Biophysical Journal, Volume *115*

Supplemental Information

Characterization of TDP-43 RRM2 Partially Folded States and Their Significance to ALS Pathogenesis

Davide Tavella, Jill A. Zitzewitz, and Francesca Massi

Basin	statistical weight (%)	R_g (nm)	β content (%)	α content (%)	# ILV contacts
A	37.3	1.15±0.02	12.8±1.7	11.8±3.1	10.2±0.7
B	11.5	1.20±0.02	7.2±1.3	11.2±2.2	5.7±0.6
C	10.7	1.24±0.03	6.7±0.6	12.6±1.3	3.4±0.4
D	10.6	1.61±0.25	1.9±0.8	14.2±1.5	1.8±0.8
E	4.8	1.20±0.02	10.0±0.7	12.9±3.0	8.0±0.4
F	0.5	1.18±0.01	13.0±0.2	11.4±2.2	7.7±0.4

Table S1: TDP-43 RRM2 Free energy landscape

native state	
	β_2:Phe 221
	β_4:Ile 249
partially unfolded states	
basin A, E, F	β_2:Phe 221
	β_3:Phe 229
	β_4:Ile 249
	β_5:Ile 253
basin B	β_2:Phe 221
	β_3:Phe 229 and Phe 231
	β_4:Ile 249 and Ile 250
basin C	β_2:Phe 221
	β_3:Phe 229 and Phe 231
	β_4:Leu 248, Ile 249 and Ile 250
	β_5:Ile 253, Val 255 and Ile 257
basin D	β_1:Phe 194
	β_2:Val 220 and Phe 221
	β_3:Phe 229, Phe 231 and Val 232
	α_2:Ile 239
	β_4:Leu 248, Ile 249 and Ile 250
	β_5:Ile 253, Val 255 and Ile 257

Table S2: TDP-43 RRM2 aggregation prone residues. List of residues with an A3D score > 1 for the native state (PDB ID: 1WF0) and partially folded conformations of TDP-43 RRM2 obtained from the RAM simulations. The residues are organized according to their location in the native state.

