>, Zuzanna Maniecka<sup>1</sup>, Andreas Plückthun<sup>2</sup>, Peer Mittl<sup>2</sup>, Paolo Paganetti <sup>3</sup>, Frédéric H.T. Allain<sup>4</sup> & Magdalini Polymenidou<sup>1</sup>

TDP-43 is a primarily nuclear RNA-binding protein, whose abnormal phosphorylation and cytoplasmic aggregation characterizes affected neurons in patients with amyotrophic lateral sclerosis and frontotemporal dementia. Here, we report that physiological nuclear TDP-43 in mouse and human brain forms homo-oligomers that are resistant to cellular stress. Physiological TDP-43 oligomerization is mediated by its N-terminal domain, which can adopt dynamic, solenoid-like structures, as revealed by a 2.1 Å crystal structure in combination with nuclear magnetic resonance spectroscopy and electron microscopy. These head-to-tail TDP-43 oligomers are unique among known RNA-binding proteins and represent the functional form of the protein in vivo, since their destabilization results in loss of alternative splicing regulation of known neuronal RNA targets. Our findings indicate that N-terminal domain-driven oligomerization spatially separates the adjoining highly aggregation-prone, C-terminal low-complexity domains of consecutive TDP-43 monomers, thereby preventing low-complexity domain inter-molecular interactions and antagonizing the formation of pathologic aggregates.

<sup>1</sup>Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. <sup>2</sup>Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. <sup>3</sup>Laboratory for Biomedical Neurosciences, Neurocenter of Southern Switzerland, Via Tesserete 46, CH-6900 Lugano, Switzerland. <sup>4</sup>Institute of Molecular Biology and Biophysics, ETH Zurich, CH-8093 Zurich, Switzerland. Eva-Maria Hock and Patrick Ernst contributed equally to this work. Correspondence and requests for materials should be addressed to M.P. (email: [magdalini.polymenidou@imls.uzh.ch](mailto:magdalini.polymenidou@imls.uzh.ch))

Amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) are two intimately related adult-onset neurodegenerative disorders<sup>1</sup>. In ALS the loss of motor neurons leads to fatal paralysis typically within 1–5 years of onset, whereas FTD causes severe atrophy of the frontal and temporal lobes leading to behavioral and language dysfunction<sup>2–5</sup>. The RNA-binding protein (RBP) TDP-43 (Trans-activation response element (TAR) DNA-binding protein 43) is mislocalized and aggregated into neuronal cytoplasmic ubiquitinated inclusions in the vast majority of ALS, as well as approximately half of FTD cases<sup>2</sup>. Even though the loss of normal nuclear localization and cytoplasmic TDP-43 aggregation correlates with neurodegeneration, the exact mechanisms of neurotoxicity remain elusive. Moreover, the molecular mechanisms triggering TDP-43 pathology in ALS and FTD remain poorly understood, in part due to lack of high-resolution structural information of TDP-43 in the physiological and pathological state.

TDP-43 is a nucleo-cytoplasmic shuttling protein comprised of two RNA recognition motifs (RRMs) that bind TG-/UG-repeat nucleic acids in a sequence-specific fashion and are indispensable for its roles in RNA metabolism<sup>3, 6–8</sup>. TDP-43 also contains a C-terminal prion-like or low-complexity domain (LCD)<sup>9</sup>, mediating protein–protein interactions<sup>10</sup> and also its incorporation into stress granules<sup>11</sup>, potentially via its property to phase separate<sup>12, 13</sup>. Furthermore, the TDP-43 LCD is crucially involved in disease, since it is proteolytically cleaved and abnormally phosphorylated, leading to its cytoplasmic accumulation in complex with the full-length protein<sup>2, 14</sup>. Finally, TDP-43 contains an N-terminal region spanning its first 80 amino acids, whose role in the function and/or malfunction of the protein remains unclear. In fact, due to the lack of sequence homology of this region with any known structures, the N-terminal domain (NTD) of TDP-43 was initially thought to be unstructured. However, it was recently reported that monomeric TDP-43 NTD can adopt a Ubiquitin-like<sup>15</sup> or DIshevelled and aXin (DIX)-domain-like<sup>16</sup> fold in solution. Intriguingly, while several recent studies highlighted the importance of NTD for functional TDP-43 dimerization<sup>17</sup> and nucleic acid interaction<sup>15, 18–21</sup>, others argued that the same domain promoted pathologic cytoplasmic aggregation and neurotoxicity<sup>20, 22, 23</sup>. Moreover, a small fraction of TDP-43 was reported to exist as dimers in cells, which led to speculation that TDP-43 dimers may initiate or “seed” the formation of high molecular weight pathologic TDP-43 aggregates<sup>17</sup>.

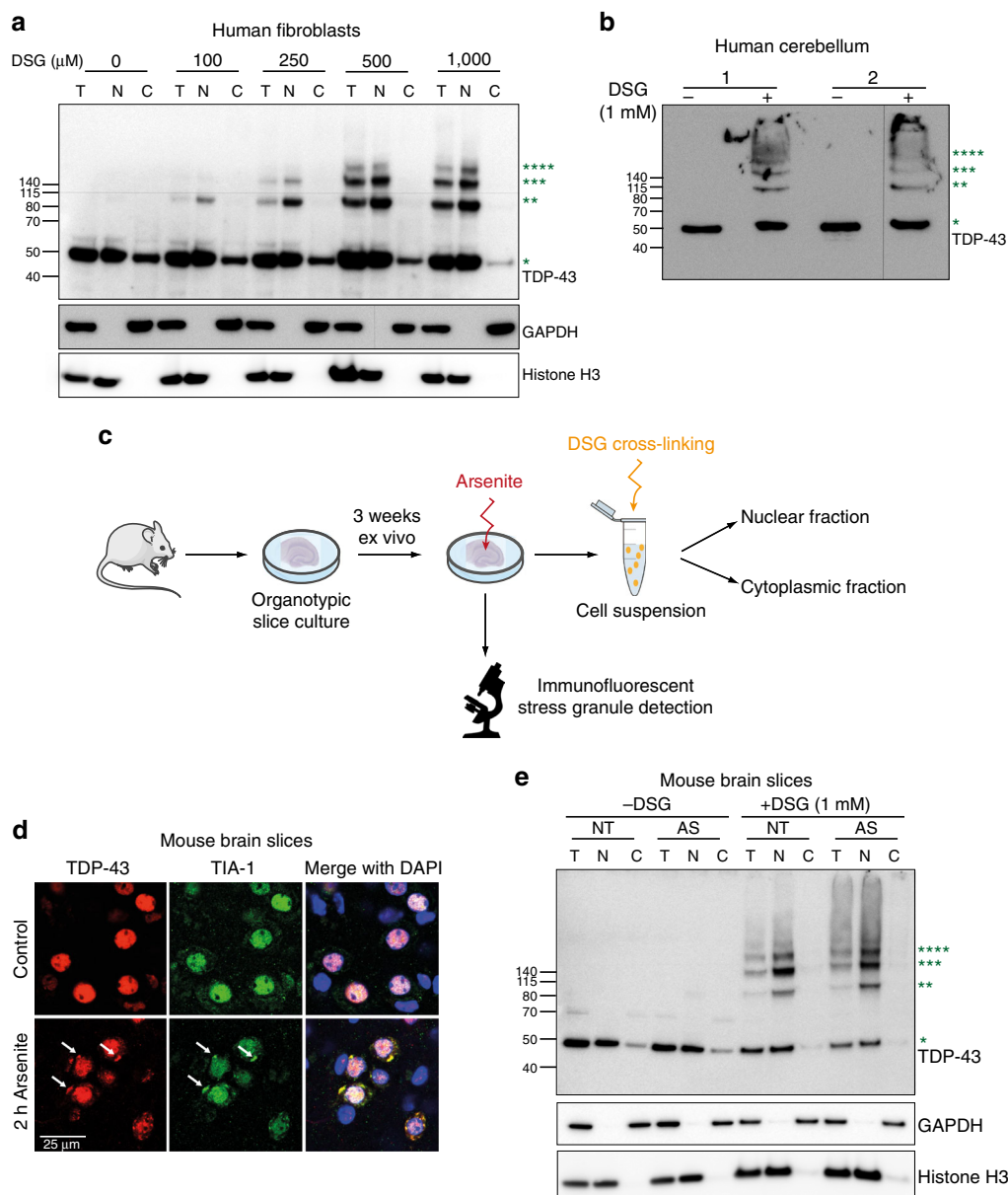
Here, we show that physiological TDP-43 exists as nuclear oligomers that are distinct from cytoplasmic complexes formed upon cellular stress or pathologic aggregates. To elucidate the molecular basis of physiological TDP-43 oligomerization, we determined the crystal structure of TDP-43 NTD at 2.1 Å resolution, which revealed an unprecedented mode of head-to-tail interactions between monomers generating solenoid-like polymers. Consistent with the crystal structure, solution NMR spectroscopy confirmed the dynamic nature of inter-molecular and low micromolar affinity electrostatic interactions that stabilize these polymers. Destabilizing oligomerization by point mutations resulted in loss of TDP-43 regulation of alternative splicing of known neuronal RNA targets, indicating that these dynamic TDP-43 oligomers are the functional form of the protein in vivo. Tripartite GFP complementation experiments in cells illustrate that physiological TDP-43 oligomerization prevents LCD intermolecular interactions. Importantly, we show that NTD-driven TDP-43 oligomerization antagonizes pathologic aggregation. This dynamic head-to-tail polymerization of TDP-43, which is reminiscent of DIX domains<sup>24</sup> involved in Wnt signaling, is unique among RBPs and broadens our understanding of TDP-43 function. Most excitingly, our findings indicate that stabilization of functional TDP-43 oligomers could have therapeutic potential by counteracting pathologic aggregation and restoring nuclear function.

## Results

**TDP-43 forms physiological oligomers in human tissues.** In order to capture the in vivo protein state, we systematically analyzed physiological TDP-43 in cells and tissues by in situ, chemically induced, cross-linking prior to cell/tissue lysis. For this purpose, we used disuccinimidylglutarate (DSG), a membrane-permeable cross-linker, which allows preservation of native protein–protein interactions<sup>25</sup>. We performed DSG cross-linking in normal human fibroblasts and brain tissue, followed by nucleo-cytoplasmic fractionation and immunoblotting, in order to analyze the subcellular localization of TDP-43 oligomers and monomers. Upon increasing the DSG concentration, a specific ladder of slowly migrating TDP-43 complexes is detected, consistent with increasing multimeric species of TDP-43 in human fibroblasts (Fig. 1a, Supplementary Fig. 1a–e) and human brain (Fig. 1b, Supplementary Fig. 1f–g). The slightly different migration pattern of TDP-43 oligomers in human motor cortex (Supplementary Fig. 1g) is due to the different gel (4–12% polyacrylamide) used to allow better resolution of autopsy human brain material, which after cross-linking migrates poorly in the regular 12% polyacrylamide gel (Fig. 1b). The higher molecular weight TDP-43 species correspond to distinct molecular masses of the dimer, trimer, tetramer and higher, indicating that native cellular TDP-43 exists not only as dimers, but rather in a spectrum of oligomeric species. The fraction of TDP-43 present in high molecular weight oligomers is comparable among various samples analyzed (Supplementary Fig. 1h). To further validate the specificity of the DSG cross-linking protocol, we investigated the cross-linking pattern of other hnRNPs such as hnRNPA1 and FUS that harbor similar RNA binding and LCDs as TDP-43, but lack the NTD. For hnRNPA1, we detect a specific dimer band predominantly in the nuclear fractions (Supplementary Fig. 2a). However, no intermediate oligomeric bands are detected as for TDP-43 (Fig. 1a). Interestingly, for both hnRNPA1 and FUS, very high molecular weight complexes that are unable to enter the 4–12% denaturing polyacrylamide gels are detected (Supplementary Fig. 2a). This is consistent to previous reports suggesting oligomerization of hnRNPA1<sup>26</sup> and FUS<sup>27</sup> upon RNA binding. In contrast, wild-type superoxide dismutase 1 (SOD1), which is known to exist exclusively as cytoplasmic homodimers<sup>28</sup>, showed only the presence of dimeric protein without any non-specific cross-linking, even at the highest DSG concentration used both in human fibroblasts and mouse brain slices (Supplementary Fig. 2b–c). Multimeric TDP-43 species are detected predominantly in the nucleus, correlating with the normal nuclear localization of TDP-43 in healthy cells (Fig. 1a).

