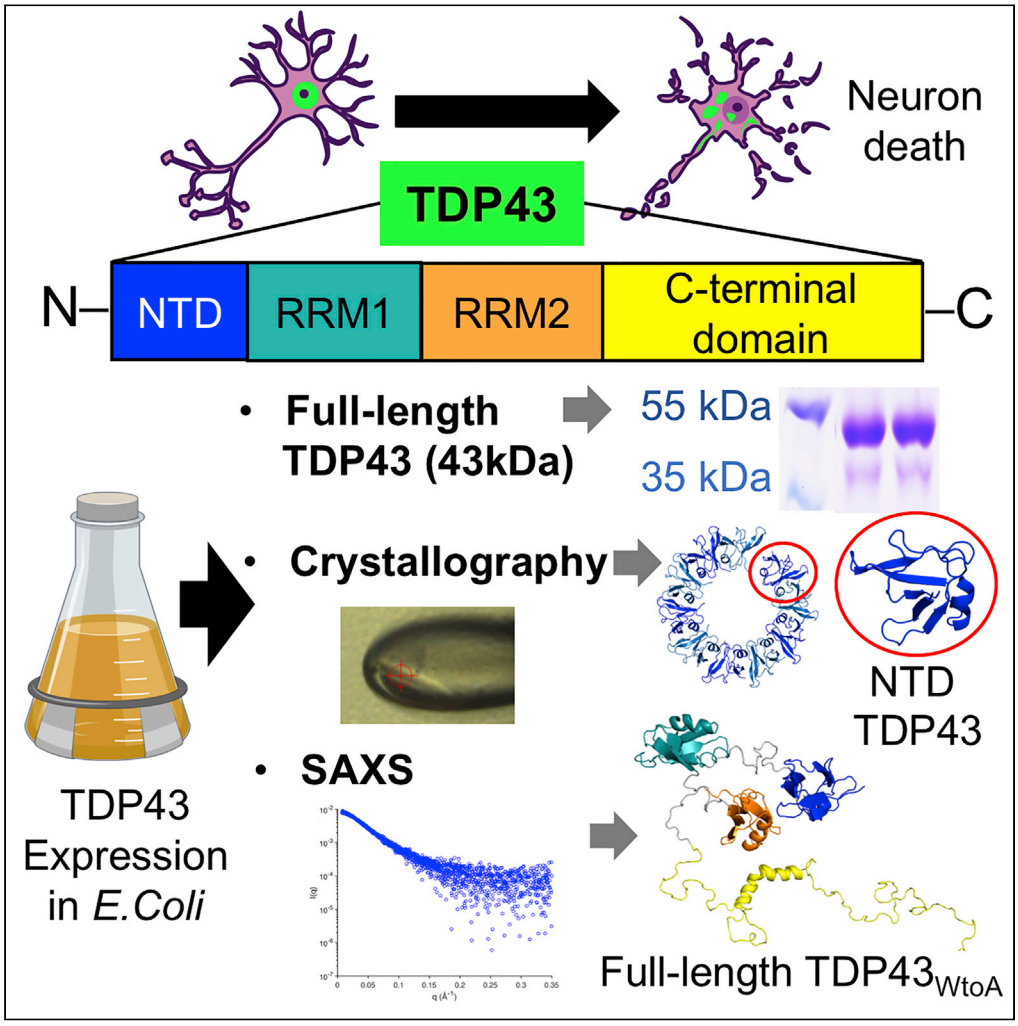
iScience

CellPress
OPEN ACCESS

Article

Purification and Structural Characterization of Aggregation-Prone Human TDP-43 Involved in Neurodegenerative Diseases



Gareth S.A. Wright, Tatiana F. Watanabe, Kangsa Amporndanai, Steven S. Plotkin, Neil R. Cashman, Svetlana V. Antonyuk, S. Samar Hasnain

s.s.hasnain@liverpool.ac.uk

HIGHLIGHTS

Purification of full-length TDP-43 with all tryptophan residues mutated to alanine

Crystallographic structure determination of the TDP-43 N-terminal domain

Small-angle X-ray scattering analysis of full-length TDP-43

Potential structural model of full-length TDP-43

Wright et al., iScience 23, 101159
June 26, 2020
https://doi.org/10.1016/j.isci.2020.101159

iScience

CellPress
OPEN ACCESS

Article

Purification and Structural Characterization of Aggregation-Prone Human TDP-43 Involved in Neurodegenerative Diseases

Gareth S.A. Wright,[1] Tatiana F. Watanabe,[1] Kangsa Amporndanai,[1] Steven S. Plotkin,[2] Neil R. Cashman,[3] Svetlana V. Antonyuk,[1] and S. Samar Hasnain[1,4,*]

SUMMARY

Mislocalization, cleavage, and aggregation of the human protein TDP-43 is found in many neurodegenerative diseases. As is the case with many other proteins that are completely or partially structurally disordered, production of full-length recombinant TDP-43 in the quantities necessary for structural characterization has proved difficult. We show that the full-length TDP-43 protein and two truncated N-terminal constructs 1-270 and 1-263 can be heterologously expressed in *E. coli*. Full-length TDP-43 could be prevented from aggregation during purification using a detergent. Crystals grown from an N-terminal construct (1-270) revealed only the N-terminal domain (residues 1-80) with molecules arranged as parallel spirals with neighboring molecules arranged in head-to-tail fashion. To obtain detergent-free, full-length TDP-43 we mutated all six tryptophan residues to alanine. This provided sufficient soluble protein to collect small-angle X-ray scattering data. Refining relative positions of individual domains and intrinsically disordered regions against this data yielded a model of full-length TDP-43.

INTRODUCTION

The deposition of intracellular TDP-43 inclusions is the hallmark of TDP-43 pathology. Initially observed in neural tissues from individuals with frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS) (Neumann et al., 2006), TDP-43 pathology is now associated with many neurodegenerative diseases. These include, but are by no means limited to, Alzheimer disease (Amador-Ortiz et al., 2007), Parkinson disease (Nakashima-Yasuda et al., 2007), hippocampal sclerosis (Amador-Ortiz et al., 2007), and chronic traumatic encephalopathy (McKee et al., 2010). Mixed pathology is common with these diseases, and TDP-43 pathology can be found not only alongside Lewy bodies (Nakashima-Yasuda et al., 2007), amyloid-β plaques, and tau tangles (Amador-Ortiz et al., 2007) in cases of neurodegenerative disease but also in clinically normal aged individuals (Wennberg et al., 2019).

TDP-43 protein has several functions, and its modular structure facilitates this multitasking. Two centrally located RNA recognition motifs (RRM) strongly bind UG-rich RNA (Lukavsky et al., 2013) or TG-rich DNA (Austin et al., 2014) directing TDP-43 to pre-mRNAs and intronic sites (Tollervey et al., 2011). Through these protein-nucleic acid interactions TDP-43 facilitates RNA transport (Fallini et al., 2012) and directly effects splicing of a multitude of RNAs including those coding for many ALS-associated heterogeneous nuclear ribonucleoprotein particles (Deshaies et al., 2018) and TDP-43 itself (Ayala et al., 2011). A low-complexity domain, situated C terminal to the RRM domains, is involved in stress granule formation following cellular stress (Colombrita et al., 2009). This domain undergoes liquid-liquid phase transitions and complexes with other TDP-43 molecules (Li et al., 2018a) or other intrinsically disordered proteins (McDonald et al., 2011), whereas RRM domains trap mRNAs to assist selective translation during and following stress. Many single amino acid substitutions within the C-terminal domain are known to cause ALS and FTLD, indicating that aberrant stress granule dynamics may lie at the heart of TDP-43 proteotoxicity (Wolozin, 2012). At the N terminus, a ubiquitin or dix-like domain (Mompeán et al., 2016; Qin et al., 2014) provides a polymerization surface that enables formation of dimer and higher-order oligomers predominantly found in the cell nucleus (Afroz et al., 2017).

[1]Molecular Biophysics Group, Department of Biochemistry & Systems Biology, Institute of Systems, Molecular and Integrative Biology, Faculty of Health and Life Sciences, Liverpool L69 7ZB, UK

[2]Department of Physics & Astronomy, The University of British Columbia, Vancouver, BC, Canada

[3]Djavad Mowafaghian Centre for Brain Health, University of British Columbia, Vancouver, BC V6T 2B5, Canada

[4]Lead Contact

*Correspondence: s.s.hasnain@liverpool.ac.uk

https://doi.org/10.1016/j.isci.2020.101159

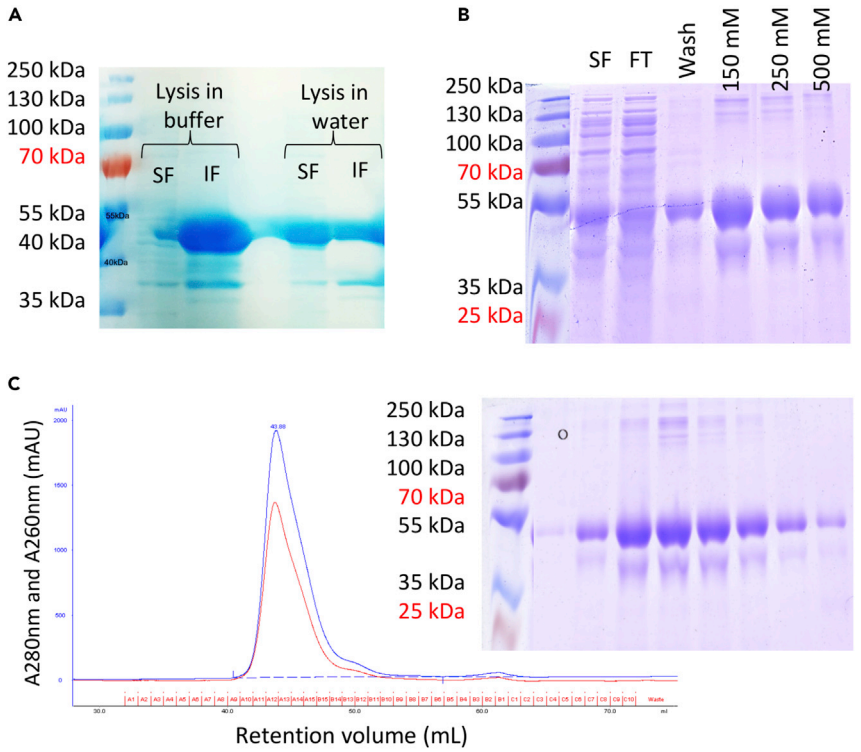


Figure 1. Purification of Full-Length Wild-Type TDP-43 Using 0.2% Sarkosyl

(A) Cell lysis in 50 mM sodium phosphate pH 8.0, 300 mM sodium chloride, 5 mM imidazole, 5 mM dithiothreitol, complete protease inhibitor cocktail, 1 mM phenylmethylsulphonyl fluoride, 50 μg/mL lysozyme, and pure water. SF, soluble fraction; IF, insoluble fraction.