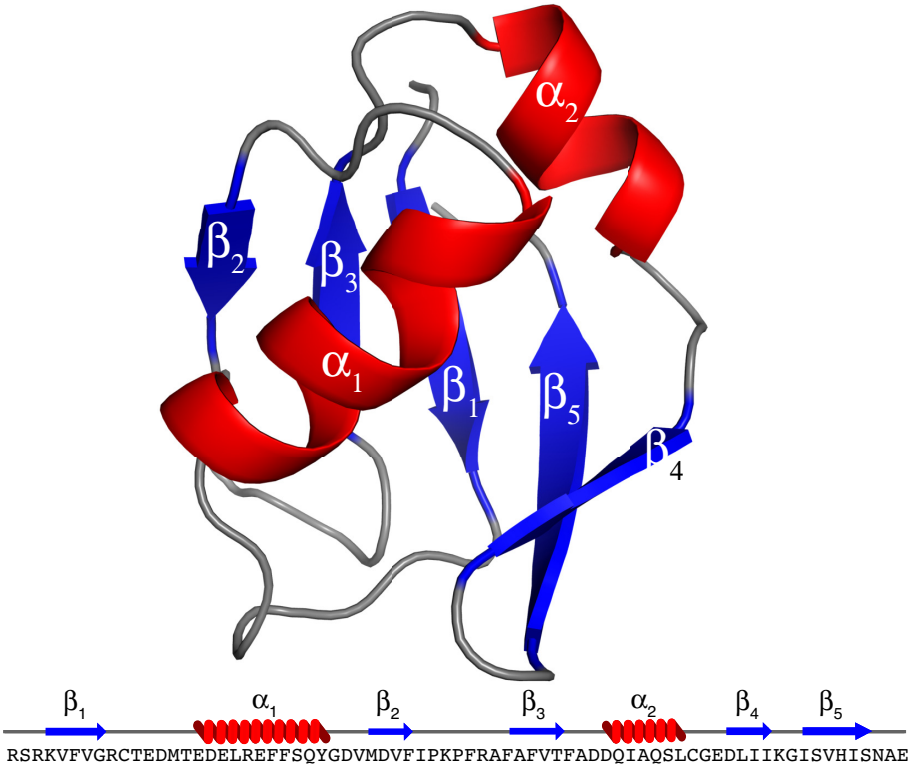


Figure S1: Structure of the native state of TDP-43 RRM2. Top: the structure of TDP-43 RRM2 with secondary structures depicted as red α-helices and blue β-strands. Bottom: sequence of TDP-43 RRM2 with schematic representation of the secondary structures.

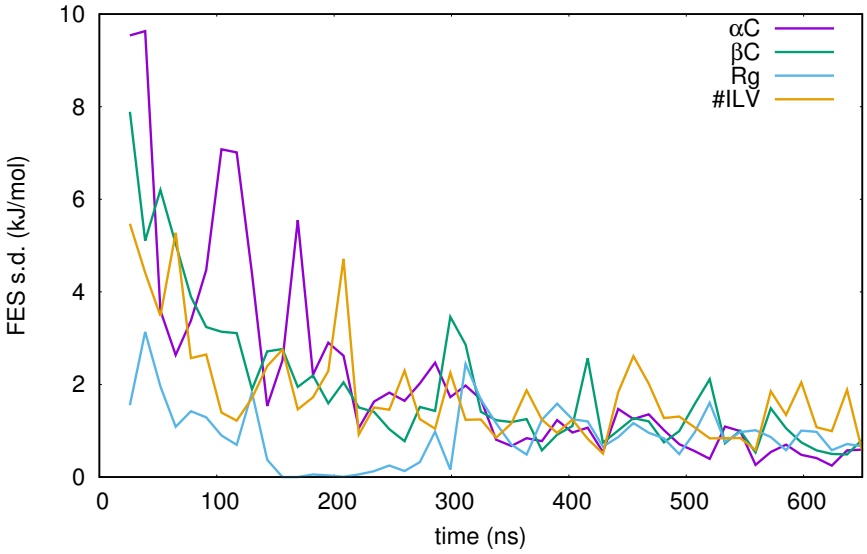
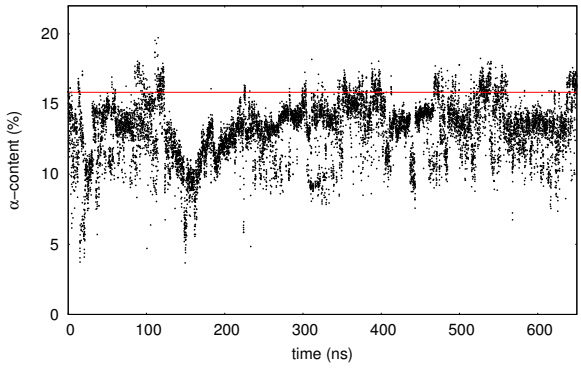
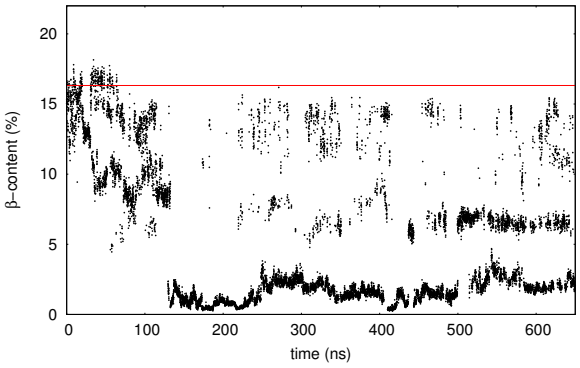