## Nuclear TDP-43 oligomers are resistant to oxidative stress.

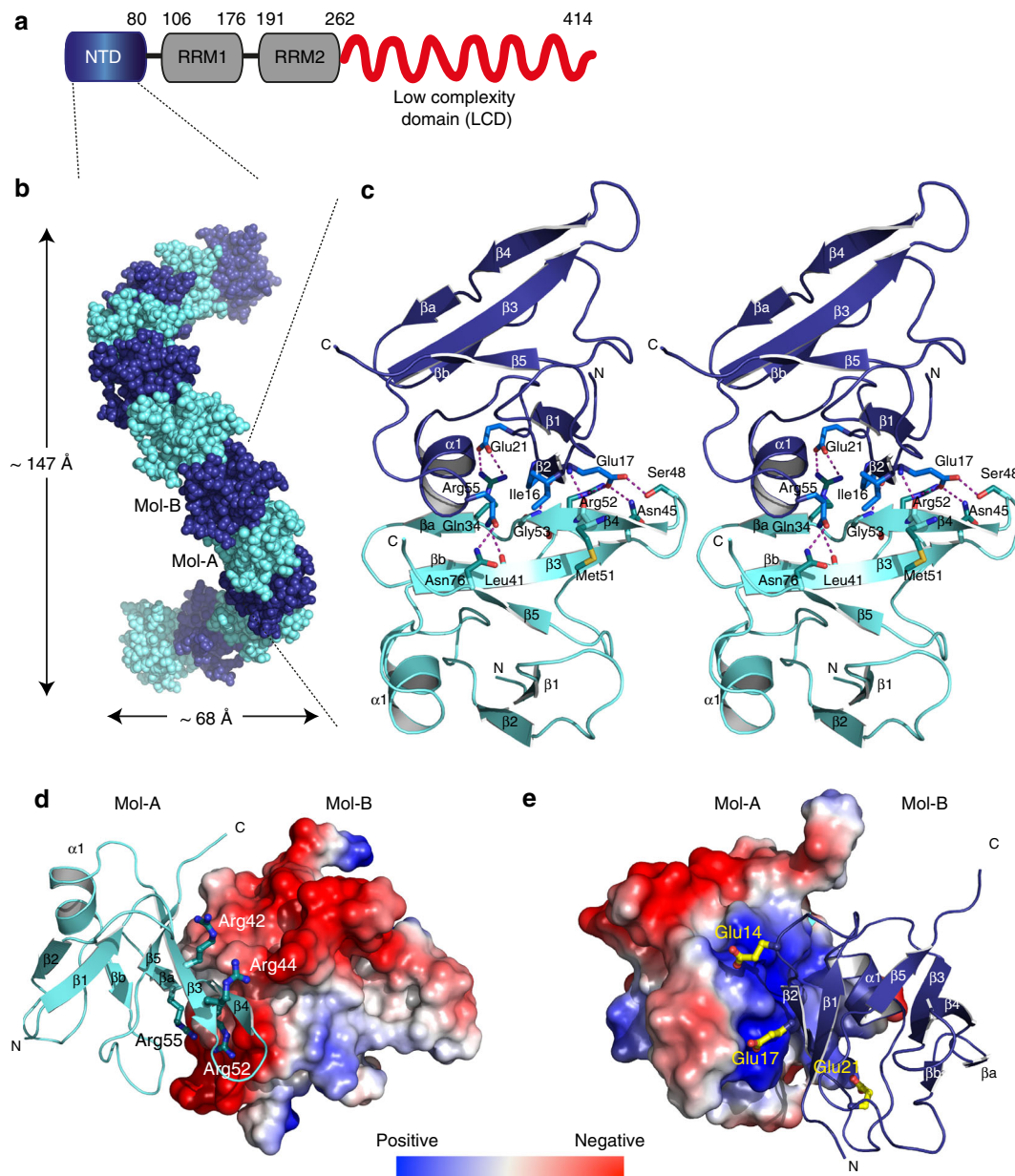
TDP-43 is recruited to stress granules upon oxidative stress and forms high molecular weight complexes in non-reducing conditions<sup>29–31</sup>. In order to determine if normal nuclear TDP-43 oligomerization is changed upon cellular stress, we treated mouse cortico-hippocampal organotypic slice cultures with sodium arsenite and then performed in situ cross-linking by DSG followed by nucleo-cytoplasmic fractionation (Fig. 1c). Arsenite induces oxidative stress by altering the redox potential in the cell. Consistent with previously published work on cell lines<sup>29</sup>, 2 h of arsenite treatment results in a small fraction of TDP-43 to localize into cytoplasmic TIA-1 positive granules (white arrows in Fig. 1d). In both control and arsenite-treated slices, the majority of oligomeric TDP-43 resides in the nuclear fraction, suggesting that this oligomeric state is distinct from the cytoplasmic TDP-43 pool residing in stress granule (Fig. 1e, Supplementary Fig. 2d–e). Furthermore, nuclear TDP-43 oligomerization is independent of



**Fig. 1** TDP-43 oligomers are expressed in human cells and brain and resist oxidative stress. **a** Immunoblots (from 12% denaturing polyacrylamide gels) of human fibroblast fractions obtained upon incubation with increasing concentration of DSG cross-linker followed by nucleo-cytoplasmic fractionation. In the *upper panel*, high molecular weight bands (marked with *green asterisks*) are detected with anti-human TDP-43 antibody at increasing concentrations of DSG in the nuclear fraction (N) and in total cell lysates (T), but not in the cytoplasmic fraction (C). *Lower panels* show immunoblots for cytoplasmic marker GAPDH, or nuclear marker histone H3. Immunoblots are representative of three independent experiments. **b** Immunoblots (from 12% denaturing polyacrylamide gels) of human cerebellum samples. High molecular weight TDP-43 bands are detected in 1 mM DSG cross-linked samples (*green asterisks*) compared to samples obtained in the absence of cross-linker. Immunoblots are representative of three independent experiments. Full immunoblots with molecular weight marker are shown in Supplementary Fig. 1. **c** Schematic diagram depicting the experimental set-up for oxidative stress treatment in mouse organotypic slice cultures followed by DSG mediated cross-linking. Subset of images (such as eppendorf tubes and microscope) in the schematic figure were adapted from the Servier medical art (<http://www.servier.com/slidekit>). **d** Confocal microscopy images of non-treated (control) and arsenite-treated (2 h Arsenite) mouse brain slices, immunostained with antibodies against TDP-43 and TIA-1. Cytoplasmic TDP-43 granules (*red*) co-localize with the stress-granule marker TIA-1 (*green*). Merged images including DAPI-staining of nuclei (*blue*) are shown in the *right panels*. **e** Immunoblots (from 12% denaturing polyacrylamide gels) of control (NT) and arsenite-treated (As) mouse brain slices after cross-linking by DSG and nucleo-cytoplasmic fractionation. DSG-cross-linked oligomeric TDP-43 (marked with *green asterisks*) remains localized in the nuclear fraction after oxidative stress. Immunoblots are representative of three independent experiments

changes in the cellular redox state and is, therefore, cysteine independent. Thus, nuclear TDP-43 oligomers are structurally distinct from the higher molecular weight species resulting from cysteine oxidation and disulfide bond formation, which are specifically formed in response to oxidative stress, predominantly in the cytosol<sup>30, 32</sup>.

**Crystal structure of TDP-43 NTD shows superhelical packing.** Incorporation of TDP-43 in cytoplasmic stress granules is mediated via its C-terminal LCD<sup>11, 22</sup>. Moreover, the nuclear localization of TDP-43 oligomers, in conjunction with their resistance to oxidative stress (Fig. 1c–e) suggested that nuclear TDP-43 oligomerization is distinct and potentially structurally independent of



**Fig. 2** Crystal structure of TDP-43 NTD at 2.1 Å resolution. **a** Schematic domain organization of human TDP-43 with the N-terminal domain (NTD) shown in blue, RNA recognition motifs (RRMs) in gray and the low-complexity domain (LCD) in red with domain boundaries indicated on top. **b** Crystals of TDP-43 NTD show helical filaments with single molecules arranged in head-to-tail fashion as seen from the side. Atoms of Mol-A and -B (comprising asymmetric unit) are shown as cyan and blue spheres, respectively. The dimensions of the crystallographic helix are indicated. **c** Wall-eyed stereo view of two TDP-43 NTD molecules (Mol-A, Mol-B) in the asymmetric unit. The two molecules are shown in cartoon representation with the secondary structure elements and the N- and C-termini of the molecules labeled. The side chains making inter-molecular contacts are shown in stick representation in the corresponding domain color and labeled. Inter-molecular hydrogen bonds are shown as dotted magenta lines. Similar color and labeling schemes are used in other figures unless stated. **d, e** Electrostatic charge on the surface of two TDP-43 NTD molecules (Mol-A, Mol-B) of the asymmetric unit. Amino-acid residues in the positively charged head **d** and negatively charged tail region **e** are shown in stick representation and labeled. Positive potential is shown in blue and negative potential in red

LCD. We, therefore, focused on the NTD of TDP-43, whose role in protein dimerization was previously suggested<sup>18, 19</sup>.

In order to understand the molecular basis of TDP-43 oligomerization, we aimed at determining a high-resolution structure of TDP-43 NTD. Based on the previously reported domain boundaries<sup>15, 16</sup>, we expressed and purified recombinant human TDP-43 NTD (amino acids 1–80) in amounts amenable for structural analysis (Fig. 2a, Supplementary Fig. 3a). TDP-43 NTD was purified under native conditions, without the use of any denaturant during the purification process, in contrast to earlier

reports where some denaturants (such as urea) were included that may lead to destabilization of oligomers in solution<sup>15, 16</sup>. Subsequently, vapor diffusion crystallization experiments of purified recombinant protein were set up that led to initial crystallization hits. These small crystals were used for microseeding experiments to grow larger crystals (70 × 70 × 200 μm) that diffracted to 2.1 Å resolution (Supplementary Fig. 3b, Table 1). Initial molecular replacement trials using the deposited TDP-43 NTD NMR structure<sup>16</sup> (PDB ID code 2N4P) failed. However, crystals were grown in sodium cacodylate buffer and consequently

**Table 1** Data collection and refinement statistics

TDP-43 NTD	
<i>Data collection</i>	
Space group	$P6_3$
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	89.52, 89.52, 49.15
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120
Resolution (Å)	44.76–2.10 (2.15–2.10)
$R_{\text{meas}}$	0.108 (2.03)
$R_{\text{merge}}$	0.103 (1.94)
$I/\sigma I$	16.7 (1.6)
$CC_{1/2}$	1.00 (0.59)
Completeness (%)	98.8 (100.0)
Redundancy	10.41 (10.92)
$\Delta_{\text{ano}}/\sigma\Delta_{\text{ano}}$	1.3 (0.7)
<i>Refinement</i>	
Resolution (Å)	49.15–2.10
No. of reflections	12,500
$R_{\text{work}}/R_{\text{free}}$	0.166/0.206
No. of atoms	1344
Protein	1192
Ligand/ion	
Acetate	1
Cacodylate	4
Water	108
<i>B</i> -factors	
Protein	48.46
Ligand/ion	54.16
Water	56.2
R.m.s. deviations	
Bond lengths (Å)	0.031
Bond angles (°)	2.72

One crystal was used for data collection. Values in parentheses are for the highest-resolution shell. Friedel pairs were treated as different reflections due to the significant anomalous signal.

starting phases were obtained by single-wavelength anomalous dispersion making use of the strong anomalous signal that was generated by the arsenic K-edge (Table 1). Using the anomalous signal, we obtained a substructure with four cacodylate ions that were covalently bound to surface exposed cysteine residues.

TDP-43 NTD crystallized in space group  $P6_3$  and two molecules (Mol-A, Mol-B) are found in the asymmetric unit, whose structures are virtually identical, indicated by a root-mean square deviation (r.m.s.d) of 0.612 Å ( $C\alpha$  atoms of residues 1–80). Crystallographic symmetry generated three intertwined superhelices of TDP-43 NTDs packed against each other, building a tubular superstructure (Supplementary Fig. 3c, d). Each of these tubules is surrounded by another six tubules, thereby forming the lateral contacts of the crystal. Within each superhelix, every turn is composed of 12 TDP-43 NTD monomers (two molecules per asymmetric unit) with a rise of 24.6 Å between two asymmetric units, a helical pitch of 147.6 Å and a radius of 34 Å to the center of mass, respectively, (Fig. 2b).