(B) Immobilized metal ion chromatography. SF, soluble fraction; FT, flowthrough; wash fraction (75 mM imidazole), elution with increasing imidazole concentration.

(C) Size exclusion chromatography (SEC) and SDS-PAGE of SEC fractions containing pure wild-type TDP-43.

In addition to the domains described earlier, TDP-43 also contains nuclear export and import sequences that flank the RRM domains. In the full-length form, this enables shuttling of the protein between the nucleus and cytoplasm (Ayala et al., 2008). Observation of TDP-43 inclusions in FTLD and ALS brain tissues has shown that it is fragmented (Neumann et al., 2006) with a 25-kDa cleavage product being the most common but 15- and 35-kDa forms are also seen. The common element present in these cleavage products is the low-complexity C-terminal domain. In this state, the nuclear localization signal is lost and the protein remains cytoplasmic and aggregates through the low-complexity C-terminal domain. The normal functions and misfunction of TDP-43 are therefore predicated by its propensity to oligomerize and aggregate. This property makes *in vitro* characterization difficult, particularly using structural techniques that require monodisperse samples at relatively high concentrations. The modularity of TDP-43 means that individual domains can be produced recombinantly to shed light on their organization and structure-property relationship. This approach has been well utilized to gain insight on dimerization (Afroz et al., 2017), RNA binding (Lukavsky et al., 2013), and thermal stability (Austin et al., 2014; Chiang et al., 2016). However, to go beyond individual domains and gain a holistic understanding of TDP-43 structure we tested several strategies to produce pure, full-length TDP-43. Using sarkosyl detergent and mutagenically removing tryptophan resides (TDP-43_{WtoA}), we were able prevent recombinant TDP-43 aggregation during protein preparation in each of these cases. We produced sufficient quantities of detergent-free TDP-43_{WtoA} for small-angle X-ray scattering (SAXS) analysis allowing us to create a model of full-length TDP-43.

RESULTS

Recombinant Expression and Purification of Full-Length and C-Terminal Truncated TDP-43

On heterologous expression of full-length wild-type TDP-43 in *E. coli* followed by cell lysis and centrifugation, we found the protein in the insoluble fraction as has previously been described (Furukawa et al., 2011).

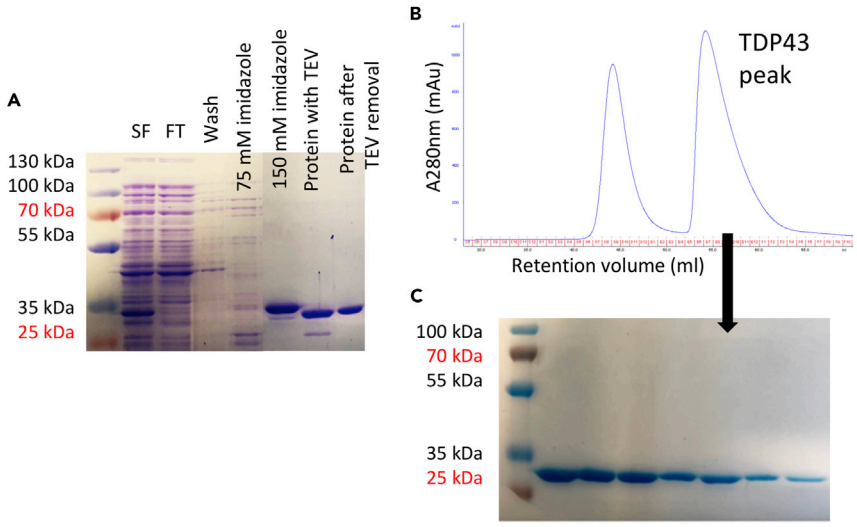


Figure 2. Purification of Wild-Type TDP-43 Residues 1-270

(A) Nickel-nitriloacetic acid (NiNTA)-immobilized metal ion chromatography (IMAC). SF, soluble fraction; FT, flowthrough; wash (5 mM imidazole); 75 mM/150 mM elution imidazole concentration.

(B) Size exclusion chromatography.

(C) Fractions from (B) containing pure wild-type TDP-43 residues 1-270.

However, low-temperature expression followed by cell lysis in pure water yielded segregation of the protein roughly equally between soluble and insoluble fractions but made affinity purification problematic. This is possibly an indication that recombinant TDP-43 is not localized to inclusion bodies when expressed in bacteria. To solve this problem, 0.2% sodium lauroyl sarcosinate (sarkosyl) was added to the cell lysate and used in all purification buffers. This yielded 5.3 mg of pure, soluble protein for a litre of culture (Figure 1). This detergent-solubilized protein is stable without visible aggregation after a week at 4°C or a freeze/thaw cycle.

Several constructs of C-terminal truncated TDP-43 (N-terminal 1-263, 1-270, 1-290, 1-320, and several single point mutations of 1-270 fragment including A90V, D169G, K263E, and N267S) were prepared and evaluated. The N-terminal 1-270 wild-type construct provided stable purified protein (Figure 2) and was pursued for further investigations. This was used for crystallization experiments from which structure of N-terminal domain (NTD) (1-80) was obtained.

Addition of soluble fusion domains is a common technique used to enable production of recombinant proteins. Wang et al. (2018) added a maltose-binding protein fusion to TDP-43 directly C terminal to the low-complexity domain where it effectively increased solubility. To obtain detergent-free TDP-43 without recourse to fusions proteins, and with the knowledge that several TDP-43 tryptophan residues are involved in folding (Prakash et al., 2018) and phase transitioning (Li et al., 2018a, 2018b), together with the role of aromatic residues in low-complexity aromatic-rich kinked segment formation (Hughes et al., 2018), we created a construct where all tryptophan residues in the TDP-43 primary sequence are replaced with alanine (TDP-43_{WtoA}). Cell lysis and protein purification with standard immobilized metal ion chromatography buffers without any detergent facilitated production of full-length TDP-43_{WtoA} that was stable for 3 days at 4°C (Figure 3), sufficient to undertake non-crystallographic structural studies.

Crystallization of the N-Terminal TDP-43 Dimerization Domain

To elucidate the molecular basis of TDP-43 pathogenesis, a stable N-terminal 1-270 fragment was used for crystallization and formed crystals when grown at 19°C. The structure provided a surprise as only the NTD (1-80) was visible. Analysis of several crystallization drops by SDS-PAGE indicated the 30-kDa original protein to have shifted to approximately 23 kDa, which would indicate cleavage of RRM2. During 1 week of crystallization the protein appears to have auto-cleaved. The TDP-43 NTD crystallized in space group $P2_12_12_1$ with five identical molecules in the asymmetric unit. The structure was solved by molecular replacement and refined to 2.55 Å resolution (Table 1). The NTD adopts a similar conformation as reported in a

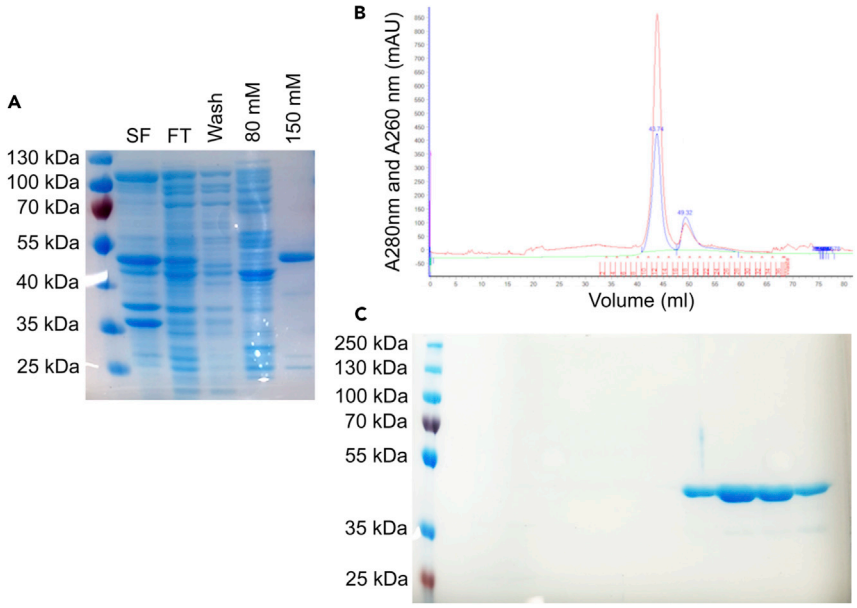


Figure 3. Purification of Full-Length TDP-43$_{WtoA}$

(A) NiNTA IMAC. SF, soluble fraction; FT, flowthrough; wash (5 mM imidazole); 80 mM/150 mM elution imidazole concentration.

(B) Size exclusion chromatography (SEC).

(C) SEC fractions containing pure TDP-43$_{WtoA}$.

recent crystallographic structure obtained in P6$_3$ space group (Afroz et al., 2017). A compact five-stranded β-sheet fold was observed with a two-turn alpha helix (Figures 4A–4C), and the negatively charged tail of each domain is packed against the positively charged C-terminal part (head-to-tail interaction).

The head-to-tail interfaces between the molecules differ from those reported previously (Afroz et al., 2017). A sulfate ion binds at the interface between the molecules and is bonded to Arg52 and Arg55, which serves as a wedge, increasing the distance between the monomers. This binding of a sulfate ion to Arg52 and Arg55 (Figure 4D) rearranges their side chains preventing strong bonding between Arg52-Glu3 and Arg55-Glu21 reported in an earlier NTD structure (Afroz et al., 2017). These change the position of the N terminus toward the interface (Figures 4D and 4E), moving the Glu3 side chain away from Arg52 and making strong salt bonding with Arg52 impossible. The changes in the position of the monomers against each other alter the superhelical bundle arrangement observed previously (Afroz et al., 2017) (Figure S1). In our structure, bundles of the tight spirals are packed against six neighboring spirals (Figure S1C). Each spiral, with radius of 46 Å, has 10 NTD molecules in a full turn (Figures 4A, 4B, S1A, and S1B) with distance between two rings being ~36 Å. Analysis of the structural changes on sulfate binding confirms that ligand binding at the molecular interface could change the nature of the helical arrangements. These head-to-tail interactions that lead to parallel spirals are consistent with the view that physiological TDP-43 oligomerization is mediated by its NTD and may be key to prevent the formation of pathologic aggregates.

