The FEBS Journal

FEBS PRESS science publishing by scientists

The TDP-43 N-terminal domain structure at high resolution

Miguel Mompeán[1,*], Valentina Romano[2,*], David Pantoja-Uceda[1], Cristiana Stuani[2], Francisco E. Baralle[2], Emanuele Buratti[2] and Douglas V. Laurents[1]

1 Instituto de Química Física Rocasolano, Madrid, Spain
2 International Centre for Genetic Engineering and Biotechnology, Trieste, Italy

Keywords
ALS/FTLD; hydrogen/deuterium exchange; protein dynamics; protein structure; TAR DNA-binding protein 43 (TDP-43)

Correspondence
E. Buratti or D. V. Laurents, Instituto de Química Física Rocasolano, CSIC Serrano 119, E-28006 Madrid, Spain
Fax: +34 91 564 2431
Tel: +34 91 745 9543
E-mails: buratti@icgeb.org or dlaurents@iqfr.csic.es

*These authors contributed equally to this work.

(Received 13 August 2015, revised 18 December 2015, accepted 8 January 2016)

doi:10.1111/febs.13651

Transactive response DNA-binding protein 43 kDa (TDP-43) is an RNA transporting and processing protein whose aberrant aggregates are implicated in neurodegenerative diseases. The C-terminal domain of this protein plays a key role in mediating this process. However, the N-terminal domain (residues 1–77) is needed to effectively recruit TDP-43 monomers into this aggregate. In the present study, we report, for the first time, the essentially complete ^{1}H, ^{15}N and ^{13}C NMR assignments and the structure of the N-terminal domain determined on the basis of 26 hydrogen-bond, 60 torsion angle and 1058 unambiguous NOE structural restraints. The structure consists of an α-helix and six β-strands. Two β-strands form a β-hairpin not seen in the ubiquitin fold. All Pro residues are in the *trans* conformer and the two Cys are reduced and distantly separated on the surface of the protein. The domain has a well defined hydrophobic core composed of F35, Y43, W68, Y73 and 17 aliphatic side chains. The fold is topologically similar to the reported structure of axin 1. The protein is stable and no denatured species are observed at pH 4 and 25 °C. At 4 kcal·mol^{-1}, the conformational stability of the domain, as measured by hydrogen/deuterium exchange, is comparable to ubiquitin (6 kcal·mol^{-1}). The β-strands, α-helix, and three of four turns are generally rigid, although the loop formed by residues 47–53 is mobile, as determined by model-free analysis of the $^{15}N\{^{1}H\}$NOE, as well as the translational and transversal relaxation rates.

Database
Structural data have been deposited in the Protein Data Bank under accession code: 2n4p. The NMR assignments have been deposited in the BMRB database under access code: 25675.

Introduction

Transactive response DNA-binding protein 43 kDa (TDP-43) is a 414 residue protein belonging to the heterogeneous nuclear ribonucleoprotein (hnRNP) family. First noted for its capacity to bind the TAR element of the AIDS virus RNA [1], the true physiological roles of TDP-43 in RNA splicing, stability and

Abbreviations
ALD/FTLD, amyotrophic lateral sclerosis/frontotemporal lobar degeneration; BMRB, Biological Magnetic Resonance Data Bank; ΔCp, heat capacity change (upon unfolding); DSS, sodium 4,4-dimethyl-4-silapentane-1-sulfonate; H/D, hydrogen/deuterium; hnRNP, heterogeneous nuclear ribonucleoprotein; HSQC, heteronuclear single quantum correlation; NTD, N-terminal domain; RRM, RNA recognition motif; TCEP, Tris (2-carboxylethyl) phosphine; TDP-43, transactive response DNA-binding protein 43 kDa.

transport were identified in 2001 [2] and have been reviewed recently [3]. Nine years ago, TDP-43 was found to form aberrant polyubiquitinated, hyperphosphorylated cytosolic aggregates that have been causatively linked to amyotrophic lateral sclerosis/frontotemporal lobar degeneration (ALS/FTLD) [4,5] and, more recently, to Alzheimer's disease [6]. The protein contains an N-terminal domain (NTD), two RNA recognition motif (RRM) RNA-binding domains and a long, intrinsically disordered C-terminal region. TDP-43 and, in particular, its N-terminal and RRM domains are well conserved in animals ranging from man to *Caenorhabditis elegans* [7]. The C-terminal region contains three segments enriched with Gly, hydrophobic and Asn/Gln residues, respectively. Considering that almost all of the amino acid substitutions arising from mutations linked genetically to pathology are found in this region [8], the C-terminus appears to be a key protagonist in the pathological processes. There is an open debate and contrasting results regarding whether the toxic aggregate is amorphous [4] or amyloid-like [9,10] and which segment is key for its formation [11–14]. We have recently shown that tandem repeats of the C-terminal segment enriched with Asn and Gln residues can reproduce almost all the disease hallmarks of the full-length protein in cells [15]. In aqueous solution, the peptide corresponding to this segment (341–366) is initially unfolded but then adopts a aggregating conformer that is rich in β-structure [16] with many amyloid-like characteristics [17]. Similar results have been reported independently for a peptide corresponding to residues 331–360 [18].

Notwithstanding the importance of the C-terminal region, the NTD, composed of residues 1–77, has recently been shown to be essential for efficient aggregation that depletes the cell of functional TDP-43 [19–21]. This result prompted us to characterize the structure, conformational stability and dynamics of this domain using NMR spectroscopy. When our work was underway, a low resolution structural model of the backbone structure was reported by Qin *et al.* [22] for the NTD construct containing residues 1–80 and a C-terminal His-tag. The model contains five β-strands and an α-helix whose topology matches the ubiquitin fold. The conformational stability of the NTD in the context of the construct reported by Qin *et al.* [22] is remarkably low, with the majority of the protein being denatured at equilibrium, and it was hypothesized that this low stability is key for the physiological and pathological roles of the protein. Qin *et al.* [22] also claim that DNA binding can stabilize the folded form of a longer construct containing TDP-43 residues 1–102. Despite the significant advance that these results represent, no structure or chemical shift values were deposited and the resolution reported was low. In addition, it is puzzling that the conformational stability of this NTD construct reported by Qin *et al.* [22] was reported to very low, with most of the protein being unfolded at equilibrium, when ubiquitin is a paradigm of a very stable, soluble protein [23,24]. Indeed, ubiquitin behaves so well that it is routinely chosen by NMR spectroscopists as an easy model protein for developing new pulse sequences. In the present study, we report the complete chemical shift assignments, high resolution NMR structure, conformational stability and backbone dynamics of the NTD. Moreover, we find that a disordered peptide corresponding to TDP-43 residues 78–102, which lies outside the folded NTD and contains several cationic residues, is able to bind DNA oligos.

Results

Production and purification of wild-type NTD and variants

As described in the Experimental procedures, we have successfully purified and produced His-tagged N-terminus sequences of TDP-43 that included residues 1–77 ($NTD_{1–77}$) or 1–102 ($NTD_{1–102}$) of this protein. The recombinant peptide was produced in bacteria and purified using a nickel-nitrilotriacetic acid resin (Qiagen, Valencia, CA, USA) in accordance with standard protocols.

Mass spectra

MALDI-TOF mass spectra of the $NTD_{1–77}$ showed masses of 9923–9929 Da for the unlabeled protein (data not shown). These values are in excellent agreement with the theoretical mass of 9927 Da calculated for the sequence of this construct: *M-11R-10G-9SHHH-5HHHG-1 S0*M1SEYI5RVTED10ENDEP15IEIPS20EDDGT25VLLST30VTAQF35PGACG40LRYRN45PVSQC50MRGVR55LVEGI60LHAPD65AGWGN70LVYVV75NY77, for which the His-tag is shown in italics. The $^{13}C,^{15}N$ protein shows a mass peak at 10 424.8 Da, which is slightly lower than the calculated mass of 10 480.9 Da for the $^{13}C,^{15}N$-labeled protein. Based on this difference, the overall incorporation of ^{13}C and ^{15}N is calculated to be 89.9%. The MALDI-TOF mass spectrum of the ^{15}N-labeled $NTD_{1–102}$ showed a major peak at 12 889.3 Da, which corresponds to its sequence, and a 90.5% incorporation of ^{15}N (data not shown). An additional peak at

13 208.4 Da, which may be attributed to mixed disulfide formation with C39, was also observed.

NMR assignment of wild-type NTD_{1-77}

The assignments are essentially complete; only the ^{15}N of prolines, the C′ of residues preceding prolines, the side chain carbonyl carbons, the H-less C of aromatic residues, all the nuclei of the first two residues of the His tag, and the ^{1}H and ^{15}N of the third residue of the His tag are missing. In the folded region of the protein; only the $^{13}C^{1}H_3\delta$ of M51 and $^{13}C'$ of Y77 (the C-terminal $-^{13}COO^-$) are missing. The assignments are unambiguous, except for the usual cases involving geminal -H (in Lys, Arg, Met, etc.), $-CH_3$ (in Val or Leu) and δ,ε signals in the Tyr and Phe aromatic rings. The six His of the His-tag share very similar chemical shift values. Two residues, R52 and G53, show two distinct ^{1}H-^{15}N correlations, which is evidence for conformational diversity. The NMR assignments have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) database under access code: 25675. The ^{1}H-^{15}N spectra at 5 °C and 39 °C and backbone and sidechain signals labeled at 25 °C show the broad chemical shift dispersion typical of folded proteins and an absence of signals characteristic of denatured protein over the range 5–34 °C (Fig. 1). A minor population of signals arising from unfolded protein is detected at 39 °C. These signals disappear upon recooling.

The NTD adopts a well folded structure that resembles axin 1 more than ubiquitin

Secondary structure of the N-terminal domain consists of six β-strands (β1, I5T8; β2, I16-P19; β3, G40-R44; β4, R55-V57; β5, I60-H62; and β6, Y73-N76) and an α-helix composed of residues L28-Q34 (Figs 2 and 3). These strands can be grouped into two β-sheets with the topologies β2 *ap* β1 *p* β6 *ap* β3 (sheet 1) and β4 *ap* β5 (sheet 2), where *p* and *ap* denote parallel and antiparallel configurations, respectively. Although the topology of the first sheet is similar to that of ubiquitin (Fig. 3E), the second sheet is a novel feature unique to TDP-43 NTD (Fig. 3B) and to the reported structure of the C-terminal Dix domain of axin 1 (Fig. 3F), a scaffolding protein implicated in axonal growth and transport [25]. Strikingly, although their sequences are dissimilar, equivalent Cα atoms in the TDP-43 NTD and the axin 1 Dix domain can be

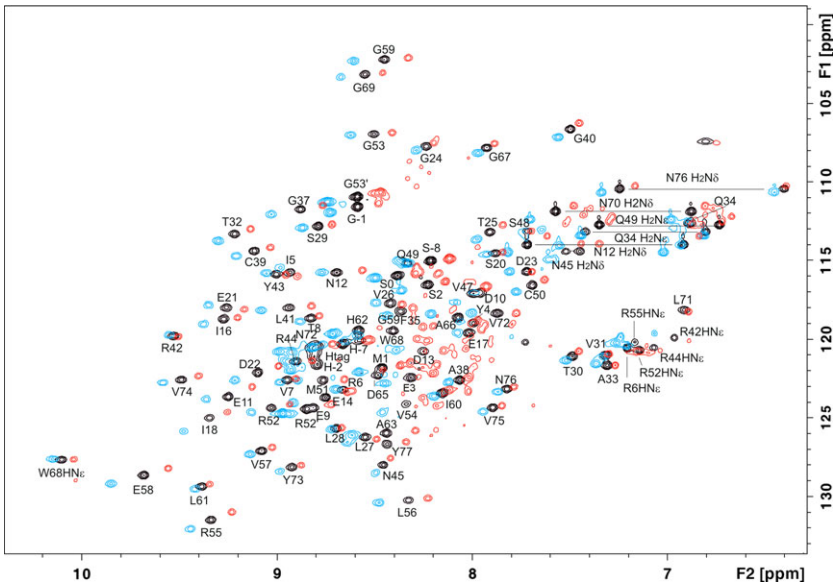

Fig. 1. Assigned ^{1}H-^{15}N HSQC spectra of the NTD at 5, 25 and 39 °C. The ^{1}H-^{15}N HSQC spectra of the NTD at 5 °C (blue), 25 °C (black) and 39 °C (red) recorded in 1 mM TCEP, 3 mM Na/DAc(d_3), 1 mM NaN_3 in 85% MilliQ H_2O and 15% D_2O are shown. The ^{1}H chemical shifts and, to a lesser extent, those of ^{15}N tend to decrease upon heating, although the changes are smaller for buried HNs forming hydrogen bonds (e.g. I5 and L61) and larger for solvent exposed HNs such as E58 and G69. Two distinct backbone HN signals were observed and assigned for R52 and G53. The second peak is labeled with a prime symbol. Although essentially no unfolded signals are seen at 5, 25 and 34 °C (not shown), a small population of resonances with denatured-like chemical shifts (between 8.3 and 7.95 ppm ^{1}H and 115 and 123 ppm ^{15}N) appear at 39 °C. These signals disappear upon recooling. The side chain HNε of W68 and $H_2N\varepsilon$ and $H_2N\delta$ of Gln and Asn residues are labeled. The side chain HNε of the Arg residues are folded and are also labeled.