**Head-to-tail interaction of TDP-43 NTD monomers.** Except for the N-terminal His-tag, all other residues were well resolved in the final electron density map. The structure of the TDP-43 NTD monomer is similar to the previously reported NMR structure with a global r.m.s.d of 1.97 Å for Mol-A and 2.12 Å for Mol-B for  $C\alpha$  atoms (Supplementary Fig. 3e), explaining the failure of solving the structure by molecular replacement. Significant differences are seen in the orientation and conformation of loops, especially the  $\beta$ 3- $\beta$ 4 loop that seems to be flexible based on the NMR data<sup>16</sup>, however, it is involved in inter-molecular interactions between monomers in

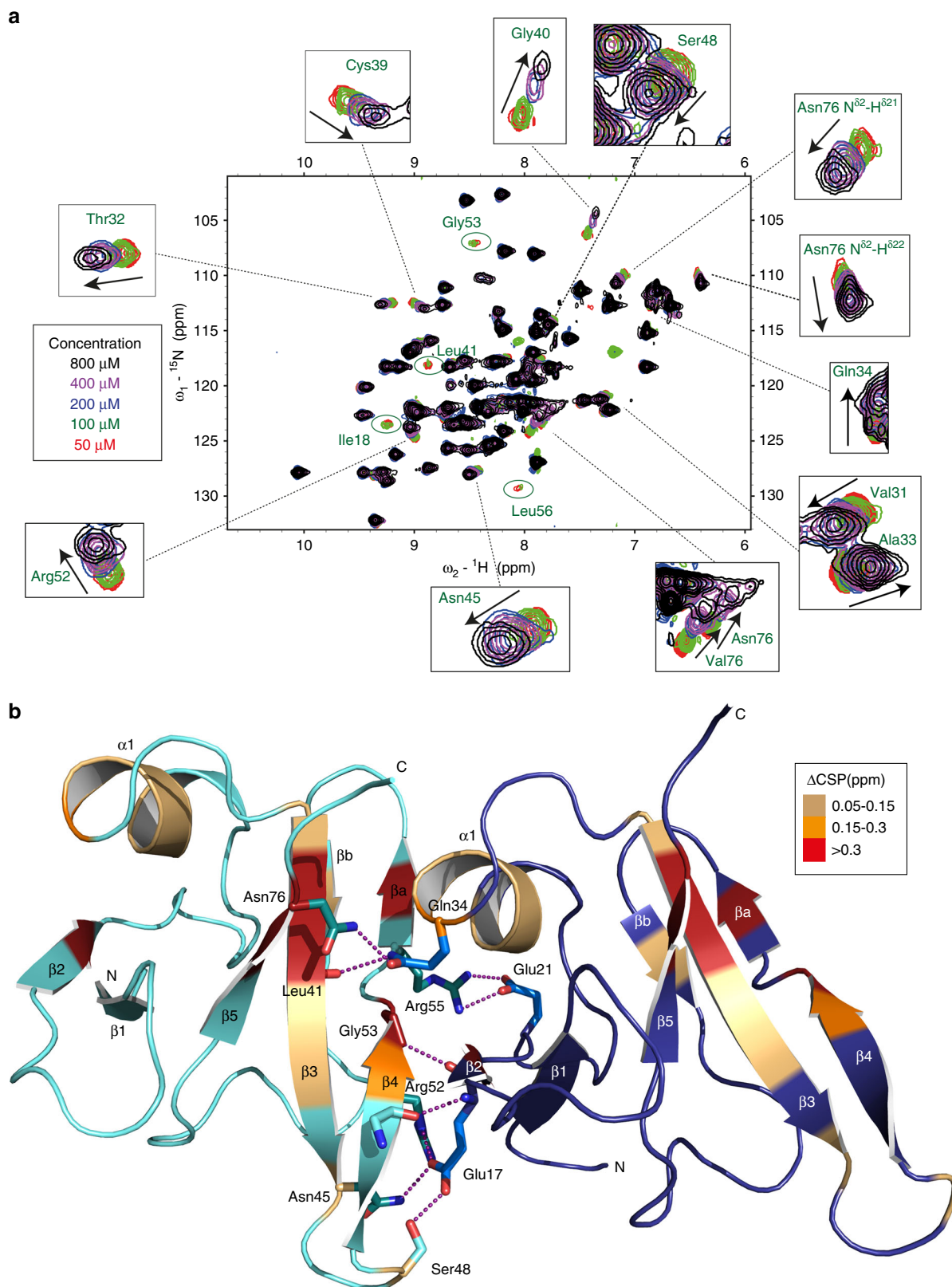
the crystal structure (Fig. 2c, Supplementary Fig. 3f). TDP-43 NTD adopts a compact fold comprised of a five-stranded  $\beta$ -sheet ( $\beta$ 2- $\beta$ 1- $\beta$ 5- $\beta$ 3- $\beta$ 4) packing against one  $\alpha$ -helix (Fig. 2c). An additional  $\beta$ -hairpin is present in the loop between  $\beta$ 4- and  $\beta$ 5-strands ( $\beta$ a and  $\beta$ b in Fig. 2c). Such domain architecture is reminiscent of Ubiquitin and DIX domains (see Discussion).

The inter-molecular interface between consecutive monomers is conferred by two distinct regions in the protein. The head region (on the right when viewing the  $\beta$ -sheet with the  $\beta$ 1-strand pointing up) of TDP-43 NTD is positively charged and interactions are made by residues mainly from  $\beta$ 3- (Asn45),  $\beta$ 4- (Arg52, Arg55),  $\beta$ 5- (Asn76) strands and the  $\beta$ 3- $\beta$ 4 loop (Ser48) (Fig. 2c, d). In contrast, the tail (on the left, viewing from  $\beta$ -sheet with  $\beta$ 1-strand going up) of the molecule is negatively charged (Fig. 2e) and is comprised of residues from the  $\beta$ 2- (Glu17) strand, the  $\alpha$ 1-helix (Gln34) and the  $\beta$ 1- $\beta$ 2 (Glu14) and  $\beta$ 2- $\alpha$ 1 (Glu21) loops that interact with the adjacent molecule (Fig. 2c, e). Such interaction between NTD molecules results in charge complementation. In addition, at each interface,  $\beta$ 2-strand of one monomer forms a parallel inter-molecular  $\beta$ -bridge with the  $\beta$ 4-strand of the adjacent monomer in which the main chain atoms of Glu17 ( $\beta$ 2-strand) forms H-bonds with the backbone of Met51 and Gly53 ( $\beta$ 4-strand) (Fig. 2c, Supplementary Fig. 4a, b). This interaction is strengthened by the guanidinium groups of Arg55 and Arg52 that form salt bridges with the carboxyl groups of Glu21 and Glu17, respectively (Fig. 2c, Supplementary Figs. 4b and 5a, b). Additionally, the carboxyl group of Glu17 forms specific hydrogen bonds with the hydroxyl group of Ser48 and the carboxamide group of Asn45 (Fig. 2c, Supplementary Figs. 4b and 5a). Moreover, the aliphatic side chains of Ile16 and Pro19 (in Mol-B) interact with the aliphatic parts of side chains of Met51, Arg52, and Arg55 (in Mol-A), respectively, to create a small hydrophobic contact at the interface (Fig. 2c, Supplementary Fig. 4b). Finally, the  $\alpha$ 1-helix of one monomer packs against the  $\beta$ -sheet of the other molecule where in the carboxamide of Gln34 ( $\alpha$ 1-helix) makes hydrogen bonds to the carboxamide nitrogen of Asn76 ( $\beta$ 5-strand) and the backbone carbonyl oxygen of Leu41 ( $\beta$ 3-strand) (Fig. 2c, Supplementary Figs. 4a and 5c).

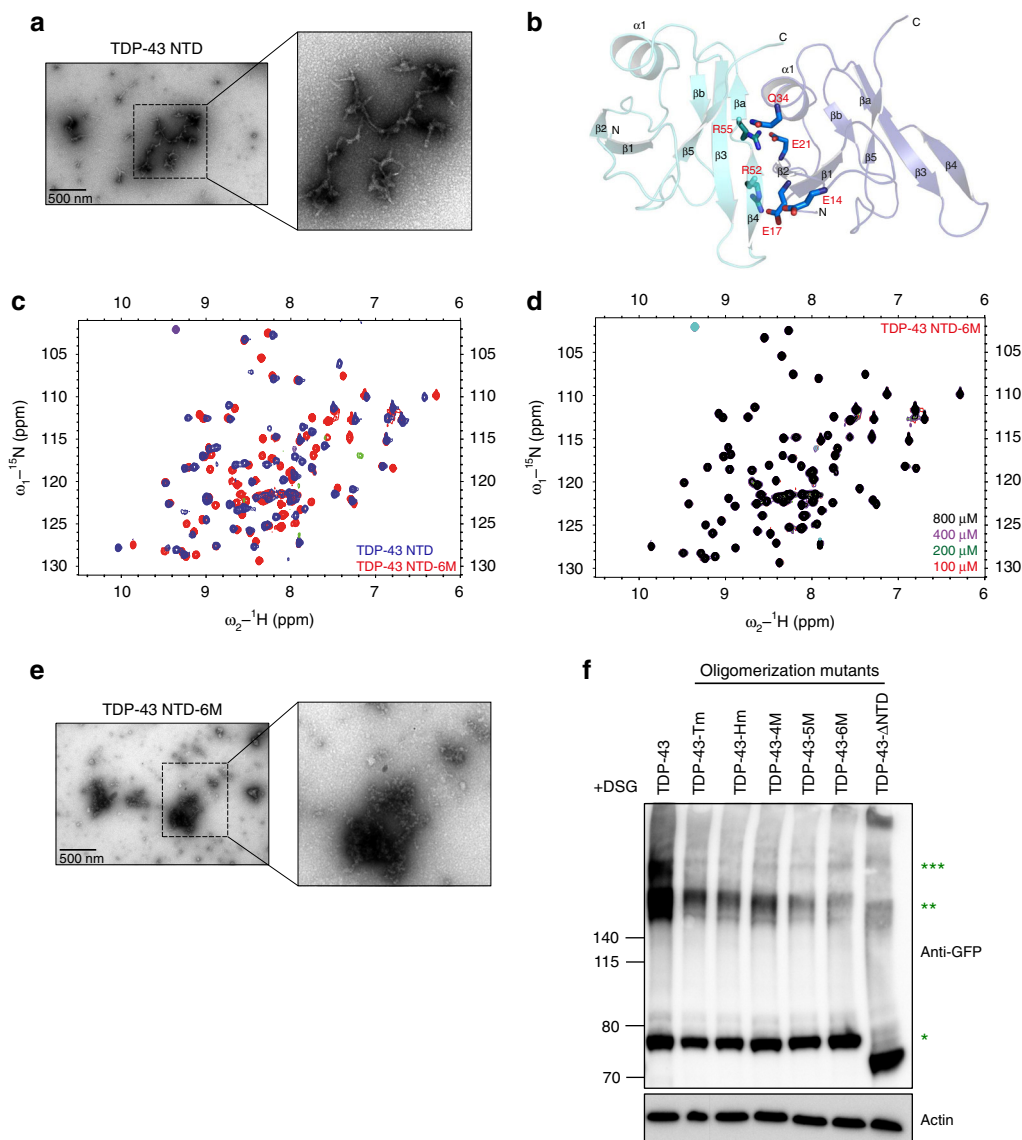
Thus, the two NTD molecules in the asymmetric unit interact extensively with one another in head-to-tail fashion (Fig. 2c, Supplementary Fig. 3f), burying a total surface area of 644.2 Å<sup>2</sup> at the interface between Mol-A and Mol-B within the asymmetric unit and 648.1 Å<sup>2</sup> between Mol-B and Mol-A' between two adjacent asymmetric units (the prime refers to a molecule related by crystallographic symmetry). This area is comparable to the buried surface area in the oligomeric structures formed by the DIX domains of both Axin-1 (768 Å<sup>2</sup>) and Dishevelled Segment Polarity Protein-2 (425–470 Å<sup>2</sup>).