Structural Characterization of Full-Length Tryptophan-Free TDP-43

Solvent-accessible tryptophan residues within intrinsically disordered regions are unlikely to contribute to individual domain structures. It is, however, important to understand the perturbations caused by mutation of tryptophans within the NTD and RRM1 domain. Available RRM1 structures (Kuo et al., 2014; Lukavsky et al., 2013) show Trp178 has a role in nucleic acid binding but in the apo form both Trp113 and Trp172 side-chains protrude into solvent. Comparing tryptophan solvent accessible surface area in various proteins taken from a non-redundant database of 27,015 protein structures (He et al., 2014) indicates the RRM1 tryptophans to be unusually exposed (Figure 5). Thus, mutation of these residues is unlikely to change the overall RRM1 structure. The side chain of Trp68, found in NTD loop 5 adjacent to the domain core, forms hydrophobic interactions with other residues within the core (Figure S2) (Mompeán et al., 2016) and may support a homodimer interface (Afroz et al., 2017). Indeed, reducing

	TDP43
Data Collection	
Space group	$P2_1\ 2_1\ 2_1$
Cell dimensions	
a, b, c (Å)	34.637, 95.224, 157.558
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)[a]	78.90–2.55 (2.62–2.55)
R_{merge}[a]	15.6 (0.986)
I/σI[a]	5.7(1.6)
$CC_{1/2}$ (%)[a]	0.985(0.549)
Completeness (%)[a]	98.7(99.1)
Redundancy[a]	3.9(4.0)
Wilson B ($Å^2$)	38.3
Refinement	
No. reflections	16,549
R_{work}/R_{free}	21.46/25.86
No. of atoms	
Protein	3,040
Ligand/ion	25
Water	235
B-factors	
Protein	50.78
Ligand/ion	94.12
Water	43.33
RMS deviations	
Bond lengths (Å)	0.0040
Bond angles (°)	1.256
PDB	6T4B

Table 1. Crystallographic Data Collection and Refinement Statistics

[a]values in paranthesis are for the outer shell of data

homodimerization propensity may have aided our recovery of soluble and relatively stable full-length protein. Trp68 retains some solvent exposure (Figure 5), is surrounded by loop regions, whose mutation is thus unlikely to affect secondary structural elements that contribute to overall domain tertiary structure.

Using the expression protocol described above, and elaborated upon in the Transparent Methods section, we were able to produce sufficient detergent-free recombinant TDP-43_{WtoA} to perform chromatographic SAXS experiments. Size exclusion chromatograms of full-length TDP-43 (Figure S3) are slightly asymmetric. This could be interpreted as a fast oligomeric equilibrium, between monomer and dimer, for example, or an interaction with column media, which delays elution of some molecules. Sampling data points in the first and last thirds of a single TDP-43_{WtoA} elution indicates a decrease in mean radius of gyration (Rg) from 41.8 ± 3.5 Å to 38.9 ± 4.0 Å (Figure S3). Owing to this small change in Rg, and lack of distinct sub-populations within each profile, we averaged the data across two full TDP-43_{WtoA} elutions. This yielded good-quality data over an angular range of 0.0084–0.35 $Å^{-1}$ (Figure 6A). Inspection of the

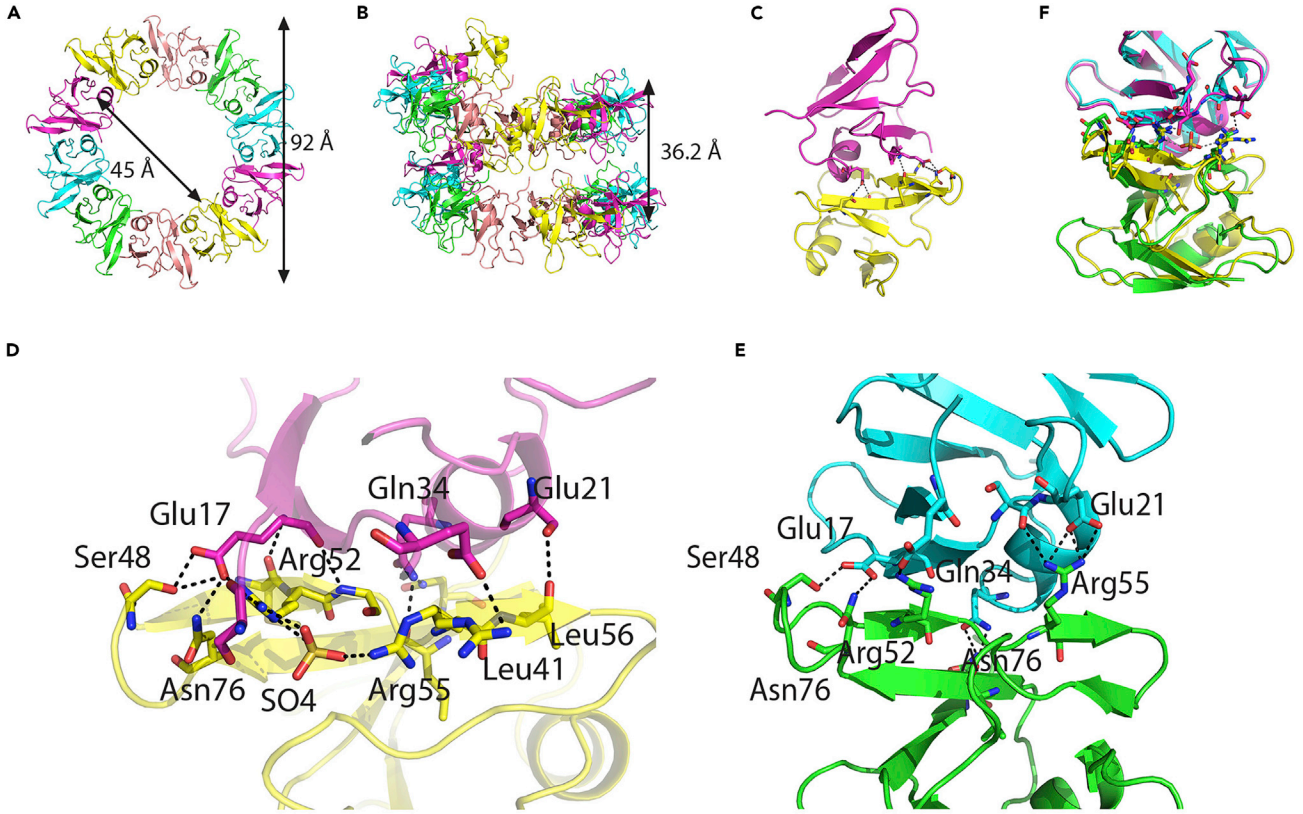


Figure 4. Crystal Structure of TDP-43 NTD at 2.55 Å Resolution

Molecules of TDP-43 NTD are arranged in the crystal as parallel spirals.

(A) Filaments from neighboring molecules arranged in head-to-tail fashion. Atoms of 10 molecules comprising two neighboring asymmetric units make one full turn of the spiral and are shown in different colors. The outside radius of the spiral is ~92 Å, whereas internal radius is 45 Å.

(B) Side view of the spiral, showing two full turns with distance between two turns of the spiral 36.2 Å.

(C) Cartoon representation of two neighboring TDP-43 NTD molecules.

(D) Expanded view of the interface between molecules shown in (C). Amino acid side chains making intermolecular contacts are shown in stick representation in the corresponding domain color and labeled. Intermolecular hydrogen bonds are shown as dotted black lines.

(E) The dimer interface described by Afroz et al. (2017) (PDB: 5MDI) and labeled as in (C).

(F) Superposition between dimer interfaces shown in (D) and (E). Structures were aligned to one molecule rather than the whole dimer.

Guinier plot (Figure 6B) indicates that TDP-43$_{WtoA}$ is not aggregated and has Rg 41.1 Å. Rendering the experimental data as a dimensionless Kratky plot shows TDP-43$_{WtoA}$ to have a high degree of unfolding but to not be completely disordered (Figure 6C). This is expected given the presence of previously characterized modular nucleic acid binding and dimerization domains (Afroz et al., 2017; Austin et al., 2014; Lukavsky et al., 2013). Assignment of largest intramolecular distances (Dmax) is difficult for disordered proteins, and this is the case for TDP-43$_{WtoA}$, with possible Dmax values ranging from 132 to 180 Å (Figure 6D).

Using available structures for RRMs, dimerization, and helical domains in conjunction with linker peptides synthesized *in silico* we generated initial models of full-length TDP-43$_{WtoA}$ (Figure S4) and refined them against our experimental SAXS data to yield a conformationally optimized model (Figure 7A). The monomeric model generated has little contact between domains except between the NTD and RRM2 (Figure 7B). Calculation of its scattering profile indicates a very good fit to the experimental data χ^2 1.1 (Figure 7C). The experimentally determined Rg agrees with the model within 0.8 Å (41.1 and 41.9 Å, respectively). Comparison of the distance distribution functions of our monomeric, full-length TDP-43$_{WtoA}$ model with that derived from experimental data also indicates exceptionally good correlation (Figure 7D). Models of TDP-43$_{WtoA}$ where dimerization was enforced through the NTD had consistently poor fit to the experimental data, χ^2 of 1.5 and Rg 49.7 Å.

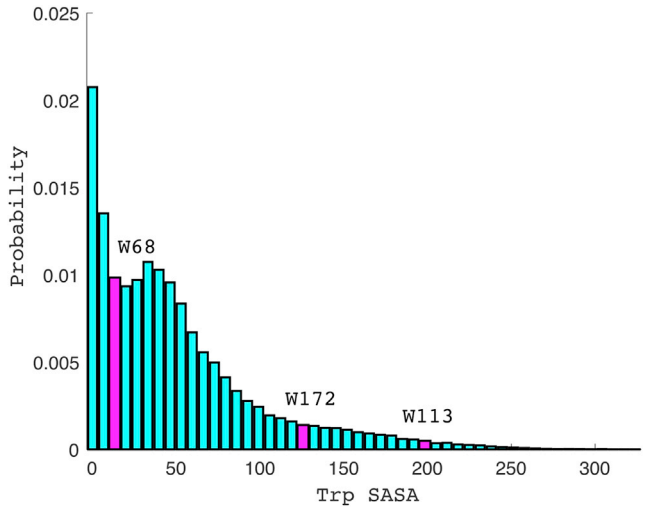


Figure 5. Distribution of Solvent Accessible Surface Area (SASA) of Tryptophans in a Non-redundant Database of Protein Structures

SASA for three tryptophan residues found in structured domains of TDP43 is shown in magenta: W68 in the NTD and W113 and W172 in RRM1. Their cumulative percentile of exposure is 23.6%, 98.5%, and 91.8% respectively.