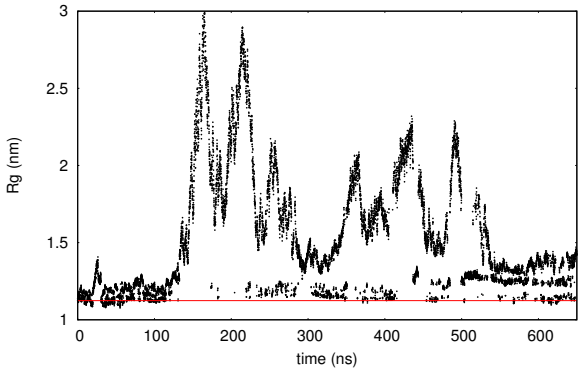
Figure S2: The convergence of the free energy calculations is evaluated by considering the free energy change as a function of the simulation time. The free energy change is calculated as the difference between the free energy at time t-t_0 and the free energy at time t, where t_0 is 13 ns.



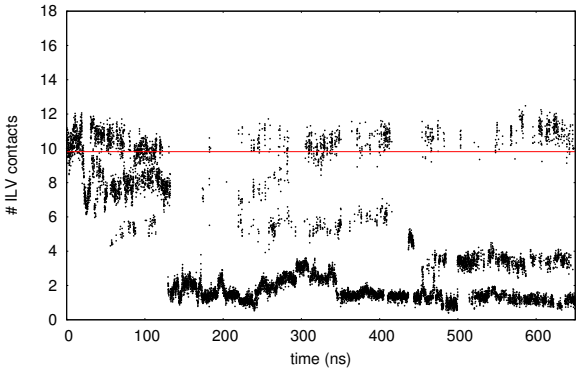
(a) CV1



(b) CV2



(c) CV3



(d) CV4

Figure S3: The values of the four collective variables employed in the RAM simulation as function of time are reported for one of the four replicas. A red line indicates the value of the collective variable for the native state in each plot.