### Dynamic polymerization of TDP-43 NTD revealed by NMR.

In order to confirm and validate the interactions observed in the crystal structure, we utilized NMR spectroscopy. At 50  $\mu$ M protein concentration, isotopically labeled TDP-43 NTD appears to be clearly folded in solution as indicated by the wide NMR chemical shift dispersion in the two-dimensional (2D) <sup>1</sup>H-<sup>15</sup>N Heteronuclear Single Quantum Coherence (HSQC) spectrum (Fig. 3a). Interestingly, upon increasing the protein concentration, a subset of cross peaks of TDP-43 NTD undergoes chemical shift perturbations (CSPs) and increased spectral line widths (Fig. 3a). Sequence-specific resonance assignments of TDP-43 NTD (Supplementary Fig. 5d) revealed that these perturbations are mainly corresponding to residues I18 (in  $\beta$ 2-strand), Val31, Thr32, Ala33, and Gln34 (in  $\alpha$ 1-helix), Cys39, Gly40 and Asn45 (in  $\beta$ 3-strand), Arg52 and Gly53 (in  $\beta$ 4-strand) and Val75 and



**Fig. 3** TDP-43 NTD oligomerization analyzed by solution NMR spectroscopy. **a** Overlay of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra at 30 °C of  $^{15}\text{N}$  isotopically labeled TDP-43 NTD at protein concentrations of 50, 100, 200, 400, and 800  $\mu\text{M}$  in red, green, blue, magenta, and black, respectively. Magnified views of cross peaks undergoing chemical shift perturbations (CSP) are shown around the full spectra overlay while peaks showing line broadening are encircled in green and labeled. Black arrows show the direction of movement of the peaks. **b** Backbones of amino acids undergoing CSP are colored on the cartoon of TDP-43 NTD crystal structure. Relative to the lowest protein concentration (50  $\mu\text{M}$ ), the changes in average chemical shift perturbations ( $\Delta\text{CSP} = \Delta((\delta\text{H})^2 + (\delta\text{N}/5)^2)/2)^{1/2}$ , in which  $\delta\text{H}$  and  $\delta\text{N}$  are in ppm) were calculated at highest protein concentration (800  $\mu\text{M}$ ) and colored on the structure based on the magnitude of perturbations as indicated. Side chains contributing to inter-molecular interface seen in the crystal structure are displayed as sticks with H-bonds in magenta



**Fig. 4** Point mutations at oligomerization interface abrogate TDP-43 polymerization *in vitro* and in cells. **a** Transmission electron microscopy (TEM) image of TDP-43 NTD fibrillar oligomers following negative staining. One region (*dotted black box*) from the image is shown magnified on the *right*. **b** Position of the six amino acids that were substituted to abolish oligomerization are shown on the cartoon of TDP-43 NTD crystal structure with side chains shown as sticks and labeled in *red*. **c** Overlay of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of wild type TDP-43 NTD (*blue*) and oligomerization mutant TDP-43 NTD-6M (*red*) at protein concentration of 100  $\mu\text{M}$  at 30  $^\circ\text{C}$ . **d** Overlay of 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of TDP-43 oligomerization mutant TDP-43 NTD-6M at increasing protein concentration of 100, 200, 400, and 800  $\mu\text{M}$  in *red, green, magenta and black*, respectively at 30  $^\circ\text{C}$ . **e** TEM image showing the lack of fibrillar structures for the oligomerization mutant TDP-43 NTD-6M. One region (*dotted black box*) from the image is shown magnified on the *right*. **f** Immunoblots (from 4–12% denaturing polyacrylamide gel) with anti-GFP antibody (*upper panel*) of lysates obtained from cells transiently expressing wild-type GFP-tagged human TDP-43 (GFP-TDP-43) or oligomerization mutants (GFP-TDP-43-Tm/Hm/4M/5M/6M/ $\Delta\text{NTD}$ ) followed by DSG-mediated cross-linking. Oligomerization mutants display significantly reduced high-molecular weight oligomeric TDP-43 compared to wild-type TDP-43. Actin is used as protein loading control (*lower panel*). Immunoblots are representative of three independent experiments

Asn76 (in  $\beta 5$ -strand, Fig. 3a). These are fully consistent with the inter-molecular interface seen in the crystal structure (Fig. 3b). Furthermore, Gly53, which is involved in an inter-molecular  $\beta$ -bridge, shows significant line broadening that goes beyond detection at higher protein concentrations (green circle in Fig. 3a). Interestingly, this process is reversible because upon dilution the cross-peaks become narrower and shift back to their original position showing that this oligomerization of TDP-43 NTD is dynamic in solution. The CSPs are in fast to intermediate exchange regime on the NMR time scale and the concentration dependence range implies that the affinity of NTD monomer self-association is in the low micromolar range, which

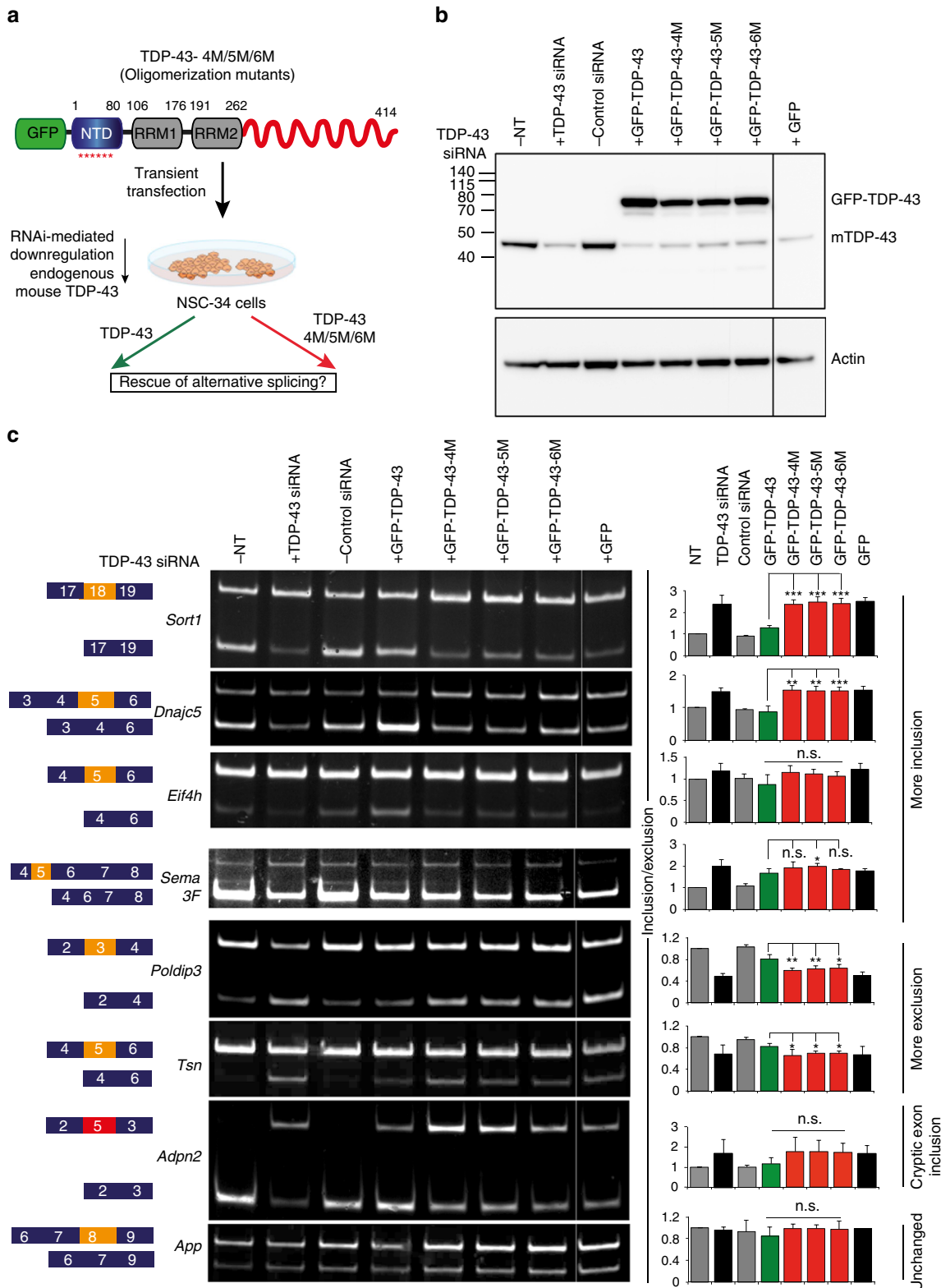
is comparable to the reported dissociation constant for DIX, Phox and Bem1 (PB1) domains (5–20  $\mu\text{M}$  for Axin-1, 200  $\mu\text{M}$  for Axin-1-Dvl-2 and 100  $\mu\text{M}$  for Par-3-PB1 domain)<sup>33–35</sup>.

#### Visualization of polymeric TDP-43 NTD by electron microscopy.

To directly visualize the TDP-43 NTD polymers under native conditions, we used transmission electron microscopy (TEM) on negatively stained samples of purified recombinant protein in the buffer conditions used for NMR experiments. TDP-43 NTD formed fibrillar structures of variable length and diameter, even at low protein concentrations of 100  $\mu\text{M}$  (Fig. 4a). Some filaments are about 7 nm in diameter, which corresponds to the width of

the helical filament seen in the crystal structure, while the observed thicker filaments may result from coalescence of individual filaments. Taken together, the head-to-tail interaction interface seen in the crystal structure and confirmed in solution by NMR and the ability of TDP-43 NTD to form fibrillar structures in native buffer may explain the intrinsic

oligomerization property of TDP-43 seen in vivo (Fig. 1a, b). The negatively charged tail (Glu14, Glu17, and Glu21) and the positively charged head (Arg52 and Arg55) interaction interface are evolutionary conserved (Supplementary Fig. 6a), suggesting a conserved mode of TDP-43 oligomerization across various species.





### Interface disrupting mutations abolish TDP-43 polymerization.

In order to investigate the functional relevance of the interaction interface observed in the crystal structure and in solution, we designed, expressed and purified recombinant proteins with specific point mutations that disrupt the oligomerization interface (Fig. 4b, Supplementary Fig. 6b–c), without affecting the folded state of the protein (Fig. 4c, Supplementary Fig. 6d). Based on the crystal structure, interference with one charged surface should be sufficient to prevent protein oligomerization *in vitro*. Indeed, both the double mutant E17A/E21A in the tail region (Tm) and R52A/R55A in the head region (Hm) and different combinations of head and tail mutants (designated 4, 5, and 6 M; Supplementary Fig. 6b) abolish oligomerization of the protein *in vitro*, as monitored by NMR spectroscopy and TEM (Fig. 4d, e, Supplementary Fig. 7a–d). Significant changes in the 2D  $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra of mutant proteins compared to wild-type NTD occurred, due to disruption of the interaction interface (Fig. 4c, Supplementary Fig. 6d). Furthermore, the narrow and uniform spectral line widths of mutant proteins compared to the wild type protein, remaining at even higher protein concentrations (800  $\mu\text{M}$ ) (Fig. 4d) strongly suggested their monomeric nature in solution. Most importantly, no changes in chemical shifts were observed upon increasing protein concentration of isotopically labeled mutant proteins (Fig. 4d, middle panel in Supplementary Fig. 7a–d). Additionally, as analyzed by electron microscopy, none of the mutant TDP-43 NTD formed any fibrils but predominantly formed amorphous deposits in the supporting matrix used in the TEM studies (Fig. 4e, right panel in Supplementary Fig. 7a–d).

To investigate the role of NTD-driven TDP-43 oligomerization in cells, we introduced interface-disrupting point mutations in the context of GFP-tagged full-length human TDP-43 in mouse motor neuron-like NSC-34 cells. To prevent oligomerization of overexpressed protein with endogenous mouse TDP-43, which is 96% identical to human TDP-43, we used point mutations in both head and the tail region (mutants designated 4, 5, and 6 M; Fig. 4b, Supplementary Fig. 6b). Oligomerization of transiently expressed protein was analyzed following DSG cross-linking in cells. The mutations severely abrogate the ability of overexpressed protein to form higher molecular weight TDP-43 oligomers with most severe disruption of oligomerization observed in TDP-43-6M, where both head and tail regions are mutated simultaneously (Fig. 4f, Supplementary Fig. 7e). These data suggest that cellular, full-length TDP-43 oligomerizes via its NTD utilizing the same protein–protein interface identified in the crystal structure and solution studies.