DISCUSSION

Although the structures of individual domains of TDP-43 have been elucidated using several techniques that have revealed some aspects of the molecular functions of this protein (Afroz et al., 2017; Chiang et al., 2016; Guenther et al., 2018; Kuo et al., 2014; Mompeán et al., 2016), the structure of full-length TDP-43 has been refractory to characterization due to the difficulty of purifying soluble and stable protein in sufficient amounts for analysis (Johnson et al., 2009; Kitamura et al., 2018; Li et al., 2017). A recent report overcame this challenge using denaturing conditions, but this could change the native TDP-43 structure (Vivoli Vega et al., 2019). In this work, we describe full-length wild-type TDP-43 and TDP-43$_{WtoA}$, both successfully purified by non-denaturing methods. SAXS data of TDP-43$_{WtoA}$ reveal the conformation of the full-length protein in solution. The region of TDP-43 comprising the NTD, RRM1, and RRM2 adopts a compact triangular structure, whereas the position of the C terminus is variable.

A fragment of TDP-43 comprising residues 1-270 including NTD and RRMs was used in crystallization experiments and the structure of the NTD (residues 1-80) was elucidated at 2.55 Å resolution. The protein appears to have been auto-cleaved at N terminal to RRM2 during the 7 days required for crystallization to produce diffraction-quality crystals, whereas RRM1 is conformationally mobile *in crystallo* and therefore not visible. The TDP43 NTD has been reported to be essential for dynamic TDP-43 oligomerization that may prevent the aggregation-prone C terminal from forming pathogenic and irretrievable TDP-43 aggregates. This crystal structure shows that the superhelical format arises from head-to-tail interactions between NTD molecules. The SAXS model implies that the peptide chain linking NTD and RRM1 domains is longer than other domains, so it is likely to be disordered when forming higher-order structures.

The ability to obtain full-length TDP-43 in a stable form without denaturing conditions opens possibilities for extensive biophysical studies on both the wild-type and C-terminal mutants, which are known to exert greater aggregation propensity (Cao et al., 2019). We propose that the approaches used here may have general applicability and may prove useful for other aggregation-prone proteins where a significant proportion of the macromolecule is classified as "unstructured," i.e., having a lower folding probability. Enabling the purification of stable complexes via mutation of residues that cause precipitation or aggregation has recently proven key to our description of the functional complexes between SOD1 and its cognate chaperone (Sala et al., 2019). Similarly, use of detergents for soluble proteins should enable high-concentration structural studies for such systems, providing an essential platform for molecular understanding of pathogenic properties and possible therapeutic solutions.

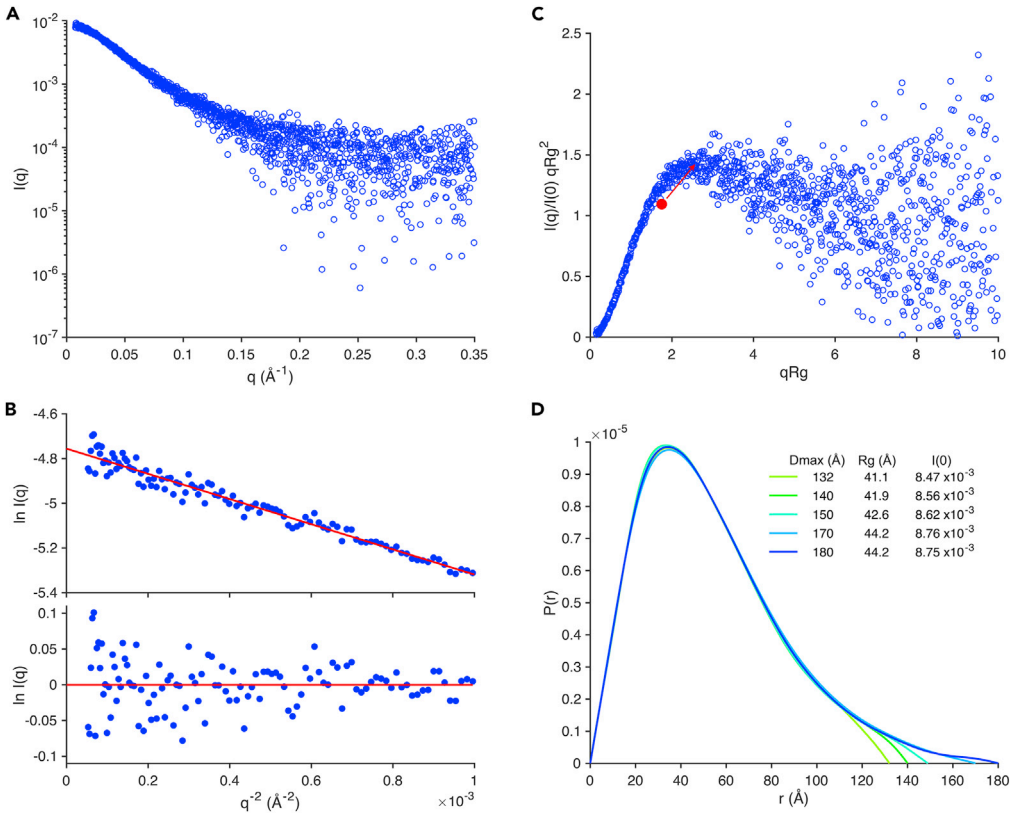


Figure 6. Full-Length TDP-43 X-Ray Scattering

(A) Intensity plot of X-ray scattering by TDP-43$_{WtoA}$.

(B) TDP-43$_{WtoA}$ Guinier plot (upper) and residuals (lower). R^2 0.974 over data range $0.3 < q.Rg < 1.29$. Rg = 41.1 ± 0.4 Å (with error stated as SEM) as determined by $m = -Rg^2/3$, where m is the gradient of the line, and $I(0) = 8.60 \times 10^{-3}$. Both plots indicate a monodisperse sample with little or no interparticle interference.

(C) Guinier-based dimensionless Kratky plot showing that the peak maximum for TDP-43$_{WtoA}$ is moved away from the point at √3 and 1.1 (highlighted in red) where globular proteins show a maximum. This is indicative of unstructured regions. The overall form of the curve, which does not return to the baseline after the initial peak but does not continue to increase with qRg, is also indicative of a protein with both folded and unfolded regions.

(D) SAXS distance distribution functions (P(r)) for TDP-43$_{WtoA}$. Multiple possible functions are plausible with a Dmax range 132–180 Å. P(r) functions with Dmax 132–150 Å have real space Rg and I(0) that correlate well with those from the Guinier approximation, whereas those with Dmax 170–180 Å show smooth transitions with the r scale. Dmax, Rg, and I(0) determined from each P(r) function are stated in the legend.

Limitations of the Study

A population of intrinsically disordered sequences are, by their nature, very unlikely to simultaneously occupy the same conformation. Although the model presented in Figure 7A is our best representation of the TDP-43 structure, it is clearly a snapshot of a very dynamic system. The model presented in Figure 7A has an unrestrained and predominantly disordered C-terminal region but, as noted, a relatively compact N-terminal region comprising the NTD, RRM1, and RRM2. Disordered parts of the protein contribute less to scattering intensity than globular domains and are poorly defined by scattering data. To determine how variable the spatial arrangement presented in Figure 7A may be, we performed long molecular dynamics simulations using our optimized model as a starting structure. Figure 8 shows that positioning of the disordered C terminus has little effect on the fit to experimental data. However, those structures that fit the data poorly exhibit an increase in the Rg value for the whole molecule and the Rg representing NTD, RRM1, and RRM2 domains (amino acids 1-258) (Figures S5A and S5B and Table S1). For these models, inter-domain linkers are also found in an extended conformation (Figures S5C–S5E and Table S1) and separation between domains is therefore maximized. Conversely, for models that fit the data well, the globular domains occupy a compact conformation, which matches that presented in Figure 7A (Figures S5A and S5B

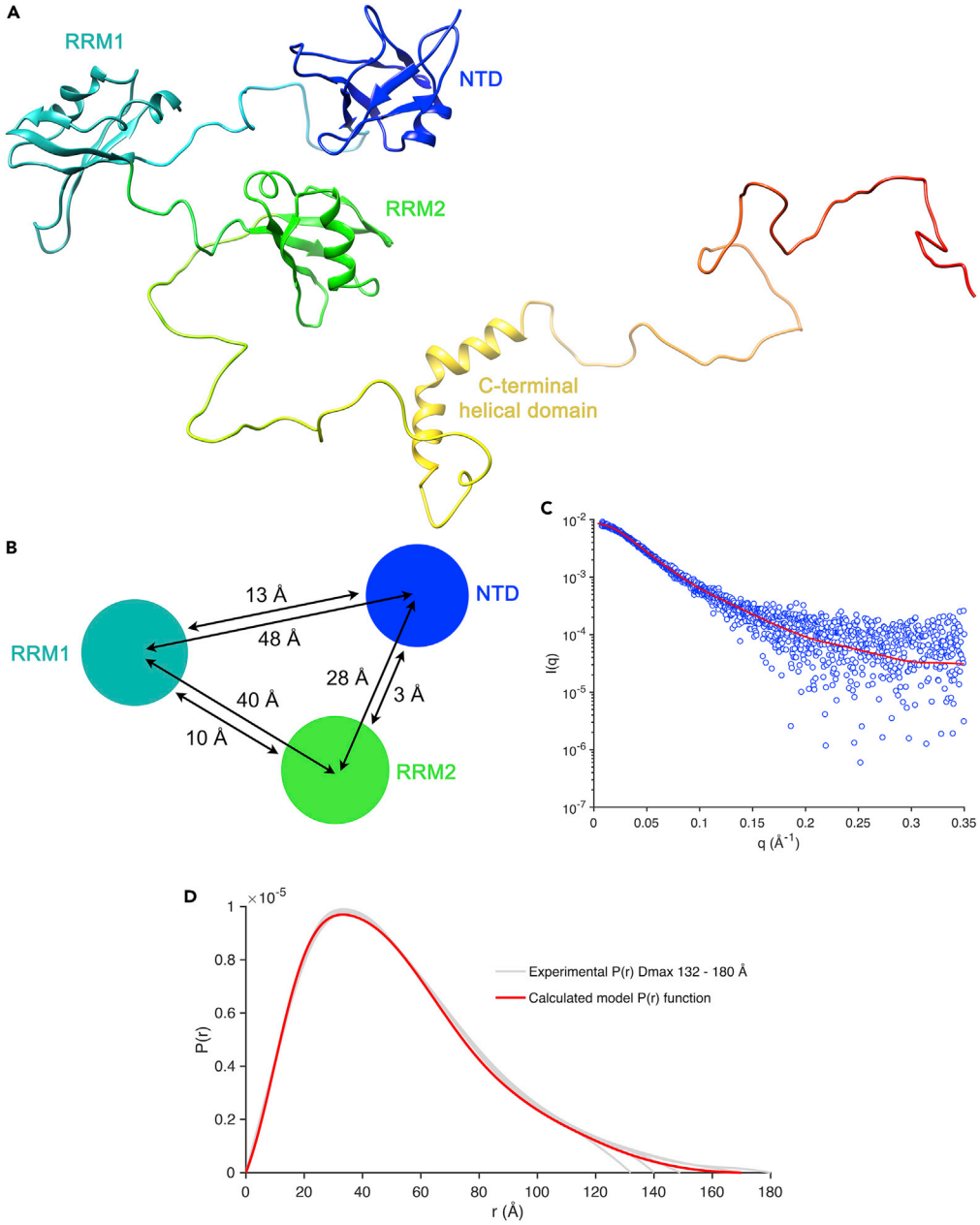


Figure 7. Model of Full-Length TDP-43 Refined Against SAXS Data

(A) Model of full-length TDP-43_{WtoA}. Model has Rg 41.9 Å and Dmax 154 Å.

