Structural Studies of TDP-43- a New Player in Neurodegenerative Diseases

Protein misfolding and accumulation of misfolded proteins in neuronal cells are associated with many neurodegenerative diseases. TDP-43 is an RNA-binding protein, and its binding to UG-rich RNA is related to the common lethal genetic disease of cystic fibrosis (CF). On the other hand, TDP-43 forms pathogenic inclusions in neuronal cells, directly linked to frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS). We have determined the crystal structure of TDP-43 RRM2 domain in complex with a single-stranded DNA. The crystal structure revealed that RRM2 has an additional unique β-strand for protein assembly. Moreover, TDP-43 RRM2 domains, part of the fragments identified in TDP-43 inclusions in pathological brain cells, were assembled into a left- handed "super β-helix" structure. This assembly of TDP-43 is likely to have broad implications for the understanding of the non-amyloid aggregates in the neurodegenerative diseases related to TDP-43 proteinopathy.

Beamlines

13C1 SW6
13B1 SW6

Authors

P.-H. Kuo, L. G. Doudeva, and H. S. Yuan
Academia Sinica, Taipei, Taiwan



Introduction

TDP-43 was originally identified as a transcriptional factor, repressing the transcription of *HIV-1* gene. TDP-43 is both a DNA-binding and RNA-binding protein, and bears multiple functions in transcriptional repression, pre-mRNA splicing, and translational regulation (1-3).

Recently, breakthrough studies showed that TDP-43 is the major disease protein in the pathogenesis of both FTLD with ubiquitin inclusions and ALS. TDP-43 was identified as the major component protein in the ubiquitinated inclusions in both FTLD-U and ALS disorders. Pathological TDP-43 in the cytoplasmic and intranuclear inclusions is hyper-phosphorylated, ubiquitinated and cleaved to ~25 kDa C-terminal fragments in affected brain regions, and these abnormal aggregates of phosphorylated and ubiquitinated TDP-43 thus define a new group of neurodegenerative diseases, the TDP-43 proteinopathies.

We have determined the crystal structure of a TDP-43 RRM2 (the second RNA Recognition Motif) in complex with a single-stranded DNA (4). The RRM2 domain has an atypical RRM fold with an additional β-strand involved in making domain-domain interactions. These studies characterize the recognition between TDP-43 and nucleic acids and the mode of RRM2 self association, and provide molecular models for understanding the role of TDP- 43 in cystic fibrosis and the neurodegenerative diseases related to TDP-43 proteinopathy.

RRM2 is a tetramer in low salt conditions

To study the biochemical properties of TDP-43, we constructed three truncated forms of mouse TDP-43, TDP-43s, RRM1 and RRM2. TDP-43s covered both RRM domains from residues 101 to 285, RRM1 covered the first RRM from residues 101 to 191, and RRM2 covered the second RRM from residues 192 to 285. The three truncated forms of TDP-43 were overexpressed in *Escherichia coli* and purified by several purification steps. RRM2 was further applied to a SDS PAGE with 98% purity and a native PAGE where it appeared as a homotetramer with a molecular weight between 40 to 55 kD. These results show that RRM2 are

homotetramers, with four copies of the RRM in each tetrameric assembly.

Overall crystal structure of RRM2-DNA complex

RRM2 was co-crystallized with a single-stranded 10-mer DNA (5′-GTTGAGCGTT-3′) with a TG binding site at the third and fourth positions (underlined). The RRM2-DNA complex crystallized in cubic F222 space group with one molecule per asymmetric unit. The structure of the complex was solved by molecular replacement using the NMR solution structure of human TDP-43 RRM2 domain (PDB entry code: 1WF0, unpublished results) as the search model. The final model contained one RRM2 molecule (residues 190 to 261) and nine out of ten nucleotides (T2 to T10) with an R-factor of 20.7% and an R-free of 24.6% for 63,632 reflections up to a resolution of 1.65 Å.

RRM2 was a tetramer in low salt conditions, however, RRM2 appeared as a dimer in the crystals grown from acidic high-salt conditions. A gel filtration analysis further confirmed that RRM2 indeed dissociated into dimers under crystallization conditions. The ribbon model of RRM2 dimer is shown in Fig. 1A, where RRM2 molecule **A** and **B** are related to each other by a crystallographic 2-fold symmetry. A typical RRM contains a four-stranded β-sheet, however, RRM2 in TDP-43 has an additional β4 next to β5 (β2-β3-β1-β5-**β4**). Thus, the RRM2 has a sandwich structure containing a five-stranded β-sheet packed with two *α*-helices. The single-stranded DNA was bound at the β-sheet of RRM2 with the three 5′-end nucleotides, T2, T3 and G4, interacting most extensively with the β-sheet residues, whereas the 3′-end nucleotides stretched away and made no interaction with RRM2. The electrostatic potential mapping onto the solvent accessible surfaces of RRM2 showed that the DNA binding surface was highly basic (Fig. 1B).