### TDP-43 polymerization is crucial for RNA splicing regulation.

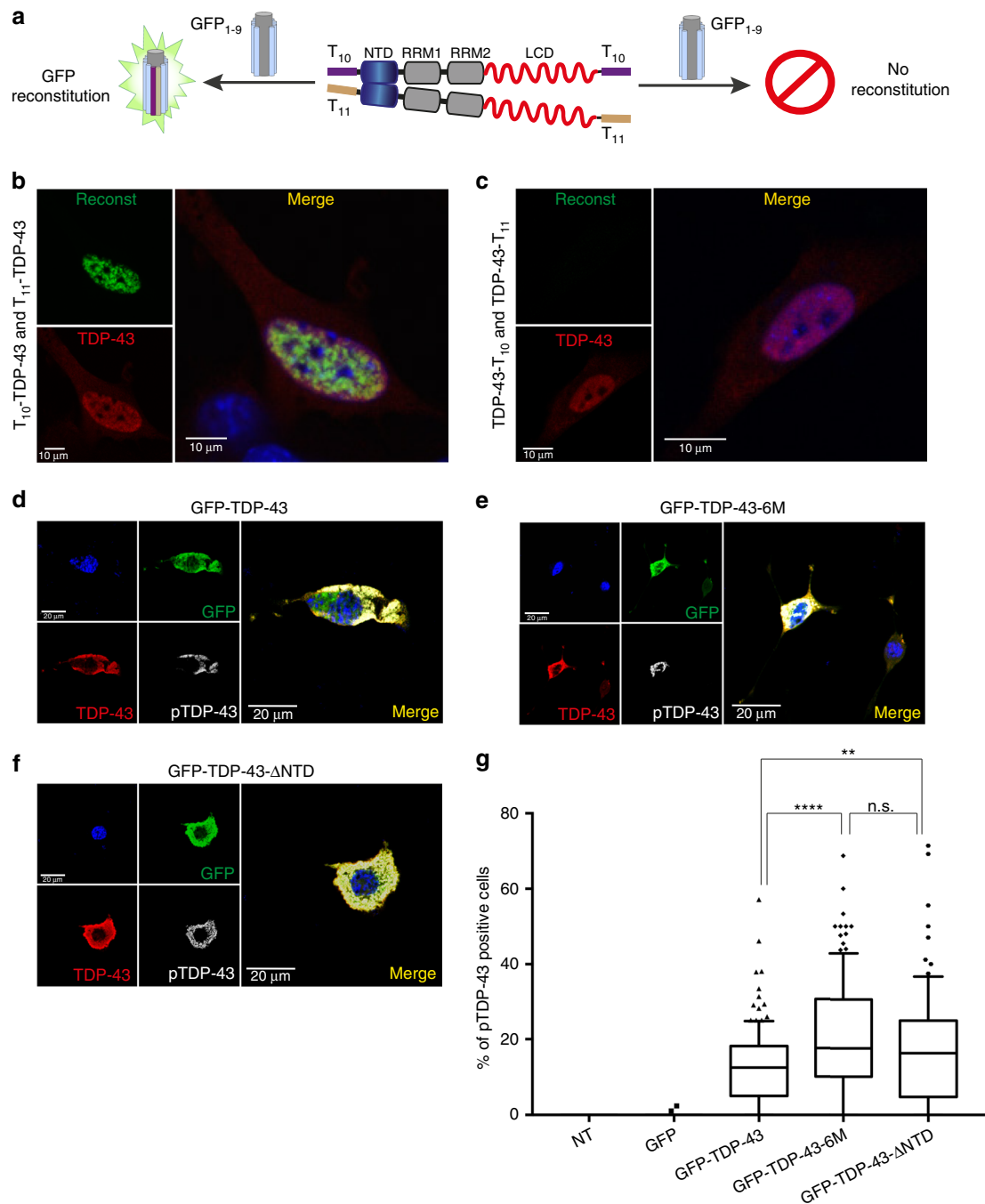
TDP-43 has been shown to regulate splicing of specific RNA targets by binding to proximal intron–exon junctions of pre-mRNA<sup>6, 36, 37</sup>. Indeed, TDP-43 downregulation in mouse brain resulted in

characteristic splicing alterations of target mRNAs<sup>6</sup>. To validate the inter-molecular interactions seen in the crystal structure and assess their functional importance, we analyzed the ability of oligomerization mutants to rescue the splicing defects induced by endogenous TDP-43 depletion in mouse NSC-34 cells (Fig. 5a). We employed a previously reported TDP-43-specific siRNA<sup>38</sup>, which reduced endogenous mouse TDP-43 levels by >90%, without interfering with expression of co-transfected RNAi-resistant wild-type or oligomerization-deficient human TDP-43 mutants (Fig. 5b). RNAi resistance was accomplished by scrambling the TDP-43 target sequence of siRNA, without altering the coded protein sequence. Transiently expressed wild-type and oligomerization-deficient TDP-43 proteins showed similar expression levels and cellular localization both in the nucleus and the cytoplasm of transfected cells (Supplementary Fig. 8a–g), allowing a comparison in their ability to rescue the RNA splicing function in the nucleus. Consistent with our previous findings<sup>6</sup>, downregulation of TDP-43 resulted in missplicing of specific RNA targets, such as *Sort1* (Sortilin1), *Dnajc5* (DnaJ homolog subfamily C member 5), *Poldip3* (polymerase delta-interacting protein 3), *Eif4h* (eukaryotic translation initiation factor 4H), *Sema3* (semaphorin-3F), and *Tsn* (translin), which was rescued by co-transfection with wild type TDP-43 (Fig. 5c). In contrast, the oligomerization-deficient mutants failed to rescue RNA missplicing, strongly indicating that head-to-tail NTD-driven TDP-43 oligomerization is essential for its function in RNA splicing. We observed a similar pattern in rescuing splicing defect in *Adnp2* that has been reported to express a cryptic exon upon TDP-43 downregulation<sup>39</sup> (Fig. 5c). As a negative control, we assessed the splicing pattern of *App*, which is a direct RNA target of TDP-43, but whose splicing is not regulated by TDP-43<sup>6</sup>. Additionally, point mutations away from the oligomerization interface could rescue the splicing defect similar to wild type TDP-43 (Supplementary Fig. 9a–c). Furthermore, consistent with *in vitro* experiments (Fig. 4f, Supplementary Fig. 7a, b), mutation of only two residues (one charged interface) is sufficient to abolish the function of TDP-43 in splicing, highlighting their role in the oligomerization-dependent function of the protein (Supplementary Fig. 9d, Supplementary Fig. 10a–h).

### Dynamic TDP-43 polymerization impedes pathologic aggregation.

Based on our structural and functional data, we hypothesized that the NTD-mediated, dynamic, physiological oligomerization might orient the LCDs of two adjoining TDP-43 molecules within the oligomer apart from each other. Such orientation would prevent inter-molecular LCD interactions and might potentially antagonize irreversible protein aggregation and phosphorylation. To experimentally test this and determine the relative spatial location of N- and C-terminal regions of adjacent TDP-43 molecules within oligomers, we used the tripartite split GFP complementation system<sup>40</sup> in mouse immortalized multipotent neural progenitor C17.2 cells (Fig. 6a). For this, we tagged either the N- or the C-termini of

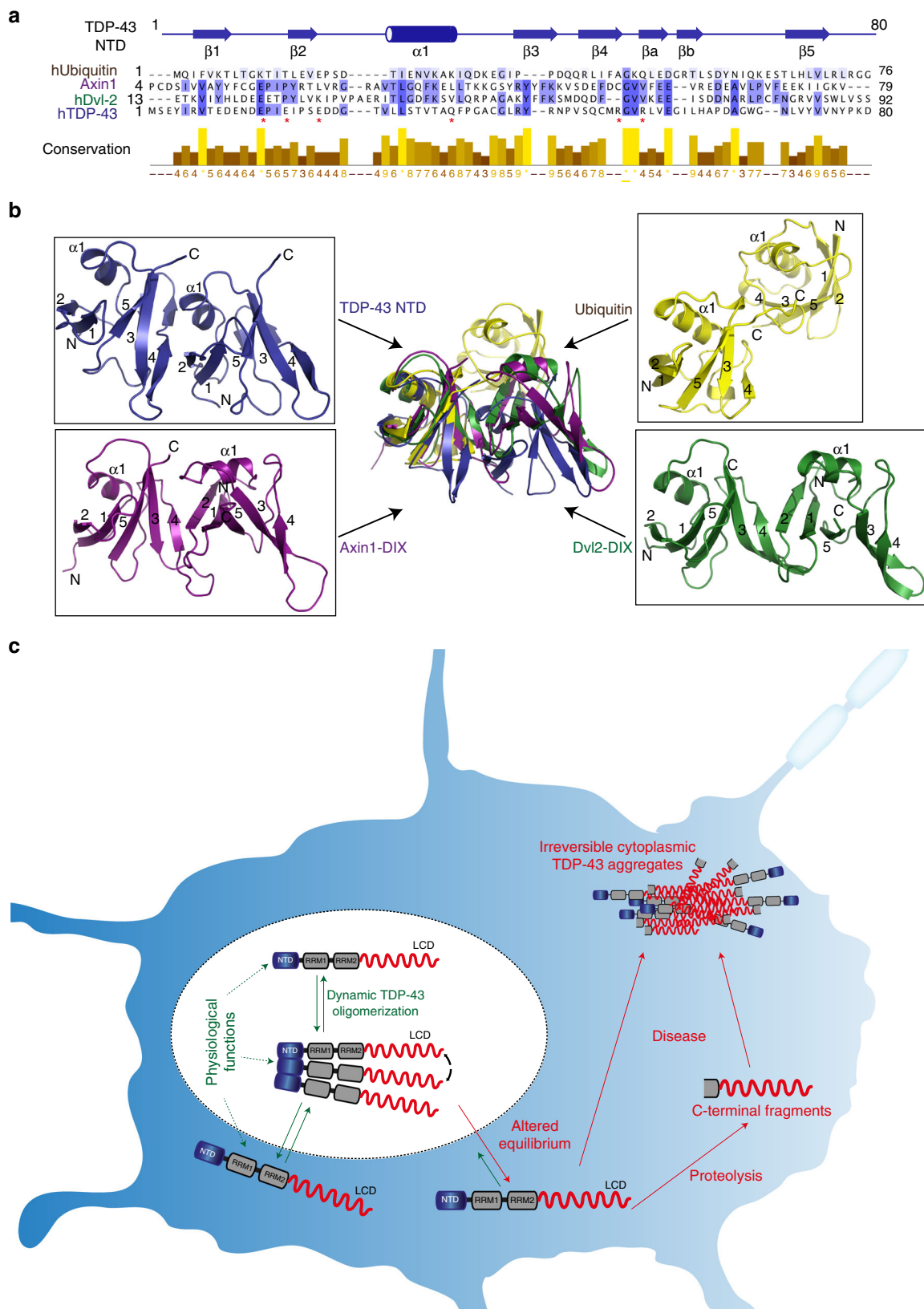
**Fig. 5** TDP-43 oligomerization is essential for its role in RNA metabolism. **a** Schematic overview of rescue experiment to assess the role of TDP-43 oligomerization in RNA splicing. Point mutations interfering with TDP-43 oligomerization are schematically marked by red asterisks below the NTD. **b** Immunoblot (from 12% denaturing polyacrylamide gel) using an antibody against mouse/human TDP-43, showing specific downregulation of endogenous mouse TDP-43 and overexpression of RNAi resistant GFP-tagged wild type and oligomerization mutant human TDP-43 (GFP-TDP-43-4M/5M/6M) in mouse NSC-34 cells (upper panel). Actin is used as protein loading control (lower panel). Immunoblots are representative of four independent experiments. **c** Semi-quantitative RT-PCR analysis of selected target's alternative splicing regulated by TDP-43. Left panel depicts alternatively spliced exons (orange) or cryptic exons (red) flanked by their constitutive exons in blue boxes. Middle panel shows representative polyacrylamide gel images of semi-quantitative RT-PCR products for all conditions. Right panel shows quantification of splicing changes plotted as the ratio of exon inclusion/exclusion (on y-axis) against different conditions (on x-axis) averaged from four biological replicates and normalized to the ratio obtained for not treated (NT) cells that is arbitrarily set to 1. In all cases, the expression of wild-type TDP-43 (green bars) rescues the changes in splicing caused by downregulation of endogenous mouse TDP-43 (TDP-43 si-RNA, black bars). The effect of each oligomerization mutant TDP-43 is shown (red bars). Statistical comparison of the inclusion/exclusion ratio between wild type and oligomerization mutants (4 M, 5 M, or 6 M) is indicated by asterisks above respective bars (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , two-tailed Student *t*-test). Standard deviation is represented by error bars. Full gels with molecular weight markers are shown in Supplementary Fig. 10



**Fig. 6** NTD-mediated TDP-43 oligomerization impedes inter-molecular LCD interactions and antagonizes pathologic protein aggregation. **a** Schematic of tripartite GFP complementation assay. TDP-43 was tagged either at the N- or C-terminus with the 10th ( $T_{10}$ ) or 11th ( $T_{11}$ )  $\beta$ -strand of GFP (schematically indicated in *magenta* and *light yellow*, respectively). Successful complementation results in fluorescence reconstitution in case of N-terminally tagged  $T_{10}$ - and  $T_{11}$ - TDP-43 **b**. In contrast, the C-terminal tagged  $T_{10}$ - and  $T_{11}$ - TDP-43 fail to reconstitute fluorescence **c**. Merged and zoomed in images are shown on the *right*. Transfected mouse C17.2 cells were counter-stained with a human-specific TDP-43 antibody (*red*) and the nuclear marker DAPI (*blue*). **d-f** Confocal microscope images of NSC-34 cells overexpressing GFP-tagged wild type **d** or oligomerization defective TDP-43 mutants **e, f**. Cells were stained with anti-GFP (*green*), anti-TDP-43 (*red*) and anti-phosphorylated TDP-43 (*gray*) antibodies. Merged images overlaid with DAPI (*blue*) are shown on the *right*. **g** The percentage of phosphorylated TDP-43 (pTDP-43)-positive cells is plotted for wild-type and oligomerization mutant TDP-43 (TDP-43-6M and TDP-43- $\Delta$ NTD). Manual cell counting was conducted on confocal images (20 $\times$  magnification). For each condition, approximately 2500 GFP-positive cells (or 1500 cells for GFP-TDP-43  $\Delta$ NTD) from three biological replicates were counted by an investigator that was blinded to the identity of the samples. Percentage of phosphorylation was determined considering the total number of GFP-positive transfected cells. An unpaired student *t*-test was performed to determine significance (depicted as *asterisks*). *P*-values: WT vs. 6 M: <0.0001 (\*\*\*\*), WT vs.  $\Delta$ NTD: 0.003 (\*\*), 6 M vs.  $\Delta$ NTD: 0.184 (n.s.)