(B) Schematic of the above showing center of mass and closest contact intradomain distances.

(C) Model fit to experimental intensity scattering data χ^2 1.10.

(D) Distance distribution functions for experimental data showing solutions with variable possible Dmax compared with that for the model.

and Table S1). For both well-fitting and poorly fitting model groups the specific orientation of globular domains cannot be accurately defined but the geometry of domain positions can.

Resource Availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, S. Samar Hasnain (s.s.hasnain@liverpool.ac.uk).

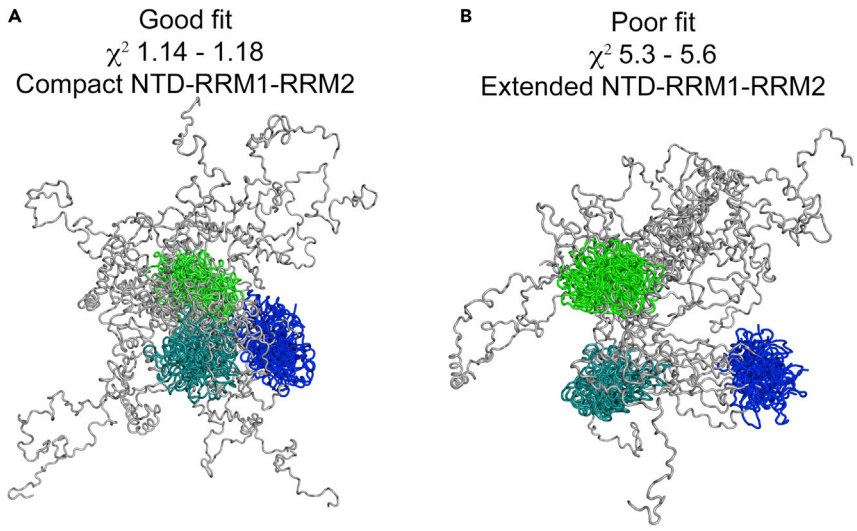


Figure 8. Comparison of Full-Length TDP-43$_{WtoA}$ Models with Both Good and Poor Fit to Experimental SAXS Data

(A) Ten TDP-43$_{WtoA}$ models with highest goodness of fit to the experimental data from a pool of 7,000 aligned to amino acids 1-258 of the model presented in Figure 7A.

(B) Ten models with lowest goodness-of-fit to experimental data aligned to amino acids 1-258 of the model with highest χ^2.

Materials Availability

All unique and stable reagents generated in this study are available from the Lead Contact without restriction.

Data and Code Availability

The atomic coordinate and structure factor of the NTD (NTD) of TDP-43 have been deposited in the Protein DataBank (http://www.rcsb.org/) under the accession code 6T4B.

Experimental SAXS data and TDP-43 models are available from the corresponding author on request.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101159.

ACKNOWLEDGMENTS

We acknowledge the use of DIAMOND light source crystallographic (I03) and SAXS (B21) facilities. The work was supported by MNDA grants: Hasnain/Apr15/833-791 and Wright/Oct18/969-799.

AUTHOR CONTRIBUTIONS

G.S.A.W. performed SAXS analysis and wrote the manuscript. T.F.W. produced protein and crystallized TDP-43 NTD, K.A. produced protein and wrote the manuscript, S.S.P. performed SASA calculations, N.R.C. suggested the idea of a Trp-free TDP-43, S.V.A. performed the crystallographic study and contributed to writing of the manuscript, S.S.H. initiated the project and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 14, 2019
Revised: March 29, 2020
Accepted: May 8, 2020
Published: June 26, 2020

REFERENCES

Afroz, T., Hock, E.-M., Ernst, P., Foglieni, C., Jambeau, M., Gilhespy, L.A.B., Laferriere, F., Maniecka, Z., Plückthun, A., Mittl, P., et al. (2017). Functional and dynamic polymerization of the ALS-linked protein TDP-43 antagonizes its pathologic aggregation. Nat. Commun. *8*, 45.

Amador-Ortiz, C., Lin, W.-L., Ahmed, Z., Personett, D., Davies, P., Duara, R., Graff-Radford, N.R., Hutton, M.L., and Dickson, D.W. (2007). TDP-43 immunoreactivity in hippocampal sclerosis and Alzheimer's disease. Ann. Neurol. *61*, 435–445.

Austin, J.A., Wright, G.S.A., Watanabe, S., Grossmann, J.G., Antonyuk, S.V., Yamanaka, K., and Hasnain, S.S. (2014). Disease causing mutants of TDP-43 nucleic acid binding domains are resistant to aggregation and have increased stability and half-life. Proc. Natl. Acad. Sci. U S A *111*, 4309–4314.

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.

Ayala, Y.M., De Conti, L., Avendaño-Vázquez, S.E., Dhir, A., Romano, M., D'Ambrogio, A., Tollervey, J., Ule, J., Baralle, M., Buratti, E., et al. (2011). TDP-43 regulates its mRNA levels through a negative feedback loop. EMBO J. *30*, 277–288.

Cao, Q., Boyer, D.R., Sawaya, M.R., Ge, P., and Eisenberg, D.S. (2019). Cryo-EM structures of four polymorphic TDP-43 amyloid cores. Nat. Struct. Mol. Biol. *26*, 619–627.

Chiang, C.-H., Grauffel, C., Wu, L.-S., Kuo, P.-H., Doudeva, L.G., Lim, C., Shen, C.-K.J., and Yuan, H.S. (2016). Structural analysis of disease-related TDP-43 D169G mutation: linking enhanced stability and caspase cleavage efficiency to protein accumulation. Sci. Rep. *6*, 21581.

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.

Deshaies, J.-E., Shkreta, L., Moszczynski, A.J., Sidibé, H., Semmler, S., Fouillen, A., Bennett, E.R., Bekenstein, U., Destroismaisons, L., Toutant, J., et al. (2018). TDP-43 regulates the alternative splicing of hnRNP A1 to yield an aggregation-prone variant in amyotrophic lateral sclerosis. Brain *141*, 1320–1333.

Fallini, C., Bassell, G.J., and Rossoll, W. (2012). The ALS disease protein TDP-43 is actively transported in motor neuron axons and regulates axon outgrowth. Hum. Mol. Genet. *21*, 3703–3718.

Furukawa, Y., Kaneko, K., Watanabe, S., Yamanaka, K., and Nukina, N. (2011). A seeding reaction recapitulates intracellular formation of sarkosyl-insoluble transactivation response element (TAR) DNA-binding protein-43 inclusions. J. Biol. Chem. *286*, 18664–18672.

Guenther, E.L., Cao, Q., Trinh, H., Lu, J., Sawaya, M.R., Cascio, D., Boyer, D.R., Rodriguez, J.A., Hughes, M.P., and Eisenberg, D.S. (2018). Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. Nat. Struct. Mol. Biol. *25*, 463–471.

He, Z., Zhang, C., Xu, Y., Zeng, S., Zhang, J., and Xu, D. (2014). MUFOLD-DB: a processed protein structure database for protein structure prediction and analysis. BMC Genomics *15*(*Suppl 11*), S2.

Hughes, M.P., Sawaya, M.R., Boyer, D.R., Goldschmidt, L., Rodriguez, J.A., Cascio, D., Chong, L., Gonen, T., and Eisenberg, D.S. (2018). Atomic structures of low-complexity protein segments reveal kinked β sheets that assemble networks. Science *359*, 698–701.

Johnson, B.S., Snead, D., Lee, J.J., McCaffery, J.M., Shorter, J., and Gitler, A.D. (2009). TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity. J. Biol. Chem. *284*, 20329–20339.

Kitamura, A., Shibasaki, A., Takeda, K., Suno, R., and Kinjo, M. (2018). Analysis of the substrate recognition state of TDP-43 to single-stranded DNA using fluorescence correlation spectroscopy. Biochem. Biophys. Rep. *14*, 58–63.

Kuo, P.H., Chiang, C.H., Wang, Y.T., Doudeva, L.G., and Yuan, H.S. (2014). The crystal structure of TDP-43 RRM1-DNA complex reveals the specific recognition for UG- and TG-rich nucleic acids. Nucleic Acids Res. *42*, 4712–4722.

Li, H.-R., Chiang, W.-C., Chou, P.-C., Wang, W.-J., and Huang, J. (2018a). TAR DNA-binding protein 43 (TDP-43) liquid–liquid phase separation is mediated by just a few aromatic residues. J. Biol. Chem. *293*, 6090–6098.

Li, H.-R., Chen, T.-C., Hsiao, C.-L., Shi, L., Chou, C.-Y., and Huang, J. (2018b). The physical forces mediating self-association and phase-separation in the C-terminal domain of TDP-43. Biochim. Biophys. Acta. *1866*, 214–223.

Li, W., Reeb, A.N., Lin, B., Subramanian, P., Fey, E.E., Knoverek, C.R., French, R.L., Bigio, E.H., and Ayala, Y.M. (2017). Heat shock-induced phosphorylation of TAR DNA-binding protein 43 (TDP-43) by MAPK/ERK kinase regulates TDP-43 function. J. Biol. Chem. *292*, 5089–5100.

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. Nat. Struct. Mol. Biol. *12*, 1443–1449.