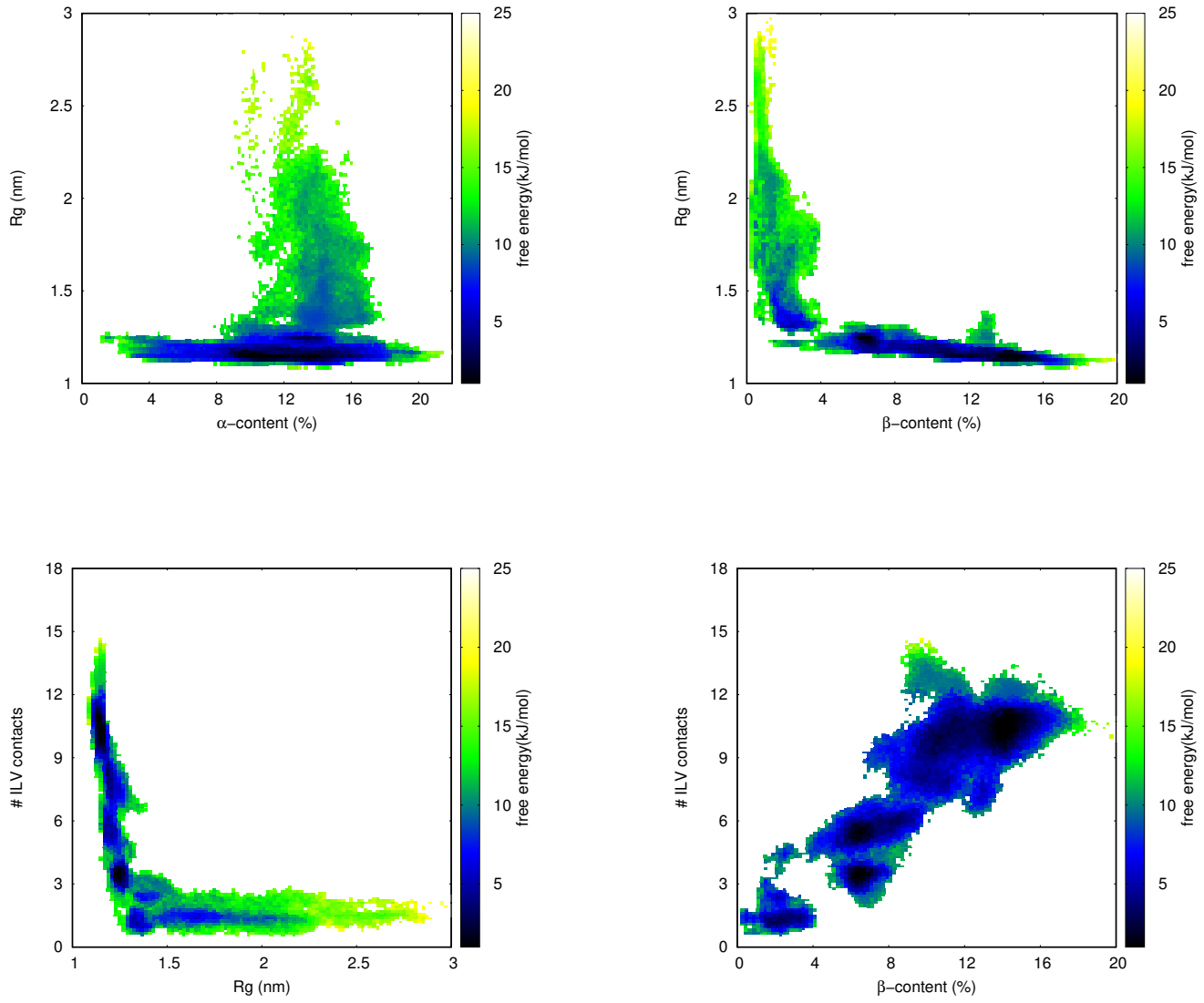


Figure S4: Characterization of the free energy landscape of the partially folded states of the RRM2 of TDP-43. Two-dimensional free energy landscapes as a function of two of the four collective variables used in the replica-exchange metadynamics simulations.

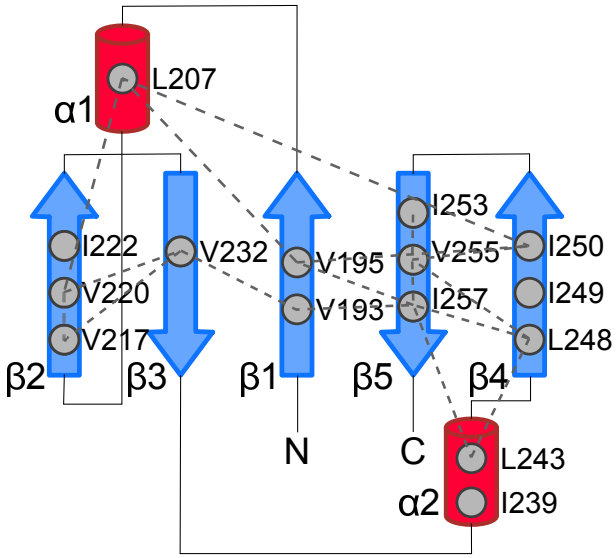


Figure S5: Native ILV cluster contact map for RRM2 is displayed on the secondary structure elements of the native state of the RRM2. α-helices and β-strands are depicted as red cilinders and blue arrows. Secondary structure elements are colored as solid if the structure is present in the microstate. Isoleucine, Leucine and Valine residues are shown as gray circles and ILV contacts are depicted as dashed lines.