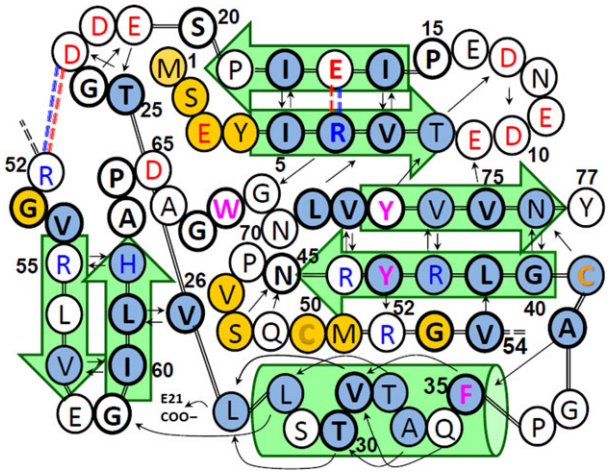

Fig. 2. Summary of N-terminal domain structure, stabilizing interactions and dynamics. Transparent green arrows or a cylinder represent the β-strands and the α-helix, respectively. The double line marks the chain trace. Bold letters and thick circles indicate buried residues (solvent accessible surface area < 25%). Blue and red letters represent positively and negatively charged residues, respectively. Blue and red dashed lines indicate favorable charge–charge interactions that may form at 4.5 < pH < 10. Solid arrows represent backbone hydrogen bonds, respectively. W, F and Y residues are labeled magenta. Cys residues are labeled amber. Every fifth residue, and some other notable residues, are numbered in black. Blue shaded circles indicate residues whose HN group exchanges slowly. Orange shaded circles indicate residues whose HN group is dynamic (^{1}H-^{15}N NOE ratio < 0.70).

superimposed with an RMSD of only 2.2 Å, and this close structural similarity is characteristic of homologous proteins [26]. Using the automatic software DALI (http://ekhidna.biocenter.helsinki.fi/dali_server), we found scores of proteins with significant structural similarity to the TDP-43 NTD. Most of these proteins mediate protein/protein interactions. However, none shared significant sequence similarity with the TDP-43 NTD and, except for axin 1, none has the same secondary structure topology. The uniqueness of the NTD sequence and its distinct structure could explain the limited success in predicting its structure via fragment-based *ab initio* modeling [22].

Long-range backbone hydrogen bonds between V26-L61, Y43-R52 and L41-V54 contribute the stability of the β-strands. In addition to these canonical structures, the residue segments E9-D13, E21-T25, F35-A38 and E58-G59 form well-defined turns that are stabilized by medium range backbone hydrogen bonds. The NTD is stabilized by the burial of numerous nonpolar groups. Trp68 lies at the heart of the hydrophobic core and makes multiple tertiary contacts. Its indole ring packs against Ala63 and Ile5, whose methyl groups experience intense ring current effects and are strongly shifted upfield. The Trp68 HNε is hydrogen-bonded to the O=C of Gly24 and shows slowed hydrogen/deuterium (H/D) exchange. The phenyl moiety of Tyr43, also buried, packs edgewise against the Hζ3 and Hη2 of Trp68, which are also shifted upfield. The side chains of Tyr43 and Tyr73, including their -OH groups, are buried in the hydrophobic core. The Tyr73 phenyl-O^{1}H signal is visible at 9.69 ppm in 1D ^{1}H and 2D NOESY spectra recorded in 85% H_2O/15% D_2O but quickly disappears in 100% D_2O. Some resonance broadening and putatively distinct ^{1}Hδ/^{1}Hδ' signals at 7.06 and 6.81 ppm reveal that the ring of Tyr73 rotates relatively slowly on a millisecond timescale.

In addition to I5, G24, A63, W68, Y43 and Y73 mentioned above, residues R6, V7, I16, I18, S20, T25, V26, T30, V31, F35, A38, G40, L41, Y43, N45, G53, V54, G59, I60, L61, G67, L71, V72 and V75 are also predominately buried (% solvent accessible surface area < 25%). Regarding the Pro residues, P15 and P64 are also buried in the hydrophobic core. All of the X-Pro peptide bonds are in the *trans* conformer, as indicated by the large difference between the ^{13}Cβ Pro and ^{13}Cγ Pro chemical shifts [27]. Although most nonpolar side chains are buried within the hydrophobic core, P36, P46, V47 and M51 are exposed and could contribute to intermolecular interactions. P46 and V47 are rather close to the E9-D13 tight turn, which bears a high density of negative charge. The structural statistics data for the final ensembles of 20 structures obtained for $NTD_{1–77}$ are provided in Table 1. The family of 20 refined structures has been deposited in the Protein Data Bank under accession code: 2n4p.

Because intradomain disulfide bond formation could contribute to the conformational stability of NTD, it is interesting to emphasize that the two cysteine residues of the domain are located far apart from each other and would not be able to form a disulfide bond in the folded, monomeric state. Because they are rather exposed, they may form disulfide bonds with other TDP-43 domains or with another NTD in a hypothetical dimer. Some studies provide evidence for the oxidized state of Cys residues in the RRM domains of full length TDP-43 contributing to aggregation [28,29]. Here, both Cys are in the reduced state as determined on the basis of their ^{13}Cβ chemical shift. These Cys remain reduced and the $NTD_{1–77}$ is monomeric even in the absence of reducing agents both at pH 4 and 7.3. In the case of the longer $NTD_{1–102}$, Cys39 appears to form a mixed disulfide with a small thiol containing molecule (see below). The lack of disulfide mediated dimerization, despite

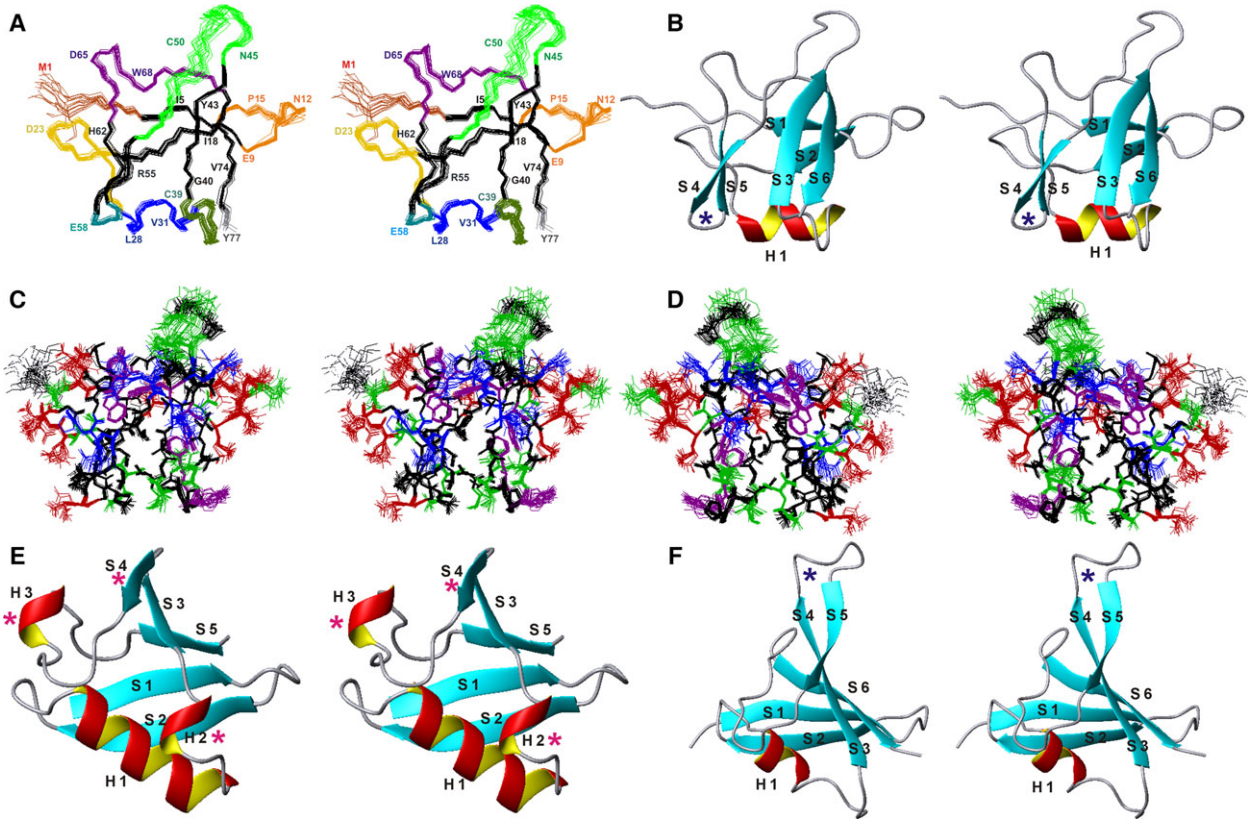

Fig. 3. Structure of the N-terminal domain. (A, C, D) Cross-eyed stereo views of all 20 structures of the N-terminal domain. (A) β-strands are shown in black, the α-helix is shown in blue and residues in turn or loop segments are colored: maroon = residues 1–4; orange = 9–15; golden = 20–27; green = 35–39; light green = 45–54; turquoise = 58–59; purple = 63–72; gray = Y77. Several residues are labeled to allow the chain trace to be easily followed. (B) Cross-eyed stereo ribbon diagram of the NTD. β-strands (S1, S2, etc.) and the helix (H) colored cyan and red/yellow, respectively, and a blue asterisk marks the distinct β-strand of NTD. (C, D) All of the heavy atoms of the NTD, with residues colored: apolar = black; polar = green; cationic = blue; anionic = red; and aromatic = purple. The structures in (D) are rotated 180° around the vertical axis with respect to (C). (E, F) Ribbon diagrams of the structures of ubiquitin (3EHV) and axin 1 (2d5g_6), respectively. The helices and distinct β-strand that distinguish ubiquitin from the TDP-43 NTD are marked with pink asterisks. Axin 1 contains a β-hairpin (blue asterisk) that resembles that of the NTD, although note that the orientation is distinct.

the favorable surface situation of the two Cys residues, suggests that additional portions of TDP-43 or other proteins or both may be necessary for dimerization.

