TDP-43 domain assembly and RNA binding specificity

H. Yuan\textsuperscript{1}, P. Kuo\textsuperscript{1}, C. Chiang\textsuperscript{1}, W. Chu\textsuperscript{2}

\textsuperscript{1}Academia Sinica, Institute of Molecular Biology, Taiwan, \textsuperscript{2}National Yang-Ming University, Department of Biomedical Engineering, Taiwan

TDP-43 is an RNA-binding protein with multiple function in the regulation of mRNA splicing and translation. However in diseased neuronal cells, TDP-43 forms aggregates and is linked to various neurodegenerative disorders, including amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD). TDP-43 contains an N-terminal domain (NTD), two RNA binding domains (RRM1 and RRM2) and a C-terminal tail region rich in glycine residues. The pathogenic forms of TDP-43 are processed C-terminal fragments containing a truncated RRM2 and a glycine-rich tail. Although extensive studies have focused on this protein, it remains unclear how the dimeric full-length TDP-43 is folded and assembled and how the processed C-terminal fragments are misfolded and aggregated. Here we present the crystal structure of TDP-43 RRM1 domain in complex with a single-stranded DNA. This structure reveals the molecular basis for the recognition of TDP-43 to UG/TG-rich nucleic acids. Combining with SAXS data, we suggest that TDP-43 is assembled into a homodimer via its NTD and interacts with long clusters of UG-rich RNA by its two RRM domains. Moreover, we found that the ALS-linked mutant D169G binds DNA efficiently but is more resistant to thermal denaturation, suggesting that the resistance to degradation is likely linked to TDP-43 proteinopathies. Our data also suggest that the proteolytic cleavage of TDP-43 within RRM2 may remove the NTD dimerization domain and produce unassembled truncated RRM2 fragments with glycine-rich C-terminal tails that can further oligomerize into high-order inclusions.


Keywords: RNA-binding protein, protein misfolding, protein aggregation