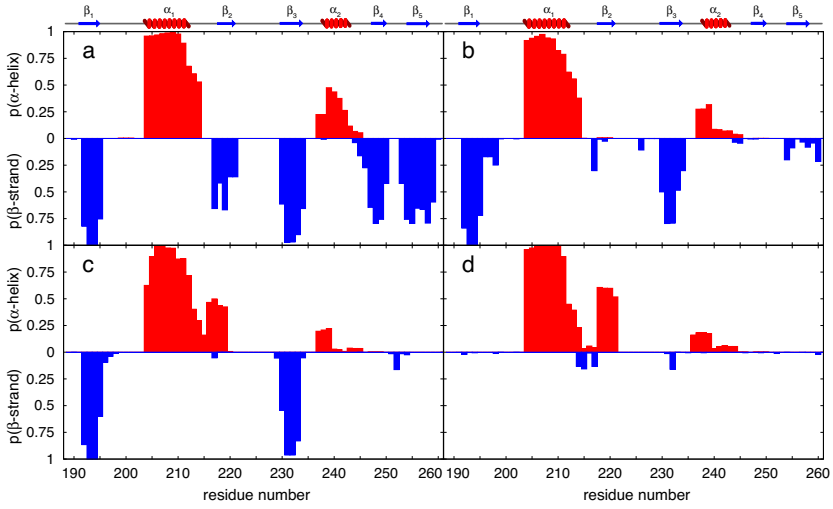
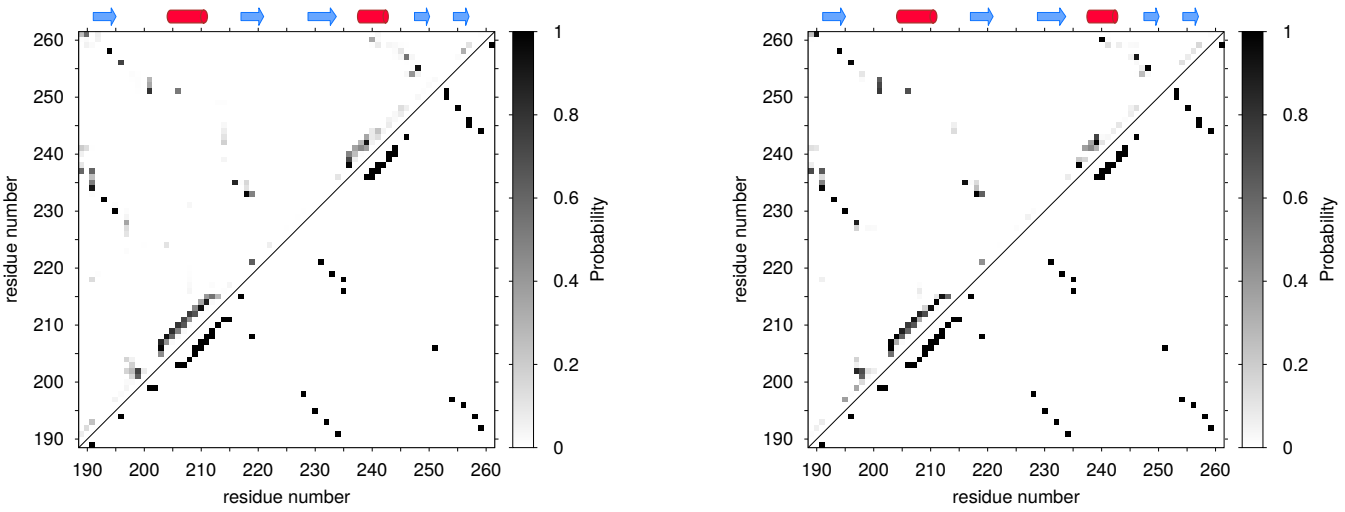


Figure S6: Secondary structures in the partially stable states of TDP-43 RRM2. The probability of each residue being in an α-helix or a β-strand is shown as computed by the DSSP algorithm(1) for the microstates in basin A (a), B (b), C (c) and D (d).



(a) Basin E.

(b) Basin F.

Figure S7: Residual structure of the microstates in the partially folded state of the RRM2 of TDP-43. Comparison of the average probability of hydrogen bonds formation in the microstates corresponding to the basins E and F (above diagonal) with the native state (below diagonal). A schematic representation of the secondary structure elements of the native state is depicted on top of each map.