H/D exchange reveals many slow exchanging HN groups

The kinetics of H/D exchange out of the NTD was followed by NMR spectroscopy (Fig. 4A). Out of a total of 72 backbone HNs in the domain, 31 exchange slowly. This is a fairly high ratio and is consistent with the elevated number of HN groups forming hydrogen bonds (Fig. 2). The indole HNε group of Trp68, which is hydrogen-bonded to the carbonyl of Gly24, also shows some protection against exchange. Some representative fits of a single exponential decay function to the peak integrals versus time are shown in Fig. 4B. These rates are up to hundreds of times slower than the corresponding 'intrinsic exchange rates' for a short, unstructured peptide with this sequence at pH 3.8 and 25 °C. Slow exchanging HN groups are observed in all elements of secondary structure, and the slowest are located in the internal β-strands or buried face of the α-helix, as is typically observed in other proteins such as ubquitin [30] or eosinophil cationic protein [31]. After exchange into D_2O was complete, the sample was exchanged back into 85% H_2O buffer and the final 1D ^1H and 2D NOESY spectra were recorded. Although the signal to noise ratio of this spectrum was only fair, the resonances of S-8, the His tag, S0, E3, M12, E17, G24, G38, M51, R52, R55, E58, G59

Table 1. Structure calculation and validation statistics.

NOE distances and dihedral constraints	
No. of short-range distances ($\|i - j\| \leq 1$)	511
No. of medium-range distances ($1 < \|i - j\| < 5$)	123
No. of long-range distances ($\|i - j\| \geq 5$)	424
No. of restraints from H-Bonds	26
No. of angular restraints (ϕ, ψ)	60
No. of total restraints	1058 + 26 + 60 = 1144
No. of restrictions/residue	13.74
Cyana target function value	1.09
AMBER refinement statistics	
RMSD from experimental restraints (NOE) (min, mean, max)	0.07, 0.13, 0.53
RMSD from dihedral restraints (TALOS) (min, mean, max)	2.54, 4.81, 7.06
Average AMBER energy (kcal/mol)	
RMSD bond lengths from ideal geometry (Å)	0.01
RMSD angles from ideal geometry (Å)	2.68
Averaged pairwaise RMSD (Å) (20 final structures)	
Mean Pairwise RMSD for the His tag (residues -9 - 0)	
Backbone	3.08 ± 0.78 Å
All heavy atoms	4.82 ± 0.72 Å
Mean Pairwise RMSD (residues 1–77, this excludes the His-tag) for the family of 20 structures	
Backbone	0.67 ± 0.18 Å
All heavy atoms	1.38 ± 0.18 Å
Mean Pairwise RMSD for Secondary Structure Elements (residues: 5–8, 16–19, 28–35, 40–44, 55–57, 60–62, 72–76)	
Backbone	0.23 ± 0.05 Å
All heavy atoms	1.01 ± 0.15 Å
Ramachandrian plot analysis	
Allowed regions, %	83.1
Additional allowed regions, %	16.9
Generously allowed regions, %	0.0
Disallowed regions, %	0.0

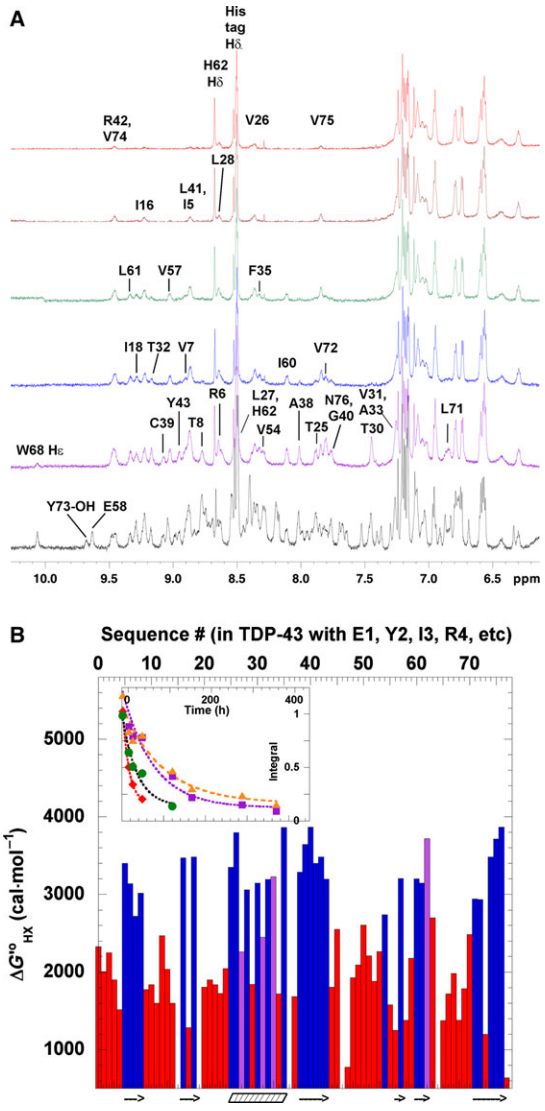

Fig. 4. Residue-level stability of the NTD measured by H/D exchange. (A) 1D ^{1}H NMR spectra of the downfield region of the NTD. The bottom spectrum (in black) was recorded in 85% H_2O/15% D_2O and the subsequent spectra were recorded after 2 h 10 m (purple), 26 h 25 m (blue), 48 h 35 m (green), 168 h 10 m (maroon) and 370 h 10 m (red). The non-exchanging side chain Hδ of His residues are also labeled. (B) Residue level conformational stability. Red bars represent upper limits of the conformational stability of fast exchanging residues. Blue bars represent the measured stability of slow exchanging residues. Purple bars represent the minimum stability of HN groups that produce peaks in a 2D NOESY spectrum recorded in D_2O (and are therefore slow exchanging) but whose exchange kinetics could not be followed because they are overlapped with non-exchanging peaks in the 1D ^{1}H spectra. Arrows and the hashed parallelogram represent the position of β-strands and the α-helix, respectively. Inset: H/D exchange of representative HN groups: T32= red diamonds; L61= green circles; I16= purple squares; V75= orange triangles. The peak integrals were normalized using non-exchanging aromatic resonances using TOPSPIN. The curves show the fit of a single exponential decay function to each data set.

and G69 could be unambiguously assigned. This confirms that these HNs have little or no protection against H/D exchange.

The NTD has a conformational stability of 3.8 kcal·mol^{-1}

The residue-level stabilities, calculated from the H/D exchange rates, are shown in Fig. 4B. The most stable residues are F35, V75 and N76, and their average stability is 3.8 ± 0.1 kcal·mol^{-1}. This value corresponds to the global conformational stability [32]. This is a fairly high conformational stability for a protein at such a low pH at 25 °C. For comparison, the conformational stabilities of remarkably stable proteins such as bovine ubiquitin and RNase A are 6.9 and 6.7 kcal·mol^{-1}, respectively, at pH 3.8 and 25 °C in low ionic strength buffer [33,34]. It is notable that the

^{1}H-^{15}N HSQC spectra for NTD remain largely unchanged over 6 months (data not shown) and that its thermal denaturation is largely reversible (see below). These are hallmarks of stable, well-folded proteins.

The N-terminal domain is stably folded over a wide range of conditions

Although most spectra were recorded at pH 3.8–4.0 and 25 °C in the presence of the Tris (2-carboxylethyl) phosphine (TCEP) reducing agent, additional 1D, 2D and 3D spectra were recorded at pH 2.0–8.6 and 5–39 °C in the absence of reducing agents. Native-like signals dominated the spectra under all of these conditions, including those of near physiological pH and temperature (Fig. 5A,B). Despite the absence of reducing agent, the Cys residues remained in their reduced state, as indicated by the ^{13}Cβ chemical shift values, and the domain remained in the monomeric state as judged by its MALDI-TOF spectra (data not shown). In addition, NMR spectral analysis and assignment of $NTD_{1–102}$ showed that residues 1–77 remain folded in the context of this longer construct, whereas residues 78–102 are disordered (Fig. 5C). By contrast, Qin *et al.* [22] observed the NTD (residues 1–80) to be less well folded in the context of a longer 1–102 residue construct.

In addition, a series of 1D ^{1}H and 2D ^{1}H-^{15}N HSQC spectra were recorded on $NTD_{1–77}$ at 5, 10, 15, 20, 25, 30, 34 and 39 °C on the pH 3.9 sample (Fig. 1). Resonances that are characteristic of the native state dominate the whole temperature range. A small population of new peaks with chemical shift values typical of denatured proteins appears at 39 °C, although this disappears when the sample is cooled. These data provide strong evidence that the native domain is folded with little or no denatured structure under physiological conditions of pH and temperature. These spectra were analyzed to determine the temperature coefficients: the change of the ^{1}HN chemical shift as the temperature is varied ($\Delta\delta^{1}H/\Delta T$). Small temperature coefficients ($|\Delta\delta^{1}H/\Delta T| < 5$ ppB·°C^{-1}) are associated with a folded structure [35]. Here, all ^{1}HN

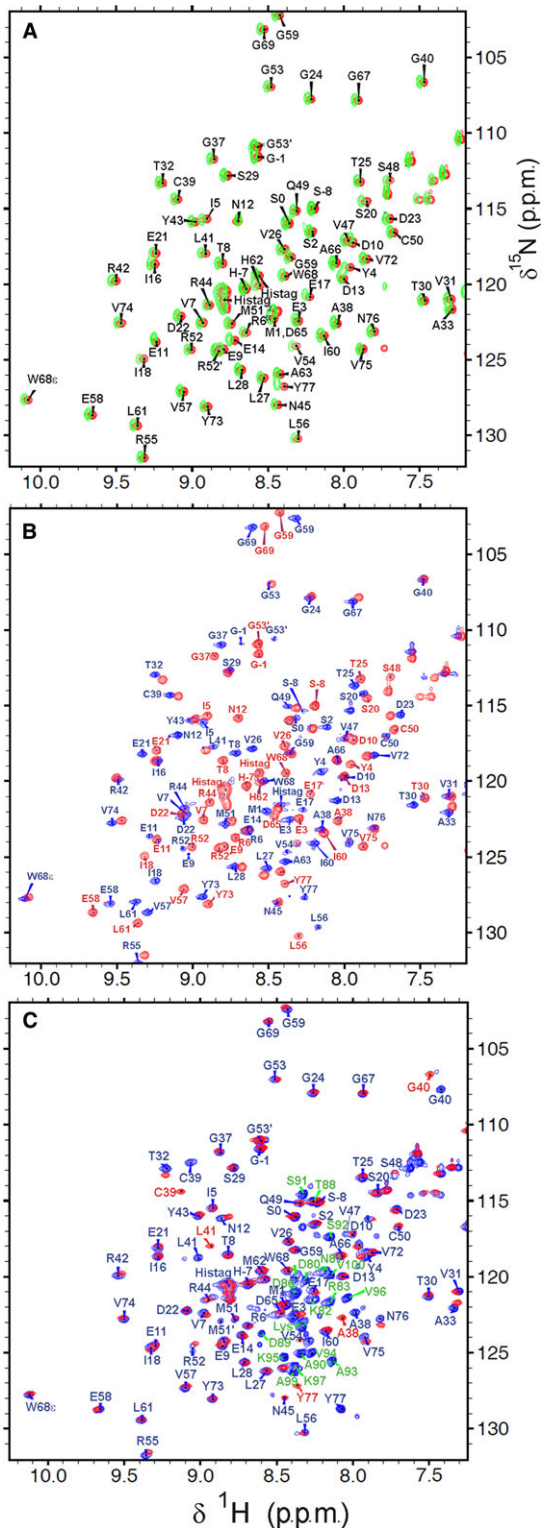

Fig. 5. The NTD is folded under a wide variety of conditions. Assigned ^{1}H-^{15}N HSQC spectra of the NTD at 25 °C. (A) At pH 4 in the presence (green) or absence (red) of the TCEP reducing agent. (B) At pH 4 (red) or pH 7 (blue) in the absence of TCEP. (C) At pH 4, in the context of the short 1–77 (red) or longer 1–102 (blue) construct, both without TCEP and with an N-terminal His tag. Peak shifts in Cys39 and nearby residues could be a result of mixed disulfide bond formation with a small molecule. Residues 78–102 in the longer construct are labeled green and they cluster in a region of the spectrum that is characteristic of disordered conformers.

showing low temperature factors are either hydrogen-bonded or buried, or both, with the sole exception of D65 (data not shown). This excellent correlation is an additional validation of the structure reported in the present study.