RRM2 domains assemble into fiber-like structure

The RRM2 dimer (monomers **A** and **B**) observed in the crystals should be highly stable since a ten-stranded antiparallel β-sheet was formed in the dimer. A close look at the crystal packing shows that RRM2 dimers were assembled into a fibril-like solenoid structure (see Fig. 2). The dimeric TDP-43 RRM2 was packed against the neighboring two-fold-symmetry related dimers. Each RRM2 domain used the β4 strand to interact with the RRM2 domain within the dimer, whereas it used the *α*1/β2 to interact with the neighboring RRM2 dimers. The ten-stranded β-sheets in the RRM2 dimer thus were wrapped into a left-handed super helix with *α*-helices buried inside forming the hydrophobic core. Overall, this "super β-helix" had diameters of 36 and 28 Å. Each turn of the super helix was consisted of twenty β-strands and eight *α*-helices from two dimers, given a helical pitch of 41 Å (Fig. 2). The single-stranded DNAs were bound in between the "super β-helices", functioning like glue to hold the super β-helices into a highly order structure. In crystal packing, each super β-helix interact each other on the edge and formed 2 dimensional extended structure.

Discussion

TDP-43 is abnormally cleaved and redistributed from nucleus to cytoplasm in the disease neurons. The C-terminal fragments of TDP-43 are then aggregated forming the pathogenic inclusions in brain cells. All these results combining with our structural study implicate that once the RRM1 domain is deleted from TDP-43, the non-natural

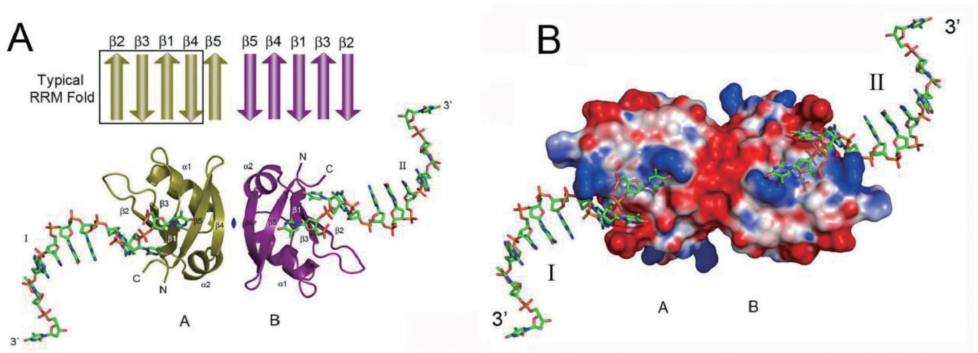

Fig. 1: Crystal structure of RRM2-DNA complex. (A) The ribbon model of RRM2 dimer bound to single-stranded DNAs. Molecule **A** (in dark yellow) is related to molecule **B** (in purple) by a two-fold crystallographic symmetry axis (marked by an ellipsoid). DNA molecules are displayed in stick models. (B) The DNA-binding surface of RRM2 is highly basic. The electrostatic potential, mapped onto the solvent-accessible surfaces of RRM2, was calculated by APBS. The color scale was set from -45 kT/e (red) to +45 kT/e (blue).

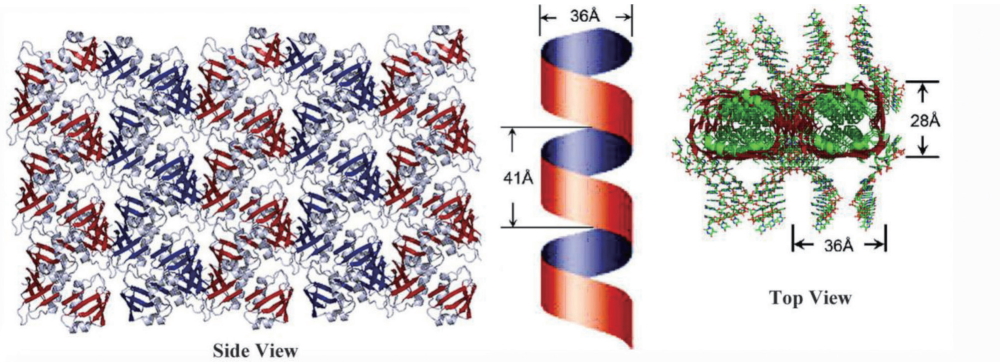

Fig. 2: The RRM2 domain assembles into a fiber-like structure. The crystal packing of RRM2 shows that the RRM2 dimers interact with the neighboring RRM2 dimers to generate a left-handed super β-helix. The α-helices of RRM2 are buried inside of the super β-helix forming the hydrophobic core to further stabilize the structure.

interface on RRM2 is exposed for protein aggregation, and the RRM2 domain likely may generate this type of "super β-helix" aggregates.

Potentially RRM2 domain of TDP-43 may assemble into this "super β-helix" structure and form pathogenic inclusions. This well-organized super β-helix structure of TDP-43 RRM2 domain thus has broad implications in the group of neurodegenerative diseases of TDP-43 proteinopathy. This well-organized RRM2 dimeric structure thus offers a testable model for the study of the pathogenic aggregates in the group of neurodegenerative diseases known as TDP-43 proteinopathies.

Experimental Station

Protein Crystallography end station

References

1. I.-F. Wang, L.-S. Wu, and C.-K. J. Shen, Trends Mol. Med. **14**, 479 (2008).
2. I.-F. Wang, L.-S. Wu, H.-Y. Chang, and C.-K. J. Shen J. Neurochem. **105**, 797 (2008).
3. I.-F. Wang, N. M. Reddy, and C.-K. J. Shen, Proc. Natl. Acad. Sci. USA **99**, 13583 (2002).
4. P.-H. Kuo, L. G. Doudeva, Y.-T. Wang, C.-K. J. Shen, and H. S. Yuan, Nucleic Acids Res. **37**, 1799 (2009).

Contacted E-mail

hanna@sinica.edu.tw