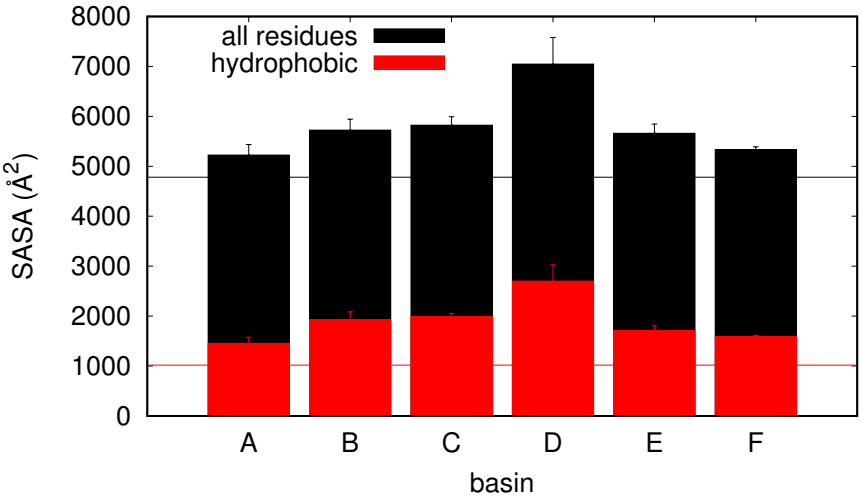


Figure S8: Solvent accessible surface area of the partially folded states of TDP-43 RRM2. The SASA of RRM2 (all residues) and of the hydrophobic residues (Ala, Leu, Val, Ile, Pro, Phe, Met, Trp) are shown for the microstates corresponding to the energy basins. The horizontal lines in the plot indicate the values for the native state.

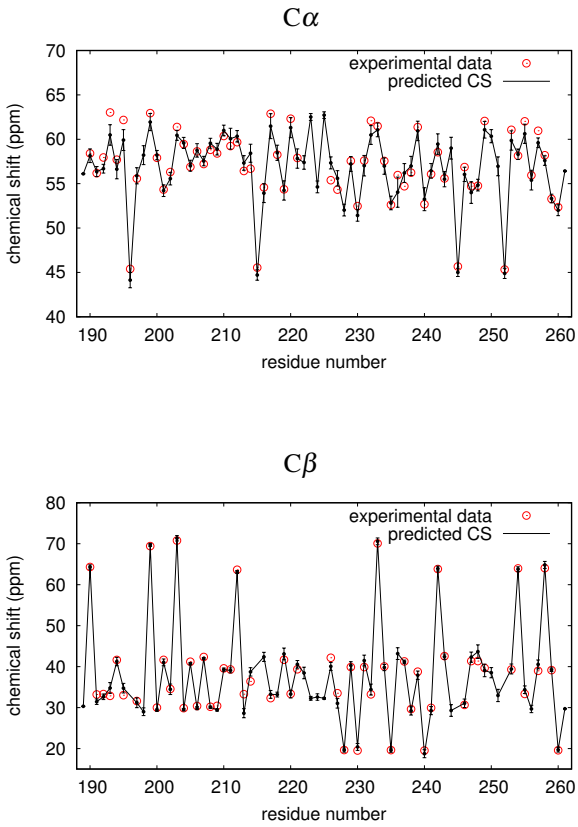


Figure S9: Comparison of the experimental Cα and Cβ chemical shifts with the values backcalculated using TALOS+(2) for the microstates obtained from the RAM simulations. The values are shown as weighted average over the entire ensemble of structures in the energy landscape.