Fluorescence spectroscopy

The fluorescence emission spectrum of the $NTD_{1–77}$ shows a maximum at 328 nm (Fig. 6A). This is typical for Trp indole moieties in hydrophobic environments and corroborates the burial of this group within the hydrophobic core of the protein as determined by NMR (*vide supra*). Upon heating, the fluorescence emission decreases and, above 44 °C, the emission maximum red-shifts reaching values over 350 nm, which are indicative of solvent exposed Trp residues by 55 °C (Fig. 6B). The variation of the wavelength of the emission maximum shows an apparent midpoint at 50 °C at pH 4.0 and at 45 °C at pH 7.0. By fitting a two state folded/unfolded model to the emission intensity data, a T_M value of 42.1 °C and a value of 3.0 kcal·mol^{-1} for the conformational stability of the NTD at 25 °C and pH 4 can be calculated (data not shown). The difference between the stability measured by fluorescence spectroscopy and H/D exchange can be ascribed to X-Pro *cis*/*trans* isomerization, which acts to increase the stability of the unfolded forms monitored by fluorescence but not those seen by H/D exchange [32]. Similar fluorescence spectra and apparent T_M values were also observed for the longer $NTD_{1–102}$ construct (Fig. 6C). Thus, the addition of residues 78–102 does not substantially affect the conformational stability of the folded NTD.

Contribution of charged residues to the stability

The NTD contains an excess of negatively charged residues; there are only 1H + 0K + 5R = 6 cationic residues and 5D + 7E = 12 anionic residues. Most of the latter are concentrated in the turns connecting β-strand 1 to β-strand 2 and β-strand 2 to the α-helix and form two clusters of negative charge on the protein surface. To explore the contribution of electrostatic interactions to the conformational stability of the NTD, the pK_a values were predicted using the PKAPRO algorithm (Advanced Analytical Technologies, Inc., Ankeny, IA, USA). Most titratable groups were predicted to have pK_a values within 0.5 pH units of their intrinsic values (Table 2). By contrast, the pK_a values of Y43, H62 and Y73 are perturbed to favor the neutral forms as a result of unfavorable charge desolvation. E9 and E21 have moderately increased pK_a values because their charged forms are predicted to experience charge repulsion from D10 and D13, and D22 and D23, respectively. In other cases, interactions favor the charged form of the titratable groups. For example, the pK_a value for D22 is lowered because its charged form is favored as a result of hydrogen-bond formation. The pK_a value for E14 is reduced and that for R6 is raised as a result of a favorable electrostatic interaction between their charges. In a minority of structures, D23 is positioned to interact favorably with R52.

On the basis of the predicted pK_a values, the pH dependence of the conformation stability ($\partial\Delta G/\partial$pH) was calculated (Fig. 6D). This calculation predicts that the the conformational stability of NTD will be highest at acidic pH and will decrease as the pH is raised. This destabilization, which is substantial but insufficient to unfold the domain below pH 10, arises mainly from mutual repulsion of negatively charged carboxylate groups. The neutralization of H62 partially compensates because the loss its positive charge favors stability. At very alkaline conditions (pH > 10), the titration of the two buried Tyr (Y43 and Y73) is strongly destabilizing.

These results contrast with those of ubiquitin, which becomes more stable as the pH is increased [33]. This occurs because, in ubiquitin, charge–charge interactions are generally stabilizing at neutral pH [33]. However in the TDP-43 NTD, anionic residues tend to be clustered together and, overall, the charge–charge interactions are destabilizing at neutral pH.

NTD backbone dynamics determined by NMR

The heteronuclear $\{^1H\}-^{15}N$ NOE and the longitudinal (R_1) and transverse (R_2) relaxation rates were measured for 64 of the 89 residues (Fig. 7A–C), including the His-tag. The longitudinal relaxation rates show little variation along the sequence of the NTD (Fig. 7A); their mean value is 1.70 ± 0.12 s^{-1}. The transverse relaxation rates are less homogeneous (Fig. 7B); they are smaller at the N-terminus and higher in the second β-strand. The mean R_2 value is 11.0 ± 2.3 s^{-1}, and this was obtained from $R_1\rho$ measurements. The average NOE ratio was 0.76 ± 0.08 for residues 1–77 of the NTD, and many residues in secondary structural elements approach the value expected for a completely rigid HN vector at 25 °C and this magnetic field strength (Fig. 7C). There were several residues with below average NOE ratios, in particular residues V47-V54; this indicates flexibility on fast time scales (picoseconds to nanoseconds). The dynamic behavior

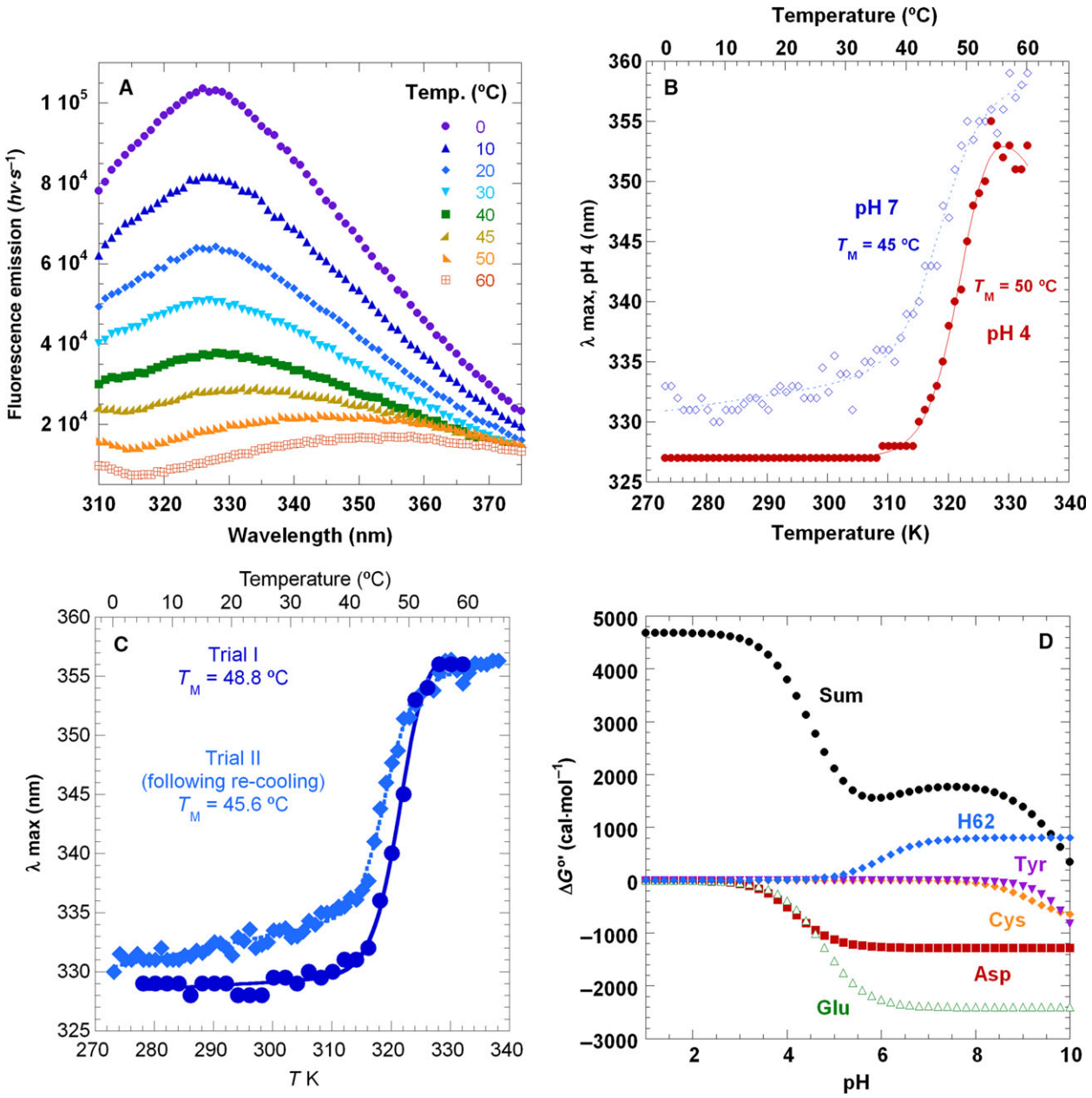

Fig. 6. Fluorescence spectra of NTD and calculation of how pH affects stability. (A) Fluorescence emission spectra of the $NTD_{1–77}$ at pH 4 recorded at temperatures ranging from 0 to 60 °C on a FluoroMax 4 instrument. The excitation wavelength was 280 nm. (B) Variation in the wavelength of maximum emission upon heating. The apparent temperature midpoints between values typical of buried and exposed Trp are indicated. (C) Variation in the wavelength of maximum emission upon heating of $NTD_{1–102}$ (a longer construct corresponding to residues 1–102 of TDP-43). The first experiment is shown in dark blue and the second experiment performed on the same sample is shown in lighter blue. The apparent temperature midpoints between values typical of buried and exposed Trp are indicated. (D) The contribution of Asp (red), Glu (green), H62 (blue), Cys (orange) and buried Tyr43 and Tyr73 residues (purple) to the pH dependence of the conformational stability. These curves were calculated using the predicted values of the pK_a values in the folded and denatured states, as outlined in the Experimental procedures. The curve in black is the sum of the conformational stability determined by H/D exchange (3.8 kcal·mol^{-1} at pH 4) plus the stabilizing/destabilizing contributions of the titratable groups as a function of pH.

of the N-terminal His-tag contrasts that of the NTD. Its average NOE ratio is much lower (0.39 ± 0.16), indicating that the His-tag is quite flexible.

The protein tumbles with an overall correlation time τ_m = 5.6 ± 0.1 ns, and behaves as an axial symmetric rotor of the oblate type (D_{par}/D_{per} = 0.82 ± 0.12). The

Table 2. Predicted pK_a values.

Titratable group	Predicted pK_a	Intrinsic pK_a	Comment
D10	4.28	3.80	
D13	4.04	3.80	
D22	**3.47**	**3.80**	Negatively charged form stabilized by hydrogen bonding to its own HN
D23	4.38	3.80	
D65	3.77	3.80	
E3	4.70	4.50	
E9	**5.23**	**4.50**	Negatively charged form experiences repulsion from charged D10 and D13
E11	4.62	4.50	
E14	4.58	4.50	
E17	**4.52**	**4.50**	Negatively charged form interacts favorably with R6 + but is destabilized as a result of partial burial
E21	**5.03**	**4.50**	Negatively charged form experiences repulsion from charged D22 and D23
E58	4.58	4.50	
H62	**5.71**	**6.30**	Charged form disfavored as a result of partial burial
C39	9.37	9.0	
C50	9.17	9.0	
Y4	10.39	10.0	
Y43	**12.69**	**10.0**	Charged form strongly disfavored as a result of partial burial
Y73	**12.51**	**10.0**	Charged form strongly disfavored as a result of partial burial
Y77	10.30	10.0	
R6	**12.80**	**12.50**	Favorable charge–charge interaction with E17
R42	12.34	12.50	
R44	12.42	12.50	
R52	12.26	12.50	
R55	12.26	12.50	

Residues with notably perturbed pK_as are shown in bold

rotational diffusion properties have been also evaluated by hydrodynamic calculations and are in good agreement with the experimental structure of a well-folded, monomeric, 77-residue domain (Fig. 3A). Thirty-three HNs have relaxation contributions from fast internal motions on the picosecond to nanosecond time scale and twelve are affected by slow exchange processes. Most of these residues are in loops or at the termini (Fig. 7D,E). Most residues in secondary structural elements or turns possess values close to one of the general order parameter S^2, indicating that these residues are quite rigid (Fig. 7F). By contrast, other segments show lower S^2 values, meaning that they are more mobile; namely, the loop connecting β-strands 1 and 2, β-strand 2, which is exposed on the edge of the β-sheet and hydrogen-bonded only on one side, the termini and the loop composed of residues R44-G53.