McDonald, K.K., Aulas, A., Destroismaisons, L., Pickles, S., Beleac, E., Camu, W., Rouleau, G.A., and Vande Velde, C. (2011). TAR DNA-binding protein 43 (TDP-43) regulates stress granule dynamics via differential regulation of G3BP and TIA-1. Hum. Mol. Genet. *20*, 1400–1410.

McKee, A.C., Gavett, B.E., Stern, R.A., Nowinski, C.J., Cantu, R.C., Kowall, N.W., Perl, D.P., Hedley-Whyte, E.T., Price, B., Sullivan, C., et al. (2010). TDP-43 proteinopathy and motor neuron disease in chronic traumatic encephalopathy. J. Neuropathol. Exp. Neurol. *69*, 918–929.

Mompeán, M., Romano, V., Pantoja-Uceda, D., Stuani, 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.

Nakashima-Yasuda, H., Uryu, K., Robinson, J., Xie, S.X., Hurtig, H., Duda, J.E., Arnold, S.E., Siderowf, A., Grossman, M., Leverenz, J.B., et al. (2007). Co-morbidity of TDP-43 proteinopathy in Lewy body related diseases. Acta Neuropathol. *114*, 221–229.

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

Prakash, A., Kumar, V., Meena, N.K., and Lynn, A.M. (2018). Elucidation of the structural stability and dynamics of heterogeneous intermediate ensembles in unfolding pathway of the N-terminal domain of TDP-43. RSC Adv. *8*, 19835–19845.

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. U S A *111*, 18619–18624.

Sala, F.A., Wright, G.S.A., Antonyuk, S.V., Garratt, R.C., and Hasnain, S.S. (2019). Molecular recognition and maturation of SOD1 by its evolutionarily destabilised cognate chaperone hCCS. PLoS Biol. *17*, e3000141.

Tollervey, J.R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A.L., Zupunski, V., et al. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. Nat. Neurosci. *14*, 452–458.

Vivoli Vega, M., Nigro, A., Luti, S., Capitini, C., Fani, G., Gonnelli, L., Boscaro, F., and Chiti, F. (2019). Isolation and characterization of soluble human full-length TDP-43 associated with neurodegeneration. FASEB J. *33*, 10780–10793.

Wang, A., Conicella, A.E., Schmidt, H.B., Martin, E.W., Rhoads, S.N., Reeb, A.N., Nourse, A., Ramirez Montero, D., Ryan, V.H., Rohatgi, R., et al. (2018). A single N-terminal phosphomimic disrupts TDP-43 polymerization, phase separation, and RNA splicing. EMBO J. *37*, e97452.

Wennberg, A.M., Whitwell, J.L., Tosakulwong, N., Weigand, S.D., Murray, M.E., Machulda, M.M., Petrucelli, L., Mielke, M.M., Jack, C.R., Knopman, D.S., et al. (2019). The influence of tau, amyloid, alpha-synuclein, TDP-43, and vascular pathology in clinically normal elderly individuals. Neurobiol. Aging *77*, 26–36.

Wolozin, B. (2012). Regulated protein aggregation: stress granules and neurodegeneration. Mol. Neurodegener. *7*, 56.

iScience, Volume *23*

Supplemental Information

Purification and Structural Characterization of Aggregation-Prone Human TDP-43 Involved in Neurodegenerative Diseases

Gareth S.A. Wright, Tatiana F. Watanabe, Kangsa Amporndanai, Steven S. Plotkin, Neil R. Cashman, Svetlana V. Antonyuk, and S. Samar Hasnain

Supplemental Information

Supplemental Figures

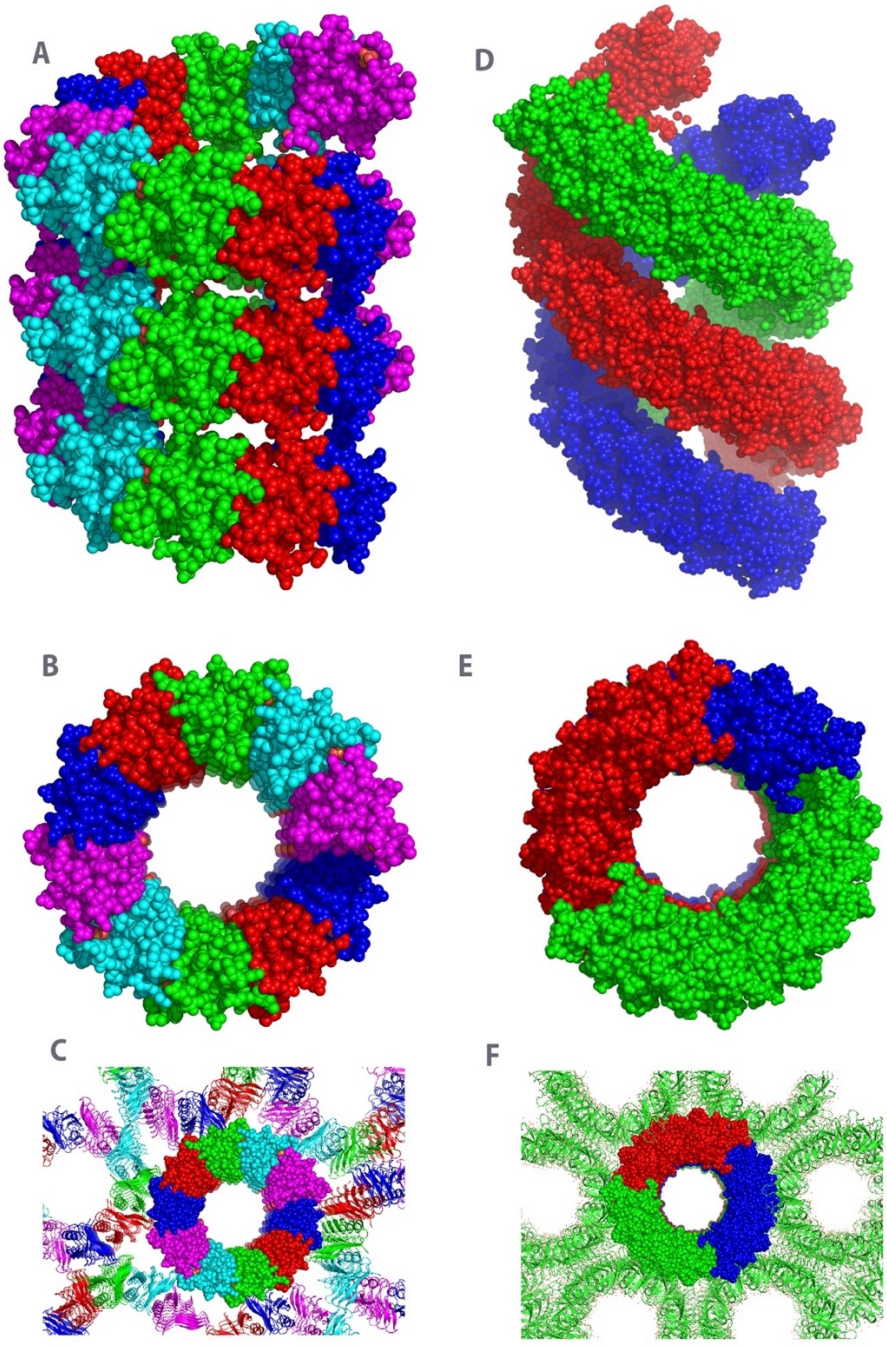


Figure S1. Arrangement of TDP-43 NTD domains within different crystal forms. Continuous spiral with 10 NDTs in one turn of the spiral A) side view of the spiral and, B) top view. C) Packing of the crystal from the same direction as B. The identical chains are illustrated in the same colour. D) Super-helical arrangement (PDB: 5MDI) from the side and, E) the top. F) Packing of the crystals containing filaments. Three different representative filaments containing 10 domains are coloured in red, green and blue.

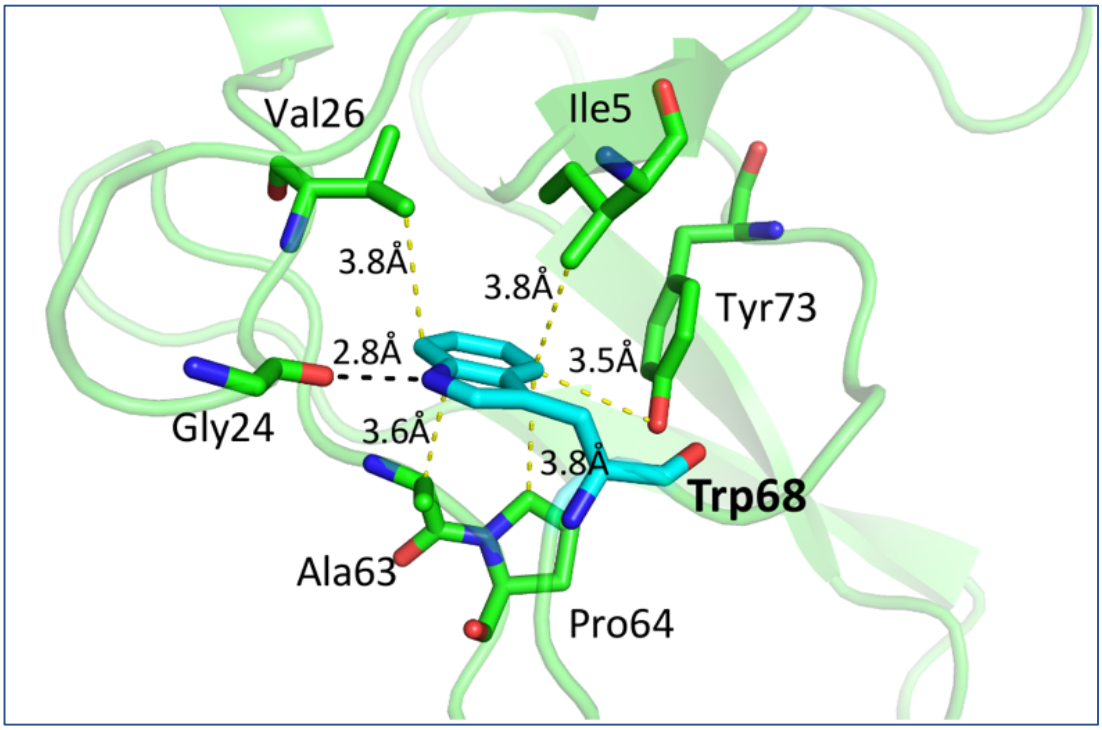


Figure S2. Hydrogen bond formation and hydrophobic interactions formed by TDP-43 N-terminal domain Trp68. Trp68 and surrounding residues are shown in cyan and green, respectively. Hydrogen bond and hydrophobic contacts are illustrated as black and yellow dashed lines, respectively.

