Supporting Information for

The folding and aggregation energy landscapes of tethered RRM domains of human TDP-43 are coupled *via* a metastable molten globule-like oligomer

Meenakshi Pillai and Santosh Kumar Jha*

Physical and Materials Chemistry Division,

Academy of Scientific and Innovative Research (AcSIR),

CSIR-National Chemical Laboratory,

Dr. Homi Bhabha Road,

Pune 411008, Maharashtra, India

*To whom correspondence should be addressed:

Santosh Kumar Jha; E-mail: sk.jha@ncl.res.in

Telephone: 91-20-25902588; Fax: 91-20-25902615

Supporting Figures:

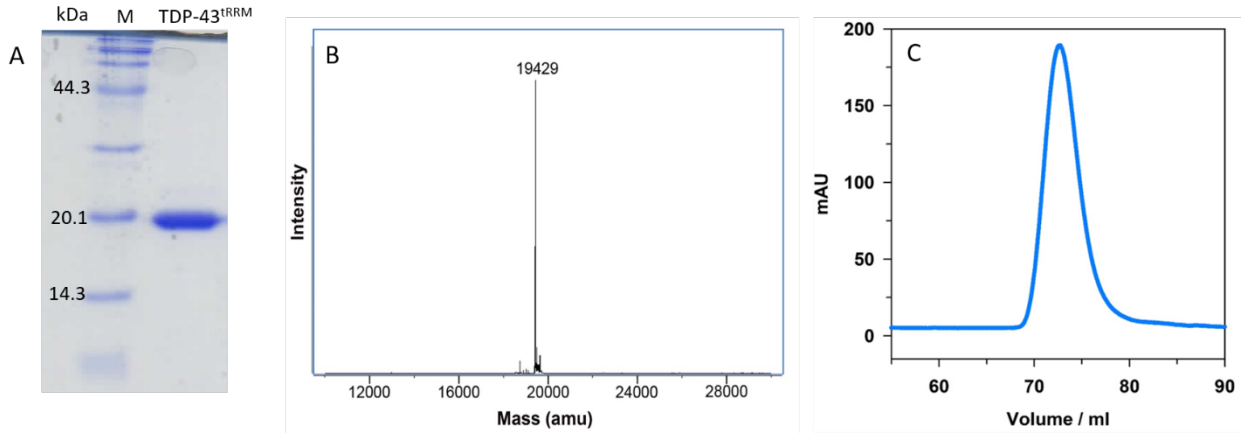

Figure S1: Recombinantly purified TDP-43^{tRRM} is pure and monomeric. Panel A shows TDP-43^{tRRM} protein as a single band near 20.1 kDa marker (M) in 20% SDS-PAGE. Panel B shows the electrospray ionization mass spectrum of TDP-43^{tRRM}. The average molecular mass of the protein is 19,429 Da corresponding to the mass of the His_6 tag cleaved TDP-43^{tRRM}. Panel C shows the size exclusion chromatography of the N form on a HiLoad™ 16/600 Superdex 75 pg gel filtration column. The protein elutes at around 72 mL, the mass corresponding to approximately 19 kDa.