A disordered peptide corresponding to TDP-43 residues 78–102 can bind DNA oligos

Qin *et al.* [22] found the $NTD_{1–102}$ to be unstable and also that the folded form could be stabilized by binding of DNA oligos. A speculative model was advanced in which DNA binding might prevent the formation of pathological TDP-43 aggregates by stabilizing the folded form of the NTD. We note that the 78–102 segment, which lies outside the folded NTD and constitutes the nuclear localization segment, contains several cationic residues that could interact favorably with those DNA oligos. Because short cationic peptides and other polycations can bind DNA [36,37], we investigated whether this segment is sufficient to bind DNA oligos.

The peptide corresponding to residues 78–102 in TDP-43 ($TDP_{78–102}$) was characterized by NMR spectroscopy (2D ^{1}H-^{1}H TOSCY and NOESY; 2D ^{1}H-^{15}N and ^{1}H-^{13}C {natural abundance} HSQC spectra) at 5 °C and pH 6.5 and the essentially complete ^{15}N, ^{1}H and ^{13}Cα chemical shift assignments were obtained (Fig. 8A,B). These conditions were chosen to enhance the population of structure formation in the peptide; nevertheless, analysis of the ^{1}Hα and ^{13}Cα chemical shifts reveals that this segment is chiefly disordered, which is consistent with its behavior in the $NTD_{1–102}$ as reported both in the present study (Fig. 5D) and previously by Qin *et al.* [22].

Importantly, we found that both the DNA oligos $(TT)_6$ and $(TG)_6$ can bind to $TDP_{78–102}$ under the conditions reported by Qin *et al.* (pH 4 and 25 °C). Chemical shift displacements and/or line broadening of the H6 of thymine bases, the H8 of the guanine bases, the H1′ of guanine ribose moieties and Arg and Lys side chains of $TDP_{78–102}$ observed in the 1D ^{1}H, 2D ^{1}H-^{1}H TOSCY and NOESY spectra indicate the formation of a complex. The binding is fast on the NMR timescale and, assuming a 1 : 1 stoichiometry, would have a K_D of ≤ 150 μM.

Discussion

Several lines of evidence have recently highlighted the role played by the TDP-43 N-terminal sequence in the aggregation of this protein and the consequent loss-of-function effects at the RNA metabolic level. For this reason, it is extremely important to understand its role in the pathological process. Previous deletion

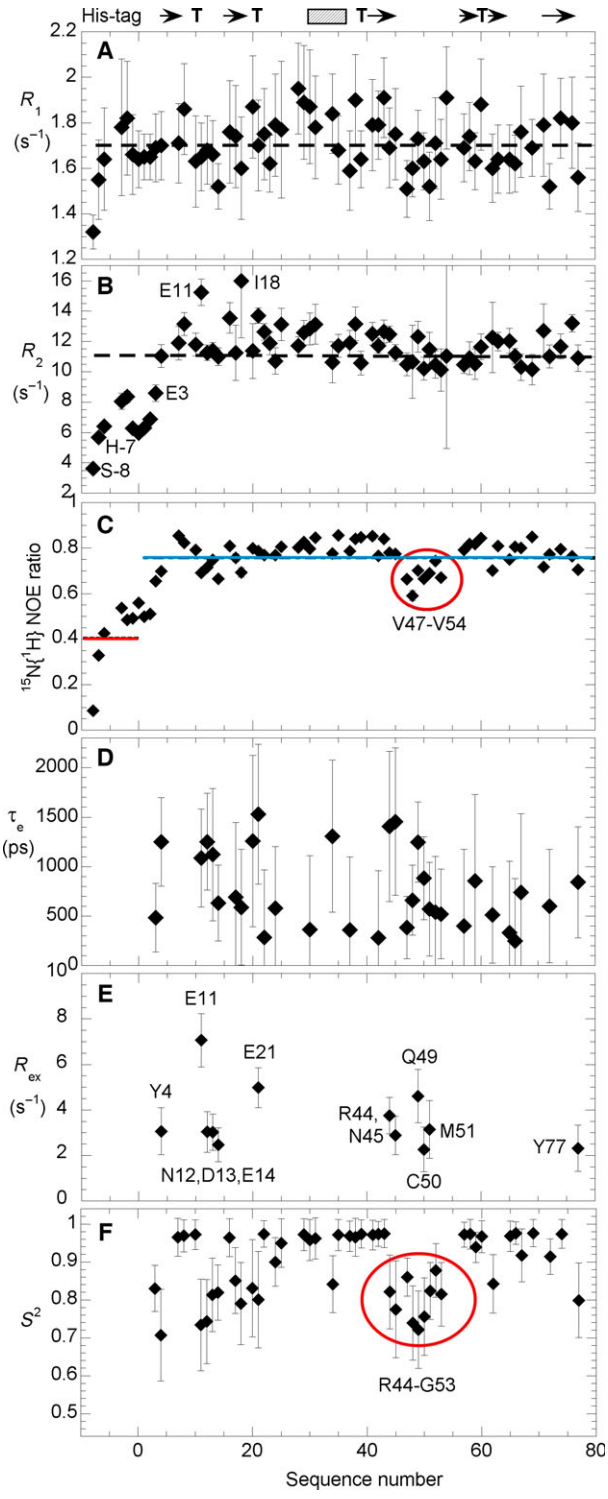

Fig. 7. NMR relaxation and NTD dynamics. The positions of the elements of secondary structure are marked by arrows (β-strands) and a hashed rectangle (α-helix), with 'T' indicating well defined turns. Note that (A) to (E) share the same *x*-axis. The dashed line in (A) and (B) indicates the mean value. Ten ^{1}H-^{15}N HSQC type spectra with different delay times were recorded and used to obtain R_1 values; another ten spectra with different delays were registered and analyzed to obtain R_2 values. In (C), the error bars reflect the experimental uncertainty as determined from the experimental signal/noise ratio. The other error bars were obtained from the uncertainties in the fitting algorithms. (A) Translational relaxation rate (R_1). (B) Transversal relaxation rate (R_2). (C) $^{15}N\{^1H\}$ NOE: the blue line is the average for the folded NTD and the red line is the mean of the flexible His-tag. A red circle highlights residues in the flexible loop. (D) Effective time constants for fast internal motions, τ_e. (E) Contributions from slow conformational exchange, R_{ex}, to the dynamics. (F) The general order parameter, S^2. Residues in the flexible loop are circled.

experiments [21,38] have suggested that structural elements could be involved in mediating the loss of function effects for TDP-43 following its aggregation within cells. For this reason, we decided to examine this connection in detail starting with a detailed characterization of its structure. Qin *et al.* [22] studied three TDP-43 NTD constructs consisting of residues 1–80, 1–102 and 10–102, all with C-terminal His-tags, and obtained a low resolution structural model. The structure of the NTD reported in the present study corrects some of the structural elements that they found, confirms others and resolves many new structural elements. Qin *et al.* [22] correctly identified most (four out of seven) of the secondary structural elements in the NTD; namely β-strands β1, β2, β3, β5 (our β6) and the α-helix. However, their predictions of the lengths of β3 and β5 (our β6) were too generous and the α-helix is significantly longer than they predicted it to be. The segment 49–52 was incorrectly identified as a β-strand by Qin *et al.* [22]. Although Q49 and C50 do donate hydrogen bonds, they do so to N45, which lies just outside the third β-strand. Thus, these hydrogen bonds serve to close a turn structure. The hydrogen bonds are too weak to afford detectable protection against H/D exchange (Fig. 5). Except for the mobile N-terminal His-tag, this segment is the most dynamic region of the domain, as revealed by both the $^{15}N\{^1H\}$NOE data of Qin *et al.* [22] and our more complete investigation of the relaxation (Fig. 7). The β-hairpin formed by residues 55–62 was not detected by Qin *et al.* [22]. This is an important element of the structure of the NTD, which is stabilized by tertiary hydrogen bonds with V26 and distinguishes the structure of the NTD from the ubiquitin fold. This β-hairpin is rigid and the β-strand formed by residues 60–62 is linked to the rest of the protein by tertiary hydrogen bonds and hydrophobic interactions (Figs 2 and 3). In addition, we have discovered and report four stable turn structures, as well as the conformation of the side chains.

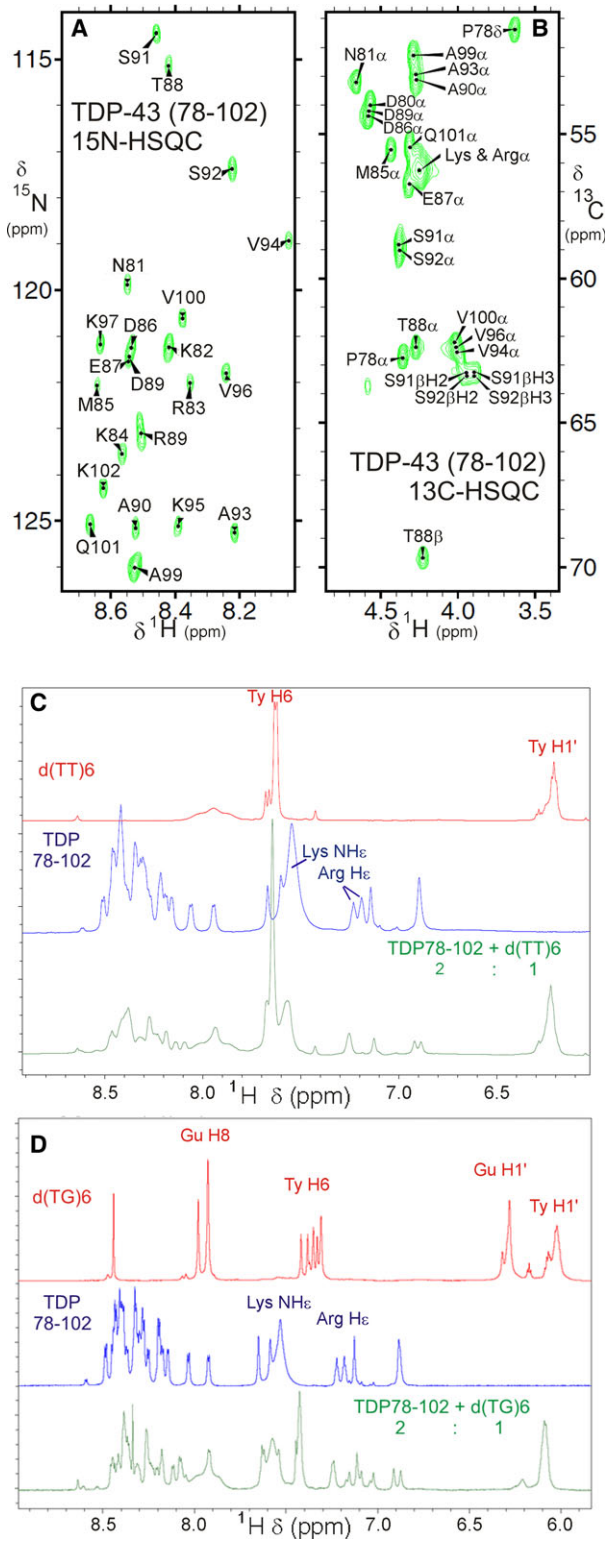

Fig. 8. $TDP_{78–102}$ binds to DNA oligos $(TT)_6$ and $(TG)_6$. (A, B) ^1H-^{15}N and ^1H-^{13}C HSQC natural abundance spectra of TDP-43(78–102) obtained at pH 6.9 and 5 °C in 10 mM K_2HPO_4. The peptide concentration was approximately 2.0 mM. The α region of the ^1H-^{13}C HSQC is shown. Resonances are labeled. (C, D) ^1H NMR spectra (downfield region) of DNA oligos $(TT)_6$ (C) and $(TG)_6$ (D) alone (red, top spectra) or mixed with polypeptide corresponding to residues 78–102 in TDP-43 (bottom, green spectra). The spectra of the polypeptide alone are shown in blue in the middle in each case. All of these spectra were recorded at 25 °C and pH 3.5–4.0 in the presence of 2 mM deuterated acetic acid. Substantial chemical shift changes are seen in the guanine H1′ and thymine H6 and Lys HNε and Arg HNε resonances and guanine H8 and thymine H6 signals are significantly broadened in the presence of the polypeptide. In addition, significant changes are seen in the HN signals of the polypeptide (between 8.5 and 8.0 ppm).