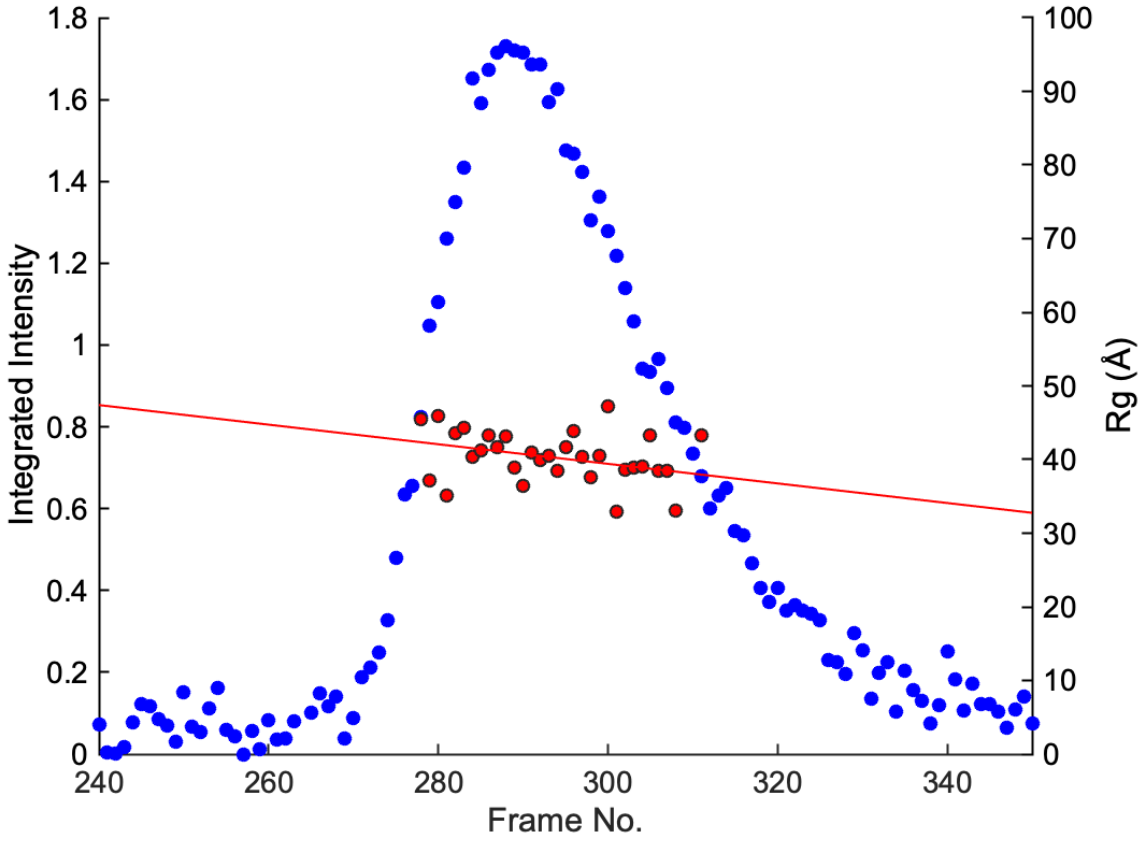


Fig S3. Chromatographic SAXS parameters for TDP-43$_{WtoA}$. Integrated intensity is plotted in blue. Radius of gyration is plotted in red. The slope of the Rg plot is shown as a red line which has gradient coefficient -0.13 and R^2 0.12. Data points in the first third of the elution have Rg 41.8 ± 3.5 Å while those in the last third have Rg 38.9 ± 4.0 Å, with error quoted as standard deviation.

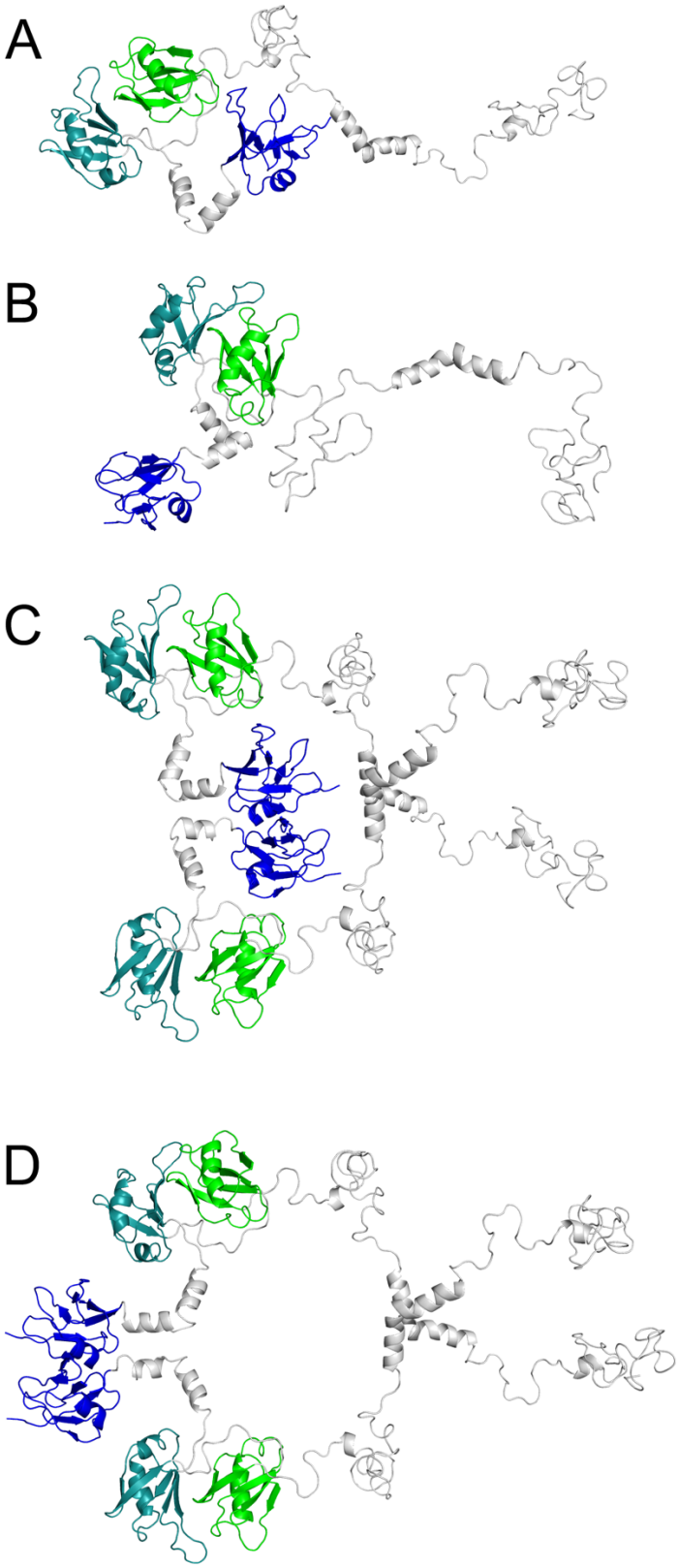


Figure S4. Comparison of initial models of full-length TDP-43. A and B) Monomeric TDP-43. C and D) Dimeric TDP-43.

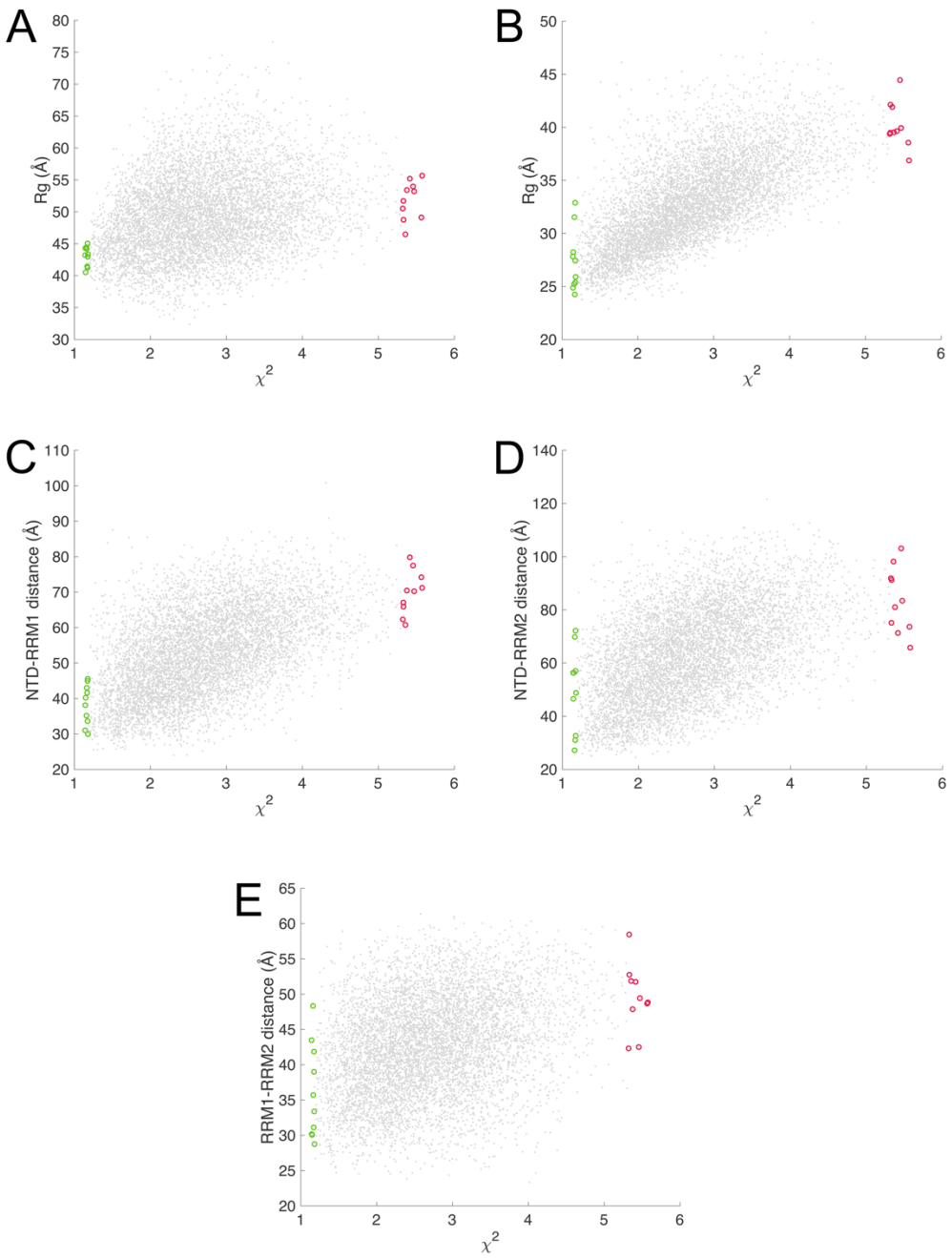


Figure S5. Analysis of goodness-off fit to experimental SAXS data versus size parameters for 7000 TDP-43$_{WtoA}$ models. A) The relationship between full-length TDP-43$_{WtoA}$ model Rg to χ^2. B) The relationship between Rg of the N-terminal region of TDP-43$_{WtoA}$ molecule comprising amino acids 1-258 (NTD-RRM1-RRM2) to χ^2. C) The relationship between TDP-43$_{WtoA}$ NTD-RRM1 distance and χ^2. D) The relationship between the TDP-43$_{WtoA}$ NTD-RRM2 distance and χ^2. E) The relationship between the TDP-43$_{WtoA}$ RRM1-RRM2 distance and χ^2. Parameters for structures presented in Figure 8A and 8B are highlighted with green and red respectively with numerical data given in Table S1.