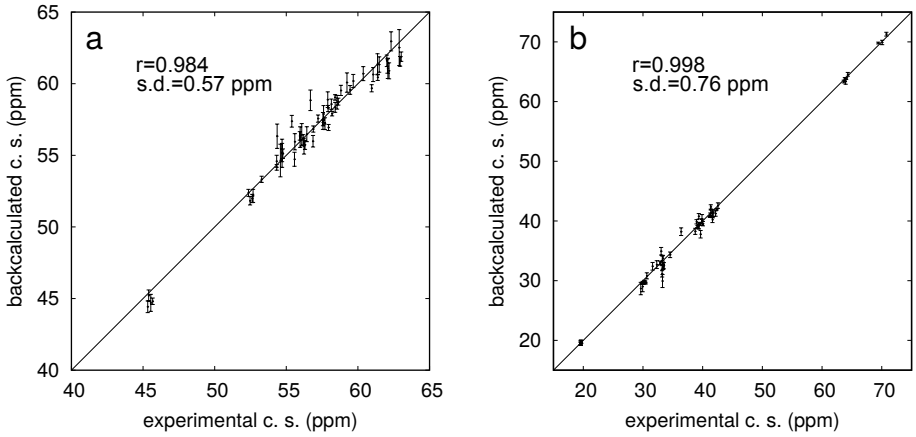


Figure S10: Correlation of the backcalculated Cα (a) and Cβ (b) chemical shifts for basin D with the experimental data collected at 6 M urea.

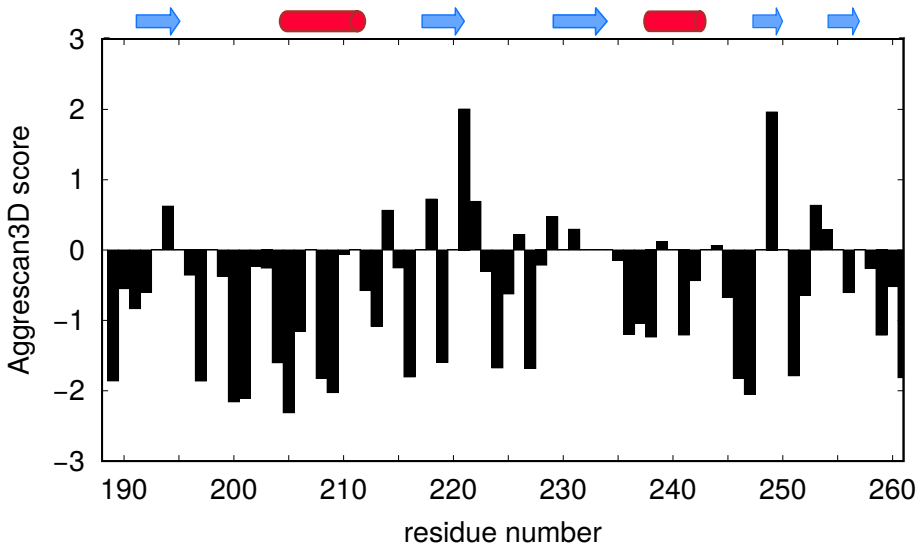


Figure S11: The aggregation propensity of each residue in RRM2 is shown as the Aggrescan3D(3) score for the native state. Positive values correspond to aggregation-prone residues, negative values correspond to soluble residue. A schematic representation of the secondary structure elements of the native state are depicted on top of the plot.