The construct of the N-terminal domain investigated in the present study contains an N-terminal HexaHis tag for affinity purification and additional residues inserted for cloning: *MRGSHHHHHHGS*-M1S 2-E3Y4R5-etc. Four lines of evidence, considered together, strongly suggest that these additional residues do not substantially perturb the conformation of the N-terminal domain. First, the His tag and its adjacent residues show a very poor chemical shift dispersion and there are no NOEs detected between this segment and the folded NTD. In the structural calculations, this segment adopts a broad ensemble of conformers; the RMSD for His tag in the final set of 20 refined structures is 3.08 Å, which is much larger than the value of 0.67 Å determined for the rest of the molecule. Second, the ^1HN temperature coefficients for the His tag are less than –6 ppB·°C^{-1}; these values are typical of an unstructured polypeptide [35]. Third, the native folded structure of the NTD is essentially the same at pH 4, where the His tag residues are predominately charged, as it is at pH 7.3, where these residues are chiefly neutral and therefore would form minimal long range charge–charge interactions with the folded NTD. Finally, the relaxation experiments show that the His tag is mobile, whereas the folded NTD is chiefly rigid.

Another remarkable difference between our findings and those reported by Qin *et al.* [22] is that we find the NTD to be stably folded, with a conformational stability of almost 4 kcal·mol^{-1} at 25 °C and pH 4, whereas they found the same domain to be in an equilibrium of folded and unfolded states under very similar experimental conditions in their 1–80 and 1–102 constructs, with the later being less stable. Our structural results do not support the proposal of Qin *et al.* [22] suggesting that the folded/unfolded equilibrium is physiologically important. In addition, the stabilization of the folded NTD conformer by nucleic acid binding observed by Qin *et al.* [22] might result from relatively unspecific favorable electrostatic interactions. Residues

1–77 of the NTD have a net charge of approximately +5 at pH 4, and the longer 1–102 construct used by Qin *et al*. [22] to study DNA binding carried a net charge of approximately +19 because the 81–102 segment contains eight additional cationic residues (6K + 2R), only five anionic residues (4D + 1E, which are largely neutral at pH 4) plus their C-terminal His-tag (6H). Our finding in the present study that a chiefly disordered peptide corresponding to residues 78–102 binds the same DNA oligos rather tightly suggests that DNA binding is not relevant for the folding and stabilization of the native NTD. Considering that the majority of proteins structurally similar to the NTD mediate protein/protein interactions, in our opinion, and at neutral pH, the NTD would be more likely to interact with cationic proteinaceous partners through its two clusters of negatively charged residues.

The high sequence conservation of the NTD of TDP-43 throughout the animal kingdom [7] strongly suggests that it plays an important physiological role. Recently, the NTD was found to be key for mediating the effective sequestering of TDP-43 into aggregative structures formed by C-terminal Q/N rich segments [21]. TDP-43 and FUS/TLS are members of the hnRNP family of proteins and both appear to form pathological aggregates implicated in ALS/FTLD, as well as functional aggregates in RNA and stress granules [39]. FUS/TLS contains multiple repeats of a motif S/G-Y-S/G that has been proposed to be key for the reversible and kinase-controlled formation of a functional amyloid 'hydrogel', which has roles in RNA regulation and transport [40]. Similar multiple S/G-Y-S/G-like motifs are also found in the C-terminal portion of TDP-43 and flank its Q/N rich motif. Although the N-terminal domain of TDP-43 appears to be unique among hnRNP proteins, it may work similar to the domains or accessory proteins that govern amyloidogenesis in functional amyloids, including curlins [41], Sup35 [42] and CPEB/Orb2A [43]. Therefore, the high-resolution structure of the NTD of TDP-43 may represent the first example of an emerging class of protein domains that regulate functional amyloid formation. For this reason, we expect that this structure and the chemical shift assignments reported in the present study will comprise valuable tools for studying these interactions between the NTD and C-terminal aggregative segments of TDP-43.

Experimental procedures

Purified and desalted DNA oligos $(TT)_6$ and $(TG)_6$ were purchased from Integrated DNA Technologies (Coralville, IA, USA). Their chemical composition was confirmed by NMR spectroscopy. A peptide corresponding to residues 78–102 of TDP-43, namely PKDNKRKMDETDASSAVKVKRAVQK, $TDP_{78–102}$, with termini blocked by acetyl and amide groups, was obtained from GenScript Corp. (Piscataway Township, NJ, USA). Its purity and identity were confirmed by HPLC, MS and NMR spectroscopy.

Sample preparation

The sequences coding for TDP-43 N-terminus residues 1–77 ($NTD_{1–77}$) or 1–102 ($NTD_{1–102}$) were cloned in the pQE306xHis plasmid using the *Bam*HI-*Hind*III restriction enzymes in accordance with standard protocols. The insert was confirmed by sequencing and the plasmid was used to transform M15 bacteria that were grown on Kana/Amp resistant plates. For expression, the bacteria were then grown in M9 media supplemented with D-glucose (U-13C, 99%; Cambridge Isotope Lab, Inc., Tewksbury, MA, USA) at 0.3% v/v and $^{15}NH_4Cl$ (Sigma, St Louis, MO, USA) at 0.1% v/v. Protein expression was induced overnight with 1 mM IPTG at 30 °C and the NTD peptide was purified using nickel-nitrilotriacetic acid Agarose Qiagen Resin in the presence of Complete Protease Inhibitor (Roche, Basel, Switzerland) in accordance with the manufacturer's instructions. The protein was eluted from the resin using 250 mM imidazole and each fraction was loaded on a 15% SDS/PAGE before staining with colloidal Coomassie Blue.

The protein was then frozen and shipped to Spain. At the Instituto de Química Física Rocasolano in Madrid, the protein samples were allowed to thaw. Then urea was added (to a final concentration of 1 M) and the pH was lowered from approximately 8.5 to 3.8 by adding HCl and CH_3COOH. The presence of urea helps dissolve the protein and prevents the formation of aggregates when the pH is near the pI of the $NTD_{1–77}$ (calculated to be 4.4 for the $NTD_{1–77}$ and 5.9 for this construct with its hexaHis tag). Next, the sample was concentrated to 1 mL using an ultrafiltration device (Amicon Corp., Danvers, MA, USA) with a 3 kDa cut-off and applied to a PD-10 gel filtration column (GE Healthcare Ltd, Little Chalfont, UK) that had been preequilibrated with buffer consisting of 85% H_2O MilliQ (Millipore, Billerica, MA, USA), 15% D_2O, 1.0 mM TCEP, as the reducing agent, 1.0 mM NaN_3, to prevent microbial contamination, and 3.0 mM $CD_3COOD/CD_3COO^-Na^+$ to buffer the pH at 3.8. Fractions of 0.5 mL were collected and the NTD-containing fractions were identified by UV absorbance or (more conveniently) fluorescence spectroscopy. MALDI-TOF MS was performed to confirm the identity of the sample. The NTD-containing fractions were concentrated to a final volume of approximately 0.25 mL and put in a water-matched NMR tube (Shigemi Inc., Allison Park, PA, USA). The concentration of the $^{13}C,^{15}N$-labeled NTD_{1-77} sample used for NMR assignment, structure and dynamics determination was 0.29 mM. For the ^{15}N-labeled NTD_{1-102} sample, the concentration was 0.86 mM. These values were determined

by UV absorbance at 280 nm using the extinction coefficient: 10 810 $cm^{-1}\cdot M^{-1}$, which was calculated on the basis of the sequence content of one Trp and four Tyr residues. NMR spectra were recorded on a 800 MHz Advance NMR spectrometer (Bruker, Billerica, MA, USA), equipped with a ^{1}H,^{13}C,^{15}N cryoprobe and Z-gradients. Although most experiments were performed at pH 3.8 and 25 °C, some NMR spectra were recorded over a range of pH (from 2 to 8.5) and temperatures (from 5 to 39 °C). Additional spectra were recorded on the buffer without protein, either alone or with excess hydrogen peroxide. By analyzing these spectra, the chemical shift values of the methylene groups of TCEP were found to be multiplets centered at 2.65 and 2.43 ppm (reduced form) and 2.50 and 2.08 ppm (oxidized form) at pH 3.8 and 25 °C. These values and 1D ^{1}H NMR spectra were utilized to confirm the presence of reduced TCEP in a sample during the course of acquiring 2D and 3D NMR spectra.

Chemical shift referencing

Because preliminary spectra revealed that the NTD has NMR signals near 0 ppm, the standard NMR chemical shift reference compound sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS) was not included in the majority of the samples. Instead, the chemical shift reference was set by measuring the DSS signal in buffer. A series of 1D ^{1}H experiments with this reference sample were recorded at several different temperatures to determine the correct reference value for the experiments used to measure the ^{1}HN temperature coefficients. The ^{13}C and ^{15}N chemical shift reference values were calculated from the ^{1}H reference by multiplying by nuclei's gyromagnetic ratios [44].

Chemical shift assignment

The assignment of TDP-43 was carried out successfully using a standard 3D approach based on sequentially connectivities between ^{1}H, ^{15}N nuclei and the $^{13}C\alpha$ and $^{13}C\beta$ of the same residue and the preceding one. This assignment process was corroborated by 2D ^{1}H-^{15}N HSQC spectra, which are ^{13}C filtered according to the side chain structure, as well as spectra that specifically detect residues proceeding or following proline. The backbone assignment process was carried out automatically using MARS [45], as well as manually using by two different operators with the aid of SPARKY [46] and CCPNMR [47]. Only a few differences were observed and these could be resolved by careful study and analysis of additional spectra. Additional verification was obtained by registering and analysis of a (2,4)D $^{1}H^{15}N$-$^{1}H^{15}N$ spectrum. Side chain chemical shifts were assigned by studying 3D HCCCNH, CCCCNH 'TOSCY-HSQC') type spectra. These assignments were checked and a Tyr-O^{1}H signal was assigned by analysis of 1D ^{1}H and 2D ^{1}H-^{1}H NOESY spectra recorded in 85%/15% H_2O/D_2O or 100% D_2O. The complete list of the NMR spectra used to assign and obtain structural constraints is provided in Table S1.