	10 well-fitting models	10 poorly fitting models	Optimised model (Figure 7A)
χ^2	1.16 ± 0.01	5.41 ± 0.10	1.10
Rg of full-length TDP43 (Å)	43.1 ± 1.5	51.8 ± 3.0	41.9
Rg of amino acids 1-258 (Å)	27.4 ± 2.9	40.2 ± 2.1	25.4
NTD-RRM1 distance (Å)	38.3 ± 5.7	70.0 ± 6.2	48.2
NTD-RRM2 distance (Å)	49.8 ± 15.7	83.5 ± 12.3	28.2
RRM1-RRM2 distance (Å)	36.2 ± 6.7	49.4 ± 4.7	40.3

Table S1. Comparison of size parameters for the models presented in Figures 7 and 8. Error measurements are given as standard deviation.

Transparent Methods

Expression and purification

Full-length wild-type TDP-43 (wtTDP-43) and TDP-43 construct containing W67A, W113A, W172A, W334A, W385A and W412A mutations (TDP-43_{WtoA}) were synthesised *de novo* with an N-terminal hexa-his tag and tobacco etch virus (TEV) cleavage site in pET-28a(+) vector. This expression plasmid was transformed into BL21 (DE3) *E. coli* and pre-cultured at 37 ºC overnight in LB media with 50 µg/ml of kanamycin. 25 ml of this pre-culture was added into 1 litre of LB media and the culture was incubated at 37 ºC in a shaker until the optical density at 600nm was 0.7. The culture was cooled for 30 min at 4 ºC. IPTG was added to a final concentration of 0.5 mM to induce protein expression and the flasks were kept in shaker overnight at 18 ºC. The cells were then harvested by centrifugation.

1 g of cells was resuspended in 5 mL water and incubated at 4 ºC for 30 minutes before sonication on ice. The cell lysate was clarified by centrifugation at 35,000 g for 1 hour before adding 50mM sodium phosphate pH 8.0 and 0.2% sarkosyl. A nickel-NTA IMAC column was equilibrated with 50 mM sodium phosphate pH 8.0, 300 mM sodium chloride, 5 mM imidazole, 0.2% sarkosyl prior to application of the soluble fraction at room temperature. Protein was eluted from the column with 150-500 mM imidazole in 50 mM sodium phosphate pH 8.0, 300 mM sodium chloride, 5 mM DTT and dialysed overnight at 4 ºC with the addition of TEV protease against the same buffer without imidazole. The cleaved protein was then filtered through the NiNTA column again with wtTDP-43 now passing directly through. It was then concentrated and applied to a Superdex 75 16x600 mm size exclusion chromatography column with 0.5 ml/min flow of 20 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM DTT, 0.2% sarkosyl. Fractions were analysed by SDS-PAGE and those found to contain full-length wtTDP-43 were pooled, concentrated and stored at 4 ºC.

Full-length TDP-43_{WtoA} and truncated amino acids 1-270 wild-type TDP-43 were expressed and purified as above protocol without using sarkosyl, but cell lysis was carried out in 50 mM sodium phosphate pH 8.0, 300 mM sodium chloride, 5 mM imidazole, 5 mM DTT, complete protease inhibitor cocktail (Roche), 1 mM PMSF, 50 ug/ml lysozyme instead of water.

Crystallisation

A 1-270 amino acid fragment of TDP-43 protein at 40 mg/ml concentration was crystallised at 19 ºC by the hanging drop method. 1 µL of protein was mixed with 1 µL of ready-made solution

F2 from PACT premier screen (Molecular Dimentions) and equilibrated over the reservoir solution, containing 0.2M Sodium bromide, 0.1M Bis-Tris propane 6.5, 20% PEG 3350. Crystals were flash frozen in liquid nitrogen using reservoir solution with additional 10% glycerol.

Data collection, structure determination and refinement

Data were collected at DIAMOND synchrotron, beamline I03 using x-rays of 0.9763 Å wavelength with PILATUS 6M detector to 2.55 Å resolution. Data were integrated with iMosflm (Battye et al., 2011) and scaled with Aimless (Evans, 2011) software as part of the CCP4 package. The structure was solved by Molrep software (Vagin and Teplyakov, 2010) with starting model TDP-43 structure (PDB:5MDI chain A) and refined using Refmac5 (Murshudov et al., 2011) with applied NCS symmetry. TLS refinement was implemented towards the end of the refinement. Data collection and refinement statistics are presented in Table 1.

Small-angle x-ray scattering data collection

Chromatographic SAXS data for full-length TDP-43_{WtoA} was collected at Diamond Light Source on beamline B21. Data was acquired following a chromatographic step where 45 μl of TDP-43_{WtoA} at 5 mg/ml was loaded onto a Superdex 200 10x300 mm size exclusion chromatography column at room temperature. The protein was eluted at a flow rate of 200 μl/min and directly exposed to x-rays. Blank frames were taken prior to protein elution for buffer subtraction purposes. Individual frames recorded over a single chromatography run were averaged based on Rg values and overall similarity in a correlation map. Two separate chromatography runs were performed and the results of each were averaged to give the final scattering profile. Averaging was performed with ScÅtter. The scattering curve of full-length TDP-43 was observed in ScÅtter, Primus (Konarev et al., 2003) and Matlab where the Guiner approximation was performed and Rg determined. GNOM (Svergun, 1992) was used to determine distance distribution functions.

Structure modelling and refinement against SAXS data

An initial model of full-length TDP-43 was constructed from NMR structures of the NTD (Mompeán et al., 2016) (2N4P); RRM1 and RRM2 (Lukavsky et al., 2013) (4BS2); and an helical structure formed by residues 321-343 in the C-terminal tail (Jiang et al., 2016) (2N3X) shown to be partially populated (Conicella et al., 2016). Pepfold (Maupetit et al., 2009) was used to generate the remaining linkers between these structured segments and Modloop

(Fiser and Sali, 2003) was used to link each component to form a continuous polypeptide. Tryptophan residues were mutated to alanine using Coot (Emsley and Cowtan, 2004).

Two different starting models were refined against experimental SAXS data using CNS (Brunger, 2007), as described previously (Wright et al., 2016, 2018). Domains defined above were initially treated as free-floating rigid bodies and inter-domain linkers allowed to move freely over 750 ps simulations at 300,000 K. Over the course of the simulation 1000 structures were written out and compared with the experimental scattering data with FoXS (Schneidman-Duhovny et al., 2016). This global refinement of domain positions was followed by fine-tuning of the position and orientation of individual domains. Here, the majority of the protein was fixed in space while one domain and its surrounding linkers were allowed to move. Each step in this process was performed independently with different initial trajectories 6 to 8 times. The structure that fit the data best was taken forward for further optimisation. To ensure that conformational space had been adequately sampled during this process, a 3.75 ns simulation was performed starting from our optimised model with all domains able to move freely. This was independently repeated with different seed trajectories 7 times. Domain centres of mass were calculated using CNS.

Experimental SAXS data and derived models are available upon request.

Solvent accessible surface area calculations

Apo structures (protein only without nucleic acid or other ligand) are used to calculate the SASA of tryptophans in a non-redundant database of 27,015 structures taken from the Protein Data Bank using MUFOLD-DB (He et al., 2014), with 70% sequence identity. The SASA of amino acids was calculated using tcl scripts in VMD (Humphrey et al., 1996) and confirmed using built-in commands in GROMACS modified with improved resolution of sampling points. Hydrogens are added using GROMACS before the SASA calculation.

Supplemental References

Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011). iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallographica Section D: Biological Crystallography *67*, 271–281.

Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nat Protoc *2*, 2728–2733.

Conicella, A.E., Zerze, G.H., Mittal, J., and Fawzi, N.L. (2016). ALS Mutations Disrupt Phase Separation Mediated by α-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. Structure *24*, 1537–1549.

Emsley, P., and Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. Acta Crystallographica Section D: Biological Crystallography *60*, 2126–2132.

Evans, P.R. (2011). An introduction to data reduction: Space-group determination, scaling and intensity statistics. Acta Crystallographica Section D: Biological Crystallography *67*, 282–292.

Fiser, A., and Sali, A. (2003). ModLoop: automated modeling of loops in protein structures. Bioinformatics *19*, 2500–2501.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J Mol Graph *14*, 33–38, 27–28.

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*, 1–11.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Cryst. *36*, 1277–1282.

Maupetit, J., Derreumaux, P., and Tuffery, P. (2009). PEP-FOLD: an online resource for de novo peptide structure prediction. Nucleic Acids Res. *37*, W498-503.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallographica Section D: Biological Crystallography *67*, 355–367.

Schneidman-Duhovny, D., Hammel, M., Tainer, J.A., and Sali, A. (2016). FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res. *44*, W424-429.

Svergun, D.I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Cryst. *25*, 495–503.

Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. Acta Crystallographica Section D: Biological Crystallography *66*, 22–25.

Wright, G.S.A., Antonyuk, S.V., and Hasnain, S.S. (2016). A faulty interaction between SOD1 and hCCS in neurodegenerative disease. Sci Rep *6*, 27691.

Wright, G.S.A., Saeki, A., Hikima, T., Nishizono, Y., Hisano, T., Kamaya, M., Nukina, K., Nishitani, H., Nakamura, H., Yamamoto, M., et al. (2018). Architecture of the complete oxygen-sensing FixL-FixJ two-component signal transduction system. Sci. Signal. *11*, eaaq0825.