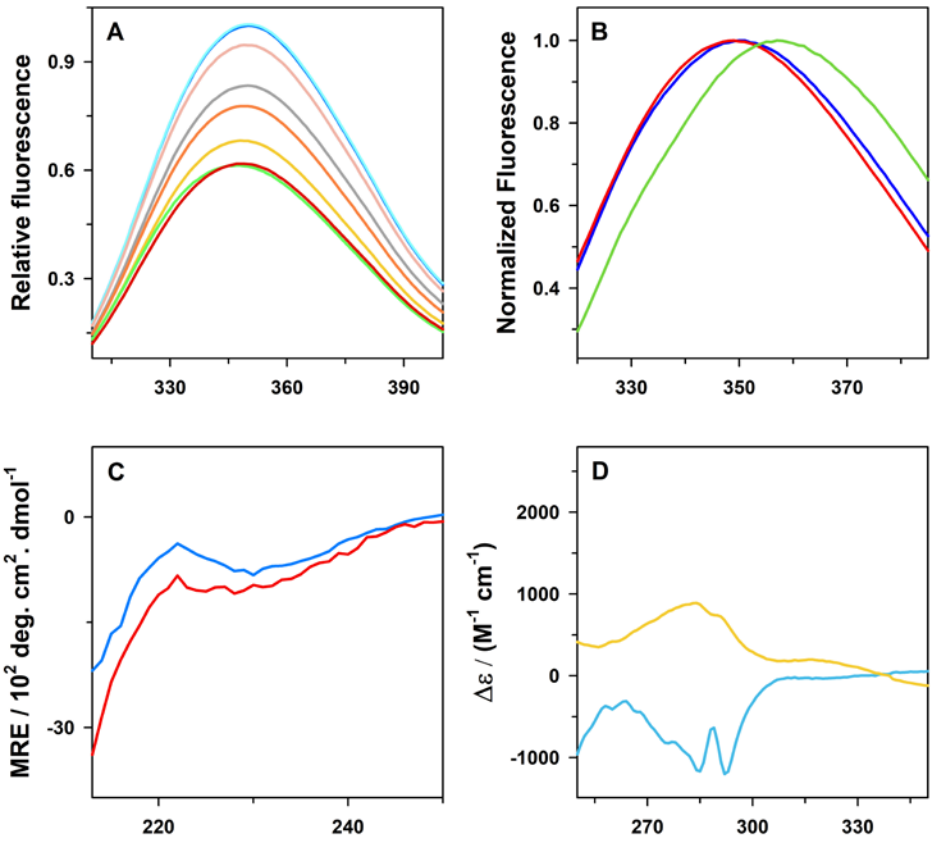

Figure S2: A form has a lower quantum yield, an N-like solvation and a different molecular arrangement of aromatic amino acids compared to the N form. Panel A shows the fluorescence emission spectra of tryptophan residues at different pH: 7.5 (blue), 7.3 (cyan), 6.9 (peach), 6.5 (gray), 5.5 (orange), 4.5 (yellow), 4.1 (green), 3.5 (red). Panel B shows the fluorescence emission spectra of the N form (blue), the A form (red) and the U form (green) at pH 7. The fluorescence values at the wavelength of the maximum emission have been normalized to 1 for comparison. Panel C shows the far-UV CD spectra of the unfolded forms at pH 7 (blue) and pH 3 (red). In panel D, the blue line shows the difference in the absorbance spectrum of the N form and the A form. The yellow line shows the difference in the absorbance spectrum of the unfolded forms at pH 7 and pH 3.