TDP-43 with the 10th (T<sub>10</sub>) or the 11th (T<sub>11</sub>) β-strand of GFP, which when in close proximity to each other can reconstitute fluorescence of the remaining GFP<sub>1-9</sub> protein, which by itself lacks fluorescence ability<sup>40</sup> (Fig. 6a). We observed fluorescence reconstitution with GFP<sub>1-9</sub> only when T<sub>10</sub>/T<sub>11</sub> were both positioned at the N-terminus of TDP-43 (T<sub>10</sub>/T<sub>11</sub>-TDP-43, Fig. 6b), but not at the C-

terminus (TDP-43-T<sub>10</sub>/T<sub>11</sub>, Fig. 6c), or any other N- or C-terminal combinations (Supplementary Fig. 11a). These data indicate that indeed the C-termini of the adjoining molecules within TDP-43 oligomers are distantly located from each other, as opposed to the N-termini that comprise the oligomerization interface within cells. To rule out the loss of fluorescence reconstitution due to



inaccessibility of the TDP-43 C-terminal region, we co-transfected either the N- or C-terminal T<sub>10</sub>/T<sub>11</sub> TDP-43 with GFP<sub>1-10</sub>, which confirmed complementation of T<sub>11</sub>-tagged TDP-43, independently of its fusion site (Supplementary Fig. 11b–d). This suggests, that a lack of fluorescence reconstitution with GFP<sub>1-9</sub> at C-terminally tagged T<sub>10</sub>/T<sub>11</sub> is indeed due to distant spatial organization of the TDP-43 C-termini in the cellular oligomers.

Collectively, the above data confirms that nuclear and functional TDP-43 oligomerization is formed via the interaction of its NTD and that within these oligomers the C-terminal LCDs of adjoining molecules are spatially separated. Since LCD interactions are necessary for cytoplasmic TDP-43 accumulation<sup>14,29</sup>, we reasoned that NTD-mediated oligomerization may antagonize its aggregation. To test this hypothesis, we compared the ability of wild-type and oligomerization-deficient TDP-43 to be recruited into cytoplasmic phosphorylated TDP-43 inclusions. Following transient overexpression of human GFP-tagged TDP-43, a small fraction of NSC-34 cells develop cytoplasmic aggregates that become abnormally phosphorylated, reminiscent to the characteristic pathology found in sporadic ALS patient tissue<sup>2</sup>. While image analysis of over 2500 control cells—either non-transfected, or GFP-transfected—failed to detect any phosphorylated TDP-43 signal, ~15–20% of wild type TDP-43-expressing cells showed phosphorylated TDP-43 inclusions, potentially driven by overexpression and increased cytoplasmic concentration of the nucleocytoplasmic shuttling protein (Fig. 6d, Supplementary Fig. 12). Remarkably, the percentage of cells showing cytoplasmic phospho-TDP-43 aggregates is significantly and consistently higher in both oligomerization-defective mutants (TDP-43-6M and TDP-43-ΔNTD), compared to wild type (Figs. 6e–g), suggesting an increased propensity of oligomerization mutants to form cytoplasmic aggregates. To test for increased aggregation in oligomerization mutants, we performed sequential insolubility assay (SIA) exposing the cell lysates to detergents of increasing strength (Supplementary Fig. 13a). Consistent with increased pTDP-43 in oligomerization mutants, the mutant protein also shows enhanced insolubility compared to the wild type, reflecting its enhanced tendency to aggregate (Supplementary Fig. 13b, c).

In conclusion, our high-resolution structural, biochemical, and cellular analyses uncover a novel NTD-mediated dynamic and functional TDP-43 oligomerization, which antagonizes its irreversible pathologic cytoplasmic aggregation.

## Discussion

TDP-43 is a multifunctional and essential RBP, whose abnormal cytoplasmic misaccumulation is integrally linked to the pathogenesis of ALS and FTD<sup>1,2</sup>. In this study, we provide evidence that the physiological TDP-43 state in the nucleus is oligomeric and that this oligomerization is indispensable for the functional role of the protein in RNA metabolism. Dimerization or oligomerization of RBPs can occur in an RNA-dependent or independent fashion<sup>41</sup>. Our crystal structure shows that

NTD-driven TDP-43 oligomerization is independent from RNA binding, although the potential increase in local TDP-43 concentration upon RNA binding within cells may well favor oligomerization. Nevertheless, oligomeric TDP-43 would be expected to increase both its affinity and specificity for its RNA targets, compared to a monomeric form. This is in agreement with the observation from our previous CLIP-seq data in mouse brain<sup>6</sup>, showing that in most TDP-43 RNA targets, the UG-repeats are longer than required for binding of TDP-43 monomer, whose “footprint” on RNA comprises of 3–4 UG repeats<sup>6</sup>. In fact, 63% of total TDP-43-binding sites identified by CLIP-seq occur within long intronic regions and this binding sustains the levels of RNAs important for neuronal function<sup>6</sup>. An attractive hypothesis that emerges from our current study is that TDP-43 binds these regions preferentially as an oligomer to increase affinity towards long contiguous UG-repeats. Direct RNA binding of the NTD may contribute to this, as suggested by the low affinity binding of this domain to TG repeats by gel shift assay (Supplementary Fig. 14a), consistent with previous reports<sup>18</sup> showing some—albeit low—contribution of TDP-43 NTD to RNA binding. Moreover, TDP-43 oligomerization may act as a “recruitment platform” for specific factors involved in splicing and RNA maturation steps important for sustaining their levels. Another potential role of NTD-driven inter-molecular TDP-43 interaction is the ability to bring distal sites on RNA in close proximity, resulting in looping-out of RNA. Depending on the binding sites, this mechanism may promote inclusion or skipping of alternative exons. Such a mechanism has previously been proposed for other RBPs such as PTB<sup>42</sup>, hnRNP A/B, hnRNP F/H<sup>43</sup>, and STAR proteins (T-star and Sam68)<sup>44</sup>.

TDP-43 oligomerization is novel and unique among the higher order structures described for other RBPs and is mediated by NTD, whose structural homologs have not been reported in any other RBP until now. However, a sub-set of protein–protein interaction domains such as DIX, PB1, and sterile alpha motif (SAM) domains undergo similar head-to-tail interactions to generate not only homo- but also heteropolymers that have been implicated in modulating signalosome assembly<sup>24</sup>. Interestingly, despite very low sequence homology (Fig. 7a), TDP-43 NTD displays high structural similarity with both the monomer structure (Supplementary Fig. 14b), and the polymer structure of DIX domains (Fig. 7b, Supplementary Fig. 14c–e), while only resembling the monomeric structure of ubiquitin as proposed earlier<sup>15</sup>. However, while the interaction interface is comprised of similarly positioned structural elements in DIX domains, the molecular interactions are different in TDP-43 NTD (Fig. 7a). Nevertheless, based on high structural resemblance to DIX domains, it is conceivable that TDP-43 NTD might as well modulate specific signaling events by regulating the high local concentration of the protein. Interestingly, TDP-43 has been shown to interact with the signaling molecule nuclear factor (NF)-κB subunit p65 via its NTD<sup>38</sup> and it would be important to clarify if this interaction depends on TDP-43 oligomerization.

**Fig. 7** Novel TDP-43 oligomerization domain and implications for disease. **a** Multiple sequence alignment of TDP-43 NTD with structurally homologous DIX domains (Axin-1, Dvl-2) and Ubiquitin. Alignment was performed by T-coffee (<http://www.ebi.ac.uk/Tools/msa/tcoffee/>) and the alignment figure was generated in Jalview (<http://www.jalview.org/>). TDP-43 NTD secondary structural elements are depicted above the sequence and residues involved in the inter-molecular interactions within TDP-43 NTD are marked by red asterisks below the sequence. Sequence conservation scores are shown below the alignment as color gradient (dark yellow—poor conservation, light yellow—high conservation). **b** Overlay of structures of TDP-43 NTD (in blue) with Axin-1 DIX (in magenta, PDB ID 1WSP), Dvl2-DIX (in green, PDB ID 4WIP), and Ubiquitin (in yellow, PDB ID 1TBE). Structures are overlaid on one molecule to reveal the differences in the orientation of the second molecule. Individual structures in the same orientation are shown around the overlay. **c** Role of dynamic nuclear TDP-43 polymerization in physiological conditions and in disease. Within the nucleus, TDP-43 oligomerizes via the NTD and such conformation prevents irreversible aggregation by spatially separating the LCDs of adjoining TDP-43 molecules. Disturbance in the equilibrium between the oligomeric and monomeric TDP-43 in the cytoplasm (shown by green-red arrows) may result in proteolytic cleavage of exposed monomeric TDP-43 and generation of C-terminal fragments lacking the NTD. These highly aggregative fragments may initiate formation of irreversible pathologic inclusions via the LCD

The mechanism of transition of physiological nuclear TDP-43 to pathologic, irreversible and insoluble protein assemblies remains enigmatic, but it is likely triggered by the proteolytic release of its C-terminal LCD, which harbors most of the ALS-linked mutations<sup>1</sup> and possesses high propensity to self-associate<sup>11, 29</sup>, phase separate<sup>12, 13</sup>, and aggregate<sup>2</sup>. Our current data strongly support the hypothesis that formation of physiological NTD-mediated, nuclear TDP-43 oligomers can counteract cytoplasmic aggregation (Fig. 7c). Any factor altering the cellular equilibrium between oligomeric and monomeric TDP-43 may trigger the initiation of cytoplasmic TDP-43 aggregation (Fig. 7c). It is conceivable that the NTD-driven oligomeric state protects TDP-43 from proteolytic cleavage, a process characteristic for pathologic TDP-43 in ALS patients<sup>2</sup>. Such cleavage, releases TDP-43 LCD, which no longer possesses the ability to form dynamic oligomers due to lack of the essential NTD. Therefore, TDP-43 LCD can act as a potent “seed” to trigger the formation of pathologic TDP-43 aggregates, which are comprised of both full-length and C-terminal TDP-43 fragments in ALS patients<sup>2</sup> (Fig. 7c).

Finally, although the relative contribution of loss of normal nuclear TDP-43 function and gain of toxicity of its cytoplasmic aggregates remains unknown, a combinatorial mechanism may well operate in disease, which precludes many possibilities for interventions that target either loss or gain of function exclusively. We propose that stabilizing TDP-43 functional oligomers could be an attractive therapeutic strategy to counteract irreversible cytoplasmic TDP-43 aggregation, while preserving its normal nuclear function.

## Methods

**Human samples.** Frozen autopsy samples of non-neurological control human cerebellum and motor cortex were provided by Dr John Ravits from University of California, San Diego and all patient consent was taken in the clinics.

**Cross-linking in cells and tissues.** Cross-linking in cells and tissues was performed by using DSG<sup>25</sup>, which is a membrane permeable molecule that can passively diffuse inside living cells. DSG stabilizes native protein–protein interactions by forming irreversible covalent bonds between primary amine groups in the proteins. For cross-linking in cells, mouse NSC-34 and human fibroblasts were first harvested using trypsin (Gibco) and its subsequent inhibition by Trypsin inhibitor (Gibco). The cells were washed twice with phosphate-buffered saline (PBS, pH 7.4) and then resuspended in 100  $\mu$ L PBS. Fresh 20 mM DSG (Sigma) stocks were prepared in dimethylsulphoxide. The reaction was started by adding an appropriate volume of the DSG stock to the 100  $\mu$ L cell suspension (with Roche protease inhibitor cocktail in PBS) at a final DSG concentration of 50, 100, 250, 500, and 1000  $\mu$ M. The reaction was incubated at room temperature with shaking for 30 min and then quenched by addition of Tris base (at a final concentration of 20 mM) for 15 min at room temperature (RT).

For cross-linking in human and mouse brain (frozen, post-mortem human cerebellum, motor cortex, and mouse organotypic slices), tissue samples were chopped twice on a Tissue Chopper (McIlwain Tissue Chopper, setting blade distance to 100  $\mu$ m). In between each chopping, the sample was rotated by 90° to ensure fine mincing of the brain tissue. The finely minced brain samples were then transferred to 15 mL polypropylene tubes (GreinerBio-One) and resuspended in 5 mL PBS/PI by gentle pipetting (3–4 times). From this tissue suspension, homogeneous aliquots were transferred to 1.5 mL eppendorf tubes (Protein LoBind, Eppendorf), that were centrifuged at 1500xg for 5 min at RT. Supernatants were rejected, and the pellets underwent cross-linking by addition of 1000  $\mu$ L of 1 mM DSG in PBS/PI for 100 mg tissue. The samples were incubated at 37 °C for 30 min with shaking followed by quenching with 20 mM Tris base (final concentration) for 15 min.