Structure calculation

The solution structure of $NTD_{1–77}$ was determined on the basis of backbone torsional angle restrictions derived from NMR parameters and NOE distance constraints. Restraints for φ and ψ dihedral angles were derived from chemical shifts using TALOS+ [48]. Distance constraints were obtained from 3D NOESY-^{13}C-^{1}H HSQC and 2D ^{1}H-^{1}H NOESY spectra. The later was acquired in 100% D_2O to achieve the most restrictive restraints. Structures were calculated following a three-step procedure. First, automatic NOE assignment of the spectra was carried out as implemented in CYANA, version 2.1 [49]. This protocol performs seven cycles of automated NOE assignment and structure calculation and a final annealing using the list of restraints obtained in the last cycle. This set of distance constraints was checked by visual inspection of the spectra. In a second step, we added a number of hydrogen-bond restrictions that were obtained from an H/D exchange experiment (see below) and we performed a standard simulated annealing to obtain 100 conformers from which those 20 with lowest target function values were selected. Finally, we used AMBER9 [50] to refine this ensemble of structures. Accordingly, 2000 steps of energy minimization with implicit solvation were performed over the selected conformers. The quality of these final structures was assessed using the Protein Data Bank validation server (http://validate.rcsb.org).

Structure comparison

The NTD structure was manually compared with known folds in the SCOP structural database [51] and the whole Protein Data Bank was surveyed automatically using the DALI server [52].

H/D exchange

To measure H/D exchange kinetics, a PD-10 column was utilized to exchange the NTD into 3.0 mM deuterated acetic acid/sodium acetate, 1.0 mM sodium azide and 1.0 mM TCEP buffer prepared with 100% D_2O. Then, a series of eleven 1D ^{1}H spectra were recorded to follow exchange over the course of 2 weeks. After approximately 12 h of exchange, a 2D ^{1}H NOESY spectrum was recorded. The analysis of this spectrum served to corroborate the identity of the slow exchanging HN groups, especially those that overlapped with other exchanging peaks or non-exchanging aromatic resonances. Between measurements, the sample was incubated in a water bath at 25 °C. The temperature and water level of the bath was checked every 2 or 3 days and the temperature variation was less than ±0.5 °C. The HN peaks and non-exchanging aromatic signals, which

were used as internal controls, were integrated using TOPSPIN, version 2.1 (Bruker) and a single exponential decay equation was fit to the data to determine the observed exchange rate for each residue. By comparison with the intrinsic exchange rates calculated using parameters reported previously [53], the conformational stability for each residue and the global stability were determined by assuming the EXII exchange mechanism [54] (which holds for low pH conditions) and in accordance with the approach of Huyghues-Despointes *et al.* [32]. After H/D exchange was complete, the sample was back-exchanged into 85% H_2O/15% D_2O buffer and a 2D ^{1}H-^{1}H NOESY spectrum was recorded to corroborate the identity of the fast exchanging HN groups.

Fluorescence spectroscopy

A Fluoromax 4 spectroflourimeter (Jobin-Yvon Inc., Edison, NJ, USA) equipped with a Peltier module for temperature control was used to record fluorescence spectra on the NTD dissolved in the same buffer (85% H_2O/15% D_2O) employed for NMR spectroscopy. Excitation and emission wavelengths were calibrated using the Xe emission line at 467 nm and the water Raman signal, respectively. The excitation wavelength was 280 nm and the emission was recorded over the range 300–400 nm with a 2 ns·s^{-1} scan speed at 25 °C. The protein concentration was 0.02–0.04 mM. For monitoring thermal denaturation, the fluorescence spectra were recorded over intervals of 1 °C. Once the temperature was raised, the sample was allowed to equilibrate for 1 min prior to the acquisition of a fluorescence spectrum over the wavelength range 310–375 nm. The slit widths were 3.0 nm for excitation and emission in all spectra. A two-state denaturation model with linearly sloping pre- and post-transition baselines was fit to the data to obtain values for T_M (the temperature midpoint of thermal denaturation), as well as ΔH_M and ΔS_M, the enthalpy and entropy changes for denaturation, respectively. The Gibbs–Helmholtz equation was employed to estimate the conformational free energy at 25 °C, $\Delta G^{\circ\prime}$ (25 °C):

$$\Delta G^{\circ\prime}(25\,^{\circ}\mathrm{C}) = \Delta H^{\circ\prime}{}_M[1 - 298.15\,\mathrm{K}/T_M] - \Delta Cp[T_M - 298.15) - 298.15 * \ln(298.15/T_M)] \quad (1)$$

using a value of 1132 cal·M^{-1}·K^{-1} for ΔCp, the heat capacity change for denaturation, which was estimated using the empirical equations of Myers *et al.* [55] for a 77-residue globular protein.

pH dependence of the conformational stability

Using the refined structures of the NTD as input, we utilized PROPKA [56] to estimate pK_a values for the domain's titratable groups in the folded state (*Ka,j,N*). These values, and standard values for the pK_a values of the same groups in the denaturated state (*Ka,j,D*) were then used to calculate the dependence of the conformational stability on pH ($\partial\Delta G/\partial$pH) as described previously [57] using the equation reported by Tanford [58]:

$$\partial\Delta G^{\circ\prime}/\partial\mathrm{pH} = -\,\mathrm{RT}\,\ln[\Pi_{j=1}^{n}(1 + K_{a,j,D}/a\mathrm{H}^+)/\Pi_{j=1}^{n}(1 + K_{a,j,N}/a\mathrm{H}^+)]$$

where $a\mathrm{H}^+$ is the activity of the hydrogen ion, which is assumed to be equal to the hydrogen ion concentration: $a\mathrm{H}^+ \approx [\mathrm{H}^+] = 10^{(-\mathrm{pH})}$. This is an excellent approximation for solutions of low ionic strength above pH 2.

Dynamics: ^{15}N relaxation

All NMR experiments were acquired at 800 MHz, pH 3.8 and 25 °C. A set of ^{1}H-^{15}N correlation spectra were registered to measure the ^{15}N longitudinal (R_1), off-resonance rotating frame ($R_1\rho$) and transverse (R_2) relaxation rates. Three sets of 10 spectra with relaxation delays ranging from 20 to 1600, 8 to 200 and 0 to 285 ms were recorded to determine R_1, $R_1\rho$ and R_2, respectively. Relaxation times were calculated via least-squares fitting of peak intensities to a two-parameter exponential function. The heteronuclear ^{15}N{^{1}H} NOEs of individual backbone HN groups were measured as the ratio of spectra recorded with and without saturation in an interleaved fashion. A long recycling delay of 10 s was employed to ensure the maximal development of NOEs before acquisition and to allow solvent relaxation, thus avoiding transfer of saturation to the most exposed amide protons of the protein between scans [59]. Uncertainties in peak heights were determined from the SD of the distribution of intensities in the region of the HSQC spectra where no signal and only noise was observed. The principal components of the TDP-43(1–77) NTD inertia tensor were calculated with PDBINERTIA (A. G. Palmer III, Columbia University, New York, NY, USA) using the lowest energy structure. Heteronuclear R_2 values were derived from $R_1\rho$ measurements. We estimated the overall correlation time from the ratio of the mean values of R_1 and R_2. The values for R_1 and R_2 were calculated from a subset of residues with little internal motion and no significant exchange broadening. This subset excluded residues with NOEs < 0.65 and also residues with R_2 values lower than the average minus 1 SD, unless their corresponding R_1 values were larger than the average plus 1 SD [60]. The diffusion tensor, which describes rotational diffusion anisotropy, was determined by two approaches [61,62], as implemented in R2R1_DIFFUSION and QUADRIC_DIFFUSION by A. G. Palmer (A. G. Palmer III, Columbia University). The ^{15}N relaxation was analyzed assuming dipolar coupling with the directly attached proton

(with a bond length of 1.02 Å) and a contribution from the ^{15}N chemical shift anisotropy evaluated as −160 ppm. Relaxation data were fitted to the Lipari and Szabo model [63] using FAST-MODELFREE [64], which interfaces with MODELFREE, version 4.2 [65]. Five models of internal motion were evaluated for each amide $^{1}H-^{15}N$ pair: (a) S^2; (b) S^2 and τe; (c) S^2 and R_{ex}; (d) S^2, τ_e and R_{ex}; and (e) S_f^2, S^2 and τe, where S^2 is the generalized order parameter, τ_e is the effective internal correlation time, R_{ex} is the exchange contribution to transverse relaxation and S_f^2 is related to the amplitude of the fast internal motions.

$TDP_{78–102}$/DNA oligo binding

The resonances of $TDP_{78–102}$ were assigned by analysis of 2D ^{1}H-^{1}H TOCSY and NOESY spectra and 2D ^{1}H-^{13}C and ^{1}H-^{15}N HSQC spectra recorded at natural abundance at pH 6.9 and 5 °C. The concentrations of the $TDP_{78–102}$ and DNA oligo were approximately 180 and 90 μM, respectively; this molar ratio of $TDP_{78–102}$ to oligo was chosen so that the number of positive charges on the peptide is approximately equivalent to the number of negative charges on the oligo. In the case of the $TDP_{78–102}/(TG)_6$ sample, 2D ^{1}H-^{1}H TOCSY and NOESY spectra were recorded on the mixtures to corroborate the assignment of the resonances that shift and broaden upon complex formation. These experiments were carried out at 25 °C and pH 4 to match the conditions employed by Qin *et al.* [22].

Acknowledgements

This work was supported by Grants CTQ2010-21567-C02-02 (FPI Fellowship to MMG), SAF2013-49179-C2-2-R (DVL), EU Joint Programme-Neurodegenerative Diseases (FPND) grants RiMod-FTD, Ministerio della Sanita', Italia (EB) and AC14/00037 (DVL), AriSLA TARMA (FB) and Theirry Latran Foundation REHNPALS (EB). We thank Professor C. González for help with $NTD_{1–77}$ structure refinement and revision of the NMR spectra of the $d(TT)_6$ and $d(TG)_6$ oligos.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Author contributions

VR and CS produced and purified protein samples. MM, DPU and DVL performed and analyzed biophysical and spectroscopic experiments. FB, EB and DVL designed the research and wrote the manuscript with contributions from VR, CS, MM and DPU. All the authors approved the final version of the manuscript submitted for publication.

References

1 Ou SH, Wu F, Harrich D, García-Martínez LF & Gaynor RB (1995) Cloning and characterization of a novel cellular protein, TDP-43, that binds to human immunodeficiency virus type 1 TAR DNA sequence motifs. *J Virol* **69**, 3584–3596.

2 Buratti E, Dörk T, Zuccato E, Pagani F, Romano M & Baralle FE (2001) Nuclear factor TDP-43 and SR proteins promote *in vitro* and *in vivo* CFTR exon 9 skipping. *EMBO J* **20**, 1774–1784.

3 Buratti E & Baralle FE (2012) TDP-43: gumming up neurons through protein-protein and protein-RNA interactions. *Trends Biochem Sci* **37**, 237–247.

4 Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou RT, Bruce J, Schuck T, Grossman M, Clark CM *et al.* (2006) Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133.

5 Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, Mann D, Tsuchiya K, Yoshida M, Hashizume Y *et al.* (2006) 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.

6 Josephs KA, Whitwell JL, Weigand SD, Murray ME, Tosakulwong N, Liesinger AM, Petrucelli L, Senjem ML, Knopman DS, Boeve BF *et al.* (2014) TDP-43 is a key player in the clinical features associated with Alzheimer's disease. *Acta Neuropathol* **127**, 811–824.

7 Ayala YM, Pantano S, D'Ambrogio A, Buratti E, Brindisi A, Marchetti C, Romano M & Baralle FE (2005) Human, *Drosophila* and *C. elegans* TDP-43: nucleic acid binding properties and splicing regulatory function. *J Mol Biol* **348**, 575–588.

8 Guo W, Chen Y, Zhou X, Kar A, Ray P, Chen X, Rao EJ, Yang M, Ye H, Zhu L *et al.* (2011) An ALS-associated mutation affecting TDP-43 enhances protein aggregation, fibril formation and neurotoxicity. *Nat Struct Mol Biol* **18**, 822–830.