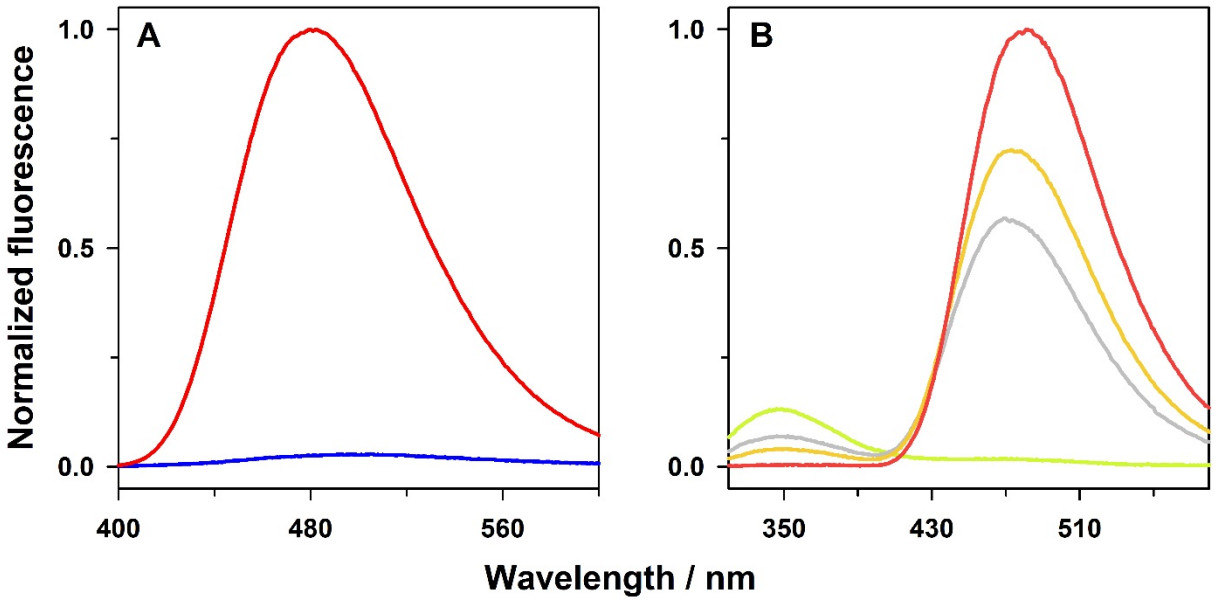

Figure S3. The A form has a loosely packed hydrophobic core and the ANS binding site is located close to the tryptophan residues. Panel A shows the fluorescence emission spectra of ANS upon excitation at 380 nm when bound to the A form (red) as compared to the N form (blue). Panel B shows FRET between the tryptophan residues of TDP-43^{tRRM} and ANS bound to the A form upon excitation at 295 nm. The green line indicates the fluorescence spectrum of the A form in the absence of ANS. The gray, yellow and red lines show the fluorescence spectra of the A form (2 µM) in the presence of 10 µM, 20 µM and 80 µM ANS, respectively.

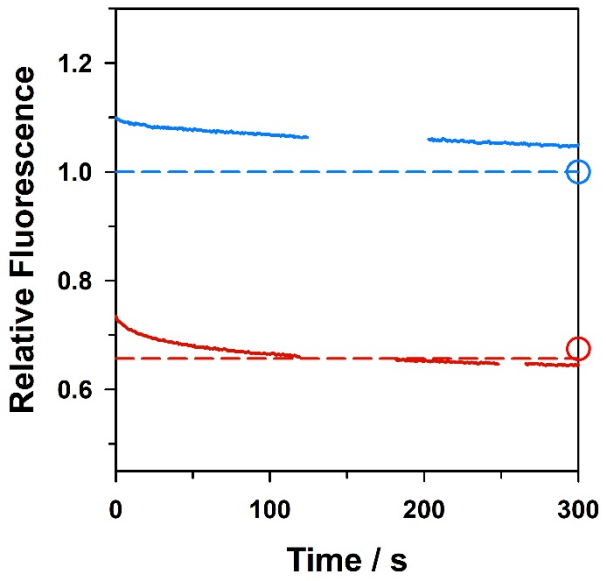

Figure S4. The N⇌A transition is fast and reversible. The dashed lines represent the reference fluorescence signal for the N form (blue) and the A form (red). The solid lines represent the kinetic signal upon transferring the A form to pH 7 (blue) and the N form to pH 3 (red). The circles are equilibrium fluorescence signals of the N form (blue) and the A form (red) taken after 1 h of incubation.