**Nucleo-cytoplasmic fractionation.** Intact cell pellet or human and mouse tissue pellet following cross-linking were lysed in lysis buffer (50 mM Tris pH 7.4, 10 mM NaCl, 0.5% NP-40, 0.25% TritonX-100, 1 mM EDTA, protease inhibitors (Roche) by incubating 5 min on ice<sup>37</sup>. After centrifugation at 3000xg for 5 min at 4 °C, the supernatant was collected as the cytoplasmic fraction and the pellet as the nuclear fraction.

**Oxidative stress treatment on mouse brain slices.** Three-hundred fifty micrometer thick brain slices were prepared from 5–8 day old C57BL/6j pups according to previously published protocols modified for cortico-hippocampal slices<sup>45</sup>. Two slices each were cultured in Millicell-CM inserts (Millipore) in forebrain culture medium (50% BME, 25% Earle’s salt solution, 25% heat inactivated horse serum, supplemented with 1% glucose, penicillin/streptomycin,

and Glutamax (Invitrogen)). Slice cultures were kept in culture for 3 weeks (37 °C, 5% CO<sub>2</sub>, and 95% humidity) before the start of the experiment to stabilize them after the cutting procedure. To induce oxidative stress, 0.25 mM sodium (meta) arsenite (Sigma) was added to the medium for 2 h. For crosslinking experiments, slices were subsequently scraped into PBS and a homogenous cell suspension was prepared by pipetting up and down. The cross-linking procedure was performed as described above with 1 mM DSG. Anti-TDP-43 antibody (Bethyl, Supplementary Table 1) was used for immunoblotting. For immunofluorescence imaging, slices were washed in PBS, fixed in 4% paraformaldehyde, blocked for 4 h with blocking buffer (10% donkey serum (DS), 0.3% Triton-X) and stained using anti-TDP-43 antibody (Proteintech, Supplementary Table 1), and anti-TIA-1 antibody (Santa Cruz, Supplementary Table 1) for 3 days. In all, 488- and 594-Alexa Fluor-conjugated secondary antibodies raised in donkey (Life technologies) were used (2 days incubation). Nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI) and slices were mounted in fluorescent anti-fade mounting medium (Life technologies). Images were taken with an inverse confocal microscope (Leica DMI6000B, SP5) in the middle of the slice tissue.

**Protein expression and purification.** For bacterial recombinant protein expression, the DNA nucleotide sequence coding for full-length human TDP-43 (414 amino acids) was PCR amplified and cloned into the expression vector—pET-28a (+). Subsequently, the region corresponding to the NTD (amino acids 1–80) was sub-cloned into the pET28a(+) expression vector. Point mutations were incorporated by site-directed mutagenesis PCR using high fidelity Phusion DNA Polymerase (New England Biolabs) with primers introducing DNA base substitutions (Supplementary Table 3). The protein was overexpressed in *Escherichia coli* host cells BL21 codon plus RIL strain (Agilent, Catalog number #230240). Bacterial cultures were induced at OD<sub>600</sub> = 1.0 by adding 1 mM isopropyl  $\beta$ -D-thiogalactoside (IPTG, Sigma) at 25 °C for 30 h in minimal M9 medium containing <sup>15</sup>NH<sub>4</sub>Cl as nitrogen source and/or <sup>13</sup>C-D-glucose as a carbon source for labeled protein samples. For unlabeled protein samples, bacterial cultures were grown in Luria-Bertani (LB) medium (Invitrogen). Following cell harvesting, the cell pellets were resuspended in lysis buffer that comprised of 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 10 mM imidazole, 1 mM dithiothreitol (DTT), 1 mM phenylmethylsulphonyl fluoride. Cell pellets were resuspended in 20 mL of lysis buffer per liter of bacterial culture pellet. The resuspended cells were then lysed by microfluidizer. Total cell lysates were then centrifuged (Sorvall SS-34 rotor, 45,000xg, 40 min). Pellet was discarded and the supernatant containing soluble protein was used for further purification. Protein purification was achieved on ÄKTA Prime purification system (Amersham Biosciences) where the N-terminal His-tagged protein was purified with Ni<sup>2+</sup>-affinity chromatography under native conditions. The bound protein was eluted from Ni<sup>2+</sup>-columns using a linear gradient of imidazole in elution buffer (50 mM sodium phosphate, pH 8.0, 1 M NaCl, 500 mM imidazole, 1 mM DTT). The fractions containing pure protein (analyzed on SDS polyacrylamide gel electrophoresis) were pooled and dialyzed overnight in a buffer without imidazole (5 mM sodium phosphate, pH 7.4, 1 mM NaCl, 1 mM DTT). Finally, the protein was concentrated using a Vivaspin 5000 MWCO concentrator (Sartorius Stedium Biotech). Further protein purification was achieved by size exclusion chromatography using a Superdex75 column (GE Healthcare). The purified protein was aliquoted and stored at –20 °C.

**Crystallization and data collection.** To screen for initial crystallization conditions, commercially available sparse matrix screens from Hampton Research (Hampton Research, California, USA) and Molecular Dimensions (Molecular Dimensions, Suffolk, UK) were used to set up sitting-drop vapor diffusion experiments. For this, purified protein at a concentration of 10 mg mL<sup>−1</sup> was mixed in a 3:1 (protein:mother liquor) ratio, dispersed on a 96 well Intelli-Plate 96-3 LVR crystallization plate (Art Robbins) using a Phoenix crystallization robot (Art Robbins) and incubated at 4 °C against 50  $\mu$ L reservoir solution. An initial crystallization hit was found in 50 mM Tris/HCl pH 8.5, 8% (w/v) PEG 8000. These crystals were crushed and used for matrix microseeding in Crystal Screen 1&2 from Hampton Research (Hampton Research, California, USA). This led to new crystals in 0.1 M sodium cacodylate pH 6.5 and 1.4 M sodium acetate and a fine screen with differing pH and molarity of the sodium acetate was set-up, which resulted in crystals with good diffraction quality.

For data collection, crystals were incubated in mother liquor containing 30% ethylene glycol and flash-cooled in liquid nitrogen. Data were collected at beamline X06DA at the Swiss Light Source (PSI, Villigen, Switzerland) with a wavelength of 1 Å using the PILATUS 2M-F detector system (Dectris, Switzerland) and processed with XDS<sup>46</sup>. Due to the formation of weak ice-rings, data between 2.235 and 2.245 Å resolution were excluded. The outer resolution limit was estimated using the CC1/2<sup>47</sup> and  $I/\sigma(I)$  criteria<sup>48</sup>. According to these criteria diffraction data were cut at 2.1 Å resolution with CC1/2 and  $I/\sigma(I)$  of 59.2% and 1.6% in the outer resolution shell, respectively.

**Structure determination and refinement and analysis.** Using the anomalous signal from the dimethyl arsinic anion (cacodylate) linked to cysteine residues 39 and 50, phases and an initial poly-alanine model were obtained using the SHELX-C/D/E<sup>49</sup> programs in the software suite hkl2map<sup>50</sup>. An initial search in SHELX-D revealed four cacodylate ions and SHELX-E was used to build an initial poly-alanine trace with

20 cycles of density modification and a solvent content of 50%. After that, phenix autobuild<sup>51, 52</sup> was used to build the initial model. Further model building was done in COOT<sup>53</sup> and refinement was done with REFMAC<sup>54</sup> in the CCP4-suite<sup>55</sup>, phenix refine<sup>56</sup> and BUSTER (version 2.10.2, Cambridge, United Kingdom: Global Phasing Ltd). Table 1 shows the data processing and refinement parameters. Structure alignments of the two monomers in the asymmetric unit and external restraints were generated by ProSMART<sup>57</sup>, buried surface area was calculated with PISA<sup>58</sup>, and the interacting residues were drawn with Ligplot<sup>59</sup>. To analyze the superhelical parameters of the structure we used the `make_symmdef_file.pl` script from the Rosetta Symmetry framework<sup>60</sup> in combination with the  $C_\alpha$  model from the asymmetric unit. All structural figures were generated in Pymol<sup>61</sup>.

**NMR spectroscopy.** The following NMR spectrometers were used for measuring the NMR experiments in this study—Bruker AVIII-500 MHz, AVIII-600 MHz, AVIII-700 MHz, and Avance-900 MHz that were equipped with cryoprobes. TopSpin2.1/TopSpin3.0 (Bruker) was used to perform data processing while the analysis was done in Sparky (<http://www.cgl.ucsf.edu/home/sparky>). For TDP-43 NTD backbone assignments, the following NMR experiments were measured (in H<sub>2</sub>O)—two-dimensional (2D) <sup>15</sup>N-<sup>1</sup>H HSQC, 2D <sup>13</sup>C-<sup>1</sup>H HSQC, three-dimensional (3D) HNCA, 3D HNCOC, 3D CACBONH, 3D HCcH TOCSY, 3D hCCH TOCSY, 3D <sup>15</sup>N-, and <sup>13</sup>C-edited Nuclear Overhauser Spectroscopy (NOESYs).

**Electron microscopy.** Purified wild type and mutant TDP-43 NTD (100–500 μM) were applied to carbon-coated grids, stained with 2% (w/v) uranyl acetate and analyzed at various magnifications on 100 kV transmission electron microscope (Philips CM 100).

**Cell culture and transient transfections and immunostaining.** Mouse motor neuron-like hybrid cell line (NSC-34, Bioconcept, Catalog number—CLU140) were grown on cell culture dishes coated with matrigel in Dulbecco's modified Eagle's/F-12 medium (DMEM) supplemented with 10% fetal bovine serum, at 37 °C in an atmosphere of 5% CO<sub>2</sub>. Mouse multipotent neural progenitor or stem-like C17.2 cells (Sigma, Catalog number—07062902) were grown on cell culture dishes in DMEM supplemented with 10% fetal bovine serum and 1% NEAA, at 37 °C in an atmosphere of 5% CO<sub>2</sub>. Both NSC-34 and C17.2 cell lines were authenticated by the respective sources and were checked for Mycoplasma contamination in our laboratory.

For RNA splicing rescue experiments, transfections for transient over-expression of each vector (5 μg of DNA/2 × 10<sup>6</sup> cells) was obtained with Lipofectamine 2000 (Life technologies) according to the manufacturer's instruction. After 12 h incubation with transfection reagents, the medium was replaced with normal growth medium. Depletion of endogenous mouse TDP-43 was performed by RNA interference as described previously<sup>38</sup>. The sequence of TDP-43 siRNA used is 5'-AGGAAUCAGCGUGCAUUAUU-3' and were used at 20 nM concentration. The sequence of control siRNA used is 5'-GUGACAUGAGUGAGAUUU-3'.

For quantification of phosphorylated TDP-43 (pTDP-43), NSC-34 cells were plated on 24-well (ibidi) cell culture-plates at a density of ~5 × 10<sup>5</sup> cells per well leading to a confluency of approximately 80% at time of transfection. Prior to transfection, the cells were checked for confluency and morphology and the media was changed. Transfection was done with Lipofectamine 2000 transfection reagent according to the manufacturer's protocol and 0.25 μg of DNA per well was added. Cells were maintained at 37 °C under a humidified atmosphere of 5% CO<sub>2</sub> for 24 h before fixing them with 4% PFA for 10 min. For immunostaining, the cells were first permeabilized and blocked for 1 h with 10% DS and 0.2% Triton X-100 in PBS. Primary antibodies in PBS with 10% DS, 0.2% Triton X-100 were added and incubated for 2 h at RT. Cells were washed with PBS before incubating them with secondary antibodies for 1 h at RT and subsequently washed with PBS. After washing, coverslips were mounted with ProLong anti-fade medium with DAPI (Life technologies) and the plate was kept for at least 24 h in the dark before imaging. For the bipartite and tripartite GFP complementation assays, C17.2 cells were transiently transfected directly on poly-D-lysine-coated 8-well microscope slides with Lipofectamine 3000 and a total of 0.3 μg of DNA per well. Cells were fixed 24 h later with 4% PFA followed by treatment with 0.3% Triton X-100 and 5% normal goat serum in PBS. Immunostaining was performed with anti-human TDP-43 primary monoclonal antibody and secondary anti-mouse IgG Alexa594. Reconstituted GFP was directly visualized by fluorescence emission.