9 Robinson JL, Geser F, Stieber A, Umoh M, Kwong LK, vanDeerlin VM, Lee VMY & Trojanowshi JQ (2013) TDP-43 skeins show properties of amyloid in a subset of ALS cases. *Acta Neuropathol* **125**, 121–131.

10 Bigio EH, Wu JY, Deng HX, Bit-Ivan EN, Mao Q, Ganti R, Peterson M, Siddique N, Geula C, Siddique T *et al.* (2013) Inclusions in frontotemporal lobar degeneration with TDP-43 proteinopathy (FTLD-TDP) and amyotrophic lateral sclerosis (ALS), but not FTLD with FUS proteinopathy (FTLD-FUS) have properties of amyloid. *Acta Neuropathol* **125**, 463–465.

11 Chen AK, Lin RY, Hsieh EZ, Tu PH, Chen RP, Liao TY, Chen W, Wang CH & Huang JJ (2010) Induction of amyloid fibrils by the C-terminal fragments of TDP-43 in amyotrophic lateral sclerosis. *J Am Chem Soc* **132**, 1186–1187.
12 Saini A & Chauhan VS (2011) Delineation of the core aggregation sequences of TDP-43 C-terminal fragment. *ChemBioChem* **12**, 2495–2501.
13 Saini A & Chauhan VS (2014) Self-assembling properties of peptides derived from TDP-43 C-terminal fragment. *Langmuir* **30**, 3845–3856.
14 Jiang LL, Che MX, Zhao J, Zhou CJ, Xie MY, Li HY, He JH & Hu HY (2013) Structural transformation of the amyloidogenic core region of TDP-43 protein initiates its aggregation and cytoplasmic inclusion. *J Biol Chem* **288**, 19614–19624.
15 Budini M, Buratti E, Stuani C, Guarnaccia C, Romano V, De Conti L & Baralle FE (2012) Cellular model of TAR DNA-binding protein 43 (TDP-43) aggregation based on its C-terminal Gln/Asn-rich region. *J Biol Chem* **287**, 7512–7525.
16 Mompean M, Buratti E, Guarnaccia C, Brito RM, Chakrabartty A, Baralle FE & Laurents DV (2014) Structural characteriation of the minimal segment of TDP-43 competent for aggregation. *Arch Biochem Biophys* **545**, 53–62.
17 Mompeán M, Hervás R, Xu Y, Tran TH, Guarnaccia C, Buratti E, Baralle F, Tong L, Carrión-Vázquez M, McDermott A *et al.* (2015) Structural evidence of amyloid fibril formation in the putative aggregation domain of TDP-43. *J Phys Chem Lett* **6**, 2608–2615.
18 He RY, Huan YC, Chiang CW, Tsai YJ, Ye TJ, Gao HD, Wu CY, Lee HM & Huang JJ (2015) Characterization and real-time imaging of the FTLD-related protein aggregation induced by amyloidogenic peptides. *Chem Commun (Camb)* **51**, 8652–8655.
19 Chang CK, Wu TH, Wu CY, Chiang MH, Toh EK, Hsu YC, Lin KF, Liao YH, Huang TH & Huang JJ (2012) The N-terminus of TDP-43 promotes its oligomerization and enhances DNA binding affinity. *Biochem Biophys Res Commun* **425**, 219–224.
20 Budini M, Romano V, Quadri Z, Buratti E & Baralle FE (2015) TDP-43 loss of cellular function through aggregation requires additional structural determinants beyond its C-terminal Q/N prion-like domain. *Hum Mol Genet* **24**, 9–20.
21 Romano V, Quadri Z, Baralle FE & Buratti E (2015) The structural integrity of TDP-43 N-terminus is required for efficient aggregate entrapment and consequent loss of protein function. *Prion* **9**, 1–9.
22 Qin H, Lim LZ, Wei Y & Song J (2014) TDP-43 N-terminus encodes a novel ubiquitin-like fold and its unfolded form in equilibrium that can be shifted by binding to ssDNA. *Proc Natl Acad Sci USA* **111**, 18619–18624.
23 Briggs MS & Roder H (1992) Early hydrogen-bonding events in the folding reaction of ubiquitin. *Proc Natl Acad Sci USA* **89**, 2017–2021.
24 Lazar GA, Desjarlais JR & Handel TM (1997) *De novo* design of the hydrophobic core of ubquitin. *Protein Sci* **6**, 1167–1178.
25 Hida T, Nakamura F, Usui H, Takeuchi K, Yamashita N & Goshima Y (2015) Semaphorin3A-induced axonal transport mediated through phosphorylation of Axin-1 by GSK3b. *Brain Res* **1598**, 46–56.
26 Subbiah S, Laurents DV & Levitt M (1993) Structural similarity of DNA-binding domains of bacteriophage repressors and the globin core. *Curr Biol* **3**, 141–148.
27 Schubert M, Labudde D, Oschkinat H & Schmieder P (2002) A software tool for the prediction of Xaa-Pro peptide bond conformation in proteins based on ^{13}C chemical shift statistics. *J Biomol NMR* **24**, 149–154.
28 Cohen TJ, Hwang AW, Unger T, Trojanowski JQ & Lee VM (2011) Redox signalling directly regulates TDP-43 via cystein oxidation and disulphide cross-linking. *EMBO J* **1–12**, 1–12.
29 Chang CK, Chiang MH, Toh EK, Chang CF & Huang TH (2013) Molecular mechanism of oxidation-induced TDP-43 RRM1 aggregation and loss of function. *FEBS Lett* **587**, 575–582.
30 Schanda P, Forge V & Brutscher B (2007) Protein folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc Natl Acad Sci USA* **104**, 11257–11262.
31 Laurents DV, Bruix M, Jiménez MA, Santoro J, Boix E, Moussaoui M, Nogués M & Rico M (2009) The (1) H, (13)C, (15)N resonance assignment, solution structure and residue level stability of eosinophil cationic protein/RNase 3 determined by NMR spectroscopy. *Biopolymers* **91**, 1018–1028.
32 Huyghues-Despointes BMP, Scholtz JM & Pace CN (1999) Protein conformational stabilities can be determined from hydrogen exchange rates. *Nat Struct Biol* **6**, 910–912.
33 Ibarra-Molero B, Loladze VV, Makhatadze GI & Sanchez-Ruiz JM (1999) Thermal versus guanidine-induced unfolding of ubiquitin. An analysis in terms of the contributions from charge-charge interactions to protein stability. *Biochemistry* **38**, 8138–8149.
34 Pace CN, Laurents DV & Thomson J (1990) pH dependence of the urea and guanidine hydrochloride denaturation of RNase A and RNase T1. *Biochemistry* **29**, 2564–2572.
35 Baxter NJ & Williamson MP (1997) Temperature dependence of 1H chemical shifts in proteins. *J Biomol NMR* **9**, 359–369.
36 Pontius BW & Berg P (1991) Rapid renaturation of complementary DNA strands modified by cationic detergents: a role for high probability binding domains

in enhancing the kinetics of molecular assembly processes. *Proc Natl Acad Sci USA* **88**, 8237–8241.
37 Diez-García F, González C, Chakrabartty A & Laurents DV (2012) An Arg-rich putative prebiotic protein is as stable as its Lys-rich variant. *Arch Biochem Biophys* **528**, 118–126.
38 Zhang YJ, Caulfield T, Xu YF, Gendron TF, Hubbard J, Stetler C, Sasaguri H, Whitelaw EC, Cai S, Lee WC *et al.* (2013) 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.
39 Bentmann E, Neumann M, Tahirovic S, Rodde R, Dormann D & Haass C (2012) Requirements for stress granule recruitment of fused in sarcoma (FUS) and TAR DNA binding protein of 43 kDa (TDP-43). *J Biol Chem* **287**, 23079–23094.
40 Kato M, Han TW, Xie S, Shi K, Du X, Wu LC, Mirzaei H, Goldsmith EJ, Longgood J, Pei J *et al.* (2012) Cell-free formation of RNA Granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767.
41 Wang X, Zhou Y, Ren JJ, Hammers ND & Chapman MR (2010) Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fibril biogenesis. *Proc Natl Acad Sci USA* **107**, 163–168.
42 Shorter J & Lindquist S (2005) Prions as adaptive conduits of memory and inheritance. *Nat Rev Genet* **6**, 435–450.
43 Hervás R, Li L, Majumdar A, Fernández-Ramírez MdC, Unruh J, Slaughter B, Galera-Prat A, Santana E, Suzuki M, Nagai Y *et al.* (2016) Molecular basis of Orb2 amyloidogenesis and blockade of memory consolidation. *PLoS Biol* **14**, e1002361.
44 Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE & Wüthrich K (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J Mol Biol* **280**, 933–952.
45 Jung YS & Zweckstetter M (2004) Mars-robust automatic backbone assignment of protein. *J Biomol NMR* **30**, 11–23.
46 Goddard TD & Kneller DG (2008) Sparky, 3rd edn. University of California, San Francisco.
47 Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J & Laue ED (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–698.
48 Shen Y, Delaglio F, Cornilescu G & Bax A (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213–233.
49 Güntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* **278**, 353–378.
50 Case D, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KJ, Onufriev A, Simmerling C, Wang B & Woods RJ (2005) The Amber biomolecular simulation programs. *J Comput Chem* **26**, 1668–1688.
51 Andreeva A, Howorth D, Chandonia JM, Brenner SE, Hubbard TJP, Chothia C & Murzin AC (2007) Data growth and its impact on the SCOP database: new developments. *Nucleic Acids Res* **36**, D419–D425.
52 Hasegawa H & Holm L (2009) Advances and pitfalls of protein structural alignment. *Curr Opin Struct Biol* **19**, 341–348.
53 Bai Y, Milne JS, Mayne L & Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75–86.
54 Hvidt A & Nielsen SO (1966) Hydrogen exchange in proteins. *Adv Protein Chem* **21**, 287–386.
55 Myers JK, Pace CN & Scholtz JM (1995) Denaturant m-values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci* **4**, 2138–2148.
56 Li H, Robertson AD & Jensen JH (2005) Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **61**, 704–721.
57 García Mayoral MF, Martínez del Pozo A, Campos-Olivas R, Gavilanes JG, Santoro J, Rico M, Laurents DV & Bruix M (2006) pH-dependent conformational stability of the ribotoxin alpha-sarcin and four active site charge substitution variants. *Biochemistry* **45**, 13705–13718.
58 Tanford C (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Adv Protein Chem* **24**, 1–95.
59 Renner C, Schleicher M, Moroder L & Holak TA (2002) Practical aspects of the 2D 15N-[1H]-NOE experiment. *J Biomol NMR* **23**, 23–33.
60 Pawley NH, Wang C, Koide S & Nicholson LK (2001) An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of 15N relaxation parameters. *J Biomol NMR* **20**, 149–165.
61 Tjandra N, Feller SE, Pastor RW & Bax A (1995) Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation. *J Am Chem Soc* **117**, 12562–12566.
62 Brüschweiler R, Liao X & Wright PE (1995) Long-range motional restrictions in a multidomain zinc-finger protein from anisotropic tumbling. *Science* **268**, 886–889.
63 Lipari G & Szabo A (1982) Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. *J Am Chem Soc* **104**, 4546–4570.
64 Cole R & Loria JP (2003) FAST-Modelfree: a program for rapid automated analysis of solution NMR spin-relaxation data. *J Biomol NMR* **26**, 203–213.

65 Palmer AG 3rd, Rance M & Wright PE (1991) Intramolecular motions of a zinc finger DNA-binding domain from Xfin characterized by proton-detected natural abundance carbon-13 heteronuclear NMR spectroscopy. *J Am Chem Soc* **113**, 4371–4380.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:
Table S1. Complete list of the NMR spectra used to assign and obtain structural constraints.