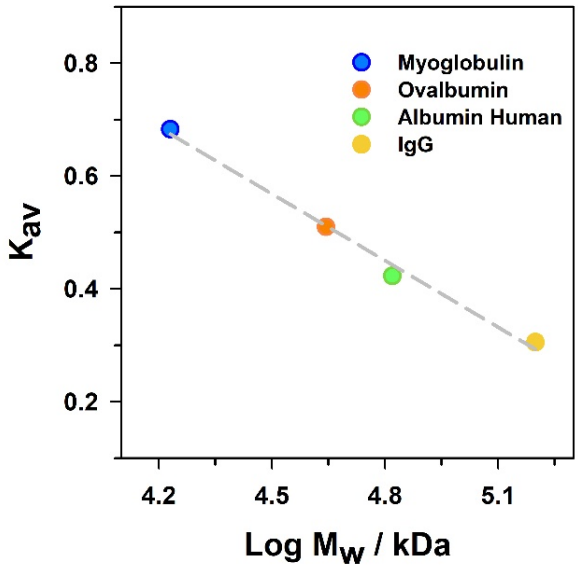

Figure S5. Calibration curve for the determination of the apparent molecular weight (M_w^{app}) of the N form and the β form from the size exclusion chromatography. Log of molecular weight of four standard biomolecules is plotted against their respective partition coefficient (K_{av}). The K_{av} for all the standard biomolecules is calculated from their respective elution volumes noted from the manufacturer provided manual for HiLoad™ 16/600 Superdex 200 pg column using equation 2. The dashed line is a linear fit to the data and used to calculate the M_w^{app} of the N form and the β form in figure 4B (main text).

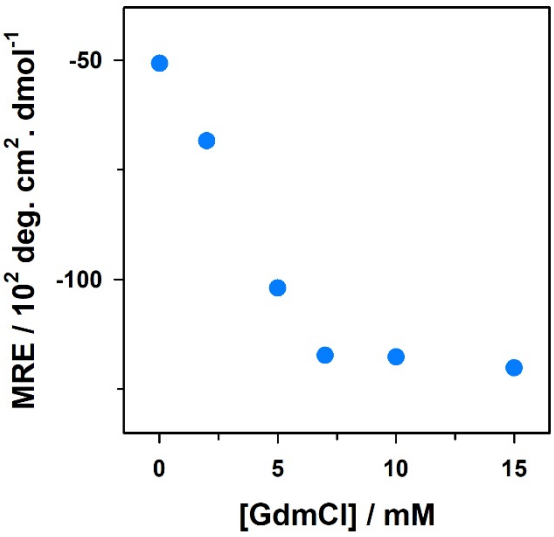

Figure S6. The A⇌β transition occurs at a very low concentration of chemical denaturant. The transition is shown between 0 to 15 mM of [GdmCl], as monitored by far-UV CD signal at 216 nm, upon incubation for 5 days.

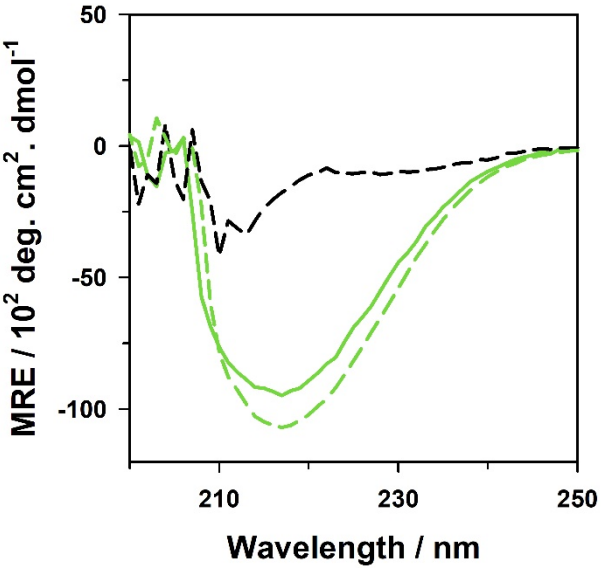

Figure S7. The β⇌U transition is reversible as monitored by far-UV CD. The fluorescence spectrum of the protein in the U form (5 M GdmCl, pH 3) before (black dashed line) and after (green solid line) transferring to a buffer containing 1.1 M GdmCl at pH 3, where the β form is predominantly populated. The green dashed line represents the fluorescence signal of the control β form at 1.1 M GdmCl at pH 3.