**RNA extraction and quantitative reverse transcription (RT)-PCR.** Total RNA from NSC-34 cells was obtained using TRIzol (Life Technologies) according to the manufacturer's instructions. Complementary DNA was generated from the total RNA (extracted from NSC-34 cells) using oligo-dT primers and Superscript III reverse transcriptase (Life technologies) as instructed by the manufacturer. Candidate TDP-43 splicing targets were tested by RT-PCR amplification using 24–27 cycles with specific primers (Supplementary Table 2) using Taq DNA polymerase (Thermo scientific). RT-PCR reaction products were analyzed on 10% polyacrylamide gels following staining with SYBR gold reagent (Life technologies). Quantification of distinct splice isoforms was done using ImageJ software (US National Institutes of Health). Finally, ratios of intensity between RT-PCR products with exon inclusion and exclusion were averaged for four biological replicates per condition.

**Sequential insolubility assay.** One 10 cm petri dish of cells transfected as described above was used for each sequential extraction of proteins with buffers of increasing stringency. Cells were harvested by scraping in solubilization buffer (SB: 10 mM Tris pH 7.4, 150 mM NaCl, 0.5 mM EDTA, 1 mM DTT, complete EDTA-free protease inhibitors (Roche)). Total protein concentrations were measured and 1 mg of total protein was used for the SIA. Samples were treated with Benzonase (250 units per sample in 2 mM MgCl<sub>2</sub>, 30 min at 4 °C (Roche)) before, and at each solubilization phase of SIA. For each step, samples were centrifuged for 30 min at 16,000 × g at 4 °C. After removing the supernatant (soluble material), pellets (insoluble material) were suspended in SB 1% v/v Nonidet P40 Substitute (NP-40, Sigma), SB 2% w/v dodecyl-β-D-maltoside (Sigma), SB 2% w/v N-Laur-oylsarcosine (Sarkosyl, Sigma) and SB 1% w/v sodium dodecyl sulfate (SDS, Sigma) sequentially. SDS insoluble material was suspended in SB 1x SDS blue loading buffer (Life Technologies).

**Data availability.** Coordinates and structure factors have been deposited in the Protein Data Bank with accession code 5MDI. The chemical shift assignments for TDP-43 NTD have been deposited in the BioMagResBank (BMRB) under the accession number 27072. All other data are available from the corresponding author upon reasonable request.

Received: 16 December 2016 Accepted: 2 May 2017

Published online: 29 June 2017

## References

- Ling, S. C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79**, 416–438 (2013).
- Neumann, M. *et al.* Ubiquitinated TDP-43 in Frontotemporal lobar degeneration and Amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
- Buratti, E. & Baralle, F. E. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J. Biol. Chem.* **276**, 36337–36343 (2001).
- Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).
- Hock, E. M. & Polymenidou, M. Prion-like propagation as a pathogenic principle in frontotemporal dementia. *J. Neurochem.* **138**(Suppl 1): 163–183 (2016).
- Polymenidou, M. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468 (2011).
- Lukavsky, P. J. *et al.* Molecular basis of UG-rich RNA recognition by the human splicing factor TDP-43. *Nat. Struct. Mol. Biol.* **20**, 1443–1449 (2013).
- Fang, Y. S. *et al.* Full-length TDP-43 forms toxic amyloid oligomers that are present in frontotemporal lobar dementia-TDP patients. *Nat. Commun.* **5**, 4824 (2014).
- Cushman, M., Johnson, B. S., King, O. D., Gitler, A. D. & Shorter, J. Prion-like disorders: blurring the divide between transmissibility and infectivity. *J. Cell Sci.* **123**, 1191–1201 (2010).
- Buratti, E. *et al.* TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. *J. Biol. Chem.* **280**, 37572–37584 (2005).
- Dewey, C. M. *et al.* TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Mol. Cell Biol.* **31**, 1098–1108 (2011).
- Kato, M. *et al.* Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
- Molliex, A. *et al.* Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
- Johnson, B. S. *et al.* TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. *J. Biol. Chem.* **284**, 20329–20339 (2009).
- Qin, H., Lim, L. Z., Wei, Y. & Song, J. 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 (2014).
- Mompean, M. *et al.* The TDP-43 N-terminal domain structure at high resolution. *FEBS J.* **283**, 1242–1260 (2016).
- Shiina, Y., Arima, K., Tabunoki, H. & Satoh, J. TDP-43 dimerizes in human cells in culture. *Cell. Mol. Neurobiol.* **30**, 641–652 (2010).
- Chang, C. K. *et al.* The N-terminus of TDP-43 promotes its oligomerization and enhances DNA binding affinity. *Biochem. Biophys. Res. Commun.* **425**, 219–224 (2012).
- Zhang, Y. J. *et al.* The dual functions of the extreme N-terminus of TDP-43 in regulating its biological activity and inclusion formation. *Hum. Mol. Genet.* **22**, 3112–3122 (2013).

- 20 Romano, V., Quadri, Z., Baralle, F. E. & Buratti, E. The structural integrity of TDP-43 N-terminus is required for efficient aggregate entrapment and consequent loss of protein function. *Prion* **9**, 1–9 (2015).
- 21 Wang, Y. T. *et al.* 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 (2013).
- 22 Sasaguri, H. *et al.* The extreme N-terminus of TDP-43 mediates the cytoplasmic aggregation of TDP-43 and associated toxicity in vivo. *Brain Res.* **10.1016/j.brainres.2016.04.069** (2016).
- 23 Suzuki, H., Shibagaki, Y., Hattori, S. & Matsuoka, M. Nuclear TDP-43 causes neuronal toxicity by escaping from the inhibitory regulation by hnRNPs. *Hum. Mol. Genet.* **24**, 1513–1527 (2015).
- 24 Bienz, M. Signalosome assembly by domains undergoing dynamic head-to-tail polymerization. *Trends Biochem. Sci.* **39**, 487–495 (2014).
- 25 Luth, E. S., Bartels, T., Dettmer, U., Kim, N. C. & Selkoe, D. J. Purification of alpha-Synuclein from human brain reveals an instability of endogenous multimers as the protein approaches purity. *Biochemistry* **54**, 279–292 (2015).
- 26 Okunola, H. L. & Krainer, A. R. Cooperative-binding and splicing-repressive properties of hnRNP A1. *Mol. Cell. Biol.* **29**, 5620–5631 (2009).
- 27 Yang, L. *et al.* Subcellular localization and RNAs determine FUS architecture in different cellular compartments. *Hum. Mol. Genet.* **24**, 5174–5183 (2015).
- 28 Strange, R. W. *et al.* Variable metallation of human superoxide dismutase: atomic resolution crystal structures of Cu-Zn, Zn-Zn and as-isolated wild-type enzymes. *J. Mol. Biol.* **356**, 1152–1162 (2006).
- 29 Colombrita, C. *et al.* TDP-43 is recruited to stress granules in conditions of oxidative insult. *J. Neurochem.* **111**, 1051–1061 (2009).
- 30 Cohen, T. J., Hwang, A. W., Unger, T., Trojanowski, J. Q. & Lee, V. M. Redox signalling directly regulates TDP-43 via cysteine oxidation and disulphide cross-linking. *EMBO J.* **31**, 1241–1252 (2012).
- 31 Zhang, T., Mullane, P. C., Periz, G. & Wang, J. TDP-43 neurotoxicity and protein aggregation modulated by heat shock factor and insulin/IGF-1 signaling. *Hum. Mol. Genet.* **20**, 1952–1965 (2011).
- 32 Bozzo, F. *et al.* Structural insights into the multi-determinant aggregation of TDP-43 in motor neuron-like cells. *Neurobiol. Dis.* **94**, 63–72 (2016).
- 33 Fiedler, M., Mendoza-Topaz, C., Rutherford, T. J., Mieszczynek, J. & Bienz, M. Dishevelled interacts with the DIX domain polymerization interface of Axin to interfere with its function in down-regulating beta-catenin. *Proc. Natl Acad. Sci. USA* **108**, 1937–1942 (2011).
- 34 Schwarz-Romond, T. *et al.* The DIX domain of dishevelled confers Wnt signaling by dynamic polymerization. *Nat. Struct. Mol. Biol.* **14**, 484–492 (2007).
- 35 Feng, W., Wu, H., Chan, L. N. & Zhang, M. The Par-3 NTD adopts a PBI-like structure required for Par-3 oligomerization and membrane localization. *EMBO J.* **26**, 2786–2796 (2007).
- 36 Buratti, E. *et al.* Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J.* **20**, 1774–1784 (2001).
- 37 Tollervey, J. R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452–458 (2011).
- 38 Swarup, V. *et al.* Deregulation of TDP-43 in amyotrophic lateral sclerosis triggers nuclear factor kappaB-mediated pathogenic pathways. *J. Exp. Med.* **208**, 2429–2447 (2011).
- 39 Humphrey, J., Emmett, W., Fratta, P., Isaacs, A. M. & Pagnon, V. Quantitative analysis of cryptic splicing associated with TDP-43 depletion. Preprint at *bioRxiv* <http://dx.doi.org/10.1101/076117> (2016).
- 40 Cabantous, S. *et al.* A new protein-protein interaction sensor based on tripartite split-GFP association. *Sci. Rep.* **3**, 2854 (2013).
- 41 Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell. Biol.* **8**, 479–490 (2007).
- 42 Oberstrass, F. C. *et al.* Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science* **309**, 2054–2057 (2005).
- 43 Martinez-Contreras, R. *et al.* Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol.* **4**, e21 (2006).
- 44 Feracci, M. *et al.* Structural basis of RNA recognition and dimerization by the STAR proteins T-STAR and Sam68. *Nat. Commun.* **7**, 10355 (2016).
- 45 Falsig, J. & Aguzzi, A. The prion organotypic slice culture assay--POSCA. *Nat. Protoc.* **3**, 555–562 (2008).
- 46 Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
- 47 Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
- 48 Karplus, P. A. & Diederichs, K. Assessing and maximizing data quality in macromolecular crystallography. *Curr. Opin. Struct. Biol.* **34**, 60–68 (2015).
- 49 Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D. Biol. Crystallogr.* **58**, 1772–1779 (2002).
- 50 Pape, T. & Schneider, T. R. HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843–844 (2004).
- 51 Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D. Biol. Crystallogr.* **64**, 61–69 (2008).
- 52 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
- 53 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
- 54 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240–255 (1997).
- 55 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242 (2011).
- 56 Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 352–367 (2012).
- 57 Nicholls, R. A., Long, F. & Murshudov, G. N. Low-resolution refinement tools in REFMAC5. *Acta Crystallogr. D. Biol. Crystallogr.* **68**, 404–417 (2012).
- 58 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
- 59 Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **51**, 2778–2786 (2011).
- 60 DiMaio, F., Leaver-Fay, A., Bradley, P., Baker, D. & Andre, I. Modeling symmetric macromolecular structures in Rosetta3. *PLoS ONE* **6**, e20450 (2011).
- 61 The PyMOL Molecular Graphics System v1.8 (Schrodinger, LLC, 2015).

## Acknowledgements

We thank Dr John Ravits (UCSD) for providing human brain tissue and Dr Shuo-Chien Ling (NUS) for providing GFP-tagged TDP-43 plasmids for mammalian cell expression. Dr Marian Hruska-Plochan for providing human fibroblasts, Dr Alexander Batyuk for help in structure determination, Sarah Erni for technical help with cell culture, Beat Blattmann and Céline Stutz-Ducommun (Protein crystallization center, University of Zürich) for help with protein crystallization screens, the staff from beamlines X06SA and X06DA from the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland) for support during X-ray data collection, Gery Barmettler for technical help in TEM, Sonu Sahadevan, Katharina Hembach, Myriam Balerna, and Julia Lüdke for helpful technical advice and discussions. This work was supported by grants from the Swiss National Science Foundation (PP00P3\_144862 to M.P. and 31003A\_166612 to P.P.), and research grants from NCCR RNA and disease to M.P. and F.H.T.A. M.P. is the recipient of a Career Development Award from the Human Frontier Science Program and further acknowledges support from the Novartis research foundation.

## Author contributions

T.A. led the study and conducted in vitro and in vivo experiments. E.-M.H. performed in situ cross-linking in mouse slices. P.E., P.M., and T.A. measured the X-ray diffraction data and performed structure determination. P.M., T.A., and A.P. directed the structural analysis of the TDP-43 NTD. L.A.B.G. and F.L. performed transient transfections, imaging and quantifications of NSC-34 cells, M.J. performed RT-PCRs, Z.M. did cloning of TDP-43, C.F. and P.P. performed GFP complementation assay, T.A. and F.H.T.A. directed the NMR and biophysical experiments and M.P. directed the whole study. T.A., P.E., F.H.T.A. and M.P. drafted the manuscript. All authors read, edited and approved the final manuscript.


## Additional information

**Supplementary Information** accompanies this paper at doi:10.1038/s41467-017-00062-0.

**Competing interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017