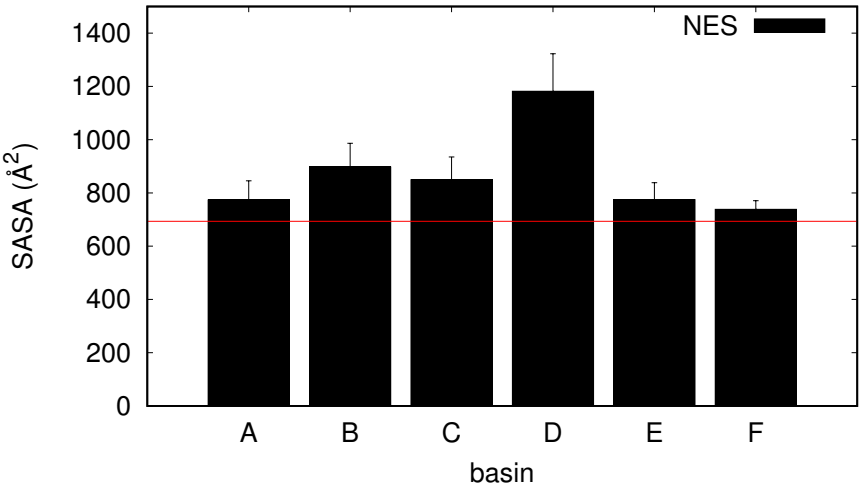


Figure S12: Solvent accessible surface area of the partially folded states of TDP-43 RRM2. The SASA of the nuclear export signal (NES, residues 239-250) region is shown for the microstates corresponding to the energy basins. The horizontal lines in the plot indicate the value for the native state.

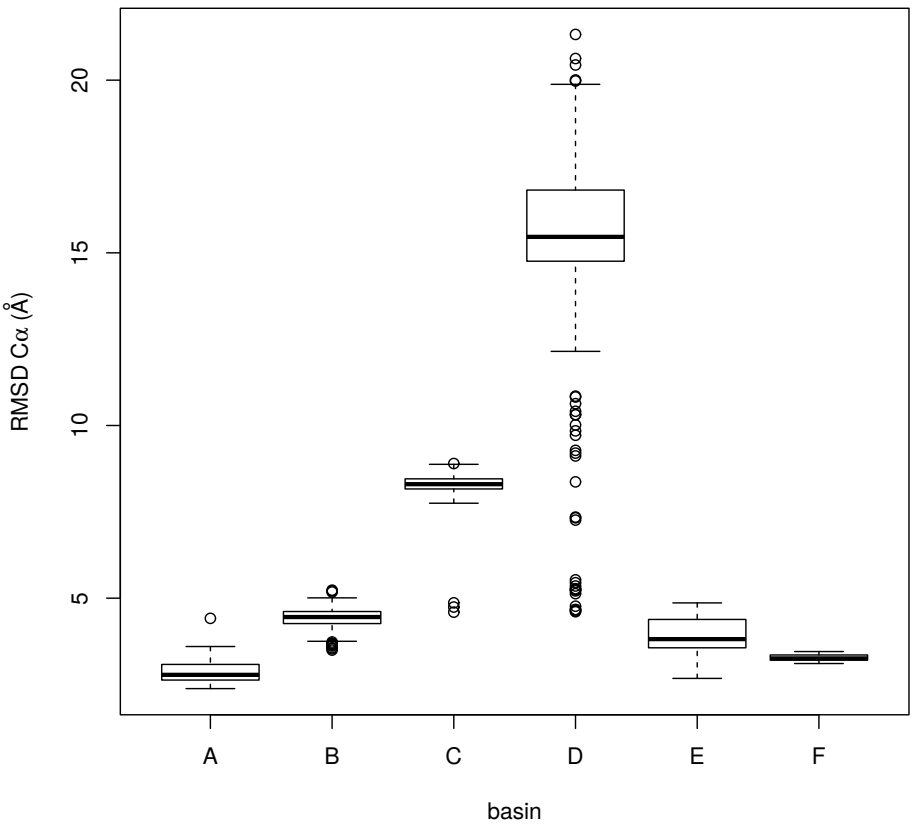


Figure S13: The C_α root-mean-square deviation (RMSD), relative to the native state, for each conformation in the energy basins is shown as a box and whisker plot.

REFERENCES

1. Kabsch, W., and C. Sander, 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22:2577–2637.

2. Shen, Y., F. Delaglio, G. Cornilescu, and A. Bax, 2009. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR* 44:213–223.

3. Zambrano, R., M. Jamroz, A. Szczasiuk, J. Pujols, S. Kmiecik, and S. Ventura, 2015. AGGRESCAN3D (A3D): server for prediction of aggregation properties of protein structures. *Nucleic acids research* 43:W306–W313.